2014 International Conference on Advanced and Agile Manufacturing Systems

ICAM-2014
Oakland University, Rochester, MI 48309, USA
May 28-30, 2014

Sponsored by the International Society of Agile Manufacturing (ISAM), International Society for Productivity Enhancement (ISPE) and Oakland University, USA.

Ramrattan, Margaret K. Joyce, Paul D. Fleming and Pavel Ikonomov	New Light Cured Media for use with Cast Prototypes	28
Technical Session 29th May, 1:30 pm to 3 pm Room 3-- Dodge Hall 203 , Session Chair: Michael Sobolewski		
Makan Taghavi Dilamani	A short review on Crystal Clear methodology and its advantages over the scrum, the popular software process model	29
Mohammed Iqbal and Aravamudan Gopal	Automated Quality Inspection of Citrus Fruits – A Review	5
Dr. Rajat Setia, Prof. K. Hans Raj and Prof. Suren N. Dwivedi	Comparison of ANN and Statistical Regression Models for Prediction of Average Equivalent Strain in Equal Channel Angular Pressing	9
Technical Session May 29th 3:30 to 5 PM-- Room 1- Dodge Hall 200, Session Chair: Suren Dwivedi		
S K Sharma, Anubha Rautela and Rajnish Kumar.	Improving the Issues in Procurement Process of Manufacturer	20
Rajiv Kumar Upadhyay, Ajay Bangar, Pawan Kumar Singh and Ashish Shastri	Enhancing the leanness of supply chain by integrated Fuzzy-QFD approach	22
Dhirendra Rana, Pravin M. Kulkarni, K.P. Karunakaran and Asim Tewari	In-Situ Property Improvements Using A CNC Integrated Pneumatic Hammer	17
Sushil Kumar Sharma, Shaarabh Muraka, Rishi Gupta and Hari Priya Choudhury	A Review on Measurement of Agility in Manufacturing System	18
Technical Session May 29th 3:30 to 5 PM-- Room 2-- Dodge Hall 201, Session Chair: Phares Noel		
Jan-Hinrich Kämper, Arne Stasch and Axel Hahn.	A fully-automated manufacturing environment realized through a flexible in house logistic system with smart transportation infrastructure	27
R S S Prasanth, Bhuvnesh Singhal, Pritam Singh and K Hans Raj	A comprehensive review on modeling and optimization of friction stir welding	19
Seyed Mirmiran	Need for an Advanced Asset Management System	41
Brandon J. Voelker, Muralidhar K.		

Ghantasala, Paul V. Engelmann and Jeff Wheeler	Manufacturing and Electroplating of Nanoengineered Polymers	36
Technical Session May 29th 3:30 to 5 PM-- Room 3-- Dodge Hall 203, Session Chair: Dr. Kai Yang		
Pavel Ikonomov and Jorge Rodriguez	3D Metal Printing/Machining	31
Pavel Ikonomov, Azem Yahamed and Dan Fleming	Application of 3d Printing for Human Bone Replacement	32
Pavel Ikonomov, Emiliya Milkova and Suren N. Dwivedi	Recognition of Moving Objects Using Sensors System for Human/Robot Teamwork	33
Pavel Ikonomov and Muralidha Ghantasala	Analyze and Determine the Forces Associated with the Nanoparticle Movement	34
Technical Session May 29th 3:30 to 5 PM-- Room 4-- Dodge Hall 202, Session Chair: Dr. Biren Prasad		
Shreyas Harish and Thirumalesh H.S.	Design and Implementation of an Unmanned Ground Vehicle	56
Ashidsha Jaleel, Rajendran T K and Lijohn P George	Cloud Manufacturing: Intelligent Manufacturing with Cloud Computing	26
Jorge Rodriguez, Charles Crouch, Joseph Dementer, Brian Guenther and Leah VanEeuwen	Mold Design for Injection Molding Using Additive Manufacturing	58
Suren N Dwivedi	Project Based Learning for STEM (Science, Technology, Engineering, Mathematics) Education	59
Technical Session May 30th 1:00 to 2:30 PM -- Room 1-- Dodge Hall 200, Session Chair: Dr. Hans Raj		
Gaurav Agarwal, Abhishek Agarwal and Shubham Agarwal	BIPV: Integration of Photovoltaic with the Construction to achieve the concept of Agility	35
Khalid Mirza, Sai Prasanna, Michael Truitt and Hudhaifa Jasim.	Intuitive 3D-Vision Based Wand for Robot Tool Path Teaching	45
Technical Session May 30th 1:00 to 2:30 PM -- Room 2- Dodge Hall 201, Session Chair: Dr. Ashok Prajapati		
Tyler Bayne, Spencer		

Hoin, Dr. Pnina Ari-Gur, Dr. Marwa Hassan, Mr. Peter Thannhauser, Dr. Roman Rabiej, Dr. Pavel Ikonomov, Mr. Jeff Johnston and Dr. Dan Litynski	Virtual Reality 3D Simulations of Concrete and Asphalt Laboratories	54
Pnina Ari-Gur,Shubram Subramanyam, Ashkan Razania, Andreas Quainoo and Sven Vogel	Formability and Crystallographic Texture in Novel Magnesium Alloys	30
Technical Session May 30th 1:00 to 2:30 PM -- Room 3-- Dodge Hall 203, Session Chair: Dr. Biren Prasad		
Mohammad Komaki, Shaya Sheikh and Behnam Malakooti	Ruled Based Approach for Scheduling Flow-shop and Job-shop Problems	48
Mohammad Komaki, Shaya Sheikh and Behnam Malakooti	Multi-Objective Scheduling Using Rule Based Approach	49
Mohammad Komaki, Shaya Sheikh and Behnam Malakooti	Rule Based Layout Planning and Its Multiple Objectives	50
Behnam Malakooti	A Synopsis of Multiplicative Z Utility Theory for Solving Risk Problems	51
Technical Session May 30th 1:00 to 2:30 PM -- Room 4-- Dodge Hall 203		
Pankaj Sharma and Ajai Jain	Performance of dispatching rules in a stochastic dynamic job shop manufacturing system with sequence-dependent setup times	7
Puneet Mangla, Ashish Agarwal, Pulak M Pandey and Subrata Das	A Study of Factors Related to Supply Chain Strategy (SCS)	23

Automated Quality Inspection of Citrus Fruits – A Review

S. Md. Iqbal[1,2], A. Gopal[1]

[1]CSIR-Central Electronics Engineering Research Institute, Chennai Centre
CSIR Madras Complex, CSIR Road, Taramani, Chennai-600113, India
[2]Research Scholar, Sathyabama University, Chennai-600119, India
Email: iqbalsmd@gmail.com, agopal@hotmail.com

Abstract

Non-destructive quality inspection of fruits provides quality products for domestic consumption as well as for export markets, increasing consumer/buyer confidence. It provides assurance of quality and subsequent value addition. A major thrust of current research is towards the development of quality inspection systems for improved sorting and automated quality control. Machine Vision inspection has received wide attention in recent years, due to various advantages compared to the conventional manual inspection. These systems typically sort the objects based on various quality features like colour, size, shape, etc. Handling of fruits for presenting them before the camera for grading and subsequent sorting plays a major role in determining the overall performance like throughput etc. The present paper reviews with special focus to citrus fruits, the evolution of traditional fruit sorting and grading techniques and the subsequent machine vision based systems with improved fruit handling and improved techniques for inspecting the fruits for various quality parameters like size, shape, colour, defects, etc. The review covers the different approaches used based on image analysis for quality inspection of citrus fruits in grading and sorting. The review aims to investigate the potential of various techniques like machine vision, infra-red, optical, robotics, mechatronics, etc in quantifying the quality parameters of citrus fruits, like colour, size, shape, etc. and explore the possibilities towards development of improved machine vision based systems for inspection of local varieties of horticultural produces, particularly citrus fruits like orange, mosambi and lemon.

Keywords - Quality inspection, machine vision, image processing, quality parameters, citrus fruits, sorting and grading.

1. Introduction

India ranks sixth in the production of citrus fruits in the world. Spain, USA, Israel, Morocco, South Africa, Japan, Brazil, Turkey and Cuba are some of the other major citrus producing countries. Citrus fruits occupy third position in the production of fruits in India after mango and banana. These fruits originated in the tropical and sub tropical regions of South East Asia, particularly India and China. North East India is the native place of many citrus species. Of the various types of citrus fruits grown in India, orange (mandarin or santra), sweet orange (mosambi, malta or satugudi) and lime/lemon are of commercial importance.

In a review made by Dattatraya Londhe et al [1] about different methods of grading for fruits and vegetables, it is reported that in India mostly the fruits are graded manually with trained operators, which is costly, affected due to shortage of labour in peak season, inconsistent, less efficient and time consuming. They have reported that farmers are in need for an appropriate grading machine to overcome such problems and improve graded products quality. According to them, grading of fruits fetches high price, improve packaging and handling. They have reported that the fruits are generally graded on basis of size, shape, colour and weight.

In another review made by Krishna Kumar Patel et al [2], it is again reported that quality inspection of agricultural produce generally in India are manual which is costly, unreliable, inconsistent, subjective and slow. Also, the accuracy of the manual inspections cannot be guaranteed [3]. Machine vision provides an automated, non-destructive, consistent, cost-effective and objective inspection technique with increased speed and accuracy. This is agreed by Morrow et al. [4], Gerrard et al. [5] and many other researchers. Gunasekaran [6] has reported that the food industry is now ranked among the top ten industries using machine vision technology.

Due to increasing population in India, the demand for consumer products including fruits and vegetables have increased and the need for automated systems have become indispensable. It is therefore necessary to have a simple, appropriate, efficient and customized system to grade the quality of such products to meet this demand.

It is very difficult for developing countries like India with limited financial resources to import very expensive automatic sorting and grading systems. So within the limited budgets, few Indian horticultural producers could import only manual and semi-automatic sorters to meet their demand.

Survey reveals that there are a number of sorters and grader equipment manufacturers abroad who manufacture systems for online sorting and grading of fruits and vegetables which are very costly. Maintenance and service, in handling these equipments are also costly. No Indian manufacturers have been reported in the literature to make MV based system and not much research work has been conducted. At present, most fruit sorting and grading systems being used in India are imported. These systems, mainly mechanical, are used for grading of horticultural produce based on their weight and size only. This has injected an urge among researchers and equipment manufacturers in India to develop low cost automated fruit sorting and grading equipments.

As both the market and the technology are constantly evolving, new fruits & vegetables grading and sorting machines are being continuously developed ensuring careful product handling as well as achieving uniformity of grading. The development of good material handling system added with image processing technique will help in the automation of such industries.

Design and development of any modern fruit quality inspection system is aimed at doing online and real-time quality parameters monitoring, satisfying customer requirements and then based on that take proper control measures for sorting. These automated quality inspection systems are very much required by the fruit industries, not only to avoid inconsistency in manual handling, but also to meet the stringent quality specifications of the product to face the global competition. Citrus fruits industry, particularly handling fruits like sweet lime, orange, lemon, etc. is one such industry, meeting some stiff competition in the global market. However, currently available imported systems are so expensive that many small and medium scale fruit inspection industries in many developing countries like India cannot afford to buy and install them. So they obviously look for affordable customized systems availability.

Development of such a customized and low cost system with suitable material handling components, with control elements for sorting and with proper material selection, is critical and essential for error free and damage free automation of fruits grading and sorting. The proposed study aims at exploring the possibilities towards development of improved machine vision systems indigenously, for inspection of wide range of local varieties of citrus fruits. This would help them to face the global market with the available manpower and materials resources, without compromising the quality parameters. The development of such system would primarily be an integration of mechanical systems, electronic interfaces and sensors with high speed computing systems and associated software components. This paper focus on the review of the work reported on grading and sorting of fruits and agricultural produce, mainly citrus fruits, since 1960.

2. Materials and methods

Grading and sorting of fruits are generally done based on external quality parameters like size, shape, colour, etc. and internal quality parameters like defects, sweetness, etc. Many researchers have developed many kinds of graders and sorters according to market requirement. In all these developments, design of the fruit handling systems with proper material selection and proper techniques play a vital role in the effective working of the online fruit grading and sorting [7]. The present study outlines the application of various techniques used for grading and sorting of fruits, with special focus to citrus fruits.

2.1 Conventional inspection techniques

According to Ayman Amer Eissa [8], product quality and quality evaluation are important aspects of fruit and vegetable production. They considered sorting and grading as the essential steps of fruits handling and also as the major processing tasks associated with the production of fresh-market fruits.

Mechanization of orange grading operations began from a couple of decades ago. In the initial stage of the mechanization, plates made with holes of orange fruit sizes were used for sorting [8]. Traditionally, fruit size used to be measured using sizing ring (Food & Agriculture Organization of the United Nations [FAO]) [9], drum-type grading machine [10] and light blocking type grading machine [11]. The above approaches enabled the classification of fruits based on their size.

Certain conveyors [12] employed a series of precision-machined stainless steel rollers for sorting fruit / vegetable in which size sorting is done by pre-setting the gap between rollers. At the beginning, objects which are smaller than the gap will pass through the rollers and exit underneath the machine. Larger objects will move forward over the rollers until exiting at the end of the machine where the gap is set wider. In another commercial sorting machine, a rotating conical drum fitted with baffle plates along its periphery preset with different gaps is used, through which fruits of different sizes are sorted [13].

According to Stroshine et al [14], sizing by weighing mechanism is recommended for the irregular shape product. Since electrical sizing mechanism is expensive and mechanical sizing mechanism reacts poorly, they have recommended dimensional method (using length, area and volume) for citrus fruit (orange) sorting. They have determined relationships between mass and dimensions and projected areas.

Tabatabaeefar et al [15] have listed several types of sizing machines existing, including perforated conveyer sizers, Greefa type belt and board sizers, rotary Greefa sizers, belt and roller sizers, Wayland-type belt and roller sizers, diverging belts Jansen-type sizers, diverging roller sizers, and weight sizers. The sizing parameters in some of these sizers are the diameter and length or a combination of these two. Dattatraya Londhe et al. [1] have added some more mechanical grading sorters to the list based on size, viz., expanding pitch rollers, inclined vibrating plate, rotary disc type, counter rotating roller having inclination type, sieve type grader.

Anonymous [16] had developed a hand operated size grading machine for orange fruits based on tapering roller principle with 80% separation efficiency. Nevkar [17] had developed and tested divergent roller type grader for lemon fruits and chiku (sapota) fruits. It was observed that the separation efficiency decreased with increase in roller speed. The overall separation efficiency for lemon fruits and chiku fruits were 71.71% and 66.75% respectively. Ghuman and Kumar [18] have developed low cost rotary disc size grader for fruits and vegetables of different diameters.

Mangaraj et al [19] have developed a stepwise expanding pitch fruit grader based on the principle of changing the flap spacing along the length of movement of fruits with provision to separate fruits into four grades by adjusting flap spacing between 45 to 140 mm. They could obtain an overall grading efficiency of 91.5% and 88.5% for sweet lemon and orange, respectively. The capacity of the grader was 3.5 TPH at grading conveyor speed of 6 m/min.

Burt and Patchen [20] have developed and tested a manual unitized machine for sorting, brushing and sizing fruits. Such machines have the advantages like minimum transfer of fruit from one section to another which reduces fruit damage, control of rate and direction of fruit rotation and less floor-space requirement than for conventional grading and sizing lines.

Bryan et al [21] have developed a mechanical separator to divert most of the unwholesome orange by differences in bouncing behaviour into a small side stream.

An online system with the use of a robotic device used by Molto et al [22] resulted in a running time of 3.5 s per fruit for the technique. According to Kondo [23], in recent ten years, operations in fruits and vegetables grading systems have become highly automated with mechatronics and robotics technologies.

2.2 Machine Vision based inspection techniques

Computer vision also known as Machine vision is a novel technology for acquiring and analyzing an image of a real scene by computers and other devices in order to obtain information or to control machines or processes [24]. The technique has been used for the automated inspection and grading of fruits to increase product throughput. Research indicates that using machine vision systems improve product quality while freeing people from the traditional hand-sorting.

2.2.1 On the basis of External Quality Parameters like Colour, Size, Shape, etc.

According to Krishna Kumar Patel et al [2], **the design for a specified machine vision system usually is uniquely structured to suit the inspection of a particular product.** Thus for example, conveyor sizes have been developed for various size range of regular shape produce by different companies depending upon the fruit sizes handled. It is reported that CVS Unisorter, [25] have developed two conveyor sizes, 4” pocket for smaller fruits – 1 ⅝” (40mm) to 4 ¾” (120mm) and 4 ¾” for larger fruits – upto 6” (155mm). Autoline Fruit sorting system [26] have developed five conveyor sizes – 3” pocket to 6” pockets under various requirements.

Ahmad et al [27] and Khojastehnazhand et al [28] have developed citrus fruits sorting system based on colour and size, using image processing with CCD cameras. Khojastehnazhand et al [28] have developed an image processing technique for estimating diameters, volume, mass and surface area of citrus fruits using two CCD cameras. According to them, though grading systems provide information such as size, color, shape, defect, and internal quality; color and size are the most important features for accurate classification and sorting of citrus fruits such as oranges, lemons and tangerines. RGB system is sensitive to lighting or other conditions. By comparing the information on the HSI (Hue, Saturation and Intensity) color values and estimated volumes of different grades of lemon during sorting phase with the available information inside the database, the final grades of the passing fruits were determined. The color of fruit was determined by calculating average Hue (H) value for the fruit.

Tao [29], Majumdar and Jayas [30], Wang and Nguang [31], Eifert et al. [32] have done studies focused on the application of computer vision to product quality inspection and grading based on colour, shape, size, textural feature, volume and surface area of fruits like apples, peaches, tomatoes and citrus fruits.

Khojastehnazhand et al [33] have developed an image processing algorithm for determination of volume and surface area of orange. The machine vision system consisted of two CCD cameras placed at right angle to each other to get two perpendicular views of the image of the orange, an appropriate lighting system and a personal computer. Initially, the algorithm segmented the background and divided the image into a number of frustums of right elliptical cone. The volume and surface area of each frustum were then computed by the segmentation method. They found that the difference between the computed volumes and surface areas obtained by the image processing method and measured by water displacement and tape method, respectively, were not statistically significant at the 5% level. The above method proposed by them was rotationally invariant and did not require fruit alignment on the conveyor. The background segmentation method adapted by them was not based on threshold values, and therefore it can be used with other fruits. They have suggested that the method may be easily integrated with HSI colour space, in an online multi-product sorting system for grading citrus fruits.

Anonymous reviewed an electronic colour sorter [34] for sorting of fruits like tomatoes, plums, papayas, pineapples, etc. Lino et al. [35] used electronic systems consisted of CCD camera and a personal computer for image capturing in quality evaluation of tomatoes and lemons. Commercial sorters frequently use a conveyor system with either shallow cups or bicone rollers that allow fruits to rotate while moving along the conveyor.

Electronic devices like colorimeters measure colours in numerical coordinates. However, these devices are limited to the measurement of small regions of a surface or only when the object has a homogeneous colour [36]. Instead, still or video cameras are more suitable where the surface has a heterogeneous colour [37] and can provide images in which the colours of the pixels are determined individually.

Gaffney determined that a particular variety of oranges could be sorted by colour using a single wavelength band of reflected light at 660 nm [38]. This technique was capable of distinguishing between normal orange, light orange and green fruits.

Slaughter and Harrel [39] devised a method to identify mature oranges based on colour images obtained using a colour camera and artificial lighting. The system used the Hue and Saturation components of each pixel and they were able to classify approximately 75% of the pixels correctly. Because of two-dimensional feature space, two thresholds were employed based on the maximum and minimum values for the saturation and the hue components.

Slaughter and Harrel [40] extended their earlier study by using the RGB components of each pixel recorded by a colour camera as features and a traditional Bayesian classifier method to segment the fruit pixels from the background pixels. They classified each pixel as belonging to a fruit or to the background without using artificial lighting or optical filters. The tests showed that 75% of the pixels were correctly classified.

In identifying external defects in citrus fruits, Blasco et al., [41] compared five colour spaces and obtained better results with HSI. Frequently, individual HSI coordinates provide simple means of colour segmentation. Anna Vidal et al [42] in their work used the Hunter L,a,b system to determine the Citrus Colour Index (CCI).

Anna Vidal et al [42] have coated the interior sides of the inspection chamber with anti-reflective material to minimize the impact of the specular reflections and used cross polarization by placing polarizing filters in front of the lamps and in the camera lenses. They have powered the fluorescent tubes by means of high frequency electronic ballast to avoid the flickering effect of the alternate current and produced a more stable light.

The US patent [43] describes an apparatus made for automatically sorting fruit and the like by colour or weight, or both, using conveyance system to move objects to be sorted past an electromechanical weighing station (comprising strain gauge) and an optical colour sensing station which, in conjunction with sequential and combinational logic, compare the colour and weight of the item to a predetermined criteria and sort accordingly.

Kondo et al. [44] investigated a non-destructive quality evaluation of oranges using fruit colour, shape and roughness of fruit surface, R G colour components ratio, Feret diameter ratio and textural features by means of machine vision system and neural networks.

Blasco et al., [45] used parameters like elongation, roundness, symmetry and compactness to describe the shape of the object.

2.2.2 On the basis of Internal Quality Parameters and Defects

It is reported that till recently, there was no imaging process commercially used to detect defects or contamination due to lack of a method for imaging 100% of the entire surface of individual fruit. Thus, manual sorting still remain the primary method for removal of fruits with defects [46]. Leemans et al [47] have mentioned that the fruits quality criteria like size, colour and shape are actually automated on industrial graders, but grading according to the presence of defects is not yet efficient and consequently remains a manual operation, repetitive, expensive and not reliable. According to Ayman, automating fruit defect sorting is still a challenging subject due to the complexity of the process [8].

In the recent one decade, machine vision and near infrared (NIR) technologies have been utilized and improved with engineering design not only to detect fruit size, shape and colour but also to measure the internal parameters like sugar content and acidity [8]. Since then, image processing techniques have been established allowing not only the size and colour measurements of the fruits but also the non-destructive determination of blemishes [48]. Nowadays, several manufacturers around the world produce sorting machines capable of pre-grading fruits by size, colour and weight.

According to Blasco et al., [45] in the field of computer vision systems for the inspection of fresh, whole fruit, most research has been focused on citrus fruits. They have developed a multispectral system to identify the 11 most common defects of citrus skin using near infrared, colour and ultraviolet with maximum success rate of 87%.

Ruiz et al [49] have studied three image analysis methods, namely, colour segmentation based on LDA, contour curvature analysis and a thinning process to solve the problem of long stems attached to mechanically harvested oranges which were wrongly identified as defects, with classification efficiency of 93%. Kondo [50] predicted the sweetness of the oranges using image processing with a quality evaluation efficiency of 87%. Using the reflectance properties of citrus fruit exhibited at the visible and NIR wavelengths, Molto and Blasco [51] were able to identify the peel and major defects.

Aleixos, et al. [52] developed a multispectral inspection of citrus fruits in real-time using Machine Vision and Digital Signal Processors with a classification accuracy of 94%. Omid et al. [53] developed an image processing based technique to measure volume and mass of citrus fruits such as lemons, limes, oranges, and tangerines. The technique used two cameras to give perpendicular views of the fruit. Naoshi et al [54] used a pair of white and UV LED lighting devices and a colour CCD camera to detect rotten citrus fruits.

Cerruto et al. [55] proposed a technique to segment blemishes in oranges using histograms of the three components of the pixel in HSI colour space. To estimate the maturity of citrus, Ying et al. [56] used a dynamic threshold in the blue component to segment between fruit and background. They then used neural networks to distinguish between mature and immature fruit.

Machine Vision systems and NIR inspection systems have been introduced to many grading facilities with mechanisms for inspecting all sides of fruits and vegetables [23]. In one of the applications, it is reported that an inspection system was developed on an orange grading line [8] where oranges were singulated by a singulating conveyor and sent to the NIR inspection system (transmissive type) to measure sugar content, acidity and inside water content of fruit, then to a X-ray imaging system for detecting biological defect and finally to an external inspection stage consisting of six machine vision sets to take colour images - four cameras for acquiring side images and two cameras for acquiring top and bottom images. After the images were processed for detecting image features of colour, size, shape and external defect, signals were sent to the judgment computer where the final grading decision was made based on fruit features and internal quality measurements.

In addition to Sugar content, Total Soluble Solids (TSS) of citrus fruits could also be measured using NIR spectroscopy [57]. Yamakawa et al [58] have developed and laboratory tested a nondestructive citrus fruit quality monitoring system. Prototype system developed by them consisted of a Light Detection and Ranging (LIDAR) and visible near infrared spectroscopy sensors installed on an inclined conveyer for real-time fruit size and total soluble solids (TSS) measurement respectively. The measuring probe used in this study consisted of ring light arranged in a concentric pattern and centrally-located light receiving fiber. Light emitted from the ring light goes through the sample and spreads inside the sample. Later, it was detected by light receiving fiber and the data was passed to the main unit. This device measured and indicated TSS of fruits with calibration curve downloaded into this device. It is reported that laboratory test results revealed that the developed system was applicable for determination of instantaneous fruit size with $R^2 = 0.912$ and TSS with $R^2 = 0.677$, standard error of prediction =

0.48 °Brix. Actual fruit sizes were also measured using the caliper and were compared with the fruit size estimates. Researchers have used fourier transform NIR [59], visible NIR spectrometers [60], [61] for nondestructive TSS measurements of citrus fruits.

Recently, Kondo et al. [62] and Kurita et al. [63] have developed technique for detecting rotten oranges by identifying fluorescence substance present in the orange skin. Similarly, Slaughter et al. [64] used machine vision techniques for non-destructive investigation of freeze damaged oranges and have reported that with the use of the UV fluorescence technique to detect freeze damage in oranges they were able to obtain an overall classification accuracy of about 70%. According to them, these technologies are being used in packing houses.

The US patent [65] describes a system designed to identify citrus fruits affected by any amount of rot and to determine the automatic expulsion of these fruits from the conveyor conveying them. The system comprises illuminating the fruits with UV-A band light in a computer vision unit and capturing images of the illuminated fruits by means of a camera to send them to a PC type computer, equipped with specific application software in order to detect fluorescences associated to the rot effect. The identified fruit is automatically expelled from the conveyor in an expulsion unit. The position of the defective fruit was determined with the aid of an encoder associated to the conveyor.

Researchers have opted various approaches for bruise detection. A prior knowledge of the properties of a round convex object was used to detect blemished oranges [66]. A region-oriented segmentation algorithm was tested by Blasco [67] for detecting the most common peel defects of citrus fruits. Much research has focused on the infrared [52] and hyper spectral imaging [68] for fruit grading. Multi or hyperspectral cameras permit rapid acquisition of images at many wavelengths. Imaging at fewer than ten wavelengths is generally termed multispectral, and more than ten termed hyperspectral [69].

The European patent [70] describes a fruit sorting apparatus for eliminating culls and unpackable fruit early in the pack line. This conveyor uses plastic conveyor rollers of singulators riding upon a passive spin track causing citrus fruit such as lemons carried thereon to rotate at approximately 1 rps to 4 rps, such that the axes of the fruit orient themselves substantially perpendicularly to the direction of travel of the singulators along their axes of symmetry. Later the fruits are spun up to 6 or more rps for brief periods using spin-accelerating belt without damaging the fruit and allowing the fruit to turn one complete revolution while being illuminated and scanned. Oblong fruit are basically unstable at these high spin rates and would tend to flip off the conveyor rollers if not previously oriented. The conveyor employs line scan camera, illuminators, digital computer and ejector mechanism comprising solenoid-controlled pistons.

Khoje et al [71] have developed a methodology for assessing the quality of fruits objectively using texture analysis based on Curvelet Transform. Being a multi-resolution approach, curvelets have the capability to examine fruit surface at low and high resolution to extract both global and local details about fruit surface. The researchers analyzed guava and lemon fruits and acquired the fruit images using a CCD colour camera. They used textural measures based on curvelet transform such as energy, entropy, mean and standard deviation to characterize fruits surface texture and investigated the discriminating powers of these features for fruit quality grading. They subjected acquired features to classifiers such as Support Vector Machines (SVM) and Probabilistic Neural Networks (PNN) and the performance of classifiers was tested for the two categories grading of fruits namely healthy and defected. The results showed that best SVM classification was obtained with an accuracy of 96%.

Early detection of fungal infections is most important in packing houses because even a very small number of infected fruits can spread the infection to a whole batch, thus causing great economic losses and affecting further operations, such as storage and transport [45]. Machine vision systems based only on RGB cameras are unable to detect decaying fruit correctly and the use of hyperspectral imaging makes it possible. The researchers have illuminated each fruit individually by indirect light from halogen lamps inside a hemispherical aluminium diffuser for imaging purpose.

Inspection of internal quality of the fruits is normally done with the aid of hyper-spectral imaging including ultraviolent fluorescence, reflective near-infrared radiations. Abdolabbas Jafari et al [72] have demonstrated a method that can be used for non destructive grading of orange or other citrus fruits to evaluate skin ratio of the fruit by finding out correlation between coarseness and thickness of the fruit skin with a simple and inexpensive machine vision system. They have reported that red component of the images were the best representatives of orange texture comparing to blue and green.

Recce et al [73] described a novel system for grading oranges according to their surface characteristics. The system handled fruits with a wide size range (55-100 mm), various shape (spherical to highly eccentric), surface coloration and defect markings. The stem and calyx was recognized in order to distinguish it from defects. Colour variation was recognized using a neural network classifier on rotation invariant transformations (Zernike moments). They have used separate algorithmic components to achieve high throughput and complex pattern recognition, together with state-of-the-art processing hardware and novel mechanical design. The grading was achieved by simultaneously imaging the fruit from six orthogonal directions as they were moved through an inspection chamber. In the first stage processing colour histograms from each view of an orange were analyzed using a neural network based classifier. Views that may contain defects were further analyzed in the second stage using five independent masks and a neural network classifier. The stem detection process was reported to be computationally expensive.

Blasco et al [45] have developed a machine to classify mandarin segments for canning. The system distinguished among sound, broken or double segments, and was able to detect the presence of seeds in the segments. The system analyzed the shape of the each individual segment to estimate morphological features that were used to classify it into different commercial categories. The machine classified correctly more than 75% of the analyzed segments.

2.3 Measurement of citrus fruit quality in fields

Currently, attempts are being made to measure the fruit quality in the field after harvest, for precision agriculture applications. Anna Vidal et al [42] in their investigation have stated that due to the restrictions of working in field conditions, the computer vision system equipped in such machine is limited in its technology and processing capacity, compared to conventional systems. They have tested oranges randomly chosen from the production line of a packing house whose colours varied from green and yellow with some green spots.

Kohno et al. [74] have developed a real-time in-field "Mobile citrus grading machine", which graded citrus fruits based on their size and colour using imaging technique, and also measured sugar content and acidity using on-board near infrared spectrometer.

Cubero et al [75] have developed a mobile platform in which a machine vision system was placed over the conveyor belts in order to inspect the fruit in the field and automatically separate those that can be directly sent to the fresh market and those that do not meet minimal quality criteria. The system was capable to estimate and separate the fruit based on colour, size and the presence of surface damages of individual fruits. They achieved very low energy consumption by using LED's in stroboscopic mode instead of conventional lighting using fluorescent tubes or incandescent lamps. Polarizing filters were used to avoid bright spots. They reported that better illumination was still needed by the system to properly detect the external defects of the fruits.

Palaniappan Annamalai [76] investigated a machine vision algorithm to identify and count the number of citrus fruits in an image and finally to estimate the yield of citrus fruits in a tree. He has calibrated and tested the yield mapping system in a commercial citrus grove and reported that the total time for processing an image was 119.5 ms, excluding image acquisition time. He has tested the image processing algorithm on 329 validation images and found that the R^2 value between the number of fruits counted by the fruit counting algorithm and the average number of fruits counted manually was 0.79.

Jimenez et al [77] have developed an automatic fruit recognition system to recognize spherical fruits in natural conditions facing difficult situations like shadows, bright areas, occlusions and overlapping fruits. They used a laser range-finder sensor giving range/attenuation data of the sensed surface. They developed a laser range-finder model and a dual colour/shape analysis algorithm to locate the fruit. After recognition, the 3-dimensional position of the fruit, radius and the reflectance were obtained.

A mobile grove-laboratory was developed by Harrell et al [78] to study the use of robotic technology for picking oranges under actual production conditions. They have designed a Citrus picking robot consisting of a single arm with a spherical coordinate system whose joints was actuated by servo hydraulic drives. In a small cavity at the end of the arm, rotating-lip picking mechanism was provided. This housed a CCD video camera, an ultrasonic ranging transducer to provide distance information to objects in front of the picking mechanism, light sources and the rotating lip to cut the stem of the fruit.

The Japanese company, Kubota [79] had developed a fruit-picking robot which used a mobile platform to approximate a small four degrees-of-freedom manipulator to the detachment area. The gripper had a mobile vacuum pad to capture the fruit and to direct it towards a cutting device. An optical proximity sensor, a stroboscope light and a colour camera were used and all were well protected.

3. Discussions

From the literature review, it is understood that most of the researchers have the common view that better results can be obtained with HSI colour space compared to other colour spaces; one or more colour cameras are being used for getting the colour images of citrus fruits and these image parameters are being used for the size estimation. The shapes of the fruits are determined using parameters like elongation, roundness, symmetry and compactness. Dynamic weight measurements are done using electromechanical load cell. The sweetness of the fruits is mostly determined using NIR spectroscopy. Defect inspection and the internal quality are assessed using machine vision, hyper-spectral imaging and near infrared (NIR) technologies. UV fluorescence technique is commonly used for finding the skin defects. Recently, inspection of citrus fruits in real time using Machine Vision in conjunction with multi-spectral imaging is found to give more classification accuracy. Of all the classifiers such as Support Vector Machines (SVM), Probabilistic Neural Networks (PNN), the performance of SVM classifiers is found to be more accurate.

4. Further Scope

Some researchers have predicted the volume of citrus fruits using image processing techniques based on single or multiple view fruit images and the shape-based analytical models [33],[53],[80]. If the weight of the fruit is measured online using load-cell or by some other method, then using this estimated volume it is easy to calculate the average fruit density of that particular batch of fruits, which may be useful for identifying the presence of hidden defects and in predicting the fruit maturity.

Mathematical modeling can be developed for estimating the mass of the fruits using fruit dimensions assuming the fruit densities to be constant for a particular batch of fruits, thereby replacing the need for a weighing device and eliminating the associated hardware components.

Fourier based shape classification technique can be used to analyze/determine the shape of the fruit. Also features such as area, eccentricity and central moments can be used to discriminate between different shapes.

Linear Discriminant Analysis (LDA) or Bayesian Discriminant Analysis in combination with a Mahalanobis distance classifier can be tried to classify the fruits based on colour. Data reduction can be done using Principal Component Analysis or Partial Least Squares coupled with multiple regression or using ANN (Artificial Neural Networks) or Wavelets or Recursive Classification Trees instead.

From the study it is also understood that many sorting and grading equipments developed use separate sections for feeding, imaging, weighing and sorting which involves careful handling during transfers from one section to another section. Only a couple of sorters and graders have been reported in the literature which uses single conveying arrangement for the entire operations. Future machine vision system may be designed around a single conveying arrangement with minimum transfer operations to improve the handling, reduce fruit damage, increase the throughput and reduce the floor space requirement. The design may be made economically to satisfy the needs of all smaller horticultural units.

5. Conclusions

Study reveals that many researchers have developed many kinds of citrus fruit graders and sorters according to market requirement which are highly challenging. Generally grading of these fruits is done on the basis of many internal and external quality parameters like size, shape, colour, weight, defects, etc. This paper explores some of the technologies of image analysis and machine vision in the development of automated machine in fruit industries. The study aims to investigate the potential of various techniques like machine vision, infra-red, optical, robotics, mechatronics, etc in quantifying the quality parameters of citrus fruits, like colour, size, shape, etc. and explore the possibilities towards development of improved and simplified machine vision systems to augment the fruit inspection capabilities for inspection of local varieties of citrus fruits like orange, mosambi and lemon. The review also covers the measurement of citrus fruit quality in the harvesting fields.

6. Acknowledgement

The authors sincerely thank the Director, CSIR-CEERI and the Scientist-in-Charge, CSIR-CEERI Chennai Centre for their valuable support.

Bibliography

[1] Dattatraya Londhe, Sachin Nalawade, Ganesh Pawar, Vinod Atkari and Sachin Wandkar, Sep 2013. “Grader - A review of different methods of grading for fruits and vegetables,” Agricultural Engineering International: CIGR Journal, vol.15, No.3, Page 217-230.

[2] Krishna Kumar Patel, Kar. A., Jha. S. N & Khan M. A., April 2012. **“**Machine vision system: a tool for quality inspection of food and agricultural products,” Journal of Food Science Technology, 49(2): Page 123–141.

[3] Park B, Chen YR, Nguyen M, Hwang H, 1996. “Characterising multispectral images of tumorous, bruised, skin-torn, and wholesome poultry carcasses,” Trans ASAE 39(5):1933–1941

[4] Morrow CT, Heinemann PH, Sommer HJ, Tao Y, Varghese Z, 1990. “Automate inspection of potatoes, apples, and mushrooms,” In Proceedings of the International Advanced Robotics Programme, Avignon, France: 179–188

[5] Gerrard, D. E, Gao, X and Tan, J, 1996. “Beef marbling and color score determination by image processing,” Journal of Food Science, 61(1), 145-148.
[6] Gunasekaran S, 1996. “Computer vision technology for food quality assurance,” Trends in Food Science and Technology, 7(8):245–256. doi: 10.1016/0924-2244(96)10028-5.
[7] Md. Iqbal S, Ganesan, Sridhar R, Gopal A and Sarma A.S.V, 2009. “Conveying Systems for Machine Vision Inspection of some Fruits and Vegetables – A Review,” National Conference on Virtual & Intelligent Instrumentation (NCVII-09), Pilani
[8] Structure and Function of Food Engineering, Edited by Ayman Amer Eissa, ISBN 978-953-51-0695-1, 404 pages, Publisher: InTech, Chapters published August 22, 2012, Chapter 10: Understanding Colour Image Processing by Machine Vision for Biological Materials
[9] Food and Agriculture Organization of the United Nations [FAO], 1989, “Prevention of postharvest food losses: fruits, vegetables and root crops,” FAO training series no.17/2, pp.157.
[10] Reyes, M. U. (Ed.), 1988, “Design Concept and Operation of ASEAN Packinghouse Equipment for Fruits and Vegetables,” Postharvest Horticulture Training and Research Center, University of Los Banos, College of Agriculture, Laguna, Philippines.
[11] Umeda S, 1976, “Automatic sorting of agricultural products,” Journal of the Japanese Society of Agriculture Machinery, 38(3): 345-351.
[12] Farmco, Inc., USA – Product catalogue
[13] Greefa, Newzealand – Product catalogue
[14] Stroshine, R, and Hamann, D. D, 1994, “Physical Properties of Agricultural Material and Food Products,” Purdue University, West Lafayette, Indiana, Course manual.
[15] Tabatabaeefar, A, Vefagh-Nematolahee, A and Rajabipour, A, 2000. "Modeling of orange mass based on dimensions," Journal of Agricultural Science and Technology 2, no. 4 (2000): 299-305.
[16] Anonymous, 1989, “Development of orange size grader,” Quoted from Annual Research Report 1988-89 for presentation to the research review Sub-Committee of Agril. Process Engg., Dr. PDKV. Akola: pp. 63-67.
[17] Nevkar, G. S, 1990, “Development and performance testing of divergent roller type fruit sorting machine,” M. Tech. thesis, MPKV, Rahuri.
[18] Ghuman B. S, and Kumar. A, 2005, “Development of low cost rotary disc size grader for fruits and vegetables,” Journal of Research, 42 (4): 497-503.
[19] Mangaraj, S, Varshney, A. C, Reddy, B. S. and Singh. K. K, 2005, “Development of a stepwise expanding pitch fruit grader,” Journal of Agricultural Engineering, 42(3): 74-79.
[20] Burt, S. W. and Patchen G, 1966, “Grading and sizing apples with brushes,” Agricultural Research Service, U.S. Dept. of Agriculture, 52-58, pp. 86.
[21] Bryan, W. L, Jenkins. J and Miller. J.M, 1980, “Mechanically assisted grading of oranges containing excessive decayed fruits,” Trans. of the ISA, 23(1): 247-250.
[22] Molto E, Blasco J, Steinmetz V, Bourley A, Navarron F, Pertotto G, 1997, “A robotics solution for automatic handling, inspection and packing of fruits and vegetables,” Proceedings of the International Workshop on Robotics and Automated Machinery for Bioproductions.BioRobotics’97, Grandia, Valencia, Spain
[23] Kondo, N., 2009, “Automation on fruit and vegetable grading system and food traceability”, Trends in Food Science & Technology, doi: 10.1016/j.tifs.
[24] Sun, D.W., 2003, “Computer Vision: An Objective, Rapid and Non-Contact Quality Evaluation Tool for the Food Industry,” Journal of Food Engineering, 61: 1-2
[25] Colour Vision systems (CVS), Australia – Product catalogue
[26] Aweta Autoline Systems, USA. – Product catalogue
[27] Ahmad U, Mardison S, Tjahjohutomo R and Nurhasanah A, 2010, “Development of automatic grading machine prototype for citrus using image processing,” Australian Journal of Agricultural Engineering, 1(5): 165-169.
[28] Khojastehnazhand M, Omid M, Tabatabaeefar A, 2010, “Development of a lemon sorting system based on colour and size,” African Journal of Plant Science 4(4):122–127.
[29] Tao, Y, 1996, “Methods and apparatus for sorting objects including stable colour transformation,” US Patent 5:533–628

[30] Majumdar, S. and Jayas, DS, 2000, "Classification of cereal grains using machine vision: IV. Morphology, colour, and texture models," Trans ASAE 43(6):1689–1694
[31] Wang, TY., Nguang, SK , 2007, "Low cost sensor for volume and surface computation of axi-symmetric agricultural products," Journal of Food Eng 79:870–877
[32] Eifert JD, Sanglay GC, Lee DJ, Sumner SS, Pierson MD, 2006, "Prediction of raw produce surface area from weight measurement," Journal of Food Eng 74:552–556
[33] Khojastehnazhand, M., Omid M, and Tabatabaeefar A, 2009. "Determination of orange volume and surface area using image processing technique," Int. Agrophysics 23 (2009): 237-242.
[34] Anonymous, 1980, "Electronic colour sorting of fruits," International fruit world. FSTA 1969-12/95, AN81-11-J1679, 38(3): 263-267.
[35] Lino ACL, Sanches J, Fabbro IMD, "Image processing techniques for lemons and tomatoes classification," Bragantia campinas, 2008;67(3):785–789. doi: 10.1590/S0006-87052008000300029.
[36] Gardner JL, 2007. "Comparison of calibration methods for tristimulus colorimeters," Journal of Research of the National Institute of Standards and Technology, 112, 129-138.
[37] Yam KL, Papadakis SE, 2004, "A simple digital imaging method for measuring and analyzing color of food surfaces," Journal of Food Engineering, 61,137–142.
[38] Gaffney J.J, 1969. "Reflectance properties of citrus fruit," Transactions of the ASAE 16(2), 310-314.
[39] Slaughter D. and Harrel R. C, 1987. "Color Vision in Robotic Fruit Harvesting", Transactions of the ASAE, Vol.30(4), 1144-1148.
[40] Slaughter D. and Harrel R. C, 1989. "Discriminating fruit for robotic harvest using color in natural outdoor scenes," Transactions of the ASAE, Vol.32(2), 757-763.
[41] Blasco J, Aleixos N & Molto E, 2007. "Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm," Journal of Food Engineering, 81, 535–543.
[42] Anna Vidal, Jose Blasco and Pau Talens, 2013. "Measurement of Colour of Citrus Fruits using an Automatic Computer Vision System," Master in Science and Engineering of Food Thesis, http://hdl.handle.net/10251/27912
[43] Warkentin, Aaron J, and George A. Mills. "Sorter for fruit and the like," U.S. Patent 4,106,628, issued August 15, 1978.
[44] Kondo, N., Ahmad, U., Monta, M., and Murase H, 2000, "Machine vision based quality evaluation of Iyokan orange fruit using neural networks," Computers and Electronics in Agriculture, 29, 135–147.
[45] Blasco, Jose, Nuria Aleixos, Sergio Cubero, Florentino Juste, Juan Gomez-Sanchis, Vicente Alegre, and Enrique Molto, 2009. "Computer vision developments for the automatic inspection of fresh and processed fruits," First International Workshop on Computer Image Analysis in Agriculture, pp. 21-24.
[46] Bennedsen, B.S., Peterson D.L, and Tabb A, 2005. "Identifying defects in images of rotating apples," Comput. Electron. Agr. 48(2): 92-102.
[47] Leemans, V., Magein, H., and Destain, M.F. 2002. "On-line fruit grading according to their external quality using machine vision," Biosyst. Eng. 83(4): 397-404.
[48] Kondo, N., Monta M, Noguchi N, Yukumoto O, Matsuo Y, Ogawa Y, Fujiura T and Arima S, 2006. "Agri-Robot (II)-Mechanisms and Practice," 26-27 & 58-59.
[49] Ruiz LA, Molto E, Juste F, Pla F, Valiente R, 1996, "Location and characterization of the stem-calyx area on oranges by computer vision," Journal of Agric Eng Res 64(3):165–172
[50] Kondo N, 1995, "Quality evaluation of orange fruit using neural networks," In: Food Processing Automation IV Proceedings of the FPAC Conference. ASAE, 2950 Niles Road, St. Joseph, MI 49085–9659, USA
[51] Molto E, Blasco J, 2008, "Quality evaluation of citrus fruits," Computer vision technology for food quality evaluation, p 247
[52] Aleixos N., Blasco J., Navarron F., Molto E., 2002. "Multispectral inspection of citrus in real-time using machine vision and digital signal processors," Computers and Electronics in Agriculture 33(2002): 121-137
[53] Omid, M., Khojastehnazhand, M., and Tabatabaeefar, A., 2010. "Estimating volume and mass of citrus fruits by image processing technique," *Journal of Food Engineering*, 100 (2010) 315-321.

[54] Naoshi, K, Takahisa N, Yoshihid M, Peter L, Mitsutaka K, Makoto K, Paolo D.F and Yuichi O, 2008. “A double image acquisition system with visible and UV LEDs for citrus fruit,” American Society of Agricultural and Biological Engineers, ASABE.

[55] Cerruto E, Failla S, Schillaci G, 1996. “Identification of blemishes on oranges,” International conference on agricultural engineering, AgEng’96, Madrid, Spain, Eur Agric Eng paper no. 96 F-017, Sept, 23–26

[56] Ying YB, Xu ZG, Fu XP, Liu YD, 2004. “Nondestructive maturity detection of Citrus with computer vision,” In Proceedings of the Society of Photo-optical Instrumentation Engineers (SPIE), 5271, Monitoring Food Safety, Agriculture, and plant Health, p 97–107

[57] Kawano, S., Fujiwara K, and Iwamoto M, 1993. “Nondestructive determination of sugar content in Satsuma mandarin using near infrared (NIR) transmittance,” Journal of Japanese Society of Horticultural Science, 62 (2): 465-470.

[58] Yamakawa M, Khot L. R, Ehsani R, Kondo N, September, 2012. “Real-time nondestructive citrus fruit quality monitoring system: development and laboratory testing,” Agricultural Engineering International: CIGR Journal Open access at http://www.cigrjournal.org Vol. 14, No.3, pp 117-124.

[59] Lu, H. S, Xu, H. R, Ying, Y. B, Fu, X. P, Yu, H. Y, & Tian, H. Q, (2006). “Application Fourier transform near infrared spectrometer in rapid estimation of soluble solids content of intact citrus fruits,” Journal of Zhejiang University SCIENCE B, 7(10), 794-799.

[60] Liu, Y, Sun X, Zhou J, Zhang H, and Yang C, 2010. “Linear and nonlinear multivariate regressions for determination sugar content of intact Gannan navel orange by Vis-NIR diffuse reflectance spectroscopy,” Mathematical and Computer Modeling, 51 (11-12): 1438-1443.

[61] Antonucci, F, Pallottino F, Paglia G, Palma A, D’Aquino S, and Menesatti P, 2011. “Non-destructive estimation of mandarin maturity status through portable VIS-NIR spectrophotometer,” Food Bioprocess Technology, 4 (5): 809-813.

[62] Kondo, N, Kuramoto M, Shimizu H, Ogawa Y, Kurita M, Nishizu T, Kiong Chong V, and Yamamoto K, 2009. “Identification of fluorescent substance in mandarin orange skin for machine vision system to detect rotten citrus fruits,” Engineering in agriculture, Environment and Food, 2 (2): 54-59.

[63] Kurita, M, Kondo N, Shimizu H, Ling P. P, Falzea P. D, Shiigi T, Ninomiya K, Nishizu T, and Yamamoto K, 2009. “A double image acquisition system with visible and UV LEDs for citrus fruit. Journal of Robotics and Mechatronics,” 21 (4): 533-540.

[64] Slaughter, D. C, Obenland D. M, Thompson J. F, Arpaia M. L, and Margosan D. A, 2008. “Non-destructive freeze damage detection in oranges using machine vision and ultraviolet fluorescence,” Postharvest Biology and Technology, 48 (3): 341-346.

[65] Blanc, Philippe Gabriel Rene, Jose Blasco Ivars, Enrique Molto Garcia, Juan Gomez Sanchis, and Sergio Cubero Garcia. “System for the automatic selective separation of rotten citrus fruits,” U.S. Patent Application 12/312,160, filed February 26, 2008.

[66] MingHui Liu, Gadi Ben Tal,Napoleon H Reyes, Andre L C Barczak, 2009. “Navel Orange Blemish Identification for Quality Grading System,” Neural Information Processing Lecture Notes in Computer Science, 5864 (2009) 675-682.

[67] Blasco J, Aleixos N, Molto E, 2007. “Computer vision detection of peel defects in citrus by means of a region oriented segmentation algorithm,” Journal of Food Engineering, 81(2007) 535-543.

[68] Peirs A, Scheerlinck N, De Baerdemaeker J, Nicolai B M, “Quality determination of apple fruits with a hyperspectral imaging system,” AgEng 02, Budapest, Hungary.EurAgEng Paper No. 02-PH-028.

[69] Abbott, Judith A. 1999. "Quality measurement of fruits and vegetables," Postharvest Biology and Technology 15, no. 3 (1999): 207-225.

[70] Mills, George A. "Apparatus for spinning fruit for sorting thereof," European Patent EP 0105114, issued April 11, 1984.

[71] Khoje, Suchitra A, Bodhe, S. K, and Adsul, A, Aug-Sep 2013. “Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet Transform,” International Journal of Engineering and Technology (0975-4024), Vol 5 No 4, pp 3251-3256.

[72] Abdolabbas Jafari, Mohammad Reza Zarezadeh, Atefeh Fazayeli, “Orange Grading Based on Visual Texture Features,” International Conference of Agricultural Engineering, CIGR-AgEng 2012, July 8-12, Valencia, Spain.
[73] Recce M., Taylor J, Piebe A, and Tropiano G, 1996. "High speed vision-based quality grading of oranges," In Neural Networks for Identification, Control, Robotics, and Signal/Image Processing. Proceedings, International Workshop on, pp. 136-144. IEEE, 1996.
[74] Kohno Y, Kondo N, Iida M, Kurita M, Shiigi T, Ogawa Y, Kaichi T, and Okamoto S, 2011. “Development of a mobile grading machine for citrus fruit,” Engineering in agriculture, Environment and Food, 4 (1): 7-11.
[75] Cubero S, Molto E., Aleixos N, Garcia-Navarrete O.L, Gutierrez A, Juste F, Blasco J, August 2010, “Machine Vision System for the In-line Inspection of Fruit on a Mobile Harvesting Platform in Field Conditions,” CIGR Workshop on Image Analysis in Agriculture, 26-27, Budapest.
[76] Palaniappan Annamalai, August 2004. “Citrus Yield Mapping System using Machine Vision,” Thesis Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Master of Science.
[77] Jimenez, A. R, Jain A. K, Ceres R, and Pons J. L, 1999. "Automatic fruit recognition: a survey and new results using range/attenuation images." Pattern recognition 32, no. 10 (1999): 1719-1736.
[78] Harrell, R. C., P. D. Adsit, T. A. Pool, and R. Hoffman, 1988. "The Florida robotic grove-lab," Transactions of the ASAE Paper No. 88-1578.
[79] Hayashi and Ueda, 1991. “Orange harvesting robot,” Kubota Co., Sakai, Japan (Mimeo).
[80] Md. Iqbal, S, Gopal, A and Sarma, A.S.V, 2011, "Volume Estimation of Apple Fruits using Image Processing," Proceedings of International Conference on Image Information Processing (Published online in IEEE Explorer; Object Identifier: 10.1109 / ICIIP. 2011. 6108909)

Performance of dispatching rules in a stochastic dynamic job shop manufacturing system with sequence-dependent setup times

Pankaj Sharma[1*], Ajai Jain[2]

[1*]Department of Mechanical Engineering, National Institute of Technology, Kurukshetra, India. Pin code: 136119, Phone Number: +919466206231(mob), Email Id.: pankajsharmagju@gmail.com

[2] Department of Mechanical Engineering, National Institute of Technology, Kurukshetra, India. Pin code 136119, Phone Number: +919412558040(mob), Email Id.: ajayjainfme@nitkkr.ac.in

Abstract

Stochastic dynamic job shop scheduling problem with consideration of sequence-dependent setup times are among the most difficult classes of scheduling problems. This paper assesses the performance of five dispatching rules in such shop from mean flow time, mean tardiness and number of tardy jobs performance measures viewpoint. A discrete event simulation model of a stochastic dynamic job shop manufacturing system is developed for investigation purpose. Five dispatching rules i.e. first come first serve (FCFS), shortest processing time (SPT), shortest setup time (SIMSET), earliest due date (EDD) and shortest sum of processing time and setup times (SPT+SIMSET) are incorporated in the simulation model. The simulation experiments are conducted under due date tightness factor of 3, shop utilization percentage of 90 and setup times less than processing time. Results indicate that SIMSET rule provides better performance for mean flow time and number of tardy jobs performance measures. The SPT rule provides better performance for mean tardiness measure.

Keywords: scheduling; stochastic dynamic job shop; sequence-dependent setup times; dispatching rule; simulation

1. Introduction

In a manufacturing system, production scheduling is associated with allocation of set of jobs on a set of production resources over time to achieve some objectives. In a job shop, a set of jobs are processed on a set of machines and each job has specific operation order. The job shop scheduling problem is a combinatorial optimization problem as well as NP-hard and it is one of the most typical and complex among various production scheduling problems [1, 2]. Dynamic job shop scheduling problem represents the environment in which jobs arrive continuously over time in the manufacturing system. Further, in a stochastic dynamic job shop (SDJS) manufacturing system at least one parameter of the job i.e. release time, processing time or setup time is probabilistic [17, 18].

In traditional approaches, in order to reduce the complexity of solving job shop scheduling problems, setup time is either neglected or included in the processing time. But this effort does not represent the realistic picture of the manufacturing system. Setup time is a time that is required to prepare the necessary resources such as machines to perform a task [3]. In many real-life situations, a setup operation often occurs while shifting from one operation to another. Sequence-dependent setup time depends on both current and immediately preceding operation [3]. Sequence-dependent setup time is encountered in many industries such as printing industry, paper industry, auto industry, chemical processing and plastic manufacturing industry. Scheduling problems with sequence-dependent setup times are among the most difficult classes of scheduling problems [4]. Manikas and Chang [5] and Fantahun and Mingyuan [6] stated that limited research on job shop scheduling problems with sequence-dependent setup times is available.

A dispatching rule selects the next job to be processed from the set of jobs awaiting processing in the input queue of a machine. Dispatching rules are also termed as scheduling rules or sequencing rules. Dispatching rules are classified into broad four categories namely as process time based rules, due date based rules, combination rules and rules that are neither process time based nor due date based [7]. This paper focuses the performance of five dispatching rules identified from literature in a SDJS manufacturing system with consideration of sequence-dependent setup times.

The remainder of the paper is organised as follows. The review of relevant literature is introduced in section 2. Section 3 describes salient aspects of configuration of the SDJS scheduling problem. The outline for development of simulation model is explained in section 4. Section 5 presents details of simulation experimentations. Section 6 provides analysis of experimental results. Finally, section 7 gives concluding remarks and directions for future work.

2. Literature review

Ramasesh [8] provided survey of simulation research in dynamic job shop scheduling. Allahverdi et al. [9] provided a comprehensive review of literature on scheduling problems with setup times (costs). Panwalkar et al. [10] provided a review of scheduling rules used in manufacturing systems. Blackstone et al. [11] presented a state-of-the-art survey of scheduling rules used in job shop manufacturing system. Holthaus and Rajendran [12] proposed two new dispatching rules for a dynamic job shop manufacturing system to minimize mean flow time, mean tardiness and percentage of tardy jobs performance measures. These rules combine Process Time and Work Content in Queue for the next operation on a job by making use of additive (Rule 1) and alternative approaches (Rule 2).

They concluded that Rule1 is quite effective in minimizing mean flow time performance measure. Jayamohan and Rajendran [13] proposed seven dispatching rules for minimization of various performance measures i.e. mean flow time, maximum flow time, variance of flow time and tardiness in dynamic shops. The proposed rules are found to be effective in minimizing different performance measures. Jain et al. [14] proposed and assessed the performance of four new dispatching rules for makespan, mean flow time, maximum flow time and variance of flow time measures in a flexible manufacturing system. Authors found that the proposed dispatching rules provides better performance than the existing rules. Dominic et al. [15] developed two better scheduling rules viz. longest sum of Work Remaining and Arrival Time of a job (MWRK_FIFO) and shortest sum of Total Work and Processing Time of a job (TWKR_SPT) for a dynamic job shop manufacturing system. These rules are tested against the existing scheduling rules i.e. First in First out (FIFO), Last in First out (LIFO), Shortest Processing Time (SPT), Longest Processing Time (LPT), Most Work Remaining (MWRK) and Total Work (TWKR) for mean flow time, maximum flow time, mean tardiness, tardiness variance and number of tardy jobs performance measures.

There have been a few attempts to address dynamic job shop scheduling problems with sequence-dependent setup times. To the best of authors's knowledge, Wilbrecht and Prescott [16] were first among researchers to study the influence of setup times on dynamic job shop manufacturing systems performance. They proposed and tested a setup oriented scheduling rule, job with Smallest Setup Time (SIMSET). Authors concluded that SIMSET rule outperforms other existing scheduling rules i.e. Random, Earliest Due Date, Shortest Run, Longest Run, Shortest Process and Longest Process for value of work-in-progress, number of processes completed in a week, number of jobs sent out of the shop in one week, number of processes completed late in one week, distribution of completion times, queue wait time of a job in a shop, number of jobs waiting in a shop, shop capacity utilized, number of jobs waiting in a queue for more than one week and size of jobs waiting in a queue for more than one week performance measures. Kim and Bobrowski [17] studied impact of sequence-dependent setup times on the performance of a dynamic job shop manufacturing system. They concluded that setup oriented scheduling rules i.e. SIMSET and job with similar setup and Critical Ratio (JCR) outperforms ordinary scheduling rules i.e. Shortest Processing Time (SPT) and Critical Ratio (CR) for mean flow time, mean work-in-process inventory, mean finished good inventory, mean tardiness, proportion of tardy jobs, mean machine utilization, mean setup time per job, mean number of setups per job and mean total cost per day performance measures when a manufacturing system with sequence-dependent setup times is considered. Kim and Bobrowski [18] extended their previous work [17] to investigate impact of setup times variation on sequencing decisions with normally distributed setup times. They concluded that setup times variation has a negative impact on system performance. Recently, Vinod and Sridharan [19] proposed and assessed performance of five setup oriented scheduling rules viz. shortest sum of Setup Time and Processing Time (SSPT), job with similar setup and Shortest Processing Time (JSPT), job with similar setup and Earliest Due Date (JEDD), job with similar setup and Modified Earliest Due Date (JMEDD) and job with similar setup and shortest sum of Setup Time and Processing Time (JSSPT) for dynamic job shop scheduling problems with sequence-dependent setup times. They concluded that proposed rules provides better performance than the existing scheduling rules i.e. First in First out (FIFO), Shortest Processing Time (SPT), Earliest Due Date (EDD), Modified Earliest Due Date (MEDD), Critical Ratio (CR), Smallest Setup Time (SIMSET) and job with similar setup and Critical Ratio (JCR) for mean flow time, mean tardiness, mean setup time and mean number of setups performance measures.

Literature review clearly reveals that there is a need to evaluate the performance of dispatching rules in a SDJS manufacturing system with sequence-dependent setup times. The present paper is an attempt in this direction. It assesses performance of existing five best performing dispatching rules identified from literature using simulation modeling for mean flow time, mean tardiness, and number of tardy jobs performance measures in a SDJS manufacturing system with sequence-dependent set up times.

3. Job shop configuration

In the present study, a job shop manufacturing system with ten machines is selected. The configuration of the manufacturing system is determined based on configuration of job shop considered by various researchers [12, 19]. It is pointed out by researchers that six machines are sufficient to represent the complex structure of a job shop manufacturing system [16, 20] and job shop size variations don't significantly affect the relative performance of dispatching rules [12, 20]. For the same reason, most of the researchers addressed a job shop scheduling problem with less than ten machines [15, 21, 22].

3.1. Job data

Six different types of jobs i.e. job type A, job type B, job type C, job type D, job type E and job type F arrive at the manufacturing system. All the job types have equal probability of arrival. Job types A, B, C, D, E and F require 5, 4, 4, 5, 4 and 5 operations respectively. The machines visited by different job types in their routes are shown in **Table 1**. The processing times and setup times of each job are stochastic and assumed to be uniformly distributed on each machine. Processing time changes according to job type and route of the job. **Table 2** presents processing times of each job on the machines according to their routes. The pattern of processing times on different machines is selected based on research work carried out by previous researcher [23]. Sequence-dependent setup times which encounters while shifting from one job type to another are given in **Table 3.**

3.2. Inter-arrival time

It is average time between arrivals of two jobs. The average arrival rate of jobs must be selected to have utilization of the machine less than 100%. Otherwise, the number of jobs in the queues in front of each machine will grow without bound [24]. Thus, inter-arrival time of the jobs is established using percentage utilization of the shop and processing requirements of the jobs. It has been observed in the literature that arrival process of the jobs follows a poisson distribution [8, 19, 24]. Thus, inter-arrival time is exponentially distributed. Mean inter-arrival time of the jobs is calculated using the following relationship [19, 21].

$$b = \frac{1}{\lambda} = \frac{\mu_p \mu_g}{UM} \qquad (1)$$

Here

b=Mean inter-arrival time

λ=Mean job arrival rate

μ_p =Mean processing time per operation (including setup time)

μ_g =Mean number of operations per job

U=Shop utilization

M=Number of machines in the shop

In the present work, μ_p is computed by taking the mean of mean processing times of all operations (from Table 2) plus mean of mean setup times (from Table 3). Thus, μ_p =19.

45. For the taken input data, μ_g is 4.5 with M=10. In the present work, experiments are carried out at shop utilization (U) = 90%. Van Parunak [25] observed that due to stochastic nature of input processes (processing times and setup times) actual shop load is approximated and fall within a range of ± 1.5% of the target value.

3.3. Due date of jobs

It is time at which job order must be completed. The due date of the arriving job is either externally or internally determined. In case of externally determined due date, due date is either established by the customer or set for a specific time in the future. In case of internally determined due date, due date is based on total work content (sum of processing times and setup times) of the job or number of operations to be performed on the job. Most of the researchers used total work content (TWK) method to assign due date of the job [12, 19, 21, 26].

$$d_i = a_i + k(p_i + n_i \times u_i) \qquad (2)$$

Here

d_i = Due date of job i

a_i = Arrival time of job i

k = Due date tightness factor

p_i =Mean total processing times of all the operations of job i

n_i = Number of operations of job i

u_i = Mean of mean setup times of all the changeover of job i

In the present study, due date tightness factor (k) = 3 is considered.

4. Structure of simulation model

The simulation modeling is one of the most powerful techniques available for studying large and complex manufacturing systems. In the present study, a discrete event simulation model for the operations of SDJS manufacturing system with each dispatching rule is developed using PROMODEL software. The job flow in the modeled SDJS manufacturing system is shown in **Figure 1**. The following assumptions are made while developing simulation model.

1. Each machine can perform at most one operation at a time on any job.
2. An operation cannot start until its predecessor operation is completed.
3. The arrival of jobs in the shop floor is dynamic. A type of job is unknown until it arrives in the shop.
4. Buffer of unlimited capacity is considered before and after each machine.
5. Processing times and setup times are stochastic and known in priori with their distribution.
6. The alternate routings for processing of the jobs is not available. A particular type of operation of the job can be performed by only a particular type of machine.

In the present study, a conceptual model of a job shop manufacturing system is developed. In order to ensure that the simulation model is correctly developed, a multilevel verification exercise is performed. For this, the simulation model is debugged and internal logics are checked. The output obtained from simulation model is compared with that obtained from a manual exercise by using same input data. Finally, the simulation model is run under different settings in order to check that the model behaves in a logical manner.

4.1. Dispatching rules

A dispatching rule (DRL) selects the next job to be processed on the machine from a set of jobs waiting in the input queue of the machine. In the present work, the following dispatching rules as identified from the literature are used to make job sequencing decision [16, 19].

(1) First Come First Serve (FCFS): The job which arrives first in the input queue of the machine is selected for processing.

(2) Shortest Processing Time (SPT): The job with shortest processing time for the imminent operation is selected for processing.

(3) Shortest Setup Time (SIMSET): The job with shortest setup time for the imminent operation is selected for processing.

(4) Earliest Due Date (EDD): The job to earliest due date is selected for processing.

(5) SSPT : Shortest (Setup Time + Processing Time)

The job with shortest value of the sum of Setup Time and Processing Time is selected for processing.

4.2. Performance measures

The following performance measures are used for evaluation purpose in the experimental investigations:

(1) Mean flow time ($\overline{F}$): It is average time that a job spends in the shop during processing.

$$\overline{F} = \frac{1}{n}[\sum_{i=1}^{n} F_i] \quad (3)$$

Here

$F_i = c_i - a_i$

F_i=Flow time of job i

c_i =Completion time of job i

a_i =Arrival time of job i

n=Number of jobs produced during simulation period (during steady state period)

(2)Mean tardiness ($\bar{T}$): It is average tardiness of a job in the shop during processing.

$$\bar{T} = \frac{1}{n}[\sum_{i=1}^{n} T_i] \qquad (4)$$

Here

$T_i = max\{0, L_i\}$

$L_i = c_i - d_i$

T_i=Tardiness of job i

L_i=Lateness of job i

d_i =Due date of job i

(3) Number of tardy jobs (*NTJ*): It is value of the number of jobs which are completed after their due dates.

$$NTJ = \sum_{i=1}^{n} \delta(J_i) \qquad (5)$$

Here, $\delta(J_i) = 1$ if $J_i > 0$ and $\delta(J_i) = 0$, otherwise.

5. Experimental design for simulation study

Using simulation modeling, a number of experiments on SDJS scheduling problem have been conducted. The first stage in simulation experimentation is to identify steady state period i.e. end of the initial transient period. The Welch's procedure as described by Law and Kelton [27] is used for this purpose. A pilot study for SDJS manufacturing system is conducted with FCFS dispatching rule. Thirty replications are considered for simulation experimentation. The simulation for each replication is made to run for 20000 jobs completion. It is observed that the manufacturing system reaches steady state at the completion of 5000 jobs. Finally, the experimental investigation is carried out to assess the performance of five dispatching rules identified from literature in a SDJS manufacturing system for 20000 jobs completion (after warm up period of 5000 jobs).

6. Results and discussion

In SDJS manufacturing system, the performance of five dispatching rules identified from literature is assessed. For each performance measure under each dispatching rule, the simulation output of 30 replications is averaged. The average values of various performance measures are shown in **Figures 2-4**.

6.1. Mean flow time

The performance of different dispatching rules for mean flow time measure is shown in **Figure 2**. It indicates that SIMSET rule is best performing dispatching rule for mean flow time performance measure. This is followed by SPT+SIMSET, SPT, EDD, and FCFS dispatching rules in that order. Thus, SIMSET rule is best performing dispatching rule for mean flow time performance measure when a stochastic dynamic job shop scheduling problem with sequence-dependent setup times is considered.

6.2. Mean tardiness

This is due date based performance measure and related to better customer service and satisfaction. **Figure 3** shows the performance of various dispatching rules for mean tardiness measure. It clearly indicates that SPT rule is best performing dispatching rule and it is followed by SIMSET, SPT+SIMSET, EDD, and FCFS dispatching rules in that order. Thus, SPT dispatching rule outperforms other dispatching rules for mean tardiness performance measure.

6.3. Number of tardy jobs

The performance of different dispatching rules for number of tardy jobs measure is shown in **Figure 4**. This figure indicates that SIMSET dispatching rule provides best performance for number of tardy jobs measure. This is followed by other dispatching rules i.e. SPT, SPT+SIMSET, EDD and FCFS in that order in minimizing number of tardy jobs performance measure.

7. Conclusions

The present work addresses a SDJS scheduling problem while considering sequence-dependent setup times. The performance of five dispatching rules taken from literature is assessed. The experimental results indicate that SIMSET rule is best performing dispatching rule for mean flow time, and number of tardy jobs performance measure. The SPT rule provides best performance for mean tardiness measure.

The present work can be extended in several ways. Further experimental work is required to address SDJS scheduling problems with sequence-dependent setup times and involving situations like limited capacity buffer between machines, machine breakdown, batch mode schedule and external disturbances such as order cancellation and job pre-emption. The development of better dispatching rules is both essential and desirable.

References

[1] Garey MR, Johnson DS, Sethi R. The complexity of flow shop and job shop scheduling. Mathematics of Operations Research 1976; 1(2): 117-129.

[2]Xiong J, Xing LN, Chen YW. Robust scheduling for multi-objective flexible job-shop problems with random machine breakdowns. International Journal of Production Economics 2013; 141(1): 112-126.

[3]Ali A, Soroush HM. The significance of reducing setup time/setup costs. European Journal of Operational Research 2008; 187 (3): 978-984.

[4]Pinedo M. Scheduling theory, algorithms and systems. Englewood cliffs: Prentice-Hall; 1995.

[5]Manikas A, Chang YL. Multi-criteria sequence-dependent job shop scheduling using genetic algorithms. Computers and Industrial Engineering 2009; 56 (1): 179-185.

[6]Fantahun MD, Mingyuan C. Job shop lot streaming with routing flexibility, sequence-dependent setups, machine release dates and lag time. International Journal of Production Research 2012; 50 (8): 2331-2352.

[7]Haupt R. A survey of priority rule-based scheduling. Operations Research Spektrum 1989; 11 (1): 3-16.

[8]Ramasesh R. Dynamic job shop scheduling: a survey of research. Omega: International Journal of Management Science 1990; 18 (1): 43-57.

[9]Allahverdi A, Gupta JND, Aldowaisan T. A review of scheduling research involving setup considerations. Omega 1999; 27(2): 219–239.

[10]Panwalkar SS, Iskander W. A survey of scheduling rules. Operations Research 1977; 25 (1): 45-61.

[11]Blackstone JH, Philips DT, Hogg GL. A state-of-the-art survey of dispatching rules for manufacturing job shop operations. International Journal of Production Research 1982; 20 (1): 27-45.

[12] Holthaus O, Rajendran C. New dispatching rules for scheduling in a job shop-An experimental study. The International Journal of Advanced Manufacturing Technology 1997; 13(2): 148-153.

[13] Jayamohan MS, Rajendran C. New dispatching rules for shop scheduling: A step forward. International Journal of Production Research 2000; 38(3): 563-586.

[14]Jain A, Jain PK, Singh IP. An investigation on the performance of dispatching rules in FMS scheduling. International journal of simulation modeling 2004; 3(2-3): 49-60.

[15]Dominic PDD, Kaliyamoorthy S, Saravana Kumar M. Efficient dispatching rules for dynamic job shop scheduling. International Journal of Advanced Manufacturing Technology 2004; 24 (1-2): 70-75.

[16]Wilbrecht JK, Presscott WR. The influence of setup time on job shop performance. Management Science 1969; 16 (4): 274-280.

[17]Kim SC, Bobrowski PM. Impact of sequence dependent setup time on job shop scheduling performance. International Journal of Production Research 1994; 32 (7): 1503-1520.

[18]Kim SC, Bobrowski PM. Scheduling jobs with uncertain setup times and sequence dependent. International Journal of Management Science 1997; 25 (4): 437-447.

[19]Vinod V, Sridharan R. Dynamic job shop scheduling with sequence- dependent setup times: Simulation modelling and analysis. International Journal of Advanced Manufacturing Technology 2008; 36 (3-4): 355-372.

[20]Baker CT, Dzielinski BP. Simulation of a simplified job shop. Management Science 1960; 6 (3): 31l-323.

[21]Yu X, Ram B. Bio-inspired Scheduling for dynamic job shops with flexible routing and sequence dependent setups. International Journal of Production Research 2006; 44 (22):4793-4813.

[22]Rossi A, Dini G. Flexible job shop scheduling with routing flexibility and separable setup times using ant colony optimisation method. Robotics and Computer Integrated Manufacturing 2007; 23 (5): 503-516.

[23]Baykasoglu A, Göçken M, Unutmaz ZD. New approaches to due date assignment in job shops. European Journal of Operational Research *2008;* 187(1): 31-45.

[24]Rangsaritratsamee R, Ferrell WG, Kurz MB. Dynamic rescheduling that simultaneously considers efficiency and stability. Computer and Industrial Engineering 2004; 46 (1): 1-15.

[25]Van Parunak H. Characterizing the manufacturing scheduling problem. Journal of Manufacturing Systems 1991; 10 (3): 241-259.

[26]Baker KR. Sequencing rules and due-date assignments in a job shop. Management Science 1984; 30 (9): 1093-1104.

[27]Law AM, Kelton WD. Simulation modeling and analysis. New York: McGraw-Hill; 1991.

Tables

Table 1

Table 1. Routes of job types.

Job type	Number of operations	Route of the job (Machine number)
A	5	1-6-10-2-4
B	4	8-3-5-10
C	4	7-9-3-1
D	5	5-7-9-2-4
E	4	2-8-1-10
F	5	6-9-1-3-5

Table 2

Table 2. Processing times of jobs on machines according to routes.

Job type	Processing times of jobs according to machines
A	U(10,11), U(14,15), U(17,18), U(16,17), (18,19)
B	U(17,18), U(10,11), U(19,20), U(13,14)
C	U(17,18), U(11,12), U(16,17), U(13,14)
D	U(12,13), U(19,20), U(16,17), U(10,11), U(17,18)
E	U(13,14), U(19,20), U(10,11), U(16,17)
F	U(19,20), U(13,14), U(15,16), U(10,11), U(14,15)

Table 3

Table 3. Job types/sequence-dependent setup times data .

Preceding job type	Follower job type					
	A	B	C	D	E	F
A	0	U(5,5.25)	U(5,5.75)	U(5,5.50)	U(5,5.50)	U(5,5.25)
B	U(5,5.50)	0	U(5,5.25)	U(5,5.75)	U(5,5.25)	U(5,5.50)
C	U(5,5.25)	U(5,5.50)	0	U(5,5.50)	U(5,5.75)	U(5,5.25)
D	U(5,5.75)	U(5,5.25)	U(5,5.50)	0	U(5,5.25)	U(5,5.50)
E	U(5,5.50)	U(5,5.75)	U(5,5.25)	U(5,5.50)	0	U(5,5.25)
F	U(5,5.25)	U(5,5.50)	U(5,5.75)	U(5,5.25)	U(5,5.50)	0

Figures

Figure 1

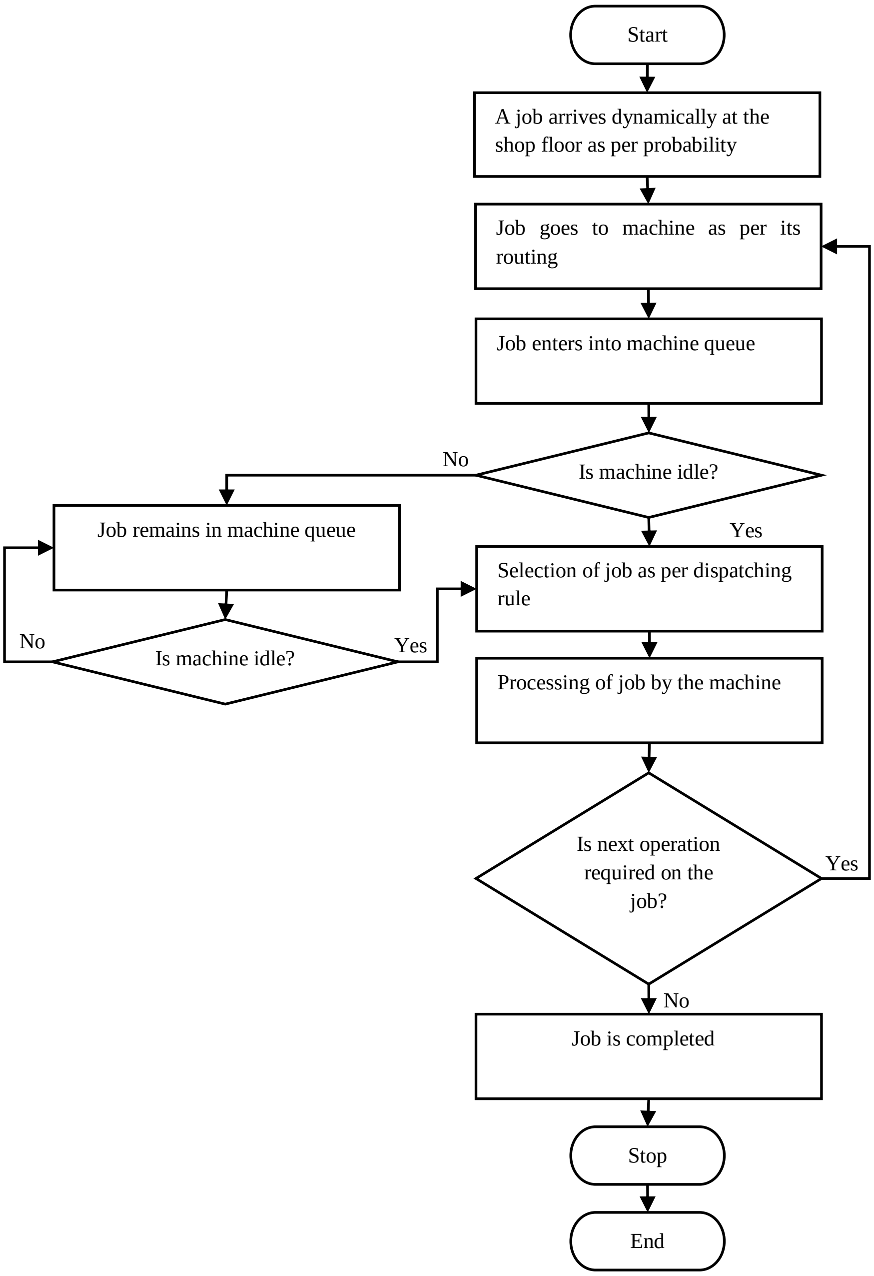


Figure 1. Job flow in a modeled job shop.

Figure 2

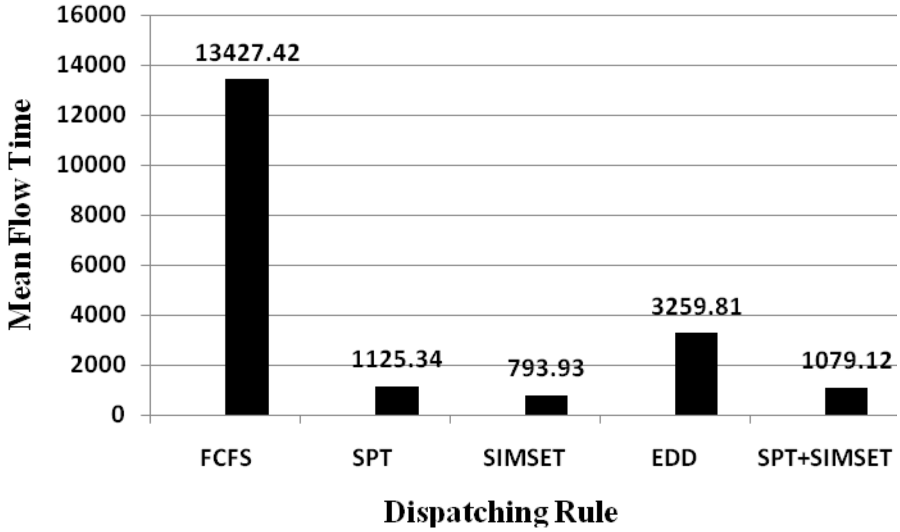


Figure 2. Performance of dispatching rules for mean flow time.

Figure 3

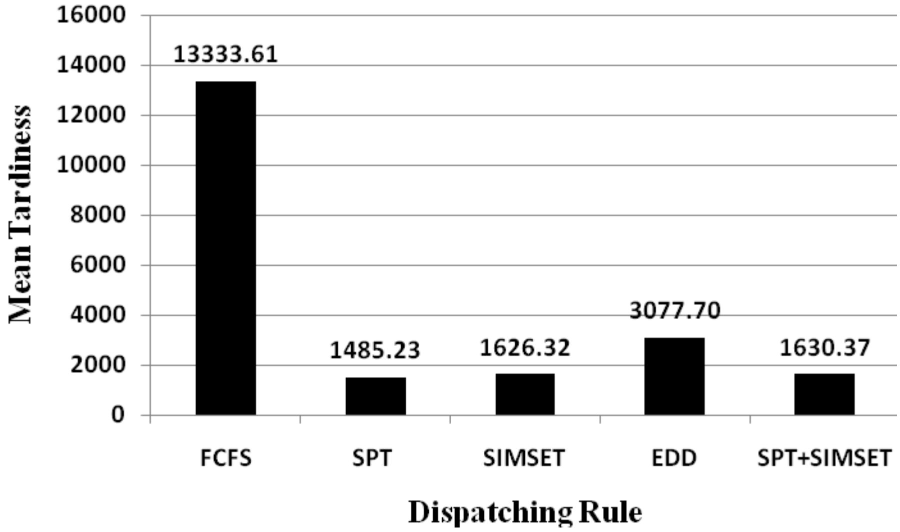


Figure 3. Performance of dispatching rules for mean tardiness.

Figure 4

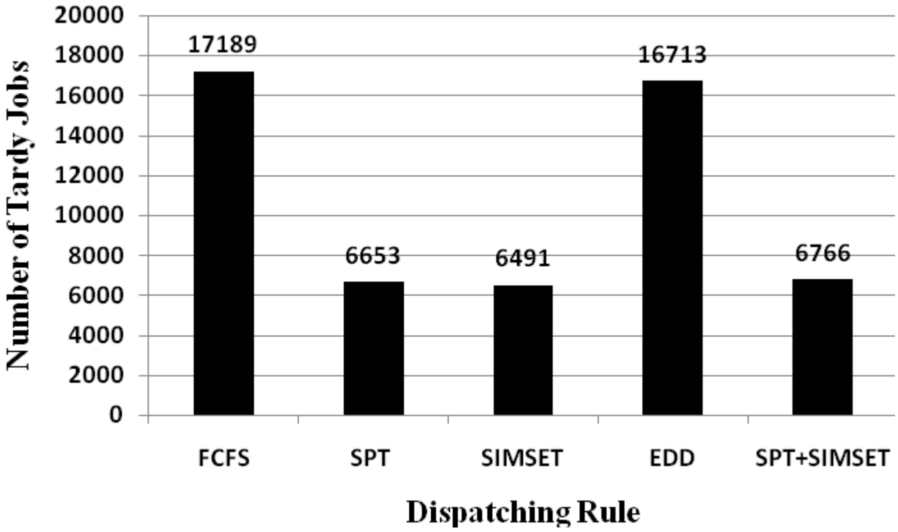


Figure 4. Performance of dispatching rules for number of tardy jobs.

Comparison of ANN and Statistical Regression Models for Prediction of Average Equivalent Strain in Equal Channel Angular Pressing

Rajat Setia [1], K. Hans Raj [2], Suren N. Dwivedi [3]

[1] U.S.I.C., Dayalbagh Educational Institute, Dayalbagh, Agra -282005, INDIA
[2] Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra -282005, INDIA
[3] Mechanical Engineering Department, University of Louisiana, Lafayette, U.S.A

Email: khansraj@redifmail.com

Abstract

In the current study, Artificial Neural Network (ANN) and regression models are developed to predict average equivalent strain in equal channel angular pressing (ECAP). Finite Element (FE) modelling in FORGE 3 environment was conducted for ECAP process with different channel angles, coefficient of frictions and processing routes. The FE modeled data are split into two sets, one for training and the other for validation of the model. The training data is used to develop the above models and the test data, which was not used earlier to develop these models, is used for validating the models. Subsequently, the models are compared and it is found that the predicted and FE results are in good agreement with simulation results. The percentage error comparison of ANN and regression model clearly indicates that the performance of ANN is much better than the statistical regression model.

Keywords: Equal Channel Angular Pressing (ECAP), Average Equivalent Strain, Artificial Neural Network (ANN) Model, Regression Model.

Introduction

Equal Channel Angular Pressing (ECAP) has emerged as most prominent Severe Plastic Deformation (SPD) technique used to produce an ultrafine grained (UFG) structure [1, 2] in metals from coarse grained metals in order to improve their mechanical and physical properties since the final grain size is related to equivalent strain in the work piece by Hall-Petch equation [3, 4] which states that the yield stress, σ_y, is given by $\sigma_y = \sigma_o + k_y d^{-1/2}$, where σ_o is termed the friction stress, k_y is a constant of yielding and d is the grain size. The ECAP parameters, viz., amount of deformation shear strain (ε), number of passes (N), rotation angle between each repetitive pressing, the strain rate monitored by movement of punch, and the temperature in process greatly influence the final microstructure and thus the properties of the final product in ECAP. A schematic of the process is exhibited in figure 1, where Φ and Ψ are the channel intersection angle and the arc curvature angle respectively. Pressure P is applied on the plunger which causes the billet to move through the channel experiencing intense shear strain responsible for producing ultrafine grained structure with improved mechanical properties such as hardness and toughness.

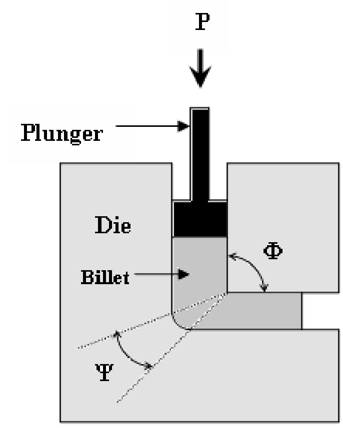


Figure 1. Typical ECAP process, where Φ is die channel angle and Ψ is angle of curvature.

Artificial Neural networks (ANN) are systems that try to make use of some of the known or expected organizing principles of the human brain. ANN's most prominent feature is to learn from examples, and then adapt themselves based on actual solution space (training data sets) [5]. ANN consists of a number of independent, simple processors - the neurons. These neurons communicate with each other via weighted connections. Learning in neural networks means to determine a mapping from an input to an output space by using example patterns. If the same or similar input patterns are presented to the network after learning, it should produce an appropriate output pattern.

In this work Back-Propagation learning methodology with Levenberg -Marquardt (LM) approximation is adopted for supervised learning of the networks and is briefly described below: The Back-Propagation (BP) neural network is a multiple layer network with one input layer, one output layer and some hidden layers between input and output layers. Its learning procedure is based on gradient search with least sum squared optimality criterion. Calculation of the gradient is done by partial derivative of sum squared error with respect to weights. This algorithm can be expressed succinctly in the form of a pseudo-code as given below.

1. Pick a rate parameter *R*.
2. For each sample input compute the resulting output until performance is satisfactory
3. Compute β (error)for nodes in the output layer using;

   $$\beta_Z = D_Z - O_Z$$

   where *D* represents the desired output and *O* represents the actual output of the neuron.
4. Compute β for all other nodes using;

   $$\beta_j = \sum_k W_{j \to k} O_k \left(1 - O_k\right) \beta_k$$

5. Compute weight changes for all weights using;

$$\Delta w_{i\rightarrow j} = rO_i O_j \left(1 - O_j\right)\beta_j$$

6. Add up the weight changes for all sample inputs and change the weights.

The standard BP algorithm suffers from the serious drawbacks of slow convergence and inability to avoid local minima. Therefore, BP with Levenberg -Marquardt (LM) approximation is used in this work. LM learning rule uses an approximation of the Newton's method to get better performance [6]. This technique is relatively faster but requires more memory. The LM update rule is:

$$\Delta W = \left(J^T J + \mu I\right)^{-1} J^T e$$

Where J is the Jacobean matrix of derivatives of each error to each weight, μ is a scalar and e is an error vector. If the scalar is very large, the above expression approximates the Gradient Descent method while when it is small the above expression becomes the Gauss - Newton method. The Gauss Newton method is faster and more accurate near error minima. Hence, the aim is to shift towards the Gauss - Newton as quickly as possible. The μ is decreased after each successful step and increased only when the step increases the error.

In the present study, ANN model is developed to predict average equivalent strain produced during ECAP process. The proposed models use data for training procedure from FE modelling in FORGE 3 environment. The channel angle and coefficient of friction were considered as the input parameters of the models. The channel angle was varied from 90^o, 105^o and 120^o and coefficient of friction from 0, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35 and 0.4.

A nonlinear regression model is also obtained from same data for comparison with the ANN model. The proposed ANN network is proven successful, resulting in reliable predictions, providing a possible way to avoid time and money-consuming experiments.

Finite Element Modelling of ECAP process

The 3D models for the ECAP simulation were developed in solidworks software with the channel intersection angle Φ = 90° and angle of curvature Ψ = 0°. The dies are assumed to be rigid pieces and the material used is an H13 tool-steel. The dimension of the upper die or punch is 10mm (width) x 10mm (breadth) and 20mm (height). The square shaped three dimensional workpiece (billet) considered has the dimensions of 10 mm (width) x 10 mm (breadth) and 60 mm (height) (refer Fig. 2). The material of the billet is taken to be Al6061 aluminum alloy.

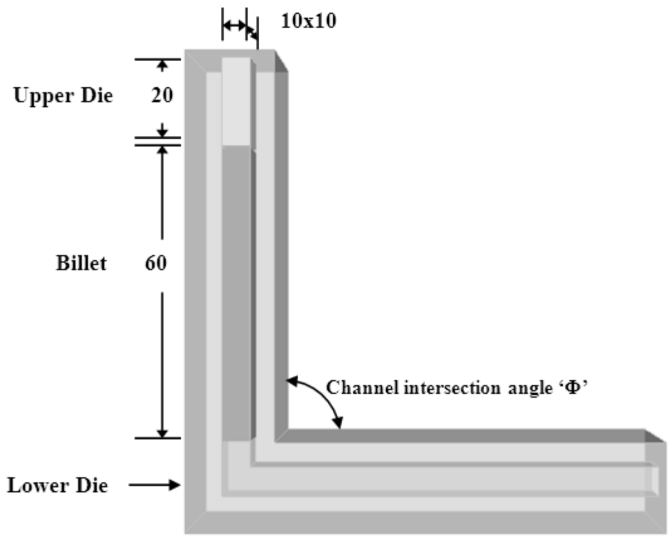


Figure 2. Geometry of ECAP dies and billet used in FE simulations

The material behavior is assumed to follow that of Norton-Hoff law written in following tensorial form:

$$s = 2K(T, \bar{\varepsilon}, \ldots)(\sqrt{3}\,\dot{\bar{\varepsilon}})^{m-1}\dot{\varepsilon}$$

where s = shear stress, K = material consistency,
T = temperature, $\bar{\varepsilon}$ = equivalent strain,
$\dot{\bar{\varepsilon}}$ = equivalent strain rate and $\dot{\varepsilon}$ = strain rate.

Generalized coulomb friction law is used in the current analysis given by:

$$\tau = \mu\sigma_n \text{ if } \mu\sigma_n < \bar{m}\frac{\sigma_0}{\sqrt{3}} \text{ and } \tau = \bar{m}\frac{\sigma_0}{\sqrt{3}}\frac{\Delta V}{\Delta V} \text{ if } \mu\sigma_n > \bar{m}\frac{\sigma_0}{\sqrt{3}}$$

where τ = friction stress tangential to the surface,
μ = coefficient of friction,
σ_n = compressive stress normal to the surface (contact pressure), $\bar{m}$ = Tresca coefficient.

Table 1 shows the FE modeling results of average equivalent strain for various values of channel angles coefficient of frictions and three different processing routes viz. route A, route B_A and route C. These results are used as training data for ANN and regression models.

	Channel Angle (Φ) 90º	Channel Angle (Φ) 105º	Channel Angle (Φ) 120º
μ	Avg. Eq. strain	Avg. Eq. strain	Avg. Eq. strain
Route ‘A’			
0.00	**6.551**	**4.579**	**2.934**
0.10	**6.958**	**4.667**	**3.096**
0.15	**7.198**	**4.675**	**3.265**
0.20	**7.322**	**4.753**	**3.315**
0.25	**7.325**	**5.126**	**3.374**
0.30	**7.328**	**5.142**	**3.508**
0.35	**7.382**	**5.664**	**3.827**
0.40	**7.605**	**6.118**	**4.453**
Route ‘B_A’			
0.00	**7.185**	**4.309**	**2.753**
0.10	**7.232**	**4.623**	**2.795**
0.15	**7.381**	**5.048**	**2.834**
0.20	**7.394**	**5.239**	**2.902**
0.25	**7.678**	**5.569**	**3.002**
0.30	**7.721**	**5.682**	**3.095**
0.35	**7.791**	**5.791**	**3.102**
0.40	**7.944**	**5.916**	**3.146**
Route ‘C’			
0.00	**6.132**	**4.468**	**2.906**
0.10	**6.235**	**4.702**	**3.272**
0.15	**6.677**	**4.737**	**3.114**
0.20	**6.699**	**4.562**	**3.482**
0.25	**6.926**	**4.940**	**3.488**
0.30	**7.057**	**5.393**	**3.870**
0.35	**7.163**	**5.628**	**3.428**
0.40	**7.188**	**5.716**	**3.274**

Table 1: ANN training data: Average equivalent strain for routes A, B_A and C

ANN Modelling of ECAP process

A three layer network with two inputs i.e. coefficient of friction, and channel intersection angle (Φ) and two outputs i.e. forming energy and equivalent strain is designed as shown in figure 3. The data obtained through FE simulations for three routes A, B_A and C is used for training the ANN model. After training, the weights are frozen and the model is validated. For this purpose, the input parameters to the network are sets of values that have not been used for training the network but are in the same range as those used for training.

This enables us to test the network with regard to its capability for interpolation. The forming energy and equivalent strain are thus obtained for this set of parameters. Then an FE simulation is performed for the same sets of parameters to determine the forming energy and equivalent strain through the FE simulation. The level of agreement between the forging force predicted by the neural network and the FE simulation indicates the efficacy of the neural model.

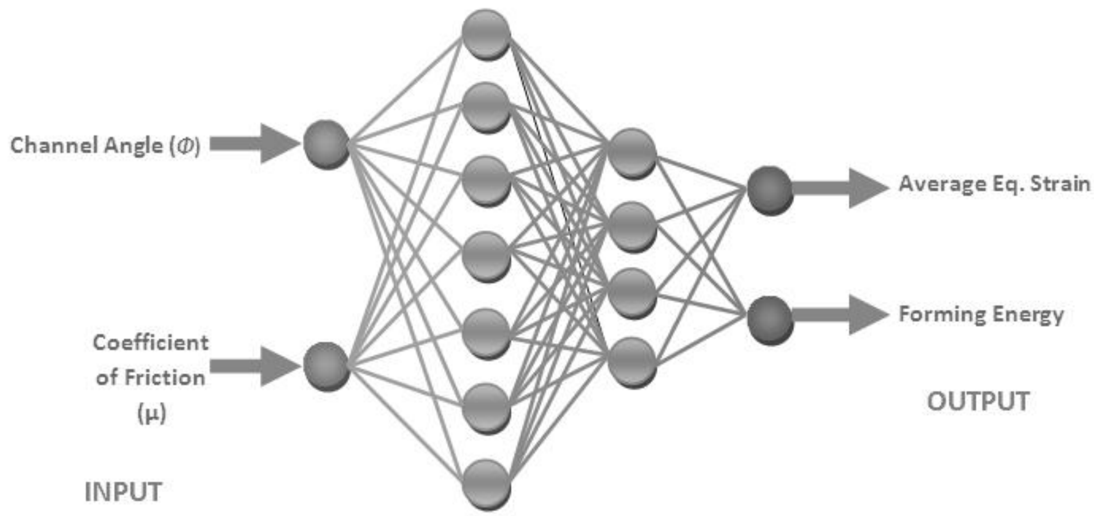


Figure 3: Two input and two output ANN architecture of ECAP process

For the training problem at hand the following parameters were found to give rapid convergence of the training network with good performance in the estimation;

- First and second layers of neurons are modelled with log of sigmoid function, and the third layer is purely linear function. Neurons taken in first and second layers are seven (7) and four (4) respectively.

- Maximum epochs considered were 1000, error goal is set at 10^{-6} and learning rate for training the network is taken as 0.2.

The results of the validation procedure described above are given in tables 2. The close agreement of the values of the equivalent strain and forming energy obtained by the neural network and the FE simulation clearly indicates that the model can be used for predicting the forging force in the range of parameters under consideration. Convergence graphs between sum squared error and number of epochs for routes A, B_A and C are shown in Figure 4.

The model is very fast and the time taken for prediction is very small. This meta model can also be used in tandem with optimizer in future for finding the optimal ECAP profile and process parameters for maximizing the equivalent strain and minimizing extrusion force.

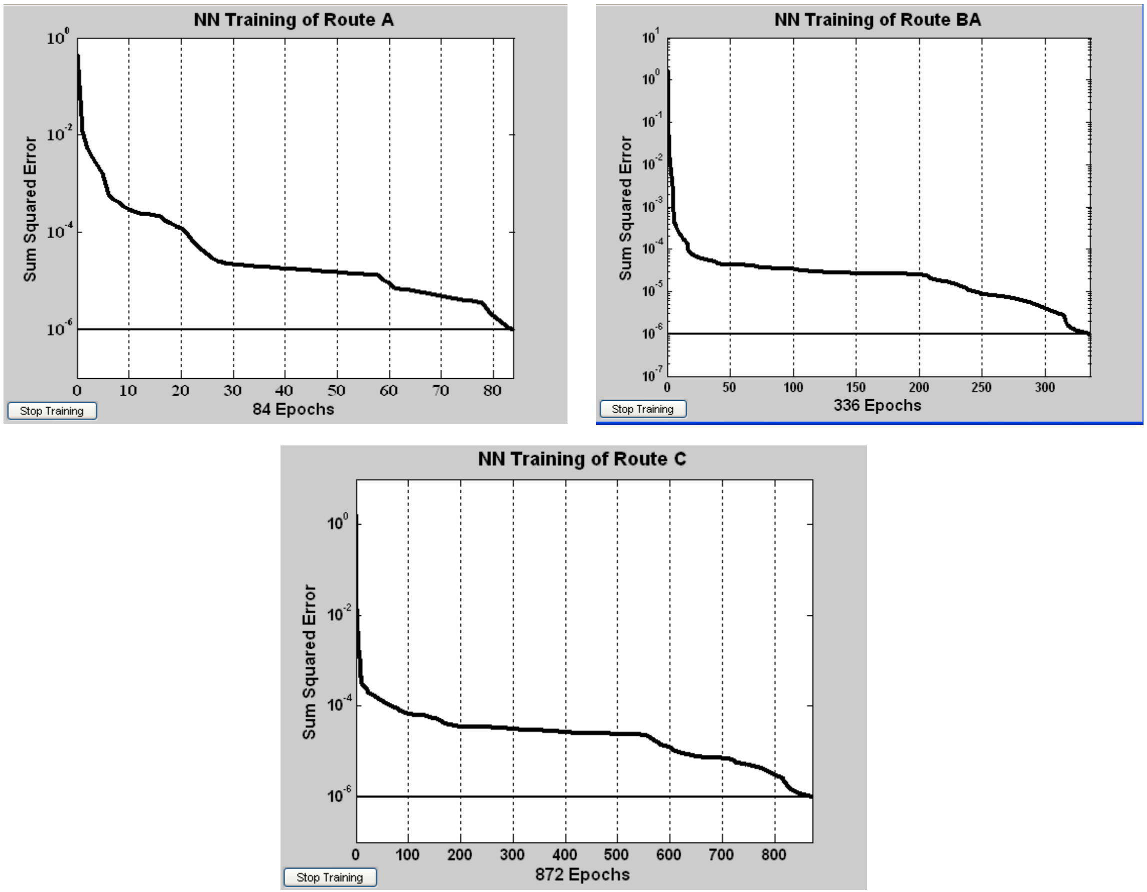


Figure 4: Convergence graph between sum squared error and number of epochs for route A, B_A and C

S. No.	Φ	μ	Average Equivalent Strain	
			FE result	ANN result
1.	90°	0.02	6.602	6.582
2.	90°	0.04	6.693	6.509
3.	90°	0.06	6.884	6.794
4.	90°	0.40	7.605	7.623
5.	105°	0.02	4.583	4.576
6.	105°	0.04	4.597	4.592
7.	105°	0.06	4.600	4.599
8.	105°	0.40	6.118	6.120
9.	120°	0.02	2.986	2.985
10.	120°	0.04	2.992	2.889
11.	120°	0.06	3.001	3.000
12.	120°	0.40	4.453	4.453

Table 2. Validation data set with FE results

Regression Analysis of ECAP Process

Statistical Regression Analysis (RA) models are developed for ECAP process and are compared with ANN models. Through the regression analysis of the results, the values of the model coefficients have been obtained and the regression equation for average equivalent strain is given as under:

$$\text{Strain} = 18.3 - 0.130\ \text{angle} + 2.69\ \text{friction}$$

In order to judge the accuracy of the regression prediction model, relative percentage error ϕ and average relative percentage error ϕ_a are used which are defined as:

$$\phi = \frac{P_{FE} - E_V}{P_{FE}} \times 100\%$$

$$\phi_a = \frac{\sum_{i=1}^{n} \phi}{n}$$

where P_{FE} is predicted FE simulation value , E_V is estimated value of the model and n is the number of observations.

The comparison of both the models i.e. ANN model and RA model in terms of relative percentage error is shown in table 3. It is clear that ANN model outperformed the RA model. Error comparison of both the models is also shown in table 4.

Channel Angle (º)	Coefficient of Friction	FE simulated Average Equivalent Strain	ANN modeled Average Equivalent Strain	RA modeled Average Equivalent Strain	Relative Error (ϕ) obtained by RA (%)	Relative Error (ϕ) obtained by ANN (%)
90	0.02	7.237	7.018	6.653	8.07	3.03
90	0.06	7.493	6.990	6.761	9.77	6.71
90	0.08	7.556	7.174	6.815	9.80	5.05
105	0.02	5.572	5.201	4.703	15.59	6.66
105	0.06	5.584	5.235	4.811	13.84	6.25
105	0.08	5.649	5.186	4.865	13.87	8.19
120	0.02	4.368	3.869	2.753	36.97	11.42
120	0.06	4.571	4.621	2.861	37.41	1.09
120	0.08	4.738	4.739	2.915	38.47	0.02

Table 3. Comparison of ANN and RA models for predicting average equivalent strain using relative percentage error

Model	Minimum relative error (%)	Maximum relative error (%)	Average relative error (ϕ_a) (%)
Artificial Neural Network (ANN)	0.02	11.42	5.38
Regression Analysis (RA)	8.07	38.47	20.42

Table 4. Error comparison of ANN and RA models for average equivalent strain

Conclusions

Two different predictive models of average equivalent strain are presented. The ANN model and the RA approach are used to construct models to evaluate the average equivalent strain in ECAP process. The results obtained with both ANN and RA are validated and error statistics are tabulated. The ANN model performance is found to be much better than the statistical regression model. The average relative error of ANN model is 5.38% whereas for regression model it is 20.42%.

Acknowledgement

We gratefully acknowledge the inspiration and guidance provided by Most Revered Prof. P.S. Satsangi, Chairman of Advisory Committee on Education, Dayalbagh.

Bibliography

1. R.Z. Valiev, T.G. Langdon, Prog. Mater. Sci., 51, 2006, pp. 881–981.
2. V.M. Segal, Mater. Sci. Eng. A , 386, 2004, pp. 269–276.
3. Hall E.O., Proceedings of Physics Society, Ser. B, 64, 1951, pp. 747- 753.
4. Petch N.J., Journal of Iron and Steel Institute, 1953, pp. 25-28.
5. Fausett, L., "Fundamentals of Neural Networks," Prentice Hall, Englewood Cliffs, NJ, 1994.
6. More, J. J., "The Levenberg-Marquardt Algorithm: Implementation and theory, Numerical Analysis", G. A. Watson (Ed), Lecture Notes in Mathematics, Springer Verlag, 630, pp.105 – 116, 1977.

Dr. Rajat Setia has a degree in Mechanical Engineering from the Institution of Engineers, Calcutta, M.Tech. in Engineering Systems and Ph.D. from Dayalbagh Educational Institute. He is Assistant Professor (Mechanical Engineering Department) in University Science Instrumentation Center, Dayalbagh Educational Institute, Agra. He has 14 publications to his credit in journals and conferences. He is a Life Member of Systems Society of India and Honorary Member of International Society for Productivity Enhancement. His research interests include Quantum-inspired Evolutionary Algorithms, Artificial Neural Networks, Neuro Fuzzy, Genetic Algorithms and Quantum Neural Networks.

Dr. Kandikonda Hans Raj was born on 31st July 63. He has received B.Sc. Engineering in Mechanical Engineering from Agra University, M.Tech. in Mechanical Engineering from I.I.T. Roorkee and Ph.D. from Dayalbagh Educational Institute. Presently he is a Professor in Mechanical Engineering Department at Dayalbagh Educational Institute, Dayalbagh, Agra. He is actively involved in research and teaching since 1988. His research interest includes Intelligent and Agile Manufacturing, Metal Forming Process Modeling and Optimization, Structural Design and Optimization, Finite Element Analysis, Soft Computing Applications in Manufacturing and Quantum Evolutionary Optimization. He has taught several courses to graduate and undergraduate students. He has supervised 25 U.G. and 27 P.G. Projects and 3 Ph.D.'s. He is currently supervising 2 Ph.D. scholars. He is a research consultant to ADRDE (DRDO), India. He has successfully completed five research projects granted by Government of India and currently having six research projects. He has been a visiting scientist to CEMEF Laboratory, Sophia Antipolis, France, University of Kiel, Germany, Old Dominion University, Norfolk, U.S.A., University of Technology, Helsinki, Finland, Mathematics and Computer Science Department, University of Maryland, MD, Rensselaer Polytechnic Institute, NY, MIT, Boston, U.S.A. He has been awarded prestigious "Production Engineering Division Medal" for 1999 & 2000 and most coveted "Institution medal" for the year 2001 of Institution of Engineers (India) for his research papers. He has also been awarded "Academic Excellence Award" 2011 by Aerial Delivery Research and Development (ADRDE), DRDO, Agra for his long standing academic and research support to ADRDE. He has 125 publications to his credit in journals and conferences. He has developed a Laboratory for development of Nano-material for Aero-Space applications with Severe Plastic Deformation (SPD) and Quantum inspired Evolutionary Algorithm Laboratory in faculty of engineering, D.E.I. He is a life member of International Society of Agile Manufacturing, Indian Society for Mechanical Engineers, Indian Society for Technical Education, Indian Society for Continuing Engineering Education, Institution of Engineers (India), Systems Society of India and Aeronautical Society of India. He is also a Fellow of Institution of Engineers, India (FIE).

Dr. Surendra Nath Dwivedi Board of Regents Eminent Scholar, Endowed Chair Professor and Director of the Virtual Reality and Product Realization Lab at University of Louisiana Lafayette, U.S.A. was born in Ballia, a small district in U.P. He graduated with a B.Sc. (honors) and a B.Sc. in Mechanical Engineering from the B.H.U. He has also earned an M.S. in Machine Design from the University of Roorkee; an M.A.Sc. in Mechanical Engineering from the University of British Columbia, Vancouver; and a Ph.D. in Engineering from the Birla Institute of Technology, Ranchi. In 2009, he received the honorable Louisiana Faculty Professionalism Award, and has received 33 awards for outstanding teaching, research, professional and community services. His research areas are Concurrent Engineering, Agile Manufacturing and Lean Manufacturing. He has developed and implemented programs in teaching CAD/CAM and robotics. He has taught several courses to graduate and undergraduate students. He has guided more than 100 students through their M.S. theses and their Ph.D. dissertations. He is the founding president of three professional journals: Concurrent Engineering: Research and Application (CERA), International Journal of Agile Manufacturing (IJAM) and The International Journal of Advanced Manufacturing (IJAMS). He has published more than 200 papers in reputed journals and at prestigious conferences.

An investigation on the deformation of Al alloy during integrated Extrusion and ECAP

Ankit Sahai[1], Shanti S Sharma[1], Rahul S Sharma[1], K. HansRaj[1], Suren N. Dwivedi[2]

[1]Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra, INDIA
[2]Mechanical Engineering Department, University of Louisiana, Lafayette, USA
Email: sahaiankit13@gmail.com

Abstract:

Bulk Nano Materials have number of applications in automobile, aero-space, medical and manufacturing applications. These are produced by subjecting materials to severe plastic deformation (SPD) and have widely emerged as a technique for grain refinement in Al, Cu, Ti, Mg alloys with improved mechanical properties. Equal Channel Angular Pressing (ECAP) is one such SPD technique employed to produce bulk ultra-fined grained (UFG) materials by introducing a large amount of shear strain into the materials without changing the billet shape or dimensions. FE (Finite Element) modeling of SPD processes has become an important tool for designing feasible production processes, because of its unique capability to describe the complex geometry and boundary conditions. In this proposed work, integrated SPD processes namely Extrusion + ECAP (Ex-ECAP) is proposed and the specimen is subjected to these processes in the same die set-up. The 3D finite element modelling of Al6061 was performed using metal forming software FORGE. The dies used in both the processes during the simulation of Al6061 billet include channel angle of 90^0 and outer corner angle fixed at 16^0 with simulation performed for different plunger velocities. The simulation results clearly depict the change in equivalent strain in the entire specimen on account of these processes. Evolution of strain at different considered cross-sections is analysed. Also, the variation in extrusion force and energy are studied for the considered process parameters. The FE simulations greatly help in designing the dies for various experimental conditions to produce bulk nano-materials.

Keywords: Severe Plastic Deformation, Nanomaterials, Equal Channel Angular Pressing, Extrusion, Finite Element Modeling, Ultra-Fine Grained Material

1. Introduction

Presently, engineering sectors like automotive and aerospace are focussing on aluminium alloys due to their good corrosion resistance, superior mechanical properties along with good machinability, weldability, and relatively low cost [1,2]. One of the key factors in changing the mechanical properties of polycrystalline materials is by controlling the grain size. At low temperatures, the yield strength is related to the grain size by the well known Hall–Petch relationship [3, 4], where the yield strength of material increases with decreasing grain size. A reduction in grain size can also lead to low temperature and/or high strain rate supeplasticity [5].

Since last decade many deformation processes are under investigation to obtain metals and alloys with ultrafine microstructures and consequently high strength. Hence, materials with nanometer or sub micrometer grain sizes are receiving greater interest because of their unique mechanical and physical properties and high performance [6‑9]. The production of materials with ultra fine grain sizes can be achieved by subjecting coarse grained metal to severe plastic deformation to improve their mechanical and physical properties [10-14]. SPD technology has become the focus of attention of many research groups and individual researchers and an analysis was conducted by Langdon [15] to evaluate the impact of the broad publications appearing over the last decade within the discipline of material science indicating it to be the most popular research area. Many SPD techniques like equal channel angular pressing (ECAP) [8-11], high pressure torsion (HPT) [16], twist extrusion (TE) [17], etc., have been developed and analysed. ECAP is a promising process because it can produce bulk, fully dense, and contamination-free UFG materials. Moreover, one can design and predict the microstructural evolution by using different routes (route A, B_A, B_C , and C). In recent years, numerous theoretical and experimental investigations on the ECAP process [9-15] have been conducted to demonstrate the effect of process parameters on material behaviour. Recently, Valiev et al. [18] discussed new concepts and principles in application of SPD processing to fabricate bulk nanostructured Al alloys with advanced properties. Many researchers are also working on Finite Element Modelling (FEM) [19-26] to understand the deformation behaviour of materials and to estimate the developed strain in the ECAP process. FE simulations help to understand and critically assess the existing ECAP process with a better insight into influence of different process parameters. Recently, Suo et al. [19] and Basavaraj et al. [20] have done some 3D analyses to trace the homogeneity during the ECAP processes after the first pass Xu et al. [22] and Jiang et al. [23] studied the distribution of strain in the cross-section of the sample of pure Al and CP-Ti, respectively, during the 3D FEM simulations for the multiple passes. Hans Raj, et al. [21,24] has analysed the influence of friction and channel angle in ECAP using FE analysis. Nagasekhar et al. [25] considered the effect of strain hardening and friction in the pure copper by 3D FEM and Balasundar et al. [26] analysed the effect of friction models on deformation behaviour of pure aluminium. Sabirov et al. [27] discussed the application of ECAP with parallel channels on deformation behaviour of Al alloys and Ahmadabadi et al. [28] analysed the changes in mechanical properties of Al alloy during ECAP with different heat treatments.

In this work, authors have tried to combine extrusion and ECAP (Fig. 2) in the same die setup and investigated the deformation of Al6061 during low friction conditions and for different plunger velocities.

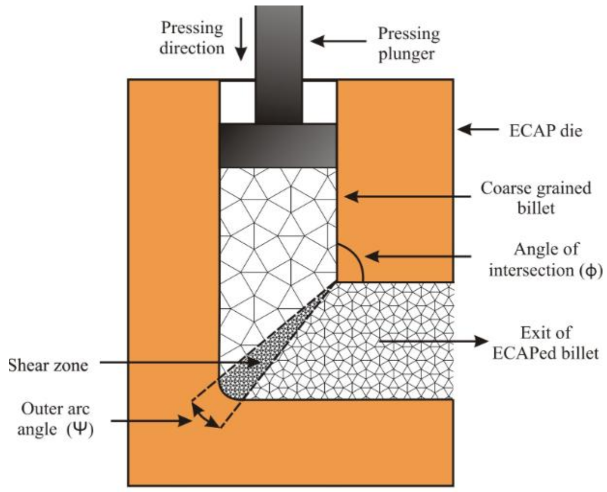


Figure 1: Schematics of ECAP process

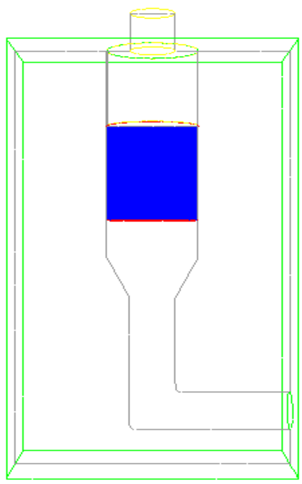


Figure 2: Extrusion-ECAP

2. Finite Element Modeling

Finite Element Method (FEM) is one of the important approaches to understand the deformation occurring in the Equal Channel Angular Pressing (ECAP) process.

In this work, the FE modeling of Extrusion-ECAP is done in FORGE environment, which is capable of modeling 3-D situations of metal forming (including thermal and friction effects) with automatic mesh regeneration. The material is assumed to be homogeneous, isotropic and incompressible. The dies are assumed to be rigid. The dimension of the plunger is 20mm (width) x 20mm (breadth). The three dimensional work piece (billet) considered has the dimensions of 20 mm (width) x 20mm (breadth) and 105 mm (height). The material of the billet is Al6061. FE simulations are carried out for ECAP ($\Phi = 90^0$ and $\psi = 16^0$) having low friction conditions ($\mu = 0.02$ and value of Tresca coefficient, $\overline{m}$ is kept constant at 0.05). All the simulations are done at 20°C under pressing velocity of 1, 5 and 10 mm/s and the material rheological behaviour is assumed to be elastoplastic.

The Hansel – Spittel equation, Eq. 1, is used to describe the behavior of the material during the deformation which is defined as:

$$\sigma_f = A \exp^{m_1 T} T^{m_9} \varepsilon^{m_2} \exp^{\frac{m_4}{\varepsilon}} (1+\varepsilon)^{m_5 T} \exp^{m_7 \varepsilon} \dot{\varepsilon}^{m_3} \dot{\varepsilon}^{m_8 T} \qquad (1)$$

Where ε is the equivalent strain, $\dot{\varepsilon}$ is equivalent strain rate, T is temperature and A, m_1, m_2, m_3, m_4, m_5, m_6, m_7, m_8, m_9 are regression coefficients.
The variation of energy required and equivalent strain in the end product with different plunger velocities for low friction conditions are obtained, Table 1.

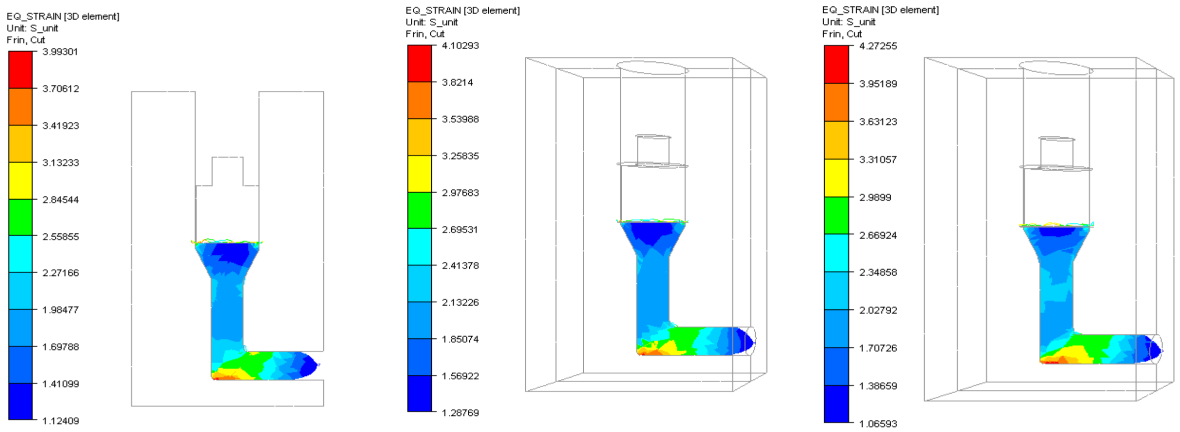
Table 1: The FE evaluation of equivalent strain, forging force and energy during Extrusion-ECAP for $\phi = 90^0$ for different plunger velocities

Plunger Velocity (mm/s)	Equivalent Strain		Extrusion Force (Tons)	Energy (KJ)
	Extrusion Zone	ECAP Zone		
1	2.22	3.1	43.37	5.53
5	2.33	3.23	44.97	5.56
10	2.66	3.34	45.1	5.72

3. Results and Discussion

3.1 Evolution of Equivalent Strain

In this process, Extrusion-ECAP, deformation occurs twice, first during extrusion and then during ECAP where billet experiences severe plastic deformation by simple shear at the region where two channels intersects. The equivalent strain contours for different plunger velocities after 1st pass are depicted in Fig. 3. The schematic end billet for the process is depicted in Fig. 4. The variation of equivalent strain along the billet length at the middle (B-B') is also evaluated and shown in Fig. 5. It can be seen that equivalent strain increases rapidly after ECAP (after 50 mm), reaching to maximum value around 3 and then starts declining once billet slides on the exit channel.



(a) (b) (c)

Figure 3: Equivalent Strain Contours during Extrusion-ECAP for plunger velocities (a) 1mm/s, (b) 5 mm/s and (c) 10 mm/sec

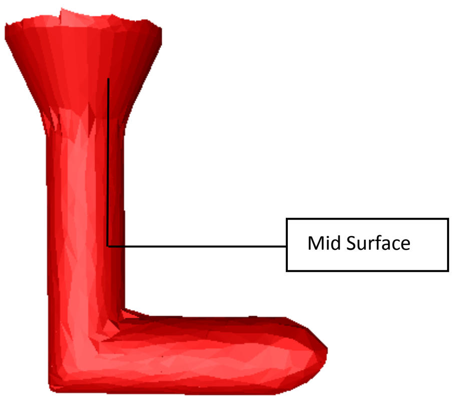


Figure 4: Schematic representation of billet after 1st Pass

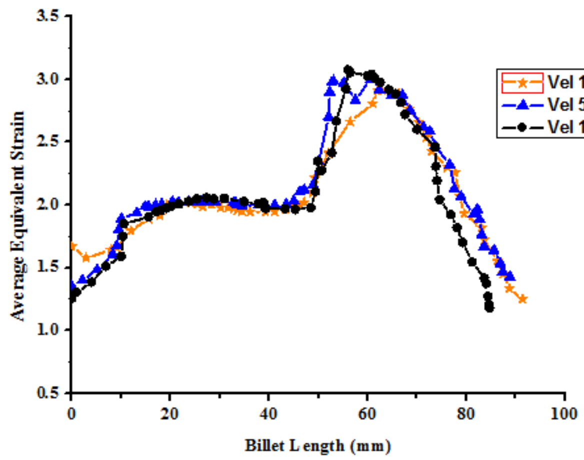


Figure 5: Strain distribution in the billet mid surface in Extrusion-ECAP for $\Phi = 90^o$ at μ = 0.02 for different plunger velocities.

3.2 Variation of Extrusion Force

Fig. 6 depicts the variation of extrusion force during 1st pass for $\Phi = 90^0$ at low friction condition. During the initial stages of deformation, the billet enters the deformation zone during extrusion and force increases to a certain limit. During further deformation, as the punch moves in the downward direction, the billet is compressed in the main deformation shear zone (ECAP) where force attains its peak value. On further pressing, billet slides on the die surface which causes shear deformation and hence the extrusion force drops. It can be seen that more extrusion force is required for plunger velocities 5 and 10 mm/sec as compared to 1 mm/sec.

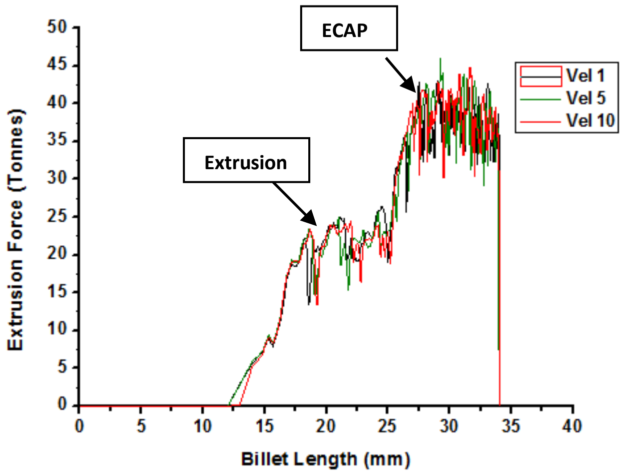


Figure 6: Variation of extrusion force during the deformation process for different plunger velocities

3.3 Variation of Energy

The variation of energy during Extrusion-ECAP is predicted in Fig. 7. It can be seen that energy is constantly increases once the deformation starts during extrusion and then decreases sharply after the end of the process. Also, plunger velocity of 10 mm/s exhibits higher values of energy as compared to other plunger velocities.

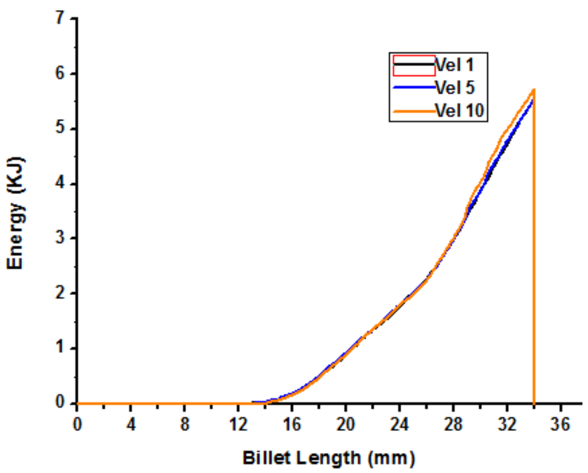


Figure 7: Variation of energy during Extrusion-ECAP for different plunger velocities

Conclusions

Severe plastic deformation (SPD) is an important process for creating bulk ultra fine grained materials. Three dimensional FE modelling of Extrusion-ECAP was carried out to study the evolution of equivalent strain on Al6061 alloy for different plunger velocities at low friction conditions.

The study revealed the following outcomes:

- During deformation, plunger velocity of 10 mm/s exhibits higher values of equivalent strain (3.34) as compared to other velocities.
- Extrusion force increases with increase in plunger velocity.
- Plunger velocity of 10 mm/s exhibits higher values of energy as compared to other plunger velocities.
- Front and back end of the billet exhibit less equivalent strain as these regions experience less deformation while high value of equivalent strain is exhibited in central part of the billet.

Acknowledgements

We deeply acknowledge the inspiration and guidance provided by Most Revered Chairman of Advisory Committee on Education, Dayalbagh, Agra, India. The authors also gratefully acknowledge the help of CEMEF Laboratory, France. Partial support for this research from All India Council of Technical Education (AICTE), under research promotion scheme (RPS), file number 8023/BOR/RID/RPS-153/2008-09 and funding from ADRDE (DRDO) and DST grant No. SR/S3/MERC/104/2008 is duly acknowledged.

References

[1] Totten, G.E., MacKenzie, D.S. (Eds.), Marcel Dekker, NY, USA, 2003

[2] Polmear, I. J., LightAlloys—Metallurgy of the Light Metals, Arnold,London, 1995

[3] Hall, E. O. The Deformation and Ageing of Mild Steel: III. Discussion of Results. Proc Roy Soc B, 1951, **64**, 747–53.

[4] Petch, N. J. Cleavage Strength of Polycrystals. J Iron Steel Inst., 1953, **174**, 25–28.

[5] Kawasaki, M., Langdon, T. G. Principles of superplasticity in ultrafine-grained materials. *J. Mater. Sci.*, 2007, **42**, 1782-1796.

[6] Segal, M., Reznikov, V.I., Drobyshevskiy, A.E., Kopylov, V. I. Plastic working of metals by simple shear. *Russ. Metall.* **1981**, *1*, 99-105.

[7] T Lowe and R Z Valiev Investigations and Applications of Severe Plastic Deformation, Dordrecht, the Netherlands Kluwer, 2000, pp. 395

[8] Furukawa, M., Horita, Z. and Langdon, T. G. Application of Equal-Channel Angular Pressing to Aluminum and Copper Single Crystals. *Mater. Sci. Forum,* 2007, **539-543**, 2853-2858.

[9] Langdon, T.G. The Principles of Grain Refinement in Equal-Channel Angular Pressing. Mat. Sci. Eng. A, 2007, **462**, 3-11.

[10] Langdon T G. Ultrafine-Grained Materials: a Personal Perspective. J. Mater. Res., 2007, **98**, 251-254.

[11] Valiev, R. Z., and Langdon, T. G. Achieving Exceptional Grain Refinement through Severe Plastic Deformation: New Approaches for Improving the Processing Technology. Metall. Mater. Trans. A, 2011, **42**, 2942-2951.

[12] Kawasaki, M. Processing of Ultrafine-Grained Materials through the Application of Severe Plastic Deformation. *Metall. Mater. Trans. A*, 2011, **42**, 3035.

[13] Valiev, R.Z., Islamgaliev, R.K., Alexandrov, I.V. Bulk Nanostructured Materials from Severe Plastic Deformation. *Prog. Mater. Sci.,* 2005, **45**, 103–189.

[14] Kim, H. S., and Estrin, Y. Micro structural Modeling of Equal Channel Angular

Pressing for Producing Ultra Fine Grained Materials. *Mater. Sci. Eng. A*, 2005, **410,** 285-289.

[15] Langdon, T. G. The impact of Bulk Nanostructured Materials in Modern Reasearch. *Rev. Adv. Mater. Sci. Res.*, 2010, **25**, 11-15.

[16] Zhilyaev, A. P. Microhardness and microstructure evolution in pure nickel during high pressure torsion. *Scr. Mater.*, 2001, **44,** 2753-2758.

[17] Orlov, D., Beygelzimer, Y., Varyukhin, V., Synkov, S., Tsuji, N. and Horita, Z. Microstructure evolution in pure AI processed with twist extrusion. J. Mater. Process. Technol., 2009, **50,** 96-100.

[18] Sabirov, I., Murashkin, M. Yu. and Valiev, R. Z. Nanostructured aluminium alloys produced by severe plastic deformation: New horizons in development. *Mater. Sci. Eng. A*, 2013, **560**, 1-24.

[19] Suo, T., Yulong Li, Deng, Q. and Liu, Y. Optimal pressing route for continued equal channel angular pressing by finite element analysis. *Mater. Sci. Eng. A*,2007, **466,** 166-171.

[20] Basavaraj Patil, V., Chakkingal, U. and Prasanna Kumar, T. S. Study of channel angle influence on material flow and strain inhomogeneity in equal channel angular pressing using 3D finite element simulation. *J. Mater. Process. Technol.*, 2009, **209,** 89-95.

[21] HansRaj, K., Sharma, R. S., Sahai, A. and Sharma, S. Finite Element Simulation of Twist Extrusion on ECAPed Al6061 Specimen. *American Institute of Physics Conf. Proc.*, 2011, **1315(507)**, 507-512.

[22] Xu, S., Zhao, G., Luan, Y. and Guan, Y. Numerical studies on processing routes and deformation mechanism of multi-pass equal channel angular pressing processes. *J. Mater. Process. Technol.*, 2006, **176,** 251-259.

[23] Jiang, H., Fan, Z. and Xie, C. 3D finite element simulation of deformation behavior of CP-Ti and working load during multi-pass equal channel angular extrusion. *Mater. Sci. Eng. A*, 2008, **485,** 409-414.

[24] HansRaj, K., Sharma, R. S., Sahai, A., Gupta, N.K. Different die designs for processing of Al Alloys: FEM Study. *Proceedings of Indian National Science Academy*,2013, **79(4),** 829-836.

[25] Nagasekhar, A. V., Yoon, S. C., Tick-Hon, Y. and Kim, H S. An experimental verification of the finite element modelling of equal channel angular pressing. *Comp. Mater. Sci.*,2009, **46**, 347-351.

[26] Balasundar, I. and Raghu, T. Effect of friction model in numerical analysis of equal channel angular pressing process. *Mater. Design.*, 2010, **31,** 449-457.

[27] Sabirov, I., Perez-Prado, M. T., Murashkin, M., Molina-Aldaguia, J. M., Bobruk, E. V., Yunusova, N. F. and Valiev, R. Z. Application of Equal Channel Angular Pressing with parallel channels for grain refinement in aluminium alloys and its effect on deformation behaviour. *Int J Mater Form*, 2010, **3**, 411-414.

[28] Moradi, M., Nili-Ahmadabadi, M., Heidarian, B. Improvement of mechanical properties of Al alloy processed by ECAP with different hest treatments. *Int J Mater Form*, 2009, **2**, 85-88.

[29] Shunqi Wang, Wei Liang, Yu Wang, Liping Bian, Kehua C. A modified die for equal channel angular pressing. *J. Mater. Process. Technol.*, 2009, **209,** 3182-3186.

[30] Segal, V. M, *Methods of Stress- Strain Analysis in Metal forming* Sc. D. Thesis, Minsk, 1974.

[31] Forge v2009 3D forging datafile, FORGE 2009. France: Transvalor Inc

Biography

Ankit Sahai is an Assistant Professor in Mechanical Engineering Department, Faculty of Engineering, Dayalbagh Educational Institute, Agra, India. He did his M.Tech. in Applied Mechanics from IIT Delhi and currently pursuing Ph.D. in Mechanical Engineering from Dayalbagh Educational Institute in collaboration with IIT Delhi. His research areas include Severe Plastic Deformations, Fatigue and Fracture Analysis, Composite Materials, Finite Element Analysis. He has few papers published in conference proceedings and journals.

Shanti Swaroop Sharma is pursuing his M.Tech. from Dayalbagh Educational Institute and working as Project Associate in UGC R&D Project in Mechanical Engineering Department at Faculty of Engineering, Dayalbagh Educational Institute, Agra. His research interest includes CAD/CAM and FE simulation of Metal Forming Processes.

Rahul Swarup Sharma has received M.Tech. and PhD in Mechanical Engineering from Dayalbagh Educational Institute Agra, India. He is a faculty member since 1999 in the University Science Instrumentation Centre of Dayalbagh Educational Institute, Agra, India. He has been visiting Assistant Professor at Rensselaer Polytechnic Institute, Troy, NY, USA in 2009. His research interests are development of Bulk Nanostructured materials through Severe Plastic Deformation. He has been awarded BOYSCAST Fellowship by Government of India in year 2009 and "Institution medal" for the year 2001 of Institution of Engineers (India) for his research paper. He has 25 publications to his credit in journals and conferences. He is a member of Institution of Engineers (India), Indian Society of Mechanical Engineers, Indian Science Congress Association, Systems Society of India and Indian Society of Technical Education.

Kandikonda Hans Raj was born on 31st July 63. He has received B.Sc. Engineering in Mechanical Engineering from Agra University, M.Tech. in Mechanical Engineering from I.I.T. Roorkee and Ph.D. from Dayalbagh Educational Institute. He is a Fellow of Institution of Engineers, India (FIE). Presently he is a Professor in Mechanical Engineering at Dayalbagh Educational Institute, Dayalbagh, Agra. He is actively involved in research and teaching since 1988. Intelligent and Agile Manufacturing, Metal Forming Process Modeling and Optimization, Soft Computing Applications in Manufacturing and Quantum Evolutionary Optimization. He is a research consultant to ADRDE (DRDO), India. He has been a visiting scientist to CEMEF Laboratory, Sophia Antipolis, France, University of Kiel, Germany, Mathematics and Computer Science Department, University of Maryland, MD, Rensselaer Polytechnic Institute, NY, MIT, MA U.S.A. He has been awarded prestigious "Production Engineering Division Medal" for 1999 & 2000 and most coveted "Institution medal" for the year 2001 of Institution of Engineers (India) for his research papers. He has supervised 26 U.G. and 28 P.G. Projects and 3 Ph.D.'s. He is currently supervising 3 Ph.D. scholars. He has 126 publications to his credit in journals and conferences. He is a life member of International Society of Agile Manufacturing, Indian Society for Mechanical Engineers, Indian Society for Technical Education, Indian Society for Continuing Engineering Education, Institution of Engineers (India), Systems Society of India and Aeronautical Society of India.

Dr. Surendra Nath Dwivedi, Board of Regents Eminent Scholar, Endowed Chair Professor and Director of the Virtual Reality and Product Realization Lab at University of Louisiana Lafayette, U.S.A. was born in Ballia, a small district in U.P. He graduated with a B.Sc. (honors) and a B.Sc. in Mechanical Engineering from the B.H.U. He has also earned an M.S. in Machine Design from the University of Roorkee; an M.A.Sc. in Mechanical Engineering from the University of British Columbia, Vancouver; and a Ph.D. in Engineering from the Birla Institute of Technology, Ranchi. In 2009, he received the honorable Louisiana Faculty Professionalism Award, and has received 33 awards for outstanding teaching, research, professional and community services. His research areas are Concurrent Engineering, Agile Manufacturing and Lean Manufacturing. He has developed and implemented programs in teaching CAD/CAM and robotics. He has taught several courses to graduate and undergraduate students. He has guided more than 100 students through their M.S. theses and their Ph.D. dissertations. He is the founding president of three professional journals: Concurrent Engineering: Research and Application (CERA), International Journal of Agile Manufacturing (IJAM) and The International Journal of Advanced Manufacturing (IJAMS). He has published more than 200 papers in reputed journals and at prestigious conferences.

Optimization of Axle NVH Performance Using the Cross Entropy Method

Glenn Meinhardt
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: gameinha@oakland.edu

Sankar Sengupta
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: ssengupta@oakland.edu

Abstract

An approach to optimization of automobile axles for noise, vibration and harshness (NVH) performance based on end-of-line testing is presented. The method used, the cross-entropy method, iteratively solves an objective function based on statistical distributions of the independent variables of the objective function. A Matlab program written by the authors is presented and discussed. The algorithm used within the method is presented along with solutions under different convergence criteria.

Introduction

Noise, Vibration and Harshness (NVH) performance is a critical quality characteristic for automobile manufacturers (original equipment manufacturers, or OEMs) and driveline component manufacturers alike. A major component of the driveline is the axle. The axle transfers torque from the engine and driveshaft to the wheels. For axle manufacturers, one of the primary NVH metrics is gear whine [1]. To ensure satisfactory gear whine performance when the automobile leaves the factory, many OEMs now require axle assemblies to be tested for gear whine performance at the end of the assembly line using an end-of-line NVH test (EOLT) prior to shipment to their assembly plants. It is in the best interest of both the OEMs and axle manufacturers to ensure that the vibration levels of axles not only meet the requirement at the EOLT, but that the levels are as low as possible [2]. One way to control the levels at the EOLT is to understand the correlation of the upstream performance variables to the EOLT result. A previous work by the authors examined one such correlation [3, 4] involving the assembly parameters of the axle and the resulting coast-side vibration. This work illustrates the use of the cross entropy method to minimize the EOLT result with the regression equation presented in [4] used as the objective function. The solution of the same problem is presented in other works by the authors using Particle Swarm Optimization [5] and a Genetic Algorithm [6].

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

The Optimization Problem

The desire is to minimize the coast-side vibration given by the regression equation from [4]:

$$
\begin{aligned}
Y(X) = \; & 486.32 \\
& - 2.8049\ a1 \\
& - 2.7890\ a5 \\
& - 0.19745\ c6 \\
& + 0.36987\ c7 \\
& - 0.29785\ d3 \\
& - 0.26230\ d6
\end{aligned}
\tag{1}
$$

Therefore, the optimization problem is written

$$\gamma^* = \min Y(X) = Y(X^*) \tag{2}$$

where γ^* is the optimum value of Y, $X' = [\ a1\ a5\ c6\ c7\ d3\ d6\ ]$, and X* are the values of X associated with γ^*. The regression equation was derived from 21 samples of data collected from the assembly line. Clearly the regression equation is only valid for the range of data from which it was derived. Therefore, the boundary conditions (constraints) for the optimization problem are the range of each variable from which the objective function was derived. The constraints are taken from range of data in [4] and summarized in Table 1. Equations (1) and (2) along with Table 1 completely define the optimization problem. Now, this problem can be solved very easily deterministically and that solution is given in Table 2. The purpose of this work is to illustrate how the cross entropy method can be used to solve optimization problems. The simple problem presented above and the deterministic solution can be used as a basis for such an illustration, the results of the optimization compared to the deterministic solution.

Table 1 – The Constraints for the Parameters, **X**, of the Regression Equation (Various Units, dB)

Variable	**Lower Bound**	**Upper Bound**
a1	66.8485	69.5424
a5	68.8496	70.9555
c6	40.0000	71.5957
c7	46.0206	78.6900
d3	46.0206	65.1055
d6	56.9020	71.3640

The Cross Entropy Method

The Cross-Entropy method (CE) was first introduced by Rubinstein [7] and is a well-known evolutionary algorithm involving variance minimization. The name “cross-entropy” is derived

Table 2 –The Solution to the Deterministic Form of the Optimization Problem

Optimization Method	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

from its use of the Kullback-Leibler divergence which is a measure of the loss of information when one probability distribution is used to approximate another. In CE, the relationship between the fitness value to be optimized, Y, and the controlling variables, **X**, is not evaluated deterministically, but with probability density functions, $f(\cdot;\mathbf{v})$ representing **X** (the associated stochastic problem). The Kullback-Leibler divergence is employed to iteratively update the parameters of the probability density functions minimizing the loss of information as the solution converges to minimum variance and the optimum value. The basic algorithm for CE is a simple two-step iterative process [8, 9]:

1. Define the associated stochastic problem by generating appropriate random samples representing each of the variables of the objective function.
2. Update the parameters of the sampling distributions for the next iteration to move the solution closer to the optimum value.

In Step 2 CE utilizes rare event estimation and importance sampling to converge each of the parameters of the probability density functions to their optimum value. A detailed derivation of CE can be found in Rubinstein and Kroese [10] with an excellent example presented by Kothari and Kroese [11]. To solve the problem with CE, it is first necessary to define the associated stochastic problem by replacing the static variables with their stochastic counterparts. It is appropriate to define the stochastic counterparts based on data collected from the assembly line. In previous work [3] it was shown that each variable can be represented by a normal distribution, except d3 due to the bi-modal nature of its distribution. It was explained in [3] that the bi-modal nature of d3 was likely due to the process producing d3 from two sources. For the purposes of this work the optimization method will assume d3 can be represented by one normal distribution. Thus Equation (1) becomes

$$\begin{aligned} Y(\mathbf{v}) = {} & 486.32 \\ & - 2.8049\, N(\mu_{a1}, \sigma_{a1}) \\ & - 2.7890\, N(\mu_{a5}, \sigma_{a5}) \\ & - 0.19745\, N(\mu_{c6}, \sigma_{c6}) \\ & + 0.36987\, N(\mu_{c7}, \sigma_{c7}) \\ & - 0.29785\, N(\mu_{d3}, \sigma_{d3}) \\ & - 0.26230\, N(\mu_{d6}, \sigma_{d6}) \end{aligned} \tag{3}$$

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

To solve Equation (2) given Equation (3), the CE method employs rare event estimation such that

$$\ell(\gamma) = \mathbf{P}_u(Y(\mathbf{X}) \le \gamma) = \mathbf{E}_u I_{\{Y(\mathbf{X}) \le \gamma\}} \quad (4)$$

where **E** is the Expected Value operator and **X** is a random vector with probability distribution functions, $f(\cdot; \mathbf{u})$, for $\mathbf{u} \in \mathbf{v}$ where

$$\mathbf{v}' = [\, \mu_{a1}, \sigma_{a1}, \mu_{a5}, \sigma_{a5}, \mu_{c6}, \sigma_{c6}, \mu_{c7}, \sigma_{c7}, \mu_{d3}, \sigma_{d3}, \mu_{d6}, \sigma_{d6} \,] \quad (5)$$

Now $\ell(\gamma) \rightarrow 0$ as $\gamma \rightarrow \gamma^*$. This is the rare event that is estimated under the importance sampling of $\mathbf{X} \rightarrow \mathbf{X}^*$. CE then adaptively updates γ and **v** until the solution converges to the tuple $(\gamma^*, \mathbf{v}^*)$. From Kothari and Kroese [11], for each iteration, *i*, and known values of $\mathbf{v}_{i-1}$ and with γ_i assigned to be a known quantile of Y(**X**) under $\mathbf{v}_{i-1}$, a value of ρ is selected such that

$$\mathbf{P}_{\mathbf{v}_{i-1}}(Y(\mathbf{X}) \le \gamma_i) \le \rho \quad (6)$$

and

$$\mathbf{P}_{\mathbf{v}_{i-1}}(Y(\mathbf{X}) \ge \gamma_i) \ge 1 - \rho \quad (7)$$

The parameter ρ defines the elite samples from the current population that will be used to estimate γ^*. For this work, ρ is chosen to be the 0.01. Again from Kothari and Kroese [11], **v** is updated in each iteration, *i*, by deriving $\hat{\mathbf{v}}_i$ from the cross-entropy program given by

$$\max_{\mathbf{v}} \hat{D}(\mathbf{v}) = \max_{\mathbf{v}} \frac{1}{N} \sum_{k=1}^{N} \mathbf{I}_{\{Y(\boldsymbol{X_k}) \le \hat{\boldsymbol{\gamma}}_{\boldsymbol{i}}\}} \ln f(\boldsymbol{X_k}; \mathrm{v}) \quad (8)$$

where

$$\mathbf{I}_{\{Y(\boldsymbol{X_k}) \le \hat{\gamma}_i\}} = \begin{cases} 1, & Y(\boldsymbol{X_k}) \le \hat{\gamma}_i \\ 0, & Y(\boldsymbol{X_k}) > \hat{\gamma}_i \end{cases} \quad (9)$$

To avoid a sub-optimal solution the convergence is slowed with a slow factor, α, such that the updated value of **v** is given as

$$\hat{\mathbf{v}}_i = \alpha\, \hat{\mathbf{v}}_i + (1 - \alpha)\, \hat{\mathbf{v}}_{i-1} \quad (10)$$

with α usually defined to be $0.7 < \alpha < 1.0$. Here α will be assigned a value of 0.75. From Kothari and Kroese [11] for normally distributed values of **X**, the solution of Equation (8) at each iteration *i* yields

$$\hat{\mathbf{v}}_i = [\hat{\boldsymbol{\mu}}_i, \hat{\boldsymbol{\sigma}}_i] \tag{11}$$

with

$$\hat{\boldsymbol{\mu}}_{ij} = \frac{\sum_{k=1}^{N} \boldsymbol{X}_{kj}}{N^{Elite}} \qquad j = 1, 2, \ldots, p \tag{12}$$

$$\hat{\boldsymbol{\sigma}}_{ij} = \sqrt{\frac{\sum_{k=1}^{N} \left(\boldsymbol{X}_{kj} - \hat{\boldsymbol{\mu}}_{ij}\right)^2}{N^{Elite}}} \qquad j = 1, 2, \ldots, p \tag{13}$$

where N is the index for the number of feasible solutions. The procedure continues until a stopping criteria is met. CE is easily adapted to a spreadsheet, but is more practical within a mathematical programming package such as Matlab. The next section will review the numerical solution of the optimization problem using CE.

Numerical Solution Using CE

The solution to the optimization problem by CE is conducted within Matlab. The algorithm used to write the program is given in the Appendix. The data used to initialize the program are the data used in [4] to derive the regression equation. These data are shown in Table 3.

The program is initialized by establishing the parameters for the CE method. These include the number of samples to generate with each generation, N, the percentile defining the elite sample, ρ, the slowing factor, α, and the stopping criteria. These values are summarized in Table 4, and are the input to the program. The program provides output of the average values and standard deviations for each parameter at each iteration, the setup parameters, the CPU Time required and the number of iterations required to converge.

The solutions from each of five runs of the CE program using the parameters in Table 4 are shown in Table 5. In addition, Figures 1, 2 and 3 show for each run the optimum value, the iterations required to converge to the optimum solution, and the solver time required. From Figure 1, it is clear that the Cross-Entropy method successfully identifies the optimum solution for this problem. Figures 2 and 3 suggest that it may not be necessary to use such strict convergence criteria since there is an impact on solver time and the number of iterations required to converge. The penalty is insignificant compared to the increased precision in the result, if increased precision is desired. In this work, it is desired to achieve precision to four decimal places. CE demonstrates the ability to do this with the more strict convergence criteria. This suggests that if high precision is desired, a good approach to optimization may be to start with more strict convergence criteria.

Table 3 – The Initial Data From the Assembly Line, Various Units, dB

Sample	a1	a5	c6	c7	d3	d6	NVH
1	67.9588	69.5134	71.5957	60.0000	46.0206	61.5836	81.03
2	66.8485	70.9555	53.9794	56.9020	58.0618	60.8279	78.18
3	67.6042	70.2377	53.9794	66.0206	49.5424	58.0618	86.26
4	68.2995	70.6296	55.5630	62.9226	49.5424	63.5218	76.49
5	67.2346	70.0212	66.4444	60.8279	52.0412	60.0000	82.29
6	69.5424	70.4228	56.9020	60.8279	46.0206	59.0849	75.86
7	68.6273	69.9662	46.0206	62.2789	62.9226	62.9226	80.95
8	68.6273	69.1273	52.0412	59.0849	53.9794	60.0000	80.22
9	68.2995	70.6040	40.0000	70.3703	49.5424	66.8485	83.64
10	68.6273	70.3703	58.0618	77.5012	49.5424	63.5218	83.44
11	68.9432	70.9309	60.0000	78.6900	46.0206	70.1030	82.09
12	67.6042	70.5526	46.0206	76.1236	49.5424	71.3640	85.15
13	68.2995	70.1030	53.9794	68.2995	52.0412	64.6090	80.36
14	67.9588	70.3703	56.9020	71.8213	65.1055	61.5836	77.54
15	68.9432	69.2180	58.0618	72.0412	63.5218	60.8279	82.33
16	68.2995	68.8496	55.5630	67.6042	64.0824	60.0000	80.61
17	68.6273	70.0212	59.0849	69.8272	64.0824	60.8279	75.29
18	69.2480	69.1273	60.0000	53.9794	64.6090	66.8485	70.63
19	67.6042	69.2480	61.5836	61.5836	46.0206	65.1055	82.28
20	67.6042	69.9109	60.0000	58.0618	46.0206	62.9226	80.53
21	68.2995	69.6575	62.9226	46.0206	59.0849	56.9020	72.87
Avg	68.2429	69.9922	56.6050	64.7995	54.1594	62.7365	79.9067
St Dev	0.6723	0.6226	6.9816	8.2071	7.3371	3.7297	3.9578
Max	69.5424	70.9555	71.5957	78.6900	65.1055	71.3640	86.26
Min	66.8485	68.8496	40.0000	46.0206	46.0206	56.9020	70.63

Table 4 – Initialization Parameters for the CE Method

Description	Parameter	Value
Number of Samples in Each Generation	N	1000
Elite percentile of the population	ρ	0.01
Slowing Factor	α	0.75
Stopping Criteria (convergence) - Trial 1	Maximum Sample Standard Deviation	0.001
Stopping Criteria (convergence) - Trial 2	Maximum Sample Standard Deviation	0.00001

Table 5 – The Average Value of the Parameters and Solution of the Optimization Problem for Each Iteration

Run	Δσ	Iterations Required to Converge	Solver Time (sec)	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
1	0.001	24	17	69.5423	70.9554	71.5940	46.0230	65.1043	71.3627	58.1427
2	0.001	24	17	69.5423	70.9554	71.5942	46.0212	65.1043	71.3631	58.1420
3	0.001	23	17	69.5423	70.9553	71.5941	46.0216	65.1041	71.3623	58.1425
4	0.001	23	16	69.5423	70.9554	71.5943	46.0215	65.1041	71.3630	58.1423
5	0.001	24	17	69.5423	70.9554	71.5945	46.0213	65.1043	71.3627	58.1420
1	0.00001	35	25	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
2	0.00001	43	31	69.3599	70.9555	71.5957	46.0206	65.1054	71.3640	58.6521
3	0.00001	35	25	69.5424	70.9555	71.5956	46.0206	65.1054	71.3640	58.1402
4	0.00001	35	25	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
5	0.00001	45	32	69.5411	70.9555	71.5957	46.0206	65.1054	71.3639	58.1439

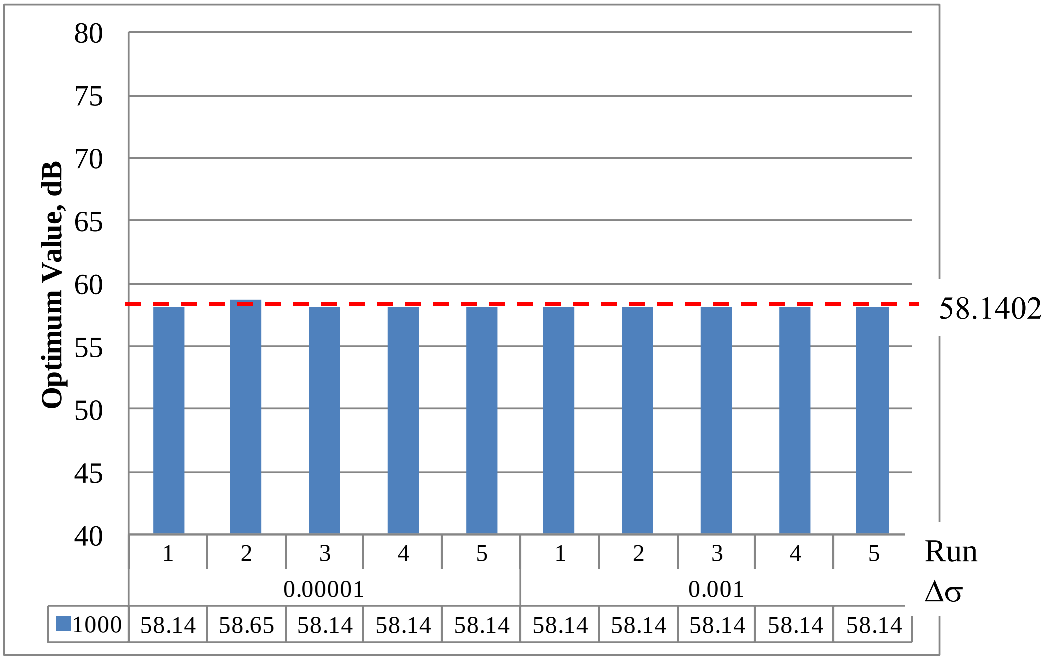


Figure 1 – The Optimum Value Identified by the Cross-Entropy Method by the Convergence Criteria and Run Number (The dashed line is the optimum value found deterministically)

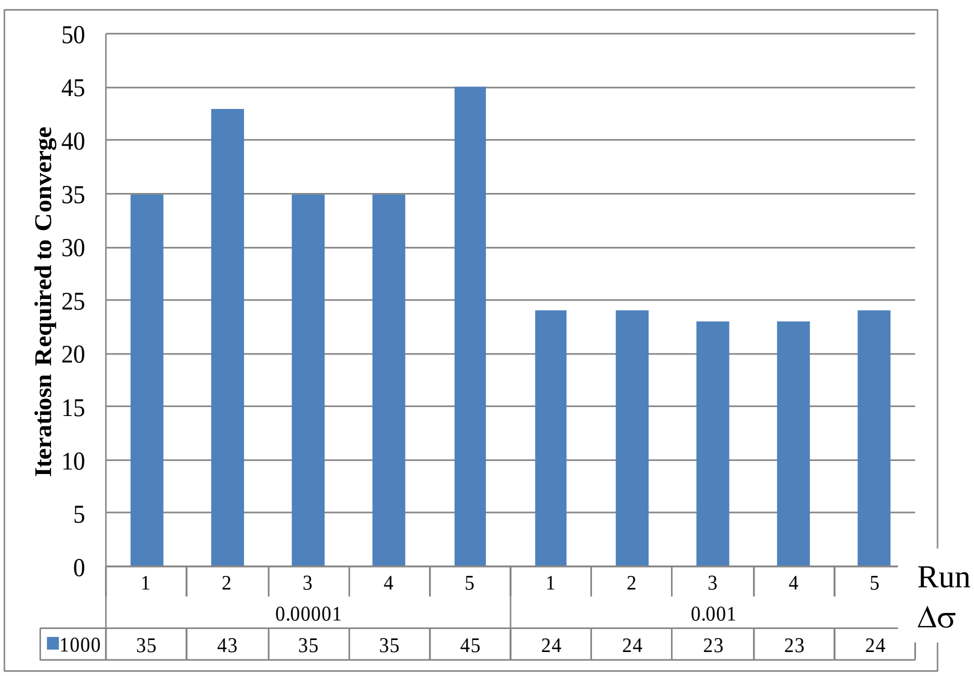


Figure 2 – The Number of Iterations Required to Identify the Optimum Value Using the Cross-Entropy Method by the Convergence Criteria and Run Number

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

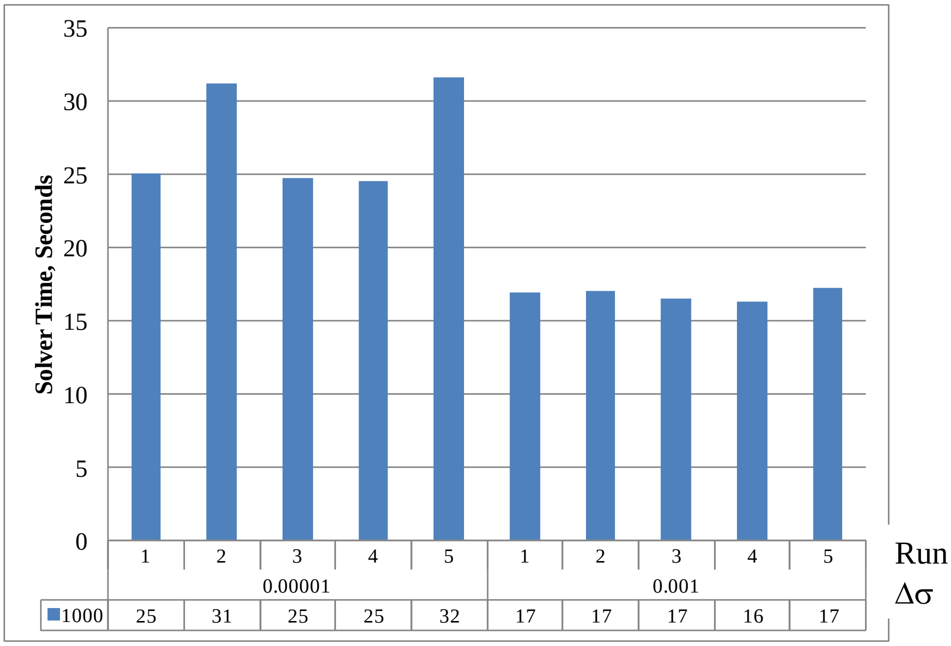


Figure 3 – The Solver Time Required to Identify the Optimum Value Using the Cross-Entropy Method by the Convergence Criteria and Run Number

Summary

The Cross Entropy (CE) method is a relatively new optimization method. This paper illustrates the application of the cross entropy method to a very simple linear regression model. The deterministic solution is used as a means of comparison of the CE results to a known solution. Table 6 shows the comparison. Table 6 shows nearly exact agreement between the classical solution and the one found through CE. As discussed above, restricting the convergence criteria further will improve the precision. It remains, for future work, to confirm that axles built to the optimum conditions indeed produce improved vibration performance. Other papers by the authors illustrate solving the same optimization problem using a Genetic Algorithm [6] and Particle Swarm Optimization [5].

Table 6 – A Comparison of the Deterministic Solution of the Optimization Problem to the Best Performance of CE

Optimization Method	N	Iterations to Solve	Solver Time (sec)	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	0	0	< 10	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
Cross-Entropy	1000	35	25	69.5424	70.9555	71.5956	46.0206	65.1054	71.3640	58.1402

Appendix

A detailed algorithm / pseudocode for the cross entropy method

1. Initialize the program:
    a. Define the number of random samples, N, for **X**.
    b. Define the percent of the solutions of Y(**X**), ρ, that will comprise the elite sample.
    c. Define the slow factor, α.
    d. Define the stopping criteria. This is chosen to be when the maximum standard deviation across all **X** and Y(**X**) is 0.001 dB or after 1,000 iterations.
2. Initialize $\boldsymbol{\mu}_{0j}$ and $\boldsymbol{\sigma}_{0j}$ for $j = 1, 2, \ldots, p$.
    a. Import the raw data from the assembly line for each parameter.
    b. Calculate $\boldsymbol{\mu}_{0j}$ and $\boldsymbol{\sigma}_{0j}$ from the data.
3. Generate N samples $\mathbf{X}_i$ from $\boldsymbol{\mu}_{(i-1)j}$ and $\boldsymbol{\sigma}_{(i-1)j}$.
    a. Evaluate $\mathbf{X}_i$ against the constraints and discard infeasible solutions.
4. Calculate $Y_i(\mathbf{X}_i)$ using Equation (1).
    a. Calculate $\boldsymbol{\mu}_{ij}$ and $\boldsymbol{\sigma}_{ij}$ and compare to the stopping criteria ($\boldsymbol{\sigma}_{ij} < 0.001$?)
    b. If True, $\gamma^* = Y_i(\mathbf{X}_i)$ and $\mathbf{X}^* = \mathbf{X}_i$
5. Sort $Y_i(\mathbf{X}_i)$ and select the ρ percentile elite solutions.
    a. The number of elite solutions = N^{Elite}
6. Calculate $\boldsymbol{\mu}_{ij}$ and $\boldsymbol{\sigma}_{ij}$ from Equations (8) and (9) using the elite samples.
7. Calculate $\hat{\mathbf{v}}_i$ from Equation (10)
8. Increment the iteration number and repeat from Step 3 until stopping criteria is met.

The Matlab code used for the cross entropy method

```
%   Open the data files
O = load('nvhdata.mat', '-ASCII');
constraint = load('const.mat', '-ASCII');
 %   Establish the number of variables (c) and the number of samples (r)
 [r,c] = size(O);
Sample(r,c)=0;
SampleStDev(1,c) = 100;
SolutionAverage(1,1) = 0;
SolutionStDev(1,1) = 0;
 %   Initialize the counters
 err = 0;
x = 0;
i = 1;
j = 1;
 %   Calculate the averages
 while j < c + 1
while i < r + 1
    x = x + O(i,j);
    i = i + 1;
end
```


Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA. Copyright © 2014, ISPE/IJAM.

```
Average(j) = x / r;
x = 0;
i = 1;
j = j + 1;
end
Average;
 %Re-iniitialize counters
 i = 1;
j = 1;
 %   Calculate the Standard Deviations
 while j < c + 1
while i < r + 1
    err = err + (O(i,j) - Average(j))^2;
    i = i + 1;
end
StDev(j) = sqrt(err / (r-1));
err = 0;
i = 1;
j = j + 1;
end
StDev;
 %  Perform the Optimization
 n = input('How many random samples shall we generate? ');
pelite = input('What percentile of feasible solutions shall we use as the Elite Sample? ');
stop = input('What is the maximum Standard Deviation desired to achieve the optimum solution? This will apply to
all variables. ');
slow = input('What weight shall we apply to the new parameters (slow factor)? ');
plots = input('Shall we create plots at the end? ', 's');
%   Begin the iterations
 numIterations = 0;
 while max(SampleStDev) > stop
 numIterations = numIterations + 1;
i = 1;
j = 1;
infeasible = 0;
 % Generate the new population
 while j < 7
while i < n + 1
    Sample(i,j) = random('norm', Average(j), StDev(j));
    Sample(i,7)= -2.8049*Sample(i,1)-2.789*Sample(i,2)-0.19745*Sample(i,3)+0.36987*Sample(i,4)-
0.29785*Sample(i,5)-0.2623*Sample(i,6)+486.32;
    i = i + 1;
end
i = 1;
j = j + 1;
end
 % Check feasibility of the solutions
  while i < n + 1
 while j < 7
 if Sample(i,j) > constraint(2,j) && Sample(i,j) < constraint(1,j)
    Sample(i,8) = 2;
     j = j + 1;
```

```
 else
   Sample(i,8) = 1;
   j = 7;
   infeasible = infeasible + 1;
 end
 end
   j = 1;
   i = i + 1;
 end
% Create the array of feasible solutions
% Sort the sample first by Column 8 (the feasible solutions, descending order) and then by
% Column 7 (the fitness value, ascending order)
 SortedSample = sortrows(Sample,[-8,7]);
NumFeasible = n - infeasible - 1;
i = 1;
j = 1;
 while i < NumFeasible + 1
   while j < 8
   Feasible(i,j) = SortedSample(i,j);
   j = j + 1;
   end
   j = 1;
   i = i + 1;
end
 % Calculate the averages of the feasible solutions
 i = 1;
j = 1;
while j < 8
FeasibleAverage = 0;
while i < NumFeasible + 1
FeasibleAverage = FeasibleAverage + Feasible(i,j);
i = i + 1;
end
SampleAverage(j) = FeasibleAverage / NumFeasible ;
SolutionAverage(numIterations,j) = SampleAverage(j);
j = j + 1;
i = 1;
end
 %   Calculate the Standard Deviations of the feasible solutions
 i = 1;
j = 1;
while j < 8
FeasibleStdError = 0;
while i < NumFeasible + 1
FeasibleStdError = FeasibleStdError + (Feasible(i,j)- SampleAverage(j))^2;
i = i + 1;
end
SampleStDev(j) = sqrt(FeasibleStdError / (NumFeasible - 1));
SolutionStDev(numIterations,j) = SampleStDev(j);
j = j + 1;
i = 1;
end
```

```
 %   Calculate the Elite Averages
 nelite = round(NumFeasible * pelite);
i = 1;
j = 1;
while j < 8
nEliteAverage = 0;
while i < nelite + 1
nEliteAverage = nEliteAverage + Feasible(i,j);
i = i + 1;
end
EliteAverage(j) = nEliteAverage / nelite ;
j = j + 1;
i = 1;
end
 %   Calculate the Elite Standard Deviations
 i = 1;
j = 1;
while j < 8
nEliteStdError = 0;
while i < nelite + 1
    nEliteStdError = nEliteStdError + (Feasible(i,j)-EliteAverage(j))^2;
    i = i + 1;
end
EliteStDev(j) = sqrt(nEliteStdError / (nelite-1));
j = j + 1;
i = 1;
end
 % Update the Average and Standard Deviation for the next population
 NumFeasible = 0;
Average = slow * EliteAverage + (1 - slow) * SampleAverage;
StDev = slow * EliteStDev + (1 - slow) * SampleStDev;
end
 % After the stopping criteria has been reached, display the results from each iteration along with the optimum
SolutionAverage
SolutionStDev
numIterations
SampleAverage
SampleStDev
% Write the results to an Excel File
 xlswrite('Solution.xls',SolutionAverage,'Averages');
xlswrite('Solution.xls',SolutionStDev,'StDevs');
 if plots == 'y'
figure
plot(SolutionAverage)
figure
plot(SolutionStDev)
end
 % End of test.  Ask to clear the memory.
 reply = input ('Do you want to clear everything? (y/n)[n]', 's');
if reply == 'y'
    clear
    clc
```


Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

```
elseif isempty(reply)
    reply = 'n';
end
```

Bibliography

1. Sun, Z., *et. al.*, "NVH Robustness Design of Axle Systems," *SAE Transactions*, v. 112, pp. 1746 – 1754, 2003.

2. Steyer, G., *et. al.*, "The Future of NVH Testing – An End-User's Perspective," *SAE Technical Paper 2005-01-2270*, 2005.

3. Meinhardt, G. and Sengupta, S., "Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part I-Preparing the Multivariate data for Regression Analysis," *SAE Technical Paper 2012-01-727*, 2012.

4. Meinhardt, G. and Sengupta, S., "Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part II-Multivariate Regression Analysis," *SAE Technical Paper 2012-01-728*, 2012.

5. Meinhardt G. and Sengupta, S., "Optimization of Axle NVH Performance Using Particle Swarm Optimization," *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

6. Meinhardt G. and Sengupta, S., "Optimization of Axle NVH Performance Using A Genetic Algorithm," *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

7. Rubinstein, R., "Optimization of Computer Simulation Models with Rare Events," European Journal of Operations Research, v. 99, pp. 89 – 112, 1997.

8. Kroese, D., *et. al.*, "The Cross-Entropy Method for Continuous Multi-Extremal Optimization," Methodology in Computing and Applied Probability, v. 8, pp. 383 – 407, 2006.

9. De Boer, P., *et. al.*, "A Tutorial on the Cross-Entropy Method," Annals of Operations Research, v. 134, pp. 19 – 67, 2005.

10. Rubinstein, R. and Kroese, D., "The Cross-Entropy Method," Springer-Verlag, 2004.

11. Kothari, R., and Kroese, D., "Optimal Generation Expansion Planning Via the Cross-Entropy Method," *Proceedings of the 2009 Winter Conference - IEEE*, pp. 1482 – 1491, 2009.

Optimization of Axle NVH Performance Using Particle Swarm Optimization

Glenn Meinhardt
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: gameinha@oakland.edu

Sankar Sengupta
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: ssengupta@oakland.edu

Abstract

An approach to optimization of automobile axles for noise, vibration and harshness (NVH) performance based on end-of-line testing is presented. The method used, the particle swarm optimization method (PSO), iteratively solves an objective function based on positioning values of the independent variables of the objective function within their corresponding constraints. A Matlab program written by the authors is presented and discussed. The algorithm used within the method is presented along with solutions under different population sizes.

Introduction

Noise, Vibration and Harshness (NVH) performance is a critical quality characteristic for automobile manufacturers (original equipment manufacturers, or OEMs) and driveline component manufacturers alike. A major component of the driveline is the axle. The axle transfers torque from the engine and driveshaft to the wheels. For axle manufacturers, one of the primary NVH metrics is gear whine [1]. To ensure satisfactory gear whine performance when the automobile leaves the factory, many OEMs now require axle assemblies to be tested for gear whine performance at the end of the assembly line using an end-of-line NVH test (EOLT) prior to shipment to their assembly plants. It is in the best interest of both the OEMs and axle manufacturers to ensure that the vibration levels of axles not only meet the requirement at the EOLT, but that the levels are as low as possible [2]. One way to control the levels at the EOLT is to understand the correlation of the upstream performance variables to the EOLT result. A previous work by the authors examined one such correlation [3, 4] involving the assembly parameters of the axle and the resulting coast-side vibration. This work illustrates the use of the cross entropy method to minimize the EOLT result with the regression equation presented in [4] used as the objective function. The solution of the same problem is presented in other works by the authors using the Cross Entropy Method [5] and a Genetic Algorithm [6].

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

The Optimization Problem

The desire is to minimize the coast-side vibration given by the regression equation from [4]:

$$\begin{aligned} Y(X) = {} & 486.32 \\ & - 2.8049\ a1 \\ & - 2.7890\ a5 \\ & - 0.19745\ c6 \\ & + 0.36987\ c7 \\ & - 0.29785\ d3 \\ & - 0.26230\ d6 \end{aligned} \tag{1}$$

Therefore, the optimization problem is written

$$\gamma^* = \min Y(X) = \mathbf{Y}(\mathbf{X}^*) \tag{2}$$

where γ^* is the optimum value of **Y**, $\mathbf{X}'$ = [a1 a5 c6 c7 d3 d6], and **X*** are the values of **X** associated with γ^*. The regression equation was derived from 21 samples of data collected from the assembly line. Clearly the regression equation is only valid for the range of data from which it was derived. Therefore, the boundary conditions (constraints) for the optimization problem are the range of each variable from which the objective function was derived. The constraints are taken from range of data in [4] and summarized in Table 1. Equations (1) and (2) along with Table 1 completely define the optimization problem. Now, this problem can be solved very easily deterministically and that solution is given in Table 2. The purpose of this work is to illustrate how the cross entropy method can be used to solve optimization problems. The simple problem presented above and the deterministic solution can be used as a basis for such an illustration, the results of the optimization compared to the deterministic solution.

Table 1 – The Constraints for the Parameters, **X**, of the Regression Equation (Various Units, dB)

Variable	**Lower Bound**	**Upper Bound**
a1	66.8485	69.5424
a5	68.8496	70.9555
c6	40.0000	71.5957
c7	46.0206	78.6900
d3	46.0206	65.1055
d6	56.9020	71.3640

Particle Swarm Optimization

Particle Swarm Optimization method (PSO) is an optimization method that was developed by Eberhart and Kennedy in 1995 [7]. The method iteratively converges to the solution of an

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

Table 2 – The Solution to the Deterministic Form of the Optimization Problem

Optimization Method	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

optimization problem by manipulating the variables in such a way that mimics a swarm or flock of birds converging on a food source or to a particular perch or other location. This work will apply PSO following the method as it was used by Robinson and Rahmat-Samii to solve an electromagnetics problem [8]. Robinson and Rahmat-Samii apply the fundamental PSO method outlined by Kennedy and Eberhart, but they also present an interesting approach to handling constraints. PSO follows a very simple algorithm:

1. Select the number of particles – usually known solutions, or random solutions.
2. Calculate the fitness of each solution.
3. Compare all of the fitness values – the solution that has the fitness value closest to optimum is the "Global Best".
4. Calculate the "velocity" of each particle – changes in each variable that will move the particle closer to the optimum.
5. Adjust the position based on the velocity and the constraints – only feasible positions are allowed.
6. Compare the fitness value of each new position of each particle to its previous position – store the position that has the fitness value closest to optimum as the new "Personal Best".
7. Compare all of the Personal Bests to the Global Best – whichever solution is closest to the optimum becomes the new Global Best.

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

8. Repeat Steps 4 through 7 until desired convergence criteria are met (e.g., small change in Global Best, maximum iterations).

The Global Best solution when the method converges is the optimum solution identified by the method.

Initializing the Swarm

Initialization of PSO begins with the definition of the sample space for each variable. Here, the sample space is defined as the range of values that were used to define the optimization equation derived by Meinhardt and Sengupta [4]. The sample space are all of the values that satisfy the constraints on **X** presented in Table 1. Next, a set of samples from the sample space of **X** are defined. Here, for the first trial, the original set of 21 samples, shown in Table 3, is used to initialize the swarm. The initial fitness value is the actual NVH result from the assembly line. This is also shown in Table 3. Another acceptable method for initialization is to select a desired number of particles, and assign the positions randomly using the initial set of 21 samples to define the sampling distributions, calculating the fitness of each particle using Equation (1). This method is also used in additional trials.

Moving the Swarm

In the PSO analog, each sample is treated as a particle. The particle analogy is appropriate because each sample has no 'mass', but is assigned 'position' and 'velocity'. The position of each particle corresponds to the current value of each variable, **X**, and is modified at each iteration according to Equation (3)

$$\mathbf{X} = \mathbf{XP} + \mathbf{V} \qquad (3)$$

with **XP** defined as the position of each particle at the end of the previous iteration. At initialization, **XP** = **X**. The velocity of the particle is defined by Equation (4).

$$\mathbf{V} = w\,\mathbf{VP} + c_1\,\mathbf{rand}()\,(\mathbf{PBEST} - \mathbf{X}) + c_2\,\mathbf{rand}()\,(\mathbf{GBEST} - \mathbf{X}) \qquad (4)$$

Equation (4) has three terms, each with a very specific purpose:

1. An inertia term

   This term, w **VP**, applies a weight (w) to the current velocity (that is, the velocity at the end of the previous iteration, **VP**)

Table 3 – The Initial Data From the Assembly Line, Various Units, dB

Sample	a1	a5	c6	c7	d3	d6	NVH
1	67.9588	69.5134	71.5957	60.0000	46.0206	61.5836	81.03
2	66.8485	70.9555	53.9794	56.9020	58.0618	60.8279	78.18
3	67.6042	70.2377	53.9794	66.0206	49.5424	58.0618	86.26
4	68.2995	70.6296	55.5630	62.9226	49.5424	63.5218	76.49
5	67.2346	70.0212	66.4444	60.8279	52.0412	60.0000	82.29
6	69.5424	70.4228	56.9020	60.8279	46.0206	59.0849	75.86
7	68.6273	69.9662	46.0206	62.2789	62.9226	62.9226	80.95
8	68.6273	69.1273	52.0412	59.0849	53.9794	60.0000	80.22
9	68.2995	70.6040	40.0000	70.3703	49.5424	66.8485	83.64
10	68.6273	70.3703	58.0618	77.5012	49.5424	63.5218	83.44
11	68.9432	70.9309	60.0000	78.6900	46.0206	70.1030	82.09
12	67.6042	70.5526	46.0206	76.1236	49.5424	71.3640	85.15
13	68.2995	70.1030	53.9794	68.2995	52.0412	64.6090	80.36
14	67.9588	70.3703	56.9020	71.8213	65.1055	61.5836	77.54
15	68.9432	69.2180	58.0618	72.0412	63.5218	60.8279	82.33
16	68.2995	68.8496	55.5630	67.6042	64.0824	60.0000	80.61
17	68.6273	70.0212	59.0849	69.8272	64.0824	60.8279	75.29
18	69.2480	69.1273	60.0000	53.9794	64.6090	66.8485	70.63
19	67.6042	69.2480	61.5836	61.5836	46.0206	65.1055	82.28
20	67.6042	69.9109	60.0000	58.0618	46.0206	62.9226	80.53
21	68.2995	69.6575	62.9226	46.0206	59.0849	56.9020	72.87
Avg	68.2429	69.9922	56.6050	64.7995	54.1594	62.7365	79.9067
St Dev	0.6723	0.6226	6.9816	8.2071	7.3371	3.7297	3.9578
Max	69.5424	70.9555	71.5957	78.6900	65.1055	71.3640	86.26
Min	66.8485	68.8496	40.0000	46.0206	46.0206	56.9020	70.63

2. A term containing knowledge of its personal ‘best’ position so far

   This term, c_1 **rand**() (**PBEST** – **X**), applies a weight (c_1) to the distance of the variable from its best position in all previous iterations (**PBEST**). This term is analogous to the personal ‘memory’ of the particle with respect to its best position in the direction of the optimum position.

3. A term containing knowledge of the global ‘best’ position

   This term, c_2 **rand**() (**GBEST** – **X**), applies a weight (c_2) to the distance of the particle from the global best position in all previous iterations (**GBEST**). This term is analogous to the ‘knowledge of the swarm’ with respect to the position of the particle relative to the best position found by the swarm.

The random number function, **rand**(), is the uniform probability density function in Matlab and returns a value between 0 and 1. The **rand**() value is different for each term and is intended to represent the unpredictability in swarm behavior.

Increasing the weights of any of the weighting factors w, c_1 or c_2 will encourage the particle to explore the sample space in the direction represented by the associated term. Conversely, decreasing the weights of any of the weighting factors w, c_1 or c_2 will discourage the particle from exploring the sample space in the direction represented by the associated term. The inertia weight, w, is generally assigned a value between 0 and 1. According to Robinson and Rahmat-Samii, a large inertia weight, w, tends to encourage global exploration even as the swarm approaches the optimum solution and for this reason they adopt the strategy of linearly decreasing w as the optimization progresses, so that global exploration reduces as the solution converges to optimum. Robinson and Rahmat-Samii decrease w linearly from 0.9 to 0.4 over the course of the optimization. This strategy is also adopted in this work. The values of the weighting factors, c_1 and c_2, chosen by Robinson and Rahmat-Samii are based on analytical analysis of PSO by Clerc and Kennedy [9] which resulted in a value of 1.49 for each. Those values for the weighting factors will also be used here.

Handling Constraints

As discussed above, the constraints are the ranges of the data from which the regression equation (the fitness function) was derived. Particles that violate these constraints are considered to be “infeasible”. Inevitably through the course of PSO the calculated velocities, if not handled accordingly, will carry some or all particles beyond the constraints into the infeasible region. Robinson and Rahmat-Samii discuss three ways of handling constraint violations. The first is to ignore the constraints with respect to the position of the particle and to not evaluate the fitness for infeasible values of **X** (the “Invisible Wall”) effectively eliminating the influence of the particle on the swarm. The second is to reverse the velocity of any particle that might otherwise violate a constraint (the “Reflecting Wall”), and finally, to set the velocity equal to zero once the particle achieves the constraint value (the "Absorbing Wall"). In this work, it is proposed that another new method (the "Arresting Wall") be considered where the velocity is set to zero if the particle would otherwise violate a constraint. This new philosophy may allow particles to explore regions closer to the constraint boundary. These methods of handling constraint violations are illustrated in Figure 1.

One of the problems with PSO can be the tendency of the particles to fly out of the solution space never to return. The control on the constraints is one way to prevent this. Another method for minimizing constraint violation is to limit the velocity within each iteration. Eberhart and Shi [10] propose limiting the velocity by introducing a maximum velocity, *vmax*, to be applied at each iteration where *vmax* is equal to 10% – 20% of the range of feasible values for each variable. Here, using a limiting factor of 10% of the range, this strategy is applied directly to the constraints in Table 1 to give the values for *vmax* for the optimization problem. These values are presented in Table 4.

Convergence to the Optimum Solution

As with any optimization method, PSO requires that the conditions for convergence be defined. Typically in an iterative optimization program convergence is defined to be the state when the optimum value has not changed, or has changed very little over the course of a pre-determined number of iterations or CPU time. In PSO, however, there is a tendency for the optimization to stagnate temporarily, sometimes for 100 iterations or more depending on the choice of parameters and the character of the fitness function, constraints and the sample space. For this reason, the researcher must take great care when defining the conditions for convergence and the parameters within PSO. In this work the solution must remain unchanged for 1,000 iterations to consider PSO converged to a solution. The next section presents the details of the solution using PSO.

Numerical Solution of the Optimization Problem by the Particle Swarm Optimization Method

The optimization problem is solved with Matlab. A detailed algorithm is shown in the Appendix. Although the number of iterations required for convergence is greater than needed for CE, as discussed earlier, there is a tendency of PSO to stagnate occasionally and there is very little penalty in allowing the routine to continue for this number of iterations. The program requires the following input:

1. Whether to use the 21 samples from the assembly line as the number of particles or to generate any desired number of particles using the values from the 21 samples to define the sampling distributions.

2. Which constraint philosophy to use throughout the run (Invisible Wall, Reflecting Wall, Arresting Wall or Absorbing Wall)

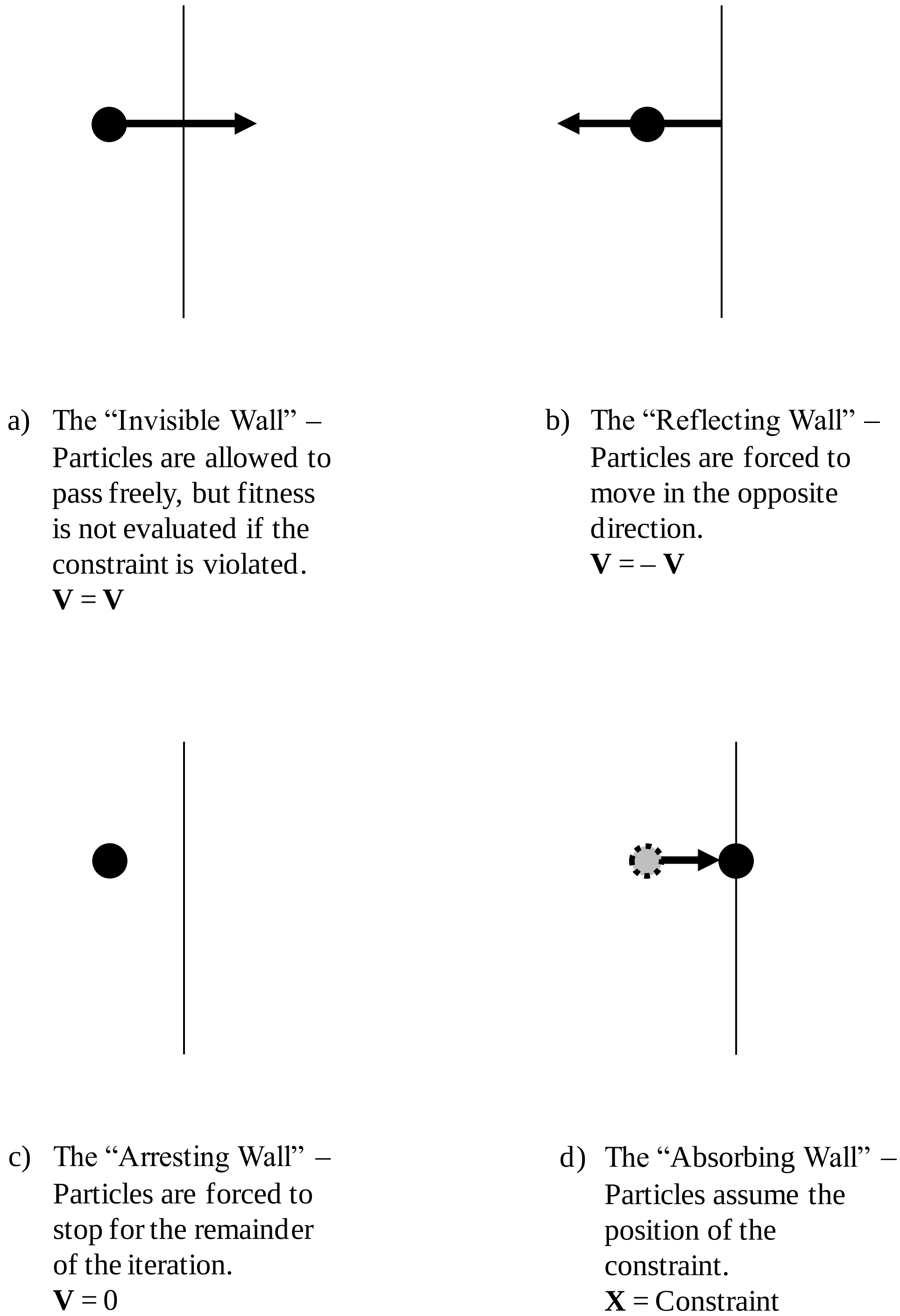


Figure 1 – Handling Constraint Violations – Four Methods (adapted from Robinson and Rahmat-Samii [8], © 2004 IEEE)

Table 4 – Values of vmax

Variable	***vmax***
a1	0.2694
a5	0.2106
c6	3.1596
c7	3.2669
d3	1.9085
d6	1.4462

3. The number of iterations required for convergence.

4. The maximum number of iterations before concluding that PSO will not converge.

5. The choice to review the detailed calculations of each iteration as they happen (otherwise only the iteration number, the number of iterations toward convergence and the current solver time used are displayed)

The program provides the following output:

1. All of the parameters for PSO including the user-input.

2. The iteration at which the optimum value was identified.

3. The optimum solution.

4. A file of GBEST for each iteration (**X** and **Y**) (how the solution progressed).

5. The average PBEST for each iteration (**X** an **Y**) (although not of much interest).

6. The average value of **X** for each iteration (the general position of the swarm).

7. The solver time used for the iteration.

Table 5 summarizes the setup parameters that are used for the optimization. The detailed Matlab code is included as an Appendix. The optimization is conducted in two trials with five runs each.

Table 5 – Setup Parameters for PSO

Description	Parameter	Value
Current Velocity Inertia Weight	*w*	0.9 - 0.4
Personal Best Weight	*c* 1	1.49
Global Best Weight	*c* 2	1.49
Stall Iterations Required	iConverge	1000
Maximum Iterations	iMax	20000
Number of Particles - Trial 1	N	21
Number of Particles - Trial 2	N	100

Due to the **rand**() function, it can be expected that the routine will not produce identical results with any subsequent run, at least in terms of the number of iterations and solver time. For this reason each run of the program is repeated four times - five runs in total. Table 6 is a summary of the runs and solutions used for solving the optimization problem with PSO using 21 particles with the data from the assembly line as the initial position. Table 7 is the same summary using 100 particles with the initial position drawn randomly from PDFs defined by the 21 samples data from the assembly line (the data from Table 2) as the initial position.

It is clear that this optimization problem did not benefit from the Reflecting Wall and Arresting Wall constraint strategies. The Absorbing Wall strategy did provide an improvement in run-time efficient with significantly fewer iterations required. The method saw no significant improvement in performance using a larger number of particles based on the original 21 samples. Figure 2 is another representation of the optimum solutions presented in Tables 6 and 7 that perhaps further clarify these points. Figure 3 is an illustration of the number of iterations required to complete the optimization for each method. Figure 4 is the solver time required for the method to converge at each run. It can be concluded that the Absorbing Wall constraint philosophy using the initial samples to define the particles identified the optimum value more efficiently than the other constraint philosophies, even with a larger population of particles.

Summary of the Solution by the Particle Swarm Optimization Method

This section presented the solution of the optimization problem using the Particle Swarm Optimization method. This section now concludes with a comparison of the best performance of the Particle Swarm Optimization Method, the Cross-Entropy Method and the deterministic solution. Table 8 shows the comparison. PSO using the Absorbing Wall constraint philosophy agrees exactly with the deterministic solution to four decimal places. However, PSO required twice the CPU time and more than 20 times more iterations. It remains, for future work, to confirm that axles built to the optimum conditions indeed produce improved vibration performance. Other papers by the authors illustrate solving the same optimization problem using the Cross Entropy Method [5] and Particle Swarm Optimization [6].

Table 6– Solutions of the Optimization Problem by the Particle Swarm Optimization Method n = 21

Run	Number of Particles	Constraint Philosophy	Iterations Required to Converge	Solver Time	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
1	21	Invisible	5728	170	69.5424	70.9555	71.5957	46.0482	65.1055	71.3640	58.1504
2	21	Invisible	7327	218	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
3	21	Invisible	5191	154	69.5424	70.9555	71.5957	46.0206	65.1054	66.6384	59.3797
4	21	Invisible	4601	137	69.5424	70.9555	71.5353	46.0206	65.1054	71.3640	58.1521
5	21	Invisible	4740	140	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
1	21	Reflecting	513	70	69.4846	70.3764	61.1302	46.1155	63.5278	66.5952	63.7399
2	21	Reflecting	222	30	69.2577	69.5087	60.8789	46.4431	65.0064	62.4810	67.6056
3	21	Reflecting	825	112	69.2236	70.9517	59.9873	48.2129	64.7831	67.1225	63.3564
4	21	Reflecting	444	60	69.4951	69.8167	59.8196	46.1429	64.1270	68.8239	64.7774
5	21	Reflecting	868	119	69.5391	70.2467	60.2865	46.1620	63.8339	68.7466	63.4770
1	21	Arresting	9538	284	69.5424	70.9534	60.5298	46.0378	64.7733	71.2373	60.4696
2	21	Arresting	18985	564	69.5424	70.9003	71.5957	46.5591	65.1054	68.2686	59.3053
3	21	Arresting	8841	263	69.5319	69.9192	71.5957	46.0206	64.3967	66.8889	62.4449
4	21	Arresting	8859	264	69.4087	70.9547	71.5957	46.0206	64.2124	66.7476	59.9944
5	21	Arresting	7656	228	69.4626	70.8970	61.2003	46.0206	65.1054	68.0458	61.4502
1	21	Absorbing	1402	42	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
2	21	Absorbing	1603	47	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
3	21	Absorbing	2124	64	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
4	21	Absorbing	2137	63	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
5	21	Absorbing	1041	31	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

Table 7 – Solutions of the Optimization Problem by the Particle Swarm Optimization Method n = 100

Run	Number of Particles	Constraint Philosophy	Iterations Required to Converge	Solver Time	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
1	100	Invisible	5444	401	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
2	100	Invisible	7532	554	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
3	100	Invisible	5247	385	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
4	100	Invisible	5165	379	69.5424	70.9555	71.5955	46.0207	65.1054	71.3640	58.1403
5	100	Invisible	5777	425	69.5424	70.9555	71.5957	46.0206	65.1054	71.3640	58.1402
1	100	Reflecting	1795	132	69.5357	70.9494	70.8944	47.0921	64.9142	68.0164	59.6458
2	100	Reflecting	830	61	69.3352	70.3283	70.3594	46.0279	64.4072	70.2936	61.2060
3	100	Reflecting	126	9	68.9376	70.5637	64.0866	46.2584	60.3654	68.8608	64.5685
4	100	Reflecting	201	15	69.4146	70.2294	58.3638	49.4231	64.8848	64.1467	66.3538
5	100	Reflecting	118	9	69.3322	70.8433	61.6919	55.2147	60.6359	68.7659	66.4116
1	100	Arresting	7579	557	69.1517	70.9369	71.5957	46.0206	60.9479	58.5511	63.8873
2	100	Arresting	12844	943	69.5424	70.9532	71.5631	46.0206	58.0015	66.1168	61.6453
3	100	Arresting	1979	145	69.5424	70.9489	53.1735	46.0749	58.5073	61.7258	66.3095
4	100	Arresting	6552	482	69.5424	69.8201	68.5383	46.0206	65.1028	61.3583	64.5359
5	100	Arresting	1830	135	67.9886	70.9242	71.5957	50.2843	64.8387	63.6681	66.2609
1	100	Absorbing	1715	127	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
2	100	Absorbing	820	60	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
3	100	Absorbing	1119	83	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
4	100	Absorbing	1220	89	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
5	100	Absorbing	986	73	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

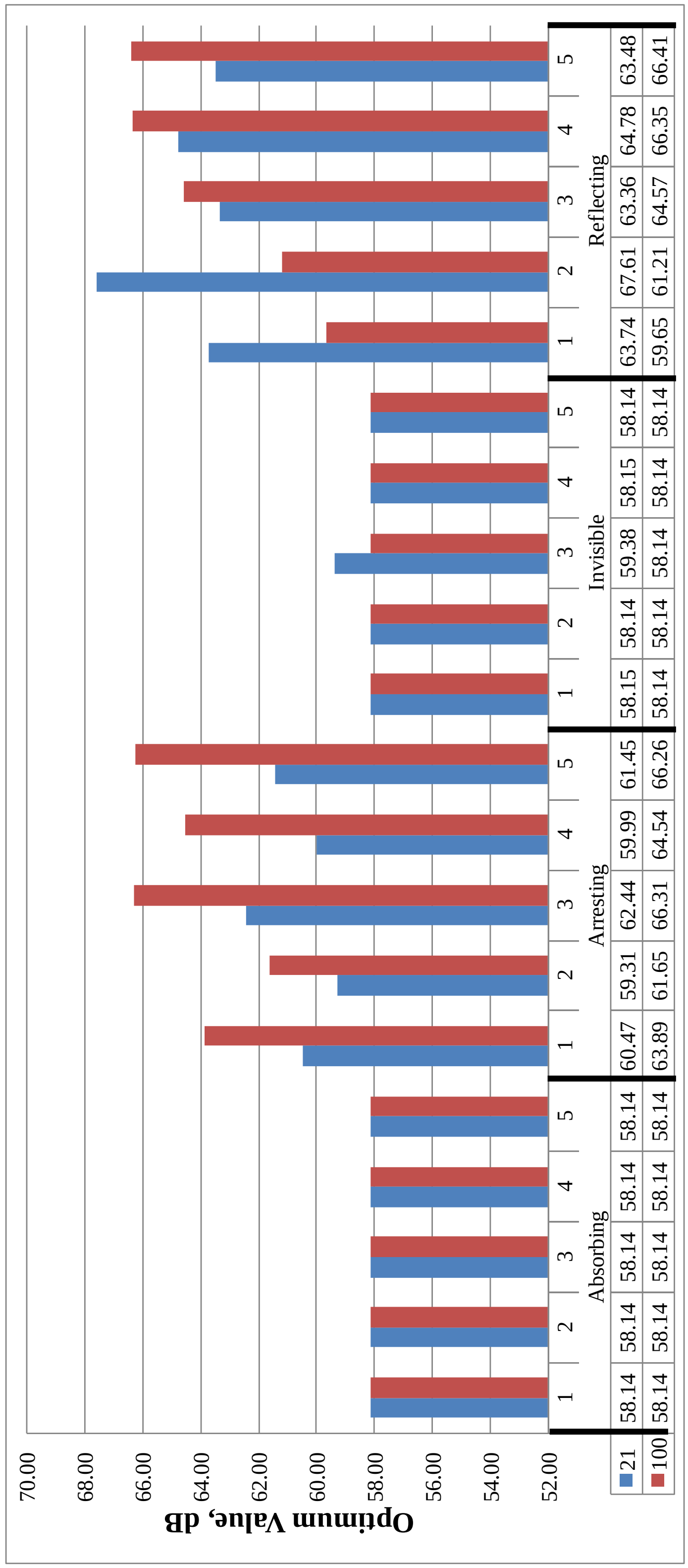
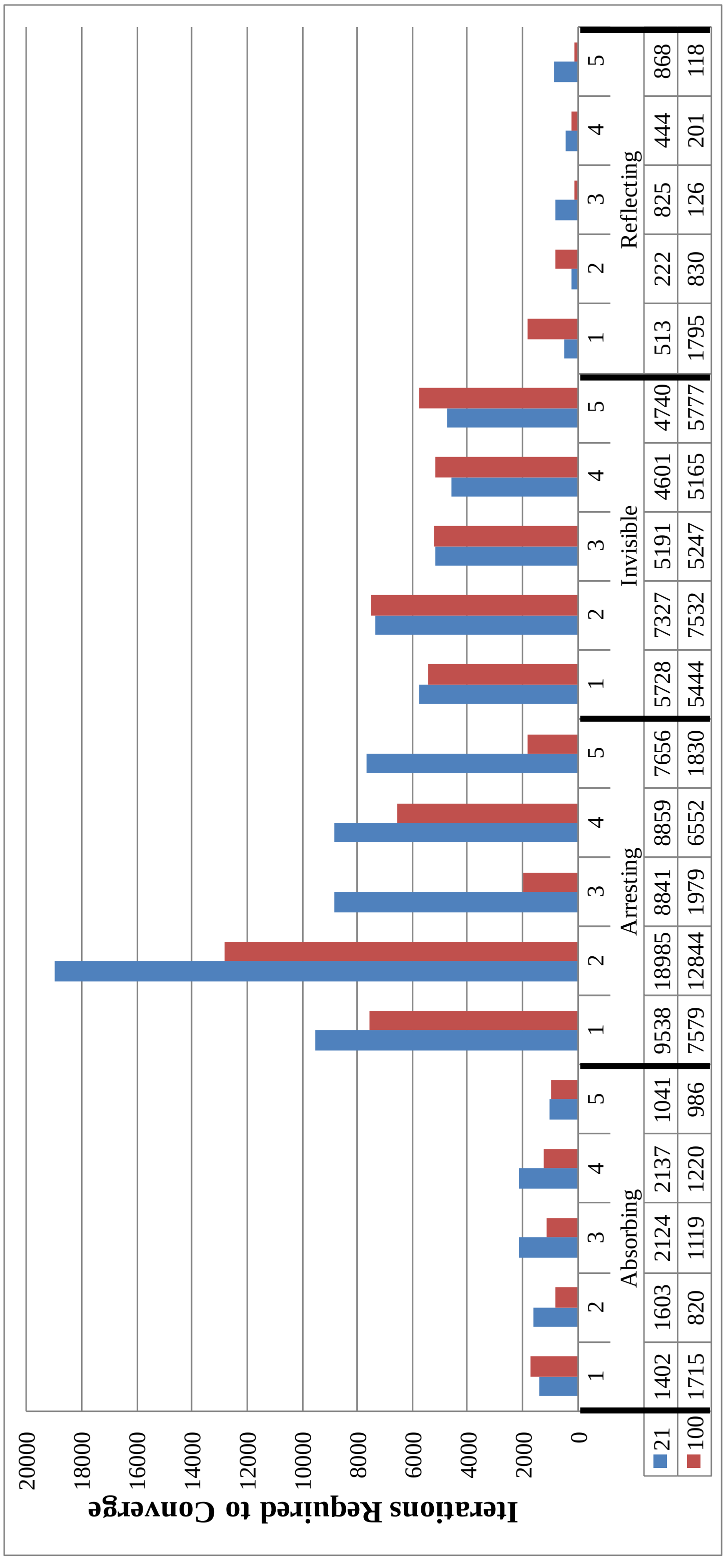
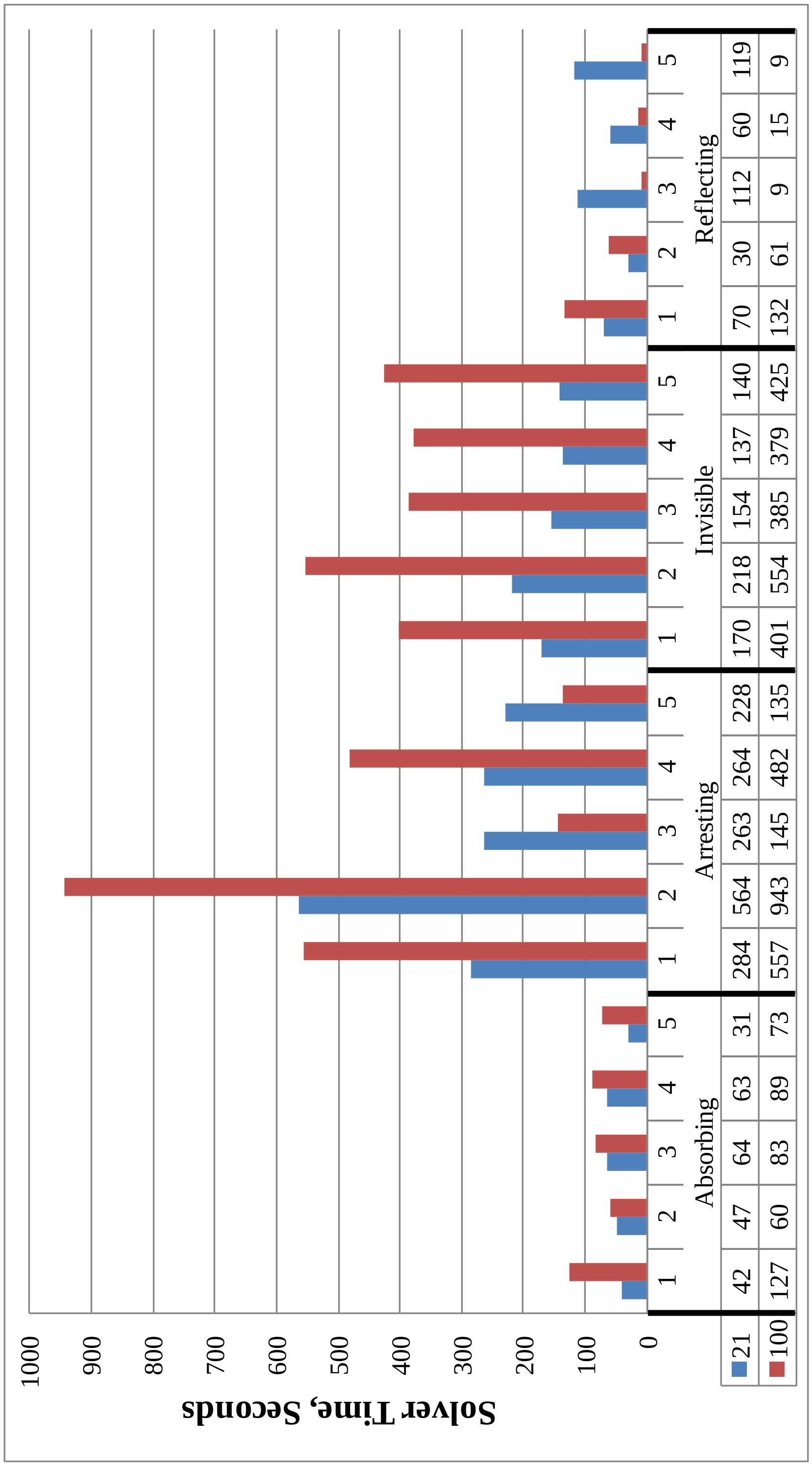


Figure 2 – Results of Particle Swarm Optimization by Run Number and Constraint Philosophy



	Absorbing 1	Absorbing 2	Absorbing 3	Absorbing 4	Absorbing 5	Arresting 1	Arresting 2	Arresting 3	Arresting 4	Arresting 5	Invisible 1	Invisible 2	Invisible 3	Invisible 4	Invisible 5	Reflecting 1	Reflecting 2	Reflecting 3	Reflecting 4	Reflecting 5
21	1402	1603	2124	2137	1041	9538	18985	8841	8859	7656	5728	7327	5191	4601	4740	513	222	825	444	868
100	1715	820	1119	1220	986	7579	12844	1979	6552	1830	5444	7532	5247	5165	5777	1795	830	126	201	118

Figure 3 – Iterations Required for Convergence of Particle Swarm Optimization (by Run Number and Constraint Strategy)



	Absorbing 1	Absorbing 2	Absorbing 3	Absorbing 4	Absorbing 5	Arresting 1	Arresting 2	Arresting 3	Arresting 4	Arresting 5	Invisible 1	Invisible 2	Invisible 3	Invisible 4	Invisible 5	Reflecting 1	Reflecting 2	Reflecting 3	Reflecting 4	Reflecting 5
21	42	47	64	63	31	284	564	263	264	228	170	218	154	137	140	70	30	112	60	119
100	127	60	83	89	73	557	943	145	482	135	401	554	385	379	425	132	61	9	15	9

Figure 4 – Solver Time Required for Convergence of Particle Swarm Optimization (by Run Number and Constraint Strategy)

Table 8 – A Comparison of the Deterministic Solution of the Optimization Problem to the Best Performance of CE

Optimization Method	N	Iterations to Solve	Solver Time (sec)	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	0	0	< 10	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
PSO - Absorbing Wall	100	820	60	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

Appendix

A Detailed Algorithm for Particle Swarm Optimization

The implementation follows the algorithm shown below.

1. Initialize the particles (**X** and **Y**) from Table 3.
2. Initialize the constraints and *vmax* from Table 2 and Table 4.
3. Set the parameters w, c_1 and c_2.
4. Calculate the new velocities for each particle.
    a. Allow w to decrease from 0.9 to 0.4 in increments of 0.005 (100 iterations)
5. Calculate the new positions based on the new velocities.
6. Check feasibility.
7. Calculate the fitness value of each feasible solution using Equation (1).
8. Check each solution against the particle's Personal Best (PBEST) and update PBEST if the new position yielded better results.
9. Check each solution against the Global Best (GBEST) and update if any of the solutions yielded better results.
10. Repeat from Step 4 until 1,000 iterations yield no change in GBEST.

The Matlab Code for Particle Swarm Optimization for Solving the Optimization Problem

```
%   Revised 03/22/2013
```

```
warning('off','all');
warning;

%   Open the data files

O = load('nvhdata.mat', '-ASCII')
constraint = load('const.mat', '-ASCII')

%   Get the range of the data and establish vmax

[ic,jc] = size(constraint);
j = 1;

while j < jc +1
    vmax(j) = 0.1 * (constraint(1,j) - constraint(2,j));
    j = j+1;
end


%   Establish the number of variables (c) and the number of samples (r)

[r,c] = size(O);


%   Make room for the iteration number in the arrays

c = c + 1;

%   Initialize the parameters

usex = 0;
detail = 0;
n = 0;
const = 1;
usex = input('Enter ''1'' to use the original data as the particles. ');
if usex > 0
    n = 21;
end
if n == 0
    n = input('How many random samples shall we generate? ');
end
const = input('How shall constraints be handled?\n( 1 = invisible, 2 =
reflecting, 3 = arresting, 4 = absorbing) ');
gbmax = input('How many iterations are required to establish convergece? ');
imax = input('What is the maximum number of iterations allowed? ');
detail = input('Enter ''1'' if you would like to see the detailed
calculations. ');

%   Default PSO velocity parameters
```

```
a1 = 1.49;
a2 = 1.49;
w = 0.9;

%   Enable these input statements to allow changes to the PSO velocity
%   parameters

% a1 = input('What weight shall be applied to the PBEST value? ');
% a2 = input('What weight shall be applied to the GBEST value? ');
% w = input('What weight shall be applied to the velocity? ');
% stop = input('What is the mimimum change in Y you are looking for before
stopping the program? ');
% stopn = input('Over how many iterations? ');

%   Initialize the arrays and variables

winit = w;
rr1(n,c)=0;
rr2(n,c)=0;
X(n,c)=0;
XP(n,c)=0;
V(n,c)=0;
VP(n,c)=0;
PBEST(n,c)=0;
GBEST(c)=0;
GBESTS(n,c)=0;
AVERAGES(n,c)=0;
PAVERAGES(n,c)=0;
avfile(1,c)=0;
pavfile(1,c)=0;
gfile(1,c)=0;
gflag = 0;
pflag = 0;
count=1;
iopt=0;
opt=0;
write=0;
gbcount=1;
dcount=imax;
d4=('Agrees with Classical Solution? ');
agree={'No'};

%   Initialize the counters for the datafile stats

err = 0;
x = 0;
px = 0;
i = 1;
j = 1;

%   Calculate the averages
```

```
AVERAGES(1,1)=0;
while j < c
while i < r + 1
    x = x + O(i,j);
    i = i + 1;
end
Average(j) = x / r;
AVERAGES(count,j+1) = Average(j);
x = 0;
i = 1;
j = j + 1;
end
Average

%   Calculate the Standard Deviations

i = 1;
j = 1;

while j < c
while i < r + 1
    err = err + (O(i,j) - Average(j))^2;
    i = i + 1;
end
StDev(j) = sqrt(err / (r-1));
err = 0;
i = 1;
j = j + 1;
end
StDev

%   Initialize the optimization values PBEST and GBEST

Ymin = 10,000;
i = 1;

while i < n + 1
    X(i,1) = i;
    XP(i,1)=i;
    PBEST(i,1)=i;
    GBEST(1)=i;
    i = i + 1;
end

i = 1;
j = 1;

if r < n
    ri = r;
else
```

```
    ri = n;
end

while i < ri + 1
while j < c

% Initialize PBEST

    X(i,j + 1) = O(i,j);
    XP(i,j + 1) = O(i,j);
    PBEST(i,j + 1) = O(i,j);
    j = j + 1;
end

% Initialize GBEST

     if X(i,8) == 0 || X(i,8) < Ymin
         Ymin = X(i,8);
         jj = 1;
         while jj < c + 1
            GBEST(jj) = X(i,jj);
             jj = jj + 1;
         end
     end
j = 1;
i = i + 1;
end

%   Write GBESTS to an array

gfile(1,1)=0;
j = 2;
while j < c + 1
GBESTS(count,j)=GBEST(j);
j = j + 1;
end

%   Write the initial run

if detail == 1
    iteration = count - 1
    X
    XP
    V
    VP
    PBEST
    GBEST
end

%
%
```

```
%   Begin Particle Swarm Optimization
%
%

count = count + 1;

while count < imax + 2

    GBESTS(count,1)=count - 1;
    AVERAGES(count,1)=count - 1;
    PAVERAGES(count,1)=count -1;
    if detail == 1
        iteration = count - 1
    end
i = 1;

confail = 0;

while i < n + 1
    j = 2;
    while j < 8

%   Generate the velocity and position for each particle

        %   Calculate the position for each particle

        %   If random samples will be the particles, generate the initial
        %   'n' sample postions

if count == 2 && usex < 1
        X(i,j) = random('norm', Average(j-1), StDev(j-1));
else

        %   Calculate the new velocity for each particle

        rr1(i,j) = rand(1);
        rr2(i,j) = rand(1);

        if w > 0.4

            w = w - 0.005;

        end

        %a1 = 2;
        %a2 = 2;
        %w = 0.9;

        V(i,j) = w * VP(i,j) + a1 * rr1(i,j) * (PBEST(i,j) - XP(i,j)) + a2 *
rr2(i,j) * (GBEST(j) - XP(i,j));
```

```
%   Don't let the velocity exceed vmax

if abs(V(i,j)) > vmax(j-1)
    V(i,j)= V(i,j) / abs(V(i,j)) * vmax(j-1);
end

%   Calculate the new position for each particle

X(i,j) = XP(i,j) + V(i,j);

%   Check the feasibility of the new position at the upper boundary


if X(i,j) + V(i,j) > constraint(1,j-1)


    %   The "invisible wall" philosophy

     if const == 1
         di = {'Invisible'};
        %display(di)
     end

    %   The "reflecting wall" philosophy

     if const == 2
        X(i,j) = XP(i,j) - V(i,j)/i;
        di = {'Reflecting'};
        %display(di)
     end

     %   The "arresting wall" philosophy

     if const == 3
        X(i,j) = XP(i,j);
        di = {'Arresting'};
        %display(di)
     end

     %   The "absorbing wall" philosophy

     if const == 4
        X(i,j) = constraint(1,j-1);
         di = {'Absorbing'};
        %display(di)
     end

end
```

```
        %   Check the feasibility of the new position at the lower boundary

        if X(i,j) + V(i,j) < constraint(2,j-1)

             %   The "invisible wall" philosophy

             if const == 1
                 di = {'Invisible'};
                %display(di)
             end

            %   The "reflecting wall" philosophy

             if const == 2
                X(i,j) = XP(i,j) - V(i,j)/i;
             end

             %   The "arresting wall" philosophy

             if const == 3
                X(i,j) = XP(i,j);
             end

             %   The "absorbing wall" philosophy

             if const == 4
                X(i,j) = constraint(2,j-1);
             end

        end

            if X(i,j) > constraint(1,j-1) || X(i,j) < constraint(2,j-1)
                confail = 1;
            end

end
        j = j +1;
    end

%   Calculate the fitness value

if confail == 0

     X(i,8) = -2.8049 * X(i,2) - 2.789 * X(i,3) - 0.19745 * X(i,4) + 0.36987
* X(i,5) - 0.29785 * X(i,6) - 0.2623 * X(i,7) + 486.32;
else
    X(i,8) = 1,000;

end
```

```
%   Check if this is the Personal Best

     if PBEST(i,8) == 0 || X(i,8) < PBEST(i,8)
         if X(i,8) > 0
            pflag = 1;
            pi = i;
         end
     end

% Check if this is the Global Best

     if X(i,8) < Ymin
         if X(i,8) > 0
            Ymin = X(i,8);
            if iopt < 1
            if X(i,8) > 58.14 && X(i,8) < 58.1405
                iopt = count -1;
                agree = {'Yes'};
            end
            end
            gflag = 1;
            gi = i;
         end
     end

     i = i + 1;
end

%   Write the Personal Best values for each iteration of the optimization

if pflag == 1
    j = 2;
    while j < c + 1
        PBEST(pi,j) = X(pi,j);
        j = j + 1;
    end

pflag = 0;

end

%   Write the Global Best Value for each iteration of the optimization

if gflag == 1
    j = 1;
    while j < c + 1
        GBEST(j) = X(gi,j);
        j = j + 1;
    end

gflag = 0;
```

```
end


%   Update GBESTS to include this iteration's Global Best

j = 2;
while j < c + 1
    GBESTS(count,j)=GBEST(j);
j = j + 1;
end

%   Calculate the averages for X and PBEST

i = 1;
j = 1;

while j < c
while i < r + 1
    x = x + X(i,j + 1);
    px = px + PBEST(i,j+1);
    i = i + 1;
end

AVERAGES(count, j + 1) = x / r;
PAVERAGES(count, j + 1) = px / r;
x = 0;
px = 0;
i = 1;
j = j + 1;
end

%   Record the X and V values as the previous values for the next iteration

i = 1;
j = 1;

while i < n +1
    while j < c +1
        XP(i,j) = X(i,j);
        VP(i,j) = V(i,j);
        j = j + 1;
    end
    j = 1;
    i = i + 1;
end

%   Display the results for this iteration (if desired)
```

```
if detail == 1
    rr1
    rr2
    X
    V
    XP
    VP
    PBEST
    GBESTS
end

if GBESTS(count,8) == GBESTS(count-1,8)
    gbcount = gbcount+1;
else
    gbcount = 1;
end

if gbcount >= gbmax

    display('Optimization has converged.')
    dcount = count;
    opt = dcount - gbmax;
    count = imax + 2;


else


if GBESTS(count,8) < 0.0,000

    display('The solution is negative.')
    dcount = count;
    count = imax + 2;

else

if count == imax +2

    display('The maximum allowable iterations has been reached.')

end
end
end


count = count + 1;

clc
Iteration = count - 2
gbcount
```

```
Optimum = GBEST(8)

end

%
%
%   End of Particle Swarm Optimization
%
%

%   Write the results

i = 1;
j = 1;

while i < dcount + 1
    while j < c +1
    avfile(i,j) = AVERAGES(i,j);
    pavfile(i,j) = PAVERAGES(i,j);
    gfile(i,j) = GBESTS(i,j);
    j = j + 1;
    end
    j = 1;
    i = i + 1;
end

%   Display the results


%AVERAGES;
%GBESTS;
%avfile
%pavfile
gfile
display('Agrees with classical solution?')
display(agree);
Iteration = dcount - gbmax
Optimum = GBESTS(opt + 1,8)
%gbcount
%imax
wall = di

%   Write to a spreadsheet?

write = input('Enter ''1'' to write this data to a spreadsheet. ');

if write == 1

    d1 = {'usex', 'n', 'iopt', 'gbmax', 'imax',
'const','vmax1','vmax2','vmax3','vmax4','vmax5','vmax6','w-init','w-
end','a1','a2'};
```

```
    d2 = ['i123456Y'];
    d3 = {'Optimum','Iteration','Agrees with Classical?'};
    xlswrite('PSO_Solution.xls',d1,'Parameters','A1');
    xlswrite('PSO_Solution.xls',usex,'Parameters','A2');
    xlswrite('PSO_Solution.xls',n,'Parameters','B2');
    xlswrite('PSO_Solution.xls',iopt,'Parameters','C2');
    xlswrite('PSO_Solution.xls',gbmax,'Parameters','D2');
    xlswrite('PSO_Solution.xls',imax,'Parameters','E2');
    xlswrite('PSO_Solution.xls',const,'Parameters','F2');
    xlswrite('PSO_Solution.xls',vmax,'Parameters','G2');
    xlswrite('PSO_Solution.xls',winit,'Parameters','M2');
    xlswrite('PSO_Solution.xls',w,'Parameters','N2');
    xlswrite('PSO_Solution.xls',a1,'Parameters','O2');
    xlswrite('PSO_Solution.xls',a2,'Parameters','P2');
    xlswrite('PSO_Solution.xls',d3,'Parameters','A3');
    xlswrite('PSO_Solution.xls',Optimum,'Parameters','A4');
    xlswrite('PSO_Solution.xls',Iteration,'Parameters','B4');
    xlswrite('PSO_Solution.xls',agree,'Parameters','C4');
    xlswrite('PSO_Solution.xls',di,'Parameters','F3');
    xlswrite('PSO_Solution.xls',d2,'GBESTS','A1');
    xlswrite('PSO_Solution.xls',d2,'PBESTS','A1');
    xlswrite('PSO_Solution.xls',d2,'X-Averages','A1');
    xlswrite('PSO_Solution.xls',gfile,'GBESTS','A2');
    xlswrite('PSO_Solution.xls',pavfile,'PBESTS','A2');
    xlswrite('PSO_Solution.xls',avfile,'X-Averages','A2');

end

% End of test.  Ask to clear the memory.

reply = input ('Do you want to clear everything? (y/n)[n]', 's');
if reply == 'y'
    clear
    clc
elseif isempty(reply)
    reply = 'n';
end
```

Bibliography

1. Sun, Z., *et. al.*, "NVH Robustness Design of Axle Systems," *SAE Transactions*, v. 112, pp. 1746 – 1754, 2003.

2. Steyer, G., *et. al.*, "The Future of NVH Testing – An End-User's Perspective," *SAE Technical Paper 2005-01-2270*, 2005.

3. Meinhardt, G. and Sengupta, S., "Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part I-Preparing the Multivariate data for Regression Analysis," *SAE Technical Paper 2012-01-727*, 2012.

4. Meinhardt, G. and Sengupta, S., "Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part II-Multivariate Regression Analysis," *SAE Technical Paper 2012-01-728*, 2012.

5. Meinhardt G. and Sengupta, S., "Optimization of Axle NVH Performance Using the Cross Entropy Method," *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

6. Meinhardt G. and Sengupta, S., "Optimization of Axle NVH Performance Using A Genetic Algorithm," *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

7. Eberhart, R., and Kennedy, J., "A New Optimizer Using Particle Swarm Theory," Sixth International Symposium on Micro Machine and Human Science, pp. 39 – 43, 1995.

8. Robinson, J., and Rahmat-Samii, Y., "Particle Swarm Optimization in Electromagnetics," *IEEE Transactions on Antennas and Propagation*, v. 52, no. 2, pp. 397 – 407, 2004.

9. Clerc, M., and Kennedy, J., "The Particle Swarm-Explosion, Stability and Convergence in Multi-Dimensional Space," *IEEE Transactions on Evolutionary Computing*, February 2002

10. Eberhart, R., and Kennedy, J., "A New Optimizer Using Particle Swarm Theory," Sixth International Symposium on Micro Machine and Human Science, pp. 39 – 43, 1995.

Optimization of Axle NVH Performance Using A Genetic Algorithm

Glenn Meinhardt
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: gameinha@oakland.edu

Sankar Sengupta
Department of Industrial and Systems Engineering
Oakland University
Rochester, Michigan 48309
Email: ssengupta@oakland.edu

Abstract

An approach to optimization of automobile axles for noise, vibration and harshness (NVH) performance based on end-of-line testing is presented. The method used, a genetic algorithm (GA), iteratively solves an objective function based on elite selections and mutation of a range of solutions coded as an analog to the chromosome in biology. The GA routine within Matlab is used; the algorithm and solutions are presented and discussed.

Introduction

Noise, Vibration and Harshness (NVH) performance is a critical quality characteristic for automobile manufacturers (original equipment manufacturers, or OEMs) and driveline component manufacturers alike. A major component of the driveline is the axle. The axle transfers torque from the engine and driveshaft to the wheels. For axle manufacturers, one of the primary NVH metrics is gear whine [1]. To ensure satisfactory gear whine performance when the automobile leaves the factory, many OEMs now require axle assemblies to be tested for gear whine performance at the end of the assembly line using an end-of-line NVH test (EOLT) prior to shipment to their assembly plants. It is in the best interest of both the OEMs and axle manufacturers to ensure that the vibration levels of axles not only meet the requirement at the EOLT, but that the levels are as low as possible [2]. One way to control the levels at the EOLT is to understand the correlation of the upstream performance variables to the EOLT result. A previous work by the authors examined one such correlation [3, 4] involving the assembly parameters of the axle and the resulting coast-side vibration. This work illustrates the use of the cross entropy method to minimize the EOLT result with the regression equation presented in [4] used as the objective function. The solution of the same problem is presented in other works by the authors using Particle Swarm Optimization [5] and the Cross Entropy Method [6].
The Optimization Problem

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

The desire is to minimize the coast-side vibration given by the regression equation from [4]:

$$
\begin{aligned}
Y(X) \;=\; & 486.32 \\
& - 2.8049\ a1 \\
& - 2.7890\ a5 \\
& - 0.19745\ c6 \\
& + 0.36987\ c7 \\
& - 0.29785\ d3 \\
& - 0.26230\ d6
\end{aligned} \qquad (1)
$$

Therefore, the optimization problem is written

$$\gamma^* = \min Y(X) = Y(X^*) \qquad (2)$$

where γ^* is the optimum value of Y, $X' = [\ a1\ a5\ c6\ c7\ d3\ d6\ ]$, and X* are the values of X associated with γ^*. The regression equation was derived from 21 samples of data collected from the assembly line. Clearly the regression equation is only valid for the range of data from which it was derived. Therefore, the boundary conditions (constraints) for the optimization problem are the range of each variable from which the objective function was derived. The constraints are taken from range of data in [4] and summarized in Table 1. Equations (1) and (2) along with Table 1 completely define the optimization problem. Now, this problem can be solved very easily deterministically and that solution is given in Table 2. The purpose of this work is to illustrate how the cross entropy method can be used to solve optimization problems. The simple problem presented above and the deterministic solution can be used as a basis for such an illustration, the results of the optimization compared to the deterministic solution.

Table 1 – The Constraints for the Parameters, **X**, of the Regression Equation (Various Units, dB)

Variable	Lower Bound	Upper Bound
a1	66.8485	69.5424
a5	68.8496	70.9555
c6	40.0000	71.5957
c7	46.0206	78.6900
d3	46.0206	65.1055
d6	56.9020	71.3640

Genetic Algorithms

Genetic Algorithms (GA) were developed by John Holland in 1975 [7] and belong to the larger class of Evolutionary Algorithms. The method is analogous to the evolution of biological species where populations of solutions "evolve" toward better solutions, and ultimately to a

Table 2 – The Solution to the Deterministic Form of the Optimization Problem

Optimization Method	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402

single optimum solution, through a selection process using an objective function which provides a measure of fitness for each solution. The general algorithm for GA is:

1. Define a population of solutions.
2. Calculate the fitness level of each solution.
3. Sort the population according to fitness level, noting the optimum value.
4. Copy an elite portion to initialize the next generation.
5. Select parents and generate offspring and mutations to complete the next generation.
6. Calculate the fitness for the new members of the next generation.
7. Sort the new population according to fitness level, noting the optimum value.
8. Compare the new optimum value to the previous optimum value.
9. Evaluate the stopping criteria, and repeat from Step 4 if required.

The detailed algorithm is available in the Appendix. The key to the algorithm lies in defining each population, coding each solution and developing the evolutionary rules. The process begins with defining the problem in terms of the fitness function and necessary constraints.

Defining the fitness function and the constraints for feasible solutions

Each solution must be assigned a fitness value based on its proximity to the optimum solution. For this problem, the optimum solution is a minimum value of the coast-side vibration. The fitness function within GA then is the regression equation given by Equation (1). The constraints are the range of values from which Equation (1) was derived as shown in Table 1. The next step is to define the initial population.

Defining the Initial Population

The population size remains constant throughout the GA optimization process. The most-fit individuals survive, and the less-fit are replaced with new individuals bred from previous populations. The researcher defines the size of the population either by a group of known solutions or by choosing a number, N, of randomly generated solutions found feasible by the boundary conditions with calculated fitness. The latter method is typically chosen since it is generally difficult to generate a large population of known solutions. For this work, population sizes of N = 21, 100 and then 1,000 individuals is examined. The initial populations for N = 21 are the initial samples. The initial populations for N = 100 and 1,000 will include the 21 samples, but the remaining individuals are generated randomly based on the boundary conditions defined by Table 1. Next GA requires selecting a coding scheme for the individual solutions which transform the real value to its "genetic" equivalent.

Coding the Individual Solutions

GA works by identifying the best solutions based on their fitness value, combining their traits to develop an even better solution. To do this, the researcher must devise a coding scheme that can be used to pass parameter values within the solution between two or more solutions. Any coding scheme can be used, but the most common approach is the binary coding scheme whereby numeric values for each parameter are coded as their binary equivalent. Each bit of the binary string is analogous to a single "gene". Each binary string is analogous to an "allele" and a group of alleles (e.g., one complete solution) is analogous to a "chromosome" or one individual. Consider the first sample from Table 3 and its value for the variable a1. The decimal value is 67.9588 dB. To code this as a binary string it is first necessary to convert it to its integer equivalent by multiplying by 10,000. This yields 679,588 with a binary equivalent of 10100101111010100100. This string is the allele for variable a1. The complete solution (individual) is defined by stringing together alleles for each parameter. Figure 1 illustrates this by showing how the alleles for the first three parameters a1, a5 and c6 for Sample 1 from Table 3 are combined. This combination is half of the solution for Sample 1. The coding scheme shown above is a 20 bit binary allele. For this problem, once combined with this coding scheme, each individual becomes a 6 allele chromosome comprised of 120 genes. It is clear from Figure 1 that with the addition of variables and precision, the solution string grows very quickly. For this reason, GA becomes increasingly inefficient with large numbers of variables. With the coding scheme defined, the population is generated, the fitness value for each solution is calculated and the solution chromosome is coded. Next, the population is evolved to a new generation by retaining an elite portion of the current population, selecting and combining "parent" solutions and generating "offspring." continually pass to future generations unless displaced by a better solution. When this optimum value changes very little over a specified number of iterations, the method is continually pass to future generations unless displaced by a better solution. When this optimum value changes very little over a specified number of iterations, the method is considered converged. Convergence is discussed in detail in future sections. The elite portion is simply some proportion, ρ, such that the number of elite samples is ρN. Here, a value of ρ =

Table 3 – The Initial Data From the Assembly Line, Various Units, dB

Sample	a1	a5	c6	c7	d3	d6	NVH
1	67.9588	69.5134	71.5957	60.0000	46.0206	61.5836	81.03
2	66.8485	70.9555	53.9794	56.9020	58.0618	60.8279	78.18
3	67.6042	70.2377	53.9794	66.0206	49.5424	58.0618	86.26
4	68.2995	70.6296	55.5630	62.9226	49.5424	63.5218	76.49
5	67.2346	70.0212	66.4444	60.8279	52.0412	60.0000	82.29
6	69.5424	70.4228	56.9020	60.8279	46.0206	59.0849	75.86
7	68.6273	69.9662	46.0206	62.2789	62.9226	62.9226	80.95
8	68.6273	69.1273	52.0412	59.0849	53.9794	60.0000	80.22
9	68.2995	70.6040	40.0000	70.3703	49.5424	66.8485	83.64
10	68.6273	70.3703	58.0618	77.5012	49.5424	63.5218	83.44
11	68.9432	70.9309	60.0000	78.6900	46.0206	70.1030	82.09
12	67.6042	70.5526	46.0206	76.1236	49.5424	71.3640	85.15
13	68.2995	70.1030	53.9794	68.2995	52.0412	64.6090	80.36
14	67.9588	70.3703	56.9020	71.8213	65.1055	61.5836	77.54
15	68.9432	69.2180	58.0618	72.0412	63.5218	60.8279	82.33
16	68.2995	68.8496	55.5630	67.6042	64.0824	60.0000	80.61
17	68.6273	70.0212	59.0849	69.8272	64.0824	60.8279	75.29
18	69.2480	69.1273	60.0000	53.9794	64.6090	66.8485	70.63
19	67.6042	69.2480	61.5836	61.5836	46.0206	65.1055	82.28
20	67.6042	69.9109	60.0000	58.0618	46.0206	62.9226	80.53
21	68.2995	69.6575	62.9226	46.0206	59.0849	56.9020	72.87
Avg	68.2429	69.9922	56.6050	64.7995	54.1594	62.7365	79.9067
St Dev	0.6723	0.6226	6.9816	8.2071	7.3371	3.7297	3.9578
Max	69.5424	70.9555	71.5957	78.6900	65.1055	71.3640	86.26
Min	66.8485	68.8496	40.0000	46.0206	46.0206	56.9020	70.63

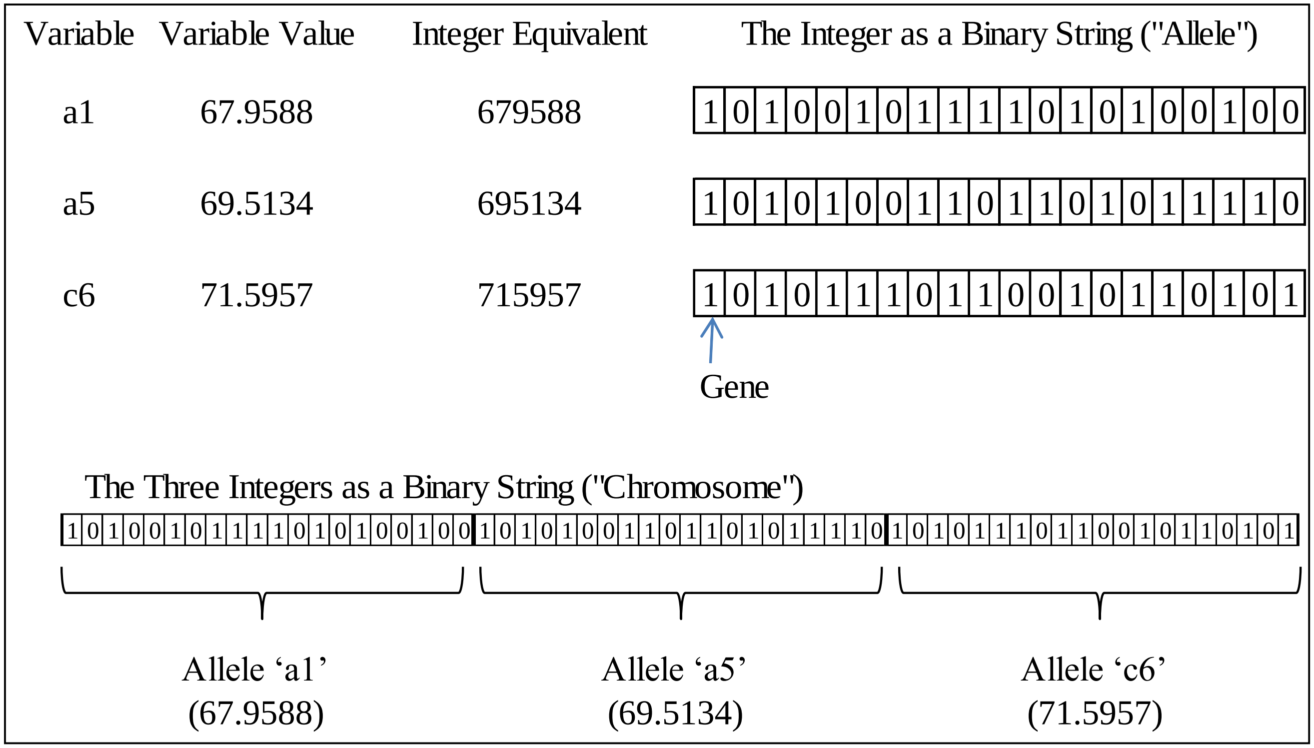


Figure 1 – Coding a Solution in GA – The First Three Alleles (a1, a5 and c6) for Sample 1 from Table 3

0.20 is chosen and the remaining (1 – ρ) N individuals will be generated by selecting parents and generating offspring or through mutation.

Selecting the Parents of Children of the Next Generation

Each individual of the current population is eligible to become a parent and produce offspring. However, to speed convergence, parents can be selected by any number of methods. Three strategies are popular in the literature. In one strategy, called Roulette Wheel Selection, the parents are selected completely at random. In another strategy, called Tournament Selection, a group is selected at random and, based on fitness level, the top two individuals become parents. A third strategy, which might be called Elitist Selection, has the parents selected randomly from the elite portion of the population. Table 4 summarizes these strategies for selecting parents. Once the parents are selected, GA then uses one of a number of strategies for generating offspring from these parents.

Populating the Next Generation with Children of the Current Generation

After the next generation is initially populated with the ρ N elite samples, the remaining (1 – ρ)N individuals are generated either from combinations of solutions from the current generation, which is discussed in this section, or through mutations which is the subject of the next section. Typically only two parents are combined to generate children, however any number of parents can be used to generate offspring. Offspring are generated by an analog of evolutionary genetic crossover whereby a portion of the chromosomes of each parent is passed along to the child. GA uses a crossover code to define the child based on the parents. With two parents, the crossover code could be simply a binary string with each bit comprising a decision for each gene (or allele) of a chromosome. In the crossover code, a value of '1' indicates that the gene from Parent 1 is passed, a value of '0' results in the gene from Parent 2 passing to the child. Figure 2 illustrates the most common methods for crossover. In the simplest case, Single Point Crossover shown in Figure 2a, crossover happens at a single point (the middle) and the first half of the chromosome from Parent 1 and the last half of the chromosome of Parent 2 are passed to the child. Two Point Crossover is illustrated in Figure 2b and works in a manner similar to Single Point Crossover. In fact, crossover can occur at any number of points and even randomly, also called Scatter Crossover. Figure 2c illustrates Scatter Crossover. In Scatter Crossover, the genes or alleles are passed according to a crossover code that is generated at random. For the problem of

Table 4 – Methods for Selecting Parents in GA

Description	Selection Method
Roulette Wheel	Completely random pairs
Tournament	Top pair from a random group
Elitist	Pairs selected from the elite portion

optimizing coast-side vibration, crossover will result in transferring alleles as a whole from the parents to the child. Figure 3 illustrates the production of one child from the parents Sample 1 and Sample 2 from Table 3. For brevity, only the first three alleles are shown. Generally the researcher specifies the percentage of the next population that is created by children of the current generation. The remainder is either generated at random (as the initial population was created) or through mutation of members of the current population.

Mutation and Other Controls on Offspring

In mutation, one or more genes may randomly change or mutate to another value according to some method defined by the researcher. In general, the researcher defines either the portion of the population that will be mutated, or the probability that a mutation will occur in an individual,

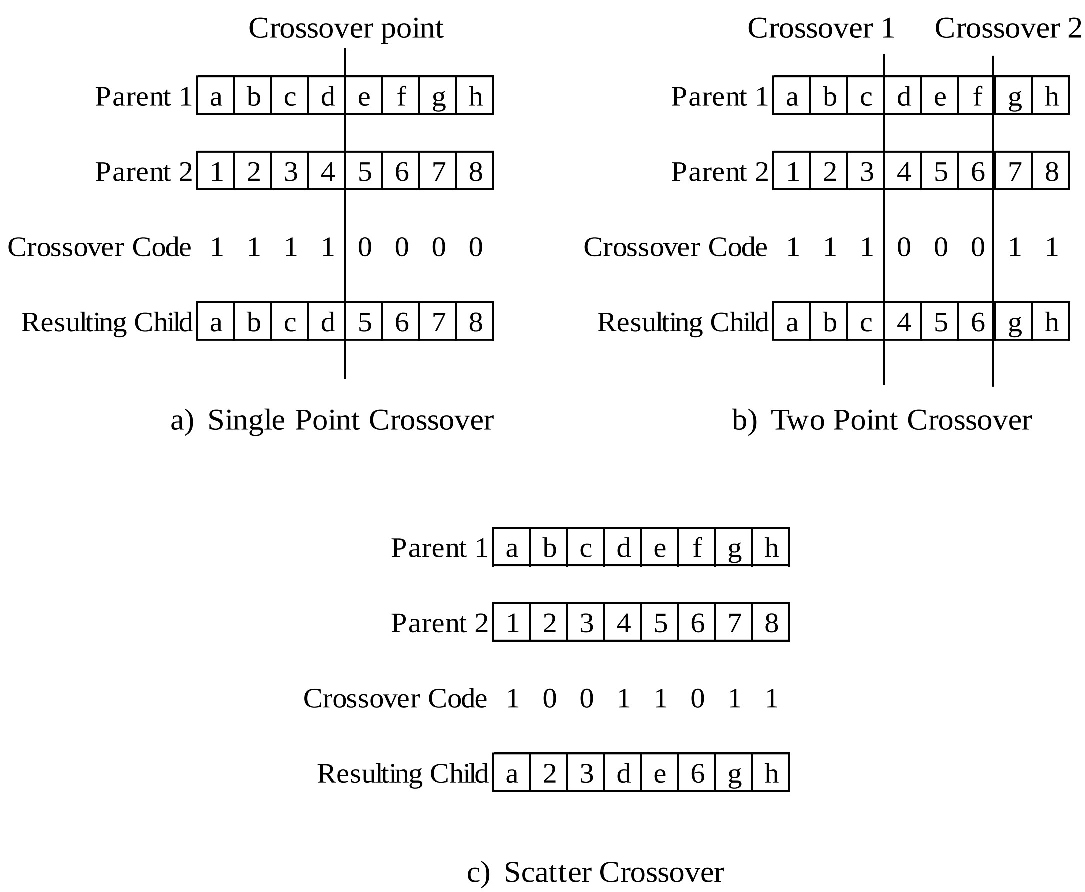


Figure 2 – Typical Methods for Crossover in GA

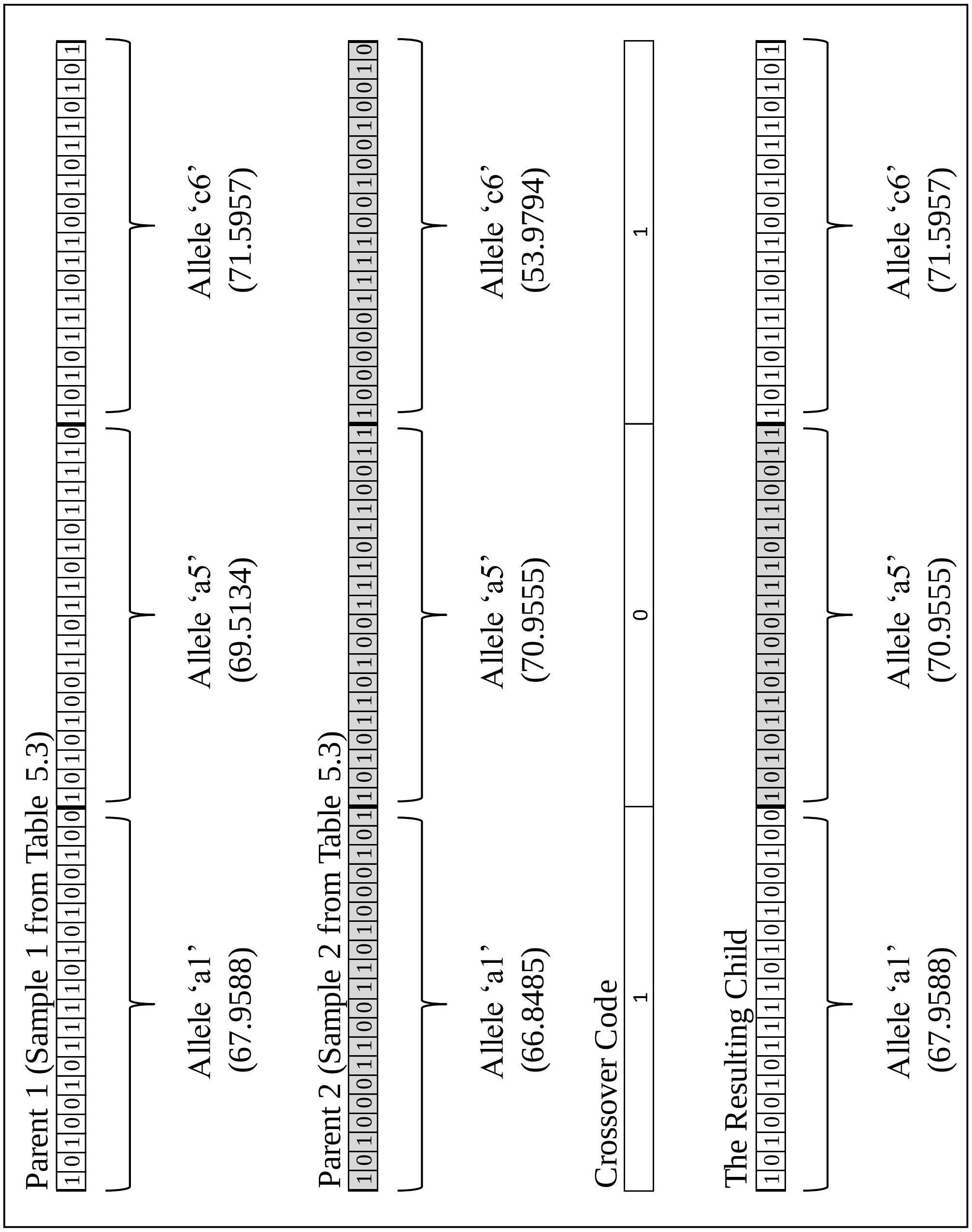


Figure 3 – Crossover of Sample 1 and Sample 2 (The First Three Alleles Only)

and the extent of the mutation. When a binary coding scheme is used, the mutation is generally the flipping of the value for a particular gene from 0 to 1 or 1 to 0 as the case may be. Typically a mutation code is generated for each child according to some probability distribution defined by the researcher. Figure 4 illustrates how a mutation code might be applied to Sample 1 from Table 3. All children and mutations must be evaluated against the constraints to ensure that the solution is feasible. Infeasible children or mutants are considered "stillborn" and are discarded. The next step in developing the GA is to consider the stopping criteria.

Stopping Criteria for GA

Stopping criteria in GA are much like those used by other optimization tools and are generally focused on small changes in the optimum value with subsequent iterations or amassed CPU time. As with the other tools it is necessary to consider that the method may 'stall' occasionally as it searches for better results. That is, there may be stretches of iterations with no change in the optimum value before the optimum value has been truly discovered. The researcher must take care to consider the number of stall iterations allowed to avoid considering the process to have converged on what may be only a local optimum. The next sections will discuss the application of GA to the coast-side vibration optimization problem using Matlab.

Implementation of GA Within Matlab

GA can be implemented easily using Matlab. Several intrinsic functions allowing conversions of decimals to binary or double-precision vectors and concatenation of multiple vectors make the implementation rather straightforward, albeit tedious. Newer versions of Matlab incorporate a Global Optimization Toolbox that has a built-in GA optimizer. The built-in GA optimizer in Matlab R2012a , OPTIMTOOL, is used here. The optimizer allows full control of the GA parameters. Here, the population is coded as a double-precision vector rather than a binary string to facilitate passing complete alleles during crossover. In Matlab, double-precision vectors are stored in memory as 64 bit binary strings. The solution of the problem will take place in three trials.

The first trial is comprised of the initial data from the assembly line as the initial population, thus the population size will be 21 throughout this optimization trial. The second and third trials are based on population sizes of N = 100 and N = 1,000 respectively. These trials are initialized with the 21 samples from Table 3 and filled out with feasible solutions generated at random within Matlab.

The GA Parameters and Strategies Specified for OPTIMTOOL

Table 5 is a summary of the GA parameters and strategies that are used to setup the GA within Matlab using OPTIMTOOL. Certain parameters are varied to focus the optimizer to the optimum solution. Before executing the optimizer, it is necessary to provide OPTIMTOOL with values for these parameters.

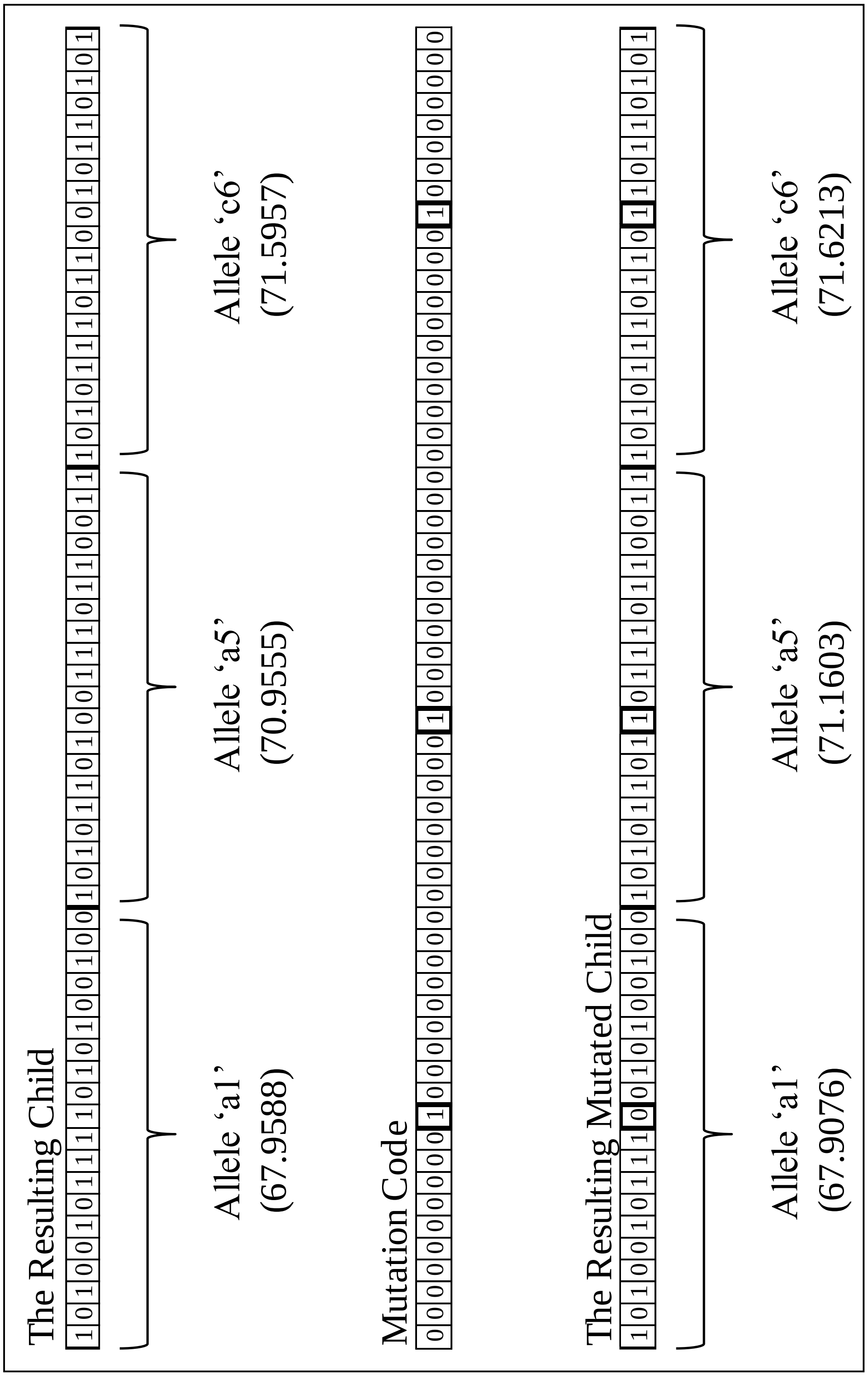


Figure 4 – Mutation of Sample 1 from Table 5.3 (The First Three Alleles Only)

Table 5 – GA Parameters and Strategies for OPTIMTOOL

Option / Parameter	Strategy / Value
Objective Function	Equation !(1)
Constraints	Upper and lower bounds defined by Table (1)
Population Type	Double Precision Decimal
Population Coding	Binary
Population Size	Group 1: N = 21 Group 2: N = 100 Group 3: N = 1000
Fitness Scaling	Rank
Selection Function	Roulette
Elite Proportion (ρN)	Group 1: ρN = 4 Group 2: ρN = 20 Group 3: ρN = 200
Crossover Fraction	0.80
Crossover Strategy	Scattered
Mutation Strategy	Adaptive Feasible
Stopping Criteria	
Maximum Number of Generations	1000
Time Limit	None
Fitness Limit	None
Maximum Number of Stall Generations	50
Maximum Stall Time	10 Seconds
Objective Function Tolerance	1E-09 dB

As discussed above, the Objective Function and Constraints are taken from Equation (1) and Table 1 respectively. The Population Type is set to Double Precision. Fitness Scaling allows the researcher use completely custom strategies to sort and select parents according to an individual fitness level. The standard within GA is to simply Rank-order and select individuals by fitness level. Therefore, the fitness scaling strategy used here is “Rank”. The Crossover fraction of 0.80 results in each subsequent population to be initialized by the elite population (ρN). 80% of the remaining (1 – ρN) individuals are born via Crossover, and the final 20% of the (1 – ρN) remaining individuals result from mutation of a random sample from the current population. Adaptive Feasible mutation causes all mutations to occur in the direction that produced the best results in previous generations while still remaining feasible. The Objective Function Tolerance is the minimum change required before declaring the routine “stalled”. In addition to the data in

Table 5, other data is required to run OPTIMTOOL such as the actual objective function, the initial sample data and the constraints all coded for Matlab. These data are read into Matlab using an input data program called GAINPUT.M that is shown in the Appendix.

Numerical Results of the Optimization Problem Using GA

The results of the optimization by GA are shown in Table 6 along with Figures 5, 6 and 7. These figures suggest that sample size from 21 to 100 did improve the likelihood that the optimum value would be identified, but that increasing from N = 100 to N = 1,000 did not improve the convergence significantly.

Summary of the Solution by a Genetic Algorithm

The best results from the optimization are compared to the deterministic solution in Table 7. The results are in excellent agreement with the deterministic solution. The agreement is precise to three decimal places. In addition the number of iterations and the required solver time is very good. It remains, for future work, to confirm that axles built to the optimum conditions indeed produce improved vibration performance. Other papers by the authors illustrate solving the same optimization problem using the Cross Entropy Method [5] and Particle Swarm Optimization [6].

Table 6 – Results of the Optimization of the Optimization Problem by a Genetic Algorithm

Run	**N**	**Generations Required to Converge**	**Solver Time (sec)**	a1	a5	c6	c7	d3	d6	Optimum Solution Y(**X***)
1	21	161	40	69.54243	70.9554802	71.5956417	46.02381	65.10277	71.36402	58.1422
2	21	145	37	69.54243	70.9554883	71.5928822	46.02653	65.10543	71.36399	58.1430
3	21	175	43	69.54243	70.9554803	64.0838908	46.02724	65.10544	71.36238	59.6263
4	21	174	44	69.54243	70.9554892	71.595665	46.02426	65.10189	71.36195	58.1432
5	21	125	33	69.54242	70.9554824	71.5956651	46.02341	65.10162	71.36048	58.1434
1	100	133	55	69.54243	70.9554897	71.5944108	46.02061	65.10545	71.36401	58.1405
2	100	112	47	69.54243	70.9554892	71.5940378	46.0206	65.10543	71.36173	58.1411
3	100	106	44	69.54242	70.955488	71.595653	46.02238	65.10545	71.3622	58.1414
4	100	111	45	69.54051	70.9554711	71.5956465	46.02063	65.10544	71.36403	58.1457
5	100	102	42	69.54242	70.9554758	71.5956367	46.02241	65.10307	71.36403	58.1417
1	1000	87	58	69.54243	70.9554867	71.5929264	46.0206	65.10544	71.3614	58.1415
2	1000	61	41	69.5424	70.9554684	71.5947246	46.02091	65.10513	71.36399	58.1408
3	1000	91	61	69.54243	70.9554885	71.595584	46.02076	65.10545	71.36403	58.1403
4	1000	86	57	69.54243	70.9554812	71.5940225	46.0206	65.10544	71.36268	58.1409
5	1000	105	70	69.54243	70.9554888	71.5956669	46.02216	65.10544	71.36403	58.1408

Table 7 – Summary of the Optimization Using a Genetic Algorithm

Optimization Method	N	Iterations to Solve	Solver Time (sec)	a1	a5	c6	c7	d3	d6	Optimum Solution Y(X*)
Deterministic	0	0	< 10	69.5424	70.9555	71.5957	46.0206	65.1055	71.3640	58.1402
Genetic Algorithm	1000	91	61	69.5424	70.9555	71.5956	46.0208	65.1054	71.3640	58.1403

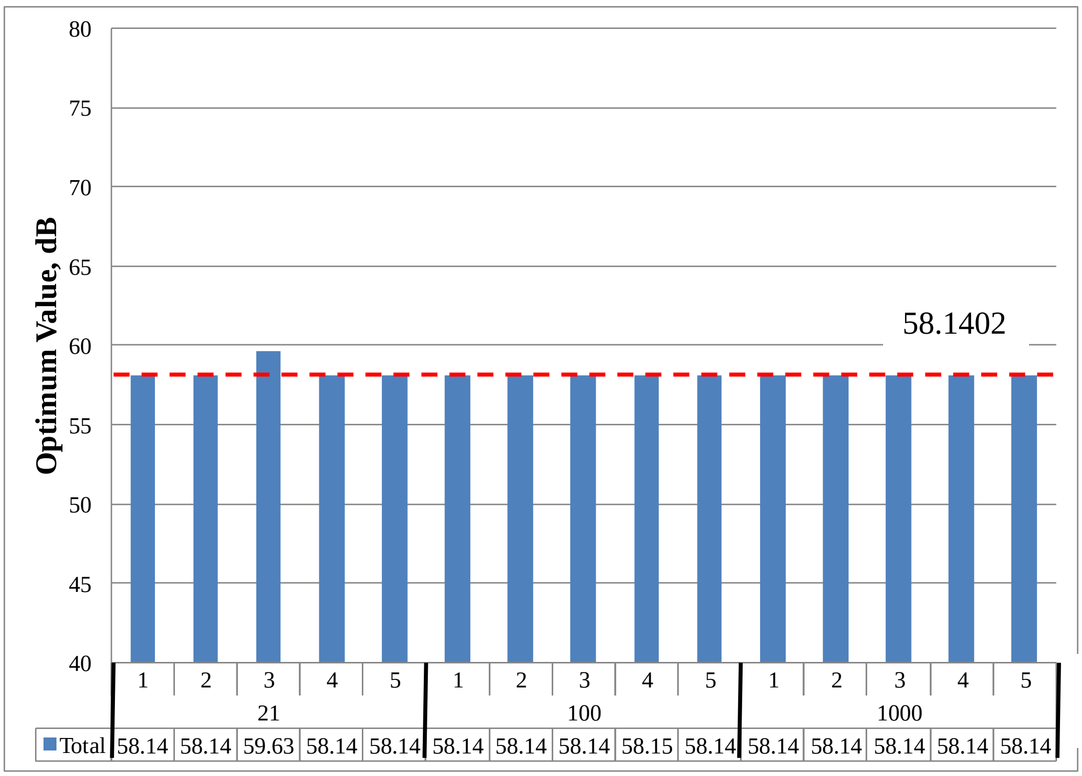


Figure 5 – Results of the Optimization by a Genetic Algorithm by Run Number and Population Size

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

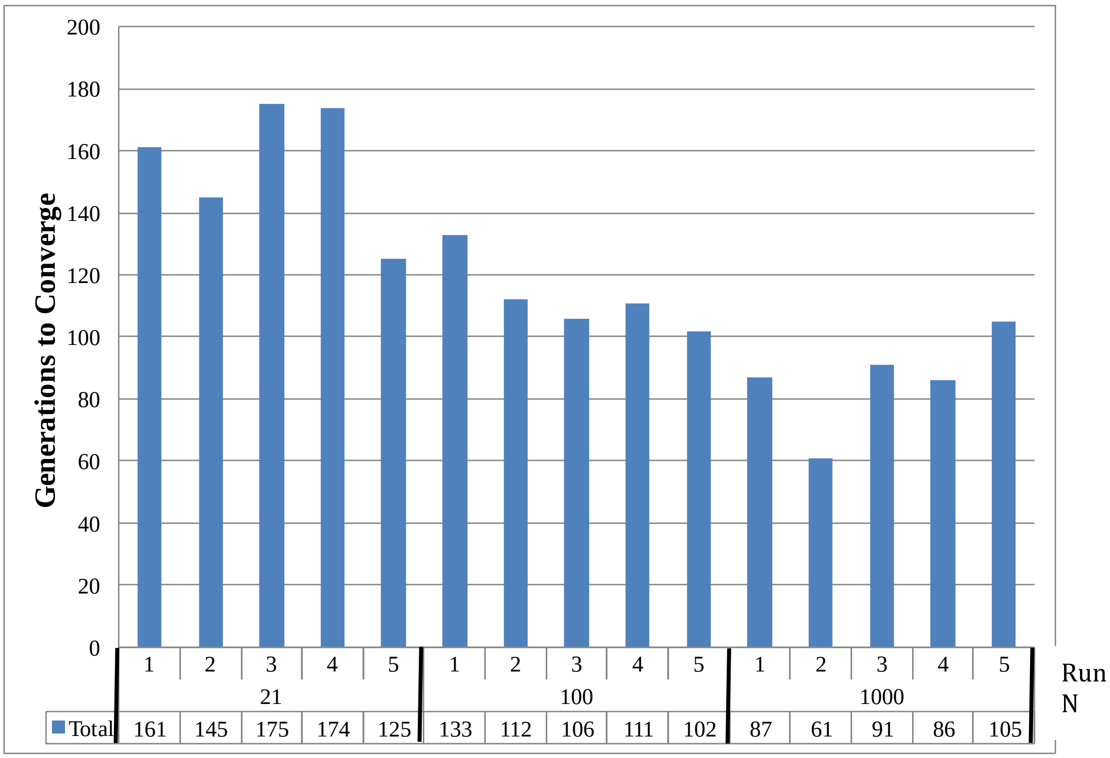
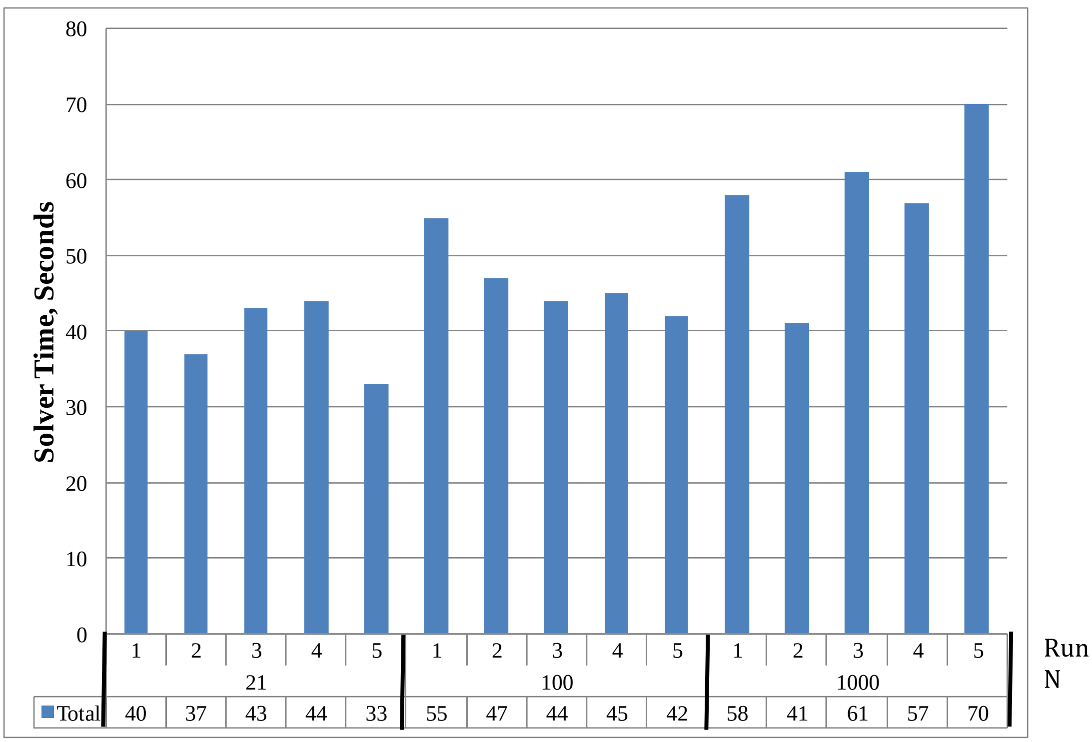


Figure 6 – Number of Generations Required to Converge by a Genetic Algorithm by Run Number and Population Size



	21					100					1000				
Run	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Total	40	37	43	44	33	55	47	44	45	42	58	41	61	57	70

Figure 7 – Solver Time Required to Converge by a Genetic Algorithm by Run Number and Population Size

Appendix

A Genetic Algorithm

1. Define the fitness function.
2. Define the boundary conditions for each parameter of the fitness function.
3. Define the initial population.
    a. A number of known solutions
    b. A number of random feasible solutions
    c. A combination of known and random solutions
4. Define the coding scheme for each solution.
    a. Binary
    b. Hexadecimal
    c. Other
5. Define the elite proportion to be retained by subsequent populations (ρN).
6. Define the rules for selecting parent solutions.
    a. Roulette – parents are chosen at random
    b. Tournament – the best solution from a subgroup
    c. Elitism – parents are chosen from an elite subgroup
    d. Other rules
7. Define the rules for crossover (birthing child solutions from parent solutions).
    a. The proportion of the population to be defined by crossover
    b. The type of crossover (single point, multi-point, scatter, etc.)
8. Define the rules for mutation (random changes to child solutions)

    a. The proportion of the population that will mutate
    b. The likelihood that mutation will occur
    c. The nature of the mutation
    d. Other rules

9. Define the stopping criteria.
    a. Minimal change in optimum fitness over a specified number of generations or
    b. A maximum number of generations or
    c. A maximum amount of CPU time
    d. Other criteria

10. Sort the current population from best to worst fitness.
11. Select the most-fit solution as the current optimum solution.
12. Select the elite proportion, ρ, of the current population and copy those to the next generation.
13. Select groups of parents from the current population.
14. Combine two parents according to the crossover rules to birth one child solution.
15. Apply the mutation rules to the child solution and mutate if required.
16. Calculate the fitness of the child solution.
17. Add the child solution to the next generation.
18. Repeat Steps 12 through 15 until the next population is filled.
19. Sort the next generation from best to worst fitness.
20. Select the most-fit solution as the new current optimum solution.
21. Evaluate the stopping criteria.
22. Repeat Steps 12 through 21 until the stopping criteria are satisfied.

Proceedings of the ICAM 2014, International Conference on Advanced and Agile Manufacturing. May 28-30, 2014, Held at Oakland University, Rochester, MI 48309, USA.

The Matlab Code for the Input Program

```
%   Revised 04/12/2013

warning('off','all');
warning;

%   Open the data files

O = load('nvhdata.mat', '-ASCII')
constraint = load('const.mat', '-ASCII')

%   Establish the number of variables (c) and the number of samples (r)

[r,c] = size(O);
[r1,c1] = size(constraint);

%   Create the initial sample values and scores

i = 1;
j = 1;
while i < r + 1;
    while j < c;
        P(i,j) = O(i,j);
        j = j + 1;
    end
    S(i) = O(i,c);
    j = 1;
    i = i + 1;
end

P
S

%   Create the boundary values

j = 1;
while j < c1 + 1;
UB(1,j) = constraint(1,j);
LB(1,j) = constraint(2,j);
j = j + 1;
end

UB
LB

% End of test.  Ask to clear the memory.

reply = input ('Do you want to clear everything? (y/n)[n]', 's');
if reply == 'y'
```



```
    clear
    clc
elseif isempty(reply)
    reply = 'n';
end
```

Bibliography

1. Sun, Z., *et. al.*, “NVH Robustness Design of Axle Systems,” *SAE Transactions*, v. 112, pp. 1746 – 1754, 2003.

2. Steyer, G., *et. al.*, “The Future of NVH Testing – An End-User’s Perspective,” *SAE Technical Paper 2005-01-2270*, 2005.

3. Meinhardt, G. and Sengupta, S., “Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part I-Preparing the Multivariate data for Regression Analysis,” *SAE Technical Paper 2012-01-727*, 2012.

4. Meinhardt, G. and Sengupta, S., “Correlation of Axle Build Parameters to End-of-Line NVH Test Performance Part II-Multivariate Regression Analysis,” *SAE Technical Paper 2012-01-728*, 2012.

5. Meinhardt G. and Sengupta, S., “Optimization of Axle NVH Performance Using Particle Swarm Optimization,” *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

6. Meinhardt G. and Sengupta, S., “Optimization of Axle NVH Performance Using the Cross Entropy Method,” *Proceedings of the ICAM 2014 May 28-30, 2014*, 2014.

7. Holland, J., “Adaptation of Natural and Artificial Systems,” University of Michigan Press, Ann Arbor, 1975.

Additive Manufacturing of Directionally Heat Conductive Objects

Pravin M. Kulkarni[1], Dhirendra Rana[1], K.P.Karunakaran[1], Asim Tewari[1], Prathmesh Joshi[1]

Rapid Manufacturing Laboratory, Department of Mechanical Engineering,

Indian Institute of Technology Bombay, Powai, Mumbai-400076, INDIA

Email: kulkarnipravin@iitb.ac.in

Abstract:

Purpose: Directionally Heat Conductive (DHC) objects are also known as thermal cables. DHC objects used in cooling solid state devices as large heat sinks. This paper presents Additive manufacturing (AM) as a new method for realization of DHC objects. AM can produce DHC objects accurately, rapidly and economically.

Design/Methodology/Approach: It is very hard to realize DHC objects through conventional manufacturing. The authors describe the realization of metallic DHC objects using their AM process called Hybrid Layered Manufacturing (HLM). HLM, which uses GMAW deposition, can, in principle, produce DHC out of any metal for which welding wire is available. . Differential conductivity in HLM is achieved by appropriately distributing the air gaps among the beads. The air gap distribution depends on the layer thickness and step over increment.

Findings: The authors have demonstrated the ability of HLM to produce DHC objects for Al alloy 4043. The thermal conductivity achieved on two orthogonal directions was 100.496 and 129.740 W m-1 K-1 as against 163 W m-1 K-1 for solid metal. These are significant variations.

Research Limitations/Implications: HLM realize near net shape of DHC objects, further post processing namely machining is required.

Originality Value: AM Specifically HLM for realization of DHC objects

Keywords: Directional heat conductive (DHC), Additive Manufacturing (AM), Hybrid Layered Manufacturing (HLM).

Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA, Copyright © 2014, ISPE USA and ISAM USA

1. Introduction

Thermal conductivity of a material has great influence on its performance in heat transfer analysis, a material is normally assumed to be isotropic. Therefore, it is a common practice to evaluate the thermal conductivity at the average temperature and treat it as a constant in calculations [1]. This assumption is realistic for most materials, but some materials, such as-laminated composites, are anisotropic. This anisotropy can be engineered to have directional conductivity. Such objects are known as *Directionally Heat Conductive (DHC)* objects. There are several applications of DHC in different fields.

The characteristics of a DHC object depend on the material composition and porosity or void along different directions. DHC objects find major applications as heat sinks in electronic packaging, especially for high voltage chips, packaging base plates, carrier plates of electronic box closures [2, 3]. In these applications, heat dissipation is important because modern integrated circuits and other powerful electronic parts generate a lot of heat.

H. Amiri et al. have reported the numerical solution for the steady state combined conductive-radiative heat transfer in an anisotropic participating medium within the irregular geometries [4]. Some of the engineering applications of such combined properties were mentioned in areas such as thermal transport in nano-materials, heat transfer through the semi-transparent, porous materials, multi-layered insulations, glass fabrication etc. [5]. Similarly DHC objects can be used as effective heat conductors in large, high speed electrical machines as housing jackets to extract the heat generated inside the stator. It can be also used for insulation in industrial furnaces to retain the heat [6].Aerospace engines uses honeycomb structures which provides directional thermal conductivity as well as weight reduction.

Temperature rise in high speed electrical machine causes mechanical failures and irreversible demagnetization of permanent magnets. Initially temperature in high speed electrical machines was maintained by using forced air-cooling with external cooling fins often by a ventilator mounted on the shaft. At higher speeds, this may be very noisy and take up a fair deal of torque, so an independent cooling is required. This independent cooling is done by DHC housings.

Figure 1 shows the current methods of manufacturing DHC objects like bonding, sheet metal forming (folding and co-corrugation process), open and closed die forging, liquid_forging, skiving, Powder Metallurgy (PM), Metal Matrix Composite (MMC) manufacture [7]. DHC objects manufactured by various process is shown in figure 1.

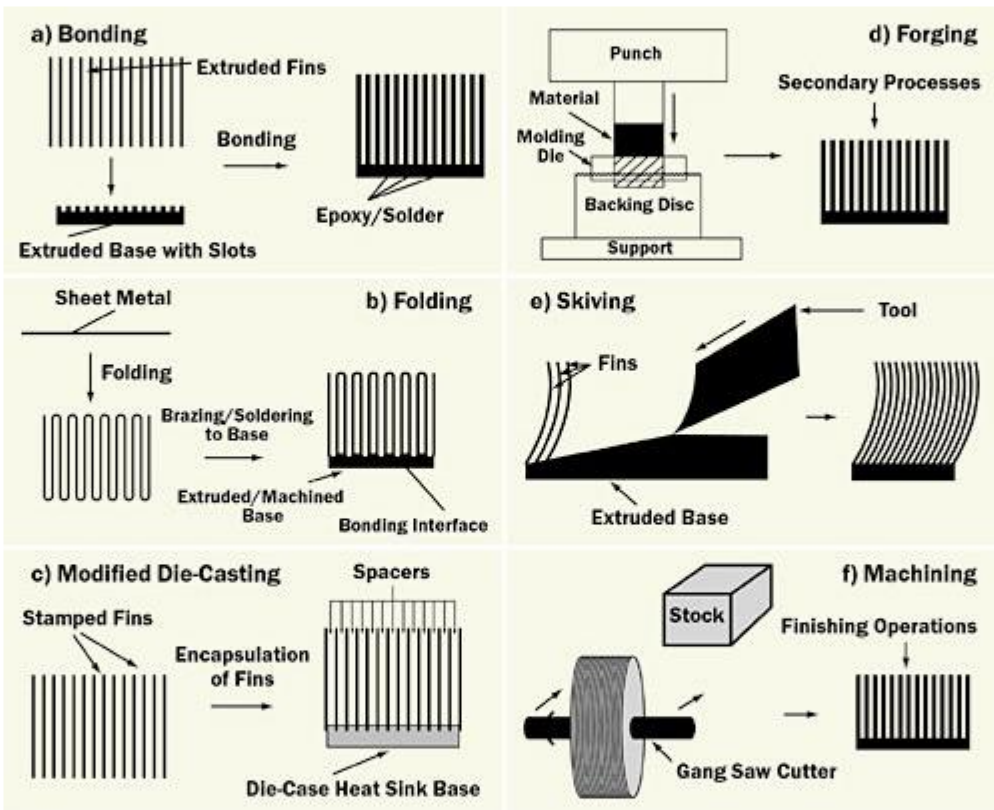


Figure 1 Various Manufacturing Methods of DHC Objects[7]

Honeycomb structure is manufactured by sheet metal forming method named as co-corrugation process. The unidirectional thermally conductive fibres are oriented in a direction parallel to the width of the formed sheet. There are various configurations of honeycomb structures, out of which hexagonal cells gives better directional thermal conductivity. Titanium honeycomb structure used for exhaust nozzles in aircraft is shown in figure 3(a) and this type of structure is manufactured by corrugation process shown in figure 3(b). This honeycomb structure finds major applications in military and commercial aircraft, automotive bodies, engine components, recreational equipment, and marine craft [8, 9].

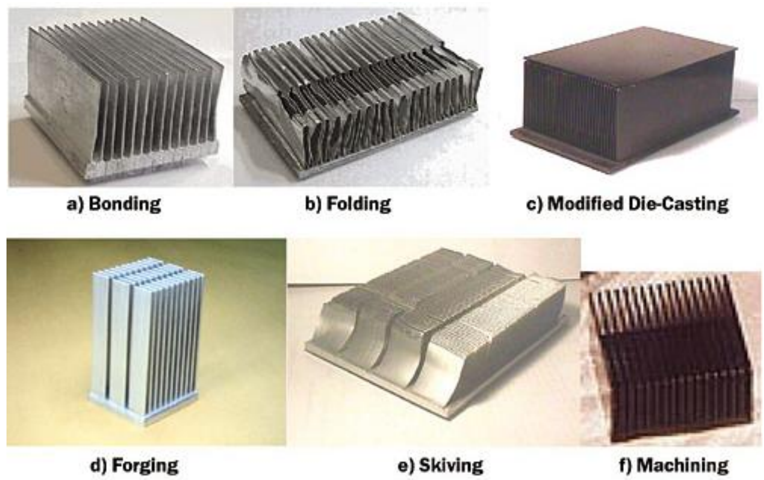
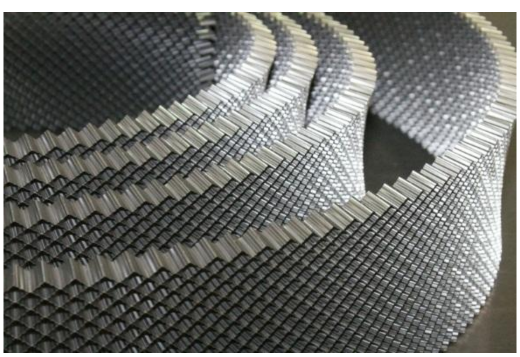
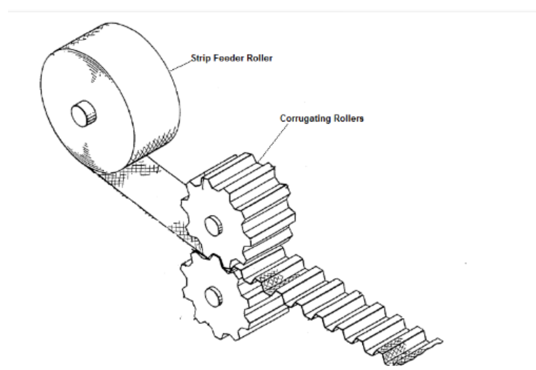


Figure 2 DHC objects Manufactured by various Process [7]



(a) Metallic Honeycomb



(b) co-corrugation process

Figure 3 Manufacturing of DHC Objects by Sheet Metal Forming Process[8]

MMC used to manufacture DHC vibration isolator imparting shock and vibration isolation. MMC can be manufactured by powder metallurgy and foil diffusion, both of them are solid based methods. Electroforming, stir casting, squeeze casting, spray deposition and reactive processing are the liquid based MMC manufacturing processes [10].

Heat sinks are manufactured by forging instead of casting, machining and extrusion. Forging is the most effective method to form complex shapes and also offers unique thermal advantages. The part is formed under high pressure which controls the grain structure and results in improved thermal performance. The cold forging process results in a heat sink that yields a 14% improvement over the extruded and a 62% improvement over the die cast heat sink [11].

The above methods to produce DHC objects are hard and expensive as they are dominated by subtractive manufacturing [12]. With the advent of Additive Manufacturing (AM), manufacturing of DHC objects has become elegant and affordable. AM is inherently anisotropy. When this inherent anisotropy is engineered to meet the functional needs of a component, it is known as Functionally Gradient Materials (FGM). DHC is a special case of FGM.

Additive Manufacturing (AM) is the term given to a group of technologies that are capable of creating physical objects from computer aided design (CAD) files by incrementally adding material such that the objects “grow” from nothing to completion. Most of today’s AM technologies start with a CAD solid model which is sliced into thin layers. Each layer comprises a 2D cross section profile of the part which is then made layer by layer.

AM involves a number of steps that move from the virtual CAD description to the physical resultant part. Different products will involve AM in different ways and to different degrees. Small, relatively simple products may only make use of AM for visualization models, while larger, more complex products with greater engineering content may involve AM during numerous stages and iterations throughout the development process. Furthermore, early stages of the product development process may only require rough parts, with AM being used because of the speed at which they can be fabricated. At later stages of the process, parts may require careful cleaning and post-processing (including sanding, surface preparation and painting) before they are used, with AM being useful here because of the complexity of form that can be created without having to consider tooling [7].

The automated approach to creating objects from CAD files in the additive manner that AM allows is beginning to have a profound effect on how objects are made. Almost all manufactured objects that we see today are made with some form of tooling at some point – parts may be machined using a cutting tool or parts may be formed from a molding tool. The use of tooling imposes numerous restrictions such as design limitations, restrictive costs to prepare for manufacture and the need to create products at a single location prior to being shipped globally. AM technologies invariably require no tooling to create parts and this profoundly affects how products can be designed, tailored, made, distributed etc. [12].

Organization of Paper:

In this paper authors give the introduction of HLM process, involvement of HLM technology in realization of DHC objects, characterization of DHC objects by using the heat conductivity testing. Finally authors give the conclusion in which they show unidirectional behavior of DHC objects.

2. Hybrid Layer Manufacturing Process

Hybrid Layered Manufacturing (HLM) is a Rapid Manufacturing process which includes layer by layer deposition using *Gas Metal Arc Welding (GMAW)*. This deposition makes near-net shape of the object to be built which is then finish machined subsequently. Face-milling is done after depositing each layer to remove oxidized surface and scallops. It also ensures z-accuracy of each layer.

Arc weld deposition is faster and more economical by an order of magnitude than laser and electron beam processes (50-130g/min as against 2-10g/min). The surface quality is in the decreasing order for laser, electron beam and arc. Even the best surface obtained through AM is not adequate enough for direct use for very serious applications except in some medical applications where rough surface is favourable for tissue growth. If finish machining is necessary, then these marginal quality differences do not matter.

Subtractive processes can produce good quality parts but are slow; although the material removal by itself is fast, human efforts required for cutter path generation is the bottleneck. On the other hand, additive processes are fast but produce poor quality parts. The additive process focuses on speed while ensuring the desired material integrity. The resulting object is only near-net as no attention to the geometric quality is paid at the time of building it in layers. In HLM, the near-net shape is obtained with total automation. However, the subsequent two stages of stress relieving/ heat treatment and finish machining have fair amount of human intervention. In other words, HLM does not strive to achieve total automation but aims at optimal/ economical level of automation.

(a) 5 Axis CNC machine

(b) Close view of 5 Axis CNC HLM unit

(c) 3-Axis CNC machine

(d) Close view of 3-Axis HLM unit

Figure 5 Integration of HLM with conventional CNC machines

2.1 Methodology of HLM

The stages involved in building the metallic object using HLM is shown in figure 6.

a. Geometric processing
b. Building the near-net shape
c. Stress relieving
d. Finish machining

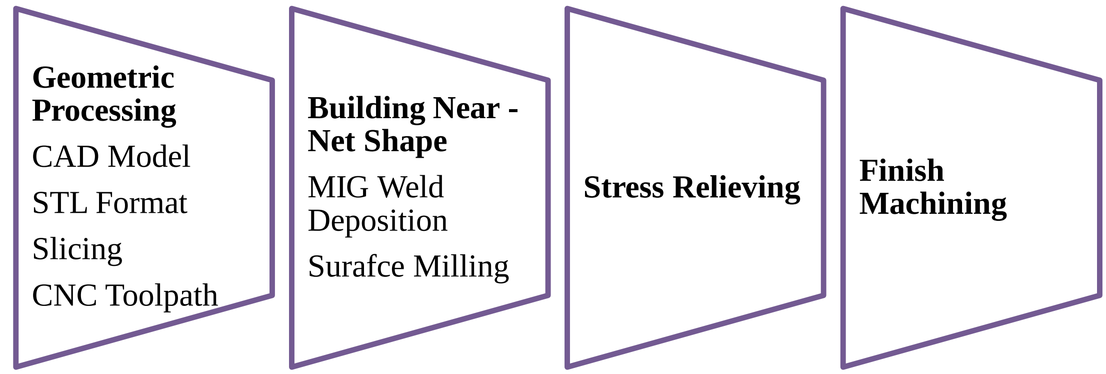
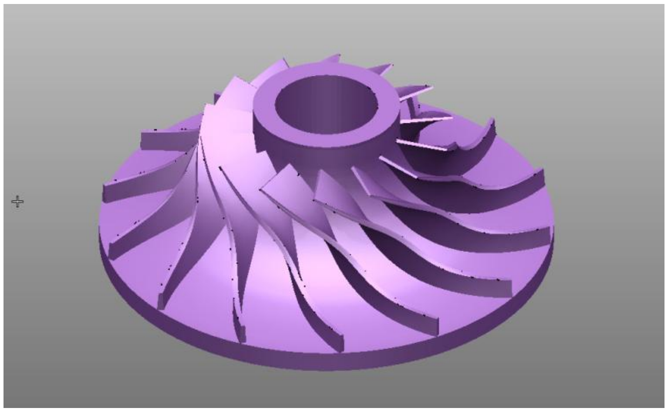


Figure 6 Process Flowchart

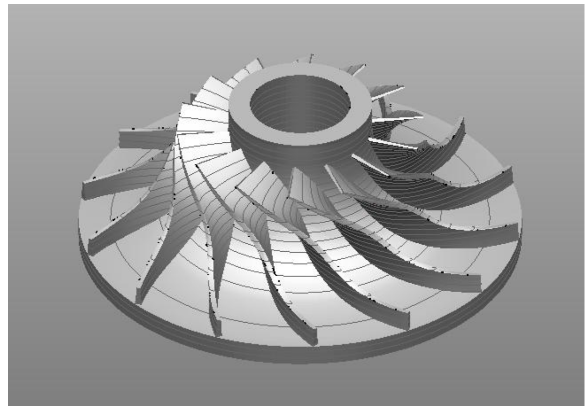
A case study of impeller which was manufactured using HLM process is shown below in figure 7.

2.2 Application

The various geometric capabilities of HLM, in its three kinematic modes of 3-axis, positional 5-axis and continuous 5-axis, are shown in figure 8. In 3 axis kinematics mode, we can manufacture parts by means of various approaches such as planar deposition, planar deposition with small undercuts, planar deposition after building the undercut, non-planar (conformal) deposition. HLM can be used to manufacture *Functionally Gradient Materials (FGM)* by using multiple welding torches at with different filler materials.



(a) CAD Model



(b) CAD Model after slicing

(c) Near net shape

(d) Finished part

Figure 7 Case Study Shows HLM Process

2.3 Advantages of HLM

i. Seamless integration of material addition and subtraction on the same platform
ii. Totally automatic manufacture of the near-net shape
iii. Use of economical, faster and safer GMAW as against laser and electron beam
iv. Retrofittable to any existing CNC machine, independent of make and age
v. Easy toggling between HLM mode and regular CNC machining mode

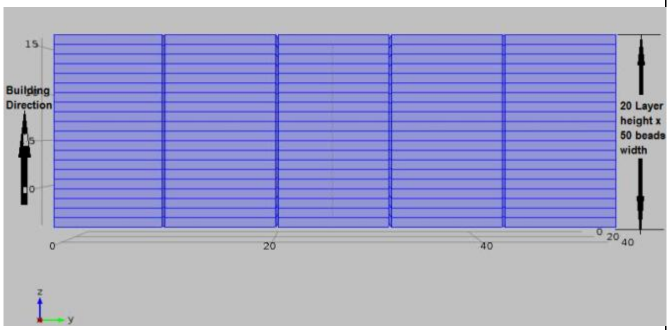
Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA, Copyright © 2014, ISPE USA and ISAM USA

Figure 8 Capabilities of HLM in Different Kinematic Modes [15]

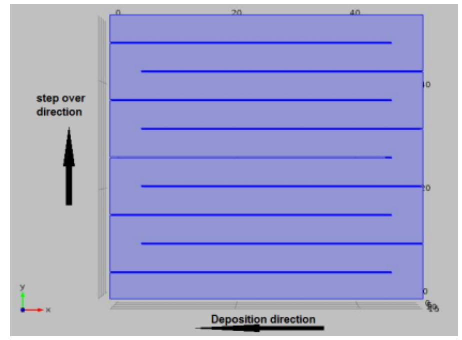
3. Realization of DHC Using HLM

A Cold Metal Transfer (CMT) 2700 welding machine and a 3-axis CNC machine were used for deposition and subtraction respectively. Three axis Cartesian coordinate system was adopted to identify the direction of deposition and directionally heat conductivity analysis. The X-axis (torch direction) is taken in the long travel of the weld torch, Z-axis is taken perpendicular to the surface of the substrate (weld direction) and Y-axis is taken in the stepover increments direction. Figure 9 illustrates the coordinate system used in deposition and characterization. The heat conduction was measured in the directions parallel and perpendicular to the air gaps in X and Y directions respectively.

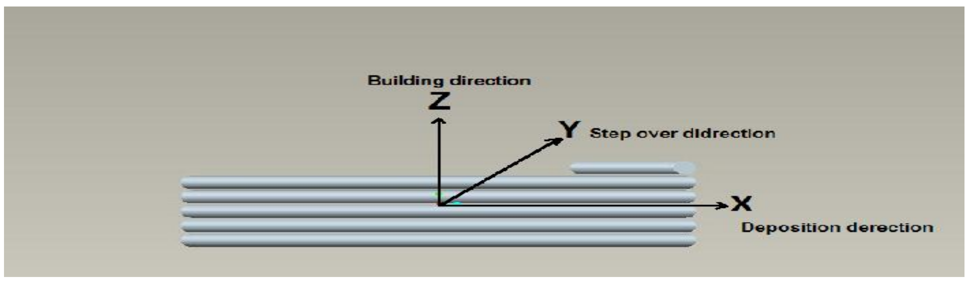
The part was built by adopting zig-zag deposition. The deposition in the longitudinal direction was built on X-axis and the pitch was provided in a perpendicular direction. The pitch size was given by considering the air gaps between consecutive beads. Complete fusion between weld beads was prevented by providing a step over which is equal to half of the bead width plus the required air gap. The CAD model of the sample, with air gaps between each strip of layers, is shown below in figure 10.



(a) Front View of Specimen



(b) Top View of specimen



(c) Coordinate system for deposition and characterization

Figure 9 Design of DHC Sample

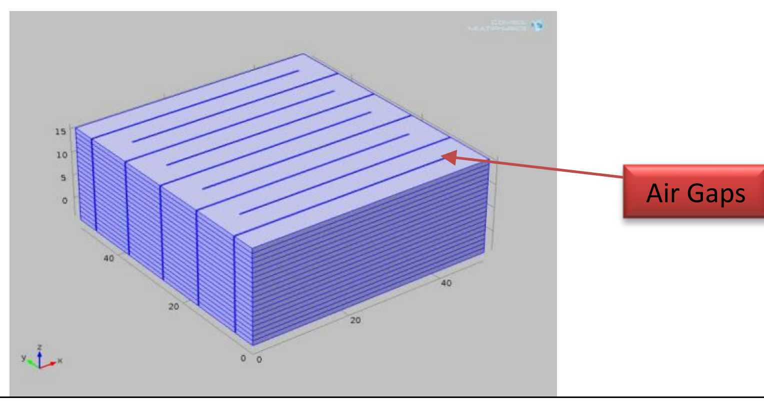


Figure 10 CAD model of a DHC sample

As shown in Table 1, during deposition different process parameters were taken into account. Wire speed which affects the bead width and thickness, and which in turn is affected by the rate of heat input, was controlled to minimize distortion of beads due to overheating during layer deposition. The torch speed which is another important welding parameter was controlled numerically to get a high precision in the weld bead size and to minimize heat distribution over a wider area [13]. If the torch speed is high, the bead width is thin and vice-versa. Therefore the lowest wire speed and the highest torch speed within their stable operating ranges are chosen to preserve the required air gap. The other and the most important parameter which was critical to our specific experiment is the stepover increment. The size of the stepover determines the size of the air gaps. If the stepover is less than half of the width of the weld bead, complete fusion between adjacent weld beads will occur. The size of stepover was determined based on the required size of air gaps.

Variation of current with layer thickness is shown in figure 11. It shows that as current increases layer thickness also increases.

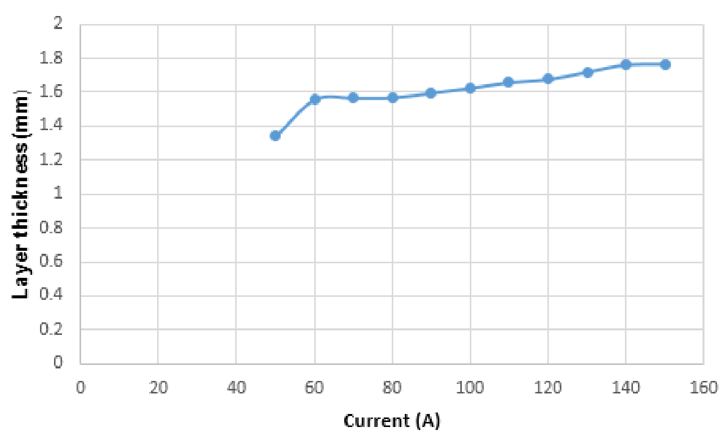


Figure 11 Variation of Layer thickness with current

The process parameters used during weld deposition are shown in table 1. The parameters are given based on the requirements of the final object.

Table 1 Welding Parameters for Al-Si

Layer thickness (mm)	1
Torch speed(mm/min)	850
Wire feed rate(m/min)	3.4
Welding wire diameter (mm)	1.2
Gas flow rate(l/min)	10-12
Shielding gas	Ar

Pitch	Bead width+ 0.25mm
Air gaps	0.25mm

4. Characterization of DHC

Characterization of DHC includes the thermocouple calibration which is used to plot a graph between temperature and voltage using which an equation is established. Furthermore a heat conduction test is performed to show the unidirectional heat flow.

4.1 Thermocouple Calibration

The necessity of the calibration is to establish an equation of curve that fits to the temperature versus voltage graph. This equation is applied to find out the unidirectional temperature drop or rise in the sample along the specified directions. Hence a calibration setup was established to calibrate a thermocouple based on a manual by ASTM [14] using fixed reference temperature technique. If properly used, it is a convenient and inexpensive way to achieve the ice point, it can be reproduced with ease and with an exceptional accuracy of 0.0001°C [14]. Ice bath reference junction setup was established for this purpose as shown in figure 12. The ice bath is made up of a mixture of melting shaved ice and water. This is one of the most common reference junction techniques used in calibration. To avoid the errors due to melting of the ice at the bottom of the bath, consumed excess water was removed periodically and more ice was added, so that the ice level is maintained safely below the reference junctions.

The water level in the bath was checked and the apparatus was connected to a power supply. The temperature in the bath was initially set to 30°C to start heating. In normal mode, the temperature shown on the digital display is the actual temperature of the bath (T_B) against which the thermocouple sensor is to be calibrated. The ends of the thermocouple are connected to the digital multimeter and which was set to read voltage in mV. The other end of the thermocouple was dipped in ice bath. A stable temperature source was kept to maintain a constant temperature. The multimeter reading in millivolts was noted down for each 5°C increment in temperature from 30°C to 80°C. The most commonly used reference temperature is 0°C [13].

The multimeter reading in millivolt corresponding to the difference in temperature between the surroundings (ice) and the bath was recorded. The measured bath temperatures values (T_B) were plotted on Y-axis against the corresponding reading of emf (mV) values on X-axis, as shown in Fig.13. The slope and intercept of the curve-fitted were obtained from the equation correlated to the graph. The linear correlation coefficient (R), which measures the strength and the direction of a linear relationship between two variables, was calculated to observe the strength of association between the millivolt reading and the temperature during calibration.

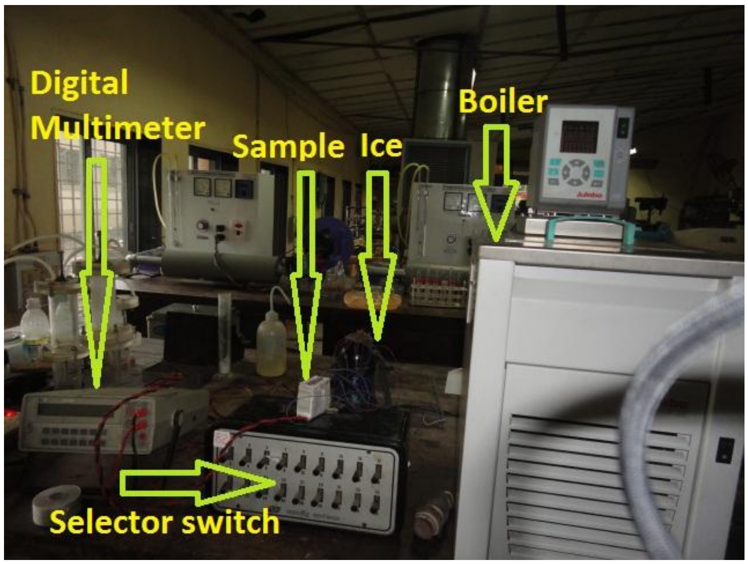


Figure 12 Thermocouple calibration setup

$$R= \frac{n(\sum T*V)-(\sum T)*(\sum V)}{\sqrt{n(\sum T^2)-(\sum T)^2}\sqrt{n(\sum V^2)-(\sum V)^2}}$$

Where

n is number of pairs of data

T is temperature in ^{0}C

V is voltage in mV

The thermocouple voltage at various known temperatures was measured. This will develop a correlation between thermocouple voltage and thermocouple temperature. This correlation is represented by the graph in figure 13. As it can be seen from the graph that the variation in temperature corresponds to the variation in voltage reading.

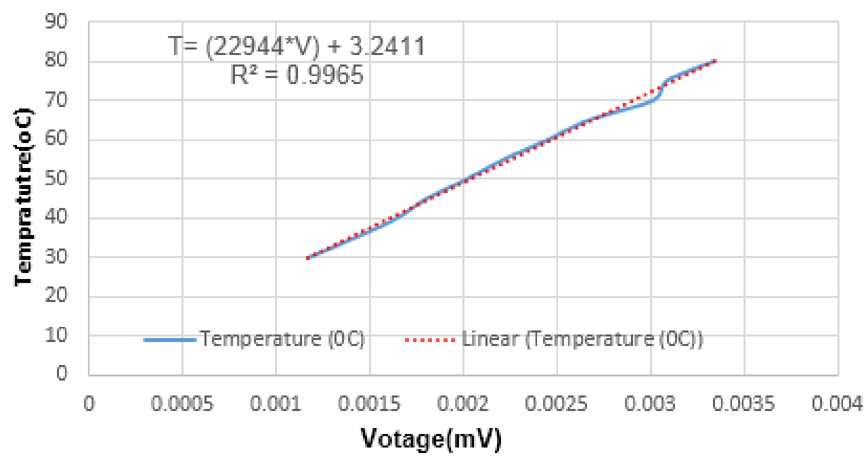


Figure 13 Correlation between temperature and voltage readings

The end result of regression analysis predicts the value of an output variable (temperature) based upon the values of input variables (mV reading). So based on the result of the regression which is 99.6% the total variation in temperature can be explained by the linear relationship between mV readings and temperature readings. It can be clearly observed from the graph in figure 13 that the regression line passes almost exactly through every point on the scatter plot. So it can be concluded that this can be able to explain all the variations.

Table 3 shows the final result of correlation coefficient which is 0.998232 and this value indicates a strong positive linear relationship via a firm linear rule.

Table 3 Correlation coefficient of the curve fit

	Voltage (mV)	Temperature(°C)
Voltage (mV)	1	
Temperature(°C)	0.998232	1

4.2 Heat Conduction Test

Aluminum alloy 4043 which has 5% Si is used in this deposition. This alloy is used to weld many similar and dissimilar cast and wrought Al parts. It is extensively used with the GTAW and GMAW processes with Argon as shielding gas. The original thermal conductivity of the welding wire is 163 $Wm^{-1} K^{-1}$.

Six thermocouples are used to record the temperature difference on the heat input and output side of opposite faces of the sample as shown in figure 14. Three small holes, for three thermocouples on each side, with 3mm depth are drilled into four sides of the sample. The three fixed thermocouples T_1, T_2 and T_3 are positioned along the heater section with a gap of 16mm in between them. The other thermocouples T_4, T_5 and T_6 are attached to the output side at similar intervals. The thermocouples on the input and output sides were connected to the selector switch for sequential selection and recording of the voltage readings.

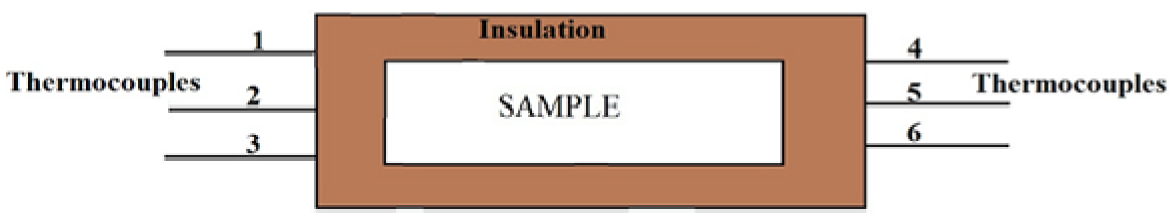
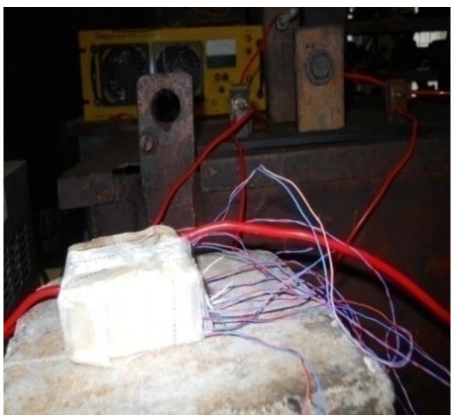


Figure 14 Schematic diagram of experimental setup

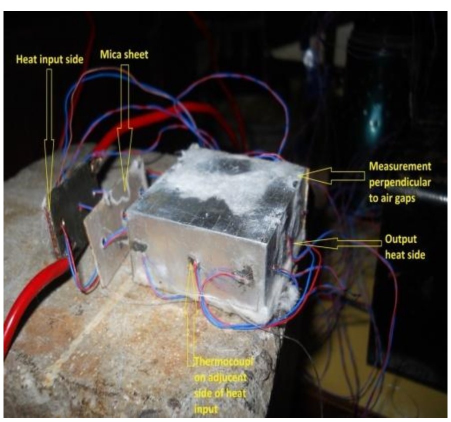
The apparatus consist of heating elements fitted on one side of the sample whose thermal conductivity has to be measured. The heater plate assembly is made up of mica sheets on which the Nichrome sheet is sandwiched in between upper and lower mica sheets for stabilizing and avoiding damage of the thermocouples. The sample was insulated on all of its faces for preventing the heat loss to the environment. When power is supplied to the heater, heat transfer will take place from the heat input side to the output side. The temperature of the sample is measured at the six positions by using a K type thermocouples and digital voltmeter. The power supply is adjusted to desired quantity by means of electronic controlled circuit by adjusting the voltage and current so as to provide the required watts of input to the heater plate. Electric power input can be measured by using digital voltmeter and ammeter multiplier [12]. The steady state temperature was ensured by the change in the voltage reading for each thermocouple which is fitted at the surface of the specimen. Steady state is reached when these voltmeter readings stop changing with time.

The rectangular sample was surrounded by a glass wool with a very high insulating capability as shown in figure 15(a). A glass tape was used to tightly hold the sample with the heating assembly and the insulating material. Temperature differences are measured by the calibrated thermocouples which are attached at the heat input and output sides of the specimen.

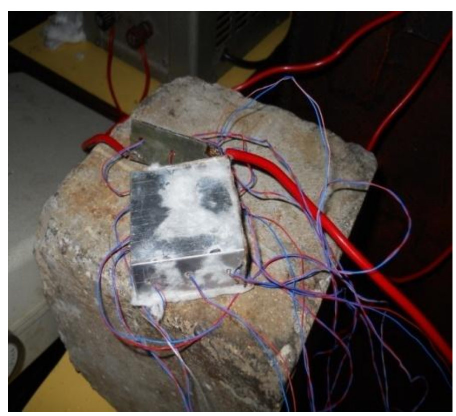


(a) Insulation of sample

(b) Power supply and selector switch



(c) Heating plate assembly



(d) Position of thermocouples

Figure 15 Sample and heat assembly for conductivity testing

Variation of input and output temperature along air gaps with respect to time is shown in figure 16, it shows that almost constant temperature difference maintained between heat input and output side.

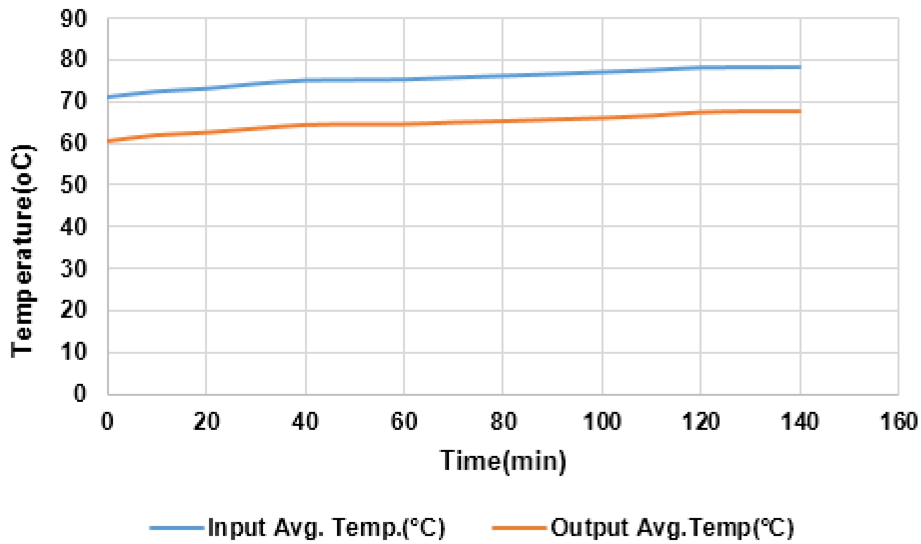


Figure 16 Variation of temperature along air gaps with time

Variation of input and output temperature across air gaps with respect to time is shown in figure 17, it shows that almost constant temperature difference maintained between heat input and output side.

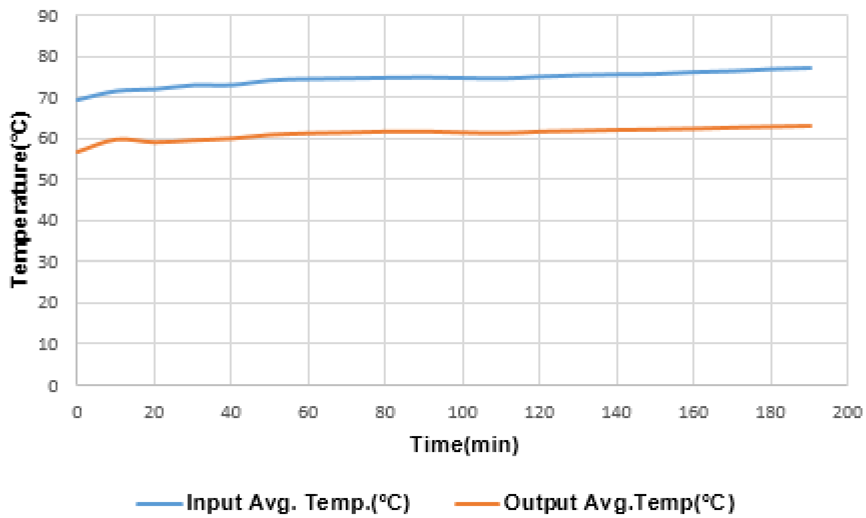


Figure 17 Variation of temperature across air gaps with time

As explained in the previous sections the main objective of this experiment is to prove the capability of *HLM* in producing directionally conductive material and to measure the directional conductivity of the sample in two different directions. It was assumed that the small variation in the air gap width due to spattering of molten metal to adjacent weld beads is negligible. For the measurement of thermal conductivity (K) one dimensional heat flow through the rectangular specimen was considered. Fourier's law of heat conduction was considered throughout the heat conduction test.

4.2.2 Experimental Results

The experimental results for the DHC specimen showed anisotropic heat conduction measured through the opposite sides. A heat input of 30.4 watt was provided from one side of the specimen and the temperature difference across the specimen was measured. The input and output measured temperatures were averaged after steady state condition was reached. The variation of average temperature with time along and across air gaps shows a similar trend.. Variation of temperature along and across air gaps with time is shown in figure 18.It shows output average temperature along air gap is more than that of temperature across air gap, which implies thermal conductivity across gap is less than that of across air gap.

Where:
Q= heat transfer rate = V.I =30.4 Watt,

A= Area normal to heat transfer, m^2

T_h = temperature of hot plate i.e. $(T_1 + T_2 + T_3) /3$ in °C

T_c = temperature of cold plate i.e. $(T_4+T_5 + T_6) /3$ in °C

L = thickness of specimen along heat flow in m.

To check the uni-directional heat flow, the temperature was measured on the adjacent side of the heat input side and a negligible heat flow was observed.

The measured value of thermal conductivity was 0.1398$Wm^{-1}\,{}^{0}C^{-1}$, which is negligible when compared to heat flow on the heat input side. The variation of temperature along adjacent side of heat input is shown in figure 19, it shows there is negligible temperature difference between heat input and adjacent side.

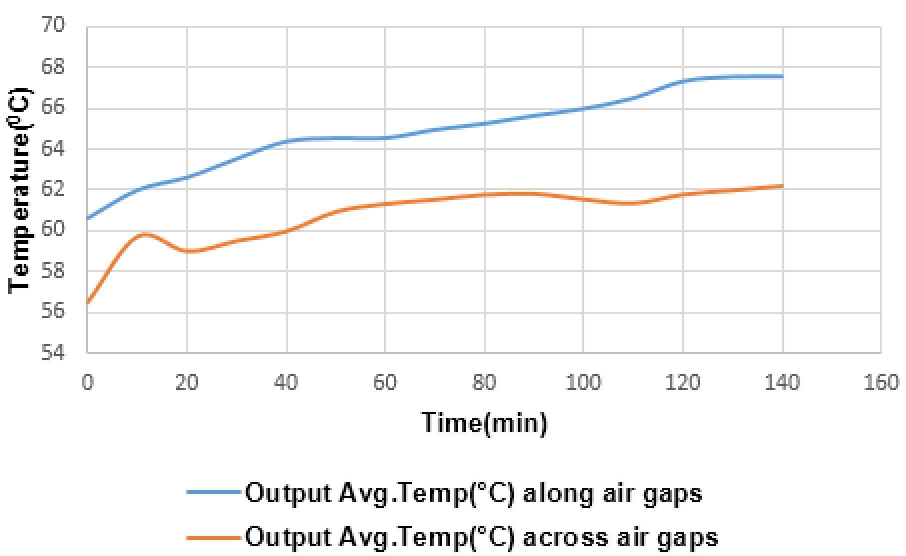


Figure 18 Variation of temperature along and across air gaps with time

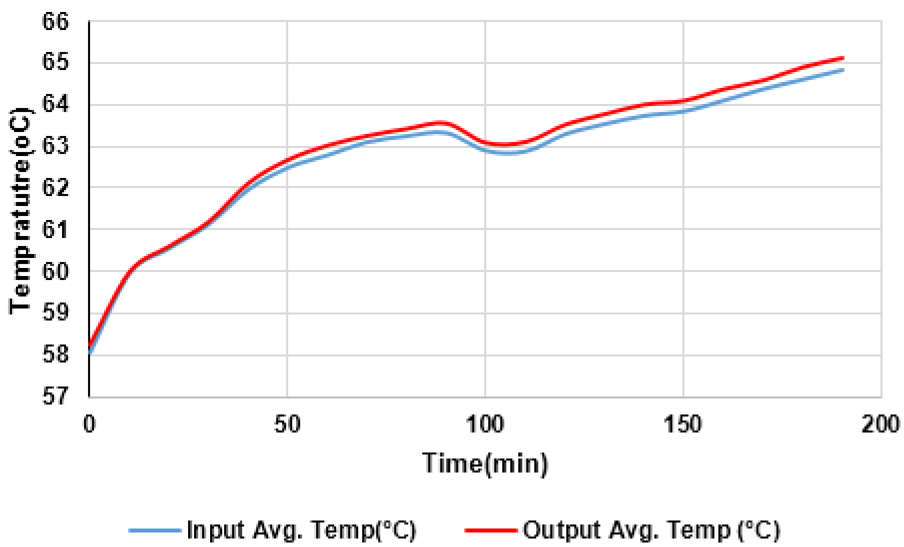


Figure 19 Variation of temperature along adjacent side of heat input

5. Conclusion

Unlike other heat sink, which radiates from all surfaces and in all directions, the new object developed through *HLM* radiate elaborately in one direction. The Al-Si alloy transfer heat away from the source along the welding direction but only minimally in a direction transverse to the welding direction. The object has demonstrated heat transfer of 130 W/m-k measured parallel to air gaps and 100 W/m-k perpendicular to air gaps. In general, the value of heat conductivity parallel and perpendicular to air gaps on the specimen was showed anisotropic thermal conductivity. The thermal conductivity was decreased by 37% in direction perpendicular to the air gaps and 20% in direction parallel to the air gaps.

References

1. Cengel, Y. A. (2003) "Heat transfer: a practical approach", 2nd edition, McGraw Hill, New York.
2. Lefrane, G., Degisher, H.P., Sommer, K.H., Mitic G. (1999), "Al-SiC improves reliability of IGBT power modules", *Proceedings of ICCM conferance*-2012, Paris, Paper ID 1335.
3. Simancik, F. (2003) “Factors influencing the development of advanced metallic material in advanced metallic materials”, proceedings of the international conference, Smolenice, 5-7 November, 2003. Bratislava.
4. Amiri, H., Mansouri, S.H., Safavinejad, A. (2010), “Combined conductive and radiative heat transfer in an anisotropic scattering participating medium with irregular geometries”, *International Journal of Thermal Sciences*, Volume 49, Issue 3, pp. 492–503.
5. Gillet, J., Djafari-Rouhani, B., Pennec, Y. (2009), “Directional Thermal Conductivity of a Thin Si Suspended Membrane with Stretched Ge Quantum Dots”, *EDA Publishing/THERMINIC* 2009, ISBN: 978-2-35500-010-2
6. Saban, D., Cloud, A., Guedes-Pinto, P., (2009), “Highly Conductive Insulation for Large, High-Speed Machines”, Presented at the *Coil Winding / Insulation & Electrical Manufacturing Exhibition*, May 5-7, 2009, Messe Berlin, Germany.
7. Iyengar, M. (2007), "Design for Manufacturability of Forced Convection Air Cooled Fully Ducted Heat Sinks", available at *http://www.electronics-cooling.com/2007/08/design-for-manufacturability-of-forced-convection-air-cooled-fully-ducted-heat-sinks/* (accessed 3 November 2013).
8. Darfler, S., Valley, C., (1996), “High Thermal Conductivity of Non-Metallic Honeycomb”, United States Patent, Patent Number: 5498462, Date of Patent: Mar 12 1996.
9. Korab, J., Stefanik, P., Kavecky, S., Sebo, P., Korb, G. (2002), “Thermal Conductivity of Unidirectional Copper Matrix Carbon Fibre Composites” Journal of *Composites,* Part A 33, pp.577-581.
10. Vos, D. L. (1999), “Thermally Conductive Vibration Isolators”, United States Patent, Patent Number: 6002588, Date of Patent: Dec 19 1999.

11. Katsui, T. (2010),"Heat sinks and information processor using heat sinks", United States Patent no-7828045B2 ,Date of Patent: Nov 9,2010
12. Hamada, Y., Otsu, W., Fukai, J., Morozumi Y., Miyatake, O. (2005) "Anisotropic heat transfer in composites based on high-thermal conductive carbon fibres", *3rd International Symposium on Advanced Energy Conversion Systems and Related Technologies, Journal of Energy,* Volume 30, Issues 2–4, February–March 2005, pp 221–233.
13. Goldstein, R.J., Ibele, W.E., Patankar, S.V., Simon, T. W. (2006) "Heat transfer—A review of 2003 literature", *International Journal of Heat and Mass Transfer,* Volume 49, Issues 3–4, pp 451–534.
14. American Society for Testing and Materials, "Manual on the use of thermocouples in temperature measurement." *ASTM Committee E20 on Temperature Measurement,* 1993 ASTM International.
15. Suryakumar, S., Karunakaran, K.P., Bernard, A., Chandrasekhar, U., Raghavender, N. Sharma, D. (2011), "Weld bead modelling and process optimization in Hybrid Layered Manufacturing", Journal of Computer- Aided Design ,vol. 43,pp. 331-344.

In-Situ Property Improvements Using A CNC Integrated Pneumatic Hammer

Dhirendra Rana[1], Pravin M. Kulkarni[1], K.P.Karunakaran[1], Asim Tewari[1]

Rapid Manufacturing Laboratory, Department of Mechanical Engineering,

Indian Institute of Technology Bombay, Powai, Mumbai-400076, INDIA

Email: dhirendra.rana@iitb.ac.in

Abstract:

Purpose: Residual stresses exist in metallic objects without the application of any service or external loads. These stresses can be devastating to stimulate the fatigue failure due to crack initiation. Hybrid Layered Manufacturing (HLM) is an Additive Manufacturing (AM) process for metallic objects using MIG welding deposition. HLM objects are functionally hindered due to preoccupation of residual stresses and porosity in it. To overcome these limitations, a new in house build in-situ property improvement technique using hammering is proposed in this paper.

Design/Methodology/Approach: HLM process is capable of manufacturing most complex parts but some manual operations are needed to make objects competent to conventionally manufactured objects. The authors describe the method of property improvement of metallic objects using an in-house hammer. After deposition of each layer, hammering operation is performed.

Findings: It was observed that the density of the sample after hammering (2.68 g/cm3) was almost equal to the original density of the sample which is 2.69 g/cm3. The samples produced without hammering showed the maximum percentage of closed porosity of 20.13% and 17.39%; whereas the closed porosity after hammering was found 2.78% and 0.08% measured in step over and weld direction respectively.

Research Limitations/Implications: HLM realize near net shape of complex objects, further post processing namely machining and hammering is required.

Originality Value: Property improvement of metallic rapid manufactured objects specially HLM objects.

Keywords: Hammering, Metal RP, Additive Manufacturing (AM), Hybrid Layered Manufacturing (HLM).

1. Introduction

Residual stresses are remained stresses inside any part, in the absence of any external loads or gradients, after all manufacturing processes are over. The nature of these residual stresses can be either tensile or compressive. For example, a welded joint will contain high magnitude residual tensile stresses in the heat-affected zone (HAZ) adjacent to the weld. Conversely, the surface of induction hardened components may contain residual compressive stresses [1]. Residual stresses influence many engineering properties particularly fatigue life, dimensional stability, corrosion resistance and brittle fracture. Welding is more prone to residual stresses, the reasons of which are non-uniform thermal expansions and solidifications.

Weld bead solidifies from liquid phase to solid phase thus volumetric contraction takes place. As a result of which base metal, on which deposition is being done, get stressed more and more after each layer is deposited. This results in the warping of the whole component. In the case of extreme warping it becomes mandatory to chop off the fixture bolts to get the final product out. Figure 1 shows these factors which imply the need of property improvement.

Fig. 1 (a) Impeller with warping

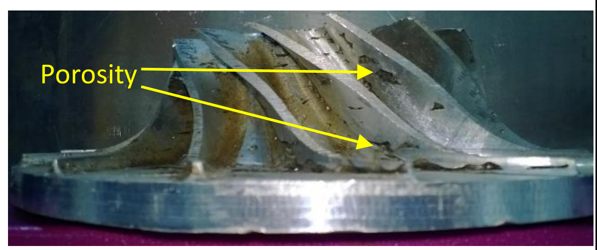


Fig. 1 (b) Impeller with porosity defects

Fatigue is basically a surface related phenomenon as the fatigue cracks usually initiate at the surface and grow from there into the material. Surface hardening is therefore often used to improve fatigue properties and an important example of this is hammering. The in-situ pressing of layers work hardens the surface layer and induces compressive residual stresses. Roughly speaking, the residual stress acts as an applied mean stress and a compressive residual stress will therefore retard fatigue crack initiation and growth [2-4]. The work hardening results in an increased dislocation density which hinders dislocation movements due to the fatigue load and suppresses localized plastic deformation which is a starting feature for crack initiation. The objective of this work is to:

i. Map the residual stress behavior
ii. Evaluate it determining parameters
iii. Develop in house build in-situ residual stress pneumatic hammer

It is a reliability check for any mechanical object during its functional operation. The various methods of stress measurement can be categorized as: destructive, semi-destructive and non-destructive. The destructive and semi destructive techniques are measured from the deformations produced. These methods are also known as mechanical methods. Sectioning and contour methods are destructive methods as these involve complete destruction of the specimen.

Organization of Paper

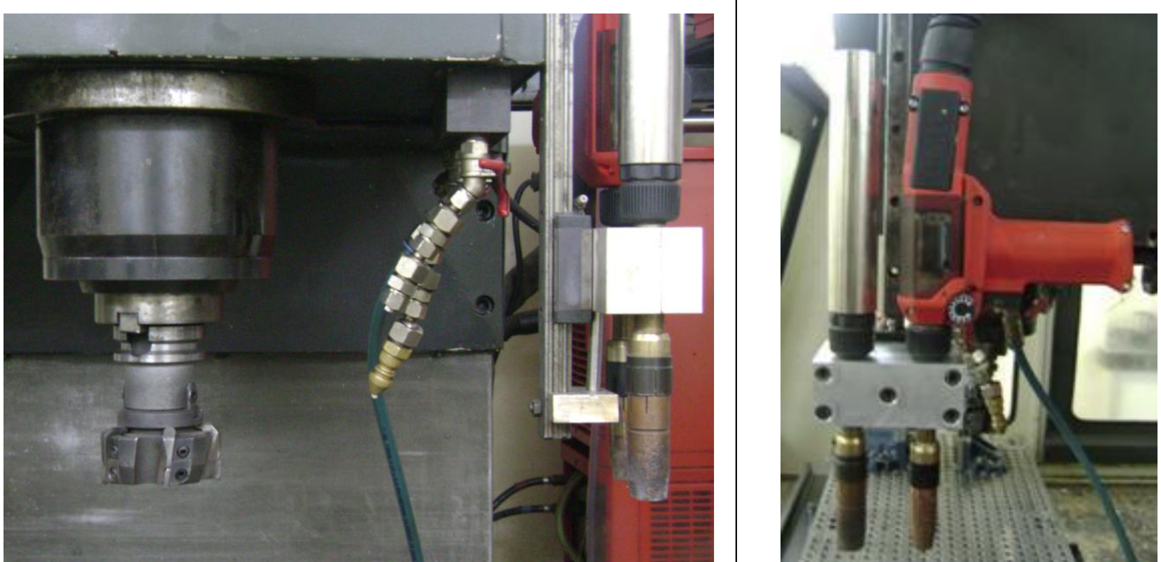
In this paper authors define the residual stress and its disadvantages. The principle of HLM process and need of property improvement in HLM parts is discussed. Then various property improvement technologies including hammering process are discussed. Finally the effects of hammering on HLM parts are discussed.

2. Hybrid Layer Manufacturing Process

HLM realizes shapes using *GMAW* deposition. Such additively manufactured matrix inherently has porosity and residual stresses [5]. Majority of the applications require homogeneous and dense structures. It is possible to achieve the desired density and dimensional stability by the appropriate selection of the property improvement parameters. In order to understand the influence of these process parameters, it is required to assess the nature and extent of porosity and residual stresses in object.

HLM is a RM process that combines the best features of CNC machining (subtractive method) and Rapid Prototyping (additive method) approaches [6]. HLM uses GMAW for layered deposition. This deposition makes near-net shape of the object to be built; the near-net shape is then finish machined subsequently. Near-net shape building and finish machining happening at the same station is the unique feature of this process. This two-level approach focuses on material integrity during material addition and geometric quality during material subtraction. One may subject the near-net shape to stress relieving or heat treatment as required before finishing. Time and cost saving of this process can be attributed to reduction in NC programming effort and elimination of rough machining. It is envisioned as a low cost retrofitment to any existing CNC machine for making metallic objects without disturbing its original functionalities. Customized software generates the NC program for near-net shape building.

Figure 2 shows the HLM unit which comprises of GMAW gun for deposition and milling cutter for subtraction.

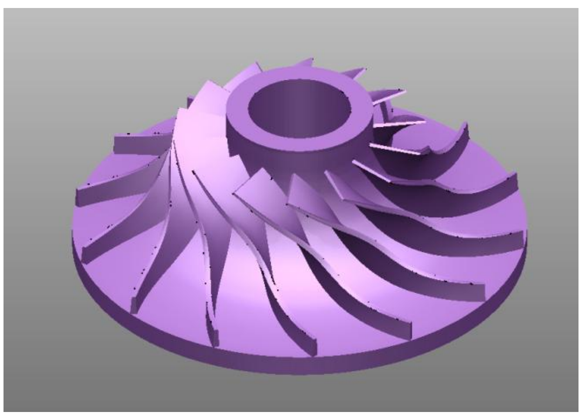


(a) Front view (b) Side view

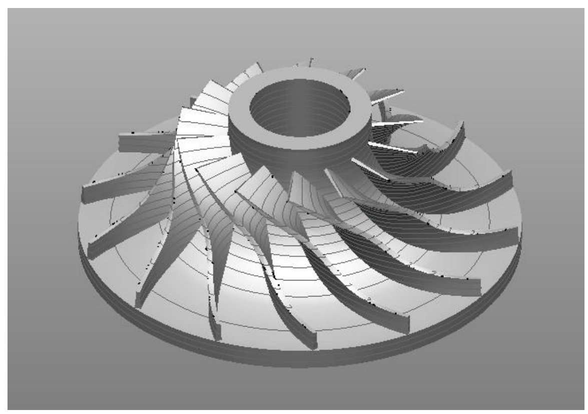
Fig. 2 HLM Process Setup [7]

Today's technology in additive manufacturing (AM) and processing has enabled fabrication of fully functional nonmetallic objects. But the AM of metallic parts is still in developing phase and fully functional metallic parts are yet to come. This paper discuss the property improvement of metallic AM parts so that they can endure the challenges same as conventionally manufactured parts. A case study of impeller which was manufactured using HLM process is shown below in figure 3.

Therefore, current work focus has been on property improvement by eliminating porosity and residual stresses with the help of a CNC integrated hammer. The results obtained from this study are very satisfying as the properties became comparable to that of conventionally manufactured objects.



(a) CAD Model



(b) CAD Model after slicing

(c) Near net shape

(d) Finished part

Fig. 3 Case Study Shows HLM Process

The deposition rate and layer thickness of HLM are an order of magnitude of the competing laser/electron beam based processes (100 g/min against 2-8 g/min; 2mm against 0.04-0.1mm). These parameters require a firm clamping mechanism during deposition and machining. Since the substrate is mechanically constrained, extensive residual stresses are built up under such severe cyclic thermal conditions. When the mechanical constraints are removed, these stresses release themselves in the form of distortions. While the residual stresses shorten the fatigue life, the distortions lead to dimensional inaccuracies. In other words, residual stresses and distortions are the two interchangeable undesirable evils, the former more harmful as the latter can be absorbed into higher machining allowance. Therefore HLM is modified to minimize residual stresses by permitting them to be relieved or distributed on a wide area. This modification was done by integrating a pneumatic hammer to 3-axis CNC machine.

3. Property Improvement using Pneumatic Hammer

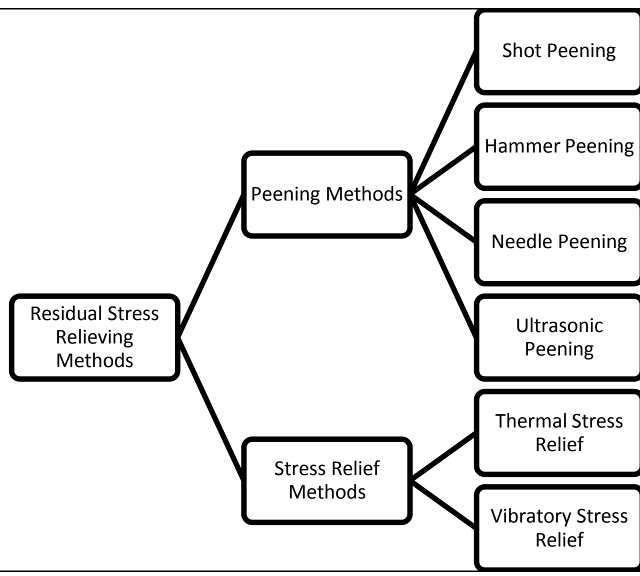


Fig. 4 Various Residual Stress Relieving Methods

The peening of layers by the CNC integrated pneumatic hammer is a hot working process used to produce a compressive residual stress layer and modify mechanical properties of HLM produced metals. It entails impacting a surface with a round tip hammer with force sufficient to create plastic deformation. Hammering on a surface spreads it plastically, causing changes in the

Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA,

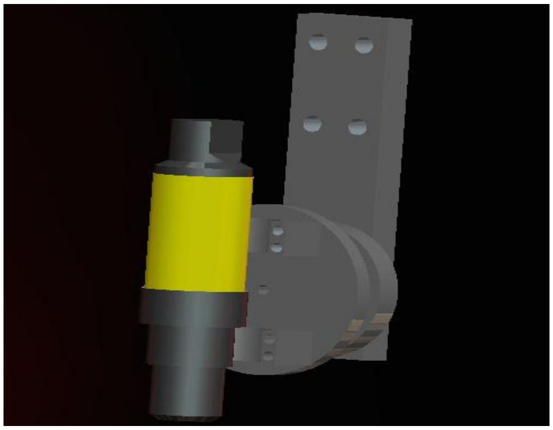
mechanical properties of the surface. Depending on the part geometry, part material, hammering intensity, hammering coverage, hammering can increase fatigue life.

Plastic deformation induces a residual compressive stress in a hammered surface, along with tensile stress in the interior. Surface compressive stresses confer resistance to metal fatigue and to some forms of stress corrosion. The tensile stresses deep in the part are not as problematic as tensile stresses on the surface because cracks are less likely to start in the interior.

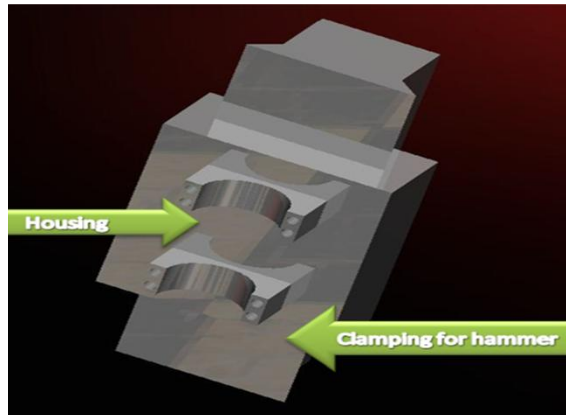
3.1 Process Description

Different design options were taken into consideration during designing the holding fixture to the CNC machine. Figure 5 shows different options for holding the pneumatic hammer in position and options to mount the assembly on the machine. The hammer was disassembled from the pneumatic hammer during milling and welding to avoid collision between the hammer and the work piece.

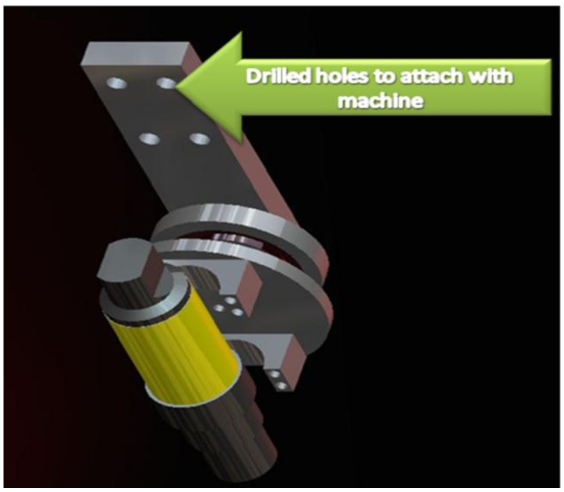
The compressive residual stress in a metal alloy is produced by the transfer of force from a vibrating hammer into the surface of a material with the capacity to plastically deform. The mechanics of the hammering involve properties like shape, and structure; as well as the properties of the work piece. Factors for process development and the control for force transfer for pneumatic hammer are: air pressure, the vibrating frequency, impact force and work piece properties.



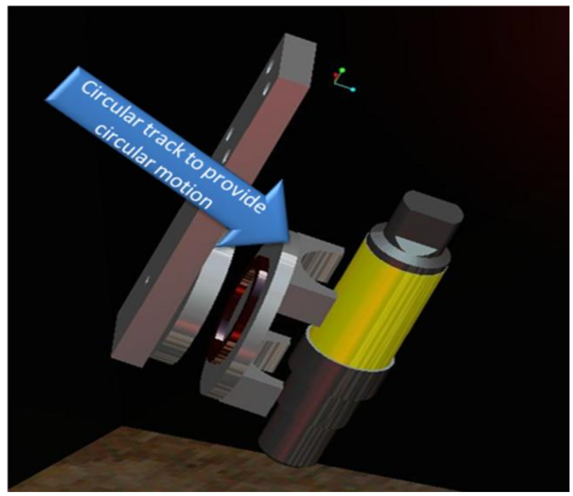
(a) First design option



(b) Housing and clamping for hammer



(c)Drilled holes to attach with machine



(d) Circular track to provide circular motion

Fig.5 Design alternatives for attaching the pneumatic hammer with the CNC machine

The coverage, the percentage of the surface indented, is subject to variation due to the step over increment to cover the work piece surface. Processing the surface with a series of overlapping passes improves coverage. Maintaining the distance between the hammer tip and the top surface of the object is an important parameter. This distance influences the force applied during hammering, if the distance between the tip of the hammer and the sample is too short the hammer will vibrate at high frequency and the applied force on the sample will be very high. The high hammering frequency will lead to excessive applied force and this may result in cracks on the sample and will be noisy for the operator. Optimizing coverage level for the process being performed is important for producing the desired surface effect.

The residual stresses and distortions are the two interchangeable manifestations of the non-uniform heat input. If the distortion is prevented by rigid clamping, it may lead to internal stresses. Hence it is preferable to deposit and peening each layer so as to allow the accumulated residual stress to be removed. The distortion is that is occurred during deposition and peening is absorbed into the machining allowance.

Stresses in HLM can be relieved by [i] pressing in-situ at the end of each layer, [ii] releasing the mechanical constraints during deposition and [iii] a post-build furnace treatment. Experiment was carried out for the first strategy.


(a) Pneumatic hammer, milling cutter and welding torch assembly on 3-axis CNC machine


(b) Step 1: Weld deposition of each layer


(c) Step 2: Milling of each layer to required layer


(d) Step3: Hammering every layer after milling

thickness	operation

Fig.6 Experimental setup and sequence of operations

Two samples were produced one without pressing the layers and one sample with pressing of each layer in-situ after deposition. When the deposition was done with online pressing of each layer, residual stresses were found to be significantly less. Each layer was subjected to a compressive load of 3kN using a numerically controlled hydraulic press. Measurements of residual stresses were carried out by X-Ray Diffraction. The in-situ peening of layers results a change from a complete tensile stress to a compressive stress.

The force applied during hammering was calculated using the basic formula of single acting pneumatic cylinder.

$F = PA = P\times\pi\times D^2/4$

Where,

F = force applied in kg-F

P = Gauge pressure in kg/mm^2

D = inner diameter of the cylinder

Now, $F= P\times\pi\times D^2/4$ P = $.07 kg/mm^2$, and D = 24mm,

$=.07\times\pi\times 24^2/4$

=316.67 kg-F =3.1 kN

3.2 Experiments before and after pneumatic peening

The following experiments were performed to observe the changes on the material after and before pneumatic peening.

1. Variations in residual stress
2. Variations in density
3. Changes in porosity

3.3 Variations in Residual Stress before and after peening

Any mechanical and thermal activities that lead to residual stress have been studied. It was observed that high residual stresses remain on the block and leading to surface cracks. To solve this problem the in-situ peening has been applied. The maximum residual stress profile can be

Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA, Copyright © 2014, ISPE USA and ISAM USA

affected by the factors of hammering parameters and part geometry, part material, hammering quality, hammering intensity, and coverage.

Two samples of Al-Si are fabricated, out of which one specimen made through conventional HLM process, and second one by peening of each layer of specimen. A total of 4 specimens were sectioned at different sections from the samples. Two samples were prepared for surface residual stress testing and other two samples were prepared to examine the sub-surface residual stress after and before peening. As shown in table 1 a surface and sub-surface compressive residual stress was measured using x-ray diffraction.

Table 1 variations in surface and sub-surface residual stress after and before peening

Sample name	Testing position	Residual Stress	
		Normal stress (MPa)	Remark
S1	Surface residual stress before peening	+123.3	Highly tensile, surface cracks
S2	Surface residual stress after peening	-20.9	Compressive residual stress, arrests surface cracks
S3	Sub-surface Residual stress before peening	+59.9	Accumulated residual stress
S4	Sub-surface residual stress after peening	-19.6	Almost equal to sample 2

It can be easily observed from table 1 that surface residual stress before peening is completely tensile (123.3 MPa) and this will lead to surface cracks. However the residual stress after peening reduced dramatically to a completely compressive residual stress of-20.9MPa which helps in arresting the initiating of surface cracks. Similarly the sub-surface residual stresses before and after peening showed a tensile and compressive residual stress respectively.

Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA,

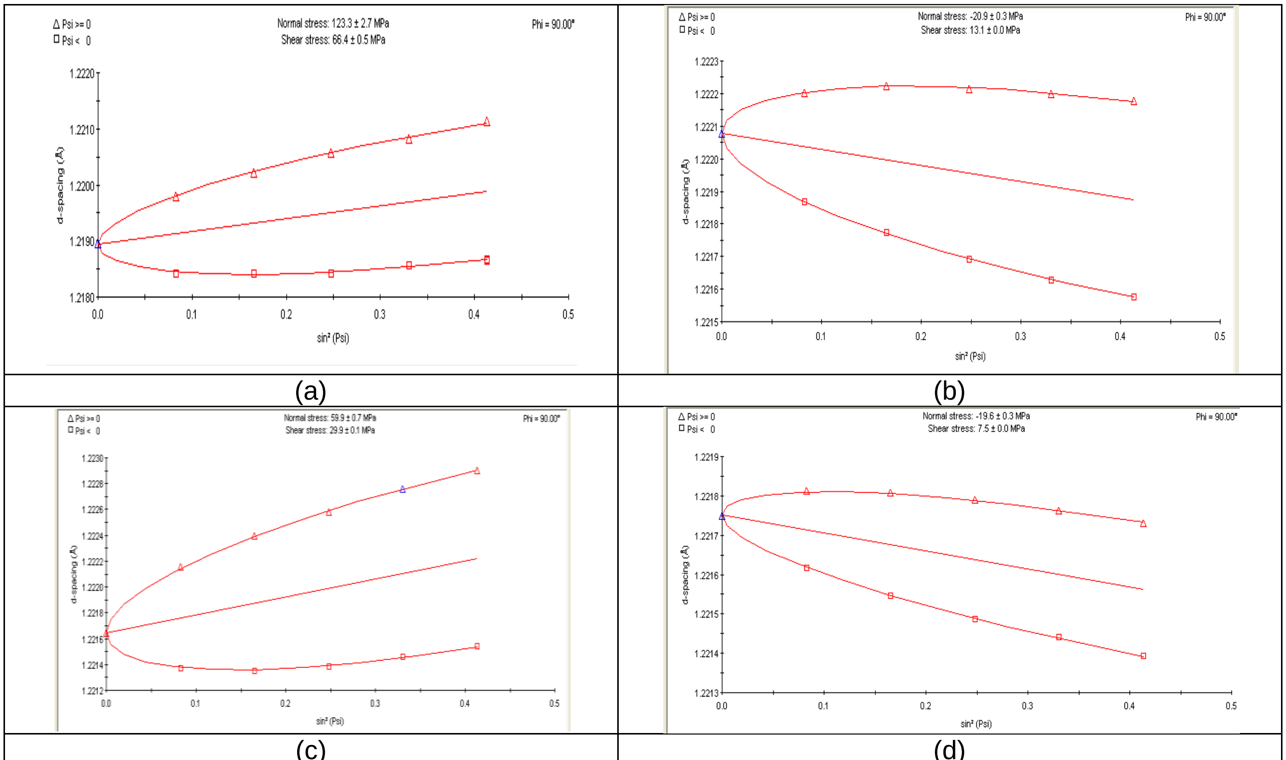


Fig.7 Measured surface and sub-surface residual stress, (a)Residual stress on the surface without hammering, (b) Residual stress on surface after hammering, (c) Residual stress across the middle section of the sample without hammering, (d) Residual stress across the middle section of the sample after hammering

It is important to note that the measured sub-surface residual stress and surface residual stress after peening of each layer for sample 2 and 4 is almost equal. It can be concluded that both surface and sub-surface residual stresses are relaxed due to the peening load.

3.4 Variations in Density before and after peening of each layer

Density of a material is a very important physical property. Essentially it is a measure of the atoms that make up the material, and crystallographic packing configuration, which establishes the density of the compound or element. The density of a material can be changed by changing the crystallographic packing even when the atom makeup of a material remains constant. Density measurement is a key tool in understanding the densification behavior of materials via different processing routes. The measurement of density is a simple process that yields a very important first step in quality control measures. With an understanding of the theoretical density of a material, the utilization of advanced density techniques can yield quantitative open and closed porosity data and an understanding of the pore structure.

HLM inherently has processing defects such as porosities and non-uniform microstructure. So it is logical to investigate the improvements of objects made by the conventional HLM and the modified HLM technique. Density is linked with many factors, including method of densification, shape and size of the part, particle size, its compressibility, and rate of pressing. Density can influence mechanical strength, hardness, electrical conductivity, and magnetic and gas permeability.

The Archimedes principle was applied to compare the density of the object before and after peening of each layer. Archimedes principle states that when an object is fully submerged, the buoyant force, or apparent loss in weight, is equal to the weight of the fluid displaced.

Once the dry weight (mass) is obtained, the submerged weight is measured. The weight loss of the material is then determined and subsequently combined with the density of the fluid to determine the volume shown in Eqn. 3.

$$\text{Volume (cm}^3) = \frac{W_{dry} - W_{wet}\ \ (g)}{\text{Fluid Density } (g/cm^3)}$$

During the use of a surfactants or hot boiling water allows penetration of surface porosity for more accurate measurements, however, it must be noted that the temperature can affect the density of both the material and the fluid thus altering the final measured values. Furthermore, the wire supported the material to be measured must be very fine, otherwise error can be introduced into the wet weight of the piece in question.

The experimental procedure during density measurement was explained as follows. Measure the dry weight of material first to obtain to the mass of the part. After the dry weight has been accurately obtained, the part was then be placed in the holder and submerged in the water bath. The submerged or wet weight of the part was measured to determine the volume. The submerged part should be subsequently re-weighed 5 times to verify the validity of the results and generate an average wet weight value. The wet weight can then be used in the Archimedes equations to generate the density in grams per cubic centimeter.

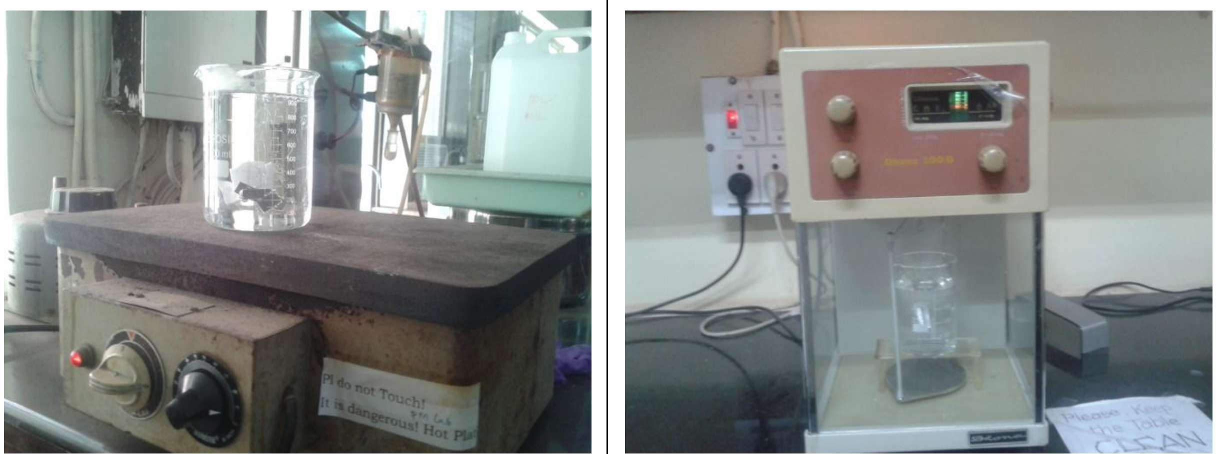


Fig.8 Actual Experimental Set-up for density variation calculation

During experiment in measuring the density of the sample, open and closed pores were taken into consideration. The following equations are driven and used in measuring the density of the sample by considering the volume fraction of open and closed pores.

$$Volume\ of\ sample = \frac{Dry\ weight - Suspended\ weight}{\rho w} \quad (1)$$

$$Volume\ of\ open\ pores = \frac{Soak\ weight - Dry\ weight}{\rho w} \quad (2)$$

$$Relative\ density\ of\ sample = \frac{Dry\ weight}{Dry\ weight - Suspended\ weight} \quad (3)$$

$$Volume\ fraction\ of\ open\ pores = \frac{Open\ pores}{Volume\ of\ sample} \quad (4)$$

$$Volume\ fraction\ of\ closed\ pores = \frac{Closed\ pores}{Volume\ of\ sample} \quad (5)$$

Based on the above formulas the open and close porosities are calculated and tabulated as shown in table 2.

Table 2 Measured densities and percentage of porosity with and without peening the layers

Sample Name	Manufacturing Method		Dry weight of sample	Suspended weight of sample	Soak weight of sample	Volume of sample	Volume of open pores	Volume of closed pores	Density	Open porosity (%)	Closed porosity (%)
	Hammering	Test direction									
1	Yes	Stepover	2.4700	1.5226	2.4728	0.9474	0.0028	0.026384	2.61	0.30	2.78
2	Yes	Along	3.3527	2.1013	3.3568	1.2514	0.0041	0.000943	2.68	0.33	0.08
3	No	Stepover	1.5358	0.8014	1.5514	0.7344	0.0156	0.147871	2.09	2.12	20.13
4	No	Along	1.2015	0.6434	1.2159	0.5581	0.0144	0.097046	2.15	2.58	17.39

It was observed that the density of the sample after hammering was almost equal to the original density of the sample which is 2.69 g/cm3. However the density of the sample produced without peening was found to be less than the original one. The other important observation is the percentage of open and closed porosities. For the sample produced without hammering and tested in stepover direction and along the weld directions showed the maximum percentage of closed porosity of 20.13% and 17.39% respectively. The percentage of closed porosities (2.78%) is high relative to the open porosities (0.3%) in the sample built with peening of each layer and tested in stepover direction in contrast, for the sample tested along the weld direction the percentage of open porosities is higher than that of closed porosity, which was 0.3% and 0.08% respectively. It is clear from the experimental results that the total percentage of porosities was found in the stopover directions for both samples produce with and without hammering of the layers.

4. Advantages and limitations of the process

The hammer integrated with CNC machine is in working after deposition. The features of the hammering process can be listed as:

Advantages

i. Improves the feasibility of HLM for producing objects for real world application, Composite Dies.
ii. Provides firm bond between layers and better material integrity.
iii. Enhances part performance and improves fatigue strength.

Limitations

i. The peening load may transfer to the machine bearings and may cause some problem.
ii. May not be applied to complex shapes with internal cavities.
iii. Might be a challenge to integrate with 5-axis CNC machine, a complex and rigid support mechanism may be required.

5. Conclusions

Peening, vibration and heat treatment were identified as the three broad methods suitable for inter-layer stress relieving. Vibration is actually stress redistribution rather than relieving. Peening too does not relieve the stresses, but only adds a compressive layer on the surface that arrests the crack propagation. Thus, complete stress relieving takes place only in heat treatment due to recrystallization. Hence, it was concluded that inter-layer stress relieving is redundant in HLM and only post-weld heat treatment is preferred. In addition, attention should be paid to reduction of heat input and management of its manifestations as follows:
Based on the experimental results the following claims can be drawn from the newly developed in-situ residual stress relieving technique.

i. Method to increase the density of metal by hammering the weld deposited layers, typically results in compacting the grains of the metal.
ii. Method to relieve residual stress by hammering each layer to distribute the accumulated residual stress.
iii. Method to eliminate internal porosity and voids in each layer, improved mechanical properties.

References

1. Jack Champaigne "Shot Peening Overview", http://www.shotpeener.com , Date: 13/04/2013
2. Shigeru Aoki, Tadashi Nishimura , Tetsumaro Hiroi , Seiji Hirai "*Reduction method for residual stress of welded joint using harmonic vibrational load"*, Nuclear Engineering and Design 237 (2007) 206–212
3. Shigeru Aoki , Tadashi Nishimura, Tetsumaro Hiroi *"Reduction method for residual stress of welded joint using random vibration"* Nuclear Engineering and Design 235 (2005) 1441–1445
4. http://www.metalimprovement.com, Date: 12/07/2013
5. Sreenathbabu Akula, K.P. Karunakaran, *"Hybrid adaptive layer manufacturing: An Intelligent art of direct metal rapid tooling process"* Robotics and Computer-Integrated Manufacturing 22 (2006) 113–123
6. Suryakumar, S., Karunakaran, K.P., Bernard, A., Chandrasekhar, U., Raghavender, N. Sharma, D. (2011), "Weld bead modelling and process optimization in Hybrid Layered Manufacturing", Journal of Computer- Aided Design ,vol. 43,pp. 331-344
7. K. P. Karunakaran, S. Suryakumar, Vishal Pushpa, Sreenathbabu Akula , *"Retrofitment of a CNC machine for hybrid layered manufacturing"* Int J Adv Manuf Technol (2009) 45:690–703

A Review on Measurement of Agility in Manufacturing System

S.K.Sharma[a], Shaarabh Muraka[b], Rishi Gupta[a,*], H.P. Choudhury[a]

[a]Department of Mechanical Engineering, Indian Institute of Technology, Banaras Hindu University,Varanasi-221005, India, Email: sksharma.mec@itbhu.ac.in, rishigupta.mec12@iitbhu.ac.in, hpchoudhury@gmail.com
[b]Department of Mechanical Engineering, Birla Institute of Technology and Science (BITS) Pilani, Rajasthan-333031, India, Email: shaarabh@gmail.com

ABSTRACT

Shifting paradigms gave rise agile manufacturing the concept of which was later applied at a broader level of organizations, workforce, enterprise as well as supply chains. As trend of agility is being recognized as a necessary step to compete in the highly turbulent environment it becomes necessary to evaluate the agility of systems. As agility means different to different people the evaluation of a system's agility has been attempted with diverse approaches. This paper reviews various agility evaluation methods proposed and attempts to provide a holistic view of every method and draws out their advantages and limitations, in order to help in further research in this field.

Keywords: Agile Manufacturing; Fuzzy Logic, Fuzzy Association Rule Mining (FARM), Manufacturing System; Manufacturing Supply Chain (MSC),

1. Introduction

21st century has witnessed humongous technological advances which has laid foundations for highly dynamic structure of today's business environment. This tendency is increasing day by day due to rising competitions. Globalization has resulted in blurring the demographic differences increasing accessibility for both customers and suppliers. This has empowered the customers to demand best products at better prices, lesser time, more customization, and arbitrary quantity. This has increased problems for enterprises who want to maintain their market share as they have to deal with highly unpredictable, dynamic and turbulent environment. Many different solutions have been put forward such as networking, reengineering, modular organizations, virtual corporations, high performing organizations, employee empowerment, flexible manufacturing, just-in-time (JIT), etc. (Sherehiy et al., 2007). However the most prominent approach the use of agility.
The term 'agility' was first introduced in the Iacocca Institute (Report 1991), describing it as one key to future competition. They defined agility as manufacturing system with capabilities to meet the rapidly changing needs of the marketplace. Kidd (1994) defined agility as a rapid and proactive adaptation of enterprise elements to unexpected and unpredicted changes. Goldmann et al., (1995) state agility as being a comprehensive response to challenges posed by a business environment dominated by change and uncertainty. According to Gunasekaran (1999), agile manufacturing (AM) is the ability of surviving and prospering in a competitive environment of continuous and unpredictable change by reacting quickly and effectively to changing markets, driven by customer-defined products and services. Yusuf et al., (1999) proposed that agility is the successful application of competitive bases such as speed,

flexibility, innovation, and quality by the means of the integration of reconfigurable resources and best practices of knowledge-rich environment to provide customer-driven products and services in a fast changing environment. Although the definitions differ from each other most of them discuss common characteristics of agility namely-time, flexibility of the system, and the ability to response (responsiveness).From these definitions we can infer that organizations which are able to react and respond to an environment that frequently or continually pressure the firm to change while continuing to satisfy customers and achieve business objectives can be considered agile.

Framework for agility has been extensively discussed and developed. Goldman et al. (1995) developed four main strategic dimensions that underline the achievement of agile competitive capabilities. These authors suggest that each company can create a program of agility achievement based on the audit that relates the agility dimensions with current and future company operations. The proposed strategic dimensions of agility are: enriching the customer, cooperating to enhance competitiveness, organizing to master changes and leveraging the impact of people and information.

Agility is based on several capabilities found in three main enterprise dimensions: manufacturing, product and market dimensions. Jackson and Johansson (2003) divided agility capabilities into four main dimensions: product-related change capabilities, change competency within operations, internal and external co-operation and people, knowledge, and creativity.

Yusuf et al. (1999) identified competitive foundations of agility as follows: speed, flexibility, innovation, proactivity, quality, and profitability. They illustrated three aspects of agility related to different levels of enterprise. Elemental agility refers to individual resources (people, machinery and management); micro-agility refers to the enterprise, and macro-agility to the inter-enterprise level. They proposed a comprehensive list of 32 agile attributes characterizing an agile enterprise.

Gunasekaran (1998) identified seven key enablers of agile manufacturing and discussed a framework taking into account customization and system integration. The key enablers of agile manufacturing include: virtual enterprise formation tools/ metrics, physically distributed manufacturing architecture and teams, rapid partnership formation tools / metrics, concurrent engineering, integrated product/production / business information system, rapid prototyping tools and electronic commerce. They illustrated a framework integrating the major capabilities of agile manufacturing, such as co-operation, value-based pricing strategies, investments in people and information, and organizational changes with appropriate enablers.

A rounded and concise conceptual framework was proposed by Sharifi et al. (2001). They proposed four main aspects of AM: agility drivers, strategic abilities, agility providers, and agility capabilities. The agility drivers represent factors of the external business environment creating turbulence and unpredictability of the changes which drive an organization to adapt agility. Strategic abilities determined through factors such as responsiveness, competency, quickness, and flexibility are considered as main attributes of the agile organization that allow adaptation to changes. The agility capabilities could be achieved by the means of agility providers. Agility providers can be derived from four manufacturing areas: organization, technology, people, and innovation .They argued that only by integrating these criteria’s agility could be achieved.

Tsourveloudis and Valavanis [2002] propose a knowledge-based framework for the measurement and assessment of manufacturing agility. In order to calculate overall agility of an enterprise, a set of quantitatively defined agility parameters is proposed and grouped into production, market, people and information infrastructures.

Ismail et al., [2006] propose an agility strategic framework combining external parameters and indicators for the enterprise's agile capability. The model consists of a business environmental audit with a number of environmental turbulence indicators and agility capability indicators covering product, process, operation, people, and organization. The product dimension is categorized into the areas of customization, re-use, re-place ability, similarity and mix. Within these dimension, concepts like product family efficiency, component commonality and process similarity are introduced to provide a rapid means to analyze vulnerabilities and plan for corrective action. The framework supports the mass customization paradigm of manufacturing.

As there is no definite method of achieving agility, defining a holistic model for it is difficult. The literature on the subject suggests various framework, models and correlations for agility measurement that can be implemented by organization to adapt to current business conditions. However assessment of organization's agility creates a problem due to inexact boundary in which the agility is defined. This is the reason various different methodology have been used to assess agility. Some measure the agility of the structure of organization, some focus on agility as purely a competitive outcome, indirect methods such as measuring complexity have also been used to measure agility of the organization. In this paper we discuss these methods of assessment and analyze them for their advantages and their shortcomings.

2. Work Done on Measurement of Agility:

A lot of research in the field of developing an agile model has been carried out but assessment of agility of such organization isn't extensively researched. This can be attributed to the fact that the metrics of agility is difficult to design due to the vagueness with which agility is defined. The term agility is understood in a broader perspective and is influenced by many characteristics which may or may not be same for all types of organizations. Accounting for the broadness the models developed are diverse and focus on certain areas. Many of the initial models develop were empirical but later more extensive and flexible models developed using more broad techniques integrating organization and work place.

Dove, [1996] identifies four change-proficiency metrics for agility: cost, time, robustness, and scope. With higher agility, the effort to change decreases, approaching an ideal where it takes no time, incurs no cost, has immediate and robust results and is not an inhibiting factor on the latitude of opportunity and innovation . Dove, (2001) concludes that an organization with change proficiency addresses all four metrics.

Sharifi et al., (2000) laid out a framework which consisted of various steps for implementing agility in various organizations. The first step in consists of determining the nature of environment the organization functioned in this was done by calculating the agility need index through a set of questionnaires. The next step was the assessment of the company itself in terms of agility, whether the company is agile or not based on certain factors such as responsiveness, customer satisfaction etc. Then a gap analysis presents the plan of action .The last step uses a set of viable tools, agility providers to attain agility, by which capabilities of the organization can be achieved. A case study was conducted by them to verify the model. The framework developed was one of the initial efforts to model agility assessment and to identify the need for agility. A similar framework for agility assessment was developed by

Jackson and Johansson, (2001) in which they mentioned a three step approach for evaluation. First being evaluating the market trends, second being analysis of the strategic objectives in order to find out whether flexibility competency is a long term objective and the areas of potential development, the final part of the analysis is to find out the capabilities that needs to be focused on. The paper mentioned the need of supplier and consumer participation to improve the results. Van Hoek et al., (2001) also proposed an empirical model based on five characteristics of agility – customer sensitivity, virtual integration, process integration, network integration and measurement. The managers need to mark their organization based on these characteristics based on 5 point Likert's scale and the overall agility is measured average of individual characteristics.

Ren et al., (2000) proposed agility assessment based on Analytical Hierarchical Process (AHP). Decision makers first apply pair wise comparisons to evaluate the agility capabilities of a company. Combining the judgements against capabilities fetched the overall agility of the organisation. Pairwise comparisons are particularly useful when direct comparison of capabilities is difficult. It is an intuitive method of comparing and gives better estimates compared to direct rating methods.

Yang and Li, (2002) proposed a multi graded fuzzy approach which was used to evaluate the agility of mass customisable (MC) organisation. This model uses weights and ratings matrices from experts who evaluate the organisation based on crisp numbers. It evaluates agility using weighted rating matrices of various agility attributes. The case study performed on MC manufacturers included a three grade structure specific to mentioned organisation.

An agility index model was developed by HP Services in conjunction with INSEAD focuses on process agility [2004]. At least twenty processes were analyzed in four vertical markets (Telecom, Financial Services, Manufacturing, and IT service providers) along three dimensions. The dimensions included in the index are time, range, and ease. Time is the amount of time required to effect change. Range is the extent of change activities across an organization. Ease measures the effort involved in effecting change.

Agrawal et al., [2007] identified fifteen variables defining the agility of a supply chain. That focuses on process operation and organization. The proposed factors that describe the agility are: Market sensitiveness, Delivery speed, Data accuracy, new product introduction, Centralized and collaborative planning, Process integration, Use of IT tools, Lead time reduction, Service level improvement, Cost minimization, Customer satisfaction, Quality improvement, Minimizing uncertainty, Trust development, Minimizing resistance to change. The model helps to understand the variables of an agile supply chain and their interrelationship.

Some process-based approaches to assessing agility concentrate on a single type of process, such as software development or product development. Sieger et al., (2000) for example, measure responsiveness of companies relative to the product development cycle time. A quantitative agility metric is developed that "provides the desired time-based performance rating which reflects the agility of a manufacturer during product development".

Arteta and Giachetti, (2004) adopted an indirect approach to measure agility. They proposed to use complexity as a surrogate measure for agility. The hypothesis supporting this proposition mentioned was that a less complex enterprise in terms of systems and processes is easier to change and consequently more agile. The model uses Petri Nets, which represent two elements needed to measure process complexity - resources in the process and the interconnection between those resources, to find the state space probabilities needed for the complexity measure. Once the complexity is measured ease of change is determined by introducing new projects, which have certain time and cost associated with them, lesser the total time and cost more is the organization susceptible to change hence is more agile. Although the model can be used to determine agility no conclusions can be derived on how agility can be achieved.

Nicola et al., (2011) presented a technique for the strategic management of the chain addressing supply planning and allowing the improvement of the Manufacturing Supply Chain (MSC) agility in terms of ability in reconfiguration to meet performance.

Another different approach was proposed by Yauch, (2005) which was later refined (Yauch et al.,(2006) Yauch, (2011)). They proposed a quantitative index of agility, based on a conceptualization of agility as a performance outcome, which captures both the success of an organization and the turbulence of its business environment. Based on the 13 factors that cause turbulence in the business environment a turbulence score is calculated. Organizational success using an empirical relation is also calculated which is then used to calculate agility index. This model integrates operational measures and external parameters to determine the agility of the any type of manufacturing organization.

These methods are relatively easy to understand and implement. However, the vagueness with which the agility enabling capabilities are defined leads to inconsistency of the models. Such conditions where the assessments are difficult to make use of linguistic variables helps the evaluators handle ambiguity. These scoring methods are frequently criticized, the main reasons for it are the evaluator’s subjective judgment as well as the multi-possibility and the ambiguity associated with assigning a number to the attributes. To overcome this problem several fuzzy based models were proposed. Research in the field of fuzzy logic has been extensive and it proves to be a capable tool in handling the vagueness and imprecision of the situation (Klir 1995, Zadeh 1965). Fuzzy logic, by making no global assumptions regarding independence, exhaustiveness, and exclusiveness, can tolerate a blurred boundary in definitions (Lin and Chen, 2004). Use of fuzzy concept enables the evaluators/experts to use linguistic terms to assess indicators related to natural language expressions, and each linguistic term is respectively associated with appropriate membership function. Even at management level fuzzy logic has found significant importance (Lin and Chen, 2004). Several methods have been proposed which utilize fuzzy logic in measuring organizational agility. One of the first such models was the IF-THEN approach proposed by Tsourveloudis et al., (2002) which was an outgrowth based on their previous work. The key idea of this approach is to combine all infrastructures namely-production infrastructure, market infrastructure, people infrastructure, and information infrastructure as agile characteristics and their corresponding operational parameters, to determine the overall agility. The agility is then evaluated numerically by an approximate reasoning method using the knowledge that is included in simple IF–THEN rules, which are conditional statements that relate the observations concerning the allocated divisions (IF-part) with the value of agility (THEN-part). These rules need to be redefined for different set of characteristics consequently for new situations. Another model based on fuzzy logic approach was proposed by Lin et al. (2006). It uses a supply chain agility evaluation model based on fuzzy logic and the multi-criteria decision-making (MCDM) to provide a means for both measuring supply chain agility and also identifying the major obstacles to improving agility levels. The model introduces a fuzzy agility index (FAI) which tells about the agility of an organization. It is computed as the weighted average of performance against agile attributes of the supply chain and the relative importance of those attributes. Once the index is calculated it can be compared to give numerical or linguistic values of agility of the supply chain using various methods. Another index, the fuzzy performance index (FPII), is calculated which identifies the agility capabilities which are the major obstacle to improving agility index. Ganguly et al., (2007) defined three metrics to measure responsiveness, market share and cost effectiveness which would help in measuring a company's agility. They proposed the use of their method along with fuzzy logic approach proposed by Lin et al. (2006) in order to arrive at a conclusion regarding the level of agility of any corporate enterprise.

Jain et al.(2008), proposed a model which along with fuzzy logic uses algorithms of association rule mining to derive associations to support decision makers by enhancing the flexibility in making decisions for evaluating agility with both quantitative (tangible) and qualitative (intangible) attributes and criteria. Recently, in data processing, relational databases and association rules have been widely used in support of business operations, and there the size of database has grown rapidly, for the agility of decision-making and market prediction for varying degree of importance for agility evaluation. Therefore, knowledge discovery from a database is very important for sustaining essential information to a business (Berry and Linoff, 1997). FARM methodology has been used for evaluation of agility in supply chains in a real time practical environment (Vinodh et al., 2011). In the model large fuzzy frameworks and effective association rules for agility evaluation are determined by fuzzy support and fuzzy confidence minimum value of which is which is given by the decision makers. In the algorithm adopted, fuzzy association rules scans a database containing both quantitative and qualitative attributes for agility and applies Boolean operators (AND, OR, XOR) to produce large fuzzy support frameworks which are then used to generate fuzzy association rules for evaluating agility. The methodology is a useful and practical method to make the fuzzy association rules from the available database for agility, which enhances the flexibility of making decisions.

3. Highlight of Evaluation Methods

As discussed the methods vary in their approaches of evaluating agility. Hence a direct comparison of each methods based on a common metric would be incorrect. Hence each of the method has been separately analysed. An attempt has been made to come up with the advantages and limitations of each method.

Authors	Description	Advantages	Disadvantages
Zhang and Sharifi, (2000)	The model consists of a frame work which consists of various steps that are- determining their agility need, assessing their current position, determining the capabilities required in order to become agile, and adopting relevant practices which could bring about the recognised capabilities.	• Familiarizes with concept of agility • Provides a holistic view of agility • Ease of computation • Provides tools for achieving agility	• Evaluation based on crisp values • A quantifiable index is not proposed

Jackson and Johansson, (2001)	The model presents a three step approach for evaluation. First is evaluating the market trends, second is analysis of the strategic objectives in order to find out whether flexibility competency is a long term objective, the final part of the analysis is finding out the capabilities that needs to be focused	• Ease of computation • Identifies capabilities needing focus to increase agility	• Evaluation based on crisp values • A quantifiable index is not proposed
Van Hoek et al., (2001)	The model is based on five characteristics of agility – customer sensitivity, virtual integration, process integration, network integration and measurement. The managers need to mark their organization based on these characteristics based on 5 point Likert's scale and the overall agility is measured average of individual characteristics.	• Ease of computation	• Evaluation based on crisp values • Comprehensiveness of the model needs to be improved
Ren et al., (2000)	The model uses Analytical hierarchy process (AHP) to determine agility capabilities which is used with the judgments of organization performance the agility index is calculated.	• Increases the ease and intuitiveness in calculating capabilities	• Assessment based on crisp numbers • Numerous pairwise comparison will be requires depending on number of capabilities considered
Yang and Li, (2002)	The model developed is for a mass customized (MC) manufacturing organization. The model utilizes multi grade fuzzy approach to calculate agility. A three grade evaluation index (specific to MC) is used to measure	• Ease of computation and implementation • Reduces fuzziness of evaluation	• Frame work specific for a MC organization • Evaluation based on crisp values • Doesn't take into account any external factors • Less number of

	agility. Index is orderly calculated as a weighted sum of companies rating against agility capabilities mentioned in the framework.		capabilities are considered
Arteta and Giachetti, (2004)	The model determines complexity and consequently agility owning to the inverse relationship between them. The model uses Petri Nets, which represents two elements needed to measure process complexity - resources in the process and the interconnection between those resources, to find the state space probabilities needed for the complexity measure.	• Determines complexity of the organization	• The model doesn't propose methods improve agility • Measurement of ease of change is difficult • The model in complex
Yauch, (2010)	This model conceptualizes agility as performance outcome of an organization in turbulent environment. Turbulence score and organization success is calculated using empirical correlation which in turn is used to calculate agility.	• Applicable to any type of industry • Includes external factors in calculation • Assessment is dynamic • A holistic approach is used	• The model doesn't propose methods improve agility
Tsourveloudis et al.(2001)	This model combines all infrastructures namely- production, market, people, and information using the knowledge that is included in simple IF–THEN rules, as agile characteristics and their corresponding operational parameters. Based on company's	• Uses fuzzy logic to determine agility	• The method suffers with inflexibility since IF-THEN rules must be redesigned to fit the new situation • Model is complex

	performance overall agility is determined.		
Lin et al., (2003)	The model introduces a fuzzy agility index (FAI) which is calculated as a weighted average of performance rating of various agility attributes and their relative importance. Fuzzy performance importance index (FPII) is used to identify attributes that need to ameliorate.	• Uses fuzzy logic to deal with ambiguity • Identifies attributes that need improvement	• Comprehensiveness of the model need to improve • Does not take external factors into account
Jain et al., (2008)	The model proposed uses Fuzzy Association Rule Mining (FARM) .The model with the help of quantitative and qualitative relational databases derives association rules for evaluating agility.	• Provides flexibility in agility assessments • Evaluation in real time practical environment	• Model is complex • Equal weight age is provided to all attributes

Table1: Advantage and Disadvantage of different Agility Measurement Methods

4. Discussion and Conclusion

As discussed the method of agility evaluation has taken different approaches. Their focus has been diverse and directed to different aspects. Some models Jackson and Johansson, (2001), Zhang and Sharifi, (2000), Van Hoek et al., (2001) have proposed a framework within which agility is calculated. While these models are simple in their approaches more extensive and accurate models were developed. Models using AHP, Ren et al., (2000), multi grade fuzzy approach, Yang and Li, (2002), Jain et al., (2008) employ more extensive approaches. Tsourveloudis et al., (2001), Lin et al., (2006) used fuzzy logic along with previously developed methods to reduce vagueness in rating hence providing a more accurate result. One of the most thorough evaluation methods developed was by Yauch, (2010) which measured the performance outcome as a measure of agility. Certain indirect methods such as measurement of complexity, Arteta and Giachetti, (2004), were also proposed. A problem related with all these model is lack of sufficient data from different organization to set a reference. Although models have covered various aspects varying from organization structure to performance outcome to measure agility need for a universal index is needed in order to set a reference point and to check the validity of models in varied field of applications. Thus due to this shortcoming it is very difficult to choose the most appropriate model for agility evaluation.

References:

[1] Arteta B M, Giachetti R E. A measure of agility as the complexity of the enterprise system. *Robotics and Computer-Integrated Manufacturing*, 2004; 20:495–503.

[2] Bohdana Sherehiy, Waldemar Karwowski, John K. Layer “A review of enterprise agility: Concepts, frameworks, and attributes”, International Journal of Industrial Ergonomics 37 (2007) 445–460

[3] Goldman S L, Nagel R N and Preiss K. Agile Competitors and Virtual Organizations – Strategies for Enriching the Customer, *Van Nostrand Reinhold*, London, 1995.

[4] Gunasekaran A. Agile manufacturing: enablers and an implementation framework. *International Journal of Production Research*, 1998; 36(5):1223-1247.

[5] Gunasekaran A. Design and implementation of agile manufacturing systems. *International Journal of Production Economics,* 1999; 62:1-6.

[6] H. Ismail, 1. Arokiam, 1. Reid, J. Poolton, V. S. Tey. "Agility Capability Indicators for Product Flexibility Assessment", Proceedings of ICAM International Conference on Agile Manufacturing, Norfolk, VA, 18-21 July 2006, (2006).

[7] Jackson M, Johansson C. Agility analysis from a production system perspective. *Integrated Manufacturing Systems*, 2003:14 (6); 482–488.

[8] Jain V, Benyoucef L, Deshmukh SG. A new approach for evaluating agility in supply chains using Fuzzy Association Rules Mining. *Engineering Applications of Artificial Intelligence*, 2008; 21:367–85.

[9] Kidd P T. Agile Manufacturing. Forging New Frontiers. *Reading, MA: Addison- Wesley*, 1994.

[10] Klir G J. Fuzzy sets and fuzzy logic: theory and applications. *New Jersey: Prentice Hall*, 1995.

[11] Lin CT, Chiu H, Tseng YH. Agility evaluation using fuzzy logic. *International Journal of Production Economics*, 2006; 101(2):353–68.

[12] N. C. Tsourveloudis, K. P. Valavanis. "On the Measurement of Enterprise Agility", Journal of Intelligent and Robotic Systems, Volume 33, Number 3I March 2002, pp. 329-342, (2002).

[13] Nicola Costantino,, Mariagrazia Dotoli, Marco Falagario, Maria Pia Fanti, Agostino Marcello Mangini 2011 : A model for supply management of agile manufacturing supply chains

[14] P. M. Swafford, S. Gosh, N. N. Murthy. "A framework for assessing value chain agility", International Journal of Operations & Production Management, Vol. 26 No. 2, pp, 118-140, (2006).

[15] Q. Zhao, H. Ding, H Lui, X. Wei. "Performance Evaluation of Logistics Service for Mass Customization Based on SCM", Proceedings of LSCM 2006 International Conference on Logistics and Supply Chain Management, University of Hong Kong, 5-7 January, 2006, Hong Kong, China, (2006).

[16] Ren J, Yusuf YY, Burns D. A prototype of measurement system for agile enterprise. *Quality Management & Technology*, 2000; 5(4):304–16.

[17] R. Dove. "Best Agile Practice Reference Base - 1994: Challenge Models and Benchmarks", 4th Annual Agility Conference, Agility Forum, Bethlehem, PA., (1995)

[18] Sanchez Luis M, Nagi Rakesh. A review of agile manufacturing systems. *International Journal of Production Research*, 2001; 39:16, 3561-3600

[19] Sharifi H, Zhang Z. Agile manufacturing in practice application of the methodology. *International Journal of Operations and Production Management*, 2001; 21 (5/6):772–794.

[20] Tsourveloudis NC, Valavanis KP. On the measurement of enterprise agility. *International Journal of Intelligent and Robotic systems*, 2002; 33(3):329–42.

[21] T. J. Kurian. "Agility Metrics: A Quantitative Fuzzy Based Approach for Measuring Agility of Software Process", Proceedings of] ICAM International Conference on Agile Manufacturing, Norfolk, VA, 18-21 July 2006, (2006).

[22] Vinodh S., Prakash N H & Selvan K E. Evaluation of agility in supply chains using fuzzy association rules mining, *International Journal of Production Research*, 2011;49(22):6651-6661

[23] Zadeh L A. Fuzzy sets. *Information and Control*, 1965; 8(3):338–353.

[24] Yauch C A and Navaresse D O. Assessing agility at 70 manufacturing plants. *International Journal of Agile Manufacturing*, 2006; 9(2):19-28.

[25] Yauch C A. Measuring agility as a performance outcome. *Journal of Manufacturing Technology Management*, 2011; 22(3):384-404.

[26] Yusuf YY, Gunasekaran A, Adeleye C and Sivayoganathan K. Agile supply chain capabilities: determinants of competitive objectives. *European Journal of Operational Research*, 2004; 159:379–92.

[27] Yang SL, Li TF. Agility evaluation of mass customisation product manufacturing. *Journal of Materials Processing Technology*, 2002; 129:640–4.

[28] Yusuf YY, Sarhadi M, and Gunasekaran A. Agile manufacturing: the drivers, concepts and attributes. *International Journal of Production Economics*, 1999; 62:33–43.

[29] Yusuf YY. Agile supply chain capabilities: determinants of competitive objectives. *European Journal of Operational Research*, 2003; 159:379–92.

A comprehensive review on modeling and optimization of friction stir welding

R.S.S. Prasanth[1], Bhuvnesh Singhal[1], Pritam Singh[1], K. Hans Raj[1]

[1]Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra – 282 110

rss.prasanth@gmail.com; bhuvanagra@gmail.com; pritamdei@gmail.com; khansraj@rediffmail.com

Abstract:
An environmentally friendly joining process called friction stir welding (FSW) is attracting academia since last decade. Currently, FSW is gradually finding its applicability in automotive and aerospace industry, as it is a solid state joining process that addresses the limitations of traditional joining of aluminum composites in particular. Many researchers have experimented on FSW and reported its process conditions and the relation between different weld properties. However, there is a need to comprehensively review the FSW process in terms of different independent and dependent process variables considered so far for the purpose of modeling and optimization of FSW process. This paper presents a comprehensive review of the experimental studies on FSW process. The review suggests that, there was greater emphasis on the study of the FSW process per se and various FSW models were developed. However, as the aforesaid joining process is complex and highly nonlinear in nature, and essentially demands a robust modeling and optimization framework. This review underlines the need of development of appropriate nature inspired algorithms for the optimization of such advanced manufacturing process.

Keywords
Solid state joining process; Friction stir welding, Joining of aluminum, Advanced manufacturing process

Introduction
Joining metals is one of the most important manufacturing processes other three being forming, machining, and casting. But, joining light weight materials such as aluminum alloys has always been a tremendous challenge. But the applications of such alloys are enormous. To reduce the weight of an aircraft, automobile or a high speed ship, it is an imperative to increase the use of lightweight materials, such as aluminum alloys. It is indeed a great challenge to join such light weight materials using conventional joining processes such as Tungsten Inert Gas (TIG) and Metal Inert Gas (MIG) welding processes, as such conventional joining processes melting can destroy the carefully engineered microstructure of composites thereby eliminating their unique properties. For aluminum plates in particular, welding issues such as oxidation, undesirable chemical reactions can cause weak welds that may contain porosity. Therefore, such welds may become more susceptible to failures associated with the metallurgical changes that occur during localized melting. In case of traditional joining processes, welding results in the presence of a tenacious oxide layer, high thermal conductivity, high thermal expansion coefficient, solidification shrinkage and, above all, high solubility of hydrogen, and other gases, in the

molten state [29]. And also, the loss in mechanical properties as compared to the base material is very significant, in case of TIG and MIG. Although the resistance welding technique may be used to join some aluminum alloys, but the difficulty is that, the surface preparation is costly and more over the surface oxide is a key concern. To address such limitations the idea first conceived was friction welding (FW). But, FW severely suffers from a limitation i.e. only small axis-symmetric parts that can be rotated and pushed against each other to form a joint. Therefore the inherent limitation in making joints of aluminum alloys with maximum strength, minimum fatigue and no fracture are identified as non-joinable using traditional welding processes lead to the development of Friction Stir Welding (FSW) at Welding Institute TWI Cambridge in 1991[56, 7]. The basic idea of joining two metals by FSW is remarkably simple as depicted in figure 1. In this process of joining, a special tool mounted on a rotating probe travels down through the length of the base metal plates in face-to-face contact. The tool serves three primary functions, (i) heating of the work piece, (ii) movement of material to produce the joint, and (iii) containment of the hot metal beneath the tool shoulder. Heating is created within the work piece both by friction between the rotating tool pin and shoulder and by severe plastic deformation of the work piece. Consequentially a good joint is produced in 'solid state'. Depending upon the geometrical features of the rotating probe, the material movement around the pin may be vary and exhibit different complex phenomenon. [34]. FSW process causes the material to undergo severe plastic deformation at elevated temperature, to generate ultra fine grain structures [23], which in turn produce good mechanical properties. And therefore, currently FSW is being considered as a most promising, efficient and environment friendly or green joining process. It consumes minimum energy, requires no additional cover gas or flux, does not involve add-on filler materials, can be applied to almost all types of joints such as butt joints, lap joints, T butt joints, and fillet joints and even dissimilar materials can also be joined.

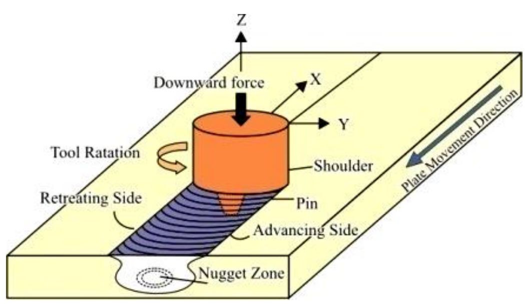


Figure 1: Principle of FSW operation

FSW has some important metallurgical advantages such as (i) solid state process (ii) minimum distortion of workpiece (iii) improved dimensional stability (iv) good repeatability (v) almost no loss of alloying elements (vi) very good metallurgical properties in the weld area (vii) very fine microstructures (viii) does not suffer from fracture (ix) may replace multiple parts joined by fasteners. Besides metallurgical benefits, FSW also has some notable environmental and energy related advantages like (i) shielding gas is not required (ii) surface preparation/cleaning not required (iii) removes grinding wastes (iv) eliminate solvents required for degreasing (v) consumable materials like rugs, wire or any other gases may be saved (vi) uses only 2.5% energy required for a laser weld (vii) low fuel consumption in light weight automotives, ships and aircraft applications [37]. Therefore, by looking at the many advantages and wide applicability of the idea of FSW researchers and engineers across the globe attracted to the study and model this green welding process to achieve better weld and metallurgical properties of different materials. This paper is further organized as follows. In section two, the review of literature of FSW on similar and dissimilar materials is presented followed by a comparison of different machine(s), tools, materials used for FSW process modeling and optimization is presented in section three. In section four modeling and optimization techniques adopted for FSW process with different process parameters on different materials of different specimen specifications have been presented. In section five conclusions are drawn accordingly.

Review of studies on joining different materials using FSW

Some important recent findings of FSW experimentation on, joining similar and dissimilar materials, that are reported in literature available through Google scholar has presented in this section. According to a report [26], friction stir welding of aluminum alloy 2219 using a milling machine is carried out successfully. It is reported that tensile strength of welds was significantly affected by welding speed and shoulder diameter whereas welding speed strongly affected percentage elongation and a maximum joining efficiency of 75% was obtained for welds with reasonably good percentage elongation. R. Karthikeyan and V. Balasubramanian [27], studied FSW parameters such as tool rotational speed, plunge rate, plunge depth, and dwell time play a major role in determining the strength of the joints. An empirical relationship was established to predict the tensile shear fracture load of Friction Stir Spot-Welded (FSSW) AA2024 aluminum alloy by incorporating independently controllable FSSW process parameters. Response surface methodology (RSM) was applied to optimize the FSSW parameters to attain maximum lap shear strength of the spot weld. Researchers [21] developed a simple three-dimensional thermo-mechanical model for friction stir welding (FSW) that allowed partial sliding between the shoulder and the workpiece. The thermal calculation accounts for conduction and convection effects. The complete thermo-mechanical history of the material during the process can then be accessed by temperature and strain rate contours. The numerical results are compared with a set of experimental test cases carried out on an instrumented laboratory device. Mohamadreza Nourani et al., [38], optimized the process parameters of friction stir welding (FSW) of 6061 aluminum alloy, using, ANOVA analysis on the L9 orthogonal array with three factors. Results indicate that among the parameters considered (*i.e.*, the tool rotational speed, transverse speed, and the axial force), the most significant parameter on the weld quality is the rotational speed, followed by the axial force and transverse speed. According to an experimental work [5], rolled plates of AA 2198 T3 aluminum alloy were used for friction-stir welding varying two fundamental process parameters: rotational and welding speeds. They developed two sets of empirical models based on regression analysis, one for welding force and the other for mechanical strength of the welded joints. Model accuracy is reported to be 95% confidence level. C. N. Suresha et al., [55], attempted to identify the most influencing significant parameter and percentage contribution of each parameter on tensile strength of friction stir welded AA 7075–T6 aluminium joints by conducting minimum number of experiments using Taguchi orthogonal array and developed a model for optimization. C. Blignault, et al., [3] developed FSW model using different tool pin geometries, and these data were statistically analyzed to determine the relative influence of a number of combinations of important processes and tool geometry parameters on tensile strength. The model reported in their study allowed the weld tensile strength to be predicted for other combinations of tool geometry and process parameters within an average error of 13%. I. Dinaharan and N. Murugan, in their study [8] used a variety of ceramic particles that are added to aluminum alloys to produce aluminum matrix composites. They attempted friction stir welding of AA6061/ ZrB2. A mathematical model was developed incorporating the FSW process parameters to predict the Ultimate Tensile Strength (UTS). It is reported that the process parameters independently influence the UTS over the entire range studied in this work. K. Hans Raj et al. used a low cost improvised vertical milling machine to study FSP on Al6061 and observed that very fine grains are formed in these sheets. It is found that the hardness has increased by 50% on the processed AL6061 sheets on account of FSP. The study clearly brings out the advantage of severe plastic deformation caused on account of FSP and HPT and is a step forward in developing Bulk Nano-structured Materials (BNM) materials in future [16]. Atul Suri, Prasanth, R.S.S. and K. Hans Raj studied the friction stir welding process using a modified vertical milling machine and compared its performance to TIG to find the improved mechanical properties of FS weld specimen when compared to TIG specimen [58].

FSW has also been applied for dissimilar materials. I. Shigematsu et al. [54] used FSW process to join 5083 and 6061 aluminum alloys successfully. Thus FSW has the potential for joining dissimilar materials such as different types of aluminum alloys. Welding properties, such as the hardness distribution and the tensile strength, were strongly influenced by the material combination. W H Jiang and R Kovacevic [66] demonstrated that friction stir welding (FSW) is a feasible route for joining 6061 aluminum (Al) alloy to AISI 1018 steel. The weld has a good weld quality and is free of cracks and porosity. Thaiping Chen and, Wei-Bang Lin [59] conducted a study on FSW for joining dissimilar materials. They attempted to join AA6061 aluminum alloy and SS400 low carbon steel, by friction stir welding (FSW). The qualities of dissimilar metals butt joints are reportedly evaluated by their ultimately tensile strength and impact value. Muhamad Tehyo et al. [40] investigated the effect of welding parameters on the microstructure and mechanical properties of friction stir welded butt joints of dissimilar aluminum alloy sheets between Semi-Solid Metal (SSM) 356 and AA 6061-T651 by a Computerized Numerical Control (CNC) machine. They used base materials of SSM 356 and AA 6061-T651. Authors reported that at higher rpm higher tensile strength may be achieved. Yuhua Chen et al. [67] applied FSW process on Titanium alloy TC1 and Aluminum alloy LF6 to butt joint and lap joint by friction stir welding (FSW), and the influence of process parameters on formation of weld surface, cross-section morphology and strength were studied. The results show that, Titanium and Aluminum dissimilar alloy is difficult to be butt joined by FSW, and some defects such as cracks and grooves are easy to occur. When the tool rotation rate is 950 rpm/min and the welding speed is 118 mm/min, the tensile strength of the butt joint is 131MPa which is the highest. R. Rai et al. [48], in a technical report reviewed and critically examine several important aspects of FSW tools such as tool material selection, geometry and load bearing ability, mechanisms of tool degradation and process economics. There is a need for concerted research efforts towards development of cost effective durable tools for commercial application of FSW to engineering alloys. Ratnesh Kumar Raj Singh et al. [51] analyzed the influence of the microstructures and mechanical properties of friction stir welded butt joint of 6101 aluminum alloy and pure copper plates in 3 mm thickness. With this aim, welds were produced using Tungsten Carbide tools, with a cylindrical pin tool having 5 mm and 20 mm diameter of pin and shoulder respectively. Authors further reported that Micro hardness in weld nugget is higher than base metal and no significant difference found in other regions. Tensile strength of weld is very poor as compare to both of the base metals and all welds were failing from nugget zone. The ductility of is also very poor and comparable to the base metals. S. Ciliberto et al. [4] studied the effect of welding parameters on the mechanical properties of friction stir-welded dissimilar aluminum alloys in overlap joint configuration. Mechanical properties were evaluated by means of both hoop stress and T-pull test. Results show that there is a great dependence of the hoop stress resistance from welding parameters; conversely, T-pull tests do not show the same dependence. Authors reported the suitability of the x in lap configuration for joining dissimilar aluminum alloys. Hamid Mohammadzadeh et al. [15] in their experimental study, a butt joint of FSW were modified by optimizing rotational speed to improve the joint strength. The formation of inter metallic compounds in 2017 Al and AZ31 Mg was investigated and discussed using the binary Al-Mg phase diagram. By changing the rotational speed from 500 to 1600 rpm metallic continuity was obtained but minimum inter metallic compound and maximum tensile strength were achieved at 500rpm.

Machines – tools – materials used for FSW process modeling and optimization

In these section different machines, tools, tool profiles, tool materials and the metals that were successfully joined together using FSW are presented. The various efforts and investigations of different researchers on different materials similar or dissimilar with respect to the machine tool and FSW tool and its profile are compared for their respective findings so as to ascertain the popular FSW machine tool, FSW tool profile employed for different similar and dissimilar materials to achieve desired mechanical properties of joined metals. The table 1 also presents a broader idea of material and profile of the FSW tool required for joining different metals for varied applications.

Table 1: Different machine(s), tools [tool profiles / tool materials], specimen materials used for FSW process modeling and optimization

Ref No.	Machine	Tool/profile	*Specimen Material*	Remarks
[64]	Custom made MTS friction stir welding machine	M2 -Tool of Steel	AA 2195-T8	It is observed that only 5% of the heat generated by the friction process flows to the tool and the rest flows to the workpiece. The ‘‘heat efficiency’’ in FSW is thus 95%.
[36]	SMARC vertical milling machine having 7.5 kW	X155CrMoV12– 1 cold work tool steel D = 12 L=2.9	Cu Brass	Different axial forces were experimented
[35]	Modified Milling Machine	H-13 steel tool heat treated to a hardness of *R*C 46–48. Tool Profile : FC L=6.35; D=8; a_s=7º	Al2024-T3	Used DAQ system to collect data
[16]	Improvised vertical milling machine 1HP	Mild steel tool D=15; D=5; L=4.5	Al6061	After FSP Hardness about 50 % increased. Grain size is reduced to 10 to 1 µm
[56]	Improvised vertical	Mild steel tool	Al6061	Compared with TIG and found

	milling machine 1HP	D=15; D=5; L=4.5		improved mechanical properties of FS weld specimen
[20]	Makino Milling Machine with mandrel power of 3.75 KW.	Tool Profile : FC D= 16/20/24/28 L= 3/4/6/8/10	Al 6061-T6	Launch Box manufactured
[21]	3-axis CNC milling machine	Tool Profile : CCT	AA6061	Instead of tilting the FSW tool, the plates are tilted using a supporting wedge underneath the plates.
[60]	Indigenously developed FSW Machine Tool speed up to 3000 rpm Axial load up to 30 KN Traverse speed up to 500 mm/min	M2 steel of 50 - 55 VHN D=15; D=5; L=4.5 Tool Profile : CT without threads	AA 5083	Confirmation test were also conducted
[49]	An indigenously developed CNC FSW 22, 6 T kW, 4 000 r/min,	Tool material: High carbon steel D=7.86/12/15/18/22 d= 2.62/4/5/6/7.37 Tool hardness 243 /450 / 600 / 750 / 956HRc	AA 6061-T6	Relationship between grain size and TS - Confirmation test
[50]	An indigenously developed CNC FSW 22, 6 –T, kW, 4000 rpm	High carbon steel	AA7075-T6	Effect of independent parameters presented very well
[45]	Milwaukee #2K Universal Milling Machine	H-13 tool steel Hardness RC 48-50. D= 0.625 in 0.250 in side-length Tool Profile: T_r L= 0.237 in a =1^o d_p= 0.0074	AA 6061-T6	It has been shown that preheating the workpiece results in a significant (43%) reduction in the axial force experienced by the tool for a wide range of welding parameter.
[57]	Mazak VQC- 15/40 vertical machining center	D=5, Pin L= 2.5 Tool Profile : Threaded pin with three flats that tapers from 3-2 mm.	Al 7075	Dynamic force model
[28]	An indigenously designed CNC FSSW machine (4000 rpm, 22 kW, 60 kN)	Tool Profile : FC D=16.2 a pitch metric of 0.8 mm, and LH threaded; d= 5.4	AA 7075	Sensitivity analysis carried out
[22]	HERMLE 3-axis CNC milling machine	Tool Profile : CT Threaded M5/ Flat pin end Tool Profile : SS a = 0◦; d= 5; L= 2.9; D= 15 Tool material Z40CDV 5.150 HRC	AA2024	Conducted 3 test cases

[41]	Friction stir welding machine made by R.V.S machine tool	Tool material HCHCr Tool Profile : SS_q without draft, D= 18, d= 6 L=5.6	AA5083-H111	Achieved maximum tensile strength at 1000 r/min
[41]	CNC Turning center and WEDM.	Tool material HCHCr steel Tool Profiles: SS_q / TS_q / SH / SO/TO	AA6351-T6 AA5086-H111	The working ranges of optimized welding parameters are reported.
[39]	Vertical milling machine	Tool material: SS310 alloy Tool Profiles: FC/CT/ T_p; d=5	AA1100	Tensile strength, percentage of elongation and nugget hardness increase with increasing welding speed.
[2]	FSP machine	Pin less tool used to cover top grooves after filling SiC particles during FSP. Square pin tool size of 3.54×3.54 mm, L=2.5; D= 15.	AZ91	A model is developed that can predict hardness and grain size as functions of rotational and traverse speeds and region types. .
[53]	RM-1 model FSW machine	Tool Profile: SS d= 12.7; D=30.5; L= 6.2	AA 2219-T87	This investigation attempts to understand the true temperature at the workpiece/weld tool interface
[10]	Friction stir welding machine	FSW tool	AA6061-T6.	A model is developed that can be effectively used to predict the UTS,YS and TE
[45]	Vertical machining centre.	Tool Profile: SH Tool material: Cold work die steel L= 4.7 D=17	AA6082-T6 AA6061-T6,	AA6082 on advancing side; corrosion rate is higher with respect to increasing welding speed of the tool while corrosion rate decreased in case of AA6061 on advancing side.
[60]	Vertical milling machine with an in- house developed FSW setup	Stainless steel grade 310 tool D= 22/19/16 d= 8/7/6	AA7039	Pin diameter of tool played important role in Tensile strength
[12]	ESAB LEGIOTM 3UL FSW machine	FSP consumable rod for coating as a tool	AA 2024-T3 AA 6082-T6	AA 6082 was successfully deposited on AA 2024 by friction surfacing.
[33]	FSW machine (FSW-RL31-010, Beijing FSW Tech. Co., China).	Tool Profile: TT D= 18 d_1 7 mm d_2 =5 (Tip) L= 5	Al 7075-T651	Joining of dissimilar metal, done effectively.
[14]	Indigenously designed FSW equip- ment.	Tool Material: High-speed steel Tool Profile: CT D= 15 d= 5 L= 2.6.	Al-Si Al-Mg	Rotational speed of 1600 rpm was preferable during FSW of dissimilar Al alloys.

[51]	ATABRIZ/4301 (15 hp, 3,000 RPM) milling machine modified	Tool Profile: FC CT/ SS_q and Straight Triangular	Al 7075	Square pin profile, rotary speed of 1,400 RPM, welding speed 1.75 mm/s, and axial force of 7.5 KN results in the highest mechanical properties. ANFIS model and SA algorithm were used.
[11]	Vertical milling machine,	Tool material: H13 hot work steel d=4 D=12	Cu	Defect free friction stir welds of pure copper are observed.
[32]	FSW-3LM-003 NRSA-FSW	Tool Profile: CT d= 4.52; L=4.80; D= 5.94/9.8	AA 2219-T6	Some defects observed during FSW.
[6]	Commercial three-axis CNC mill (HAAS TM-1) welding fixture that are attached to a three-axis force dynamometer (Kistler model 9265A)	Tool material: H13 tool steel a_s =4.4° Tool Profile: T_c with three flats. D= 15; d= tapers from7.0 -5.2 L= 4.7	AA 6061-T6 AA7075-T6	Good weld are produced UTS of a negative offset, dissimilar AA6061 weld observed less weakening adjacent to the joint than the similar weld.
[61]	ESAB LEGIO 3UL NC machine	The probe is a 9^0 conical with 4 mm bottom diameter and left-handed threads along its length. The shoulder is a plane with two spiral striations scrolling an angle of 180^0 with inner and outer diameters of 5 and 16 mm.	AA2024-T351	Optimum condition within the selected parameters, values, according to the algebraic model developed. Taguchi method is suitable to optimize the mechanical behavior of the AA2024-T351 aeronautic aluminum alloy joints.

Note: *D is Shoulder diameter in mm; d is Pin diameter in mm; L is Pin length in mm; a is tilt angle in* ^{o}*;* a_s *is shoulder angle in* ^{o}*; FC is Flat Cylindrical; SS is Scrolled Shouldered;* SS_q *is Straight Square;* TS_q *is Tapered Square; SH is Straight Hexagon; SO is Straight Octagon; TO is Tapered Octagon; T is threaded;* T_r *is Trivex; CCT is Concave Cylindrical Threaded; CT is Cylindrical Taper;* T_p *is Trapezoidal; TT is Taper Threaded;* T_c *is Threaded Conical*

Modeling and optimization techniques used for FSW process optimization

In this section different modeling techniques employed to model FW processes experimented on different similar and dissimilar materials is presented along with the optimization methodologies that were adopted to predict optimal or suboptimal process parameters of FSW process. This section also presents the efforts of different researchers who studied the interaction between various independent parameters like Tool rotational speed in rpm, Feed rate in mm/min, Axial load /force in KN, Tool shoulder diameter in mm, Pin diameter in mm, Tool tilt angle, different tool materials with varied HRC etc and dependent process parameters such as Ultimate Tensile Strength (UTS), Yield Strength (YS), Hardness, Fatigue, Fracture, Grain size in µm, %

elongation etc, of FSW. The following tabulation of efforts of researchers to model and optimize similar and dissimilar metals is drawn with respect to the specimen specifications such as length, width, thickness of different specimens employed. Many researchers used Response Surface Methodology (RSM), Design of Experiments (DOE), Generalized Reduced Gradient (GRG), Artificial Neural Networks (ANN), Finite Element Method (FEM) etc for modeling and Taguchi, Genetic Algorithm (GA), Pattern Search Algorithm (PSA) for optimization of FSW processes.

Table 2: Modeling and optimization techniques adopted for FSW process with different process parameters on different materials of different specimen specifications

Ref No.	Material	Specimen specs.	Independent Parameters	Dependent Parameters	Modeling	Optimization
[65]	Al	Aluminum plates	N= 710/900/1400 f= 40/70/140	% EL Tensile Stress Bending stress	ANN	-
[34]	Al2024-T3	100mm×100mm×20mm	f =0.45 mm/sec F= 0 -25 N=300	Tool wear	FEM / ABAQUS	-
[30]	Al RDE-40 ≈ AA7039	300 x 150 x 6 mm	N=1200 – 1600 f= 22 – 75 F= 4 – 8	TS	-	Taguchi
[44]	Acrylic sheets ASTM D 638	25×100×6 mm	Vibrating Tool frequency 8-20 Hz f=12-24 CF= 4 – 20	TS	Orthogonal Array ANN	Taguchi GA
[19]	AA6061	6.3-mm thick	a= 2.25^o; d_p= 0.25 f= 30 -210 N= 31.4 – 125.6 rad/sec	YS	ANSYS FEA/ 3-D	-
[24]	A319	175 x 75 x 6 mm	N= 1000 - 1400 f= 22 – 75 F= 2 -4	TS	Regression	Taguchi
[31]	AA7039	300 x 100 x 6 mm	N= 1200 - 1600 f= 22 – 75;F= 4 - 8	TS; UTS; % EL; H	RSM ANN	-
[62]	AA 5083	100 x 10 x 5 mm	N= 500 - 800 f=115 – 155 ; F= 9 -17	YS; UTS; % EL	-	Taguchi
[48]	AA 6061-T6	300 x 150 x 5 mm	N= 924 - 1875 f= 12.43 – 107.56 F= 5.62 – 10.37	GS; TS	RSM	-
[20]	Al 1080	150 x 50 x 5 mm	N= 500–1500 f= 6.25–20 F= 20	TS 75–115 YS 45–90 %EL 3–12	ANN	GA
[25]	AA 2219-T87	250×50×5 mm	N= 250-400 f= 60-180	Axial force UTS	RSM	-
[26]	AA2024-T3	75 x 25 x 2.7mm	N = 600/800/1000/1200/1400	S_t FL TS	RSM	

			f= 4/8/12/16/20 d_p= 4.4/4.6/4.8/5.0/5.2 d_t=3/4/5/6/7			
[18]	Al 2024	250 x 150 x 4 mm	N= 800 -1100 rpm f= 21 Thermocouples at different distances	3D Thermal Model	FEA COSMOL	-
[46]	AA 6061-T6	9 x 3 x 0.25 mm	f= 5/5.5/7/8 in/min Initial welding temp 22 º /50 º /75 º /100 º /150 º /200 º /250 º /300 º F= 1000-6000 N= 2000	YS	ANSYS-FLUENT FEM	-
[57]	Al 7075	76 x 101x 6.35 mm	N= 2000/2250/2500 f= 120/180/240 d_p= 0.02/0.11/0.2	Force	RSM	-
[1]	AA6061	-	N= 900/1200/ 1500 f= 1.25; F= 30 Density (kg m_3) 2700	YS UTS	3D heat transfer and visco-plastic flow model	-
[5]	AA 2198 T3	200×100x3.2 mm	N=500/700 /900 f=150/ 225/300	YS UTS	RSM	DOE
[37]	Al 6061	-	N= 315 - 385 f= 1.55 – 1.9 F= 9.9 – 12.1	HAZ distance	COMSOL	Taguchi
[27]	AA 7075 rolled sheet	2.7-mm	N =500 /600 700/ 800 /900 f= 8 /10/12/14/ 16 d_p=4.4/4.6/4.8/5.0/5.2 t_d = 3/ 4 /5 /6/ 7	TS;FL	RSM	Hooke–Jeeves pattern search method
[21]	AA2024 -T351 temper state	70x 600x3.2 mm	N=400-800 f=100- 400 3D sensor (Kistler 912 4A) Thermocouples	HAZ Thermo-mechanical model	Simulation Exp.	-
[40]	AA6351	100 x 50 x 6 mm	N= 600-1500 f= 0.45 - 2.4 F=1- 2	UTS; YS; % EL	RSM	DOE
[63]	AA 5083	-	N=650; f=115; F=9	UTS	GRA	Taguchi
[8]	AA6061 ZrB2	100 x 50 x 6 mm	N=1000/1075/1150/1225/1300 f = 30/40/50/60/70 F=4/5/6/7/8 ZrB_2.= 0.2/2.5/5/7.5/10	UTS/ductility/wear resistance	RSM	GRG
[8]	AA6061-T6 AA7075-T6	Square Butt welds; 6-mm thickness to L	F=6/8/10 f=30/60/90. N=800/1000/1200	UTS YS	RSM CCD	DOE

		L=100				
[42]	AA6351-T6 and AA5086-H111	100 mm X 50 mm X 6 mm The	N=950; f= 63; F=1.5	Wear resistance	RSM	DOE
[38]	AA1100	6 mm thickness plate	N=710/1000/1400 f= 80/160/212 F=3.5	UTS; %EL; H	RSM	DOE
[25]	AA (6061)-B_4C	100 mm x 50 mm x 6 mm	N=800/900//1000/1100/1200; f= 3/.8/1.3/1.8/2.3 F= 6-14	UTS	GRG	DOE
[2]	AZ91	Square pin with dimensions of 3.54×3.54 mm L= 2.5	N= 710-1400 f= 12.5-63; a=3°.	H; GS	ANN	DOE
[10]	AA6061-T6	(100 x 50 mm) Butt joint	F= 6/8/10 N= 800/1000/1200 f= 30/60/90	UTS, Y,S %EL	CCD RSM	DOE
[47]	EN-GJS-400-15	ductile iron	N= 1125	UTS	RSM	GA
[13]	Al-Si Al-Mg	100 x 930 x 93 x mm	a=3^0; N=1000; f=70/80/130/190/ 240	TS	Experimental	DOE
[51]	AA 7075	200 x100 x 6 mm	N= 1400; f= 1.75 F=7.5;N=800/1000/1200/ 1400/1600 f= 0.25/0.75/1.25/ 1.75/2.25 ; F=5-9	TS YS H	CCD ANFIS	Optimization
[17]	Cu	4 -mm thickness	N= 563/700/900/ 1100/1236 f=33/50/75/100/117 F=1.16/1.5/2/2.5/2.8	TS/UTS/H	CCRD RSM	Optimization
[32]	AA 2219-T6	5.0- mm thick	f= 100 mm/min, N=600/700/800/900/1000	TS; H	Experimental study	-
[43]	AA6351 T6−AA5083 H111	100 x50 x6 mm	N = 600/775/950/ 1125/1300 f= 36/49.5/63/76.5/ 90 F=9.8/12.25/14.7/ 17.15/19.6	UTS	RSM	Optimization
[59]	AA2024-T351	200 x 145 x 4 mm	Variable parameters F=850/900/950 f= 120 250 500; t_d =8; a= 0^0 ; N= 1000 Plunge speed 0.1 mm/s;	UTS; %EL	CCD	Taguchi method

Note: N is Tool rotational speed in rpm; f is feed rate or welding rate in mm/min; F is axial force in KN; a is Tilt Angle in ^{o}; t_d is Dwell Time in seconds; CF is Clamping Force in Mpa; d_p is Plunge depth in mm;

TS is Tensile strength; UTS is Ultimate Tensile Strength; %EL is % of Elongation; H is Hardness; S_t is Shear Tensile; FL is Fracture Load; GS is Grain Size

Conclusions on literature review of FSW

This review of FSW literature reveals that there are number of studies conducted on FSW process on different similar and dissimilar Al alloys, Cu, Mg, steel, titanium. The review also suggest that they employed different FSW tools (such as Scrolled Shouldered, Straight Square, Tapered Square, Straight Hexagon, Straight Octagon, Tapered Octagon, threaded, without threaded, Trivex, Concave Cylindrical Threaded, Cylindrical Taper, Trapezoidal, Taper Threaded, Threaded Conical etc.) with variety of independent parameters (such as Tool rotational speed in rpm, feed rate or welding rate in mm/min, axial force in KN, Tilt Angle in ^{o}, Dwell Time in seconds, Clamping Force in Mpa, Plunge depth in mm etc.) for the want of different dependent parameters (such as Tensile strength, Ultimate Tensile Strength, % of Elongation, Hardness, Shear Tensile, Fracture Load, Grain Size etc.).

It is also a key point to note that, during FSW process the material flow is very complex but quite crucial that needs to be further understood. From this review it is clear that complete understanding of material movement in the region of rotating tool is vital to the optimization of FSW process parameters and design of tool geometry. Therefore, it is required to develop suitable tool geometries and come up with more robust FSW models of material flow for its eventual optimization. And thus, it may be necessary to optimize FSW process parameters and geometry simultaneously, will be an imperative to improve in welding properties and productivity. Shape and material of FSW tool are also important factors that govern the welding and metallurgical properties of joining metals. Tool wear is generally not considered as a severe issue in FSW of Al alloys. But for welding of high melting point materials such as steel and titanium and wearable materials like metal matrix composites using FSW tool wear becomes a critical issue that needs to be addressed. But, a few studies on the FSW tool wear have been reported.

This review suggest that besides some rudimentary or intuitive tool designs for FSW there are a necessity of computational models and algorithms that may provide suboptimal or near optimal design parameters of FSW tools. The appropriate selection of FSW tool material is also a key concern that may be required to address, more particular there is a need of comparative studies on steel, titanium, and their composites for joining different metals using FSW. From the review of the literature of FSW process it appears that, although few studies report the effect of preheating and post weld heat treatment on the development of microstructure, grain size, mechanical properties of the joint, it required to conduct further experimental studies independently for more number of similar and dissimilar metals to form a generalized opinion on the pre and post heat treatment of joinable metals. Form the above presented review of process modeling and optimization of FSW, it is evident that many researchers applied response surface methodology, to study the complex interaction of different independent process parameters of FSW to join similar and dissimilar metals, but there are no studies that report the application systems engineering concepts such as meta-modeling techniques for FSW process.

As regard to the optimization FSW process largely researchers applied only Taguchi technique. Only a few studies are reported that present the effectiveness of the Nature Inspired Algorithms (NIAs). Therefore, the need of the integration meta-modeling techniques and meta-heuristics such as NIAs is an area for research so as to develop robust and integrated models which may be appropriately optimized using NIAs for different highly complex, nonlinear and dynamic FSW process environments. More particularly, as the complex phenomenon of interplay among various parameters that govern and effect the material flow during the FSW and the impact of the dynamic variation soft e independent process parameters on various dependent parameters such as UTS, YS, microstructure, grain size etc. This literature review outlines that the seamless integration of NIAs would be a great benefit to predict optimal ranges of various dynamic parameters to achieve excellent mechanical properties of the friction stir welding.

Acknowledgement

Authors gratefully acknowledge the inspiration and guidance of Most Revered Prof. P. S. Satsangi, the Chairman, Advisory Committee on Education, Dayalbagh, Agra, India.

References

[1] Arora, A., Deb, A., Deb Roy, T.: Toward optimum friction stir welding tool shoulder diameter, Scripta Materialia, 2011, 64, pp. 9–12. doi:10.1016/j.scriptamat.2010.08.052

[2] Asadi P. &et.al.: Predicting the grain size and hardness of AZ91/SiC nano-composite by artificial neural networks, Int J Adv ManufTechnol DOI 10.1007/s00170-012-3972-z

[3] Blignault C.., Hattingh D.G., and James M.N.: Optimizing of Friction Stir Welding via Statistical Design of Tool Geometry and Process Parameters

[4] Ciliberto, S., Astarit, A, Squillace, A.: FSW of T joints in overlap configuration: process optimization in joining dissimilar aluminium alloys for the aeronautic application, Surf. Interface Anal. 2013, 45, pp. 1631–1637, DOI 10.1002/sia.5214

[5] CiroBitondo, Umberto Prisco, AntoninoSquilace, Pasquale Buonadonna, Gennaro Dionoro: Friction-stir welding of AA 2198 butt joints: mechanical characterization of the process and of the welds through DOE analysis, Int J AdvManufTechnol, 2011, 53, pp. 505–516. DOI 10.1007/s00170-010-2879-9

[6] Cole E. G. & Fehrenbacher A. & Duffie N. A. & Zinn M. R. & Pfefferkorn F. E. & Ferrier N. J.: Weld temperature effects during friction stir welding of dissimilar aluminum alloys 6061-t6 and 7075-t6 Int J AdvManufTechnol, DOI 10.1007/s00170-013-5485-9

[7] Dawes, C.J.; Thomas W.M.: Friction stir process welds aluminum alloys. Welding Journal, 1996, 75(3), pp.41 – 45.

[8] Dinaharan,I., Murugan N.: Optimization of Friction Stir Welding Process to Maximize Tensile Strength of AA6061/ZrB2 In-Situ Composite Butt Joints, Procedia Engineering 38 (2012) 105 – 110, doi: 10.1016/j.proeng.2012.06.015

[9] Elatharasan G.,Kumar V.S. Senthil: Modeling and Optimization of friction stir welding parameters for dissimilar aluminium alloys using RSM Procedia Engineering 38 (2012) 3477 - 3481International Conference on Modeling, Optimization and Computing, doi: 10.1016/j.proeng.2012.06.401

[10] Elatharasan G., Kumar V.S. Senthil.: An experimental analysis and optimization of process parameter on friction stir welding of AA 6061-T6 aluminum alloy using RSM, Procedia Engineering 64 (2013) 1227 – 1234 International Conference On Design and manufacturing, IConDM 2013 doi: 10.1016/j.proeng.2013.09.202

[11] Farrokhi H., Heidarzadeh A. and Saeid T.: Frictions stir welding of copper under different welding parameters and media, Science and Technology of Welding and Joining 2013 VOL 18 NO 8, 697-702 DOI 10.1179/1362171813Y.0000000148

[12] Gandra, J., Pereira, D., Miranda, R. M., Vilaça P.: Influence of process parameters in the friction surfacing of AA 6082-T6 over AA 2024-T3, Forty Sixth CIRP Conference on Manufacturing Systems 2013,Procedia CIRP 7 (2013) 341 – 346 doi: 10.1016/j.procir.2013.05.058

[13] Ghosh M., Gupta R.K., and Husain M. M.: Friction Stir Welding of Stainless Steel to Al Alloy: Effect of

Thermal Condition on Weld Nugget Microstructure, Metallurgical and materials transactions A, DOI: 10.1007/s11661-013-2036-9
[14] Ghosh, M., Husain Md. M., Kumar K., and Kaila S.V.: Friction Stir-Welded Dissimilar Aluminum Alloys: Microstructure, Mechanical Properties, and Physical States, JMEPEG (2013) 22:3890–3901 DOI: 10.1007/s11665-013-0663-3
[15] Hamid Mohammadzadeh1, Mohammad Ali Hajibagheri: Influence of Rotation Speed on Microstructure and Mechanical Properties of Dissimilar FSW Between Al2017 and Mg AZ31, Journal of Applied Science and Engineering Management, 2013, 8-17.
[16] Hans Raj, K., Rahul Swarup Sharma, Pritam Singh, Atul Dayal: Study of friction stir processing (FSP) and high pressure torsion (HPT) and their effect on mechanical properties, Procedia Engineering 10 (2011) 2904–2910, doi:10.1016/j.proeng.2011.04.482
[17] HeidarzadehA., SaiedT., KhodaverdizadehH., Mahmoudi,A.., and Nazari, E.: Establishing a Mathematical Model to Predict the Tensile Strength of Friction Stir Welded Pure Copper Joints, Metallurgical and materials transactions , Vol. 44B, 2013,175-183 DOI: 10.1007/s11663-012-9755-y
[18] Hilgert, J., Schmidt, H.N.B., dos Santos, J.F., Huber, N.: Thermal Models for Bobbin Tool Friction Stir Welding, Journal of Materials Processing Technology, 2010, pp. 1-16. DOI: 10.1016/j.jmatprotec.2010.09.006
[19] HoseinAtharifar, Dechao Lin, and Radovan Kovacevic: Numerical and Experimental Investigations on the Loads Carried by the Tool During Friction Stir Welding, Journal of Materials Engineering and Performance, 2009, 18 (4), pp. 339–350. DOI: 10.1007/s11665-008-9298-1.
[20] Hsing-Ta Hsieh, Jahau Lewis Chen: Influence of Welding Parameters on Mechanical Properties of Friction Stir Welded 6061-T6 Launch Box, Materials Transactions, 2008, 49 (10), pp. 2179 to 2184. DOI:10.2320/matertrans.L-MRA2008829.
[21] Ibrahim N. Tansel, Mustafa Demetgul, HasanOkuyucu, AhmetYapici: Optimizations of friction stir welding of aluminum alloy by using genetically optimized neural network, Int J AdvManufTechnol, 2010, 48:95–101. DOI 10.1007/s00170-009-2266-6
[22] Jacquina, D., Meesterb, de B., A. Simar, A., Deloison, D., Montheillet, F., Desrayaud, C. : A simple Eulerianthermomechanical modeling of friction stir welding, Journal of Materials Processing Technology 211 (2011) 57–65. doi:10.1016/j.jmatprotec.2010.08.016
[23] Jata, K.V., Semiatin, S.L., Scripta Mater. Vol. 43, 2000, p. 743.
[24] Jayaraman, M., Sivasubramanian, R., Balasubramanian, V., Lakshminarayanan, A.K.: Optimization of process parameters for friction stir welding of cast aluminum alloy A319 by Taguchi method, Journal of Scientific & Industrial Research, 2009, 68, pp. 36-43.
[25] Kalaiselvan K. and Murugan N. :Optimizations of friction stir welding process parameters for the welding of A1-B_4C composite plates using generalized reduced gradient method International Conference on Modeling, Optimization and Computing (ICMOC 2012)Procedia Engineering 38 (2012) 49 - 55doi:10.1016/j.proeng.2012.06.008
[26] Kanwer S. Arora, Sunil Pandey, Michael Schaper, Rajneesh Kumar: Effect of process parameters on friction stir welding of aluminum alloy 2219-T87, Int J AdvManufTechnol, 2010, 50, pp. 941–952. DOI 10.1007/s00170-010-2560-3
[27] Karthikeyan, R., Balasubramanian, V.: Predictions of the optimized friction stir spot welding process parameters for joining AA2024 aluminum alloy using RSM, Int J AdvManufTechnol, 2010, 51, pp.173–183. DOI 10.1007/s00170-010-2618-2
[28] Karthikeyan, R., Balasubramanian, V. : Statistical Optimization and Sensitivity Analysis of Friction Stir Spot Welding Process Parameters for Joining AA 7075 Aluminum Alloy, Experimental Techniques, 2011, DOI:10.1111/j.1747- 1567.2011.00746.x
[29] Lakshminarayanan, A. K.; Balasubramanian, V., Elangovan K.: Effect of welding processes on tensile properties of AA6061 aluminium alloy joints. Springer, 2007, pp. 286-296.
[30] Lakshminarayan, A.K., Balasuramanian, V.: Process parameters optimization for friction stir welding of RDE-40 aluminium alloy using Taguchi technique, Trans. Nonferrous Met. Soc. China, 2008, 18, 548-554.
[31] Lakshminarayan, A.K., Balasuramanian, V.: Comparison of RSM with ANN in predicting tensile strength of friction stir welded AA7039 aluminum alloy joints, Trans. Nonferrous Met. Soc. China, 2009, 19, pp. 9-18.
[32] Li ,J. Q. & Liu H. J.: Design of tool system for the external non-rotational shoulder assisted friction stir

welding and its experimental validations on 2219-T6 aluminum alloy Int J AdvManufTechnol (2013) 66:623–634 DOI 10.1007/s00170-012-4353-3
[33] Li W.-Y. & Li J.-F. & Zhang Z.-H. &Gao D.-L.& Chao Y.-J.: Metal Flow during Friction Stir Welding of 7075-T651 Aluminum Alloy Experimental Mechanics (2013) 53:1573–1582 DOI 10.1007/s11340-013-9760-3
[34] London, B., Mahoney, M., Bingel, B., Calabrese, M., Waldron, D.: Proceedings of the Third International Symposium on Friction Stir Welding, Kobe, Japan, 27–28 September, 2001.
[35] Mandal, S., Rice, J., Elmustafa, A.A.: Experimental and numerical investigation of the plunge stage in friction stir welding, Journal of materials processing technology, 2008, 203, pp. 411–419. DOI:10.1016/j.jmatprotec.2007.10.067
[36] Meran, C., Kovan, V.: Microstructures and mechanical properties of friction stir welded dissimilar copper/brass joints, Mat.-wiss. u. Werkstofftech, 2008, 39 (8), pp. 521-530. DOI: 10.1002/mawe.200800278.
[37] Mishra, R.S., Ma, Z.Y.: Friction stir welding and processing, Materials Science and Engineering, 2005, R 50, pp. 1–78. doi:10.1016/j.mser.2005.07.001
[38] Mohamadreza Nourani, Abbas S. Milani, Spiro Yannacopoulos: Taguchi Optimization of Process Parameters in Friction Stir Welding of 6061 Aluminum Alloy: A Review and Case Study, Engineering, 2011, 3, pp. 144-155.DOI:10.4236/eng.2011.32017
[39] Mohanty, H. K., Venkateswarlu D., Mahapatra M. M., Kumar Pradeep, Mandal N. R.: Modeling the Effects of Tool Probe Geometries and Process Parameters on Friction Stirred Aluminum Welds Journal of Mechanical Engineering and Automation 2012, 2(4): 74-79 DOI: 10.5923/j.jmea.20120204.04
[40] Muhamad Tehyo, Prapas Muangjunburee, Somchai Chuchom: Friction stir welding of dissimilar joint between semi-solid metal 356 and AA 6061-T651 by computerized numerical control machine, Songklanakarin J. Sci. Technol. 33 (4), 441-448, Jul. - Aug. 2011
[41] Palanivel, R., Koshy Mathews, Murugan, N.: Development of mathematical model to predict the mechanical properties of friction stirwelded AA6351 aluminum alloy, Journal of Engineering Science and Technology Review, 2011, 4 (1) pp. 25-31. ISSN: 1791-2377
[42] Palanivel R., MathewsP.Koshy, Murugan N., Dinaharan.I: Prediction and optimization of wear resistance of friction stir welded dissimilar aluminum alloy, Procedia Engineering 38 (2012) 578-584 doi:10.1016/j.proeng.2012.06.072
[43] Palanivel, R, Mathews P. Koshy , Murugan N. : Optimization of process parameters to maximize tensile strength of friction stir welded dissimilar aluminum alloys using response surface methodology Optimization of process parameters to maximize ultimate tensile strength of, J. Cent. South Univ. (2013) 20: 2929−2938 DOI: 10.1007/s1177101318151
[44] Panneerselvam, K., Aravindan, S., Noorul Haq, A.: Hybrid of ANN with genetic algorithm for optimization of frictional vibration joining process of plastics, Int J AdvManufTechnol, 2009, 42:669–677. DOI 10.1007/s00170-008-1641-z
[45] Patil,H., Soman S.: Effect of weld parameter on mechanical and metallurgical properties of dissimilar joints AA6082–AA6061 in T6 condition produced by FSW, S. N. FratturaedIntegritàStrutturale, 24 (2013) 151-160; DOI: 10.3221/IGF-ESIS.24.16
[46] Paul C. Sinclair , William R. Longhurst , Chase D. Cox , David H. Lammlein , Alvin M. Strauss, George E. Cook, Heated Friction Stir Welding: An Experimental and Theoretical Investigation into How Preheating InfluencesProcess Forces, Materials and Manufacturing Processes, 2010, 25:11, pp. 1283-1291. DOI: 10.1080/10426914.2010.496122
[47] Radosław Winiczenko, Robert Salat, Michał Awtoniuk: Estimation of tensile strength of ductile iron friction welded joints using hybrid intelligent methods, Trans. Nonferrous Met. Soc. China 23(2013) 385−391 DOI: 10.1016/S1003-6326(13)62474-7
[48] R. Rai, A. De, H. K. D. H. Bhadeshia and T. DebRoy, Review: friction stir welding tools, Science and Technology of Welding and Joining, 2011, VOL 16, No. 4, pp. 325-342.
[49] Rajakumar, S., Muralidharan, C., Balasubramanian, V.: Establishing empirical relationships to predict grain size and tensile strength of friction stir welded AA 6061-T6 aluminium alloy joints, Trans. Nonferrous Met. Soc. China, 2010, 20, pp. 1863-1872. DOI: 10.1016/S1003-6326(09)60387-3
[50] Rajakumar S., Muralidharan, C., Balasubramanian, V.: Influence of friction stir ss and tool parameters on strength properties of AA7075-T6 aluminium alloy joints. J Mater Design, 2010,

doi:10.1016/j.matdes.2010.08.025
[51] Ratnesh Kumar Raj Singh, Rajesh Prasad, Sunil Pandey: Mechanical properties of Friction Stir Welded dissimilar metals, Proceedings of the National Conference on Trends and Advances in Mechanical Engineering, YMCA University of Science & Technology, Faridabad, Haryana, Oct 19-20, 2012, pp. 579-583.
[52] Roshan S. Babajanzade & Jooibari M. Behboodi &Teimouri R., Asgharzadeh G. Ahmadi, M. Falahati-Naghibi, H. Sohrabpoor, S. Babajanzadex, M. Behboodi Jooibari, R. Teimouri, G. Asgharzadeh-Ahmadi, M. Falahati-Naghibi, H. Sohrabpoor: Optimization of friction stir welding process of AA7075 aluminum alloy to achieve desirable mechanical properties using ANFIS models and simulated annealing algorithm, Int J AdvManufTechnol (2013) 69:1803–1818, DOI 10.1007/s00170-013-5131-6
[53] Schneider, J, Stromberg, R., Schilling, P., Cao, B., W. Zhou, W., J. Morea, J., O. Myres, O.: Processing Effects on the Friction Stir Weld Stir Zone, Welding Technology Journal,January 2013, Vol. 92,pp..11-19
[54] Shigematsuy, I. Y. Kwon, J, J K. Suzuki, J. K., T. Imai, T., N. Saito, N.: Joining of 5083 and 6061 aluminum alloys by friction stir welding, Journal of Materials Science Letters, 22, 2003, 353– 356.
[55] Suresha C. N., Rajaprakash B. M., Sarala Upadhya: A Study of the Effect of Tool Pin Profiles on Tensile Strength of Welded Joints Produced Using Friction Stir Welding Process, Materials and Manufacturing Processes, (2011), 26(9):1111-1116.
[56] Suri, Atul., Prasanth, R.S.S., Hans Raj, K.: Comparative study of friction stir welding and tungsten inert gas welding of pure aluminum, International Conference on Energy Efficient Technologies for Sustainability (ICEETS), pp.:929-935, DOI: 10.1109/ICEETS.2013.6533512. ISBN: 978-1-4673-6149-1
[57] Thaiping Chen and ,Wei-Bang Lin: A Prime Study On FSW Joint Of Dissimilar Metals, Proceedings of the XIth International Congress and Exposition June 2-5, 2008 Orlando, Florida USA ©2008 Society for Experimental Mechanics Inc.
[58] Thomas, W.M., Nicholas, E.D, Needham, J.C. Murch, M.G.: Temple- Smith P. Dawes C.J., G. B. Patent, 1991, Application No. 9125978.8.
[59] Tyler A. Davis, Yung C. Shin,Bin Yao: Observer-Based Adaptive Robust Control of Friction Stir WeldingAxial Force, 2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics Montréal, Canada, July 6-9, 2010, pp. 1162 - 1167
[60] Venkateswarlu, D., Mandal N. R., Mahapatra M. M., and Harsh. P., Tool Design Effects for FSW of AA7039[Gan13]Gandraa J.,Pereirab D., Mirandab R. M., Vilaça P.: Influence of process parameters in the friction surfacing of AA 6082-T6 over AA 2024-T3, Forty Sixth CIRP Conference on Manufacturing Systems 2013,Procedia CIRP 7 (2013) 341 – 346 doi: 10.1016/j.procir.2013.05.058
[61] Vidal C. and Infante V., Optimization of FS Welding Parameters for Improving Mechanical Behavior of AA2024-T351 Joints Based on Taguchi Method, JMEPEG (2013) 22:2261–2270 DOI: 10.1007/s11665-013-0499-x
[62] Vijayan, S., Raju, R., Subbaiah, K., Sridhar, N., Rao, S.R.K.: Friction stir welding of Al-Mg alloy optimization of process parameters using Taguchi method, Experimental techniques, 2010, pp. 38 -44.DOI: 10.1111/j.1747-1567.2009.00563.x
[63] Vijayan Sundaravel, Raju R.& Rao S. R. K.: Multi-objective Optimization of Friction Stir Welding Process Parameters on Aluminum Alloy AA 5083 Using Taguchi-Based Grey Relation Analysis
[64] W. H. Jiang, R. Kovacevic: Feasibility study of friction stir welding of 6061-T6 aluminum alloy with AISI 1018 steel, Proc. Instn Mech. Engrs Vol. 218 Part B: J. Engineering Manufacture, 2004, pp. 1323-1331.
[65] Yousif, Y. K., Daws, K. M. , Kazem, B. I. , Prediction of Friction Stir Welding Characteristic Using Neural Network, Jordan Journal of Mechanical and Industrial Engineering, 2008, 2(3), pp.151-155. ISSN 1995-6665.
[66] Yuh J. Chao, X. Qi W. Tang, Heat Transfer in Friction Stir Welding—Experimental and Numerical Studies, Transactions of the ASME, 2003, Vol. 125. DOI: 10.1115/1.1537741
[67] Yuhua Chen, Changhua Liu and Geping Liu, Study on the Joining of Titanium and Aluminum Dissimilar Alloys by Friction Stir Welding, The Open Materials Science Journal, 2011, *5*, 256-261

Improving the Issues in Procurement Process of Manufacturer

Prof S K Sharma[1], Mr. Rajnish Kumar[2], Ms. Anubha Rautela[3]

Department of Mechanical Engineering, IIT (BHU)-Varanasi, Uttar Pradesh

India- 221005

e-mail- sksharma.mec@itbhu.ac.in[1] , rajnishkumar1971@gmail.com[2]
arautela.rs.mec12@itbhu.ac.in[3]

Abstract

Today's manufacturing scenario has shifted focus from complete in house manufacturing to collaborative production where major share of manufacturing is outsourced and only core activities are handled exclusively by a Locomotive Manufacturer henceforth referred to as the Original Equipment Manufacturer (OEM). The OEM is a typical organization where it has to deal with growing demands of outsourcing but has to follow strict rules of public procurement. Issues of material/item categorization, types of contracts and structural changes in system are critical. Presently all items are dealt with same kind of tender and procedure but it has been found by research that such a strategy is unscientific and unproductive. Even determination of optimal number of suppliers considering the probability of supply disruptions, affects the cost of operating a supplier and loss. In supply chain, evaluation of supplier's performance has a major role. The four salient criteria which are important in supplier evaluation and selection are quality, on-time delivery, price and service. The expected benefits are reduction in the cost of items by 2-5% by optimizing the number of suppliers and long term relationships. This will also ensure reduction in the potential loss due to supply disruptions, which will mean reduction in financial loss.

Keywords: Procurement, Original Equipment Manufacturer (OEM), Supply Disruptions

1.1 Introduction

The supply chain for a manufacturer is of paramount importance as the industry all over the world is turning to strategic outsourcing, and concentrating only on core activities. Such outsourcing of normal raw material, semi finished products make the process even more complex, since the manufacturer now depends totally on efficacy and efficiency of the supply chain. The Locomotive Manufacturer, henceforth referred to as the Original Equipment Manufacturer (OEM) is the only one of its kind in India. It manufactures locomotives primarily for the Government owned Railways. Three issues, the paper focused are first classification of items, second determining optimal number of suppliers and last one evaluation of new and existing suppliers. The most famous procurement strategy was formulated by Kraljic (1983), is used. In 1983, Kraljic introduced the seminal paper, "Purchasing must become supply management", the first comprehensive portfolio approach for the determination of a set of differentiated purchasing strategies. Kraljic's approach includes the construction of a portfolio matrix that classifies products on the basis of two dimensions: profit impact and supply risk (low and high). As optimal number of suppliers was never attempted at the OEM and this was an unexplored area for the OEM. Thus relevant research in this area was reviewed. Supplier Evaluation is a very contentious issue, mainly because the OEM is a government organization and has to follow all rules of public procurement of equal opportunity, fairness, objectivity and is subject to statutory audit as well as scrutiny from departmental vigilance wings.

1.2 Classification of Items

In this section there is classification of contract types and classifying material as per contract/sourcing strategy type. For each class of material a different procurement strategy has been proposed with qualification criteria for suppliers being suitably modified. The items have been classified in the Vendor/Supplier Directory on the basis of location and use. As far as the design of loco and production is concerned it is a scientific system, but from procurement point of view no distinction is made between items. ABC analysis for items classification has been used in the study to qualify certain items as strategic, critical which require special attention. Within each sub group there are several items which are critical and strategic or have been historically troublesome with respect to Quality and Supply.

Kraljic has defined four types of material, **Strategic items** (high profit impact, high supply risk), **Leverage items** (high profit impact, low supply risk), **Bottleneck items** (low profit impact, high supply risk), **Non-critical items** (low profit impact, low supply risk). In Kraljic's model, Firstly analysis of products and their classification in a matrix is done. Secondly, the supplier relationships are analyzed and thirdly, strategies are discussed to bridge the gap between current and desired supplier relationships. There are a few decision variables in categorization of items for defining their purchase strategy. The four variables that directly impact the classification of item are:

1. Disruptions in supply of the item
2. Failure Rate of the item
3. Price of the item
4. Complexity/Technological intensiveness of the item

The supplier base is not considered as a decision variable since in the case of OEM, this is still in a nascent stage and there are not enough suppliers for all items. In fact this categorization will be one of the tools used for determination of the optimal number of suppliers. All the mentioned four variables are to be compared in a matrix so that the item can be categorized for defining the purchase strategy. The aim of the exercise is to provide decision maker with least amount of discretion. As Kraljic's model has classified four items, in the typical case of the Locomotive manufacturer, Leverage is considered as development item. For OEM there is recommended categories of items, shown in table 1.1.

Table 1.1 Recommended categories of items for OEM

I	Strategic: *Complex, Cost*
II	Development: *Complex, Subject to supply disruptions*
III	Bottleneck : *Failure Rate, Subject to supply disruptions*
IV	Normal: *Cost, Failure Rate*

The four VARIABLES have been quantified as defined in Table 1.2. Here all variables/dimensions have been compared on a scale of 100. The disruptions in supply for a particular item are compared to total loss in days by all items in the decision period. The highest priced item in the organization is taken as the base for comparing all other items. The average price of the item is expressed as % of the highest priced item, thus placing it on a scale of 100.

Table 1.2 Definition and quantification of Variables for classifying items

Disruptions in supply	S	Number of loco days lost compared to total days lost in the Decision Period by all suppliers for the subject item, expressed as %
Failure Rate	F	% of total supply of items by all suppliers
Price	P	Average price of item expressed as % of highest priced item
Complexity/Technological intensiveness	C	Scale 0 – 100, in steps of 10, where 10 is simplest to 100 which is most complex

Since there are four variables the centroid can be plotted on the chart. The area in which the centroid falls, will determine the category of the item. The plotting is done as shown in Fig 1.1.

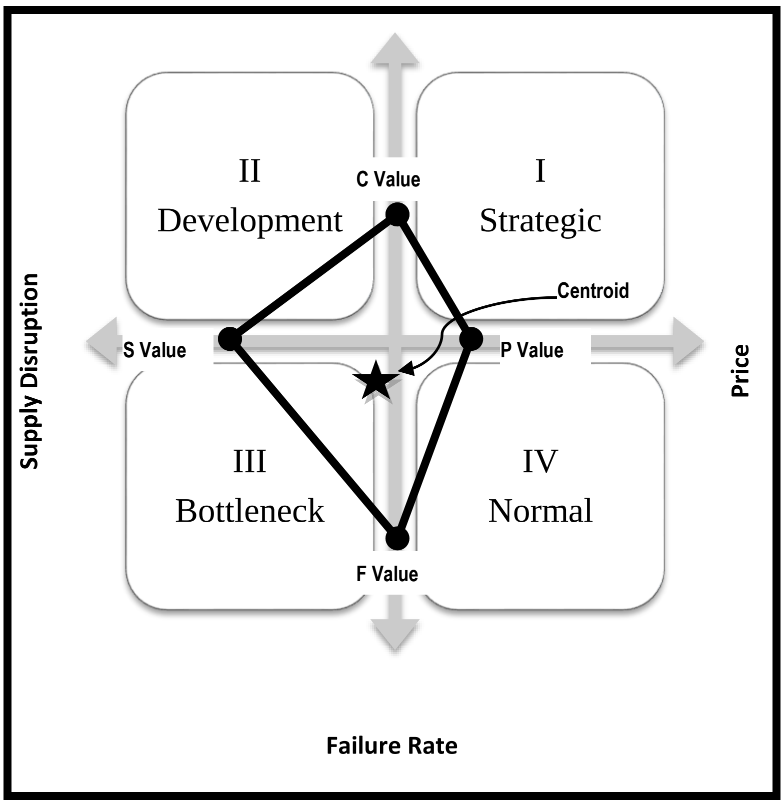


Figure 1.1 Classification Model

For the polygon that is formed by the four variables, the points are,$(x_0, 0), (0, y_1), (x_2, 0), (0, y_3)$
Using these values in the equation for calculation of centroid of polygon, we get

$$A = \frac{1}{2}[x_0 y_1 + x_2 y_1 + x_2 y_3 + x_0 y_3] \quad (1.1)$$

$$C_x = \frac{1}{6A}(y_1 + y_3)(x_0^2 - x_2^2) \quad (1.2)$$

$$C_y = \frac{1}{6A}(x_0 + x_2)(y_1^2 - y_3^2) \quad (1.3)$$

Here, C_x and C_y are coordinates of the centroid of the polygon and A is the area of the polygon

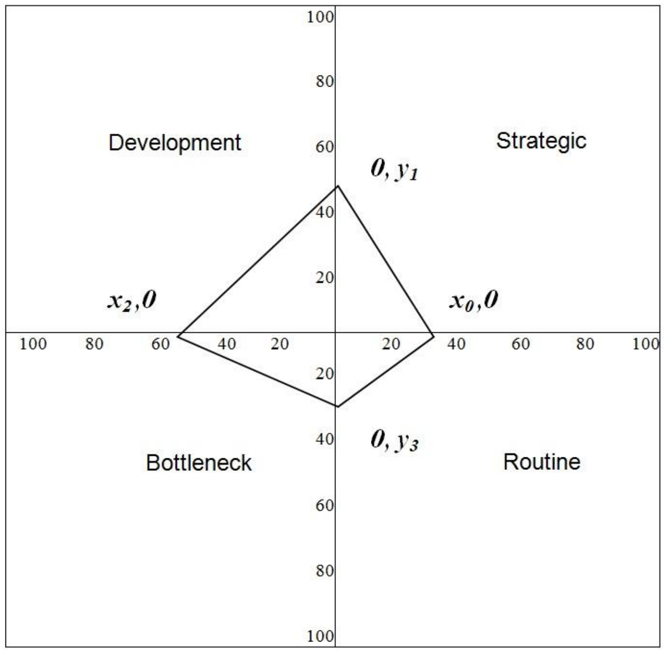


Figure 1.2 Graphical Representation of Classification of Items

Here the values of all variables are absolute and x_0 : P value, y_1: C value, x_2: S value, and y_3: F value. The centroid can be calculated and region in which it falls will define the item's category.

Table 1.3 Classification of Items according to model

ITEM	RATE in Rs	Total loss in Loco days(3 years)	Price P	Complexity C	Supply Disruption S	Failure Rate F	Category
			x_o	y_1	x_2	y_3	
AC-AC Traction System	25836115	2211	100	90	7.49	3	**I**
TCC	19061099	1035	74	90	3.51	6	**I**
Traction Motor	2671481	4390	10	90	14.87	15	**II**

The model has been applied to Locomotive Manufacturer, specifically to the items which caused disruptions in last three years 2009 to 2011 and items which figure in the traditional "A" category of the ABC analysis. . As examples few of the cases are discussed below, and Table1.3 depicts these sample items and their classification according to this model. Also the main problem is all items are procured with the same set of rules, and the procedure for negotiation of rates is subject to public scrutiny. Thus the main issue is devising an extremely transparent system of negotiation of rates.

Table 1.3 Proposed contract type for class of items

Class of items	Procurement scheme
Strategic Items	Incentive to supplier by having 3 year contracts and long term i.e. 10 year commitment.
Development Items	Request for Proposal must be floated, globally and locally. After RFP is floated and list of potential suppliers is received, a development tender will be floated to get the exact price and delivery schedule.
Bottleneck Items	Limited Tendering to approved suppliers, and the goal should be developing them to enter into long term agreements. For such items, supplier training and motivation is must.
Normal Items	Limited tenders to approved suppliers. This class of suppliers should be graduated to long term relationships.

1.3 Estimation of Optimal Number of Suppliers

There are a number of original equipment manufacturers (OEM) who have been traditionally manufacturing all the required components in house. This serious problem has, therefore, given way to outsourcing of major production activities instead of manufacturing all the items in house. This has led to the problem of having a number of suppliers for an efficient functioning of the organization. The improvement in supplier performance and supplier-manufacturer relationships has provided several opportunities for cost reduction and the performance of supply chain. It is therefore focusing on the determination of optimal number of suppliers by taking into consideration various factors including the probability of supply disruptions, both at the global as well as local level. The issues being faced by OEM are: few suppliers for the costly and complex items, no objectivity in supplier evaluation and the supply disruptions in case of low value items. In the present work we have attempted to develop a model by taking into consideration various factors outlined earlier. This model has then been applied to the Locomotive manufacturer to help it in the decision making process. It can be seen from Table 1.4 that there are products for which there are no vendors, which means in-house manufacture.

Table 1.4 Vendor Status at OEM

Type of Product	Vendors (V) Number of Vendors				Total Items
	V=0	V=1	V=2	V > 2	
HHP	191	1102	780	185	2258
OLDER	335	610	375	139	1459
HARDWARE	33	422	234	153	842
RAW MATERIAL	59	52	20	0	131
TOTAL	618	2186	1409	477	4690
% of TOTAL ITEMS	13%	47%	30%	10%	100%

The present model takes into consideration the events discussed by Berger *et al.* (2004). However, for the sake of simplicity, we take into account only two events which can disrupt the supply chain (i) Global event (ii) Local event. The Global events affect all suppliers whereas local events affect only some suppliers. A very important factor which must be taken into consideration is the quantity discount offered by suppliers for bulk orders. It is possible to express the quantity discount (*D*) as an exponential function:

$$D=A\left\{\theta e^{-\lambda(n-1)}\right\}$$

Where, $\boldsymbol{A}$ is the cost of item, $\boldsymbol{\theta}$ is a parameter for highest discount (estimated as 0.05 for present study), $\boldsymbol{\lambda}$ is a parameter for rate of decrease of discount. One very important aspect is to model the probability of the occurrence. For this we define the parameters, $\boldsymbol{P_g}$ the probability of Global event and $\boldsymbol{S_i}$ the probability of Local event (for supplier *i*). Further it is assumed that the probabilities of local events for the different suppliers are independent. The probabilities of global event and local event are also not dependent on each other. Analysis is first done by taking a case where the quantity of discounts offered by suppliers is not accounted for as adopted by Berger *et al.* (2004). The analysis has to be performed under the following conditions:

1. The number of suppliers is chosen from *i*= 1 to *n*.
2. There are two conditions, all suppliers fail, or some fail.
3. If all the suppliers fail, there is a loss to the company denoted by $\boldsymbol{L_t}$
4. The cost of operating a supplier, *i,* is denoted by ***C(i)***, *i*=1,2…*n*.

$$E(n+1)-E(n)=v+L_t\left(1-P_g\right)\left(S^{n+1}-S^n\right) \tag{1.1}$$

This is called the Model-B as developed by Berger *et al.*(2004). . The exponential function defined earlier for ***D*** will be introduced in Eq.1.1 to take care of quantity discount. The final equation (mathematical representation of the present model) is obtained as following:

$$E(n+1) - E(n) = -A\theta\left[e^{-\lambda n} - e^{-\lambda(n-1)}\right] + v + L_t\left(1 - P_g\right)(S^{n+1} - S^n) \qquad (1.2)$$

The procedure is to increase ***n,*** till ***E***(***n+1***) **-** ***E***(***n***) > 0, which would mean that it is costlier to operate (***n+1***) suppliers than ***n***. The number ***n*** thus obtained will be the optimal number of suppliers, ***n****. The calculations may be carried out either by assuming the values or by using empirical values.

We now apply the present model to OEM's case to arrive at the optimal number of suppliers. The parameters for the calculations to be used for Eq.1.2 have either been assumed or estimated on the basis of experience. Therefore the optimal number of suppliers, ***n**** will be calculated for a range of values of *A*. The discount offered by suppliers for 100 % order is of the order of 5% so the discount parameter θ, is estimated at 0.05. With each increase in supplier the discount reduces by ~ 1% to 1.5%. However, we have to determine a suitable value of parameter λ.

Using $D=A\left\{\theta e^{-\lambda(n-1)}\right\}$

Keeping $\boldsymbol{\theta}$= 0.05, and by varying $\boldsymbol{\lambda}$, the % discount will be estimated. The value of $\boldsymbol{\lambda}$ is assumed to be the one which best compares with practice.

Table 1.5 The value of percent discount for two values of λ

Θ	0.05		*ϑ*	0.05
Λ	0.6		*λ*	0.4
N	% discount		*n*	% discount
1	5.000		1	5.000
2	2.744		2	3.352
3	1.506		3	2.247
4	0.826		4	1.506
5	0.454		5	1.009

It is seen that $\boldsymbol{\lambda}$=0.4 gives % discount value very near to actual one. Hence $\boldsymbol{\lambda}$=0.4 is the estimated value to be used in further calculations. The calculations for the optimal number of supplier will be done for both, the present model as well as for Model-B.

1.3.1 Calculation of optimal number of suppliers by Model-B

The optimal values, **n*** are calculated by taking into consideration the following values for different parameters, $\mathbf{P}_g$, the probability of Global event is assumed to be 5%, $\mathbf{S}_i$, the probability of Local event for supplier i is 10% ,$\mathbf{L}_t$, the loss to company due to supplier's failure is being estimated as Rs 300,000 per day, ***v*** , the variable, is the cost of operating supplier and is estimated as Rs 2000. Therefore we have only ***v,*** $\boldsymbol{P_g}$, ***S*** and $\boldsymbol{L_t}$

V	P_g	*S*	L_t
2000	0.05	0.1	300000

Substituting the values of parameters in equation (1.1), we get optimal number of suppliers, ***n**** = 3. This value of ***n**** is the value of optimal number of suppliers corresponding to Model-B.

1.3.2 Calculation of optimal number of suppliers by present Model

The present model which has been developed here accounts for supplier's discount and will have a direct bearing on the optimal number. Eq (1.2) will be used to estimate the optimal number of suppliers, ***n****. The optimal number ***n**** will depend upon the value of ***A*** which will have a direct role to play in the present model. The value of λ, as estimated earlier, has been taken as 0.4. All other parameters are the same as used in Model-B.

Table 1.6 Optimal number of suppliers for various values of *A*

v	Θ	Λ	P_g	S	L_t
2000	0.05	0.4	0.05	0.1	300000
Result					
A, cost of item in Rs			***n****, optimal number of suppliers		
5000			3		
10000			3		
50000			3		
100000			2		
200000			2		
500000			2		
1000000			2		
5000000			1		
10000000			1		
25000000			1		

In case of items having price less than Rs 100,000, the optimal number of suppliers, ***n****=3. It is suggested that for these items three suppliers may be considered as the optimal number for the OEM in order to maintain proper supply of items and production. Although ***n**** =1 for higher values of ***A***, it may not be feasible, therefore, it is suggested, to have two suppliers.

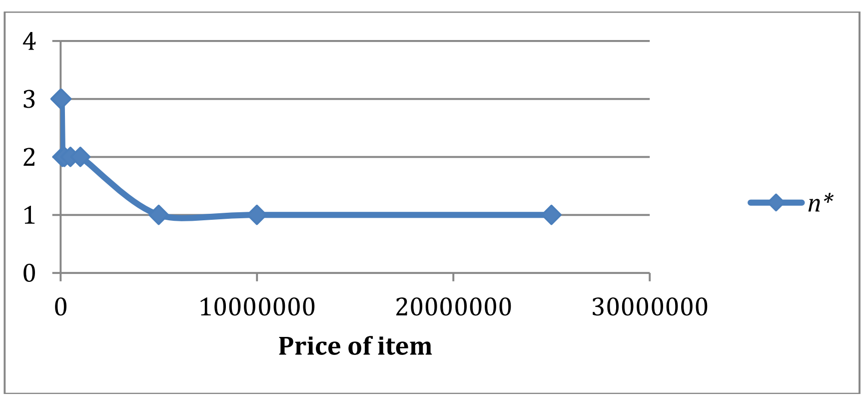


Figure 1.3 Dependence of optimal number of suppliers on price of item

In the case where, quantity discount by suppliers is offered by suppliers, the cost or price of item has an indirect relationship on the optimal number of suppliers as depicted in figure1.3. Other dependencies like optimal number of suppliers on loss due to supply disruption, optimal number of suppliers on probability of supplier's failure etc can also be depicts like figure 1.3. The results show that the optimal number of suppliers varies mainly on the cost of operating the supplier, cost of item and the probability of supplier's failure. In the present case, it is recommended that the OEM should have different number of suppliers depending upon the cost of item under consideration.

1.4 Evaluation of Suppliers

There are various integrated approaches for supplier selection. It was noticed that the integrated AHP approaches are more prevalent. The wide applicability is due to its simplicity, ease of use, and great flexibility {(Ho (2008)}. In supplier selection problem, besides the weightings of alternative suppliers, the decision makers also need to consider the resource limitations (e.g., budget of buyer and capacities of suppliers). For this reason, the GP can compensate for AHP. There are various techniques used for Supplier selection in which the most widely accepted are the MCDM (multi Criteria Decision Method) techniques.The supplier selection is a formalized system, governed by strict rules and policies, basically on the concept of selection of 'lowest priced technically suitable offer'. But for certifying an offer technically suitable, the methodology is quite subjective. However, the modern concepts of supplier selection, supplier development, long term contract, buyer-supplier relationship maintenance system, total cost of ownership are not visible in the organization.

The standard stipulates Quality, Price, Delivery, Service and System as five key factors to work out rating system in supplier selection criteria.The important point in the selection of criteria is that they it must be possible to quantify them and the data also be available. In the typical scenario of the Locomotive Manufacturer which is required to abide by rules of public procurement and also exhibit fairness at all stages, any criteria which cannot be quantified will be subject to the scrutiny. It is also important that the most prevalent criteria used by the researchers and practitioners be identified. Based on past research, and evaluation of criteria applicable and manageable in present context of the locomotive manufacturer, the four salient criteria Quality, Delivery, Price and Service are identified as important in supplier evaluation and selection.

1.4.1 Taguchi Loss Function

In traditional systems, the product is accepted if a product measurement falls within the specification limit; otherwise, the product is rejected. Taguchi suggests a more narrow view of quality by indicating that any deviation from the target value results in a loss. If a measurement is the same as the target value, the loss is zero. Otherwise, the loss can be measured by using a quadratic function, after which actions are initiated to bring back the system close to the target value.

Three types of loss functions are used in the Taguchi loss function. The first, the nominal value, is the best value. Second is smaller is better and third one is larger the better. The loss functions are given by eq. 1.3, 1.4 and 1.5, respectively:

$$L(y) = k\,(y - m)^2 \qquad (1.3)$$

$$L(y) = k(y)^2 \qquad (1.4)$$

$$L(y) = k/y^2 \qquad (1.5)$$

Where, ***L(y)*** is the loss associated with a particular value of quality character ***y***, ***m*** is the nominal value of the specification, ***k*** is the loss coefficient, whose values are constant. The four criteria as stated earlier have been tabulated in Table 1.7 with their recommended target value, range, specification limit and Taguchi loss coefficient.

Table 1.7 Target Value of Losses and Specification Limits

Criteria	**Target Value**	**Range**	**Specification limit**		**Taguchi Function**	**Loss**	**Value of k**
Quality	0%	0-5%	5% rejection	Lower the better	ky^2	$100=k \times (0.05)^2$	40000
Delivery	0	0-15	15 days	Lower the better	ky^2	$100=k \times (15)^2$	0.4444
Price	Lowest	0-10%	10% higher	Lower the better	ky^2	$100=k \times (0.10)^2$	10000
Service	100%	100%-50%	50% lower	Higher the better	k/y^2	$100=k / (0.50)^2$	25

1.4.2 Weighting the Criteria

After the quality losses of all the criteria for all the suppliers are calculated by Taguchi loss functions and the weights of all the decision criteria are obtained by the AHP process, the total loss of all the criteria to each supplier can be calculated:

$$\text{Loss} = \sum_{i=1}^{n} w_i L_{ij}$$

Where, Loss is the total loss of supplier for all evaluation criteria, $\boldsymbol{w_i}$ is the weight of criterion i calculated from the AHP and $\boldsymbol{L_{ij}}$ is the Taguchi loss of criterion i of supplier j.

An integrated method employing Taguchi loss method, AHP for pair wise comparison and TOPSIS for final ranking would suffice the requirement of the study for finding an objective, acceptable and simple procedure for supplier evaluation. After the quality losses of all the criteria for all the suppliers are calculated by Taguchi loss functions, the weights of all the decision criteria are obtained by the AHP process. The matrix with relative importance of criteria is formed (

Table 1.8 Relative importance **matrix**).

Table 1.8 Relative importance matrix

	Quality	Delivery	Price	Service
Quality	1	**3**	**7**	**9**
Delivery	1/3	1	**3**	**5**
Price	1/7	1/3	1	**1/3**
Service	1/9	1/5	3	1
SUM (col)	**1.59**	**4.53**	**14.00**	**15.33**

From calculation for the AHP, the relative weights of each criterion have been defined. Quality: 62%, Delivery : 22%, Price : 8% and Service: 8%. From a holistic view, if the quality, delivery and service of a supplier are not as per specifications or within range, it can cause loss and irreparable damage to the organization. The Taguchi Loss is calculated as (Value of Taguchi constant, ***k***) ***** (*Loss*, taking the value of each supplier)^2. The data on each attribute/criterion for every supplier and Taguchi loss is shown by the formula *k*(Value)*^{2} in table 1.9 with bold font:

Table 1.1 Data for Suppliers of particular item and calculation of loss

Supplier	**Quality % rejection**	**Delivery delay in days**	**Price compared to lowest**	**Service level opinion**
A	**3% = 36**	5 = **11.11**	0% = **0**	85% = **34.60**
B	**3% = 36**	6 = **16.00**	6.50% = **42.25**	75% = **44.44**
C	2% =**16**	7 = **21.78**	8.40% = **70.56**	80% = **39.06**
D	4% = **64**	2 = **1.78**	4.20% = **17.64**	65% = **59.17**

The normalized decision matrix is produced by dividing each element with the square root of the sum of squares of all values in each column.

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^{m} x_{ij}^2}}, \text{ where, } \boldsymbol{i} = 1, 2, ..., \boldsymbol{m} \text{ and } \boldsymbol{j} = 1, 2, ..., \boldsymbol{n}$$

The weights as determined from the AHP method can be deployed with the help of formula $v_{ij} = r_{ij} \times w_j$.

In the next step the Positive and Negative Ideal solutions are determined:

	Q	D	P	S	
v^-	0.4762	0.1637	0.0671	0.0523	Negative Ideal
v^*	0.1190	0.0134	0.0000	0.0306	Positive Ideal, lesser loss

The separation measures are now calculated with the help of formula:**Error! Reference source not found.**

$S_i^- = \sqrt{\sum_{j=1}^{m}(v_{ij} - v_j^-)^2}$ ... (Negative ideal solution) $S_i^* = \sqrt{\sum_{j=1}^{m}(v_{ij} - v_j^*)^2}$ … (Positive ideal solution)

The final step is estimation of Relative Closeness to Ideal for each supplier, ranking depends on closeness of this value to 1. Table 1.10 depicts the value and ranking of suppliers.

Table 1.10 Relative closeness to ideal and final ranking

	$C_i^* = \frac{S_i^-}{S_i^* + S_i^-}$	Ranking
A	0.587315	2
B	0.5337	3
C	0.68471	1
D	0.306774	4

Based on data, it is recommended that the Manufacturer should have the classes of suppliers like a. Approved – The first 3 of the ranked suppliers from past performance record, b. Enlisted – Under active consideration, c. Potential suppliers – Those who have supplied in past but not ranked.

The advantages of being approved for more than 3 years are, a. Self certification for quality,
b. Purchase preference at least 10% assured business, c. Preference for entering into long term contracts.

1.5 Conclusion

The procurement system at the Locomotive Manufacturer was analyzed with the aim to suggest quantitative methods to improve the process. Optimizing the supplier base will reduce the cost of items by 2-5% as due to increased volumes, the suppliers will reduce the price. By focusing on supplier issues and solving the issues the potential loss due to supply disruptions can be reduced, which will mean reduction in financial loss.

Bibliography:

1. Abginehchi, S., & Farahani, R. Z. (2010). Modeling and analysis for determining optimal suppliers under stochastic lead times. Applied Mathematical Modelling,34(5), 1311-1328
2. Adler, N., Friedman, L., & Sinuany-Stern, Z. (2002). Review of ranking methods in the data envelopment analysis context. European Journal of Operational Research, 140(2), 249–265.
3. Agrawal, N. and Nahmias, S, (1997) Rationalization of the supplier base in the presence of yield uncertainty, Production and Operations Management 6(3), 291-308, 1997
4. Berger, P. D., Gerstenfeld, A., and Zeng A.Z. (2004); How Many Suppliers are best? A Decision-Analysis approach, Omega: The international Journal of Management Science, 32(1), 9-15
5. Bharat Heavy Electricals Limited (BHEL), supplier evaluation, approval & review procedure (2010)
6. Bhattacharya, A., Geraghty, J., & Young, P. (2010). Supplier selection paradigm: An integrated hierarchical QFD methodology under multiple-criteria environment. Applied Soft Computing Journal, 10(4), 1013–1027
7. Chai, J., *et al*. (2013). Application of decision-making techniques in supplier selection: A systematic review of literature.Expert Systems with Applications: An International Journal, Volume 40 Issue 10,3872-3885
8. Chin-Nung Liao, Hsing-Pei Kao,(2010) Supplier selection model using Taguchi loss function, analytical hierarchy process and multi-choice goal programming, Computers & Industrial Engineering, Volume 58, Issue 4, May 2010, Pages 571-577
9. De Boer, L., Labro, E., Morlacchi, P., (2001) A review of methods supporting supplier selection. European Journal of Purchasing and Supply Management 7, 75–89
10. Feng, B., Fan, Z., & Li, Y. (2011). A decision method for supplier selection in multiservice outsourcing. International Journal of Production Economics, 132(2), 240–250
11. Gelderman, C.J., (2000),‘Rethinking Kraljic: Towards a Purchasing Portfolio Model, Based on Mutual Buyer-Supplier Dependence’, *Danish Purchasing & Logistics Forum* (*Dansk Indkøbs- og Logistikforum)*, Vol. 37, No. 10, October, 9-15
12. Ho W., Prasanta K. Dey, Martin Lockström, (2011),"Strategic sourcing: a combined QFD and AHP approach in manufacturing", Supply Chain Management: An International Journal, Vol. 16 Iss: 6 pp. 446 - 461
13. Kauffman, R. G., & Popkowski Leszczyc, P. T. (2005). An optimization approach to business buyer choice sets: how many suppliers should be included?. *Industrial Marketing Management*, *34*(1), 3-12.
14. Kraljic, P., (1983),Purchasing must become supply management, Harvard Business Review, 16(5), 109-117
15. Lee, A. H. I. (2009). A fuzzy supplier selection model with the consideration of benefits, opportunities, costs and risks. Expert Systems with Applications, 36(2), 2879–2893
16. Liao Chin-Nung, Kao Hsing-Pei, (2010), Supplier selection model using Taguchi loss function, analytical hierarchy process and multi-choice goal programming, Computers and Industrial Engineering, Volume 58, Issue 4, Pages 571-577
17. Marquez A Crespo, Blanchar C. (2004); The procurement of strategic parts, Analysis of portfolio of contracts with suppliers using a system dynamics simulation model, International Journal Of Production Economics, 88(1), 29-49
18. Narasimhan, Ram; Talluri, Srinivas; Mendez, David, (2001), Supplier Evaluation and Rationalization via Data Envelopment Analysis: An Empirical Examination, Journal of Supply Chain Management 37(3): 28-37
19. Saaty, T. L. (1990). How to make a decision: the analytic hierarchy process,.*European journal of operational research*, *48*(1), 9-26.
20. Sarkar A., and Mohapatra Pratap K.J., (2006), Evaluation of supplier capability and performance: A method for supply base reduction, Journal of Purchasing and Supply Management, Volume 12, Issue 3, Pages 148-163
21. Yang, G., Qian, C. (2008); The optimal number of suppliers considering the quantity discount and supplier failure, IEEE International Conference on Service Operations and Logistics, and Informatics, 12-15 Oct, 2327–2329

Enhancing the leanness of supply chain by integrated Fuzzy-QFD approach

Rajiv K Upadhyay[1] , Ajay Bangar [2] , Pawan Kumar Singh [3] , Ashish Shastri [4]
[1,2,3,4] Department of Mechanical Engineering
[1] Hindustan College of Science & Technology (HCST), Mathura, UP, India, INDIA

Email: drrajeevkrupadhyay@gmail.com, abangar1000000@gmail.com
pawanme48@gmail.com, ashishshastri2000@yahoo.co.in

Abstract

Lean attributes and lean enablers, i.e. the capabilities which allow promptly responding to reduce the waste in the business environment and the available leverages to achieve the leanness. Few studies, however, provide empirical evidence about the main characteristics of lean enterprises and tools practically exploited by companies to achieve leanness. This study attempts to improve the existing knowledge on leanness. On the other hand, many researchers have focused on the supply chain in recent years because of the high importance of this issue in the supply chain management context. By linking Lean Attributes (LAs) and Lean Enablers (LEs), this study used Quality Function Deployment (QFD) to identify viable LEs to be practically implemented in order to increase the leanness of the supply chain. FQFD (First QFD)) are used to rank lean attributes Furthermore, fuzzy logic is used to deal with linguistic judgments expressing relationships and correlations required in QFD. To illustrate the practical implications of the methodology, the approach is exemplified with the help of a case study in the some processing industry.

***Keyword:*-supply chain, Lean supply chain, Quality Function Deployment (QFD), House Of Quality (HOQ), Fuzzy logic.**

Introduction

The need for holistic modeling efforts that capture the extended supply chain at a strategic level has been clearly recognized first by industry and recently by academia. In addition, in today's world of global markets and stiff competition for every product along with increasing consumer demand, it becomes imperative for companies to explore ways to improve their productivity in terms of implementing flexible and standardized technology and adopting proven management principles. Lean production is one of the initiatives that many major businesses all around the world have been trying to adopt in order to remain competitive in the increasingly global market. Lean production promises significant benefits in terms of increased organizational and supply chain communication and integration. The core thrust of lean production is the capability of working synergistically to create a high-quality system that produces finished products at the pace of customer demand with little or no waste.

A supply chain consists of all parties involved, directly or indirectly, in fulfilling a customer request. The supply chain includes not only the manufacturer and suppliers, but also transporter, ware house, retailer, and even customers themselves. Leanness as a concept has been firstly stated by John Krafcif as a term for the new production system applied by Toyota. Herron and Hicks argue that the reason behind Toyota implementation for such concept has been the fact that it couldn't afford the huge capital- based mass production systems applied by the US companies. As a result, it has searched for means to reduce waste in all its operational activities, and hence, Lean "production" has been born.

Quality Function Deployment (QFD) is an effective tool for planning attributes of new products based on customer demands and involves all members of the producer or supplier organization. QFD can be used to integrate an organization's diverse sources of information during product and process development, so that the goal of Total Quality Management (TQM) and Concurrent Engineering (CE) inside the organization can be facilitated.

The aim of this study is to develop an integrated approach to increase the leanness of the supply chain. Furthermore, a practical tool is introduced that can be easily adopted to implement lean strategies. By linking Lean Attributes

(LAs) and Lean Enablers (LEs), this study, which is based on the House of Quality (HOQ) of Quality Function Deployment (QFD) methodology, aims to identify the most appropriate LEs for implementation by supply chain management. LAs represent company requirements and appear as ‘‘what’s” in the HOQ, while LEs are listed as ‘‘how’s”, since they are considered as practical tools that the company can use to achieve leanness. Moreover, in the proposed approach, fuzzy logic is used to deal with linguistic judgments expressing the relative importance of LAs as well as the relationships and correlations required in the HOQ.

Contribution In Study

A first outcome of the literature analysis is that none of the approach proposed in the literature ground on the QFD methodology. This study, and specially the HOQ represents a particular tool, which allows directly assessing the impact of lean attributes on competitive bases and of lean enablers on lean attributes, through the relationship matrixes. Clearly in practical cases, it would also be possible that a company directly identifies a set of suitable lean enablers to be implemented, without linking them with lean attributes and competitive priorities. However, in the case, the risk is that the selected strategic leverages do not marketing objectives. Moreover, QFD allows identifying correlations between attributes or enablers, which are not examined in the methodologies available in the literature.

As a further contributions, existing methodologies ground on crisp assessments, and in this regards, we believe that fuzzy approach could provide useful advantages. A main reason why our approach is based on fuzzy set theory is that, as a result of the literature analysis, it emerged that leanness assessment is often dealt with through fuzzy logic, due to the imprecise and vague definition of leanness indicators consequently, linguistic expressions are often used to estimate both companies performance against lean attributes and enablers, and the relative importance of such parameters. When vague or ill-defined issues should be examined, the adoption of fuzzy logic is recognized to substantially improve the capabilities of traditional crisp approaches. As a matter of fact, the main strengths of fuzzy logic have to be found in the opportunity to take into account the different meanings that decision makers (DMs) may give to the same linguistic expression, as well as to translate linguistics judgments into numerical values. Methodology

The framework for achieving a lean supply chain by Fuzzy- QFD comprises four main parts. It has a stepwise description as shown in Fig.1 and below. The fuzzy HOQ whose specific structure is detailed in Fig. 2 is adopted here.

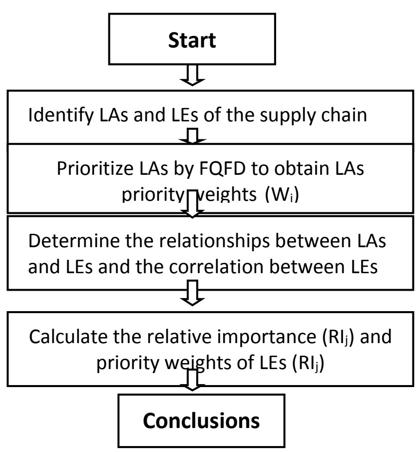


Fig.1.Schematic representation of the methodology.

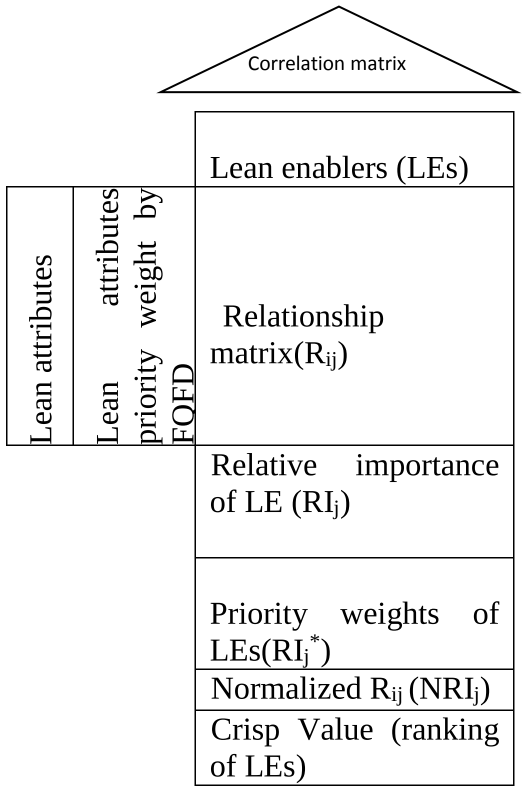


Fig.2. The proposed integrated Fuzzy-QFD approach

A. Identify LAS and LES of the supply chain

To be truly lean, a supply chain must possess a number of distinguishing attributes and enablers. LAs here after are defined as elements which constitute the underlying structure of a lean organization. They were originally conceived as core concepts of lean manufacturing. Accordingly, LEs are enabling tools, technologies, and methods critical to successfully accomplish lean supply chain management.

LAs enhancing supply chain leanness and LEs to be exploited in order to achieve the required LAs, as accepted by several authors [6], were identified. On the basis of a review of the normative literature some LAs and LEs were defined for the lean supply chain, as shown in Table 1. Furthermore, suggestions to identify viable sets of lean attributes and enablers can be found in literature, and different or additional LAs/LEs may be listed.

Table 1

Lean attributes and enablers defined for lean supply chain from related to food industry.

Service level improvement Conformance quality
Pull production Delivery reliability
Eliminate obvious wastes Low buffering cost
Cost efficiency Low variability in process time JIT manufacturing Low variability in delivery time
Continuous improvement Low variability in demand rates
Human resource training Delivery speed
Quality improvement Low variability in process time
Vendor management inventory

Total quality management
Supplier management

B. Prioritize LAs by FQFD to Obtain LAs Priority Weights (Wi)

Due to its wide applicability and ease of use, the FQFD, has been studied extensively for the last twenty years. It has been widely used to address multi-criterion decision- making problems. The FQFD consists of three main operations: Identifiy criteria(competetive base CB_i) and Subcriteria Lean Attributes LAs), Find Out relationship and correlationship between CB_i and LAs, and Find Out relative importance and crisp value by applying triangular fuzzy logic. Moreover, QFD is one of the five tools that are commonly combined with the Fuzzy Logic. In this study, the FQFD was deployed to prioritize LAs. After defining LAs, their priority weights were computed by using FQFD for this purpose, first, the pair-wise assessment matrices were prepared to evaluate the eight alternatives, i.e. LAs with respect to criteria; the criteria were then evaluated with respect to the goal. Effective management of the supply chain is viewed as the driver of decreasing the cost of material, services, and manufacturing, reducing lead times and improving product quality and responsiveness. Therefore, after evaluating the related project, five criteria

Were identified: speed, cost, responsiveness, competency, and quality. To identify the appropriate criteria to rank LAs of supply, those attributes that help meet the rapidly changing needs of the business environment of supply were considered. Thus, LAs were evaluated on the basis of quality as well as the speed of providing products, services, and information to customers, suppliers, and employees, as quality and speed are significant criteria in the leanness assessment of a supply chain. Cost was another criterion identified for the evaluation of LAs because the focus of the lean approach is on cost reduction. The responsiveness ability of a lean supply was also considered. Another notable criterion in pair-wise assessment of LAs was the competency ability obtained in a lean supply by achieving the mentioned LAs. These five criteria are important in evaluating the degree of leanness achieved in a supply chain; therefore, the mentioned alternatives were evaluated with respect to each criterion to determine how much achievement of the LAs may lead to a leaner supply. Output of FQFD represented as Wi, is the input of the Second Fuzzy-QFD (SQFD) component of the proposed model.

C. Determine The Relation Between LAS And LES And Correlation Between LES

Because of the qualitative and ambiguous attributes linked to lean implementation, most measures are described subjectively using linguistic terms that cannot be handled effectively using conventional approaches. However, fuzzy logic provides an effective means of dealing with problems involving imprecise and vague phenomena . It was exploited to translate linguistic judgments required for relative importance of LAs, relationships, and correlations matrices into numerical values. In this step, the degree of relationship between LAs and LEs was stated by the corresponding TFNs and placed in the HOQ matrix. Moreover, the degree of correlation between LEs was then expressed by TFNs in the fuzzy HOQ. Both of these correspondences are shown in Tables 2 , 3 and 4.

TABLE 2
Degree of relationships, and corresponding fuzzy numbers[1]

Degree of relationship	**Fuzzy no.**
Strong	(0.7; 1; 1)
Medium	(0.3; 0.5; 0.7)
Weak	(0; 0; 0.3)

TABLE 3
Degree of correlations, and corresponding fuzzy numbers[1]

Degree of correlation		Fuzzy no.
Strong positive	(SP)	(0.3; 0.5; 0.7)
Positive	(P)	(0; 0.3; 0.5)
Negative	(N)	(-0.5; -0.3; 0)
Strong Negative	(SN)	(-0.7;-0.5;-0.3)

TABLE 4
The 4 point linguistic scale for importance judgment

Importance judgment		Fuzzy No.
Very High	(VH)	(0.7; 1; 1)
High	(H)	(.5; 0.7; 1)
Low	(L)	(0; 0.3; 0.5)
Very Low	(VL)	(0; 0; 0.3)

1) TFN (TRIANGULAR FUZZY NUMBER)

The TFN can be denoted as a triplet (a, b, c), as shown in Fig. 3., where, $a \le b \le c$. When $a = b = c$, it is a non-fuzzy number by convention. The membership function can be defined as[2] :

(x-a)/(b-a), x is a function of [a,b]
$\mu_n(x)$:(c-x)/(c-b) x is a function of [b c]
0 otherwise 1

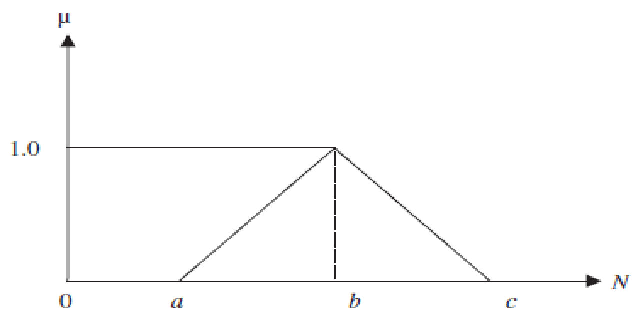


Fig 3 Triangular fuzzy number (TFN).

If M = (a1, b1, c1) and N = (a2, b2, c2) represent two TFNs, then the required fuzzy calculations are performed as given below.

Fuzzy addition : $M + N = (a_1 + a_2; b_1 + b_2; c_1 + c_2)$ …………2
Fuzzy multiplication : $M \times N = (a_1\times a_2; b_1\times b_2; c_1 \times c_2)$ …………..3
$M \times 1/N = (a1/c2; b1/b2; c1/a2)$ ………….4
Fuzzy and natural number multiplication: $r \times M = (r.a; r.b; r.c)$……... 5

D. CALCULATE RELATIVE THE IMPORTANCE (RI_J) AND PRIORITY WEIGHTS OF LEs (RI_J^*)

The aim of computing these two parameters was to determine which LE has the most effect on the lean supply chain. RIj was computed by fuzzy multiplication of W_i to R_{ij}.

$RI_j = \sum W_i \times R_{ij}$ j=1,2,3……m 6

$RI_j^* = RI_j + \sum T_{kj} \times RI_k$ j=1,2,3……m 7

T_{kj} was shown in the roof part of HOQ. The mentioned parameters are shown in Fig. 3. Furthermore, normalization was performed by dividing each RI_J^* by the highest one according to the fuzzy set algebra . Then, in order to rank the LEs, the normalized scores of RI_J^* were de-fuzzified. Suppose M (a, b, c) is a TFN; then, the de-fuzzified value is computed as (a+4b+c)/6.

8LEs with high crisp values indicate that they can be usefully exploited to enhance relevant LAs. Thus, such enablers must be selected for implementation.

RESULT AND DISCUSSION

This section presents an example of the proposed approach through a case study in the liquor industry to illustrate the usefulness and ease of application of the method as well as considering the practical implications of the approach. Focusing on the methodological point of view, the definition of a specific set of and LEs for applying the approach was not dealt with in this project; they should be identified according to the special characteristics of the company under consideration.. Then, 8 LAs and 11 LEs were identified these are shown below in Fig.4

Step 1: identifying the competitive bases a company is willing to achieve competitive advantage;

Step 2: identifying lean attributes enhancing the selected competitive bases and filling the first HOQ;

Step 3: identifying lean enablers to be exploited in order to achieve the required lean attributes, and filling the second HOQ.

APPLICATION STEP 1

As the starting point of the approach proposed, a company should identify the relevant competitive bases. For illustration purpose, the set of competitive bases used in the numerical example has been ground on existing studies and information available in literature. Specifically, a viable list of five competitive bases (CB_i, i = 1,.....,5), namely “speed”, “cost”, “responsiveness”, “competency” and “quality”, was derived . They are listed as “whats” in fig. 4.

In real case applications identifying the relevant competitive bases of a company would require direct contacts with company’s members (in particular, marketing manager), either in the form of interviews or roundtable discussion. To support the application of the methodology in practice, and to quickly collect the required information, it is suggested to setup an appropriate workgroup, headed by academics and including firm’s executives, reporting to the main business functions involved in the development of lean strategies.

APPLICATION STEP 2

Step 2 requires filling the first HOQ, which, in turn, involve the following sub-steps:

Defining the fuzzy linguistics scales;

Assessing the relative importance of competitive bases;

Listing the lean attributes

Assessing the relationship between lean attributes and competitive bases;

Identifying possible correlations between lean attributes.

Sub-step i. Fuzzy linguistics scales to be used to assess weights of CB_s, relationships between CB_s and LA_s and correlations between LAs could be either defined by the workgroup or derived from the literature. In this example, they were taken from Bottani and Rizzi (2006).

Sub-steps ii. In this example, the relative importance w_i of CBs was defined based on the work by Ren et al. Starting from findings by the authors, w_i (i = 1,....5) were pondered based on a normalised 4-point fuzzy linguistic scale, ranging from “very low’’ (VL) to “very high’’ (VH), as shown in table4. Relative importance of CBs is listed in the second column of Fig.4.

Whenever the procedure is applied to a real case, the same information can be derived asking company’s members to express their judgment against the relative importance of competitive

bases with regard to the over all strategy of the company. Judgment will be thus translated into fuzzy numbers according to the scale defined in the previous sub-step.

Sub-step iii. A viable list of LAs should be defined depending on the specific case in example. For the purpose of this example 8 lean attributes suggested by Zarei et al. Where listed as LA1,...., LA8 in columns in the HOQ, and their acronyms are detailed more in fig 4.

Sub-step iv. Selecting the set of LAs proposer by Yusuf et at. Greatly simply the assessment of the relationships.

Since the impact of those lean attributes on the above mentioned set of competitive bases was investigated by Ren et al. Accordingly, as well grounding on the fuzzy scale proposed in table 2 to4, the relationships matrix of this example was built as shown in the centre of fig.4.

The same assessment, in real case, requires inner viewing the workgroup, to drive information concerning the impact of each LAs on the CBs identified in sub-step ii. Specifically, company's member should be asked to assess how, and to what extent, an LA has potential to enhance a given CBs; judgments could be expressed on a linguistics scale and translated into fuzzy numbers according to previous sub-step i.

Sub-step v. Outcomes proposed in table 5 were used as a guideline to derive correlations required in the roof of the HOQs, as all the studies cited in the table suggest possible links between LAs. According to such findings, the roof of correlations of the first HOQ was built as shown in fig.4. Then, the relative importance RI_j (j = 1,.....,8) and the final scores of lean attributes were computed applying Eqs. (6) and (7). Outcome of the computation are presented in the last row of fig.4.

The application of the methodology to a real case would require interviewing company's members, to get, based on their in-field experience, information concerning the possible impact of lean attribute on another. Such information should be expressed following the linguistics scale defined in sub-step i, and translated in fuzzy numbers for computational purpose. None the less, findings from the literature could be useful to suggest possible interactions between lean attributes to interviewees, thus helping in the evaluation

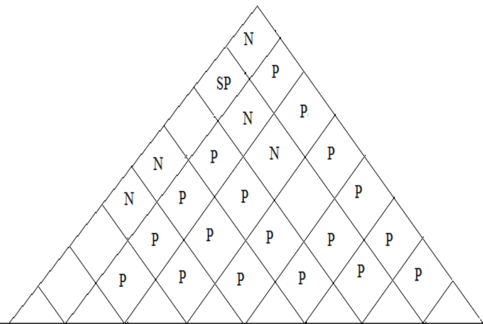
Sub-step vi. To find the normalized value of LAs so this can be use in second fuzzy QFD as the relative weight (wi) as shown in fig 4 the normalized value shown in table 5

APPLICATION STEP 3

Requires building the second HOQ, which involves the following sub-steps:
defining the fuzzy linguistics scale;
assessing the relative importance of lean attributes;
listing the relevant lean enables;
assessing the relationship between lean attributes and lean enablers; and
identifying possible correlations between lean enablers.

Sub-step i. To be consistent with results of the previous application step, in this example we adopt the same fuzzy linguistics scales to assess weights of LAs, relationships between LAs lean LEs in the second HOQ.

Sub-step ii. Depending on the specific case study, the second HOQ can be built starting from all LAs examined in the previous step, or the analysis can be limited to those attributes which got the highest score in the first HOQ. In order to thoroughly illustrate the application of the methodology, in this example all LAs previously examined will be considered as "whats'' in the second HOQ. Importance weights of LAs can be derived from the final score obtained in the first HOQ. However , since a normalized fuzzy scale has been adopted to express importance judgment of competitive bases in the previous step, a preliminary normalization of fuzzy scores of lean attributes is suggested before they are used as importance weights in this step. Normalization is performed by dividing each score by the highest one, i.e. score of LA4, according to the fuzzy sets algebra. Lean attributes and related normalized



CBi \ LAj	Wi	LA1	LA2	LA3	LA4	LA5	LA6	LA7	LA8
Speed	VH	W	S	S	W	M	W	W	S
Cost	L	W	M	M	W	S	W	W	W
Responsivness	H	S	S	W	M	W	M	M	S
competency	L	S	S	M	S	S	S	M	S
Quality	VH	S	S	M	W	W	W	W	W
RELATIVE IMPORTANCE (RIj)		1.47 3 3.45	1.96 4.15 4.35	0.91 2.15 3.05	0.70 1.5 2.45	0.70 1.5 2.60	0.70 1.50 2.39	0.42 1 2.09	1.47 3 3.39
SCore j (RIj*)		0.161 1.70 4.913	1.75 6.745 11.290	0.70 5.345 10.895	-0.035 3.390 10.595	-0.035 4.590 11.46	0.70 5.490 11.355	-0.574 2.410 8.695	0.735 5.640 12.06
NORMALIZED VALUE (NRIj*)		0.013 0.252 2.807	0.145 1 6.450	0.058 0.792 6.225	0.00 0.502 6.054	0.00 0.681 6.548	0.058 0.814 6.488	-0.048 0.357 4.968	0.061 0.836 6.891
Crisp Value		0.831	2.148	1.967	1.76	1.978	2.044	1.408	2.156

Fig 4 First Fuzzy QFD approach For finding W_i of LAs

TABLE 5
Normalized value of lean attributes

Lean Attributes	Relative weight ($W_{i)}$ or normalized score
LA1	0.058
LA2	0.150
LA3	0.138
LA4	0.123
LA5	0.140
LA6	0.143
LA7	0.098
LA8	0.151

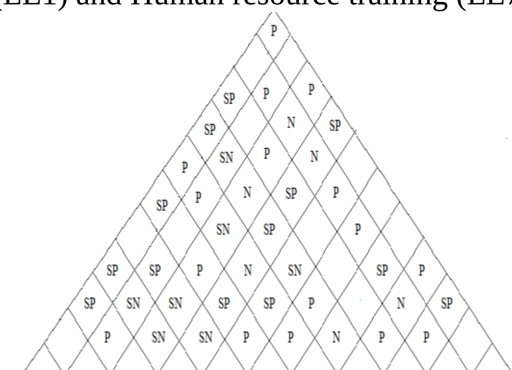
importance weights are thus listed in the first columns of the second HOQ, shown in fig.5.
Sub-step iii. As per the previous steps, LEs to be implemented by companies in order to acieve leanness should be defined according to the case study company considered. Interviews with company's members could be useful to

derive a possible list of enablers. In this example, starting from the works of Gunasekaran, and gunasetaran and yusuf, 11 viable lean enablers were identified and listed as LE1,.....,LE11 in columns in fig.5. Acronyms are explained in note to the figure.

Sub-step iv. To our knowledge, no specific studies are currently which thoroughly describe the impact of lean enablers on lean attributes, thus directly providing the relationships matrix of the second HOQ. Such relationship, however, are partially dealt with by scientific literature. The resulting values are proposed in the centre of fig.5 .Findings from the literature could be useful even in a real case application, to suggest possible interviewees to help in the evaluation.

Sub-step v. As per the previous sub-step, in this example possible correlations between LEs were derived from the literature. On the basis of the literature examined, as well as on the degree of correlations proposed by company, the roof of correlations of the second HOQ is built as shown in fig. 5 Findings from the literature could also be useful in the case the second HOQ is built based on experts' opinion, to suggest possible interactions between enablers, as well as correlations between enablers, thus helping in the assessment.

The relative importance RI_k (k = 1,.....,7) and the find $score_k$ of LEs were computed according to Eqs. (6) and (7). Eq. (8) was finally adopted to derive crisp scores. Outcomes of the computation are presented in the last rows of fig. 5. AS a result of the computation, supply chain management practices Quality improvement (LE8) got the highest crisp score, due to both the wide number of positive correlations with others LEs and strong relationships with several LAs. Thus, such an enabler has the highest implementation priority in order to achieve leanness, followed by Service level improvement (LE1) and Human resource training (LE7) shown in table 6.



LAj \ LEk	Wi	LE_1	LE_2	LE_3	LE_4	LE_5	LE_6	LE_7	LE_8	LE_9	LE_{10}	LE_{11}
LA1	0.058	W	S	W	W	S	M	S	M	M	S	M
LA2	0.15	S	W	S	S	M	M	W	M	S	W	S
LA3	0.138	S	W	S	S	M	M	M	S	S	W	S
LA4	0.123	S	W	S	S	M	M	W	W	M	W	M
LA5	0.14	S	S	S	W	S	M	S	W	M	S	M
LA6	0.143	S	W	S	S	M	M	W	W	S	W	S
LA7	0.098	S	W	S	S	S	M	S	S	M	S	M
LA8	0.151	S	W	S	S	M	M	W	W	S	W	S
Rij (Relative Importance)		0.659 0.942 0.959	0.139 0.198 0.139	0.659 0.942 0.959	0.561 0.802 0.861	0.419 0.649 0.790	0.30 0.50 0.70	0.249 0.365 0.563	0.228 0.340 0.549	0.533 0.792 0.875	0.207 0.296 0.508	0.708 0.792 0.875
SCOREj (RIj*)		1.343 2.977 4.438	-0.288 0.432 1.584	-0.221 1.512 2.624	0.209 1.392 2.975	-0.478 0.127 1.385	0.666 2.149 3.659	-0.419 0.718 2.233	-1.286 1.110 2.679	1.111 2.535 3.770	-0.442 0.494 2.304	1.038 2.256 3.324
Normalized Value (NRIj*)		0.204 1 3.304	-0.065 0.145 1.179	-0.05 0.508 1.954	0.047 0.468 2.215	-0.108 0.043 1.031	0.15 0.722 2.724	-0.094 0.241 1.663	-0.290 0.373 1.995	0.25 0.852 2.807	-0.099 0.166 1.715	0.234 0.758 2.475
Crisp Value		1.377	0.504	0.730	0.799	0.252	1.079	0.513	0.613	1.190	0.487	1.056

FIG 5 Second Fuzzy QFD For ranking of Lean enablers

TABLE 6

RANKING OF LEAN ENABLERS

Rank	LEs	Crisp Value
1	LE1	1.377
2	LE9	1.190
3	LE6	1.079
4	LE11	1.056
5	LE4	0.799
6	LE3	0.730
7	LE8	0.613
8	LE7	0.513
9	LE2	0.504
10	LE10	0.487
11	LE5	0.252

CONCLUSIONS AND SUGGESTIONS

In this study, an integrated Fuzzy-QFD approach was proposed to enhance the leanness of the supply chain. The approach showed the applicability of the QFD methodology, especially of the HOQ, to identify viable lean enablers for achieving a defined set of LAs. The FQFD was used to prioritize lean attributes. In order to cope well with the vagueness of linguistic judgments required in building the HOQs, relationships, and correlations, Wi, relative importance (RI_J), and priority weights (RI_J^*) of LEs were all defined with TFNs. A case study was presented to illustrate the ease of application of the proposed approach.

The focus of attention of future researches can be on the integration of useful methods with QFD to prioritize LEs in order to enhance supply chain leanness. Future researches can also consider utilizing other ranking methods instead of the FQFD, such as TOPSIS, AHP, to prioritize the LAs and compute their priority weights. Moreover, W_i, i.e. LA priority weights, obtained from different ranking methods, can be compared. In this project, a case study of a company in the liquor industry was presented. More case studies for other supply chains should be presented.

Acknowledgment

The author wish to express their sincere appreciation to the top management of belsend sugar company (distillery division). DR. Ghanshyam Singh for his useful advice and comment.

REFERENCE

[1] Bottani E. Rizzi, AStrategic management of Logistics service: “A fuzzy-QFD approach”. International Journal of Production Economics 103 (2), 2006, pp. 585–599.

[2] Chamodrakas I., Alexopoulou N., Martakos D.,Customer, “evaluation for order acceptance using a novel class of fuzzy methods based on TOPSIS”. Expert Systems with Applications 36, 7409–7415.

[3] Chien C.J, Tsai H.H.,Using fuzzy numbers to evaluate perceived service and quality. Fuzzy Sets and Systems 116, 2000, pp. 289–300.

[4] Dohnal.M, Vystrcil.J, Dohnalova.J, Marecek.K, Kvapilik.M, Bures.P, Fuzzy food engineering. Journal of Food Engineering 19 (2), 1993, pp. 171–201.

[5] Georgiadis, P, Vlachos. D, Iakovou. E. “A system dynamics modeling framework for the strategic supply chain management of food chains”. Journal of Food Engineering 70 (3), 2005 , pp. 351–364.

[6] Hopp, W.J., Spearman, M.L., To pull or not to pull: what is the question? Manufacturing and Service Operations Management 6 (2), 2004, pp. 133–138.

[7] Hunt, I., Wall, B., Jadgev, H., "Applying the concepts of extended products and extended enterprises to support the activities of dynamic supply networks in the agri-food industry". Journal of Food Engineering 70 (3), 2005, pp. 393–402.
[8] Nambiar, A.N., Mahalik, N.P., 2010. Trends in food packaging and manufacturing systems and technology. Trends in Food Science and Technology 21 (3), 2010, pp. 117–128.
[9] Schonbergerm, R.J., Japanese production management: an evolution with mixed success. Journal of Operations Management 25, 2007, pp. 403–419.
[10] Scherrer-Rathje, M., Boyle, T.A., Deflorin, P., Lean, take two! Reflections from the second attempt at lean implementation. Business Horizons 52, 2005, pp.79–88.

A Study of Factors Related to Supply Chain Strategy (SCS)

Puneet Mangla [1], Ashish Agarwal [2], Pulak M Pandey [3] , Subrata Das [4]

[1,2,3,4] Department of Mechanical Engineering
[1,4] Hindustan College of Science & Technology (HCST), Mathura, UP, India, INDIA
[2] Indira Gandhi National Open University, New Delhi, INDIA
[3]Indian Institute of Technology, New Delhi, INDIA
mangla_p@rediffmail.com, ashisha@ignou.ac.in, pmpandey@mech.iitd.ac.in, skd1316@gmail.com

Abstract

This paper presents comprehensive inspection of factors related to supply chain strategy (SCS) literature. From a narrowed sample of 143 SCS related articles, 54 were qualified as the juncture of SC strategy factor and inventory management relevance. Themes found within this literature include the developing nature of strategy, channel functioning, inter-firm relationships and strategic organizational processes involving the SCS function, its evolving nature and adoption of SCM technology. The study is seeking to ramp up on the current body of theory and practices with direct relevance to supply chain using several agenda setting articles as a footing. Additional SCS research, both conceptual and empirical would be regarded in contemporary economic climate.

Keywords : Supply chain, SC strategy, SCS, SCM technology, SCM.

Introduction

Supply chains are critical to all business customer propositions. The supply range is often overlooked in the council chamber; however, businesses with excellent supply chains that support and enforce customer offers are considerably more profitable, through driving additional sales. SC strategies are a blend of ingredients that assure that the supply chain is in tune with the business requirements. Its sensitivities have been measured and investigated context of changing nature of the job, i.e., future is becoming more uncertain and there is the need for flexibility [10]. However, Chase is of view that in the new global paradigm firms are driven to locate alternative and flexible ways to meet customers need [4]. It is obvious that effective supply chains improve business performance. Nevertheless, areas of focus are cost, inventory, and carbon reduction and delivery methods. A more important angle is that excellent supply chain increase sales through better shelf availability and improve products speed into the market.

For maintenance of competitive advantage exhaustive exploration of Supply Chain Strategy (SCS) factors is possibly a pivotal necessity [10]. Although, supply chain strategy is emerging research area of SCM and needs further consolidation in the context of construct, characteristics and application methods. The SCS research focuses on different streams such as marketing/downstream [11], e-supply chains [4], demand based [11] and SC integration [8]. Topics often focused on a priority basis within SCSF, includes, knowledge management, self-service technologies [29], supply-chain management, e-commerce, relationship technology, and enterprise resource planning (ERP). Although an improper supply chain, is a consequence of inadequate supply chain innovation, therefore holds the organization in the back seat, eventually fell behind. [39, 27]. Therefore, review of SCS factor research is substantive and critical for both academic and practitioner's perspectives. In particular for practitioners, to ascertain current knowledge and practices on SC and to identify gaps in theory and practices. This study attempts to conceptualize systematic synthesis of current literature.

Literature attributes that alignment of supply chain strategies with business strategies is essential that perhaps lead to the achievement of overall business goals. The SCS factors for maintaining a high degree of delivery reliability and high delivery flexibility has been the hallmark of business strategy [20, 22, and 25]. Firms often witnessed major gaps between business strategy and supply chain strategy. Therefore, incorporating powerful supply chain models, results in the achievement of the goal. Nevertheless, SCS factors and competitive strategy must go together [9]. SCS is defined as a process of material procurement, materials transportation, product fabrication and distribution. In this context, Mitra indicated that a highly responsive SCS is required to address uncertain demand and supply scenario, whereas efficient strategy is needed to address certain demand and supply issues [39, 26]. As such axiomatic design methodology often used where supply chain decomposition is done and each unit is designed that eventually contributes to the accomplishment of the overall business objectives. Subsequently, the articles for this research were selected by searching online database 'EBASCO' with keywords 'supply chain' and 'strategy' 'factor' in title.

Factors positioned to Supply Chain Strategies

This report synthesizes the literature on SCS factors to find out the current practice for academic inquiry. SCS factors discussed in this paper identify strengths and weaknesses, and also identify research gaps. In this context, taxonomy of SCS factors is suggested to enable academicians and practitioners to evaluate possible optimal alternatives for SCM. On the basis of lack of theory to explain SCS factors in a holistic setting, including its connection to overall business strategy, suggestions for future research directions is offered.

In recent time, supply chains have experienced a figure of major changes in customer sourcing behaviors due to:

- Increased growth in sourcing which lead to increased lead-times and stock levels,
- Variability in demand has increased the total range of merchandise and increased associated promotional activities,
- Growth of E-commerce and multi-channel retailing strategies,
- Exponential increasing cost of fuel and the demand for environmental sustainability.

Today, more than ever, there is a need to focus attention on the supply chain to secure and defend the future demands of customers. Understanding the benefits, a scheme is to secure its implementation and this must be judged whether it delivers the scheme on time, with expected benefits. In this context, effective preparation and budgeting needs to be extended to ascertain that the benefits are tracked, realized and then incorporated in the business plan.

We realize the demand for flexibility

Successful strategies are a blend of factors that ensures that the supply chain is in tune with the job prerequisites. They are as individual as your customer offer, website or retail release and their alignment with business objective is critical. While sensitivities have been measured and investigated the changing nature of the business means as the future is getting more uncertain and therefore flexibility needs to be built in [19, 21]. Despite the existence of supply chains for thousands of years, there is not sufficient and serious academic inquiry until the 1980s, after that the field developed rapidly [6]. The comparative growth of in this field of study can also be differentiated from the numerous competing definitions for supply chain management and related constructs proposed in the literature [33, 38]. Some researchers focused on the supply chain, instead of the individual company, as the unit responsible for the development and sustainment of competitive advantage [28, 30]. This phenomenon, if straight, has profound implications, and needs in-depth survey of SCS factors. It is emerging research area of supply chain management, nevertheless to build consensus in terms of definitions, constructs, characteristics and methods of application in complex issue. The SCS factors research focuses on different streams: manufacturing/upstream [2], marketing/downstream [11, 30], product based [14], demand based [18, 31], SC integration focused [15], and the strategy for electronic-supply (E-supply) chains [8]. A review of SCS factors research is important at this juncture both from academic and practitioner's perspectives. This is supported by Hauguel and Jackson, who in a survey of 300 large European and U.S. companies found that 68% companies indicated that making improvements to their supply chains was one of their three main priorities. [22,29]. It is observed that around 20% of the firms stated supply chain improvements as a concern, while more than half of the respondents rated its supply chain efficiency as 3 or lower on a 5-point scale. From this viewpoint, it is worthwhile to learn the current state of research on SCS factors and to identify gaps in literature, if any, and this paper undertakes to accomplish this through a synthesis of current literature.

Definitions

Supply Chain

A "Supply Chain (SC) is defined as a set of three or more entities (organizations or individuals) directly involved in the upstream and downstream flows of products, services, finances, and/or information from a source to a customer " [38, 33]. For the purposes of this paper, supply chain will be defined broadly so as to include Mentzer's definition as well as, third-party logistics providers, consulting companies, marketing companies and other types of firms that may be involved in the value chain but not necessarily directly involved in the manufacturing process. Although the term supply chain appears throughout the paper, the concept should be conceived as a network, web or system that functions as a whole to provide value from the highest upstream source, through various value-adding or modifying nodes, to the end customer [19].

Supply Chain Management

"Supply Chain Management (SCM) integrates the full exchange of information and movement of goods between suppliers and end customers, including producers, distributors, retailers, and any other enterprises within the extended supply chain" [20, 32].While this definition mentions information as components of most supply chains, it states nothing about management. In this paper, supply chain management is defined as planning, implementing, monitoring, controlling, where possible, and commuting, where necessary, all the processes, resources and flows within an expanded concept of the supply chain which includes all workforce in a network, web or system that functions as a whole to provide value from the highest upstream source through various value-adding nodes, to the end customer.

A strategic factor perspective

Strategy

Through an extended critique of the strategy literature, Chaharbraghi and Willis observed that strategy is driven by the need to be competitive, to adapt change, and to survive [7, 36]. They posit that the increasingly dynamic nature of environmental context has led to an increasing range of responses by organizations. As such, there is a plethora of strategy models and accompanying language which are still developing [12]. At this stage, there is no consensus in the literature as a definition of strategy or its accurate nature. Through a search of the strategy literature, we have amassed a lengthy list of notions relative to the nature of strategy and states that strategy is presented in the literature as everything from a pattern or plan, in a state of mind, and a learning procedure. There are 37 descriptors related to strategy in the literature, including differentiation and price leadership, adhering fit and scope, intent, and core competencies. Further, the literature states that strategy can be generic, deliberate or emergent, rational, incremental, implicit, explicit, prescriptive, descriptive or configurationally. Although scholars stated that this series of unconnected and often contradictory models do little, in fact, confuse practicing managers who are often fighting for the survival of their organizations and the livelihoods of their employees. Nevertheless, the authors have presented their own model which view strategy as a natural process involving constant evolution and continuous rotation [5, 36 and 40]. The factor 'Development' refers to strategies which allow the system to evolve as markets and 'revolution' refers to strategies aimed at actually making new market that can be exploited. It is remarkable that there is no consensus in the literature as to whether overt strategy making has any impact on organizational performance [47] and [13]. In recognition of the many views of strategy that exist in the literature, the broad definition that will be used in this paper is "judgmental designing, intuitive visioning, and emergent learning. It is about transformation as well as perpetuation; that involve individual cognition and social interaction, cooperative as well as conflictive. Therefore, it has to include analysis before and after sourcing as well as during negotiation and all of these activities must respond to what may be a demanding environment"[38].

Supply Chain Strategy

The term supply chain strategy is used in the literature to cover a broad range of concepts ranging from tactical to strategic. Sometimes the term is used simply to achieving minimum cost and predictable flow [22]. As a more strategic concept, it can refer to “strategic decisions made by companies including whether to outsource or perform a supply chain function in-house, the location and capacities of production and warehousing facilities, the products to be manufactured or stored at various locations, the modes of transportation to be made available along different shipping legs, and the type of information system to be utilized [44, 46]. A firm must ensure that the supply chain configuration supports its strategic objectives” and “Supply chain strategy” defines not only processes that firm should use but also what role is played by each supply chain entity [9]. Although, Katz is of view that the goals of appropriate SCS factors includes open and effective sharing of information, order coordination, predictable price changes, and effective long-term material planning [28, 43]. Schnetzler define SCS factors as a series of prioritized objectives focused on enhancing logistics success, thus enhance business performance [48]. They have listed general targets of SCS factors as improved quality and a means for increasing customer demands, enhancing delivery reliability, punctuality, improving delivery fill-rates and lead times. Due to increasing flexibility in the face of uncertainty and rapid change, reduce capital investments, as well as transportation, inventory, infrastructure and other supply chain costs [4]. They observed that the availability and exchange of information, and ongoing collaboration are also strategic targets. These authors seem to view SCS factors as the targets and the planning involved meeting goals defined by each sourcing vendors [48]. However, many of these targets might be viewed as tactical objectives. For example, if a particular company was trying to set and achieve many of these targets internally, they would generally fall into a lower level of planning or perhaps viewed as part of a continuous improvement approach.

Supply Chain Strategies

While supply chains may be formed and managed in an ad hoc manner, there are a number of strategies that may be formulated and implemented in an effort to make the chains more efficient and effective. It does not mean that these strategies will certainly bring desired results, but without a coherent strategy among the supply chain partners, it may be difficult to coordinate activities or focus on common goals. For instance, if a particular company’s focus is ‘cost leadership’, it may have quite different demands of its supply chain if its goal is ‘innovation’. Supply chains need to be customized and dependent upon the unique circumstances involved [1] and [9]. This may be dependent upon variables such as product characteristics [14, 46], context and integrative practices [23], product life-cycle stage [10], and contractual issues such as cost distribution and terms [41]. In a literature review of 37 studies published from 1996 to 2003 on international supply chain management, underline the necessity to implement clear and effective SCS factors that encompasses both the supply and customer sides of the chain [3]. This approach emphasizes that all links in the supply chain are a part of the value chain that leads to customer satisfaction.SCS factors can be studied from a number of perspectives, including the portion of the supply chain involved (e.g., upstream or downstream), and design philosophies including focus on the product or the customer, lean and agile in an attempt to integrate all portion of the chain. In this section, the paper classifies and discusses different perspectives found in the literature relative to SCS factors. This classification is made under major research fields of business management, i.e., supply chain management (plan, source, make, deliver, return), organization theory (levels of integration), marketing (strategies based on product, demand, market segmentation, differentiation, cost leadership and market focus), operations perspective (lean, agile, and flexible), financial legal fields (contracts), and strategic perspective. The discussion is followed by taxonomy of SCS factors based on extant literature.

Supply Chain Management Perspective

Supply chain management broadly concerns itself with "plan, source, make, deliver, return" processes (Supply Chain Council, 2012), which provide us with a valid perspective to discuss about SCS factors.

'Plan' – Creating Strategies

Strategic planning in supply chain management is a nascent academic area, and forms the primary subject matter of this paper. Research on strategic planning process in supply chain management is relatively sparse. Schnetzler in a supply chain research, as part of a large two-year international research project called 'ProdChain', found that the methodology provided a structure to the development of SCS factors that was previously lacking in the five case studied companies [47, 49]. Further, they found that a detailed planning process approach was useful in identifying the root causes of problems such as excess inventories, and in tying SCS factors to corporate strategy and objectives. However, there has been little discussion about the relative advantages of a formal strategic planning process and emergent strategies, and whether adoption of different planning methods has any relationship with evolution or the formation process of specific supply chain itself.

'Source' – Upstream or Supplier Side Strategies

Although sourcing or upstream (supplier side) SCS factors are focused on the sourcing functions in the firm and the various tiers of suppliers. There is an overlap with production and downstream SCSFs. For instance, a marketing-related SCS factors (i.e., downstream) allows customers to order customized products may lead to an upstream SCS factors that focuses on suppliers capable of supporting such activity.

'Return' – Reverse Logistics Strategies

When visualized as a 'forward flow', value addition is the main concern of a supply chain. On the other hand, for Reverse Supply Chain Management (RSCM), the objective reversed and recovery of value. Reverse flows of products may not only cover return of damaged products, but also undamaged unsold product and out-of-season or obsolete products. Such a reverse flow scenario may be further confounded by short product life-cycles, upgrade offers, ever-new product ranges, seasonal launches, and leasing of products. Focus on recovery of value has shifted RSCM from the ‘cost of doing business’ sphere into being generator of competitive advantage [37] and [42]. This underlines the importance of strategic planning for RSCM, covering a range of processes, ranging from recovery of ‘end-of-life’ products to remanufacturing.

Product-Based Strategies

The complexity of the operating environment for many firms has increased dramatically due to the interactive nature of the network formed by a wide array of customers and suppliers. This has necessitated the customization of SCSFs. Aitken suggest that this customization must be aligned with product characteristics and at stage of a particular product in its life cycle [1]. Crossovers with resultant efficiencies can be achieved, but each opportunity must be decided upon on a case-by-case basis to ensure that gains on some products are not offset by problems on others [1, 51]. Katz proposes that modularizing, appending and innovating SCSFs be based upon the characteristics of the products involved [28]. Although, modularizing, or providing assemblies to a customer, should increase the insulation of a firm’s competitive position because it will be more closely tied to its customer as a member of a smaller supplier community. In addition, the information sharing required allowing a supplier to provide modularized assemblies more tightly integrates with a supplier and its customer. Modularizing should result in fewer suppliers and therefore fewer transactions, communication nodes and relationships, thereby reducing the complexity of the network and the effort required towards its coordination. In that matter, ‘appending’ refers to a supplier’s intent to add higher margin products or services to their offering for existing customers [28]. Examples include delivery companies who have added logistics consulting to their service offering. Katz suggest that the appending strategy leads to tighter integration among supply chain partners, which in turn, leads to more accurate material planning and fewer unexpected issues such as material shortages or sporadic ordering [28]. However, there are risks to the supply chain when this strategy is employed as it involves firms moving away from their core competencies [50] and diverting resources away from existing products and services. An innovating SCS factors refers to offering new products or services in an attempt to access first-to-market advantages [28].

Demand-Based Strategies

Customers are becoming increasingly demanding in terms of value and this has, at times, been viewed as an issue to be addressed solely by marketing. The changes in demand patterns must be creatively addressed through the whole supply chain [18, 52]. They further suggest that the strategies employed must be specific to the particular environmental and competitive context, and aligned with the end-customers perceptions of required value. They state that demand fulfillment must be aligned with demand creation. This has also been referred as demand chain management [24]. Specifically, Godsell suggest that it is end-customer buying behavior that should be the focus of SCS factors formulation [18]. Further, they posit that end-customers should be segmented based on their buying behavior, and appropriate SCS factors should be developed to target the segments. This clearly means a segmented marketing strategy aligned with a corresponding differentiated set of SCS factors.

Strategic Perspective Analysis

- Though a large number of researchers have focused on different perspectives of SCS factors, these perspectives broadly deal with operational or tactical domains. There is an urgent demand for embracing a strategic perspective for SCS factors, that is, a long-term perspective that helps alignment of long-term costs, risks and incentives for supply chain partners [37]. For long it has been presumed that the interests of supply chain partners are somehow one and the same, or at least aligned, and they generally run without conflict of interests. In the real world, it is not so simple. This strategic aspect calls for immediate attention both from academia and the supply chain management practitioners.

- The proper synchronization of the business strategy aligned with SCSF renders a firm to reach targets. Further, to achieve the SCSF and business strategy, location plays an important role. As an instance, TISCO supply chain strategy in the year 2000 take intervention steps and changed to a robust supply chain. The design resulted in decrease of product delivery time from 3-4 weeks to 2 weeks [39].

- Further, the SCS factors literature contains approaches that set out from strategic to tactical. The predominant focus of the research on the medium to short-term ‘tactical end’ of the continuum can be traced to origin of supply chain management research as a part of production and operations management arena, where the primary focus is on operational or tactical levels of preparation and decision making.

- The capacities of SCS factors are enhanced by adding reference to the internal capability analysis that is common with the deliberate strategy formulation approach. In the case of supply chain, this internal capability analysis would include the capabilities of the various members of the chain and the hoped-for synergistic nature of their interrelationships.

- Continuous improvement is certainly viewed as a component of the firms overall strategy, but the tactical targets and plans perhaps simply be a part of the implementation. In the context of this paper, SCS factors will be defined as the deliberate and/or emergent conceptual framework by which a firm involves its supply chain and supply chain members in efforts to reach its own corporate strategic objectives. This broad definition allows the incorporation of policies that are related to tactical efforts, such as reducing the inventory levels of firm, and strategic actions such as deciding to partner with, and making investments in other firm at distant country to enable access to a new market or resources.

- As the fact presented is based on the current state of research in the area, it covers a continuum of perspectives, ranging from strategic to tactical and up to operational.

- If the strategy is successful, the entire supply chain gets benefitted as a new market niche. There is also the risk of fragmenting SCS factors to an ineffective level as strategies are changed when products change or when product moves along in life cycles.

Conclusion

While strategies are focused on sections of the string that result in improvements to a specific issue, they are restricted in their ability to enhance overall performance. This is because strategies can contribute toward imbalance in supply chain characteristics and consequences. While these SCS factors have the advantage of focusing on the needs of the product or customer, they also have weaknesses. The key is to align the philosophy with the corporate strategy, merchandising strategy and product strategy of a particular product, and be prepared to change the SCS factors if other strategies change. A focus on supply chain integration can create close links between all supply chain partners and the functional departments within the partner organizations toward achieving the supply chain and corporate objectives of the firm. While there is little focus on the literature relative to how SCS factors are formulated. This paper discusses processes that resemble those used to create corporate strategy [48]. From a high-level strategic perspective, most parts of a corporate strategy should be impressed by its relationships with suppliers, distributors and service providers, or in other words by its SCS factors. Thus, it would appear prudent for a firm to ensure that its SCS factors are well aligned with its corporate strategies to enable the diverse portions of corporate strategies to mutually align with one another. However, this paper is valuable to academics in ascertaining the current stage of research in SCS factors and identification of promising future research areas. Although, it is also valuable for practitioners, as it offers a systematic taxonomy of the latest SCS factors that can be immediately applied in SC area. The paper also brings out the need for a strategic view of SCS factors, in contrast to the current focus on tactical and operational aspects of supply chain management.

Bibliography

[1]. Aitken, J., & Todeva, E. (2011). Co-Alignment of Supply Chain Strategies and the Knowledge Outcomes for Buyer-Supplier Network Relationships. *In System Sciences (HICSS), 44th Hawaii International Conference*, 1-12. IEEE 2011.

[2]. Abdel-Malek, Layek, Tarathorn Kullpattaranirun, and Suebsak Nanthavanij. (2005). A Framework for Comparing Outsourcing Strategies in Multi-Layered Supply Chains. *International Journal of Production Economics* 97.3, 318-328.

[3]. Ahlstedt, D., & Hameri, A. P. (2004). Review of Supply Chain Management Research: Practical Business Value And International Aspects. *Supply Chain Forum: an International Journal* (Vol. 5, No. 1, pp. 38-48). BEM-Bordeaux Management School.

[4]. Banham, R. (2008). Forging Stronger**.** *Treasury & risk.*

[5]. Bernstein, M. (2001). New Product Strategy Drives New Supply Chain. *World trade*, 0–3.

[6]. Burgess, Kevin, Prakash J. Singh, and Rana Koroglu. (2006). Supply Chain Management: A Structured Literature Review and Implications for Future Research. *International Journal of Operations & Production Management* 26.7, 703-729.
[7]. Chaharbaghi, K., & Willis, R. (1998). Strategy: The Missing Link between Continuous Revolution and Constant Evolution. *International Journal of Operations & Production Management*, *18*(9/10), 1017-1027.
[8]. Chen, F. T. S., & Chan, H. K. (2010). An AHP Model for Selection of Suppliers in the Fast Changing Fashion Market. *International Journal Advance Manufacturing Technology*, 1195–1207. doi:10.1007/s00170-010-2683-6.
[9]. Chopra, Sunil, and Peter Meindl. (2007). Supply Chain Management. Strategy, Planning & Operation. *Gabler*, 265-275.
[10]. Childerhouse, P., & Towill, D. R. (2002). Enabling Seamless Market-Orientated Supply Chains.
[11]. Cox, Nancy M. (2005). Does Size Matter? Small Firms' Use of E-Business Tools in the Supply Chain. *Electronic Markets* 15.2, 94-105.
[12]. Eric Harris, C. M. (2011). Supply Chain Strategy: Convincing The Board That Investment Can Add Real Value. *Focus*.
[13]. Falshaw, J. R., Glaister, K. W., & Tatoglu, E. (2006). Evidence on Formal Strategic Planning and Company Performance. *Management Decision*, *44*(1), 9-30.
[14]. Fisher, Marshall L. (1997). What is the Right Supply Chain for Your Product? *Harvard Business Review*, 75, 105-117.
[15]. Frohlich, Markham T., and Roy Westbrook. (2001). Arcs of Integration: An International Study of Supply Chain Strategies. *Journal of Operations Management* 19.2, 185-200.
[16]. Gardner, J. T., & Cooper, M. C. (2003).Supply Chain Reviews and Agendas. *Journal of Business Logistics*, Vol. 24. No. 2. 2003 37.
[17]. Gill, B., & Logistics, T. (2012). The Principles of Supply Chain Strategy. *Total Logistics*, 1–6.
[18]. Godsell, J., Storey, J., Emberson, C., & Harrison, A. (2006). Supply Chain Management: Theory, Practice And Future Challenges. *International Journal of Operations & Production Management*, *26*(7), 754-774.
[19]. Golicic, S. (2008). Supply Chain Strategy For Nascent Firms In Emerging Market. *Journal of Business Logistics*, 29 (1), 75–92.
[20]. Gunasekaran, A., Bayraktar, E., Lenny Koh, S. C., Sari, K., & Tatoglu, E. (2008). The Role of Forecasting on Bullwhip Effect For e - SCM Applications. *International Journal of Production Economics*, *113*(1), 193-204.
[21]. Golicic, S. L., & Sebastiao, H. J. (2011). Supply Chain Strategy in Nascent Markets : The Role Of Supply Chain Development In The Commercialization Process. *Journal of Business Logistics*, 32 (3), 254–273.
[22]. Hauguel, A., & Jackson, K. (2004). Linking SCOR Planning Practices to Supply Chain Performance: An Exploratory Study. *International Journal of Operations & Production Management*, *24*(12), 1192-1218.
[23]. Ha, a. Y., & Tong, S. (2008). Contracting and Information Sharing Under Supply Chain Competition. *Management Science*, 54 (4), 701–715. doi:10.1287/mnsc.1070.0795
[24]. Heikkilä, J. (2002). From Supply To Demand Chain Management: Efficiency and Customer Satisfaction. *Journal of Operations Management*, *20*(6), 747-767.
[25]. Hillman, G. (2012). Warehousing – the heart of your supply chain. *www.scemagazine.com*.
[26]. Ivanov, D. (2010). An adaptive framework for aligning (re) planning decisions on supply chain strategy, design, tactics, and operations. *International Journal of Production Research*, 48 (13), 3999–4017. doi:10.1080/00207540902893417
[27]. Jackson, N., & President, V. (2001). Outward-looking supply-chain strategy. *European Business Journal*, 13 (3), 113–119.
[28]. Katz, J. P., Pagell, M. D., & Bloodgood, J. M. (2003). Strategies of supply communities. *Supply Chain Management: An International Journal*, *8*(4), 291-302.
[29]. Knight, P. (2005). How IBM Linked Its Supply Chain to Its Business Strategy. *Frontline Solutions*, (June), 16–18.
[30]. Lassar, Walfried M., and Jeffrey L. Kerr. (1996). Strategy and Control in Supplier–Distributor Relationships: An Agency Perspective. *Strategic Management Journal* 17.8, 613-632.
[31]. Lee, H. L. (2002). Aligning Supply Chain Strategies with Product Uncertainties. *California Management Review*, 44.

[32]. Li, J., & Sheng, Z. (2011). A Multi-Agent Model For The Reasoning Of Uncertainty Information In Supply Chains. *International Journal of Production Research*, 49 (19), 5737–5753. doi:10.1080/00207543.2010.524257
[33]. Lummus, Rhonda R., Leslie K. Duclos, and Robert J. Vokurka. (2003). Supply Chain Flexibility: Building a New Model. *Global Journal of Flexible Systems Management* 4.4, 1-13.
[34]. Mintzbcrg, Henry, and Joseph Lampel. (1999). Reflecting on the Strategy Process. *Sloan Management Review*.
[35]. M M Sodhi. (2003). How to do Supply Chain Planning. *MIT Sloan Management Review*.
[36]. Madhani, P. M. (2010). SCM and Marketing Management: Mutual Integration. SCMS, *Journal of Indian Management*, 16–25.
[37]. Mann, H., Cao, Y., Jit, I., & Mann, S. (2011). Strategy Implementation Tool: Supply Chain Contract. *IUP Journal of Operations Management,*, VIII (2), 34–49.
[38]. Mentzer, John T., et al. (2001). Defining Supply Chain Management. *Journal of Business logistics* 22 vol. 2, 1-25.
[39]. Mitra, A., & Bhardwaj, S. (2010). Alignment of Supply Chain Strategy with Business Strategy. *IUP Journal of Operations Management*.
[40]. Parmar, D., Wu, T., Callarman, T., Fowler, J., & Wolfe, P. (2010). A Clustering Algorithm for Supplier Base Management. *International Journal of Production Research*, 48 (13), 3803–3821. doi:10.1080/00207540902942891
[41]. Peleg, M. E., & Whang, S. (2002). E-Business and Supply Chain Management: An Overview and Framework. *Production and Operations management*, *11*(4), 413-423.
[42]. Price-White, C. (2002). Check it out and back in again. *Frontline Solutions*, 24-29.
[43]. Richardson, H. L. (2005). Think Supply Chain Security Think Strategy. *Logistics Today*.
[44]. Rodrigues, A. M., & Lynch, D. F. (2004). A Review in Supply Chain Management, *Journal of Business Logistics*, Vol. 25. No. 2, 26-25.
[45]. Rose, W., Singh Mann, I. J., & Rose, S. (2012). A Strategic Perspective and Taxonomy of Supply Chain Strategies. *IUP Journal of Operations Management*, 11(3).
[46]. Scott, C., & Westbrook, R. (1991). New Strategic Tools For Supply Chain Management. *International Journal of Physical Distribution & Logistics Management*, *21*(1), 23-33.
[47]. Schäffer, U., & Willauer, B. (2003). Strategic Planning As a Learning Process. *Schmalenbach Business Review (SBR)*, *55*(2).
[48]. Schnetzler, M. J., Sennheiser, A., & Schönsleben, P. (2007). A Decomposition-Based Approach For The Development Of A Supply Chain Strategy. *International Journal of Production Economics*, *105*(1), 21-42.
[49]. Trends, G. (2007). Yesterday's Supply Chain Networks are Costing Businesses Today. *www.scemagazine.com*.
[50]. Vickery, (1991). Inventory Management: Is There a Knock-on Effect? *International Journal of Production Economics*, 93-94, 129-138.
[51]. Visvanathan, C., & Kumar, S. (1999). Issues For Better Implementation Of Cleaner Production In Asian Small And Medium Industries. *Journal of Cleaner Production*, 7, 127-134.
[52]. Wadhwa, S., Kanda, A., & Bhoon, K. S. (2006). Impact of Supply Chain Collaboration on Customer Service Level and Working Capital. *Global Journal Flexible Systems Management*.

MODELING OF AGILE MANUFACTURING SYSTEM

Priyank Srivastava[1], N.C.C.E, Panipat, India (priyank.cim@gmail.com),
Dr. Dinesh Khanduja, N.I.T, Kurukshetra, Haryana, India (dineshkhanduja@yahoo.com),
Dr. V.P Agrawal, T.I.E.T, Patiala, Punjab, India.

Abstract

Researchers have been using different methodologies for modeling, analyzing and designing different aspects of manufacturing system. During past two decades various modeling methodologies have evolved.

We know that most of the manufacturing systems are complex. Most of the modeling methodologies are not well defined and therefore, open to misinterpretations and inconsistencies. Modeling needs a method which is simple and able to support different levels of abstraction.
This paper summarizes modeling methodologies which have been used for modeling Agile Manufacturing System. This paper also introduces a new modeling methodology, Graph Theory for modeling of Agile Manufacturing System.

Keywords: Agile Manufacturing, Agility, Modeling, Analysis

Introduction

Over the past two decades a powerful drive by enterprises and academic institutions has boosted the development and adoption of new manufacturing initiatives to enhance business in an increasingly competitive market. The need for a method of rapidly and cost-effectively developing products, production facilities and supporting software, including design, process planning and shop poor control system has led to the concept of Agile Manufacturing. Modeling and analysis of AGMS is difficult task as system is complex. Actually all the manufacturing systems are complex. The dynamic nature of the manufacturing environment greatly increases the number of decisions that need to be made and system integration makes it difficult to predict the effect of a decision on future system performance. Quantitative analysis of system is under development. It may be limited to productivity behavior of individual component which can be equipment, process etc. Metrics for measuring and analyzing the productivity of manufacturing operation from equipment level or from individual level to system level are of increasing importance to the companies seeking to continuously optimize the existing component. There is need of systematic, coherent technology for productivity measurement and analysis at system level. As already stated today's manufacturing systems have become more and more complex and advanced. It involves higher levels of automation and integration, greater demands on performance, and various forms of human supervisory controls. Understanding a manufacturing system, now vows for understanding of the complex layout and interactions between the components of system. Word modeling means different to different people, even within the context of advanced manufacturing technologies. System analysis can be defined as a stage in the development cycle in which real system problems are studied and examined to gain an understanding of system requirement. During the last two decades several modeling methods and methodology have been used for modeling, analyzing and designing different aspects of manufacturing system.

[1] Priyank Srivastava, Department of Mechanical Engineering
N.C.C.E, Israna, Panipat. +919896840098, priyank.cim@gmail.com

.

This paper summarizes different aspects of Agile Manufacturing System and techniques of modeling and analysis of the system.

Agile Manufacturing

The term 'Agile Manufacturing' appeared at the beginning of the 90s.In 1991, a group of more than 150 industry executives participated in study. Their effort culminated in a two volume report titled "21st Century Manufacturing Enterprise Strategy". As a result, the Agile Manufacturing Enterprise Forum (AMEF), affiliated with Iacocca Institute at Lehigh University was formed and the concept of agile manufacturing was introduced.

Since the evolution of buzz word "Agile Manufacturing", academicians and industry executives have viewed it in different ways. It can be defined as, a manufacturing system with capabilities (soft technologies, human resource, educated management, information) to meet rapidly changing needs of the market place (speed, flexibility, customers, competitors, suppliers, infrastructure and responsiveness) [1]. One given by Goldman defines agility as an overall strategy focused on thriving in an unpredictable environment. Focusing on the individual customer, agile competition has evolved from unilateral producer centered customer responsive companies inspired by lean manufacturing refinement of mass production to interactive producer customer relationships [2]. Other definition may be , Agile manufacturing system is a system that is capable of operating profitably in a competitive environment of continually and unpredictably changing customer opportunities [3].To understand Agile manufacturing, we should consider the Iacocca Institute report, which makes three key points:

1. A new competitive environment is emerging, which is acting as a driving force for change in manufacturing.
2. Competitive advantage will accrue to those enterprises that develop the capability to rapidly respond to the demand for high quality, highly customized product.

To achieve the agility that is required to respond to these driving forces and to develop required capability, it is necessary to integrate flexible technologies with highly skilled, knowledgeable, motivated and empowered workforce.

It is evident that reconfigurability is an important aspect of AGMS. Therefore rapid modeling capabilities are required.

Agile Manufacturing Structure

Agile Manufacturing can be considered as a structure with in which every company can develop its own business strategies and products. The structure is supported by three primary resources: innovative management structure and organization, a skill base of knowledgeable and empowered people and flexible and intelligent technologies. Agility is achieved through the integration of these resources into a coordinated and interdependent system. The four key dimensions of AM system can be listed into enabling subsystems.

The four key dimensions of AM system can be listed into enabling subsystems. These key dimensions are [4]:

1. People

2. Strategy
3. Technology
4. Strategy

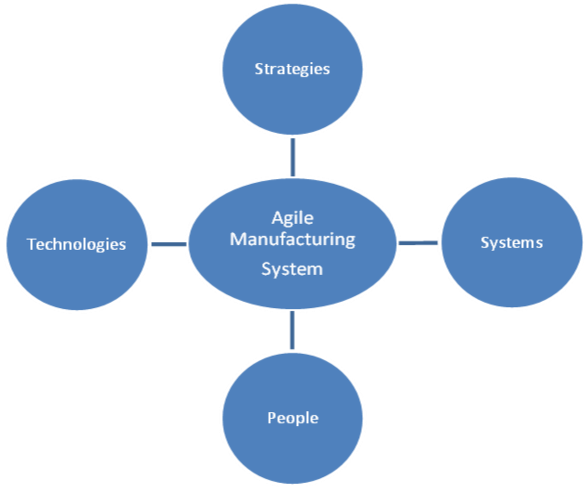


Figure 1: Agile Manufacturing Structure

A more detailed structural diagram for these components may be:

People	Technology	Systems	Strategies
• Flexible Work Force • Multi Lingual • Empowered Workers • Top Management Support	• Rapid Hardware Changeover by robots, Flexible part feeders, Flexible Fixturing, Modular Grippers • Multimedia, Information Technology(EDI, Electronic Commerce) • Flexible manufacturing Cells	• MRPII • Internet • CAD/CAE • ERP • Kanban • CIM • ABC/ABM • JIT	• Concurrent Engineering • Virtual Enterprise • Rapid Partner Formation • Strategic Alliances • Physically Distributed Manufacturing Systems

Figure 2: Component diagram

Agile manufacturing can be seen as integration of Lean Manufacturing and Flexible Manufacturing systems. At this point it important to understand that tools of Lean Manufacturing and Flexible Manufacturing systems can be used for successful implementation of Agile Manufacturing. Also Agility covers different areas of manufacturing, from management to shop floor. It is a top down enterprise wide effort. The agile manufacturing company needs to integrate design, engineering, and manufacturing with marketing and sales, which can only be achieved with information and communication technology (ICT).

From the above literature it is quite obvious that if we are able to identify key components and subcomponents of a manufacturing system, we can do modeling and subsequently its analysis.

Current Modeling Methodology

There are some important modeling methodologies used in manufacturing system such as queuing network, activity cycle diagram, IDEF0, IDEF3, Petri Net, object-oriented. Analysis and design of complex systems often involve two kinds of uncertainty: randomness and fuzziness [5].Randomness refers to describing the behavior of the parameters by using probability distribution functions. In other words, the randomness models stochastic variability. Fuzziness models measurement imprecision due to linguistic structure or incomplete information.

Here in this paper more stress is given on modeling technique by Queuing, PetriNets and Graph theoretic Method.

Activity Cycle Diagram

Activity cycle diagram represents the changes of system status in graphical manner, to taking advantages of understanding and analyzing manufacturing system. Activity cycle diagrams (ACD) are the natural way to represent the activity paradigm of discrete event simulation. They are particularly appropriate for problems with a strong queuing structure. Their utility however is not limited to queuing systems and the logical internal structure of the method, makes ACD's a natural choice for model design specification. Activity Cycle Diagram is representation of the states that an entity goes through. ACD are one way to modeling interactions of the entities and are useful particularly with a strong queue structure [6].The relationship of conditions and events in the discrete events dynamic system is not represented by the activity cycle diagram.

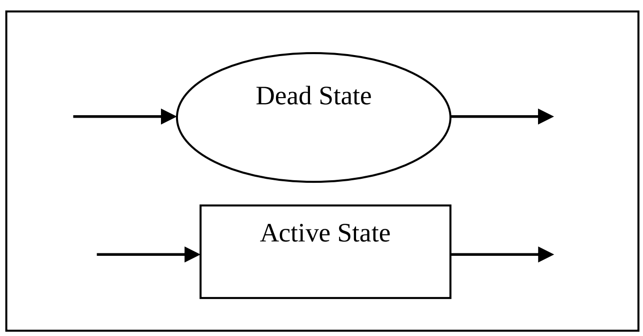


Figure 3: Symbols of Activity Cycle Diagram [7]

IDEF0

IDEF0 was first developed by the US Air Force on its integrated computer aided manufacturing (ICAM) programme [8]. IDEF0 is modeling methodology for designing and documenting hierarchic, layered, modular system. The IDEF0 methodology has been widely used. The building block of this modeling approach is the activity box. The activity box defines a specific activity in the organization that is being modeled. The activity may be a decision-making or information converting activity or a material-converting activity. Four types of arrows are associated with an activity box. They are input, output, control, and mechanism, respectively. Inputs are items (material, information) that are transformed by the activity. Outputs are the results of the activity acting on the inputs. A control is a condition that governs the performance of the activity.

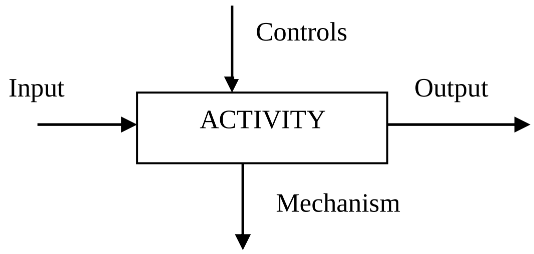


Figure 4: IDEFO Diagram

A mechanism is the means by which an activity is realized. The activity box and the four entities provide a concise expression: an input is transformed into an output by an activity performed by a mechanism and governed by a control.
The specific activity, its inputs, outputs, mechanisms, and controls must be defined for the situation being modeled. An activity box can be further broken down into several sub-activities. Nevertheless, the disadvantage of IDEF0 methodology is its static nature, that is, it can only represent the static behavior of a system. It does not explicitly represent the real-time status and sequences of activities.

IDEF3

IDEF3 diagram is a process-flow modeling method describing how activities work together to form a process [9] [10]. IDEF3 diagram identifies the behavior of the system. It builds structured descriptions about "what" a system actually does and "how" activities work together to form a process. There are two description modes: "Process Flow Diagram and Object State Transition Network".

An IDEF3 process flow describes a process and the relations that exist between processes. The activities of the process appear as labeled boxes. The term for elements represented by boxes (activities, processes, events, operations, and procedure) is a Unit Of Behavior (UOB). The boxes are connected by arrows that define the logical flows. The arrows are the same as in IDEF0 diagrams, but there are no bottom arrows (mechanisms). There are also the smaller boxes that define logic junctions: AND (&), OR (O), and exclusive OR (X). Logic junctions could present asynchronous or synchronous behavior among UOBs (they present inputs that proceed and outputs that follow the UOB). Each UOB can be decomposed and can be associated with a description of the objects and their relations, called elaboration.

The Object State Transition Network summarizes all the transitions an object may undergo throughout a particular process that is very useful for simulation modeling.

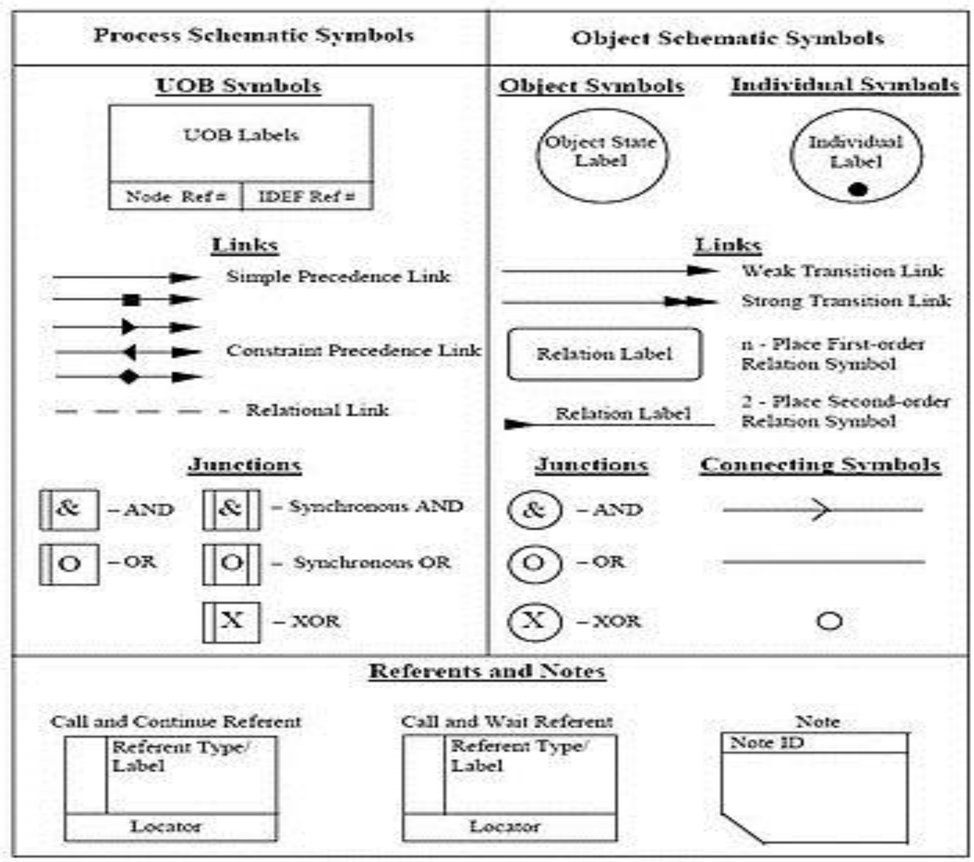


Figure 5: IDEF3 [11]

Petri Net

Petri nets (PN) introduced by Carl Adam Petri in his PhD thesis in 1962, are a formal graphical and mathematical tool which can be used for analyzing complex systems which are characterized by synchronous, parallel, simultaneous, distributed, resource sharing, non deterministic/ stochastic system. A Petri net is a directed bipartite graph, in which the nodes represent a transition (signified by bars) and places (signified by circles) [12]. The directed arcs describe which places are pre and/or post conditions for which transitions (signified by the arrows) occur.

A Petri net consists of a number of places and transitions with tokens distributed over places [13]. When every input place of a transition contains a token, the transition is enabled and may fire. The result of firing a transition is that a token from every input place is consumed and a token is placed into every output place. Petri nets have been generalized by allowing multiple tokens arcs, inhibitor arcs, place capacity, colored token etc [14]. Formally, a five tupled PN is defined as given below.

PN = (P, T, I, O, Mo); where:

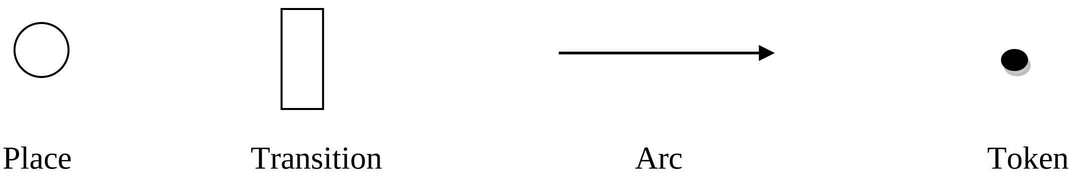
P = {p1, p2, ………,pm} is a finite set of places, $m \geq 0$.

T = {t1, t2, ……….. ,tn} is a finite set of transitions, $n \geq 0$, $P \cup T \neq \varphi$ and $P \cap T = \varphi$.

I : $(P \times T) \rightarrow N$ is an input function that defines directed arcs from places to transitions, where N is a set of non-negative integer i.e. N = {0, 1, 2, ...}.

O : $(T \times P) \rightarrow N$ is an output function which defines directed arcs from transitions to places.

Mo: $P \rightarrow N$ is the initial marking. It is a | P | dimensional vector with M0(p) being the initial token count of a place p.



Petri nets and their variations, such as colored Petri nets and timed Petri nets, have been widely applied to model, control and analyze the system's dynamic behaviors due to the characteristics of the graphical representation and mathematical analysis of control logic.
Petri nets are powerful tool in describing event driven systems these systems may be synchronous, contains sequential and concurrent operations, and involve conflicts, mutual exclusion and non-determinism, merging and priorities. These characteristics are represented using a set of simple constructs:

1. **Sequential Execution**: In Figure 6, transition t2 can fire only after the firing of t1. This impose the precedence of constraints "t2 after t1." Such precedence constraints are typical of execution parts in an AMS. Also, this construct models the casual relationship among activities.

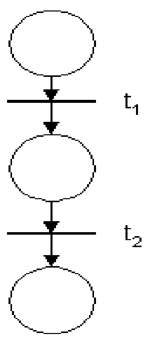


Figure 6

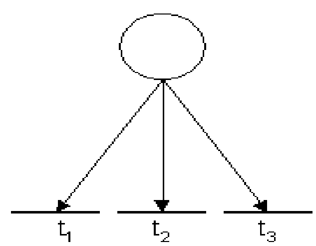


Figure 7

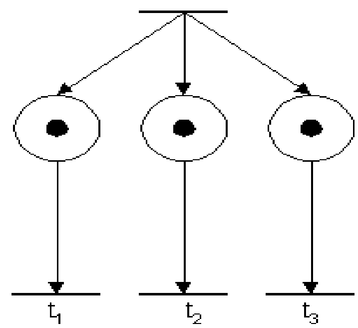


Figure 8

2. **Conflic**t: Transitions t1, t2 and t3 are in conflict in Figure 7. All are enabled but the firing of any leads to the disabling of the other transitions. Such a situation will arise, for example, when a machine has to choose among part types or a part has to choose among several machines. The resulting conflict may be resolved in a purely non-deterministic way or in a probabilistic way, by assigning appropriate probabilities to the conflicting transitions.

3. **Concurrency:** In Figure 8, the transition t1, t2, and t3 are concurrent. Concurrency is an important attribute of AMS interactions. A necessary condition for transitions to be concurrent is the existence of a forking transition that deposits a token in two or more output places.

4. **Synchronization:** Often, parts in a AMS wait for resources and resources will wait for appropriate parts to arrive. The resulting synchronization of activities can be captured by transitions of the type shown in Figure 9. Here, t1 will be enabled only when a token arrives into the place currently without token. The arrival of a token into this place could be the result of possibly complex sequence of operations elsewhere in the rest of the Petri net model. Essentially transition t1 models the joining operation.

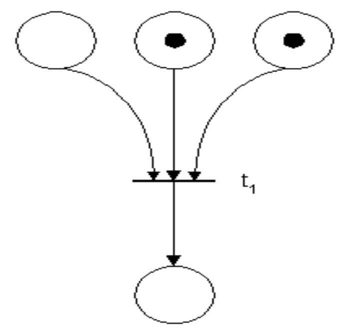


Figure 9

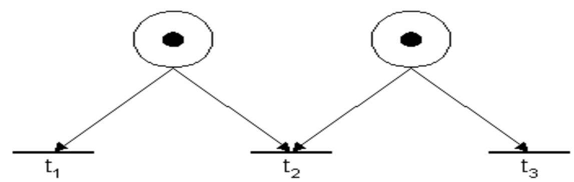


Figure 10

5. **Merging:** When parts from several streams arrive for service at the same machine the resulting situation can be depicted as in Figure 10. Another example is the arrival of several parts of several sources to a centralized warehouse.

6. **Confusion:** A situation where concurrency and conflicts co-exist as in Figure 11. Both t1 and t3 are concurrent while t1 and t2 are in conflict, and t2 and t3 are also in conflict.

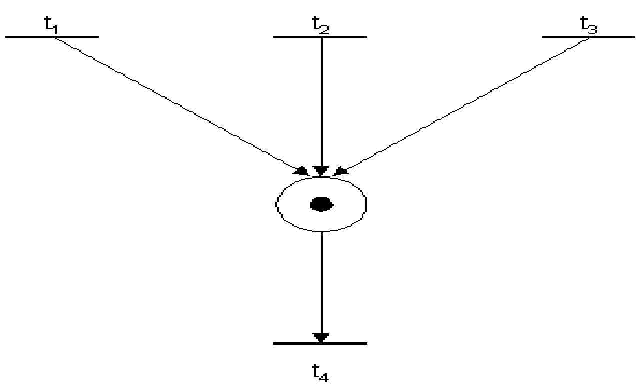


Figure 11

Petri net based models (ref) for design and performance evaluation of manufacturing system are useful for addressing integration issues. Under this model, different model parts are merged to create a complete model of manufacturing system. A typical petri net diagram have been shown in figure 6

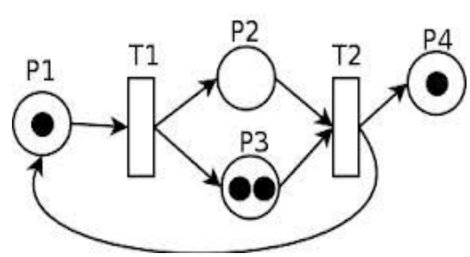


Figure 12: Petri net

P_1, P_2, P_3, P4 =Place, $T_{1,}$ T_2= Transition

Petri nets and its type have been used to model life cycle of production system [15], production process [16], manufacturing system [17], and flexible manufacturing system [18], supply chain inventory [19], modeling and analysis of tool sharing system [20], scheduling of process industry using hybrid petrinets [21], shop floor scheduling and control in FMS [22][23], modeling logistics network using coloured petrinets [24], deadlock preventions from FMS [25] [26] and automated system [27], simulation analysis [28]and cyclic scheduling [29] of FMS.

Queuing Network

Queuing network models have emerged as the most widely studied analytical models of manufacturing systems. Most network models used are of the Jackson-Gordon-Newell type that possesses a product form equilibrium (steady-state) distribution. In many instances, the assumption of exponential service times is invalid and can lead to misleading results.

Queuing Theory in manufacturing process involves the study and simulation of models to predict the behavior of a manufacturing process which attempt to provide services for randomly arising demands in manufacturing work station. Engineers have applied results of queuing theory to show how cycle time is related to utilization of machine and statistic of inter-arrival time and service. These analyses provide means and predicting the average cycle time in steady state condition. By knowing the cycle time, work in progress can be determined.

Queuing Theory was developed to provide models to predict behavior of systems that attempt to provide service for randomly arising and not unnaturally demand. The earliest problems studied were those of telephone traffic congestion (Syski 1986)[30]. Mital (2010)[31] carried out queuing analysis in his case study which provides a basis for estimating medical staff size and number of beds, which are two very important resources for outpatient and inpatient services in a large hospital, and all other hospital resources in one way or another depend on them.

As discussed by Hideaki Takagi (1993)[32], presented a new approach in designing of semiconductor equipment based Queuing Theory to reduce cycle time. They applied Queuing Theory in calculating optimum batch size for their processing equipment. They have showed the batch size required to achieve short time under different production situations. As presented by Cooper (2000)[33], Hoover and Bartlett have also applied the results of Queuing Theory to show how the cycle time is related to large and small production. In large production, cycle time is important to determine the amount of work in progress, and it can be determined by using Queuing models. However, their study had limitation which it can be applied only in large production. Their simulation has showed greater fluctuations in cycle time compared to the value predicated by Queuing Theory.

Ullah (2011)[34] presented a comparison between Petri net (PN) and queuing network tools to determine the optimum values for flexible manufacturing system (FMS) measures of performance. A queuing theory was presented by Tsarouhas (2011)[35] to calculate the total processing time for the processing time per pizza line at workstation in food production lines. McGuire (2010)[36] proposed and tested a model which defines the psychological processes that mediate the relationship between perceived wait duration (PWD) and satisfaction.. Mehmood and Lu (2011)[37] reported that Markov chains and queuing theory are widely used analysis, optimization and decision-making tools in many areas of science and engineering. Real life systems could be modeled and analyzed for their steady-state and time-dependent behavior. Gudmundsson and Goldberg (2007)[38] developed a model to study a commercially available industrial part feeder that uses an industrial robot arm and computer vision system. The problem of optimizing belt speeds and hence throughput of this feeder are addressed that avoid starvation, where no parts are visible to the camera and saturation, where too many parts prevent part pose detection or grasping. There have been a number of books focusing on the application of queuing theory on manufacturing systems, such as those by Papadopoulos et al. (2013)[39], Guy et al. (1997)[40], Gershwin (1994)[41], Yao et al. (1994)[42], Buzacott and Shanthikumar (1993)[43] and Narahari (1992)[43].

Koo et al. (1995)[45] proposed a manufacturing system modeling approach using computer spreadsheet software, in which a static capacity planning model and stochastic queuing model are integrated. Most stochastic performance measures such as throughput time or work in process as well as deterministic measures can be captured directly from the proposed model. Several special manufacturing features such as machine breakdown and batch production can be included in the model. The performance of the proposed model was evaluated by comparing its results with those obtained from other existing approaches. Their finding for this comparison stated that the maximum allowed a relative error was 10%. Sukhotua and Peters (2005)[46] discussed a number of approaches in the facility design for modelling material flow using queuing networks. In these approaches, Poisson arrival or Markovian job routing assumptions were used. However, for many manufacturing environments, these assumptions lead to an inaccurate estimation of the material handling system's performance and thus lead to poor facility designs. The proposed modeling approach has showed to provide more accurate results than previous methods used in facility design based on numerical comparisons with results from discrete-event simulation. Marcheta et al. (2012)[47] presented an analytical model to estimate the performances (the transaction cycle time and waiting times) for product tote movement. The model is based on an open queuing network approach. The model effectiveness in performance estimation was validated through simulation.

The use of simulation in improving cycle time had been discussed by many researchers like Sivakumar and Chong 2001[48], Domaschke and Brown 1998[49], Wang et al. 1993, Toh et al. (1995)[50]. Based on their study on many manufacturing systems, they concluded that simulation can improve cycle time varies from 15% to 45%. The use of simulation is a powerful technique that helps decision maker to solve difficult problems in the design, control, or improvement of complex systems to reduce cost, improve quality or productivity, and shorten time-to-market. However, the technology is still underutilized due to several reasons: (1) simulation modeling is a time-consuming and knowledge-intense process that requires knowledge not only about simulation but also application and implementation tools; (2) most simulation models developed with current technology are customized “rigid” ones that cannot be reused or easily adapted to other similar problems; and (3) transforming related knowledge and information from application domain to simulation is a unstructured or ill-defined process dependent on the skill and experience of individual modelers (Zhou et al. 2010)[51].

Graph Theory and Matrix Approach

A graph G = (V, E) consists of a set of objects V = {v1, v2, ….} called vertices or nodes, and another set E = {e1, e2, ….}, of which the elements are called edges, such that each edge e_k is identified with a pair of vertices. The vertices v_i and v_j associated with edge e_k are called the end vertices of e_k.

The most common representation of a graph is by means of a diagram, in which the vertices are represented by small points or circles, and each edge as a line segment joining its end vertices. The application of graph theory was known centuries ago, when the longstanding problem of the Konigsberg bridge was solved by Leonhard Euler in 1736 by means of a graph.

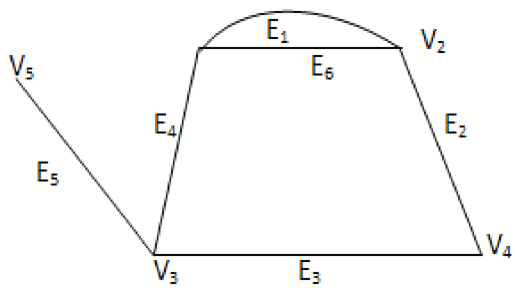


Figure 13: Graph (5V, 6E)

Here V is vertices and E is edges.

Since then, graph theory has proved its mettle in various fields of science and technology such as physics, chemistry, mathematics, communication science, computer technology, electrical engineering, sociology, economics, operations research, linguistics, internet, *etc.* Graph theory has served an important purpose in the modeling of systems, network analysis, functional representation, conceptual modeling, diagnosis, *etc.* Graph theory is not only effective in dealing with the structure (physical or abstract) of the system, explicitly or implicitly, but also useful in handling problems of structural relationship. The theory is intimately related to many branches of mathematics including group theory, matrix theory, numerical analysis, probability, topology, and combinatorics. The advanced theory of graphs and their applications are well documented in several studies by researchers [52] [53].

The matrix which is invariantly used for the modeling a system by a graph theoretic method is incidence and adjacency matrix of a digraph. A digraph is nothing but a directed graph.

The incidence matrix of a digraph with n vertices, e edges, and no self loop is a n by n matrix A= [a_{ij}], whose rows corresponds to vertices and columns corresponds to edges such that

a_{ij} = 1, if the j^{th} edge is incident out of i^{th} vertex,

= -1, if the j^{th} edge is incident into of i^{th} vertex,

= 0, if the j^{th} edge is not incident on of i^{th} vertex,

Another important matrix used in the representation of digraph is the adjacency matrix. The adjacency matrix X = [x_{ij}] of the digraph G is a n by n (0,1) matrix whose element

x_{ij} = 1, if there is an edge directed from i^{th} vertex to j^{th} vertex,

= 0, otherwise.

A digraph is a visual representation so it helps in analysis to a limited extent only. For the establishment of the expression for factors affecting the performance of the system under study, the digraph is represented in matrix form, which is a convenient in computer processing also. Matrix representation of the digraph for evaluation gives one to one representation. This matrix is called performance evaluation matrix or variable permanent matrix.

Both digraph and matrix representations are not unique as they are changed by changing the numbering of nodes. To develop a unique representation, a permanent function of the matrix VPMDP is used. Permanent is a standard matrix function and is used in combinatorial mathematics [54]. The permanent function is obtained in a similar manner as its determinant but with all signs positive.

Some of areas and the potentials of the graph theory and matrix approach in Manufacturing Systems and Process is in Failure Analysis, Selection of Manufacturing System or Process, Modeling, Analysis and evaluation of Manufacturing Systems and Processes.

Graph theoretic Method has been successfully used to model systems and the interaction of the sub components of the systems. GTM has been used for wear evaluation analysis [55], reliability analysis and evaluation of system [56], TQM analysis and evaluation [57].

Some of the areas where GTM techniques has been successfully used:

Table 1

SNo	Author	Problem	Attribute
1	V.P.Agrawal and J.S.Rao [58], 1989	Identification and isomorphism of kinematic chains.	Links and joints.
2	O.P.Gandhi and V.P.Agrawal [56], 1991	Reliability analysis of mechanical and hydraulic system.	Reservoir, pump, pressure relief valve, metering out unit, control valve, and actuator.
3	O.P.Gandhi and V.P.Agrawal [55], 1994	Wear function characterizes and compares the subsystems' wear of different types.	Macro and micro geometry, loads, motions, environmental aspects, interface media, stress state , lubrication condition, formation of oxide, surface temperature, and friction
4	M.F.Wani and O.P.Gandhi [59], 1999	Evaluation of maintainability of mechanical systems through maintainability index	Accessibility, simplicity, disassembly, standardization, identification, test equipment, and documentation. diagnosability, system environment, tribo-concepts, ergonomics, tools and
5	Sandeep Grover et al [57], 2004	TQM evaluation of industries.	Behavioral factors, Non-behavioral factors, use of tools and technique/methodologies, human factors, and functional areas.
6	Sushma Kulkarni [60], 2005	Performance evaluation of TQM in Indian industries.	Infrastructure, top management support, strategic planning, employee empowerment, and customer satisfaction
7	Sandeep Grover et al [61], 2005	Performance measurement of human resource in TQM environment	Employee, employer, customer, and Supplier.
8	R.K.Garg et al [62], 2006	Evaluation, Comparison and Selection of power plants.	Capital cost, electricity generation cost, and plan load factor
9	R.T Duari Prabhakaran et al [63], 2006	Structural Modeling and Analysis of Composite product system.	Resin system, reinforcement system, processing equipment, tooling system, and product design.
10	R.T Duari Prabhakaran et al [64], 2006	Design of composite products through concurrent design approach..	Design for minimal weight, DFE, DFM, DFR, design for material inserts, and design for quality.
11	R.T.Durai Prabhakaran et al [65], 2006	Modeling and analysis of Polymer composite products.	Resin system, reinforcement system, processing equipment, tooling system, and product design.
12	R.Venkata Rao [66], 2006	Evaluation of flexible Manufacturing system	Total cost involved, floor space required, number of employees, throughput time, product mix flexibility, and routing flexibility.
12	Mohd Nishat Faisal et al [67] 2007	Mitigation of risk in supply chain environment.	Information sharing, supply chain agility, aligning incentives,

			strategic risk planning, risk sharing supply chain, etc
14	Variinder Singh [68], 2007	Structural Modeling and Integrative Analysis of Manufacturing Systems using Graph Theoretic Approach	Elements constituting the manufacturing plant and the interactions between them have been identified through a literature survey and have been represented by graph-based model. The matrix models and the variable permanent function models are developed for carrying out decomposition, characterization and the total analysis

Conclusion

This paper has demonstrated various techniques for modeling of a manufacturing system. We can see that authors have applied modeling techniques into various application, may be scheduling and control of FMS or Automated system by identifying components of system, or modeling of system (may be manufacturing or facility design or material flow system) by markow chains again by defining system, sub-system and sub-sub-system.

As already pointed out in paper that Agile Manufacturing is integration of Lean Manufacturing and Flexible manufacturing, therefore by assuming this paradigm as system and identifying its sub-system , modeling of Agile Manufacturing as a system can be done by any one of the techniques.

Bibliography:

1 Iacocca Institute 1991. “21st Century Manufacturing Strategy: An Industry led view”, Vol 1&2, Lehigh University, Bethlehem, PA.

2 Goldman, S. L., Nagel, R. N. and Preiss, K. (1994). “Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer” (New York: Van Nostrand Reinhold).

3 Goldman, S. L., Nagel, R. N. and Preiss, K. (1995). “Agile Competitors and Virtual Organizations—Measuring Agility and Infrastructure for Agility”, Van Nostrand Reinhold, International Thomas Publishing, London.

4 Kidd, P. T. (1994). “Agile Manufacturing: Forging New Frontiers” (Reading, Mass: Addison Wesley)

5 Liangjian, Hu., Rangquan, Wu., Shihaung, Shao. (2002). “Analysis of dynamical systems whose inputs are fuzzy stochastic processes”, Fuzzy Sets and systems, V. 129n.1, p 111-118.

6 Paul, R.J. (1993). “Activity Cycle diagram and the three phase method”, proceeding of the 1993 Winter Simulation Conference, NY: IEEE, 123-131.

7 Pidd, M. (1998). “Computer Simulation in Management Service”, Wiley UK, Chichester.

8 Marca, D.A. and McGowan, C.L. (1988). "SADT", McGraw-Hill, New York.

9 deWitte, P. and Pourteau, C., (1997). "IDEF enterprise engineering methodologies support simulation", Information Technology for Manufacturing Managers, 70-75.

10 Mayer, R.J., Benjamin, P.C., Carawaz, B.E. and Painter, M.K., 1999. "A Framework and Suite of Method for Business Process Reengineering", http://www.idef.com/articles/framework.

11 Richard J. Mayer et al. (1993) Information Integration for Concurrent Engineering (IICE): IDEF3 Process Description Capture Method Report. Logistics Research Division, Wright-Patterson AFB, OH 45433

12 http://lib.bioinfo.pl/courses/view/471

13 Daniela Coman, Adela Ionescu, Mihaela Florescu. "Manufacturing System Modeling Using Petri Nets", International Conference on Economic Engineering and Manufacturing Systems, Brasov, 26 – 27 November 2009, pg. 207-212.

14 Peterson, J.L.1981." Petri net theory and modeling of systems", Prentice-Hall Inc., ISBN: 01336619835.

15 Silva M., Tervel E., 1997. "Petri nets for design & Operations of Manufacturing System", European Journal of Control, 3(3), pg. 82-199.

16 Gradisar D., Music G. 2007. "Production-process modeling based on production-management data: a Petri-net approach" International Journal of Computer Integrated Manufacturing, Vol 20, Issue 8, pg. 794-810.

17 Alan A. Desrochers and Robert Y. AI-Jaar. 1995."Applications of Petri Nets in Manufacturing Systems; Modeling , Control, and Performance Analysis", IEEE Press, ISBN 0879422955.

18 Jain, A., Jain, P.K. & Singh, I.P.2006. "Performance Modeling of FMS with Flexible Process Plans- A Petri nets Approach", International Journal of Simul. Model., 5(3), pg. 101-113.

19 Chen, H., Amodeo, L., Chu, F., Labadi, K. (2005). " Modeling and performance evaluation of supply chains using batch deterministic and stochastic Petri nets", IEEE Transactions on Automation and Engineering 2(2), 132–144

20 Chen, J., Chen, F.F. (2003). " Performance modelling and evaluation of dynamic tool allocation in flexible manufacturing systems using coloured Petri nets: an object-oriented approach", International Journal of Advanced Manufacturing Technology 21(2), 98–109

21 Ghaeli, M., Bahri, P.A., Lee, P.L. (2008). " Scheduling of a mixed batch/continuous sugar milling plant using Petri nets", Computers and Chemical Engineering 32, 580–589

22 Huang, B., Sun, Y., Sun, Y.M. (2008). “Scheduling of flexible manufacturing systems based on Petri nets and hybrid heuristic search”, International Journal of Production Research 46(16), 553–4565

23 Tuncel, Gonca. (2008). "An integrated modeling approach for shop-floor scheduling and control problem of flexible manufacturing systems." The International Journal of Advanced Manufacturing Technology 59.1127-1142.

24 Wang, Z., Zhang, J., Chan, F.T.S. (2005). “A hybrid Petri nets model of networked manufacturing systems and its control system architecture”, Journal of Manufacturing Technology Management 16(1), 36–52

25 Uzam, M. (2002). “ An optimal deadlock prevention policy for flexible manufacturing systems using petri net models with resources and the theory of regions”, International Journal of Advanced Manufacturing Technology 19(3), 192–208

26 Uzam, M. (2004). “The use of the Petri net reduction approach for an optimal deadlock prevention policy for flexible manufacturing systems”, International Journal of Advanced Manufacturing Technology 23(3-4), 204–219

27 Maione, G., DiCesare, F. (2005). “Hybrid Petri net and digraph approach for deadlock prevention in automated manufacturing systems”, International Journal of Production Research 43(24), 5131–5159

28 Lee, J.K., Korbaa, O. (2006). “Scheduling analysis of FMS: an unfolding timed Petri nets approach”, Mathematics and Computers in Simulation 70(2), 419–432

29 Hsu, T., Korbaa, O., Dupas, R., Goncalves, G. (2008). “ Cyclic scheduling for FMS: modelling and evolutionary solving approach” European Journal of Operational Research 191(2), 464–484

30 Syski, R. (1986). “Introduction to Congestion Theory in Telephone Systems”, 2nd Edition, Elsevier Science Publisher, Amsterdam.

31 Mital, K.M. 2010). “Queuing analysis for outpatient and inpatient services: a case study”, Management Decision, Vol. 48 No. 3, pp. 419-439,

32 Cooper, R.B., Queuing Theory: Chapter 10 in Stochastic Models, Heyman and M.J. Sobel Edition, North Holland, 1990.

33 Hideaki Takagi, Queuing Analysis, 1st Edition, vol. 2, Elsevier Science Publisher, Netherland, 1993.

34 Ullah, H. (2011). “Petri net versus queuing theory for evaluation of FMS”, Assembly Automation, Vol.31, No. 1, pp. 29–37.

35 Tsarouhas, P.A. (2011). “A comparative study of performance evaluation based on field failure data for food production lines”, Journal of Quality in Maintenance Engineering, Vol. 17 No. 1, pp. 26-39.

36 McGuire, A.M. (2010). "A framework for evaluating the customer wait experience", Journal of Service Management, Vol. 21 No. 3, pp. 269-290.

37 Mehmood, R. and Lu, J.A. (2011). "Computational Markovian analysis of large systems", Journal of Manufacturing Technology Management, Vol. 22 No. 6, pp. 804-817.

38 Gudmundsson, D. and Goldberg, K. (2007). "Optimizing robotic part feeder throughput with queuing theory", Assembly Automation, Vol. 27 No. 2, pp. 134–140.

39 Papadopoulos, H.T., Heavy, C. and Browne, J. (2013). "Queuing Theory in Manufacturing System Analysis and Design, Springer Verlag Gmbh.

40 Guy L.C. and, Feldman, R.M., "Manufacturing Systems Modeling and Analysis", Second Edition, Springer, 2010.

41 Gershwin.S.B, Manufacturing Systems Engineering, Englewood Cliffs, NJ: Prentice-Hall, 1994.

42 Yao, D.D., "Stochastic Modeling and Analysis of Manufacturing Systems", Berline: Springer-Verlag, 1994.

43 Buzacott,J.A. and shanthikumar J.G., "Stochastic Models of Manufacturing Systems", Englewood Cliffs, NJ: Prentice- Hall, 1993.

44

44 Koo, P.H., Moodie, C.L., and Talavage, J.J. (1995). " A spreadsheet model approach for integrating static capacity planning and stochastic queuing model" International Journal of Production Research, vol. 33, no. 5, pp. 1369-1385.

45 Sukhotua, V. and Peters, B.A. (2005). "Modeling of material handling systems for facility design in manufacturing environments with job-specific routing", International Journal of Production Research, vol. 50, no. 24, pp. 7285-7302.

46 Marcheta, G., Melacinia, M., Perottia, S. and Tappiaa, E. (2012). " Analytical model to estimate performances of autonomous vehicle storage and retrieval systems for product totes", International Journal of Production Research, Vol. 50, No. 24, pp. 7134-7148.

47 Sivakumar, A.I. and Chong, C.S. (2001). "A simulation based analysis of cycle time distribution, and throughput in semiconductor backend manufacturing", Computer in Industry, Vol. 59, pp. 78–45, 2001.

48 Domaschke, J. and Brown, S., Effective implementation of cycle time reduction, Proceeding of the 1998 Winter Simulation Conference, USA, 1998.

49 Wang, M., Sun, G. and Wang D., "Manufacturing simulation – An effective tool for productivity improvement productivity and reducing manufacturing cycle time through simulation modeling", Proceeding of 3rd International Microelectronic & Systems Conference, Malaysia, 1993.

50 Toh, G.K., Teck, U.W., Lie, A., Sun, G., Ming, W., and Kok, K. (1995). "Reducing manufacturing cycle time of wafer fab with simulation", World Scientific, Vol. July 1995, pp. 889-896.

51 Zhou, M., Chen, Z., He, W. and Chen, X. (2010). "Representing and matching simulation cases: a case-based reasoning approach", Computers & Industrial Engineering, Vol. 59, pp. 115–125. 2010.

52 Deo N. 2000. "Graph theory with applications to engineering and computer science", Prentice Hall India, New Delhi.

53 Biswal P.C. 2005. "Discrete mathematics and graph theory", Prentice Hall India, New Delhi.

54 Jurkat WB, Ryser HJ. (1966). "Matrix factorization of determinants and permanents", J Algebra 3:1–27.

55 Gandhi O.P., Agrawal V.P. (1994). "A digraph approach to system wear evaluation and analysis", Journal of Tribology 116:268–274.

56 Gandhi O.P., Agrawal V.P., Shishodia K.S. (1991). "Reliability analysis and evaluation of systems", Reliability Engineering and System Safety 32:283–305.

57 Grover S, Agrawal V.P., Khan I.A. (2004). "A digraph approach to TQM evaluation of an industry", International Journal of Production Research 42:4031–4053.

58 V.P.Agrawal and J.S.Rao, "Identification and isomorphism of kinematic chains and mechanism", Mech.Mach.Theory, Vol.24, No.4, 309-321, 1989.

59 Wani, M.F., Gandhi, O.P. (1999). "Development of maintainability index for mechanical systems", Journal of Reliability Engineering and System Safety, 65, 259-270.

60 Kulkarni, S. (2005). "Graph theory and matrix approach for performance evaluation of TQM in Indian industries", The TQM magazine, 17(6), 509-526.

61 Grover, S., Agrawal, V. P., Khan I. A. (2005). "Human resource performance index in TQM environment." International Journal of Management Practice 1.2, 131-151.

62 Garg, R.K., Agrawal, V.P., Gupta, V.K. (2006). "Selection of Power plants by evaluation and comparison using graph theoretical methodology", Electric power and energy systems, 28, 429-435.

63 Durai Prabhakaran, R.T., Babu, B.J.C., Agrawal, V.P. (2006). “Structural modeling and analysis of composite product system: A Graph theoretic Approach of Composites materials”, Vol.40, No.22, 1987- 2007, 2006.

64 Durai Prabhakaran, R.T., Babu, B.J.C., Agrawal, V.P. (2006). “Concurrent Engineering: Research and Application”, SAGE publication, Vol.14, No.2, 151-161.

65 Durai Prabhakaran, R.T., Babu, B.J.C., Agrawal, V.P. (2006). “Quality modeling and analysis of polymer composite”, J.of Polymer composite, 329-340, 2006.

66 Rao, R. V. (2006). “A decision-making framework model for evaluating flexible manufacturing systems using digraph and matrix methods”, The International Journal of Advanced Manufacturing Technology, 30(11-12), 1101-1110.

67 Faisal, M. N., Banwet, D. K., & Shankar, R. (2006). “Supply chain risk mitigation: modeling the enablers”, Business Process Management Journal,12(4), 535-552.

68 Singh, V., & Agrawal, V. P. (2008). “Structural Modelling and integrative analysis of manufacturing systems using graph theoretic approach”, Journal of Manufacturing Technology Management, 19(7), 844-870.

Biography: (12 Point Times New Roman Bold)

Priyank Srivastava is Research Scholar at National Institute of Technology, Kurukshetra. His research interest lies in Computer Aided Manufacturing, Manufacturing Systems and Operation Management.

Dr. Dinesh Khanduja is Professor at National Institute of Technology, Kurukshetra. His research interest lies in Production Engineering, Quality Management, Productivity Management, Entrepreneurship, Operation Management and Energy

Dr. V.P Agrawal is visiting Professor at T.I.E.T, Thapar University, Patiala, India. His research interest lies in Structural Modeling and Analysis, Reliability Engineering, Concurrent Engineering.

Cloud Manufacturing: Intelligent Manufacturing with Cloud Computing

Ashidsha Jaleel[1] Prof. T.K Rajendran[2] Lijohn P George[3]

[1,2,3]Department of Production Engineering

University of Calicut

Thrissur, Kerala, India, 679533

Email:ashiktd@gmail.com

Abstract

Cloud Manufacturing is the newly service oriented networked manufacturing system. The existing manufacturing models are improved using the cloud computing, Internet of things (IoT), virtualization, service oriented technologies and advanced high speed computing. Cloud manufacturing (CM) aims at a flexible, self adaptive manufacturing foundation which is capable of dynamically exploiting manufacturing resources and services distributed across the network. As cloud computing is the major enabler for the manufacturing industry, it can transform traditional manufacturing business model to product innovation business model with the help of intelligent factory networks. The service provided by cloud computing for cloud manufacturing, layers, flow of production service, layout of production service and key features are discussed and investigated. A flow criteria in a production service for a typical small, medium enterprise (SME) is introduced. A solution for the security barrier of the CM network is introduced with the help of a recent computing technology in this paper.

Keywords— cloud manufacturing; cloud computing; service oriented architecture; networked manufacturing; internet of things; small, medium enterprises.

Introduction

Cloud manufacturing is the newly service-oriented, network based manufacturing mode. The resource sharing technology of cloud computing is the central core of Cloud manufacturing. Cloud manufacturing vision is mainly designed targeting the small, medium enterprises (SME). Networked manufacturing is the present used way of connecting methodology in industry. When the ASP (Application Service Provider) and Manufacturing grid technologies[1] combines a new revolutionary vision Cloud manufacturing is acquired.
The services provided in cloud manufacturing system are production service, Design service, Testing and Simulation service and Operational management service. In this paper the production service is taken into study. The flow and layout of production service is explained in this paper. The paper gives a glimpse about the advanced security technologies that will make the system more secure.

History

The history of Cloud manufacturing begins when the Hitachi Seiki, USA, Inc [2] brought forth the idea that the CNC machine could be attached to the network (and therefore available to the Internet as well) by use of an on-board Ethernet port. This Ethernet port would allow the owner to simply plug the CNC machine into a network jack and then send or receive part programs at network speeds. The Hitachi Seiki UUP machine interface was designed to plug into any network connection anywhere in the plant. If machines require relocation, the user could simply relocate a network cable and plug it into the nearest available network jack quickly putting the machine back online. This technology innovation can be stated as the first internet enabled manufacturing step. The UUP approach was discontinued due to financial crisis in the company. As the concepts were strong it was estimated that this concept may re-emerge in the industry in any non-proprietary format.

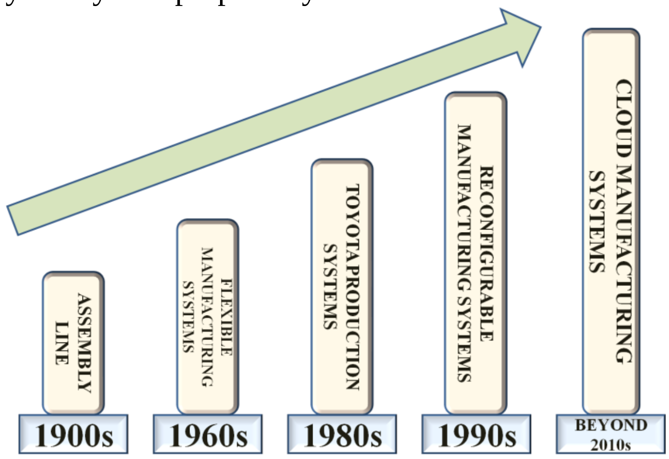


Fig1.1 A history of manufacturing systems development

The concept of Cloud manufacturing was initially explained by the scholar Bohuli[3] in 2010 in one of his paper. It gave an initial outlook to the Cloud manufacturing (CM) concept and many supporting papers were published.
In 2010 European Commission initiated the FP7 program named Manucloud[13][14] . Its objective was to investigate, develop and evaluate a suitable IT infrastructure to provide better support for on-demand manufacturing scenarios. In this path, ManuCloud seeks to implement the vision of a cloud-like architecture concept. It provides the users necessary capability to utilize the manufacturing capabilities of configurable, virtualized production networks, based on cloud-enabled, federated factories, supported by a set of software-as-a-service applications. ManuCloud is part of EU FP7 program[4].

Cloud Computing

Dynamic resource sharing, on-demand resource provisioning, virtualization are the key building entities of cloud computing. “Everything is a service” when it comes into cloud. The concept of cloud computing was initially introduced by John McCarthy. However the term “cloud” came into prominence when large enterprises began to use the Virtual Private network (VPN). The concept of VPN uprooted the point-to-point data circuit technology. We can say VPN as the beginning factor of cloud computing[5][6].
Customers do not have to pay for infrastructure, installation, manpower etc. to handle these activities including maintenance. The recent trend in IT envisage, moving the computing and data from desktop to large data centres.
In 2006 Amazon web services (Amazon EC2) was launched giving world a new revolution of resource sharing concept. The open source platform Eucalyptus which is used to create

Amazon configured cloud technology paved a new way for the development of cloud networks in large scale.

Services provided by Cloud Computing for CM

"Everything is a Service" (XaaS) is the key motto of cloud computing. It is the core enabling technology of CM. The cloud computing technology is provided as services for the user in the following three models:-

1. *Software as a Service* (SaaS). The application software is offered as a service. Here the application runs on a cloud platform. It doesn't require installing on clients computer; examples are Gmail, Facebook etc.
2. *Platform as a Service* (PaaS). The platform is provided as a service here. This enables development and deployment of software without buying hardware and software layers for it; examples are Google App Engine
3. *Infrastructure as a Service* (IaaS). The storage and computing capabilities are made as a service; examples are Gogrid, Rackspace servers

Rather than IT services, CM is mostly concentrated in the areas of manufacturing resources and its allied services. So we can derive a new set of cloud computing models for CM utilising SaaS, IaaS, PaaS including the following models [7]:-

1. *Production as a Service* (PROaaS) – The production activities of a product and its abilities are provided as a service.
2. *Design as a Service* (DaaS) – The design resource and ability are provided as a service.
3. *Testing as a Service* (TaaS) – The testing resource and ability are provided as a resource.
4. *Simulation as a service* (SIMaaS) – The simulation resource and ability are provided as a service.
5. *Management as a Service* (MANaaS) – Managerial resource and ability are provided as a service

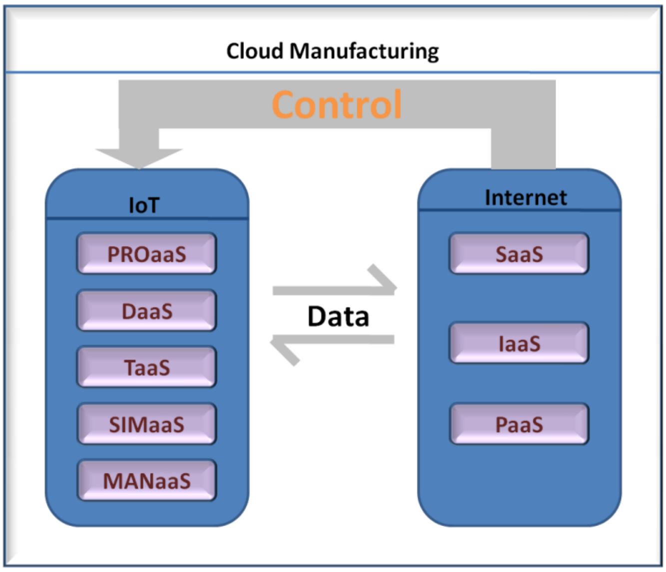


Fig1.2 Relationship between cloud computing and CM

Cloud Manufacturing

Cloud manufacturing is the newly service oriented networked manufacturing system[8]. This manufacturing vision is introduced to provide the on-demand architecture with flexibility and reliability based on cloud computing.
Cloud manufacturing provides a flexible and adaptive infrastructure for industries to share and use various manufacturing resources or services on-demand under the dynamic, complicated and large-scale business environment.
The concept of cloud manufacturing was initially introduced by academician Bo-hu Li in 2010.
Cloud manufacturing mixes network-based manufacturing, service technology with cloud computing, IOT[9] and other advanced technology to achieve all kinds of manufacturing resources, centralized management and intelligent business.
CM centralizes an intelligent management to supply low cost, global service to cloud customers, and its main clients are small or medium enterprises.

A. *Layers of Cloud Manufacturing*

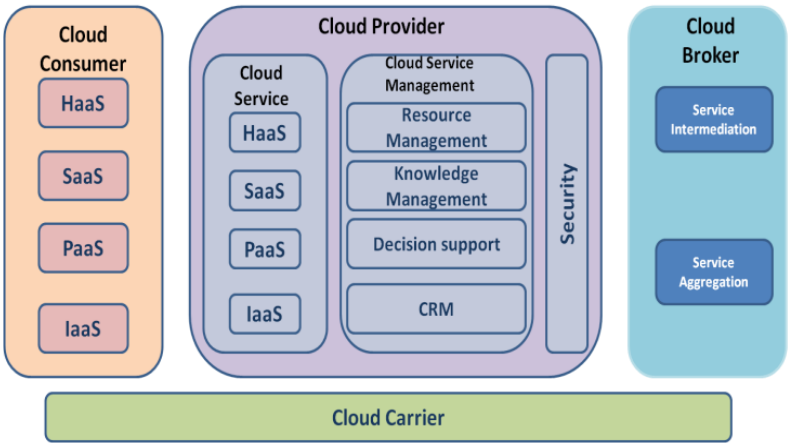


Fig1.3 Layers of Cloud Manufacturing

Layers [10] of a CM consist of cloud consumer, Cloud provider, Cloud broker and the Cloud carrier. Cloud provider is the core player in the CM. Cloud broker acts as a middle person between the consumer and the service provider.

Flow of Production Service in Cloud Manufacturing System

The flow of production service in a cloud manufacturing system is carried out in basically seven steps:-

a) **Client defining the needs**
b) **Registering with Cloud Agent**
c) **Raw material (RM) searching & Remote procurement**
d) **Plant identification**
e) **Transfer of RM**
f) **Manufacturing of Product**

g) **Transportation of product to client**

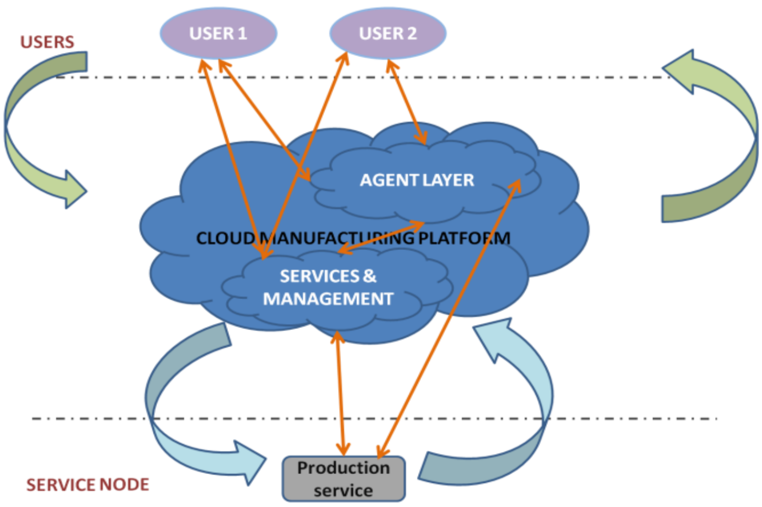


Fig1.4 Production Service model of Cloud Manufacturing

a) *Client initiates the production process by defining his needs*

Client plans his product or production process which needs to be outsourced into the central network. Pricing, product specification and designs are documented into a standard format (Standard cloud format).

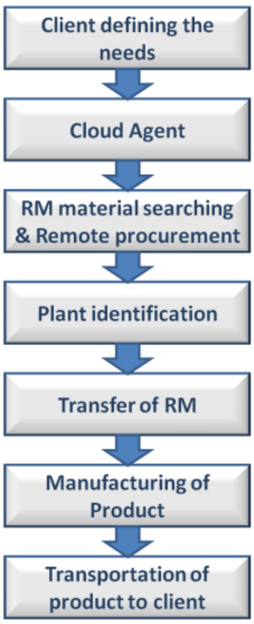


Fig1.5 Flow of production service in CM

b) *Registering with Cloud Agent*

The agents in the cloud are the major controllers in the cloud manufacturing system. Agents are equipped with intelligent data warehousing systems for the storage of information regarding plants and supporting services.

c) Raw material searching and Remote procurement

The RM suitable for the production process (as specified in documentation) is searched by the agent. The suitable RM providers are shortlisted. The suitable providers are screened with strict specification checking. The selected RM is temporarily procured to the deploying unit.

d) Plant identification by agent in association with client

The plant which is suitable for the manufacturing of the clients design is identified. Database with the agent is utilized for this purpose. The intelligent computer systems will compare the plants available and the documentation provided by the client to identify suitable, cost effective, time bound and completing plants.

e) Transfer of RM to Plant

The procured RM at the RM provider is temporarily stocked at their warehouse. The agent will transfer these RM to the plant warehouse unit using their logistics techniques.

f) Manufacturing of Product

When the RM arrives at the plant, quality testing is carried out. This helps the plant to deliver quality outputs. The qualified and accepted RM is utilized for production.

g) Transportation of product to client specified destination by agent.

The finished products are kept at the temporary inventory of the plant. It is transported to the specified location as per the delivery schedule.

Layout of a Production Service in Cloud Manufacturing System

In a production service Cloud Manufacturing system, the SME users are the service requestors. The requested service is available to the agent cloud through common public platform. The agent cloud analyses the request, decomposes, schedule and do some fault tolerance in the data.

The agent cloud is frequently linked with the RM provider cloud and the plant cloud. Agent cloud acts as a moderator between the requestor and utility provider. The plant cloud is connected to each and every production plant in the CM system. This is achieved using an interface. Plant private network is the controller grid of the machining center and equipment’s connected to it. The commands for these machines are generated in the plant server. The command for the machines is extracted /created from the documentation provided by the agent cloud.

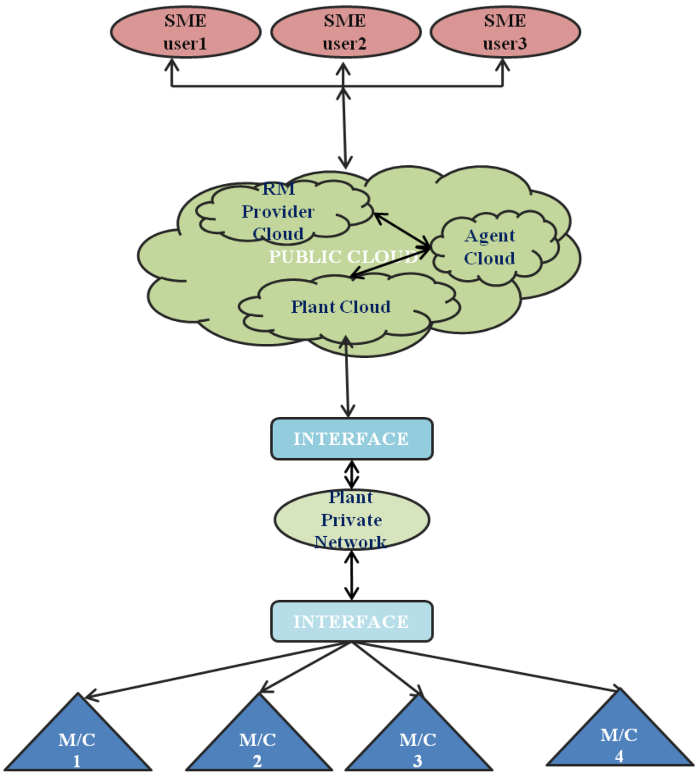


Fig1.6 Layout of Production Service CM system

Security Enhancement in Cloud Network for Enterprises

The major barrier in cloud technology implementation is lack of sufficient security in cloud network[11]. Lack of appropriate security technologies resisted the growth of cloud systems. As cloud computing is the enabling technology of CM, the security barrier retarded its growth. By 2011 various technologies with wide acceptance began to catch the IT market.

The past barriers are given below:-

- Lack of traceability
- Lack of intelligent monitoring
- Lack of standard encryption techniques

The emergence and usage of security technologies like AES 256 bit encryption, FIPS 140-2 make the cloud functionality of enterprises to real. Examples are PerspecSys, CipherCloud. The solutions for the past barriers can be categorized as:-

- *Traceability*

  The traceability of cloud applications can help enterprises in tracing their data and how they are used in the enterprise. By this technology we can discover who is accessing and updating records from various locations. Example- AppProtex Discover

- *Activity monitoring*

Monitoring of cloud systems represents operational, legal or compliance related risk to the enterprise. This helps in auditing and external compliance related tasks. Example- AppProtex Analyze

- *Cloud encryption gateways*

The Encryption methods use a cipher algorithm for mathematical transformation of data. These encrypted data can be reverted to original value using a key. Tokenization replaces a sensitive data field and replaces it with a surrogate value for security purpose. These two technologies (encryption and tokenization) can be effectively used and enabled as cloud encryption gateways [12]. This makes the IN and OUT data secured. Examples are AppProtex Protect, CipherCloud.

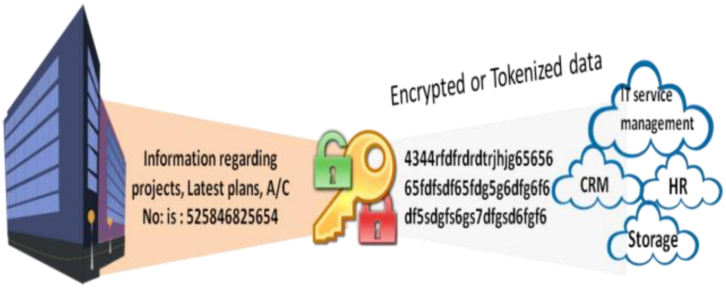


Fig1.7 Encryption/decryption of enterprise data

Conclusion

Cloud Manufacturing offers a processing infrastructure for colossally collaborative manufacturing with agility, adaptability and cost effectiveness.Currently, the enterprise informational integration is not perfect, as it hinders the application of cloud-manufacturing process to some extent.The emergence of cloud computing provides a new way to complete the mission of networked manufacturing. In contrast to the conventional networked manufacturing approach, the cloud manufacturing vision introduced in this paper promises flexibility, elasticity and adaptability through the on-demand provisioning of manufacturing resources as a utility by blending the principles of cloud computing.

Acknowledgement

We would like to acknowledge the sincere support provided by Dr. Sudha Balagopalan (Principal, Vidya Academy of Science and Technology, Thrissur, Kerala) in completion of the paper.

References

[1] Foster I, Zhao, RaiculI "cloud computing and grid computing 360 degree compared. In :*grid computing environments work- shop; 2008.*

[2] http://www.mmsonline.com/articles/connecting-your-legacy-cnc-machine-tool-to-the-internet

[3] Bohu Li, Lin Zhang, XudongCai,"Introduction of Cloud Manufacturing" [J].*Zhongxing Communications Technology, 2010, 4(16):5-8.*
[4] Matthias Meier, Joachim Seidelmann and IstvánMezgár ,ManuCloud: The Next-Generation Manufacturing as a Service Environment, *ERCIM News 83*
[5] YashpalsinhJadeja, KiritModi, "Cloud Computing - Concepts, architecture and Challenges", *2012 International Conference on Computing, Electronics and Electrical Technologies [ICCEET]*
[6] Pallis G, "Cloud computing: the new frontier of internet computing". *IEEE Internet Computing 2010. [14:5:5562494:70-73].*
[7] F Tao, L Zhang "Cloud manufacturing: a computing and service oriented manufacturing model" *Proc. IMechE Vol. 225 Part B: J. Engineering Manufacture*
[8] Fanghua Ning1, Weizong Zhou "THE ARCHITECTURE OF CLOUD MAUFACTURING AND ITS KEY TECHNOLOGIES RESEARCH" *Proceedings of IEEE CCIS2011*
[9] Yunhao Liu, "Introduction of the Internet of Things" [M]. *Beijing, The Science Press, 2010.*
[10] Dazhong Wu, Matthew J. Greer "Cloud Manufacturing: Current status and future trends", *ASME 2013 MANUFACTURING SCIENCE AND ENGINEERING CONFERENCE*
[11] Subashini S, Kavitha V. "A survey on security issues in service delivery models of cloud computing". *Journal of Network and Computer Applications 2011; 34(1):1–11.*
[12] Gerry Grealish, "Gartner Highlights Cloud Data Encryption Gateways - A means to protect sensitive corporate data", http://virtualization.sys-con.com/node/2464514
[13] http://www.strath.ac.uk/dmem/research/researchprojects
/manucloud
[14] http://www.manucloud-project.eu/index.php?id=186

A fully-automated manufacturing environment realized through a flexible in house logistic system with smart transportation infrastructure

Jan-Hinrich Kämper[1], Arne Stasch[2] and Axel Hahn[1]

[1]Department of Computing Science, Carl von Ossietzky University
[2]OFFIS Institute for Information Technology
Oldenburg, Germany
Email: jan-hinrich.kaemper@uni-oldenburg.de,
stasch@offis.de, hahn@wi-ol.de

Abstract

Modern industrial production and transportation systems require agile material handling systems. Changing requirements demand flexibility in transportation layout and control systems that are adaptable to the actual needs. Decentralized control systems promise complexity reduction and dependability while centralized approaches improve planning and allow general optimization. By using the use case of a cross docking logistic system the paper describes how this approach can be combined and implemented by using an agent-based control system on a decentralized lightweight microcontroller infrastructure. The scenario is applicable to manufacturing environments with modularized assembly lines. These modules are connected by different transport units which merge into an intelligent infrastructure including AGVs, conveyors and storages.

1. Introduction

Highly dynamic markets in a globalized world and the increased demand for individualized products make flexible concepts for in-house logistics indispensable in modern manufacturing environments. Companies need to optimize their processes in order to sustain their position on very competitive markets with high volatilities in supply and demand. Optimization in this context not only relates to costs but also and more importantly to “soft factors” such as reliability, agility and transparency.

Traditional manufacturing systems with a strict sequential order of assembly lines and machines tend to impose a very inflexible structure onto the underlying logistical processes. Often manufacturing systems are designed and the planning of the corresponding material flow infrastructure including conveyors, cranes, belts etc. is tailored to the needs of the manufacturing system. The big downside of this approach is that neither the manufacturing system itself nor the logistical system can be modified without modifying the other. This is not dramatic for minor changes, but as soon as essential parts of the manufacturing system are subject to updates or modifications the logistical system has to be adapted appropriately. Scalability is another crucial criterion which can not be satisfied with the traditional approach, e.g. suppose that the throughput of a production site has to be increased by a certain factor then the entire logistical chain must be re-engineered as well. Also traditional systems suffer from a lack of reactivity and efficiency. Sudden deviations in demand and supply can not be accounted for on a short term basis (time horizon of minutes or hours) by accelerating or reducing performance of transport as it is hardly possible to turn of some of the conveyor belts or transport hubs without affecting the entire system due to connections and dependencies between the components. In the worst case each unit of transport can be a single point of failure disrupting the whole network when it breaks down.

Hence, the aim here is to rethink the idea of in-house logistics as an adequate service for the underlying manufacturing environment instead of being an inseparable part of the system. This approach shall support three high level quality goals of agile manufacturing:

1. Flexibility with respect to external influence factors on a short term basis, such as demand and supply, energy prices, security threads, quality issues etc.

2. Robustness and high availability by proactive behavior of decentralized organizational structures in very dynamic environments

3. Enabling high performance in mass-customization through a highly adaptable logistical setup

The approach described in this paper is based on decentralized control of all mobile and stationary transport entities in an agile manufacturing system (compare [1]). All these entities work autonomously to accomplish their primary goal, i.e. the completion of transportation jobs of the right good at the right time to the right place. They interact in order to organize the material flow between each other and communicate via a wireless network to this end. Properties of all entities can quite accurately be captured by the Agent metaphor [2]. Therefore all entities in the system are modeled as modularized agents that manage the material flow. The change of paradigm is thus to go from centralized planning of production processes and transport between production steps to decentralized behavior driven system control. Typical optimization goals in this context concern energy saving or a high overall throughput. To preserve goal-driven optimization a hybrid system which allows reactive balancing between the new decentralized kind of control and the traditional central algorithms is deployed. This is done by introducing dynamic hierarchies in agent societies.

This approach has been realized in a small use case consisting of four AGVs, four intermediate storage buffers and three stationary conveyors as well as a coupled Multi-Agent Simulation Framework that enables scaling up the size of the system in a virtual environment. The system basically is a simple cross-docking storage scenario, but when thinking of the intermediate storage buffers as points of production, then the whole scenario represents a complete manufacturing environment.

The remainder of this paper is organized as follows: Chapter 2 will now first give a short overview over the state of the art and similar in-house logistic systems employed in industry and research. Chapter 3 explains the techniques applied to realize dynamic balancing between central and decentralized control. Technical aspects regarding the realization of agent software on microcontroller hardware used to control the conveyors and the AGVs with an autonomous drive unit are treated in chapter 4. The design and the architecture of the use case are described in detail in chapter 5. Chapter 6 gives an outlook on future research questions and possible further enhancements of our system.

2. State of the art

The basic principles that the approach proposed here is inspired from are not new. However the combination of them and especially their application in an autonomous in-house logistics system that fits into arbitrary manufacturing chains with possible tight integration in a smart grid is innovative.

The idea of substituting fixed conveyor lines, manually driven vehicles or forklifts with a fleet of mobile autonomous transport robots emerged in recent decades as a result of the ongoing automation of industrial systems. The company *Kiva Systems* develops and deploys such systems for warehouse logistics in eCommerce companies [4]. Their approach focuses on order fulfilment solutions with self-organization of products and high flexibility of the material handling equipment used. *Fraunhofer IML* built up an experimental hall in the last years where 80 AGVs manage the transportation of goods from commissioning stations into a storage rack where they can move vertically as well [5]. Their research focus has been on intelligent swarm behavior of the vehicle fleet. *Karis* is a system developed by *Karlsruhe Institute of technology* where AGVs can team up autonomously in order to reach a higher transport capacity [6].

Lewandowski et al. (see [22]) developed a cyber physical system equipped with RFID to realize a material handling infrastructure where items track and trace themselves through the environment based on modular software agents. However they use a centralized repository to link the tracking data with the agent layer. Their approach tackles the issues posed by "Industrie 4.0" a German Universities Excellence Initiative which aims to improve automation technologies by integrating physical and IT components in complex manufacturing environments through deep pervasion of resources with software and their interconnection through the internet. The main methods used to create this kind of a digital factory are self-optimization, interoperability, cognition and intelligent support of workers in their increasingly complex work (see for example [7]).

A multi-agent system (MAS) is an appropriate approach for distributed control. Agents can work independently in decentralized systems. The loss or gain of agents does not affect the other agents' functionality. MAS make decentralized systems flexible and reliable. Requirements, like real-time operations, industrial standards and small footprints sort out most of the existing MAS frameworks for real-time microcontroller-based applications. JADE [8] is a very common MAS framework but microcontrollers with little processing power do not run Java. Mobile C uses C/C++ as programming language and C meets industrial standards [9]. However Mobile C needs a General Purpose Operating System (GPOS) and even the low memory embedded GPOS, e.g. Embedded Linux, need a larger footprint than 259 KB Flash Memory [10] and this does not account RAM usage. This is not applicable for low power/processing power microcontrollers which can be used in any environment. One part of this research is to create an Agent Framework which can run on such electronic control units (ECU) and is even applicable on any control unit.

Having multiple agents collaborate in teams is a concept as old as the agent metaphor itself. Different variants exist that implement such a collaborative behavior. Anderson et al investigate the question whether central or decentralized control is preferable in production systems and inspire their argumentation from nature. Ogston et al. describe a method to cluster agents in teams based on properties and goals [11]. Marcolino et al. discuss the trade-off between diversity and uniformity in designing MAS [12]. One way to mediate between central and decentralized control are holonic manufacturing systems (HMS). HMS were first mentioned by Koestler [13]. Fletcher et al. describe a holonic model for manufacturing environments aimed towards robustness [14]. Gerber et al. use HMS in a logistic scenario [15].

3. Realization of the fully automated logistic system

A decentralized in-house transportation system is a proper use case to evaluate the approaches of distributed problem solving. The aim is to create a modularized material flow system (MFS, [16]) that uses a MAS as control method. Dividable parts of the MFS are modules which have their own actuators and sensors to interact with the environment (See Figure 1). Each module has its own ECU and can act on its own behalf. They communicate and create a cyber-physical system (CPS) in bulk which acts as an extensive MFS [17]. In order to preserve the option of applying global optimization methods the MAS is enhanced with the ability to dynamically reconfigure between central and decentralized control. The resulting system thus has a hybrid control structure which is described in detail in the following section.

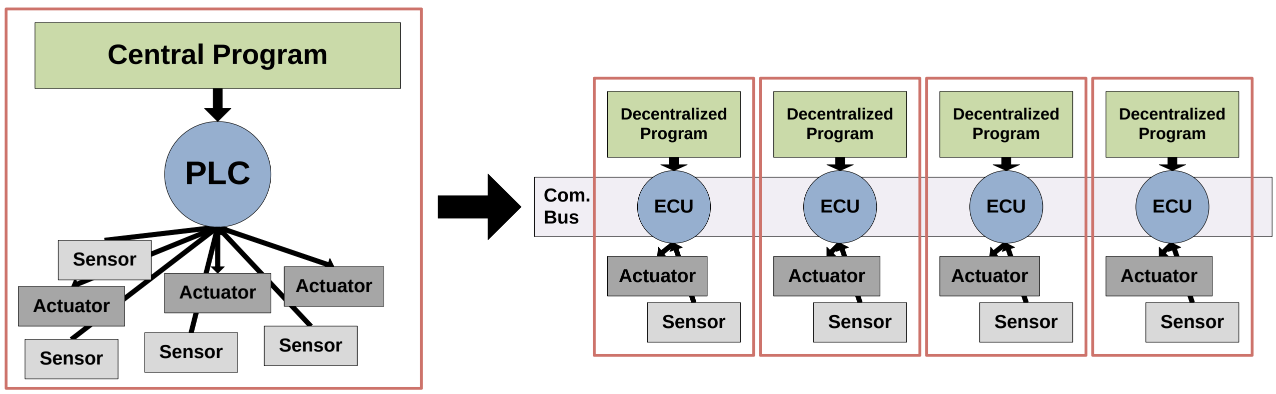


Figure 1: Change from centralized to distributed structure

3.1. Collaboration Strategies

In addition to the modelling of the agent architecture as described in the previous chapter a collaboration layer is provided which enables all agents from different conveyors to form larger teams orientated from the principle of HMS (see [15]). Typically the agents being part of teams are only platform agents representing the entire conveyor that they belong to including the corresponding routing, order and package agents (the precise agent architecture will be discussed in chapter 4). So for example a team could consist of two platform agents from AGVs and two platform agents from ramps which team up because they are located close to each other and complement one another with respect to their skills. So they should work on the incoming jobs collaboratively and not each one on its own. Such a team acts in the agent society as aggregated agent. Aggregated agents also have a proactive and behavior-driven control just like single agents. They also provide the same communication interface towards their environment as single agents. The crucial difference between a team of agents and a single agent lies in the internal control. Whilst a single agent only has to manage itself and consequently there is no such issue of control, teams of agents must somehow organize their collaborative behavior.

3.1.1. Central vs decentralized control driven by team structure

Internal control in these teams is intentionally designed to be easily adjustable. Unlike in a true MAS there always exists some kind of hierarchy in a team. Thus the single agents within a team are not as free as they would be without the team. To some degree they obey to the superior control of their team.

The exact degree of freedom they have is defined on a dynamic basis. One extreme is a strict hierarchy where one designated managing agent assigns orders to all other agents that only perform actions to complete these assigned orders (“authority”). There is no interaction or proactive behavior at all among the other agents. The managing agent takes the role of a central instance leading the team, so that distributed control no longer exists in that team. The other extreme would be to have a team of entirely autonomous agents which only talk to one very weak central instance when interacting with the environment (“autonomy”). The central instance in that case only represents a communication interface towards the surrounding world.

Many intermediate steps between these two extremes “authority” and “autonomy” are possible. See the 3 graphics in the figure below for an illustration. The first setup (Figure 2

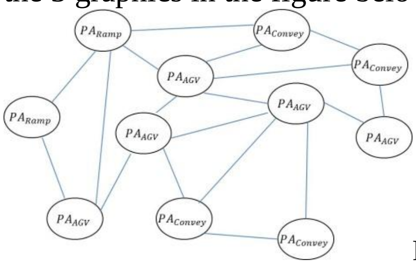


Figure 2) represents the per default purely decentralized control structure where all agents can communicate in a mesh but all have a locally restricted view of the system (a true MAS). In the second graphic (Figure 3) 3 teams have been formed consisting of 3-4 single agents. In these teams one agent always takes the lead and organizes the internal behavior of its team. This configuration therefore is a mixed setup with central control on a local level (intra-team) and distributed control on a global level (inter-team). In the last graphic (Figure 4) the system is reconfigured now depicting the case of a traditional centrally controlled system with one authority managing all other agents.

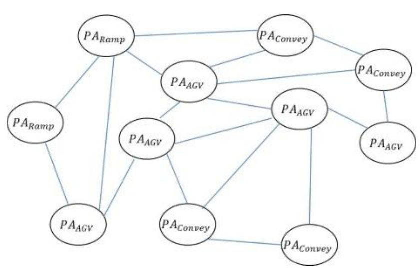


Figure 2: Full Autonomy

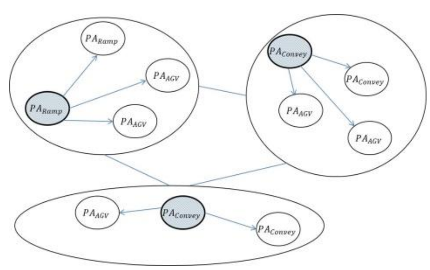


Figure 3: Aut-Balanced System

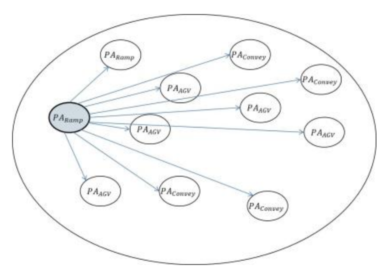


Figure 4: Full Authority

Depending on the number and size of the teams and the configuration of their internal control the described approach enables a flexible switching between central and decentralized control in operating mode (in the remainder called “aut-balancing”). This ability is a true asset since it allows for flexibility regarding different external influence factors. A purely decentralized control is preferable when efficient and robust transport processes are required and the demand remains far under the maximum bound of the systems transport capacity. However when the throughput and speed have to be increased due to peaks in demand or failure of machines it becomes necessary to employ optimization methods which can only be executed with a central instance having a global view over parts of the system. Also exceptional circumstances such as security threads, error recovery or obliged audits require a central control in order to make the system behavior more transparent and be able to detect problems and risks.

3.1.2. Aut-Balancing in the simulation and issues

We realize this approach only in the virtual simulation environment since there is no additional value in doing this in the physical world where the number of entities is too small to investigate effects of large swarm dynamic behavior. This is still work in progress so evaluation results can not yet be presented.

There are several issues regarding the realization of aut-balancing but which will not be discussed here but are topics for future work:

1. Selecting the managing agents
2. Choosing the internal hierarchy in a team
3. Deciding when to reconfigure the system
4. Supplying enough computing power for managing agents

4. Microcontroller infrastructure for MAS

In this chapter the modularization of MFS is addressed. The networked ECUs control their modules and a MAS should be implemented as decentralized control concept for the collaboration strategies. In previous work [19] the design for an Agent Framework for microcontrollers with low processing power was presented. The concept is refined and the adaptability of the concept will be proven through use of different platforms. Inspired from the model of a modular SW-architecture for ECUs from AUTOSAR [20] the SW should be adaptable to different microcontrollers. The focus of this part of the research is to design an Agent RTE which delivers in- and output abilities and flawless communication between agents. The layer architecture shown in Figure 5 is an example for a MICAz node (see [18]) on a ramp module which is described in chapter 5.4. Through unitary interfaces the Agent RTE can use different HW for the same functions. Even the RTOS in this example Contiki OS should be replaceable through another. The Agents are embedded with the same environment on every platform.

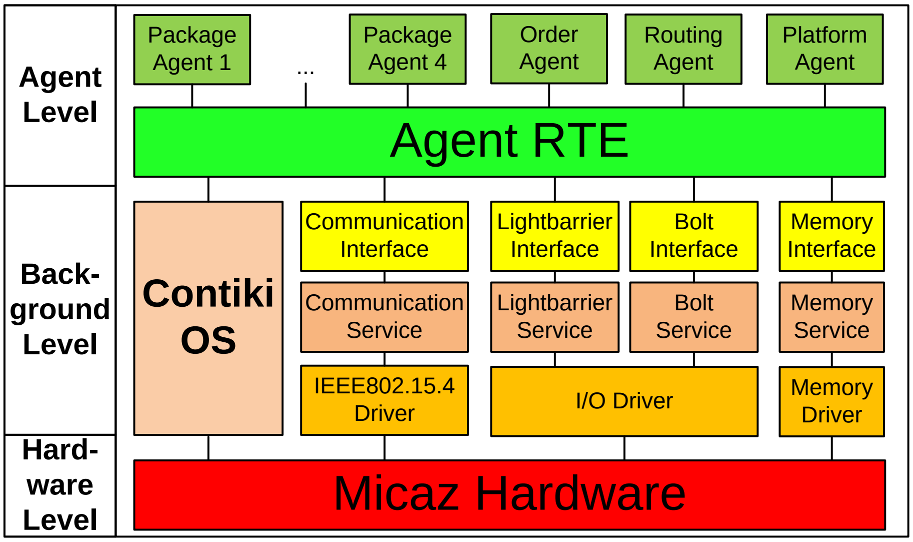


Figure 5: Modular architecture of a ramp

The difficult part is to construct a lightweight RTE. The possibility for an agent to communicate with every other Agent with a reduced set from FIPA messages [21] is an RAM consuming task. Allowing all agents to use the same external communication channel per platform and the ability to communicate on the same platform, the use of message queues or token methods are inevitable. Although the Agent RTE sets many boundaries to capabilities and the design of the agents, it makes it easier to add new agents or reconfigure, add or remove modules in appreciation of the MFS view. This enables “traveling” agents like the Package Agents who can “travel” with the items from module to module. This use case of a MFS affiliates to real-time safety critical systems which are sensitive against transferred executable code and this is not object of this research. The Package Agents will be only transferred via parameters and the Agent RTE activates and deactivates the existing code of the agent.

5. Construction of the overall system

The aim was to create a highly scalable fully autonomous MFS that can benefit from the technologies introduced in the previous chapters 3 and 4. This approach to a MFS (see Figure 6) includes the transport units: AGVs, conveyors and storages. The AGVs transport items individually from point to point. The conveyors take over fixed high traffic routes. The storage is represented by autonomous ramps with FIFO logic. Due to scalability, costs and workload a realistically scaled plant has not been realized. On an area of 50 m^2 four AGVs, three short conveyors and four ramps are representing a realistic extract of the plant. The virtual simulation of the whole plant works in hybrid mode to interact with the physical parts. It allows in single mode tests of new algorithms before they are used in the real environment, this will be described in more detail in chapter 5.5. The systems design allows creating any layout combination with the modules to adapt to necessary structures. A modularized infrastructure consisting of different resources assures the cooperation of the modules and is described in the next section.

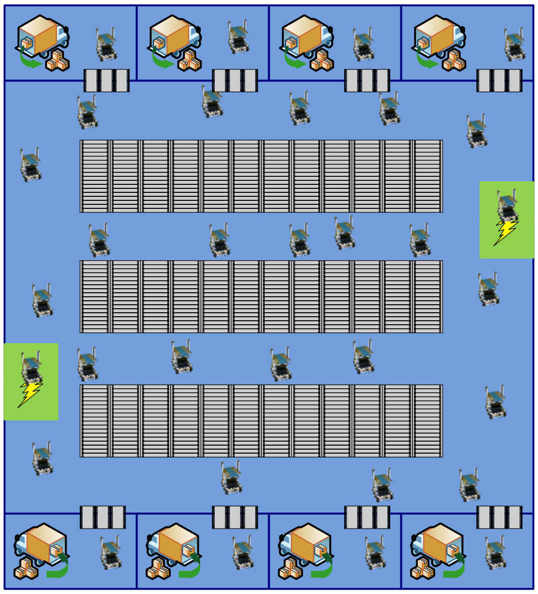


Figure 6: Draft of an entire MFS site

5.1. Infrastructure

The infrastructure of the realized parts of the system is designed in a modular setup (see [3]). The described module types (see Figure 7) are physically very different but they all have the same vertical system layout [16]. Every module has its actuators, sensors and an ECU that carries a multi-agent platform described in chapter 4.

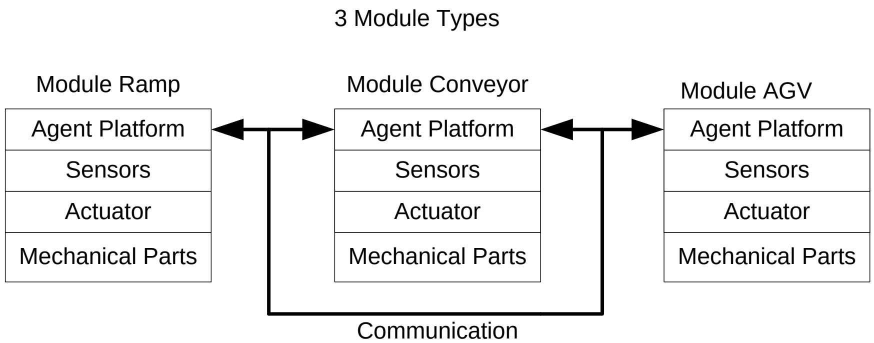


Figure 7: Module Types

The ECU controls the module and communicates with other modules (See Figure 8). The AGVs and ramps are equipped with MICAz nodes (see [18]) which are based on the ATMEL ATmega128 chip and use wireless IEEE 802.15.4 radio standard. They are often used as sensor nodes in distributed measurement applications. The “STASH” (Shortened and Transformed Agent System Handler) controllers are custom ECUs with Texas Instruments MSP430 low power microcontrollers. They communicate over the fieldbus standard Profibus. A Gateway between those two communication standards is necessary. The gateway is realized through a MICAz node with a Profibus adapter. In the remainder of this chapter the three module types will be explained in more detail.

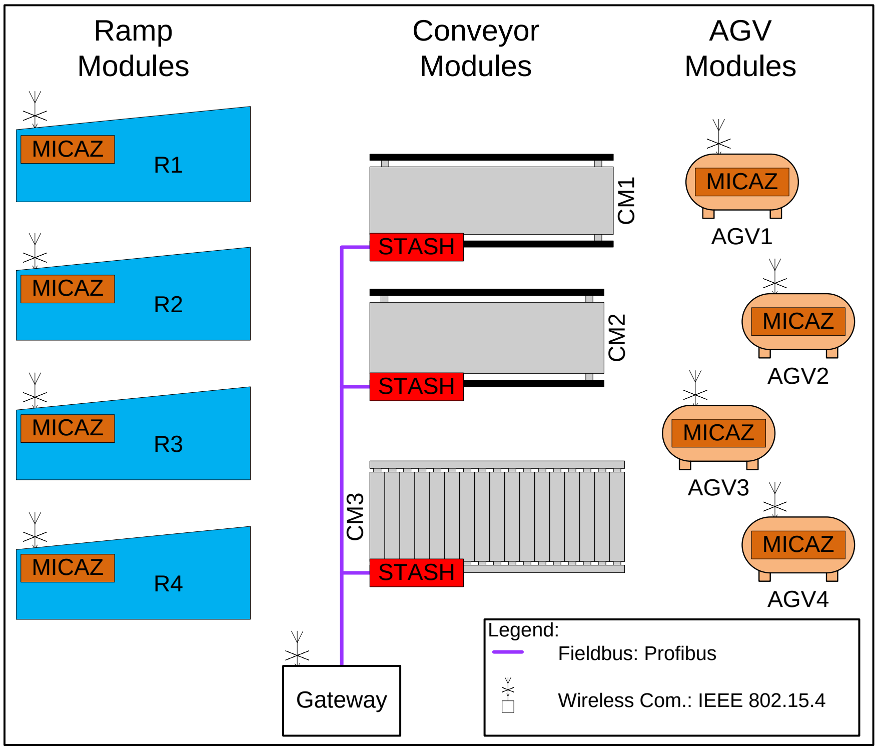


Figure 8: Infrastructure with ECUs and communication types

5.2. AGVs

The AGVs are VolksBots from the Fraunhofer IAIS (see www.volksbots.de) with a custom lifting platform (See Figure 9). The bot has three wheels on each side which are coupled with a belt. They allow a driving speed of about 2 m/s and turning on the spot by rotating the right and left sides into opposite directions. The lifting platform can take in and take out one item with a motorized belt. They can also lift the item to the height of the interacting module with a motor driven arbor. Two light barriers are monitoring the items positions.

Figure 9: Volksbot with Lifting Platform

Laser scanners in front and back ensure positioning and coordination with other AGVs. Analyzing laser scanner data and path planning needs more processing power then a small low power controller could provide. Subnotebooks provide the power and are connected to MICAz nodes. The MICAz node assigns higher commands and the subnotebook acts in lower levels of execution.

Algorithms used for path finding are the A* search algorithm from Hart et al. (see [23]) and Lee's algorithm (see [24]) depending on the traffic situation and current goals of the AGVs (high speed, energy saving etc.). The software architecture of the AGVs is layered in the following Order: On the lowest layer the hardware (sensors and actuators) are controlled. The operational layer above contains all important algorithms for analyzing data and driving as well as decision making (routing, collision detection, positioning etc). The layer on top ensures the connection to the MFS infrastructure via a MICAz gateway. The infrastructure indicates the destinations and handles the MAS communication and representation in the module network.

5.3. Conveyors

The conveyors are short roll and belt conveyors. Through historical use of these conveyors and "STASH" controller everything works on Profibus (see Figure 10). The ECUs are Profibus Slaves and cannot communicate with other electronics in the Profibus network then a programmable logic controller (PLC) as Profibus master. The PLC remains in the system as a Profibus network switch to transfer messages (ECU to ECU), actuator commands (ECU to frequency converter) and light barrier state (light barrier to ECU). The higher logic and decision making is realized in the ECU or the ECUs can be substituted by the PLC to evaluate decentralized system control versus central control.

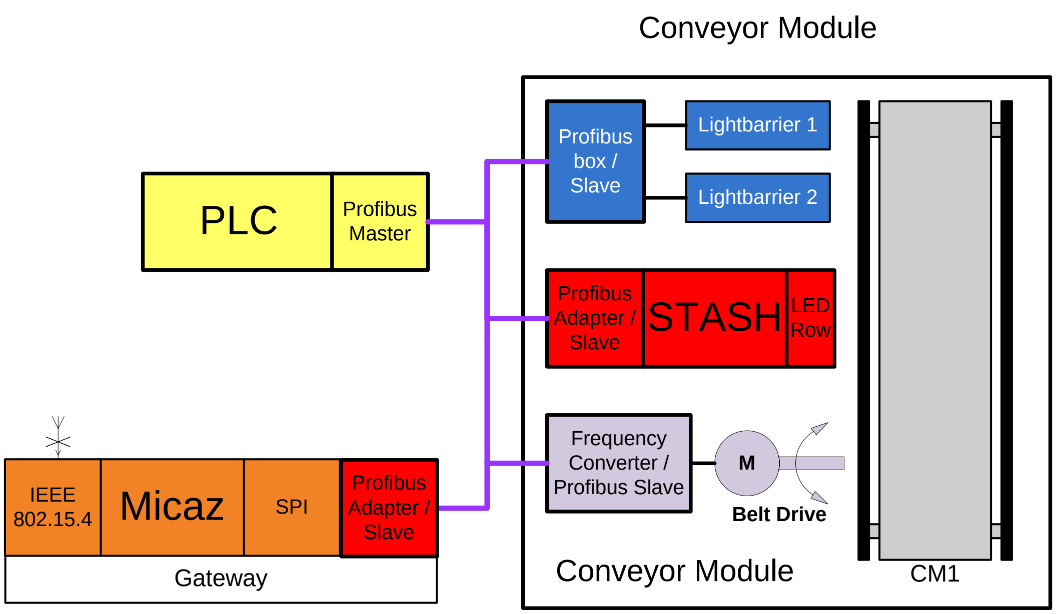


Figure 10: A conveyor module structure with necessary

5.4. Ramps

Storing items in a self-controlled shelf with only a minimum of electronics that allow easy docking to AGVs were the requirements for the storage. The ramps were custom made to correlate with these requirements. The items which are loaded on the higher side of the ramp roll down till magnetic bolts stop them (see Figure 11). One ramp can store four items at the same time and checks the slots with light barriers. They can come and go only in FIFO order. The magnetic bolts allow isolating the item on the low outgoing side. So the receiving AGV can take the item easily. The reason for this storage design is to be able to store items with a minimum of required energy consuming actuators. In this case we only need a lifting platform on the AGV and bolts on the ramps. Most common storage systems employed in warehouses or production environments are less space consuming but need a lot more mechanical parts and consume more energy to actually move items, e.g. fork lifts, automated racks, cranes etc.

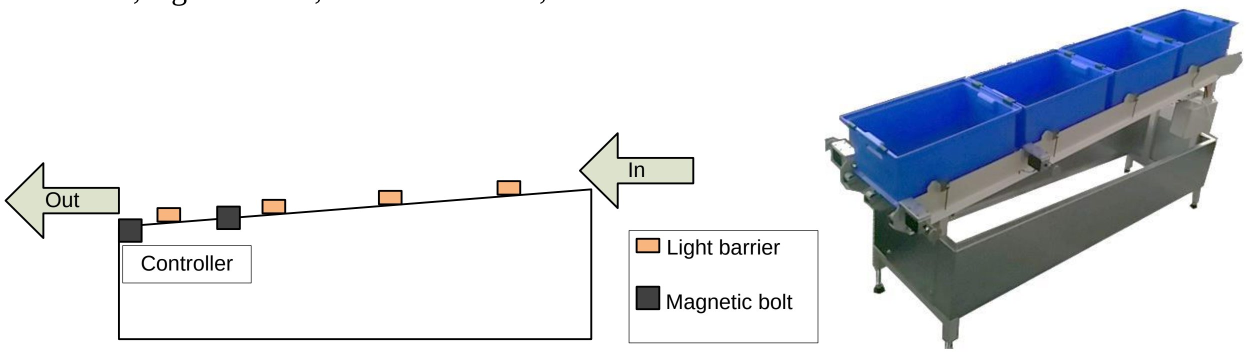


Figure 11 draft of ramp setup (left) picture of a ramp (right)

5.5 Coupled Simulation

The eleven modules of this MFS setup are only a small part of realistic plant environment. A simulation of an entire plant is developed in order to scale up the MFS to a size of a realistic scenario. In this scenario the simulations works as a test bench for different strategies and algorithms for the swarm of entities, module positioning and layout planning. The simulation enables tackling research questions, e.g. concerning the optimal amount of vehicles in a certain layout to achieve a predefined throughput or the best strategies regarding team formation. Besides the technical aspects and the roll out of the architecture on a real physical infrastructure the virtual simulation can go beyond and allows examining agent behavior in larger swarms. In a hybrid mode the simulation works in real-time and synchronizes with the physical entities by sending and receiving messages from the physical infrastructure. This allows to test the simulated results and to get a realistic evaluation of tested designs etc. The coupled simulation is implemented with the JADE agent framework.

6 Conclusion and Future Work

The work described in this paper shows how a flexible in-house logistic system for agile manufacturing environments can be created using a lightweight multiagent framework for microcontrollers running on heterogeneous material flow components. A system designed in this way can easily be fitted to arbitrary manufacturing environments. The modularity of the agents enables fast transfer to different platforms and simplifies information exchange between the various components via a wireless network. Items define their own route through the MFS and are thus able to react quickly to sudden changes. Failures of single entities can be compensated easily by substituting the affected hardware and creating a new agent. The same holds for adaptivity with respect to the number of entities active in the system which can be increased or decreased with no additional overhead other then registration.

This approach has been implemented in a physical system consisting of 11 entities and will be connected to a virtual simulation allowing macroscopic investigations on team behavior and different strategies in the future. The goal is to find solutions for an agile balancing between decentralized control driven by distributed single agent behavior and central control performed by clustered agents running optimization methods and exact algorithms for decision making on a global level. The system is designed to reconfigure itself in this respect depending on external influence factors like availability of goods, demand situation, energy price volatility, security risks and management decisions. The next steps in further developing the system will be to model these external factors in the coupled simulation and to finally realize the coupling of the agents running on the physical entities and the JADE agents from the simulation. Then it will be possible to evaluate the efficiency of the whole concept and compare it with the performance of traditional material handling systems using centralized control and fixed conveyor lines or even manually operated vehicles.

Bibliography

[1] Gunasekaran, A.: *Agile Manufacturing: The 21st Century Competitive Strategyultiagent systems: a modern approach to distribute artificial intelligence.* Elsevier Science & Technology. 2001.

[2] Weiss, G.: *Multiagent systems: a modern approach to distribute artificial intelligence.* MIT press. 1999.

[3] Günthner, W., et al*: Internet der Dinge in der Intralogistik.* Springer, Heidelberg. 2010.

[4] D'Andrea, R.: *Guest editorial: A revolution in the warehouse: A retrospective on KIVA systems and the grand challenges ahead.* IEEE Transactions on Automation Science and Engineering, Vol. 9(4), p. 638-639. 2012.

[5] Kamagaew, A., et al.: *Concept of cellular transport systems in facility logistics.* 5th International Conference on Automation, Robotics and Applications (ICARA), IEEE, p. 40-45. 2011.

[6] Seibold, Z., et al: *Layout-optimized sorting of goods with decentralized controlled conveying modules*. Systems Conference (SysCon), IEEE. 2013.

[7] Horbach, S., et al: *Building blocks for adaptable factory systems.* Robotics and Computer-Integrated Manufacturing, 27(4), 735-740. 2011.

[8] Bellifemine, F., Poggi, A., & Rimassa, G. 1999. `JADE–A FIPA-compliant agent framework`, In Proceedings of *The Practical Application of intelligent Agents and Multi-Agent Technology (PAAM),* vol. 99, pp. 97-108.

[9] Chou, Y., Ko, D. & Cheng, H. 2010. `An embeddable mobile agent platform supporting runtime code mobility, interaction and coordination of mobile agents and host systems`, *Information and Software Technology,* vol. 52, no. 2, pp. 185-196. Elsevier B.V..

[10] http://www.lynuxworks.com/products/whitepapers/xp-vs-linux.php3: conclusion last visited 15.04.2014

[11] Ogston, E., et al.: *A method for decentralized clustering in large multi-agent systems.* Proceedings of the second international joint conference on Autonomous agents and multiagent systems. ACM, 2003.

[12] Marcolino, L., et al.: *Diversity beats strength?-Towards forming a powerful team.* 15th International Workshop on Coordination, Organisations, Institutions and Norms (COIN), 2013.

[13] Koestler, A: *Beyond atomism and holism—the concept of the holon*. 1969.

[14] Fletcher, M., Deen, M.: *Fault‐tolerant holonic manufacturing systems.* Concurrency and computation: practice and experience, Vol. 13(1), p. 43-70. 2001.

[15] Gerber, C, et al.: *Holonic multi-agent systems.* 2011.

[16] ten Hompel, M., et al: *Engineering of Modular Material Flow Systems in the Internet of Things*. Automatisierungstechnik, vol. 59, no. 4, pp. 248-256. Oldenbourg Wissenschaftsverlag. 2011.

[17] Lee, E.: *Cyber Physical Systems: Design Challenges*. 11th IEEE Symposium on Object Oriented Real-Time Distributed Computing (ISORC), pp. 363-369. 2008.

[18] Hill, J.L., Culler, D.E.: *Mica: a wireless platform for deeply embedded networks*. IEEE Micro, 22(6).

[19] Stasch, A., Hahn, A.: *Towards a Multi-Agent Platform for Cyber-physical Systems based on Low-power Microcontroller for Automated Intralogistics*. ICINCO 2013 - Proceedings of the 10th International Conference on Informatics in Control, Automation and Robotics, Volume 2, Reykjavik, Iceland, 2013.

[20] Heinecke, H., et al: *AUTomotive Open System ARchitecture-an industry-wide initiative to manage the complexity of emerging automotive E/E-architectures.* Convergence International Congress & Exposition On Transportation

Electronics, Detroit, Michigan, United States, 2004.
[21] Nicol, R.C., O'Brien, P.D.: *FIPA — towards a standard for software agents*. BT Technology Journal 51-59, 1998.
[22] Lewandowski, M., et al.: *Agent-based Control for Material Handling Systems in In-House Logistics*. European Conf. on Smart Objects, Systems and Technologies. 2013.
[23] Hart, P. E., et al.: *A Formal Basis for the Heuristic Determination of Minimum Cost Paths*. IEEE Transactions on Systems Science and Cybernetics SSC4 4 (2): 100–107, 1968.
[24] Lee, C. Y.: *An Algorithm for Path Connections and Its Applications*. IRE Transactions on Electronic Computers, EC-10 (2): 346–365, 1961.

New Light Cured Media for use with

Cast Prototypes

H. Bohra [1] 1, S. N. Ramrattan [2] 2, M. K. Joyce [1] 3, P. D. Fleming III [1] 4 and P. Ikonomov [2] 5

[1]Department of Chemical and Paper Engineering
[2]Department of Industrial and Manufacturing Engineering
Western Michigan
University,
Kalamazoo, MI 49008
Email:hemant.bohra@wmich.edu

Abstract

Continuously increasing rapid casting demands in the production of functional or test prototypes have stirred the recent developments reported in this work. These technological advancements help to reduce the time till casting by introducting new techniques to produce pattern-less molds. Three-dimensional (3D) printing[1] offers the flexibility and ease of production of sand molds directly from CAD models. The new technology offers the advantages of minimal processing steps and higher precision; it also provides the capabilities to readily produce complex shaped sand molds. However, wide industrial use of thia technology has been suppressed by some limitations and concerns associated with this technology such as low process throughput, environmental impact concerns, and increased logistic costs. An alternate methodology of developing 3D cured sand molds has been proposed in this study through the introduction of a hybrid rapid prototyping[2,3] approach to overcome these limitations. The resin coated sand particles are bonded layer-by-layer after being exposed to a high intensity pulsed light source, to enable complex shapes to be obtained by precision machining[4,5] which follows the curing cycle. Although hybrid rapid prototyping techniques have been employed in the past, the methodology of developing sand molds through the integration of light curing and granular media is new. The purpose of this study was to evaluate the cured molding media for castability.

Keywords: 3D Printing, Additive Manufacturing, Prototyping, Rapid Casting.

Introduction

Manufacturing technology[6] refers to the processes involved in the production of structures with definite shapes identified by material and geometric characteristics. Various processes are involved in achieving the desired parameters, which can be classified into different groups, forming, cutting, joining, etc.

Casting can be considered the backbone of most shaping processes for various applications where metalcasting is a key component. Metal casting[7] can be explained as the process of pouring molten metal into a shaped cavity in order to transform it into a desired shape upon solidification. The primary step of this process involves the creation of a shaped cavity, also known as a mold. A mold can be prepared using various heat resistant materials such as sand, which provides the primary advantage of cost effectiveness. The molding processes can be classified into permanent and non-permanent where the mold is destroyed to remove the casted metal structure (e.g. sand molds). This step is referred to as the shakeout process. The casting process offers many advantages over other primary shaping processes, such as freedom of castability over a wide range of metals of desired shapes and versatile mechanical properties.

Achieving net near shape and integral castings leads to the reduction in the time and costs associated with any additional machining and assembly requirements. Altogether, the advantages of the process include savings in material and energy use along-with recycling and ecological benefits.

The selection of the mold material and mold making process plays a vital role in governing the casting characteristics. Sand being a refractory material can withstand very high metal pouring temperatures, which imparts the required dimensional stability at elevated temperatures. Sand also provides an economical advantage over other comparable refractory materials. Dry sand molding requires the introduction of a binder system to hold the loose sand particles during the mold shaping process. The finished mold’s physical, chemical and thermo-chemical properties depend on the sand-binder system[8] interactions and the process of creating the mold as well.

Additive manufacturing[9,10], also known as layered manufacturing, has attained acceptance in rapid prototyping and rapid manufacturing with the integration of computer aided manufacturing[11] by imparting additional opportunities for attaining complex shapes with higher precision and tolerance levels as compared to conventional molding techniques. Rapid prototyping[12] refers to the production of prototypes of actual design. It is used for the product development phase of a production process because it is able to impart the characteristics in close proximity to the finished product, which enables the further investigation and analysis of the product before finalizing the end product features. The rapid casting[13] process defines the integration of traditional metal casting techniques with additive manufacturing approaches to achieve either functional prototypes or end products.

Various rapid casting solutions have been developed during the last few decades to implement concurrent engineering approaches for the development of functional prototypes and customized production of metal castings for applications in various sectors of manufacturing industries. These rapid casting solutions include unique production approaches, such as selective laser sintering, fused deposition modeling, sterolithography, 3-D printing and rapid tooling[14-18].

Experimental Methodology

The procedure consisted of 4 major steps:

1) Characterization of granular media.
2) Preparation of 50.0 mm diameter x 8.0 mm thick disc-shaped specimens.
3) Testing of the various properties of the specimen.

Characterization of Granular media:

Resin coated[19] shell sand with 3% binder loading will be evaluated and characterized for density, Grain Fineness Number, pH, shape and sphericity or roundness (Refer Table 1).

The bulk density of sand plays a vital role in the estimation of the finished sand mold properties from a handling and logistical prospective. Bulk density is also significant in tracking any lot-to-lot variations in the sand systems employed. The density of a system is representative of mass/volume ratio and was measured by weighing a known volume of material and thereafter calculating the ratio.

The Grain Fineness Number[7] is an indicative of how small or large is the average grain size of a sample of sand. Though smaller grains of sand are capable of imparting better surface finish, due to their closely packed structure, they hinder the venting characteristics of the mold, which can result in various casting defects. The higher the GFN, the smaller the sand grain size. Very low GFN refers to larger sand grain sizes compromising the packing efficiency consisting of larger interstices between the grains and hence resulting in rough casting surfaces. To calculate the GFN, a known sample weight was placed into an arrangement of sieves, which was then placed in a sieve-testing device with continuous sieve shaking for 15 minutes. Thereafter, the weight of sand retained in each sieve and pan was recorded followed by calculating GFN as shown in Table 2:

Table 1. Properties of Sand

Shape	**Roundness/ Sphericity (Krumbein)**	**pH**
Sub-angular	0.5 / 0.7	7.4

Table 2. AFS GFN calculations for the resin coated sand

Sieve	**Amount Retained**	**%**	**Multiplier**	**Product**
6	0	0	0.03	0
12	0	0	0.05	0
20	0.01	0.01	0.1	0.00
30	1.94	1.95	0.2	0.39
40	3.99	4.01	0.3	1.20
50	15.64	15.71	0.4	6.28
70	26.14	26.26	0.5	13.13
100	33.08	33.23	0.7	23.26
140	16.01	16.08	1	16.08
200	2.57	2.58	1.4	3.62
270	0.17	0.17	2	0.33
PAN	0	0.00	3	0.00
Total=	99.54		**AFS GFN**	**64.30**

Preparation of 3D CURED Disc-Shaped Specimen:

A laboratory scale assembly designed at Western Michigan University and integrated with the light source was used for preparing specimens. A 50 mm diameter mold cavity was adjusted and filled with resin coated sand particles to achieve 8 mm-thick disc specimens. The excess sand was screened off to achieve uniform layer and exposed under the light source (Xenon Sinteron 2000 photonic light curing) for 90 sec. duration. For 2-pass 8mm thick specimens the mold cavity was adjusted to 4mm thickness, filled and exposed for 45 sec, thereafter the first layer was dropped down to 4mm from the mold reference point so as to again obtain the 4mm thick cavity and was filled and exposed with the same procedure as followed for the first layer.

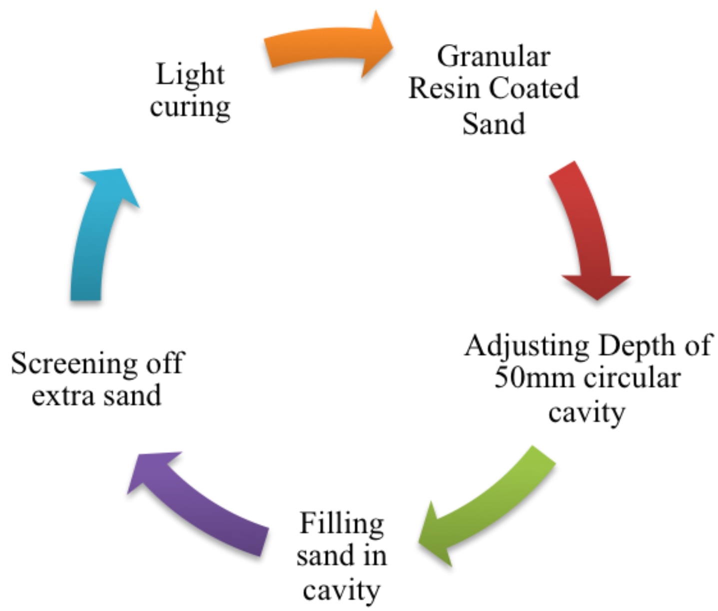


Figure 1: Steps to produce specimens

Physical Properties:

The thicknesses of the specimens were measured and recorded. Disc-shaped specimens were weighed prior to conducting any tests using a digital balance and the thicknesses were measured using a digital caliper.

The porosity of a material can be defined, as the measure of total void volume present, as such is dimensionless. Darcy's Equation[30] defines the relationship between permeability and flow rate as shown in Equation 1:

$$Q = (k\ A\ \Delta P)\ /\eta\ L \qquad \text{Eq. 1}$$

Where, k = Permeability Coefficient (Area)

A = Area of cross-section

L = Thickness of Material

ΔP = Pressure drop

Q = Flow rate (Volume/s)

η = Viscosity

AFS Permeability is a measure of gas flow through a porous medium, such as a sand mold or core. Permeability and MQI tests were performed to study the venting characteristics of specimens as good venting parameters play a major role in elimination of casting defects originating due to gas evolution at mold-metal interface.

A Gerosa Simpson permeability tester was used to measure AFS permeability. MQI was measured using Disa George Fisher Mold Quality Indicator.

Mechanical Properties:

Disc transverse strength is a measure of transverse tensile strength of bonded sand particles. The disc transverse strength is used in the foundry industry to study the parameters required for mold or core against transverse tensile stresses acting on mold-metal interface being resultant of the hydrostatic pressure. Therefore, a disc transverse strength test (DTS) was used to measure the strength of the disc shaped sand specimens using a disc transverse testing machine (Dietert Model 490-A) and was supported on its ends. During the test run, the specimen is subjected to a linearly increasing load and the maximum load to failure is then determined.

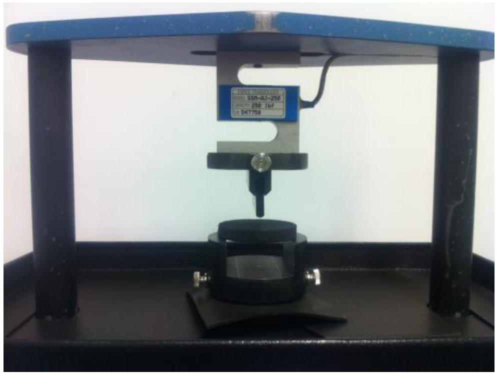


Figure 2: DTS Tester

Thermo-mechanical Properties:

Thermal Distortion Testing (TDT)[31-33] was used in this experiment to expose 50 mm diameter, 8 mm-thick disc specimens to 1000ºC for up to 90 seconds. The purpose of this investigation was to test the 3D cured sand system used for rapid prototyping manufacturing The TDT method was used for application on foundry mold and core media, and to further compare the data.

The longitudinal distortion of specimens can be differentiated vide expansion (DE) and plastic distortion (DP) considering the observations from the thermal distortion curve. The total longitudinal distortion (TD) can be represented as the sum of expansion and plastic distortion components (TD = DE + DP). The radial distortion (DR) indicating expansion in the radial direction will be studied in further studies using next generation thermal distortion tester (TDTng) developed at WMU.

Each disc shaped specimen was weighed prior to the thermal distortion testing and was blown with 20 psi (0.14 MPa) air pressure to remove any loose sand grains at the completion of the distortion testing. The change in mass of the specimens before and after thermal distortion testing was recorded.

Casting Trial:

For conducting casting trial so as to evaluate the castability of molding media, a 56 mm thick and 50 mm diameter disc shaped specimen was prepared using layer-by-layer approach with each layer thickness being 8mm. The procedure for preparing the specimen was similar as stated in Fig. 1. Two holes with approx. 13 mm diameter were drilled into the specimen and refractory coating[34] was thereafter applied on one using the dip coating method to analyze the effect of refractory coating application on castability. Aluminum was thereafter melted in a small crucible and poured at a pouring temperature of 760°C. A manual shakeout procedure was followed after the solidification of the casting to separate the metalcasting from the mold media.

Results & Discussion

The weight for each disc shaped specimen was recorded prior to conducting further testing to keep track of any variability in each sample set.

Table 3. Average weight of disc-shaped specimens

Specimen	Weight(g)	Density(g/cm3)
3D Cured 1-Layer	16.73	1.07
3D Cured 2-Layer	15.95	1.02

A light cured single pass 8mm thick disc specimens tested using the custom made specimen holder and plug, had a permeability number of 170 ±3 and MQI of 157 ±3; while 2-pass 8mm thickness specimens had a permeability number of 173 ±3 and 149 ±5 respectively (Fig. 3). An inverse relationship between permeability and MQI has already been established in the foundries for other sand systems such as green sand. The results from permeability and MQI tests suggest that the proposed sand binder system also follows an inverse relationship trend (Refer Fig. 3).

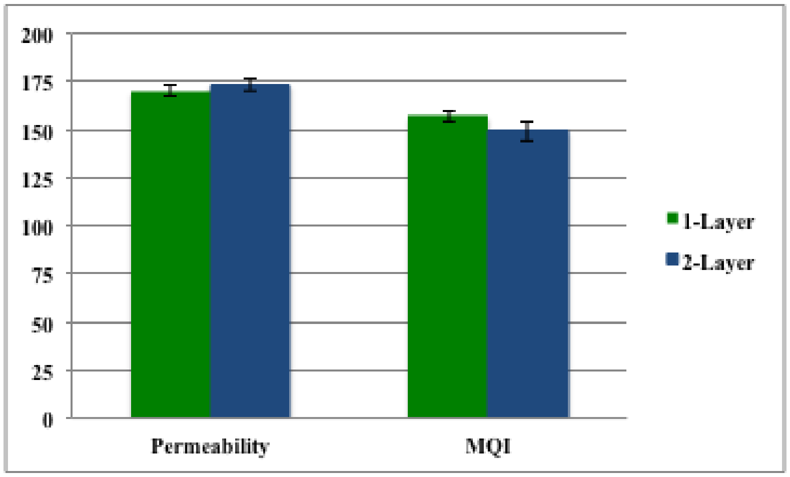


Fig 3. Permeability and MQI data for disc shaped sand specimens

The 3D cured disc-shaped specimens prepared with 1-Layer and 2-Layers were tested at elevated temperature over a 90 second interval for thermal distortion testing. The TDT curves for both specimens (Fig. 4) showed thermo-mechanical and thermo-chemical changes in these systems at the elevated temperature. The longitudinal distortion curves showed expansion (DE) followed by plastic deformation (DP). The TDT curves for longitudinal distortion depict an average for 3 specimens.

The 3D light cured 1-layer specimens expanded for approx. 46 seconds followed by plastic deformation for approx. 44 seconds whereas 3D light cured 2-layer specimens expanded for ~48 seconds followed by plastic deformation for the remaining ~42 seconds of the test.

Heat induced thermo-chemical reactions occurring in both sets of specimens were evident from the surface cracks found on tested specimens and percent change in mass values (Table 4).

Table 4. Physical, Mechanical, and Thermo-Mechanical Properties of the Discs Specimens

Sample	Blow Pressure (psi)	(DE) Longitudinal (mm)	(DP) Longitudinal (mm)	(TD) Total Dist. (mm)	% Change in Mass	Cracks & Fractures
1-Layer	20	0.049	0.041	0.090	0.64	present
2-Layer	20	0.028	0.020	0.048	0.64	present

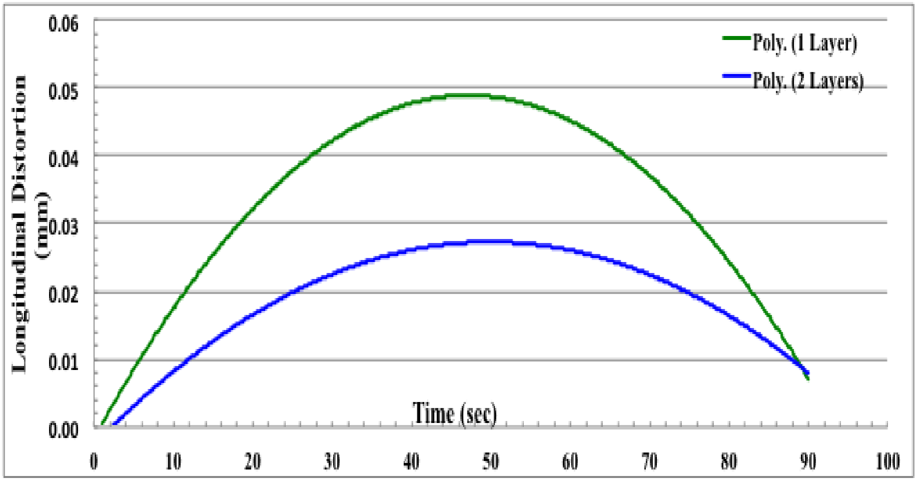


Figure 4: TDT Curves for 3D Light Cured Sand Specimens

The observations from the casting trial (Fig. 5) show the evidence of presence of surface casting defects such as penetration with the absence of the refractory coating applications. The casted metal surface obtained with the application of refractory coating imparted better surface finish and did not incorporated any surface defects subjected to penetration and veining.

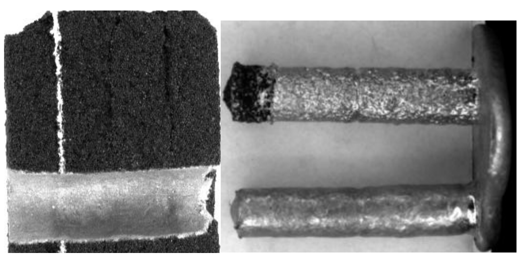


Figure 5: Coated and uncoated mold cross-sections shown on the left and metal casting obtained from corresponding surfaces is shown on the right.

Conclusions

The observations from the experimental analysis suggest that the proposed media for the light cured system possess the desired physical, mechanical and thermo-mechanical properties for castability. These light cured sand specimens are quite stable at elevated temperatures as revealed from the thermal distorting testing. Since, the proposed sand binder system may not be compatible with certain casting alloy chemistries; a refractory coating interface may be required.

Acknowledgments

The authors gratefully acknowledge Glenn Hall, Matt Stoops (Western Michigan University), and Xenon Corporation for their technical support.

References

1. R. Hochsmann; "Method for the layered construction of models"; US Patent 8,020,604 B2, Sep. 20, 2011.
2. I. Brensons, N. Mozga; "Research on rapid prototyping influence in manufacturing technology"; Trends in Development of Machinery and Associated Technology, 2011.
3. J Hur, K Lee, J Kim; "Hybrid rapid prototyping system using machining and deposition"; Computer-Aided Design, 2002.
4. K.P. Karunakaran, S. Suryakumar, V. Pushpa, S. Akula; "Low cost integration of additive and subtractive processes for hybrid layered manufacturing"; Robotics and Computer-Integrated Manufacturing, 2010.
5. C.G. Basila, Z. Baki-hari; " The main applications of rapid tooling"; International Conference on Economic Engineering and Manufacturing Systems, 2007.
6. Ed. K. H. Grote, Erik K. Antonsson; "Handbook of Mechanical Engineering", 2009.
7. F. P. Schleg, F. H. Kohloff, J. Radia, D. Oman, J. Schifo, M. Gwyn; "Technology of Metalcasting", American Foundry Society, 2008.
8. Carey, R. P. et al., "Sand Binder Systems," Foundry Management & Technology, A Penton Publication, Bulletin 2292-1 (1998).
9. R. Beaudoin, P. Carey, T. Sorovetz; "Direct Shell Sand Rapid Prototyping: From CAD to Casting in Days", Modern Casting, Nov. 1997.
10. B. Jang, E. Ma; " Layer-additive method and apparatus for freeform fabrication of 3-D objects"; US Patent 2002/0062909 A1, May 30, 2002.
11. A. Bernard, Jean C. Delplace, N. Perry, S. Gabriel; "Integration of CAD and rapid manufacturing for sand casting optimization", Rapid Prototyping Journal, 2003.
12. Ortiz, et al. "Rapid Prototype Casting", United States Patent 7,448,433, Nov. 11, 2008.
13. M. Chhabra, R. Singh; "Rapid casting solutions: a review", Rapid Prototyping Journal, 2011.
14. G. N. Levy, R. Schindel, J. P. Kruth, K. U. Leven; "Rapid Manufacturing and Rapid Tooling with Layer Manufacturing (LM) Technologies, State of the art and Future Perspectives", CIRP Annals- Manufacturing Technology, 2003.
15. J. W. Barlow, N. K. Veil; "Binder composition for laser sintering process", US Patent 6,048,954, Apr. 11, 2000.
16. J. P. Kruth, G. Levy, F. Klocke, T. H. C. Childs; "Consolidation phenomena in laser and powder-bed based layered manufacturing", 36th International MATADOR Conference Proceeding, 2007.
17. J. Kawola; "Zcast Direct Metal Casting from Data to Cast Aluminium in 12 hours", April 2003.
18. S. S. Gill, M. Kaplas; "Comparative Study of 3D Printing Technologies for Rapid Casting of Aluminium Alloy", Material and Manufacturing Processes, 2009.
19. Okubo, T. Murakami, F. Takeuchi, K. Shiduka; "Resin – coated sand", US Patent 7,267,876 B2, Sep. 11, 2007.
20. Hentschel, et al.; "Method for milling casting moulds", United States Patent 7,025,108, April 11, 2006.
21. S. Wetzel; "Offering Customers Rapid Manufacturing", Modern Casting, 2012.
22. D. Kochan, Chua C. Kai, Du Zhaohui; "Rapid prototyping issues in the 21st century", Computers in Industry, 1999.
23. http://www.xenoncorp.com/uv-curing/Press/sinteron2000_nr.html
24. C.M. Cheah, C.K. Chua, C.W. Lee, C. Feng, K. Totong; "Rapid prototyping and tooling techniques: a review of applications for rapid investment casting", International Journal for Advanced Manufacturing Technology, 2005.
25. "Analysis of Casting Defects"; American Foundrymen's Association, 1947.
26. "Atlas of Defects in Castings", Institute of British Foundrymen, Second Edition, 1963.
27. "Casting Design and Performance", ASM International, 2009.
28. Guyer, O.B., Emptage, R.C., Ramrattan, S.N., "Controlling Heat Transfer with Refractory Coating", Modern Casting, Feb. 2006.
29. Guyer, O.B., Emptage, R.C., Ramrattan, S.N., "The Effect of Refractory Coating Thermal Heat Transfer on Phenolic Urethane Cold Box Core Distortion at Iron Temperature and Pressure," AFS Transactions, Paper 05-192 (2005).
30. L. Pal, M.K. Joyce, P.D. Fleming; "A Simple Method for Calculation of the Permeability Coefficient of Porous Media", TAPPI Journal, 2006.

31. Ramrattan, S., Vellanki, S., Jideaku, O. Huang, C., “Thermal Distortion in Process Control of Chemically-Bonded Sands,” AFS Transactions, vol 152, pp. 161-165 (1997).
32. Iyer, R., Ramrattan, S., Lannutti, J., Li, W., “Thermo-Mechanical Properties of Chemically Bonded Sands,” AFS Transactions, vol 109, pp. 1-9 (2001).
33. Rebros M., Ramrattan S. N., Joyce M. K., Ikonomov P., “Behavior of 3D Printed Sand at Elevated Temperature”, 111th Metalcasting Congress, 15-18 May (2007), Houston, TX
34. Guyer, O.B., Emptage, R.C., Ramrattan, S.N., “Controlling Heat Transfer with Refractory Coating”, Modern Casting, Feb. 2006.

Biographical Information

Hemant Bohra graduated with B.Tech in Printing Technology from G.J. University of Science and Technology, INDIA in 2005. He received his master’s degree from Department of Paper Engineering, Chemical Engineering and Imaging, Western Michigan University in December 2009. He is currently pursuing Ph.D in Engineering and Applied Sciences, Western Michigan University. He has worked on various projects during this period involving paper manufacturing, coating formulations & its applications, printed electronics, ink formulations and conventional & digital printing technologies. He is currently working on development of a new rapid casting/prototyping technique implementing hybrid manufacturing approach.

A short review on Crystal Clear methodology and its advantages over scrum, the popular software process model

Makan Taghavi Dilamani

Department of Computer Science

Islamic Azad University Central Tehran Branch

Tehran, Iran

Makan.taghavi@ieee.org

Abstract

In recent years agile software development methods have gained considerable popularity among managers and computer science specialists. Versatility of this methods compared to traditional models has caused wide range of companies to switch to Agile processes. Since the establishment of agile manifesto in 2001, numerous models have been offered in order to pave the way for companies to find the best fit Agile process to adopt with. This paper briefly sheds light on Crystal Clear software process model and compares it with Scrum, the most approached Agile method. The purpose of this is to draw the attention of software engineers and project managers to the noteworthy advantages of Crystal Clear methodology and qualities it offers in terms of adaptability and stability especially for small development teams.

1-Introduction

There have always been endemic problems in the software development process since the term of software engineering was first popularized by Friedrich L.Bauerin in Nato conference in 1968. Namely, that systems cost too much, take too long to deliver, projects go over budget and when they eventually get delivered they do not serve their intended purposes(1). In order to solve these issues a considerable number of software development models have been introduced over last 40 years, but none have proven to be absolutely reliable(2).

Recently agile methods have emerged as a popular approach towards software development. They claim to solve the chronic problems which exist in the field more effectively compare to previous models. However, switching to this model requires notable flexibility and risk taking. Each of the agile methods is designed to fit development teams based on their situation and needs. The aim of this article is to draw the attention of software engineers and managers to different tailoring ability of two of the agile methods prior to switching to this process model. We focus on Crystal Clear method and explicate its characteristics and its highly flexible features in terms of tailoring and fitting in different software teams especially small ones. Also we compare this method to Scrum, the most approached agile process, and point out the subtle advantages of Crystal Clear method over this model.

Proceedings of the 2014 ICAM, International Conference on Advanced and Agile Manufacturing, Held at Oakland University, Rochester, MI 48309, USA

2- Crystal Clear Method's Advantages

Crystal methodologies were created by Alistair Cockborn, one of the founders of Agile manifesto. These methods are set of processes that can be applied to different projects depending on the size and the complexity of it(2). One of the most flexible methods of Crystal family is the Crystal Clear which is mainly focuses on projects consisting of 6 to 8 developers. However this method takes advantage of its small team size and proximity to strengthen **Close Communication** into a more powerful **Osmotic Communication** which leaves it with enough tailoring ability to fit on almost any kind of project, not just small teams(3). Crystal Clear basically relies on three main properties:

1- Frequent Delivery
2- Reflective Improvement
3- Close Communication

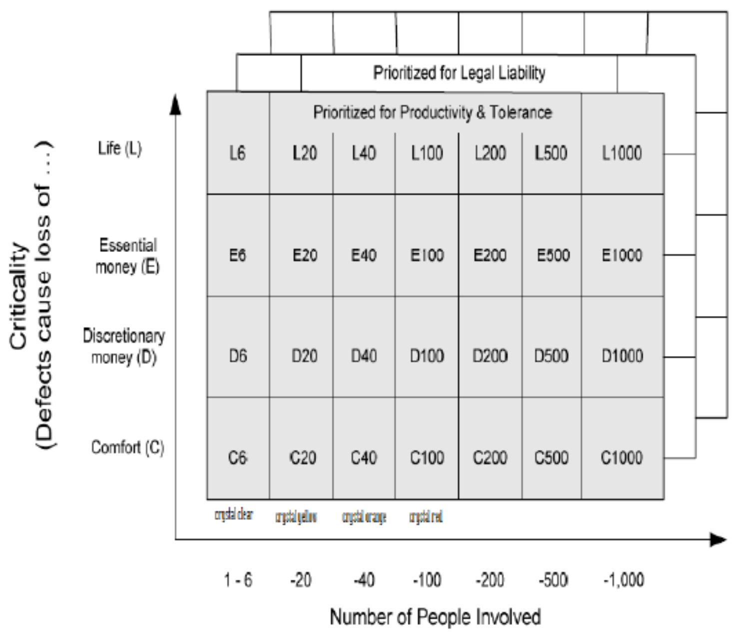


Figure1. *Crystal Methods(Taken from(2))*

2-1 Frequent Delivery

Frequent delivery is the main common feature in all the agile methods. In Crystal Clear the maximum length of each time box is 4 month. During each iteration, having easy access to expert users and plans for frequent meet ups is crucial in order to provide both the customers and developers enough assurance to follow with the rest of the iteration. According to this method development teams gather after the working software is delivered to measure the team's velocity and success ratio. In Scrum method, iteration length is fixed to 30 days(4). There is no doubt that having a fixed end date is helpful for the team because with this strategy they get an idea of how much they can achieve in a certain period of which is a useful feedback and also motivates the team with and early victory. However it can gradually impose a considerable amount of stress and pressure on development teams. As well probable failure of having delivery and meeting checkpoints on time while having a fixed deadline of 30 days can lead to failure in the iterations. This can damage the team's spirit and productivity and will consequently increase the amount of undone work which is precisely against the properties of the Agile manifesto(5, 6).

2-2 Reflective Improvement

The nature of software development process is tied to rapid changes in technology and requirements during the project. The people on the team are the best equipped to say what is best suited to their situation, which is why Crystal Clear leaves so many details unstated, but for the team to finalize(3). The Reflective Improvement mechanism allows them to make those adjustments. The main point of difference between Crystal Clear and Scrum in this case is that in the latter, all the members have to meet after every iteration in the Reflection Workshops to discuss the existing problems whereas Crystal Clear leaves it all to the team members to reflect their situation by not requiring them to spend great deal of time on long hour meetings(4).

2-3 Close Communication

One of the unique properties of Crystal Clear method is its focus on the way the team members communicate with each other. It strongly recommends developers to have Osmotic Communication with each other, in other words the information flows into the background hearing of members of the team, so that they pick up relevant information as though by osmosis. This is normally accomplished by seating them in the same room. Then, when one person asks a question, others in the room can either tune in or tune out, contributing to the discussion or continuing with their work(3).
On the scrum method, the main communication between the team members occurs during daily stand up meetings for 15 minutes(4). Since in small teams each of the members may have various roles and dedication to different parts of the project having the constant presence of all the members is not always feasible and limiting their communication to just a few minutes can negatively affect the code integrity of the development team. Moreover, osmotic communication reduces the cost of communications and increases the feedback rate in a way that bugs get detected and removed quickly and more effectively.

	Customer Collaboration	**Time to market**	**Responding to Change**	**Documentation**
Crystal Clear	Direct user involvement	1-4 months	Multiple tuning technique Short release Direct user involvement	Object Models User Manuals Test Cases
Scrum	Create and priorities product backlog	30 Days	Daily Scrum Meeting Short Release	Product backlog list Sprint backlog list

Figure 2. *An overview of Crystal Clear and Scrum methodologies' main properties*(Taken from(2))

4- Conclusion

In this paper we tried to briefly explicate Crystal Clear methodology and introduce it as a possible solution for companies struggling with their adaptation on heavy weight methodologies. There has been very little research to date on method tailoring. What we believe is one of the main reasons many companies mistakenly sought after popular methodologies and fail to adapt with them.

As a future work, a rigorous methodological testing approach in comparing different Agile methods with respects to system critical variables may need to be included.

Bibliography

1. Conboy K, Fitzgerald B. Method and developer characteristics for effective agile method tailoring. ACM Transactions on Software Engineering and Methodology. 2010;20(1):1-30.
2. Qasaimeh M, Mehrfard H, Hamou-Lhadj A. Comparing Agile Software Processes Based on the Software Development Project Requirements. 2008:49-54.
3. Cockburn A. Crystal Clear
A Human-Powered Methodology For Small Teams2004.
4. J.Sutherland KSa. The Scrum Guide2013.
5. [cited 2014 12/03/2014]; Available from: www.agilealiance.com.
6. Bohner MCaS. The Impact of Agile Methods on Software Project Management. 2005.

Formability and Crystallographic Texture in Novel Magnesium Alloys

Pnina Ari-Gur [1], Andreas Quojo Quainoo[1], Shubram Subramanyam[1], Ashkan Razania [1], Sven Vogel [2]

[1]Department of Mechanical and Aerospace Engineering, Western Michigan University, Kalamazoo, MI 49008, [2]Materials Science and Technical Division, Los Alamos National Laboratory, MST-8, MS G755, Los Alamos, NM 87545
Email: pnina.ari-gur@wmich.edu

Abstract

Magnesium alloys offer many advantages. They offer very low density and good strength. They also offer good damping properties. One of the industries where reducing component weight is the automotive industry. That makes the magnesium alloys good candidates for these applications. Reduced weight of an automobile means also lower fuel consumption. The hexagonal closed packed structure of magnesium, lends itself to strong mechanical anisotropy. In the current work, neutron diffraction was used to study the crystallographic texture developed in novel magnesium alloys during cold rolling operations. The texture was compared with that developed in the commercial AZ-31 magnesium alloy. Tests were run at the High-Pressure-Preferred-Orientation (HIPPO) beam line at Los Alamos National Lab. The texture was then analyzed using pole figures, created using the Material Analysis Using Diffraction (MAUD) software.

Introduction

Magnesium alloys have high strength-to-weight ratio and that makes them an attractive substitution to heavier alloys, including the light-weight aluminum. In addition, magnesium alloys offer good damping properties [1, 2]. Plastic deformation in Mg alloys is strongly affected by the initial texture, and by process conditions, since the operation of non-basal slip systems require thermal activation [3]. The insufficient number of slip and twinning systems, results in the magnesium alloy demonstrating poor formability at room temperature. This low ductility has prevented their wide-spread application as cold-formed components [4]. The hexagonal closed packed (HCP) structure of magnesium, lends itself to strong mechanical anisotropy, especially when a significant crystallographic texture is present. Other properties, such as damping strongly depend on the texture [1, 2]. *HCP* metals can be classified into two groups according to their principal slip systems: basal slip metals (e.g. Mg and Zn), and prismatic slip metals (e.g. Ti and Zr). When the *c/a* ratio of the *HCP* lattice differs from the ideal value (1.633), the relative close packing of the crystal planes will vary. These changes influence the slip behavior during plastic deformation [5]. In some *HCP* metals like Mg, with limited ductility, twinning only works by the most common types of twins [6]. Operating a particular twinning system depends not only on the magnitude of the shear stress, but also on the *c/a* ratio. For pure Mg the ratio is 1.624 (close to the ideal 1.633). The *c/a* ratio may be altered by alloying elements; and this may change the active slip systems at RT. Work by Gao and Liu [7] on AZ-31 alloyed with small addition of Sn and Pb showed improved ductility. The present work was aimed at analyzing the crystallographic texture developed in these novel magnesium alloys during cold rolling operations. The baseline for comparison was commercial AZ-31.

Experimental description

The alloys were prepared at the Light Metal Centre, University of Auckland, New Zealand. [Gao et al. Private communications]. They were made by melting commercially pure metals in a low carbon steel crucible under RJ-2 fusing agent. The melt was stirred to ensure homogeneity and cast into pre-heated (523 K) iron molds. The casting temperature was 983 K.

Three Mg alloys were studied; their compositions are described in Table 1.

Table 1: Alloys composition

Alloy	Composition
AZ31 (reference)	96%Mg-3%Al-1%Zn
AZ31-2Sn	94%Mg-3%Al-1%Zn-2%Sn
AZ31-2Sn-1Pb	93%Mg-3%Al-1%Zn-2%Sn-1%Pb

The geometry and procedure for cold rolling of the samples is described in Fig. 1

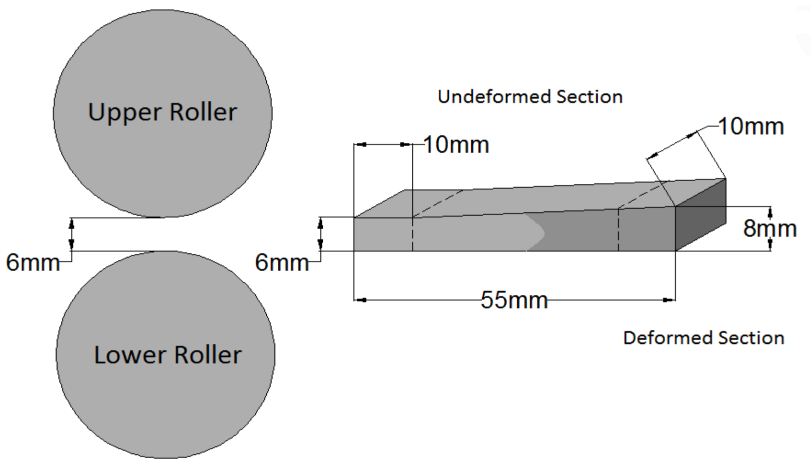


Figure 1. Sample geometry prior to the cold-rolling process

Table 2: Samples nomenclature

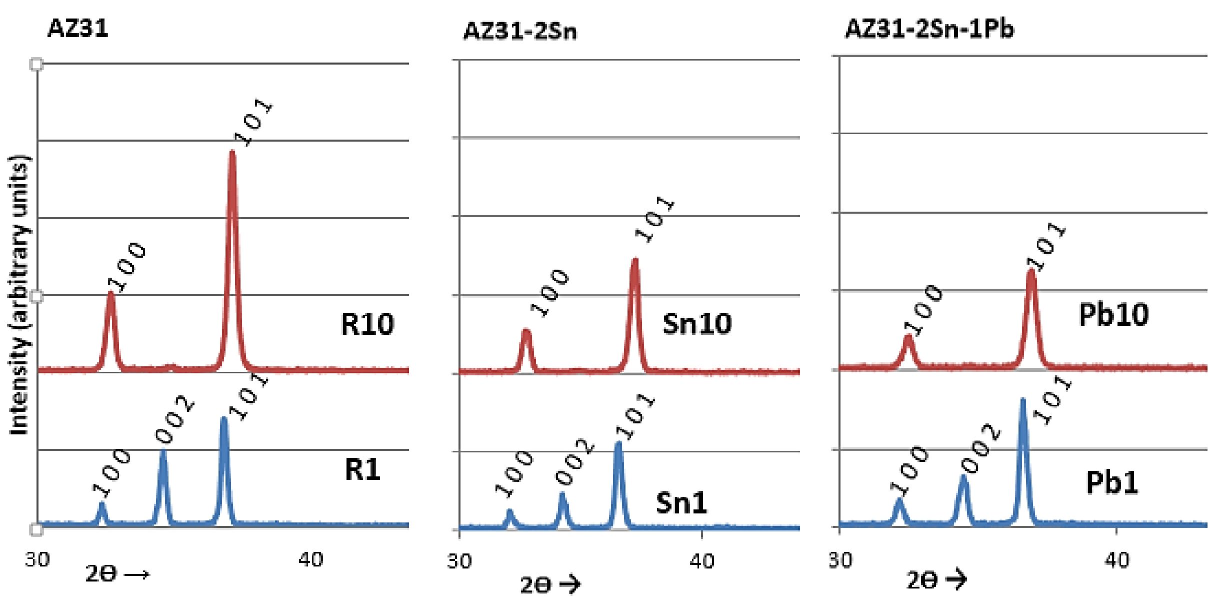
Alloy	Section No.	Sample ID	% RA
AZ31	1	R-1	8.80
AZ31	10	R-10	23.19
AZ31-2Sn	1	Sn-1	7.85
AZ31-2Sn	10	Sn-10	21.10
AZ31-2Sn-1Pb	1	Pb-1	7.20
AZ31-2Sn-1Pb	10	Pb-10	21.11

X-ray diffraction was conducted on the surface perpendicular to the RD-TD plane (i.e. on the RD-ND plane). Panalytical X-ray diffractometer with CuKα radiation was used. The system is equipped with a monochromator, a stepping motor and an MDI control and digital output.

Neutron diffraction tests were conducted at the High-Pressure-Preferred-Orientation (HIPPO) beam line at Los Alamos National Laboratory. It is designed to collect data for texture analysis. The crystallographic preferred orientation was analyzed using pole figures, created from the data using the Material Analysis Using Diffraction (MAUD) software [8].

Results and discussion

X-ray diffractograms of the samples are shown in Fig. 2 (a-c). Comparison was made between the diffractogram of the slightly deformed alloys and the highly deformed ones, as well as between the different alloys. Most notable in all is the absence of the basal plane reflection in the highly deformed sample. This is a result of the strong texture developed that causes the basal plane to rotate to the rolling plane. Some small differences between the three alloys can be seen too. Most notable is the strengthening of the pyramidal plane in the RD-ND plane in the reference alloy (AZ-31). Because the information obtained from the X-ray diffraction is limited to some 50 μm under the surface, and it provides data only about crystallographic planes parallel to the surface, neutron diffraction was performed that provides information over the whole volume as well as complete pole figures.



(a) (b) (c)

Figure2. X-ray diffraction of the three alloys before rolling and after heavy deformation

Pole figures of the three alloys are shown in Fig. 3 (basal plane {0001}) and in Fig. 4 (prismatic plane$\{10\bar{1}0\}$).

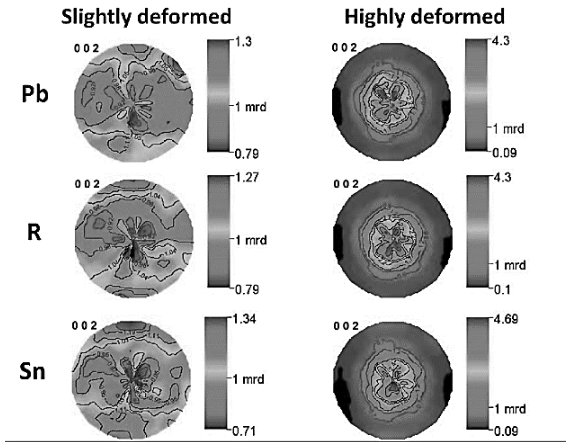


Figure 3. Basal plane pole figures (For sample nomenclature refer to Table 1)

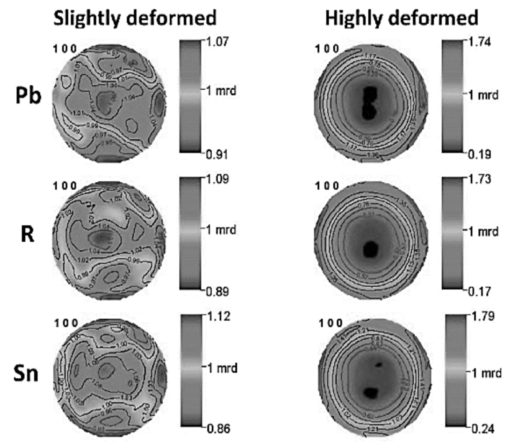


Figure 4. Prismatic plane pole figures (For sample nomenclature refer to Table 1)

From the pole figures, it is evident that the initial texture is far from being random. Comparing with Fig.1 at Wang and Huang [9], it is seen that after more than 20% engineering strain ('highly deformed'), the texture of the three alloys is mostly a {0001} fiber texture. The strength of the texture somewhat differs between the three alloys, but its nature remains the same. The slight deviation from a perfect fiber texture, results from the initial texture at the "slightly deformed" state. At that state, it can be characterized as a weak $\{10\bar{1}0\}$<0001> texture. During rolling, this texture is first "destroyed" and then the fiber texture gradually develops.

Conclusions

Rolling ductile Mg alloys (AZ31 plus Sn and Pb), develops {0001} fiber texture similar in nature to the one found in unalloyed Mg [2]. It is concluded that the same slip systems and twinning mechanism operate, and the source of the observed difference in mechanical behavior [10] has other sources, such as microstructure and precipitates formed with the addition of the new elements.

Acknowledgements

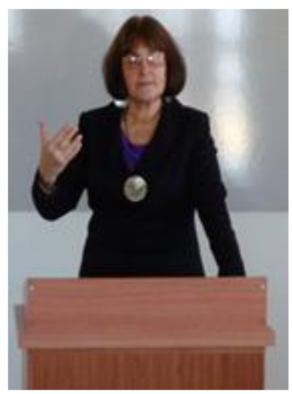
The authors are grateful for the samples received from the Light Metals Center, at the University of Auckland, New Zealand (Drs. Wei Gao and Tainping Zhu), and to students Andreas Erbisch, Ernesto J Matos Ojeda, Alejandro Carrasco Pena, and Francisco a Yaporgenao for their assistance.

Bibliography

1. HU, X. S., WANG, X. J., HE, X. D., WU, K., & ZHENG, M. Y. (2012). Low frequency damping capacities of commercial pure magnesium. Transactions of Nonferrous Metals Society of China, 22(8), 1907-1911.
2. Suzuki, K., Chino, Y., Huang, X., & Mabuchi, M. (2011). Elastic and Damping Properties of AZ 31 Magnesium Alloy Sheet Processed by High-Temperature Rolling. Materials Transactions, 52(11), 2040-2044.
3. Al-Samman, T., & Gottstein, G. (2008). Room temperature formability of a magnesium AZ31 alloy: Examining the role of texture on the deformation mechanisms. Materials Science and Engineering: A, 488(1), 406-414.
4. Barnett, M. R., Jacob, S., Gerard, B. F., & Mullins, J. G. (2008). Necking and failure at low strains in a coarse-grained wrought Mg alloy. Scripta materialia, 59(10), 1035-1038.
5. Honeycombe, R.W.K., *The plastic deformation of metals*. 2nd ed. 1984: London: Edward Arnold.
6. Rosi, F. D., Dube, C. A., & Alexander, B. H. (1953). Mechanism of plastic flow in titanium-determination of slip and twinning elements. J. Metals, 5(2), 257-264.
7. Gao, W., & Liu, H. (2009, August). A highly ductile magnesium alloy system. In IOP Conference Series: Materials Science and Engineering (Vol. 4, No. 1, p. 012003). IOP Publishing.
8. Wenk, H., Lutterotti, L., & Vogel, S. (2010). Rietveld texture analysis from TOF neutron diffraction data. *Powder Diffraction*, 25(3), 283-296.
9. Wang,Y.N., and Huang J.C. (2003). Texture analysis in hexagonal materials. *Materials Chemistry and Physics* 81, 11–26
10. Wei, S., Zhu, T., Hodgson., Gao, W. (2013). Effects of Sn addition on the microstructure and mechanical properties of as-cast, rolled and annealed Mg–4Zn alloys. *Materials Science and Engineering*, A 585, 139-148

Biographical Information

Professor Pnina Ari-Gur, received her Doctor of Science degree in Material Engineering from the Technion, Israel Institute of Technology. After a post doctorate research at the University of British Columbia, as well as R&D at the aerospace industry, she started her career at WMU engineering. She has active research collaboration around the globe. Her work is published in engineering and science journals. In addition to her basic research, she conducts some applied projects for the benefit of providing experience to her students.



Andreas Quojo Quainoo, Shubram Subramanyam, and *Ashkan Razania* are students in the Mechanical and Aerospace Department at Western Michigan University. They came to WMU from Ghana, Malaysia and Iran.

3D Metal Printing/Machining

[1]**Pavel Ikonomov and** [2]**Jorge Rodriguez**
[1,2] Industrial Manufacturing and Engineering Department,
Western Michigan University
Kalamazoo, Michigan 49008
e-mail:pavel.ikonomov@wmich.edu

A

Abstract

The goal of this research is to design and build a hybrid 3D metal printer/machining system capable to produce functional high quality parts. It is well know that 3D printing became affordable but still could not provide high precision and finish quality, on the other hand machining process can produce high quality parts, but at higher cost, waste of material, and time. This 3D metal printing/machining system takes advantages of the both methods providing fast, efficient, and low cost production. As an additional benefit the proposed system is user friendly, so person without experience and knowledge how machine works can produce functional, high quality products. The 3D metal printer/machining system is easy to build, maintain, and affordable so average user could even make it themselves. The machine works by depositing metal (3D printing), one layer at a time, then precisely machining the material on that layer. Next, a new layer is added on the top of the previous one and machined again; these operations are repeated multiple times to produce a part. Both 3D metal printing and machining are digitally controlled from computer program using common Computer Numerical Control (CNC) programming and digital interface for 3D printing. User uploads the 3D design, the same way as common 3D printers are operated.

Introduction

1.1 Background

At present the majority of 3D printing technology has been developed around termoplastics and even this is still a very new technology. While there are many options in plastics, metal 3D printing, with a few exceptions, has been largely confined to costly pieces of equipment like one used in the jet propulsion industries [1]. There are some major developments in the industrial grade 3D metal printers using laser sintering and post machining technology, to increasing the strength of the material created [2]. Most of them are applied for highly specialized products, have very limited application, and use very expensive machines (from $800k to several millions dollars) and processes (exotic materials, power consumption, maintenance). This would not permit this technology to be brought to the common user to produce functional, ready to use products. There are several attempts to produce affordable 3D metal printers [3], but the quality and finish of the products make them unsuitable for application without additional machining. Some of the special concerns for 3D metal printing over plastics printing is the need to refine the surface finish in order to maintain the level of detail desired and capability to produce complete parts, without open surfaces on the top. The application of metal 3D printing offers more challenges, because of the need for high temperature, so the melted metal bonds together.

1.2 Additive and subtractive manufacturing

At present additive manufacturing [AM] is one of the fastest growing manufacturing systems. The widespread is became the usage is simple and it is often called 3D printing, implying that it is an extension of the 2D printing in one more dimension, but allowing 3D object to be produces. The explosive increase of the 3D printer usage is based on the fact that everybody, without any engineering knowledge can produce their object anytime at any place, and mostly at home. Another factor for widely usage is that the cost of the AM type 3D printer and materials became very affordable. Additive manufacturing is type of so called 3D printing process, when a material is added layer by layer to build 3D structure. We will explain this general process to reveal the difference with the proposed system explained later.

At first the 3D model is created using CAD, 3D modeling software or ready design is acquired [4]. Then the 3D model is converted to STL (Stereolithography) file or other format, commonly used by AM/3D printers. The STL format is converted to slices that have two dimensions outline path for controlling the tool and material delivery motion, based on the AM/3D printer parameters, high of the layer, size of the material, precision required. The tool path is created from each slice, material is deposit on the layer and the tool or table moves vertical one step equal to the distance between slices, depending on the height of the material place on each layer. Special coded program (most common it is type of CNC program, so called G code) drives automatically the table/tool on horizontal plane each layer, then moves one step on vertical direction and drive the tool to deposit material on the previous layer. This process is repeater multiple times till the 3D object is created.

1.3 Proposed 3D Metal Printing/Machining

Traditional AM have many limitations, like materials, mostly plastics, small working volume, and low quality could not even get closer to the industrial created products. One if the biggest problem is the functionally of the object produces. In most of the cases object is produced just to prove the design concept with limited application. There is no question that high volume and quality products still will be made by industrial manufacturing from plastics or metal. So the subtractive process of metal cutting will continue to be one of the most common processes in the future. Looking at the advantages and limitations of AM/3D printing and subtractive manufacturing process we can derive a new method the take the best of both processes without drawback of them. To distinguish our process from traditional AM, and subtractive manufacturing (mostly metal cutting) we call our system 3D metal printing/machining, since is combines both process but is applied exclusive for metals.

To achieve the quality of subtractive manufacturing (machining) and the flexibly and easy usage of AM/3D printing we propose 3D metal printing/machining system that combines the advantages of both systems. While there are industrial grade machines and examples of laser sintering and subsequence machining process available for high end manufacturing, the cost of such machine is so high so make it practically inapplicable for common company and user.

Proposed system is using the flexibility of 3D printing by applying two well-established technology, welding and Computer Numerical Controlled machining to create a fully functional finished metal products. This research is based on our previous successful experience to create 3D printing of special sand molds, by creating layer and machining them [5, 6, 7]. While this is quite successful in creating very complex, near net shape molds, we still need to do secondary machine operations on some surfaces after casting the functional part. This experience led us to this new idea to produce directly functional metal parts that can be used right after the process is finished.

1.4 3D Metal Printing/Machining operations

The proposed system machine operates similar to a 3D printer and 3 axis CNC machine, see Figure 1. There are several main components. The Gas metal arc welding (GMAW), also called metal inert gas (MIG), has retractable welding head attached to the CNC machining head. During the 3D printing part of the process the welding head is moved close to the tabletop fixture. Welding machine delivers the build welding wire and weld a layer on the build platform fixture. To format the shape of a layer CNC machine controls the movement of the table in X, Y direction, while the welder deliver and welds the material on each layer, next welding head retracts, the cutting tool is placed in position and follow the path for cutting one of the sides of the. After layer is machined, the welding head is put down back in welding position, while the high of the is adjust by moving in up one step and the next layer is deliver. Again this welding/machining sequence process is repeated over and over again until the part is completed. A standard CNC program provides the control of the machine for both, welding and machining. Automation of the vertical position of the welding head and the cutting tools are also defined by in the program.

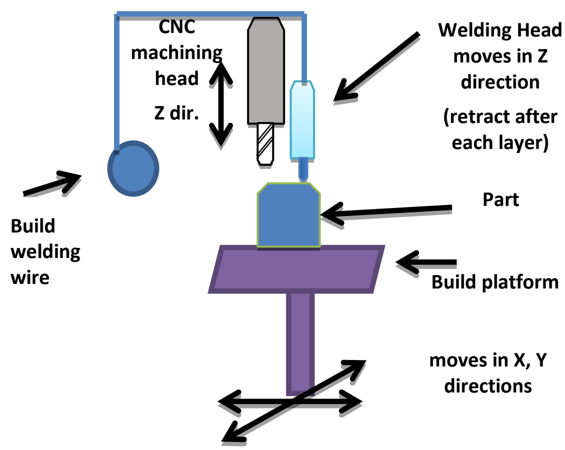


Figure1. 3D Metal printing/machining

2. Methodology

2.1 Feasibility study and testing of the 3D metal printing/machining system concept

This project defines parameters for a 3D metal printer with machining capability as an alternative to the more expensive options on the market currently. This machine would have the capability to deposit and remove metal material from the work piece. The machine would follow the 3D design, created using a 3D solid modeling program and uploaded to it. The machine would need to be able to withstand higher temperatures needed to work with metals, while still maintaining the accuracy desired. A set of specifications would need to be created in order to keep the project on track and set a set of benchmarks that should be accomplished.

We are investigating suitable metal 3D printing technology for metal deposition. We have examined several methods for 3D metal printing to deliver material layer by layer. One of the common way to build material is gas metal arc welding (GMAW), also called metal inert gas (MIG) welding, provides a sufficient method to build material structure but lacks a building precision. At present we have complete prototyping using MIG welding and machining controlled by CNC machine. Although the testing process so far is semi-automated and the prototype machine is not fully completed, we have been investigating the process of weld metal deposition and CNC machining to optimize the input parameter for the final 3D printer/machine system that will function completely automatically.

2.2 Operation process of the 3D metal printing/ machining system

The operation process of the 3D metal printing/machining system is show in the following figures bellow. At the initial stage of the process a CAD model is created and then converted to STL data, see Figure 2 (left) and (middle). Most of the CAD or 3D modeling software have capabilities to convert 3D parametric model to a mesh STL format. The STL model data format is then imported to 3D slicing software to create individual slices, based on desired parameters such as layer height, solid layer, infill, fill pattern, pattern spacing, diameter of the wire, size of the objects, starting point, etc., as show in Figure 2 (right). Some of the slicing software can produce directly CNC G-code, so additionally speeds, feed rates, cooling needs to be defined for a specific machine.

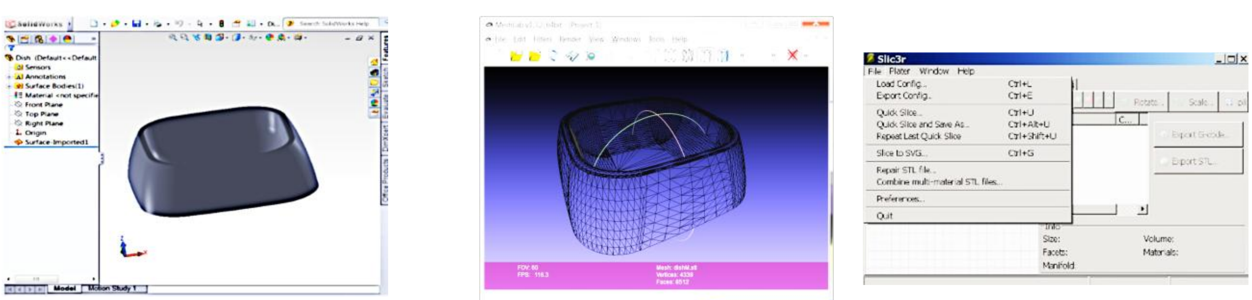


Figure 2. CAD models (left) STL model (middle), Sliced and CCN code generation (right)

The actual 3D printing/machining proces start after the program is loaded in the CNC system. At first the welding head, in down position, deposit a thin layer of material on the fixture plate attached to the machine table, see Figure 3 (left). The movement of the table is controlled by the CNC program while the welder power and wire feed is controlled by the MIG welder. Starting and stopping of the welding process is also controlled by the CNC program. After the layer is finished the welding head retracts and the cutting tool cut the path on the chosen side of the wall guided by the CNC program, see Figure 3 middle and right. The process is repeated multiple times till the product is completed. Depends on the requirements, the system program may also need to control machining the wall on the top of some layer and on the other side. After testing that prove of the concept gives good result, the new 3D printer/machining system has been designed and build.

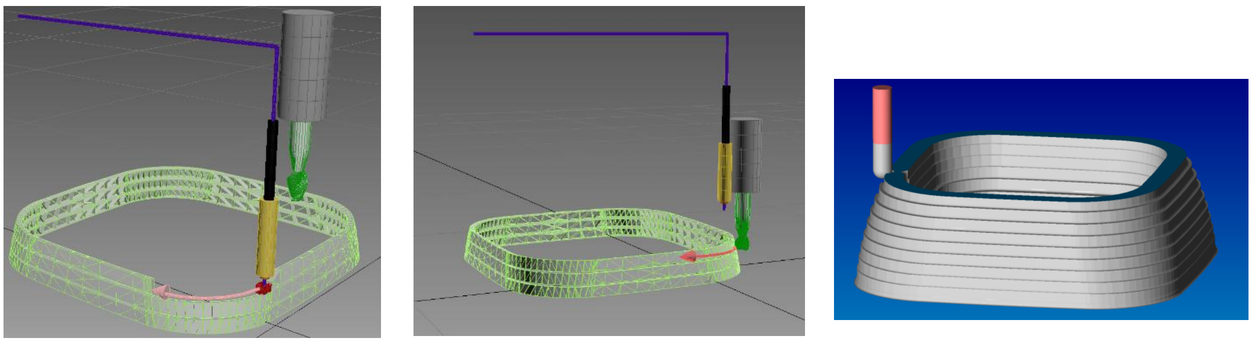


Figure 3. 3D metal printing (left) and machining processes

2.3 Design and build 3D metal printing/ machining system

The goals of this project are to design and build a personal hybrid 3D printer/machining system capable of producing functional high quality parts. The proposed 3D metal printing/machining machine should be something that the average user could afford to build and keep in their house. This machine would have the capability to deposit metal, one layer at the time, then precisely machine metal material on same layer. These operations are repeated many times till the complete part is produced. Both 3D printing and machining are automatically controlled from computer program using common Computer Numerical Control (CNC) programming and digital interface for 3D printing. From user perspective, the machine will create a part from the 3D design uploaded to the computer, the same way as common 3D printers are operated. The machine would need to be able to withstand higher temperatures needed to work with metals, while still maintaining the desired accuracy during the machining part of the process. At present the design and bill of materials is completed, building of the prototype machine has started, and the goal in the following months is to finish and test the actual machine.

Another method that we are been investigated is using 3D printing process with low temperature melting alloys (bellow 450 C^0 degree). Again, the same problem with low precious and quality as welding will persist. In order to make quality surface, we will machine each layer, as soon as it is deposit on its place, using CNC machine head, and repeat the process till the work piece is complete. Since we already have a CNC machine from our previous project [8], this will be relatively easy task.

3 Results and Discussion

To investigate the feasibility of proposed 3D printing/machining system we completed several sample feature build processes with different complexity to test different parameters settings and to optimize speed and quality of 3D printed and machined parts. The mockup setup for the 3D printer /machining system is using an industrial grade GMAW welder (Mid States -Tri Arc 160) with welding head attached to the spindle of the 3 axis CNC milling machine (Milltronics Partner MB 18).

This system operates in similar matter as the proposed system. The welder arm is attached to the CNC spindle with retractable arm, after the welder prints one or more layer it is retracted and the milling tool remove the excess material on one or both side of the build wall. On some of the layers the milling tools also machine the top of the layer to clear it for next deposition or to finish the surface. Since both the welder arm and tool are controlled from the CNC computer, high quality can be achieved matching on of part produced with industrial CNC machine. Next we will explain in detail procedure how this process was realized.

At first we build several simple straight walls to investigate the feasibility of building thin layers on the top of each other and machine them with the setup and procedure described above. As it is shown in the Figure 4 (left), the wall was built with predefined feed speed rate and welding parameters. Although the wall was build successfully, it is clear the higher quality of the surface of the wall was achieved only after the machining, see Figure 4 (right). The second group of experiments with 3D welding was building of several overlapping strips of the base layer, it was done mostly to test the parameters, width and height of each strip, before we make to more complex shape such as box or cylinder.

Figure 4. 3D printed wall before (left) and after (right) machining

Later, walls of a box were successfully 3D printed and machined with high quality, without cracks and gaps in the wall material, as shown in Figure 5, (left). Our goal was to build a hollow object, instead of complete solid that can be easily achieved but will mean a lot of wasted material. Next challenge was covering the pocket gap inside the box. Since the gap between walls was quite wide (1.5”x1.5”) the welding layer cannot overlap enough when placing next to each other. To fill the gap, we placed a tiny piece of sheet metal to provide connectivity for the arc, and welded on the top of the surfaces with two layers to provide structural rigidity. This process allow us to build a hollow box rather than one filled with the metal. We also found out that if the welder head is incline under certain angle, this pocket gab can be covered by adjusting the welding and feed rate parameters, as normally welders in industry are doing. Since we need full control of all parameters, including welding head position and rotation control, this will be tested after completing our own prototype machine.

Finally we tested a process of creating of the small cylinder, see Figure 4 (right), but the temperature increased considerable, since we started building the layers on top of each other without stopping, and we decided that some cooling feature to the welding layer, to avoid overheating, need to be added. We are continuing to test and optimize the process with goal to have completed and actual part to make final adjustments of the building procedure. Successfully completed these tests allow us to finalize the design of the prototype machine and working parameters that can perform 3D metal printing/machining of complex shapes.

Figure 5. 3D metal printing and machining of hollow box

The material was tested for hardness and to see if it was cracked. From the Rockwell hardness test the hardness of the material was found to be Rockwell A scale 48.1. Another test was done to see if there is any crack between the layers, which is a common problem with welding parts. A magnetic particle inspection was done and no cracks were found.

We believe this project will be successful, as it is build based on previous experience gain by our research group, who have successfully build three 3D printer and maintain them, as well as several commercial one working at Computer Integrated Design lab. One of the 3D printer FDM (fused deposition modeling) we built has large work envelope about 3 x 3 x 3 feet, while the latest one can do 3D plastic printing and machining [8]. We also have successfully produced 3D casting parts using light-cured sand for rapid casting technology. The process involves layer curing with high intensity pulsed light source and CNC precious machining to produce high quality functional casting mods and parts. As a prove of the concept we were able not only to produce functional, high quality parts but also to make parts that are typically impossible or very difficult to produce using existing casting technology [5,6,7].

This 3D metal printing/machining method, utilizing precision machining, has two major advantages. First, it allow precise machining at each layer to follow dimensional and tolerance requirements. Second, it proved building of very complex structure with undercuts, overlaps, etc., with a simple 3 axis CNC machining. Third, as an additional benefits, the requirements to machining for power, size, and speed are minimal as we are machining only a thin layer, one layer at the time.

4. Conclusions

We have proposed and successfully complete 3D metal printing/machining for function parts and designed a prototype system. One of the best advantages of the proposed technology its capability to produce high quality finished functional part in one setup. By using 3D printing process and machining, each layer quality is controlled precisely, thus make it possible creating of complex sculpture surfaces with intricate internal and external shapes that cannot be achieved with traditional CNC machines or 3D printers. In addition, it make possible to make multi-compound products and in some case eliminate assembly and finishing operations. Further, the amount of the material used can be reduced significantly as it makes possible creating of hollow part since only the functional loading walls need to me made. Finally, 3D printing and machining process, as derivative of existing 3D printing, will allow users without special knowledge to produce functional parts in his own. As the technology evolves, it can be also used to create features on the top of existing parts, or in some case to rebuild existing part to original functionality. At last but not least, the proposed system is designed and build from affordable materials and devises allowing everyone with interest to build one for himself.

5. References

1. Beth McKenna, 3-D Printing in the Aerospace Industry: How General Electric and United Technologies Are Using This Technology, retrieved on 4/28/2014, http://www.fool.com/investing/general/2014/02/28/3-d-printing-in-the-aerospace-industry-how-general.aspx
2. Kyle Maxey, Mitsubishi to Introduce Metal 3D Printer in North America, retrieved 04/20/2014, http://www.engineering.com/3DPrinting/3DPrintingArticles/ArticleID/6903/Mitsubishi-to-Introduce-Metal-3D-Printer-in-North-America.aspx
3. Gerald c. Anzalone,Chenlong Zhang, Bas Wijnen, Paul G. Sanders, and Joshua M. Pearc, A Low-Cost Open-Source Metal 3-D Printer, IEEE Access, VOLUME 1, 2013
4. Thomas Campbell, Christopher Williams, Olga Ivanova, Banning Garrett, Could 3D Printing Change the World?, retrieved on 4/28/2014, http://www.atlanticcouncil.org/publications/reports/could-3d-printing-change-the-world
5. Hemant Bohra, Sam Ramrattan, and Margaret Joyce, Evaluation of a 3D Light Cured Sand for Rapid Casting Technology 118th Metalcasting Congress, American Foundry Society, April 8-11, 2014 – Schaumburg, IL USA
6. H. Bohra, S. Ramrattan, & M. Joyce, Evaluation of a 3D Light Cured Sand for Rapid Casting Technology, AFS Transactions 14-062, 2014 American Foundry Society
7. H. Bohra, S. N. Ramrattan, P. Ikonomov, P. D. Fleming III, M. K. Joyce, Integration of CAD/CAM in a Hybrid Technique for Rapid Manufacturing Applications, ICAM 2014, Rochester, MI 48309, USA, May 28-30, 2014.
8. Grieve, A., Acuna, M., Klingler, K., Ikonomov, P, and Rodriguez, J., Dual Purpose Desktop Machine – 3D Printer and CNC Milling, Proceedings of the 2014 ASEE North Central Section Conference, Rochester, MI, April 2014..

Dr. Pavel G. Ikonomov is Associate Professor of Industrial and Manufacturing Engineering at Western Michigan University, Kalamazoo, Michigan, USA. His main focus has been 3-D modeling design and VR simulation in manufacturing and assembly, nano-technology, medical application, robotics, and large scale dynamic simulation in various research organizations in Japan such as Hokkaido University (Vis. Researcher) , Tokyo Metropolitan Institute of Technology (Vis. Assoc. Prof), 3D Incorporated and Virtual Reality Center Yokohama (CTO), UCLA (2001-3) and NIST (2002-3)-Vis. Prof. At NIST he was responsible for industrial Virtual Reality Assembly (VADE) and worked with VR simulation for the optical nano-tweezers. Dr. Ikonomov has more than 100 journals and refereed conference proceedings publications, three books and a book chapter.

Application of 3D Printing for Human Bone Replacement

Azem Yahamed[1], Pavel Ikonomov[2], and Paul D. Fleming[3]

[1,3] Department of Chemical and Paper Engineering,
[2]Department of Industrial and Manufacturing Engineering
Western Michigan University, Kalamazoo, MI 49008
Azemkhalifa.yahamed@wmich.edu
pavel.ikonomov@wmich.edu
dan.fleming@wmich.edu

Abstract

This paper aims to test the mechanical properties of thermoplastic materials that can be used for 3D printing of human bone structure substitute. In many cases bone replacement is required. That's why bone substitute is a significant issue in such situations. As a result, the artificial bone substitute materials are widely used in medical applications. The most common method to produce 3D printed items out of thermoplastic materials is Fused Deposition Modeling (FDM). The most used thermoplastic materials are ABS, PLA, PCL, PVA and Nylon.

1. Introduction

1.1. Building bone structure

3D geometry of bones can be acquired using MRI or CT scan of the actual body organs [1]. After that, the 3D modeling software is used to produce the new part or the model of the missing bone structure. The 3D model then is imported to the 3D printing software for building the substitute bone structure [1]. There are many successful tries recently to print 3D items for human bone substitutes using 3D printing technology [2-3]. Therefore, many biocompatible materials have been studied and researched to investigate their strength and the possibility of their use in bone structures by 3D printing.

1.2. Fused Deposition Modeling (FDM)

FDM is a modern method widely used to produce 3D printed items from thermoplastics. It takes the plastic filament and drives it through the extruder heats it up and deposit the molten plastic on the build platform to start building a 3D item layer by layer. The first step is to make 3D model and then convert it to STL format to produce the 3D object. STL format has advantages and disadvantages. The advantage of this format is that it facilitates the geometry of the object by reducing to it is initial components. The disadvantage of this format is that the object loses some of its resolution because it uses just triangles to represent the geometry [4].

Proceedings of the 2014 ICAM, International Conference on Advanced and Agile Manufacturing, Held at Oakland University, Rochester, MI 48309, USA

Once the STL file format is imported to Makerware software to be prepared for 3D printing, it is sliced into many similar thin slices that become layers. These layers describe the two dimensional lanes that the 3D printing process will build, which when stacked upon one another, will create directly a 3D item matching the original design. It is obvious that, the thinner is the layer, and higher is the precision of the lane movement, the higher is the precision that can be carried out for the item [5].

The working mechanism of the FDM technique is shown in figure 1. It takes plastic filament from a coil and drives it through the extruder. Then the plastic is heated and melted by the heat extrusion nozzle and the molten filament flows through the nozzles and is deposited on the building plate to form layers. The heads move on the X-Y axes to form the specific shape of the layer and the platform moves vertically on the Z axes [5].

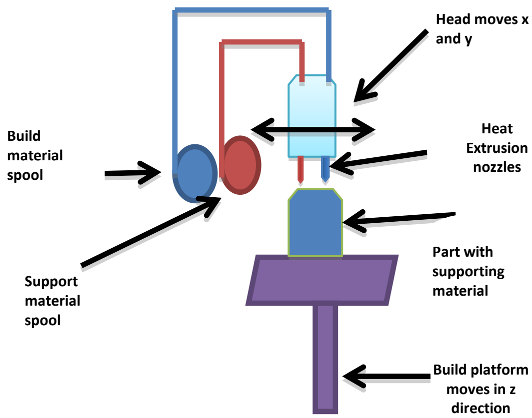


Figure1. Fused Deposition Modeling

3D printing is used to produce fulfilled products because the thermoplastic has become a common material widely used and in some applications, it is the only choice to create parts with specific properties [6].

1.3. Thermoplastics

Thermoplastics have been used successfully as a replacement for certain metals for many years and recently the biomaterials used widely in medical application. 3D printing has a very significant role in improving these materials with high performance, minimum cost and enhancing resistance to environment conditions. The low melting temperature used in 3D printing is considered as the advantage of the technology to create high quality parts for manufacturing and in medical applications also allowing precise replacement of tissue, specifically bone structures [7 -10].

2. Methodology

2.1. 3D printing of test samples

Using 3D printing technology four different samples of thermoplastic materials were printed. ABS (Acrylonitrile-Butadiene-Styrene), PVA (Polyvinyl Alcohol), PLA (Polylactic Acid) and Nylon 618 were selected. The mechanical properties and structure of these thermoplastic materials are different as shown in table 1 below [11-14].

Material	Tensile strength	Elongation	Melting Point
ABS	70 MPa	25%	220°C
PVA	65MPa	3%	185°C
PLA	70 MPa	3.8%	178 °C
Nylo 618	66.0 MPa	300%	218 °C

Table 1. Mechanical properties of thermoplastic Materials

We use 3D slicer or OsiriX software to design 3D models and then we convert it to STL format for 3D printing. Figure 2 below shows the test sample imported by the software in STL format. We can control the sample size and dimensions as we need as well as we can control the other parameters such as temperature, extruder speed, travel federate, plate temperature, infill percentage and resolution. In addition, figure 3 shows the 3D printer running during printing the test sample.

Figure 2. Test sample imported using Maker ware software

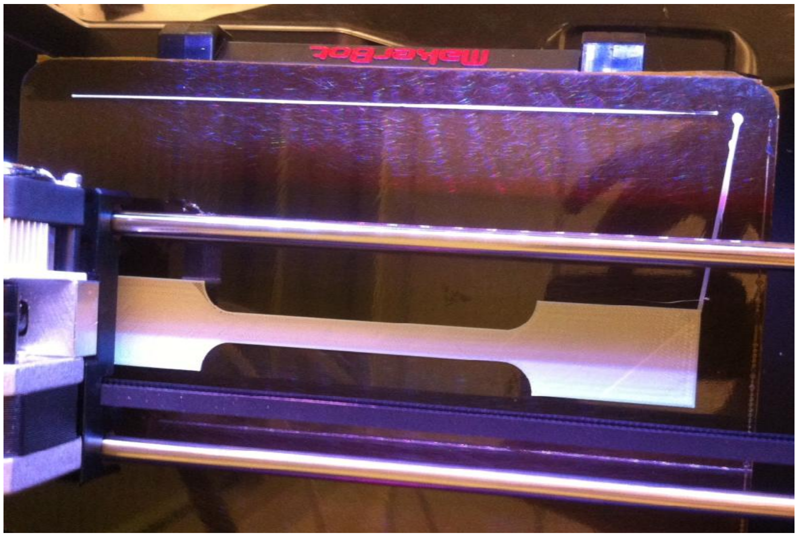


Figure 3. MakerBot during the printing of the tensile test sample

2.2. Creating of 3D bone structure model using OsiriX

OsiriX is free open software used to create 3D models of human organs from CT and MRI scans. Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and ultrasound scans are high quality images used for medical applications including surgeries. To create 3D models for 3D printing bone structure in addition to CT and MRI we can use DICOM images and slice information from actual patients [15].

For creating 3D model there are several steps required to be achieved. The initial step when we get started is the region of interest (ROI) must be selected on the image. After that the segmentation should be performed to separate the borders of the organ. An example of using OsiriX to make 3D model is shown in figure 4.

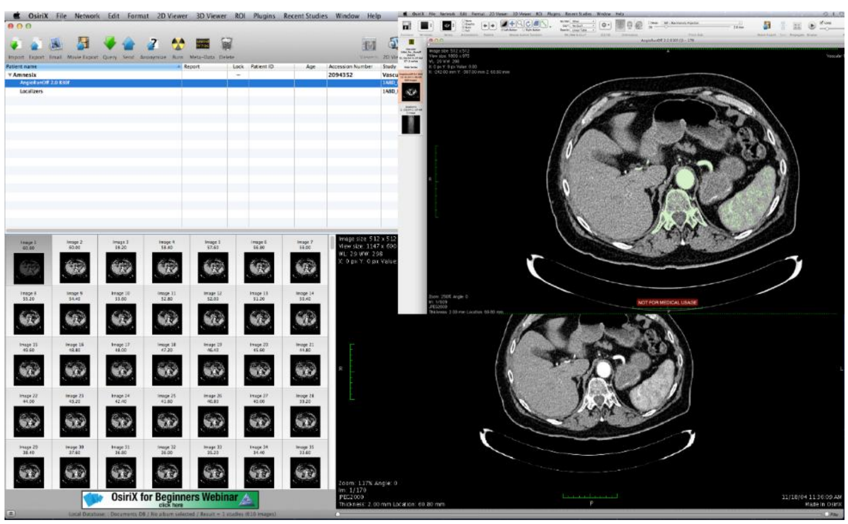


Figure 4. ROI and segmentation in OsiriX

OsiriX aims to view, approximate, read and post processing the images. OsiriX software presents very developed post processing techniques for 2D imaging, database, and 3D models. Figure 4 illustrates collection of images used to describe the ROI (region of interest) and segmentation as illustrated in the upper right corner to create the 3D model. Once the segmentation is finished through all the slides, the volumization is carried out to make the 3D shape. As shown in figure 5, the 3D model is visualized by OsiriX. Then the model is exported to 3D format, which is STL in our case to be printed by 3D printer.

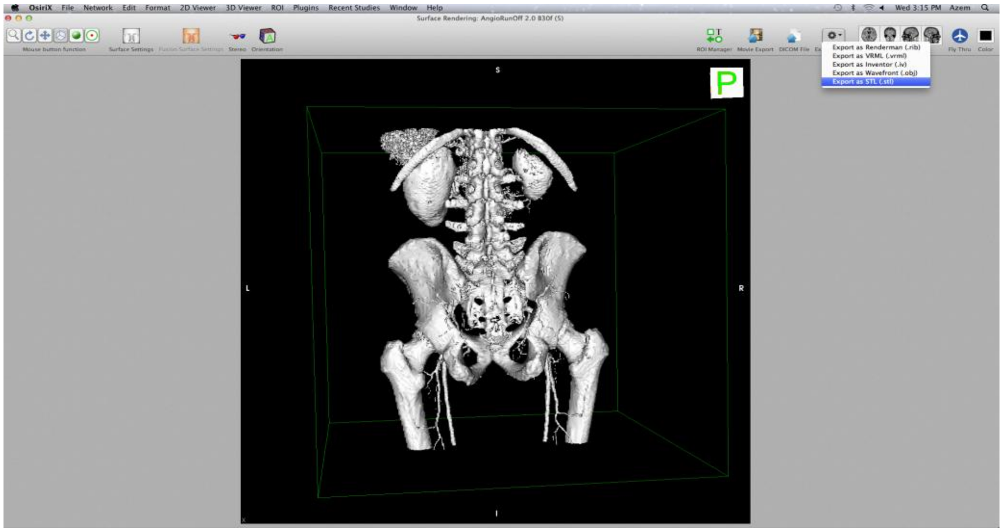


Figure 5. 3D model created after segmentation in Osirix

The mechanical properties of the thermoplastic material can be investigated after printing the test sample. Then we will be able to make a decision if the thermoplastic biomaterials can be used for bone structure. The mechanical properties of the samples are ready to be tested once the samples are printed by using tensile test machine (MTS) at ambient temperature.

2.3. Cleaning 3D model using MeshLab

We use MeshLab software for cleaning a mesh, which means trying to remove all the tiny geometrical irregularities that may be found in shelf meshes. Common problems we usually check, such as duplicated vertices, unreferenced vertices, null faces, self intersecting faces, non manifold faces and small holes. For filling holes, we use the holes filler tool that allows us to select holes and edit them in different ways. The basic filling algorithm uses a technique inserts a face between two adjacent border edges. This algorithm selects every time the best pair of adjacent borders edge into the hole. Then we smooth the model and the smoothing process can be performed by several different criteria as shown in figure 6 below [16].

Figure 6. 3D model of femur bone cleaned and smoothed using Meshlab

We printed bone structure sample to test the accuracy of 3D printer as shown in figure 7 below. OsiriX was used to design 3D models from CT and MRI scans using segmentation techniques to create specific bone structures. 3D models were exported to STL format to be printed by the 3D printer. Designing and building specific fixtures for bone structures mechanical properties being tested and investigated is our goal.

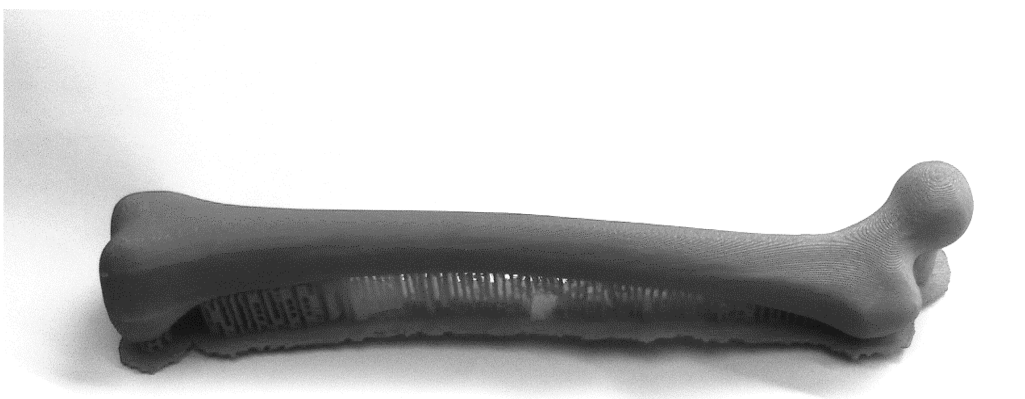


Figure 7. 3D printed bone femur structure

3. Results and Discussion

We use FDM techniques to print four different thermoplastic materials. Two samples were printed from each category and tested. We use an MTS machine to test the tensile strength of the thermoplastics. The software of MTS in figure 8 shows the breaking point of ABS is 727.8 N after 17.58 sec at 0.2 mm/sec and 15% infill.

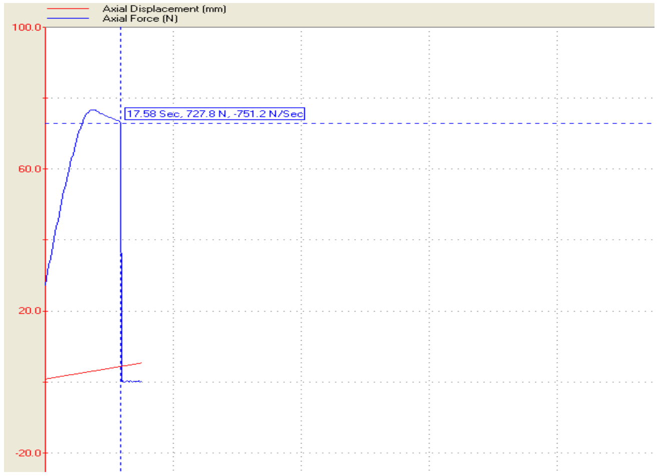


Figure 8. ABS tensile strength test

Consequently, the breaking point of PVA was 563 N after 65.92 sec, PLA was 394 N at 14 sec, Nylon 618 was 239 N after 80.8 sec all the samples were tested with 15% infill at the same speed 0.2 mm/sec and the same room temperature 20 C°.

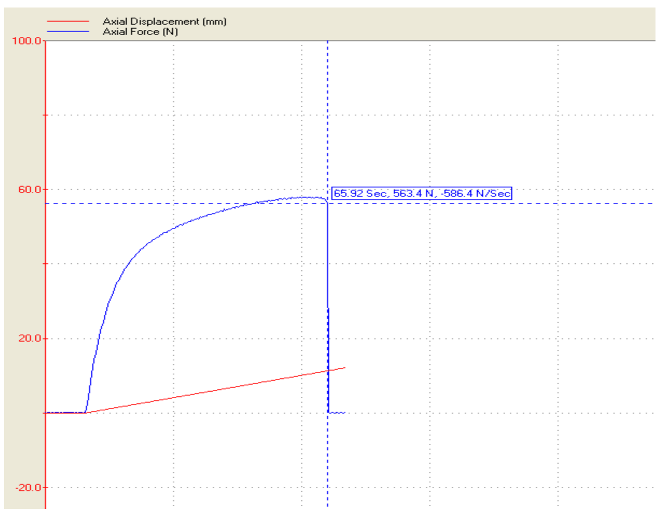


Figure 9. PVA tensile strength

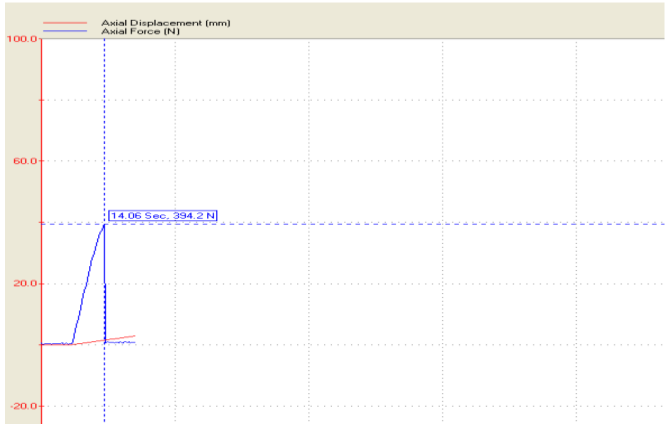


Figure 10. PLA tensile strength

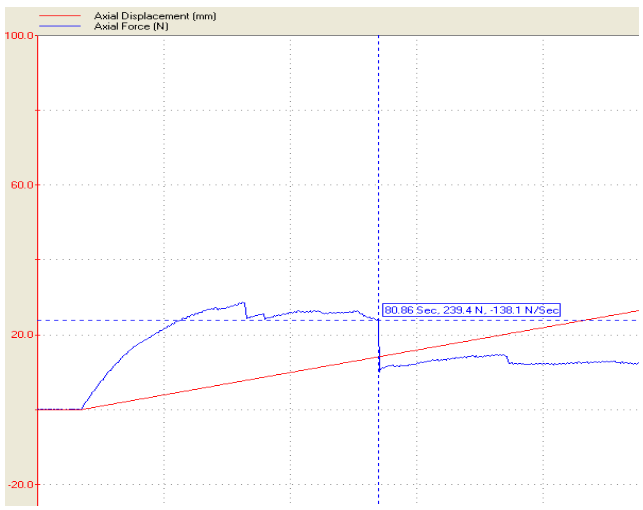


Figure 11. Nylon 618 tensile strength

Table 2 shows the results of the tensile strength after testing with the MTS machine.

Tensile strength (N/mm²)			
ABS	PLA	PVA	Nylon 618
14.6	48.1	9.27	4.8

By making a comparison between the results of the tensile strength of 3D printed samples that we get and the values of the material safety data sheet (MSDS) of the manufacturer, we found that our results are less than the ones we get from the manufacturer. This difference was obviously marked due to the structure filling of the tested thermoplastic samples, which was 15% rather than a 100% solid filling structure. Further study of the actual bone structure substitutes will be tested and investigated including bone strength, porosity and biodegradability.

4. Conclusion

In brief, four different tested samples of thermoplastic materials were printed by using fused deposition modeling (FDM) and investigated by using an MTS machine. The thermoplastic materials were tested ABS, PVA, PLA and Nylon 618. The results show that the tensile strength of these thermoplastics shrink and were less than the tensile strength that we get from manufacturers data sheets, but is still durable enough to be used as a bone replacement.

5. References

[1] Leukers, B., Gulkan, H., Irsen, S.H., Milz, S., Tille, C., Schieker, M., Seitz, H. Hydroxyapatite scaffolds for bone tissue engineering made by 3D printing. *J. Mater. Sci. Mater. Med*, 16(12), 1121-1124, 2005.

[2] Rachel Ehrenberg, Plastic implant replaces three-quarters of man's skull, *Science News*, retrieved on 04/20/2014.https://www.sciencenews.org/article/plastic-implant-replaces-three-quarters-mans-skull

[3] Transplant jaw made by 3D printer claimed as first, *BBC New technology*, retrieved on 04/20/2014. http://www.bbc.com/news/technology-16907104

[4] Ahn, S. H., Montero, M., Odell, D., Roundy, S., & Wright, P. K. (2002). Anisotropic material properties of fused deposition modeling ABS. *Rapid Prototyping Journal*, *8*(4), 248-257.

[5] Hutmacher, D. W., Schantz, T., Zein, I., Ng, K. W., Teoh, S. H., & Tan, K. C. (2001). Mechanical properties and cell cultural response of polycaprolactone scaffolds designed and fabricated via fused deposition modeling. *Journal of biomedical materials research*, *55*(2), 203-216.

[6] Fischer, F. (2011). Thermoplastics: The Best Choice for 3D Printing. *White Paper, Stratasys Inc., Eden Prairie, MN*.

[7] Jia, N., & Kagan, V. (2001). Mechanical Performance of Polyamides with Influence of Moisture and Temperature-Accurate Evaluation and Better Understanding. *Plastics Failure Analysis and Prevention*, *1*, 95.

[8] N. Jia and V. A. Kagan, "Compatibility Analysis of Tensile Properties of Polyamides Using ASTM and ISO Testing Procedures", *SPE Annual Conference Proceedings/Antec'98*, Vol.2, Materials, pp. 1706-1712, (1998).

[9] N. Jia and V. A. Kagan, "Effects of Time and Temperature conditions on the Tensile-Tensile Behavior of Short Fiber Reinforced Polyamides", *SPE Annual Conference Proceedings/Antec'97*, Vol. 2, pp. 1844-1848, (1997).

[10] V. Kagan, "Good Vibrations Join Thermoplastics Fast", *Machine Design*, August 5, 1999.

[11] ABS Material Data Sheet, retrieved 03/10/2014 http://teststandard.com/data_sheets/ABS_Data_sheet.pdf

[12] Taulman 618 Fact Sheet, (Nylon 618) retrieved 04/10/2014, http://3dprinterhub.com/download/Taulman_618_Fact_Sheet.pdf

[13] Futerro Polylactic Acid Oligomers (PLA), retrieved 04/10/2014, http://www.futerro.com/documents/msds_loligomer_en.pdf

[14] DuPont Elvanol 52-22 (PVA), retrieved 04/10/2014, http://www2.dupont.com/Elvanol/en_US/assets/downloads/elvanol_52_22.pdf

[15] OsiriX Imaging software- OsiriX DICOM viewer, retrieved 04/10/2014, http://www.osirix-viewer.com/

[16] MeshLab Documentation, retrieved 04/15/2014 http://www.cyi.ac.cy/system/files/MeshLab%20Documentation1.pdf

Biography

Azem Yahamed is PhD Student in the Department of Chemical and Paper Engineering at Western Michigan University, Kalamazoo, Michigan, USA. I am working on 3D printing for Medical Applications, specifically Bone Structures.

Dr. Pavel G. Ikonomov is Associate Professor of Industrial and Manufacturing Engineering at Western Michigan University, Kalamazoo, Michigan, USA. His main focus has been 3-D modeling design and VR simulation in manufacturing and assembly, nano-technology, medical application, robotics, and large scale dynamic simulation in various research organizations in Japan such as Hokkaido University (Vis. Researcher) , Tokyo Metropolitan Institute of Technology (Vis. Assoc. Prof), 3D Incorporated and Virtual Reality Center Yokohama (CTO), UCLA (2001-3) and NIST (2002-3)-Vis. Prof. At NIST he was responsible for industrial Virtual Reality Assembly (VADE) and worked with VR simulation for the optical nano-tweezers. Dr. Ikonomov has more than 100 journals and refereed conference proceedings publications, three books and a book chapter.

Recognition of Moving Objects Using Sensor System for Robot Teamwork

Pavel Ikonomov[1], Emiliya Milkova[2], Suren N. Dwivedi[3],
[1,2] Industrial Manufacturing and Engineering Department,
Western Michigan Univ., Kalamazoo, MI, USA
[3]University of Louisiana at Lafayette, LA, USA
E-mail: pavel.ikonomov@wmich.edu

Abstract

Robots are used more and more in industries; it inevitable that that they are going to share the same environment in which humans work. The aim of this research is to sense and recognize the human or any other moving object that appears in it the working envelope of the robot and to calculate the precise location and direction of motion of the human or object. An approach has been proposed, to collect the precise information about the position and movement of the human or object, with the help of multiple sensors data fusion. This information about the location of human or object is then used by the Intelligent Robot Controlling System (IRCS), which controls the movement of robot, to recognize new obstacles in path of movement of robot and modify trajectory accordingly. This would make human and robot interaction safe in shared workplace and can be used for cooperate with each other on some manufacturing of assemble process to work in an efficient manner.

Keywords: robot, human-robot interaction, coordination, path adjustment, space motion, virtual reality

1. Introduction

People are working in shared human-robot spaces. According to the World Robotics Organization [1], approximately industrial 174,600 robots were in operation throughout North America in 2010, after significant increase of 33 percent, with annual rate of 9 percent increase for the following three years and beyond. There is exponentially increase of the number of robots in industry, households, children's toys, and space and sea exploration is increasing. Worldwide industrial robots in operation are 1,027,000. In addition, one of the fastest growing groups of robots are service robots: domestic use - 5.6 millions and entertainment and leisure use - 3.1 millions for 2009.

At present the existing industrial practice is to keep humans and robots completely separate. Each robot is placed inside a metal cage to keeps people away and safe. Safety can also be provided by a so-called three-level safety system that includes a light curtain, pressure mats, and ultrasonic sensors. The existing systems provide safe workplace by eliminating any close contacts between humans and robots is based on technological assumptions from the last century. Most robots used in daily life already use the newest technologies (sensors, computers and software) to provide safe work-at-home and services sector. Therefore the application of robots in industry needs to be modernized to be able to perform in human presence and to collaborate with them.

Most of robots used today in industry do not have much intelligence in sensing and computing capabilities, but new robots shall be able to react and adjust to any changes in their environment in real-time; thus working together with humans is shared space should be safe and possible.

At present robots are not only replacing human in attaining difficult tasks, but also they can pose new risks to workers. For example, while a robot may appear to be idle a remote signal may cause sudden motion, which may lead it to change its path and respond to a change in conditions [2]. Existing robot with their control systems can harm the human or objects which are in the working space of the robot. So for the robots to be able to work effectively in a human environment they should have clear perception of the environment in which it operates. They should extract enough information about the environment, so that it can interact with human that are present in the surroundings of the robot. We have proposed an approach, with the help of sensors data fusion in collecting the information of the human present in an environment. The main aim is to detect, recognize the human present in the sensing range of the robot.

1.1 Application of Sensors in Robots

To make robots to work effectively and to perform assigned tasks in human environment, there is a need for comprehensive information about surroundings that can be supplied by multiple sensors. The environments in which sensors operate can produce broad range information such as temperature, humidity, light levels, and change in position. Therefore sensing any changes in the environment the robot controller shall be able to modify the motion path.

For the robot completing an assigned task is important, while at the same time it should ensure the human safety. Sensors are being developed to meet this need by making robot systems adaptive and intelligent. This can be attained by obtaining information about the robot's workplace and the object present in the environment.

1.2 Using Virtual Reality for robot simulation and control

The rationale for choosing virtual reality (VR) is its unique features and flexibility. A 3-D virtual environment [3] includes a map of the environment such as the walls etc., the location and orientation of the robots, which are marked with panoramic images, and the virtual location of the other users present in the system. The user is not only able to explore the environment, but also to control the robots to ensure that a collision does not occur and to avoid obstacles. It gives us real life experience to interact with the life like model or environment in safety and convenient times. We can build any environment virtually and control the objects as we like. It provides us control over the simulation that is usually not possible in real-life situation [4].

Virtual Environment (VE) is the VR space where the simulation is done. Within VE, we can visualize, simulate, and analyze the human-robot interaction from each side respectively. On the other hand, if we perform the same interaction in real situation we can face limitations. For example, the robot cannot resume work after an improper operation that can also harm the human or robot itself. If the operation is not conforming to the accepted standards we have to test it again. As a result, the real simulation would be very expensive and sensors could not use be tested in advance. However in VR we can easily change the location, orientation, and program path of the virtual robot and the placement of virtual work tools and manufactured products. The added advantage is that we can make desired modifications immediately and easily and test the simulation number of times.

In general, experimenting with a working robot and human is too dangerous. In addition testing the simulation practically with real sensors is very expensive. On the other hand, industrial standards do not allow for human-robot interaction because it can be risky for the human. Safeguarding devices and practices are used to minimize any hazards associated with robot experimentation but cannot be completely eliminated. In this research we have used VR for the detection and recognition of the human or object and its position from the robot with the help of sensors. We believe that VR simulation is the best choice for testing human-robot interaction using sensors and in the future could be used for some functions in real environment. Once the simulation is tested within VR simulation with affirmative results the actual implementation system using real robot and sensors can be build.

2. Methodology

This study proposes an approach to detect and recognize any human or object present in the vicinity of the robot. It gives a concept of how to explore the unknown environment in the surroundings of the robot. For this approach, we have used VR as a tool, using group of sensors to do the simulation. It allows any robot to work in the human environment without any risk involved to human or robot. The required data is collected in Virtual environment using selected group of sensors. These sensors are mounted on robot to collect the required data set. Each data set is a collection of different types of information such as position, orientation, shape, size etc.

In this research, sensors are mounted on the top of the robot. The set of sensors used are proximity sensors, video camera, and laser sensor (or ladder sensor). The data collected is updated continuously by providing the latest information about the surroundings of the robot.

The robot currently used in this research is Motorman UPJ. It has six axes: sweep, lower arm, upper arm, rotate, bend, and twist. Each joint has its respective axis namely S-axis, L-axis, U-axis, R-axis, B-axis, T-axis. The robots' parts were measured, then modeled in Pro/Engineer and Catia, and finally imported to Virtual Reality software Eon Reality. The individual robot parts are assembled with respect to their axis in Eon Reality and appropriate degree of freedom as placed on each axis. Behavior model, control algorithm, including input/output from sensors signals, and intelligent logics for interaction with human were developed and run within the VR software.

2.1 Virtual Reality simulation, algorithm, and function control system

Virtual Reality simulation and control system developed is shown in Figure 1. Sensors are activated whenever the human or moving object is detected within the range of the proximity sensor of the robot. It main objective is to locate and measure the distance to the human or object.

When the proximity sensor detected new object it sends the measured data to the Intelligent Optimized Control System (IOCS) that activates a video camera.

The video camera records the human or objects approaching and captures motion frames. The data captured is processed by MATLAB algorithm that calculates the two dimensional (2D) silhouette (outline and shape) of the captured human or object. The silhouette points of the human or objects are converted to three dimensional (3-D) space points and send the IOCS. As a result a laser sensor is activated that scans only the area defined by the silhouette algorithm of the human or object, present in the sensing area, and determines the exact position and the three dimensional coordinates (x, y, z). This collected data points from the laser sensor is send to the IOCS.

Alternatively to the laser sensor, a lader sensor can be activated (it is designed as a group of laser beams). It scans the area of the human or objects that is in sensing range of the robot and gives the volume of the data points of the human or object. The main advantage of the ladder sensor over laser sensor is it scans larger area with high speed and accuracy. It also gives more data points of the human or object than can be directly converted to volumetric information.

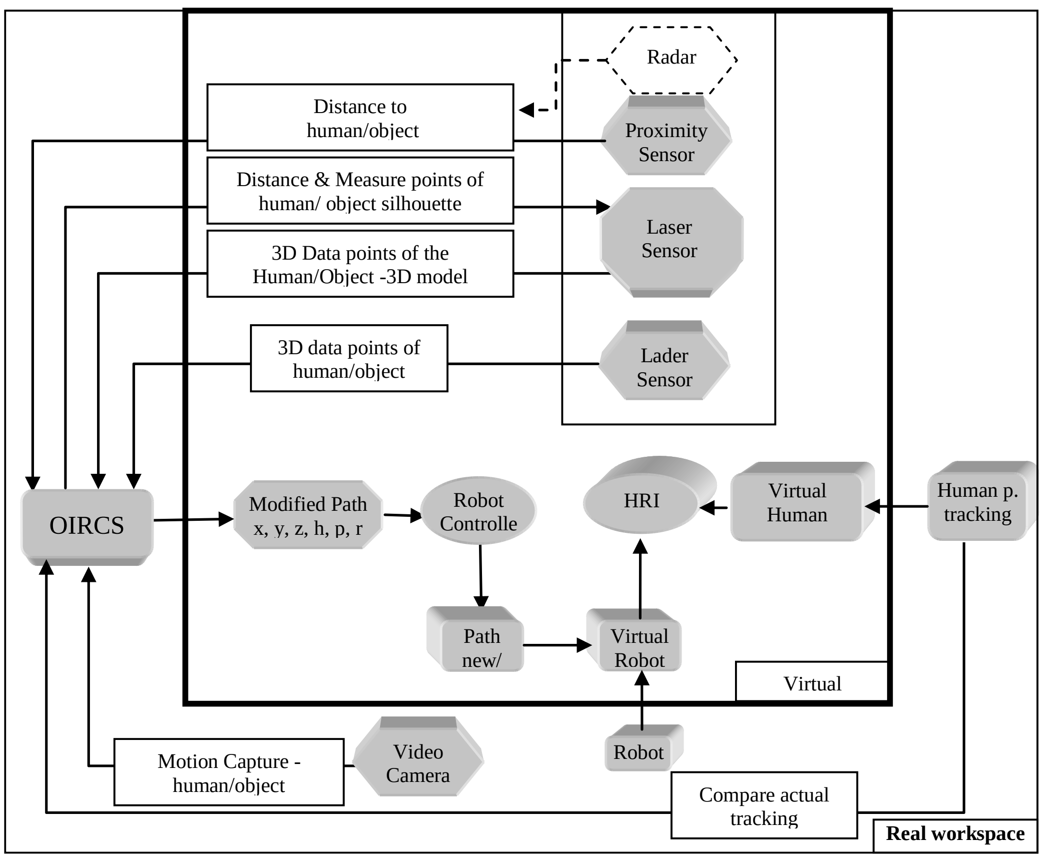


Figure 1. Human robot interaction and control system

The IOCS analyzes the collected data from proximity sensor, camera and laser sensor (or ladder sensor) and determines exact position and orientation of the human or object and its 3-D measurements. It also can define if the approaching human or object is near to the robot working envelope then it can modify the robot arm motion and avoids any further collision from the human or object. Intersection of calculated 3-D space motion envelopes of robot and approaching human/object is used to calculate new the modified path that is sent to robot controller. Further precise calculation of those 3-D envelopes would allow safe collaboration between humans and robots, when a human can give/ take a part to/from robot.

2.2 Distance measurement with proximity sensor

The proximity sensors are mounted on the robot with a sensing area that is in the form of semi-sphere. We will explain the concept of the sensors operations with an example, shown in Figure 2. The sensing range can be divided into three (or more) areas namely proximity area 1, proximity area 2, and proximity area 3. The main objective of the proximity sensor is to locate an approaching human/object and detect the distance to it. To cover completely the surrounding area several proximity sensors are used. The sensing range detects any obstacles within the predefined range from the center of the robot.

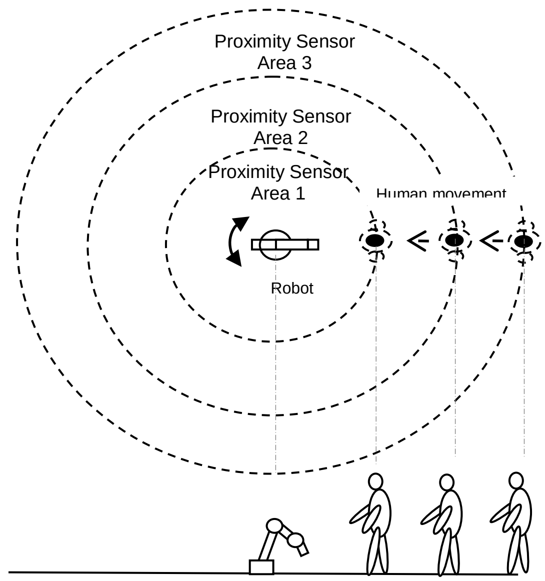


Figure 2. Proximity Sensor measurement of the human approaching the robot

Figure 2 shows the function of the proximity sensor, where the human approaching the robot. It gives the distance of the object from the robot, whenever any obstacle is detected in the proximity range. Proximity sensor gives the signal of presence of a human or object is in the sensing range and the distance to it. This signal is sent to the IOCS, indicating that the human is in the sensing range; the distance is measured and is collected in the data collection file.

Whenever a human or object enters the vicinity of the proximity sensor, it gives the updated distance of the object with respect to the time. The updated distances detected are saved in data collection file through IOCS to be used for calculation and control of other sensors.

2.3 Video camera silhouette algorithm

The video camera used in this research work is a CCD camera. These types of cameras are inexpensive and widely available. The camera is mounted at the base of the robot. Whenever any human or object enters in the proximity area of the robot, measured by the proximity sensors, a signal is sent to activate the video camera and records the information of any objects in the vicinity of the robot. We used a background segmentation algorithm [5], with MATLAB to obtain the silhouette. The main advantage of this algorithm is that it captures only the moving objects. It does not capture any static objects like walls, doors etc. This algorithm captures the approaching human or objects frame by frame and processes the movie in number of frames. The result is a capture of human or object image and that is represented in 2D data points. The algorithm essentially calculates the silhouette of the human or objects and detects the corner points of the image and gives it in pixel values. This algorithm builds the silhouettes by taking the pixel color from every fourth frame and clears the data above and below a pair of thresholds. From the remaining frames it estimates the mean and variance of each pixels color assuming to be a normal distribution.

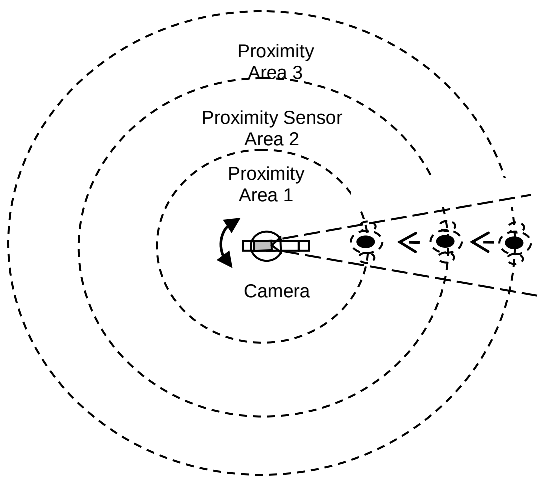


Figure 3. Video camera captures the motion of human approaching the robot

In this research, we have used a video of human approaching the robot. An example of how the algorithm processes the video of approaching human is shown in Figure 3. Here we have taken three different positions where the human is at three different distances. These three different distances coinciding with a distance measured by the proximity sensor areas explained above. By synchronizing the time, we know the distance (from proximity sensor) of the human from the robot and obtain the silhouette data with respect to that frame. The information then is processed by IOCS and used to control the laser sensor movement.

2.4 Measurements with laser sensor

Laser sensor is a device which emits the light beam. The laser beam to be used in real environment is of type not harmful to the human. An example of the simulated laser that scans the surface of the object and collects the data points is shown in Figure 4. The laser scanning points controlled IOCS are calculate directly from video camera silhouette algorithm.

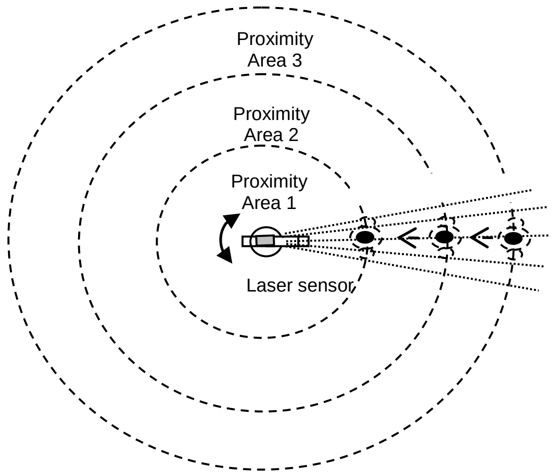


Figure 4. Laser sensor measurement of human approaching the robot

When the laser sensor is activated, it scans only the human or object silhouette (calculated from video foreground algorithm) that is in the vicinity of the proximity range. It measures precisely the distance to the human or object, from it x, y, z coordinates data points are calculated. The laser sensor collected data points set with respect to the time. Further 3-D surface models of the human or object is calculated as show later in this paper.

2.5 Measurements with Primesence sensors

Instead of exiting sensors that are very expensive and computer intensive we are investigating new technologies that became available recently. This technologies can be applied for measurements of the movement of robot and human in real time.

One of the promising technology we are testing now is using Kinetic for Windows that employ Primesense sensors [6, 7].

Real time motion depth data measurement is accomplished by as system with IR emitter and camera. The IR emitter radiates infrared light beams then the depth sensor reads the IR beams reflected back to the sensor. The reflected beams are measuring the distance between an object and the sensor and convert it a depth image (3D) In addition this system has RGB camera with high resolution 1280x960 resolution that maps the real image on the 3D measured data sensors [8]

The fast measurements and processing algorithms working with several moving humans/objects make it possible to be applied to propped human/robot interaction system. Since the processing of the 3D and picture image is in real time this make it possible to be implemented in control system. This technology was successfully tested to measure and calculate the movement of human that control robot in real time [9]. We are investigating the capabilities to be applied for high precision detail measurements to replace the array of sensors proposed, proximity, laser, radar, with one single combined sensor system.

3. Results and discussions

Sensors in this application are mounted on the arms of the robot or inside its working space. The robot control and sensor data will be collected for human/objects approaching the robot. The data from each type of sensors is used as input to the next sensor while at the same time accuracy increases while at the same time the built-in redundancy in measurement with each sensor allows safe operation.

The proposed OIRCS software/hardware system is built based on the following concept. A robot in the environment has the data for positions, sensory direction data, collision detection, or the point of collision information within the environment. Input data for the system of initial and target points for the robot are used by the OIRCS for its feedback system to control the robot and avoid collision (by calculating the intersection motion space envelopes) with the external source. A new position is self-determined by the intelligent system to make sure of any new encounters with an external object. When the system detects the approach or collision (close proximity) of humans or objects, it modifies the motion of the robot for a collision-free path until the robot reaches its goal point. From human movements, measured by sensor, 3-D space motion envelopes can be constructed, and from their intersections with the robot motion envelope a safe solution can be easily found.

3.1 Distance measurement with proximity sensors

The first group of sensors, several short range proximity sensors, is attached to the robot arms to provide safety measures with high fidelity and accuracy. Those sensors similar to the existing industrial applications allow safety work. Sensors measure the range to the human body/hand or object at any stage of the motion. If no external objects are sensed within a predefined minimum (safe) distance, the robot works normally. If a human body/hand or object is sensed, then the intelligent control system modifies the trajectory, leading to a collision-free path for the robot.

The second group of long range proximity sensors is used to measure distance to the human/object to be used by the next group of sensors. It is calibrated to detect motion toward the robot. When a human or object is on the path approaching the working robot, the video camera starts automatically at predefined distance measured by the long range sensor.

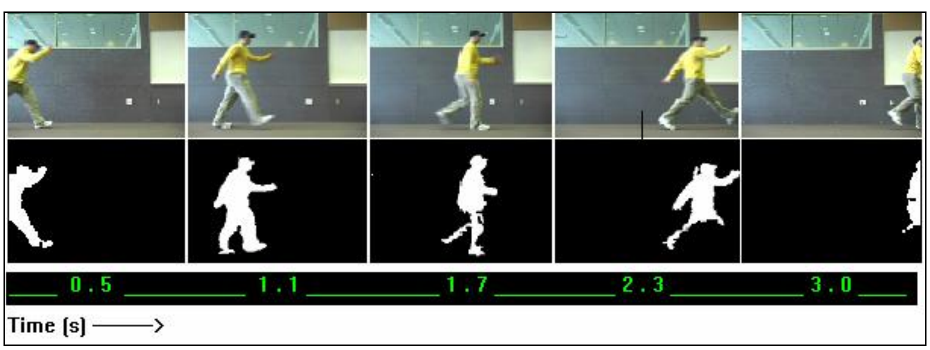
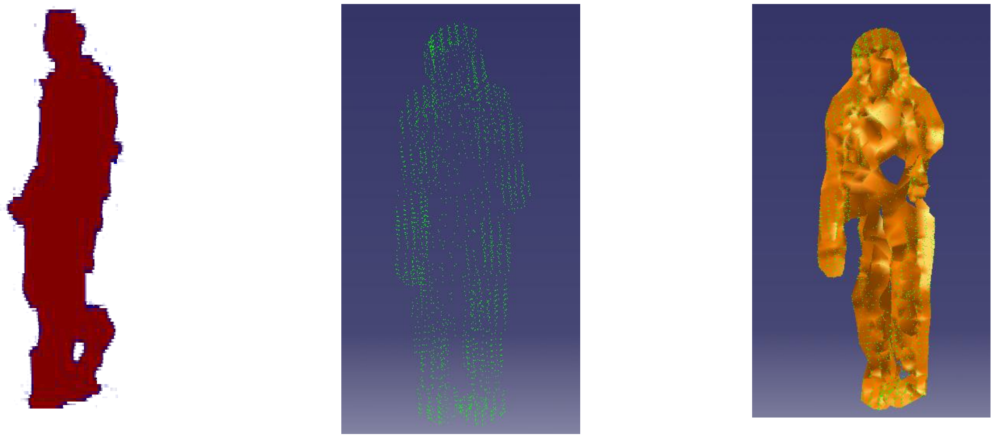


Figure 5. Obtaining the Silhouette from video camera of human approaching the robot

The camera captures the human in the first frame, where the human is at specified distance (e.g. 5m), and it gives the silhouette of the human and the pixel values of the human silhouette. The pixel values will be in the form of P (y, x). Figure 5 shows the silhouette of the human at several different frames at different distances with respect to the time.



(a) (b) (c)

Figure 6. (a) Silhouette from video camera (b) 3-D cloud points cloud (c) 3-D model

3.2 Human silhouette reconstructed from video capture

All the information of the human silhouette along with the pixel values P(y,x) and the distance with respect to the time is continuously updated sent to the IOCS. By analyzing the different frames of the human silhouette and the distances the IOCS have the updated information of the human or object approaching the robot. It was tested and verified that the algorithm is independent on color changes of moving human objects or static background and obtains the silhouette only of the moving human or objects.

IOCS system analyzes and controls the robot-sensor interaction; it takes the information of set of human silhouettes and their respective distances from the camera and changes the 2D image plane coordinates (y, x) into world coordinates of the object plane. Subsequently, conversion of the image coordinates to object coordinates x, y, z points in real world are done. Then the laser sensor is activated and scans the human or object which is approaching the robot.

3.3. 3-D data points obtained from laser/lader/ Primesence sensors

Depending on the size of the object derived from the silhouette (Figure 6 (a)) and distance, laser sensor scans and collects the data points. IOCS systems control the laser sensor scans to move on defined pattern, with distinct single step in one direction, motion speed and number of turns steps in other direction required to measure the approaching human or object. The laser beam scans and send measured distance and data points back to OICS. From collected data points, Figure 6 (b), the surface, Figure 6 (c), of the human or object is calculated. By plotting the data points, a point cloud is formed in the surface shape of the measured human. Point cloud is the collection of the three dimensional (3-D) coordinates of the data points. With this point cloud a 3-D model of the measured surface is constructed. Figure 6 shows the intermediate stages of obtaining of completed 3-D models of the human. With this comprehensive information of the sensors, and 3-D models robot can react to the human or object motion and can modify its path to avoid collision and behavior accordingly.

4. Conclusion

This study proposes an approach for a robot to detect and recognize the human or object in the unknown environment with the help of fusion of different sensors data. With the comprehensive information obtained from the sensors, robot can react to the human or object motion and can modify its path to avoid collision and behavior accordingly. The propose method is designed to work with industrial or other robots that work in unknown and dynamic changing environment. The modifications can be made using already available off-the-shelf, cost effective, components and Artificial Intelligent systems coupled with advanced Virtual Reality, which can be adapted to safely increase the workspace human/robot automation environment.

The system provides the robot/automated system a way to predefine a possible unsafe human/machine interaction by predicting in real time such things as human/object motion direction and position, the probability of dangerous conditions, and the execution of a safe alternative path. Early detection gives the OIRCS system timely modification of a collision-free path adjustment. This will allow development of a system that is safe, dependable, and easy to use. Robots could perform autonomous tasks and be cable of collaboration with humans. At present applications of actual sensors and their integration with robot control has been studied. New sensors, such as Kinect from Microsoft X-box can be used for real time calculation of motion envelops of both, the robot and the human [6, 9]. Future work will continue on implementation of the actual sensors with existing robots and building of OIRCS capable to work in real time.

5. References

[1] World Robotics 2010 - Executive Summary www.worldrobotics.org (retrieved on 5/25/2011)

[2] Jianming Xu, An implementation of multiple sensors for robot safety Master Thesis, University of Louisville, Louisville,1996,

[3]Fagg, A. H., Ou, S., Hedge, C.s, T. R., Brewer, M., Piantedosi M., Amstutz P., Hanson, A., Zhu, Z., Grupen, R., and Riseman, E. (2002), Human-Robot Interaction Through a Distributed Virtual Environment, Proceedings of the Workshop on Intelligent Virtual Environments and Human Augmentation (WIHAVE), Chapel Hill, NC, October 17-19.

[4] Odashima, T., Onishi, M., Luo, Z. and Hosoe, S., 2003, Immersion Type Virtual Environment for Human-robot Interaction, Proc. of IEEE International Conference on Systems, Man & Cybernetics (SMCC03), Washington, DC, pp. 3651-3656.

[5] Howe, N., Silhouette Lookup for Automatic Pose Tracking, IEEE Workshop on Articulated and Nonrigid Motion, June 2004. http://maven.smith.edu/~nhowe/research/code/

[6] Kinect Hacks from MIT: Interview with Philipp Robbel, Boston Innovation, retrieved on 05/25/2011, http://bostinnovation.com ()

[7] Norman Villaroman, Dale Rowe, Bret Swan, eaching Natural User Interaction Using OpenNI and the Microsoft Kinect Sensor, retrieved on 04/20/2014 from: http://sigite2011.sigite.org/wp-content/uploads/2011/10/session14-paper02.pdf

[8] Kinect for Windows Sensor Components and Specifications, retrieved 04/20/2014 from: msdn.microsoft.com/en-us/library/jj131033.aspx

[9] NASA Uses Kinect and Oculus Rift to Control a Robotic Arm, retrieved on 04/20/2014, www.ign.com/articles/2013/12/31/nasa-uses-kinect-and-oculus-rift-to-control-a-robotic-arm

Biography:

Dr. Pavel G. Ikonomov is Associate Professor of Industrial and Manufacturing Engineering at Western Michigan University, Kalamazoo, Michigan, USA. He has been working on Virtual Reality simulation for more than 17 years. His main focus has been 3-D modeling design and VR simulation in manufacturing and assembly, nano-technology, medical application, robotics, and large scale dynamic simulation in various research organizations in Japan such as Hokkaido University (Vis. Researcher) , Tokyo Metropolitan Institute of Technology (Vis. Assoc. Prof), 3D Incorporated and Virtual Reality Center Yokohama (CTO), UCLA (2001-3) and NIST (2002-3)-Vis. Prof. At NIST he was responsible for industrial Virtual Reality Assembly (VADE) and worked with VR simulation for the optical nano-tweezers. Dr. Ikonomov has more than 100 journals and refereed conference proceedings publications, three books and a book chapter.

Analyze and Determine the Forces Associated with the Nanoparticle Movement

Pavel Ikonomov[1] and Muralidhar Ghantasala[2]

[1] Industrial Manufacturing and Engineering Department,
[2]Mechanical and Aeronautical Engineering Department
Western Michigan University
Kalamazoo, Michigan 49008
e-mail:pavel.ikonomov@wmich.edu

Abstract

The primary research objective of this paper is to analyze and determine the forces associated with the nanoparticle movement for use in *in-vivo* drug delivery by considering the fundamental physical and bio-mechanical principles. It requires a thorough study and understanding of the complex inter-relationships of the forces created from interactions among different nanoparticles, cells, and the biological environment. Lab-on-chip for nanoparticle drug delivery are the two innovative concepts that have evolved from recent investigations in the area of micro/nanosciences and technology. Lab-on-chip allow *in-vitro* research to be tested before, *in-vivo* techniques are applied. These studies will lay the foundation for future use in drug targeting to the diseased site or organ using externally applied forces.

1. Background

Lab-on-chip and nanoparticle drug delivery are the two innovative concepts that have evolved from recent investigations in the area of micro/nano sciences and technology. While lab-on-chip is mainly utilized for *in-vitro* research, the nanoparticle applications of biotechnology more relate to *in-vivo* techniques. These include the area of experimental anti-cancer therapeutics wherein nanotechnology has the potential to rein in systemic toxicity of chemotherapeutic drugs by encapsulating the cytotoxic agents [1]. Further, nanoparticle based therapeutics, in case of brain diseases, is reported to cross the blood brain barrier and carries out multiple tasks in a predefined sequence [2]. This is one of the most promising applications of nanotechnology in clinical neuroscience. Therefore, 'nano materials' have created a whole new area now being termed "nano-biotechnology" and this has brought a major paradigm shift in applied biology.

Targeting of drugs encapsulated in nanoparticles to a diseased organ or site will lead to a decrease in cytotoxicity and an increase in efficacy by delivering the drug payload directly to the area where it is needed the most. Although, nanoparticles are widely investigated as the drug carriers, there are only limited means to deliver them to a specific site within a body. Therefore, novel alternative modes of targeting nano-drug-vehicles is needed. One such approach, which we are investigate is the guiding of magnetic nanoparticles to a specific region with an external magnetic field. Without the knowledge of the forces experienced by nanoparticles, it will be

Proceedings of the 2014 ICAM, International Conference on Advanced and Agile Manufacturing, Held at Oakland University, Rochester, MI 48309, USA

difficult to predict the field requirements to navigate the particles. Further, magnetic field gradients at a branching point can create difficulties in determining the path of the particles along a particular branch [4, 5]. Hence, a thorough understanding of the forces experienced by these nanoparticles at various stages is critical to develop the ability to steer them in a chosen path. *In-vivo* experimental studies on these aspects are complex and quite expensive. Simulation methods can provide basic understanding of various forces involved in the interaction of nanoparticles with regions of body during their transport. This research aims to provide control and visual representation for nanoparticles steering using virtual reality (VR) tool with dynamically changing input parameters. Concerted collaborative efforts by the experts having research experience encompassing disciplines like physical sciences, engineering and biology is needed for such analysis.

2. Methodology

The main research objectives of are (i) Determine the forces associated with *in-vivo* nanoparticle transport, by considering the nanoscale physical and biomechanical parameters (ii) Formulate analytical and numerical models using the identified forces to predict the nanoparticle motion within the body (iii)Steer the nanoparticles and provide a real time view of the nanoparticle tracks using Virtual Reality (VR) simulation methods and (iv) Validate the steering and retention of nanoparticles at a target site using an *in-vitro* microfluidic system.

2.1 Analyses of forces that influence nanoparticle transport in vivo

The proposed research will provide a better understanding of the nanoparticle transport mechanism, their possible interaction with the bio-environment and physiological to biophysical/mechanical forces experienced by nanoparticles. A validated model that can prognosticate the forces experienced by nano drug delivery platforms in vivo along with virtual reality (VR) simulation has the potential to reduce the number of experiments involving either human subjects and/or experimental animals. Currently, active drug delivery is mainly achieved using systemic injection of nanoparticle vehicles through the complex biological system and specific retention at the site of action [3, 5]. Mechanically, these nanoparticles diffusing towards the target site experience a number of intrinsic and extrinsic forces, depending on the environment and application.

Modeling of the blood circulation in arteries, veins, and in other parts of the body is attempted by a number of laboratories [6-8]. The transport of blood in the arterial media is described by the convection-diffusion equation

$$\frac{\partial C}{\partial t} + V_r \frac{\partial C}{\partial r} + \frac{V_\theta}{r}\frac{\partial C}{\partial \theta} + V_z \frac{\partial C}{\partial z} = D_r\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial C}{\partial r}\right)\right] + D_\theta\left(\frac{1}{r^2}\frac{\partial^2 C}{\partial \theta^2}\right) + D_z\left(r\frac{\partial^2 C}{\partial z^2}\right) \quad (1)$$

Where, *C* denotes local tissue drug concentration. Both Diffusion *(D_r, D_θ, D_z)* and convection (V_r, V_θ, V_z) are directionally dependent.

This is one of the simplistic models of the blood circulation based on physiological forces. However, there are many other factors that affect nanoparticle transport in drug delivery such as the particle size relative to the blood vessel, pressure drop due to the collisions of these particles to the walls, variation in buoyancy forces based on the density, frictional, adhesive and different contact and non-contact forces. Therefore, it is critically important to assess these factors for calculating the net driving force of the particles that defines its direction and velocity. In addition to these intrinsic parameters for magnetic nanoparticle, the influence of externally applied magnetic forces needs to be evaluated.

Drug targeting using magnetic field, has been previously attempted to increase the concentration of the drug at a defined target site and away from the reticular endothelial system [9]. Normally, nanoparticles are injected through an artery in the presence of an external magnetic field of sufficient strength and gradient to retain the carrier at the targeted site [10]. Generally, magnetic field gradient is used to steer the particles towards an identified target [11]. In a preliminary experiment, it was shown that the nanoparticle can be driven few centimeters in the presence of sufficiently strong enough fields, even against the diffusion and blood stream velocities [12]. Thus there exist enough preliminary evidence that magnetic field can be used steer and retain iron oxide harboring multifunctional nanoparticles. However, a thorough and systematic study to evaluate forces needed and their implications in cancer therapy outcome is lacking. This may be due to the lack of understanding of the methods and their theoretical underpinnings with respect to engineering the nanoparticles with specific characteristics and steering/retaining these at the required site using localized field configurations and applied fields.

Magnetic nanoparticles are preferred in drug delivery systems because of their biocompatibility and non-contact drivability characteristics. Magnetite (Fe_3O_4), magnemite (Fe_2O_3), and other ferrite particles are commonly synthesized in different sizes ranging from few nanometers to several tens of nanometers [13]. Material, particle size and surfactant play an important role in determining their magnetic properties [14]. They behave as free paramagnetic particles in a solution without attracting each other, but, when an external field is applied, their very large intrinsic ferromagnetic moments allow them to polarize easily to form collectively small magnets, approaching bulk ferromagnetic strength. These characteristics can be effectively exploited for *in-vivo* imaging or drug delivery of nanoparticles moving towards the tumors or other specific physiological targets. When the field is turned off, the nanoparticles exhibit a decaying magnetic field, the remanence field. The remanence field magnitude is strongly dependent upon the magnetic characteristics and the distribution of diameters of the nanoparticles; typically the characteristic time for decay depends on the duration of the polarizing field for a polydisperse size distribution [15-17].

The injection of these particles at a specific area of the body can be accomplished either through intravenous methods or into a specific location where the interaction is required. Typically the blood cells are of few microns size (white blood cell are nearly 10 -14 μm, red blood cells are nearly 8 μm), through which nanoparticles can easily be maneuvered by diffusion. For example, Brightman demonstrated that 9 nm ferritin particles could diffuse rapidly through intercellular spaces, aided by blood pressure gradients, to achieve near uniform distribution in a few minutes [15, 18]. However, it was also shown that larger particles greater than 100 nm in size did not seem to diffuse in this way and tended to remain in the circulation or attached to the walls of the vascular system [19]. The delivery of nanoparticles to a targeted location through blood pressure based diffusion may not provide enough control to facilitate a precise delivery, which necessitates the investigation of alternative means of achieving this. As it is possible to move the magnetic particles in static magnetic field gradients, one can employ static fields in most of the situations. However, inside the human body, within the constraints provided by the presence of complex fluids, the use of oscillatory magnetic fields may be necessary to overcome the particle mechanical arrest/containment. This can ensure the alignment of the particle magnetic moments with the rotating field vector, which provides the alternate pathways to maneuver away from any obstacles, especially in the presence of a complex assembly of fluids and semisolids, as is seen in the human body. It was shown by Dinh and his co-workers that this could be achieved with a reasonable success [20].

2.2 Virtual reality simulation to visualize and track nanoparticle movement

The movement of nanoparticles can be easily visualized and tracked using VR simulation. Though the forces are computed using numerical modeling, none of the existing software packages are capable of providing a real time visualization of the particle movement and their tracks. Further, one of the main objectives of this research is to be able to control the particle movement using a feedback mechanism. The proposed VR techniques fit very well into these requirements especially in allowing real time view of the particle movement, at the same time providing the controlling capability. So far, the VR models were applied to situations, where Atomic Force Microscope (AFM) tips were used as the tools for moving the nanoparticles. The forces between the particle and the tip were represented using different mathematical models [21-26].Some of the simulation data obtained in the modeling of these forces were verified experimentally [23, 24, 26]. This experience helped us in formulating a new model for better simulation of nanoparticle transport.

2.3 Numerical Modeling of nanoparticle transport in magnetic fields and VR application

We have been working on the analysis and modeling of forces on nanoparticles for nanoassembly for the last ten years. The forces derived from these analysis were in agreement with experimental data, confirming the validity of the modeling approach [25, 26]. In addition to all possible conventional forces experienced by the nanoparticles in their manipulation in a given direction using tools like Atomic Force Microscopy and optical tweezers, this study also considered externally applied magnetic forces. For this work, a number of different nanoparticles such as Poly methyl methacrylate (PMMA), gold, Fe_3O_4 and SiO_2 were considered. This constitutes the first part of the investigations in calculating forces experienced by nanoparticles in different environments. The next stage of this work would be the determination of forces experienced by the nanoparticles in more complex solutions (e.g. blood) and cellular milieu (e.g. tumor).

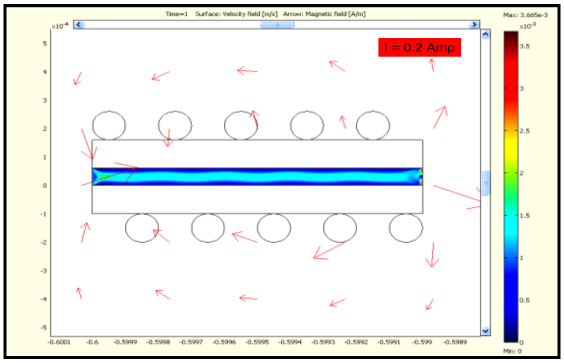


Figure 1. Preliminary results of COMSOL simulation on the ferrofluid velocity profile in a microchannel.

In one of our preliminary numerical analysis studies, we coupled the multiphysics module (incompressible Navier-Strokes equations) with magnetostatics module using COMSOL software. An electromagnetic coil current is assumed to drive the ferrofluid within the microfluidic channel. Dimensions of the channel are selected closer to that of blood vessel. The results of the simulation is shown in Figure 1, which clearly indicates the velocity profile of the fluid at an electromagnetic coil current of *0.2 A*, which corresponds to an approximate field strength of 250 G. These simulations were carried out at different coil currents (up to 2 Amps) and for channels with different dimensions, which helped us in understanding the effect of increasing field (up to 3000 G) on the velocity profiles. The velocity is of the same order as was observed in other studies [27] confirming the qualitative agreement in values. However, the model requires refinement, in terms of achieving relatively accurate predictions. One of the main concerns with this kind of treatment is that the single phase flow assumed in this simulation may not be accurate in the case of nanoparticles transporting through the blood. A more in-depth study into the quantitative analysis may require the consideration of **two phase flow**.

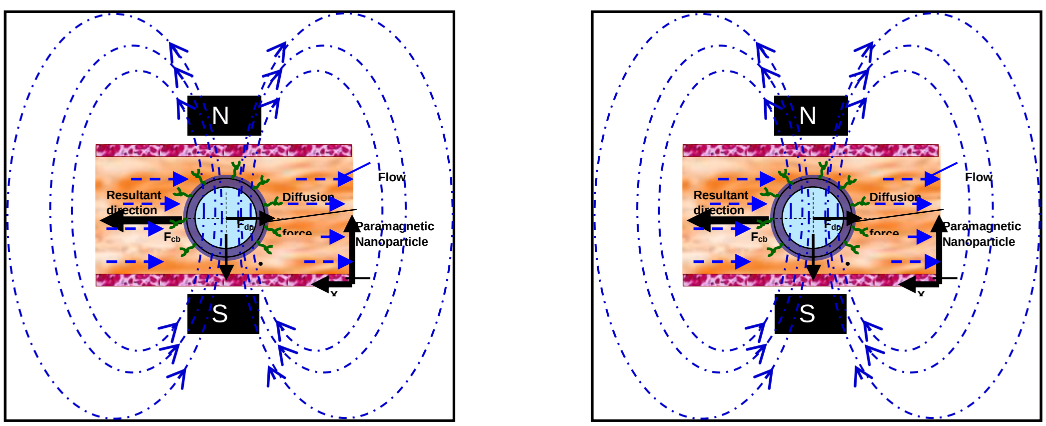


Figure 2 Representation of forces of the nanoparticles within the blood vessel for particle moving midstream

Figure 2 shows typical situations where the nanoparticle is moving within the blood vessel. The outer wall of the particle shown in the figure is mainly representing a surface functionalized layer thickness, while the particle itself is simply paramagnetic. The drug to be delivered or a typical antibody and antigen pair attached to the surface functionalized layer are seen in this figure. The particles in the midstream can potentially result in different friction (F_{fs}) and adhesive forces (F_{as}) compared to those moving closer to the vessel wall, which can affect the particle velocity and the overall resultant force. In addition to these forces, the projected model should also include the capillary forces (F_{cb}), gravitational forces (F_{gp}), and diffusion force (F_{dp}) and Van der Waal's forces (F_{vs}). Among these forces, convective/advective and diffusive forces appear to affect the nanoparticle movement more than the other forces, which will determine the magnetic force required in *in-vivo* particle transportation [6-8, 19].

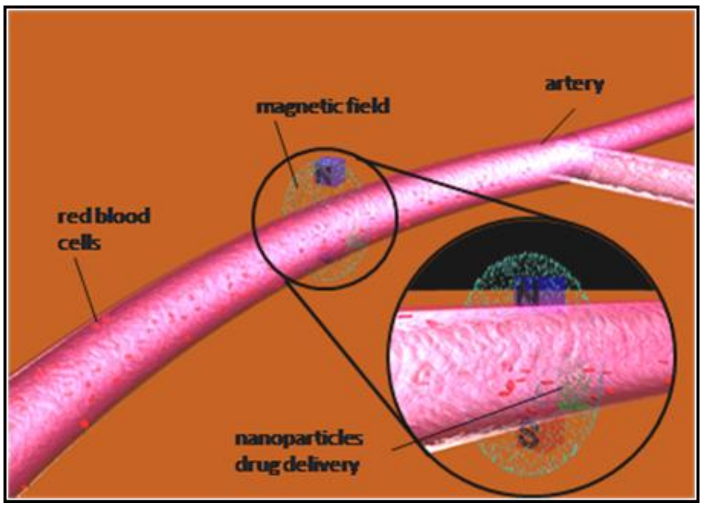


Figure 3. VR simulation of forces and magnetic field tracking of the nanoparticles in an artery during the delivery at a given point (red particles represent the red cells while the green particles shown in enlarged window represent the drug being delivered at an identified location).

As mentioned above, **one of the major goals of the project** is to study and understand the role of different forces involved in the nanoparticle drug delivery. A reasonable understanding of these forces and their inter-relationships will help in deriving an analytical model, which can be used to estimate the forces required for determining the resultant direction and movement of the magnetic nanoparticles *in-vivo*. The analytical force model will be evaluated using numerical simulations to ensure the correctness of the model. **Inputting these forces into a VR simulation, we propose to track the direction and movement of the particles.** Different forces acting on and in between the magnetic nanoparticles determined using the analytical and numerical models are input to the VR simulation for visualization. Inside the VR simulation one can see the nanoparticles represented along with force vectors. A control strategy will be developed in a following step, which helps in steering the particle in a given direction using computer interface and special devices (e.g. haptic force feedback device or data glove) to a desired level, as shown in Figure 3. The output of VR simulation provides force feedback to control and steer the particle. Actual position of the nanoparticles is thus continuously monitored and updated on the VR simulation set-up. This algorithm will repeat to update the new position, forces and direction, so most recent values will be used for next iteration of the calculation.

Although the use of magnetic nanoparticles as carriers for *in vivo* targeted therapeutic drug delivery was first proposed over 15 years ago, there were only few applications that proceeded to Phase I/II clinical trials in spite of successful of preclinical animal studies. The slow progress in this research and application may be attributed to the lack of understanding of the methods and their theoretical underpinnings with respect to engineering the nanoparticles with specific characteristics and steering/retaining these at the required site using localized field configurations and applied fields. The proposed research would be attempting to resolve the two major issues through applying innovative modeling methodologies (combining analytical, numerical and virtual reality simulation) by considering all possible particle forces and developing magnetic nanoparticles with suitable functionalized (biological) and surfactant (to prevent agglomeration) layers. Development and application of these simulation methods will provide the path of engineering and biological innovation for any targeted drug delivery techniques in future.

3. Results and Discussion

3.1 Force Evaluation, Modeling and VR simulation of nanoparticle transport

To utilize engineering modalities to increase therapeutic efficacies of anticancer agents harbored in multifunctional nanoplatforms. We propose to use magnetic field to steer and retain drug carrying nanoparticle to a tumor site. In order to design an effective magnetic transport and retention system one need to have a thorough knowledge of forces experienced by the particles which then provides a way to determine the magnetic field configuration. In addition, the virtual reality simulation of the particle transport will provide a real time tracking and steering capabilities.

3.2 Analytical Modeling of Forces

The proposed work involves the computation of both long- and short-range forces among nanoparticles and those between particle and the surface of the body through which it comes into contact during its motion. In particular, the forces that are important for the simulation include Van der Waal's, Lennard-Jones forces (non-contact), adhesive and capillary forces, repulsive contact forces, and frictional forces [21, 23, 26]. Capillary and adhesive forces are one of the important contributors to the overall force experienced by the particles during their motion. Capillary force can be determined using a simple formula:

$$F_c = 4\pi\gamma R \cos\theta \qquad (2)$$

where γ is the surface tension, R is the particle radius (assuming the spherical nanoparticles) and θ is the contact angle. This will be modified to include the particles of different radii as well as surface (arterial wall) characteristics. As surface tension is a function of particle size up to certain value, this will be taken into account for the particle sizes in the range of few *nms* to 10's *nm*. For larger particles (1 μm and above), surface tension becomes almost negligible. Adhesive forces will also be determined following a similar treatment.

As magnetic particles will also be involved, we need to consider the magnetic forces into the simulation and take the related effects into account, while determining the overall effective force experienced by these particles. These forces will be calculated using appropriate equations and relationships. The magnetic force on given nanoparticles may be expressed mathematically using fundamental electromagnetic equation. Considering the standard Maxwell's electromagnetic equations and following simple analytical derivations, the force experienced by the magnetic nanoparticles can be expressed as:

$$\vec{F} = \vec{m}\sum_{i=1}^{3}\hat{x}_i \frac{\partial B_z}{\partial x_i} - \vec{\nabla}(\vec{m} \bullet \vec{B}\ ) \quad (3)$$

where m is the magnetic moment and B is the induced field. In this, the induced field tends to align the net magnetic moment of the particle in a fixed direction, while the gradient leads to a force that may move the particle. As the magnetic moment of the material, m is proportional to the applied field, H and the intrinsic magnetic susceptibility of the material χ_m, ($m = \chi_m H$), it is possible to increase the field by using appropriately sized magnetic particles with large m. The magnetic force (using a number of assumptions for magnetite particles) is found to be in the order of 10^{-9} to 10^{-12} Newtons.

One of the important forces that the magnetic field will be acting against, during such a nanoparticles movement in the blood stream, is the hemodynamic force, represented by the equation:

$$\vec{F} = 6\ \pi\, \eta\, v\, r \quad (4)$$

where $\vec{F}$ is the drag force, η is the viscosity of the blood (viscosity of the blood is known to be ~ four times that of water), v is the velocity of the nanoparticles and r is the radius. The velocity of the particle relative to the blood flow can be determined by equating the hemodynamic force and the magnetic forces. Different mathematical models can be used to determine particle trajectories depending on the field, particle size and blood vessel sizes [28]. As the particle approach the vessel wall, Stoke's law may not be applicable for determining the drag force. This will have to be treated separately.

This clearly shows that it is important to understand the blood flow characteristics and its variation in different locations of the human body and under different conditions (e.g. blockage in the intravenous system or heart valve malfunction etc.). Blood consists of plasma that almost accounts for *55%* of the volume, which constitutes the red cells, white cells and platelets. The plasma consists of a solution of larger molecules. Blood is also known to be a shear thinning fluid with its viscosity varying as a function of shear rate. Flow in larger blood vessels is thus more uniform and termed as plug flow, while it is more parabolic in the smaller blood vessels. All these factors will have to be taken in to consideration while computing the magnetic forces required to steer the magnetic particles. Though the fluid flow in this context is more complex, simple approximation of the force represented by Equation 3 can be considered as a good starting point.

A detailed analysis of the flow and the calculation of different forces have been carried out in this work. This study involves the consideration of the diffusion forces, Lennard-Jones forces (non-contact), capillary forces, Van der Waal's forces (between the particle and the blood vessel surface and between the particles and cells), friction forces, adhesive forces and magnetic forces. Some of these forces may have a relatively marginal effect compared to others. These will be calculated in the first stage using simple known mathematical relationships. We have extensive experience in calculating these forces for nanoparticles being maneuvered to achieve nanoassembly based manufacturing technologies [21, 22].

In order to determine the forces experienced by the particles due to the internal diffusion within the blood, the diffusion coefficient of these particles will have to be calculated [13].

$$D = \frac{K_b\, T}{3\,\pi\,\eta\, d} \tag{5}$$

Where D, K_b, T and d are diffusion coefficient, Boltzmann's constant, temperature, and particle diameter respectively.

Knowing D, one can calculate the time taken to diffuse to a given distance, thereby able to estimate the velocity and forces experienced by the particle.

3.3 Numerical Modeling of Forces and Nanoparticle velocity profiles

The second stage in the calculations will involve a more detailed and thorough numerical analysis using COMSOL software, which facilitates the modeling of physical, chemical as well as mechanical forces involved in the given situation.
In numerical simulation, we could use the Navier-Stokes equations for incompressible flow [29]:

$$\rho\frac{\partial u}{\partial t} - \nabla\cdot\eta(\nabla u + (\nabla u)^T) + \rho(u\cdot\nabla)u + \nabla p = F \tag{6}$$

where F involves gravity and magnetic forces F_m+F_g; ∇p is the pressure term;
$\eta(\nabla u+(\nabla u)^T)$ is the viscous contribution

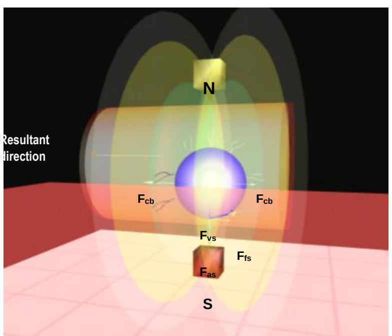
At this stage we can consider two-phase flow approach instead of a single phase flow that was used in our initial work depending on the results. In such a two phase flow model, in addition to the liquid phase, the particle nature of the nanoparticle will be recognized and resolved in the simulations. For a liquid phase, the blood carrier flow will be assumed laminar with small Reynolds numbers. For a rectangular microchannel, Cartesian type of rectangular meshes will be used with mesh refinement near the channel wall. The immerse boundary method (IBM) [30] will be applied in cases where the deformation of the channel wall can become important. The IBM method can account for both small and large dynamic deformations of the physical boundaries of the fluid motion. The boundary motion will be prescribed, as a fully coupled computational fluid-structure interaction approach is beyond the scope of the work proposed here. It can be considered in future research. For the solid phase, the above mentioned driving force mechanisms, such as Van der Waal's, adhesive, capillary, and frictional forces, for nanoparticle transport have been examined and evaluated using the fluid phase solutions. The interactions between the nanoparticles and the constituents of the blood, such as the plasma and the particulate matter, will be accounted for by adopting an effective blood viscosity for the nanoparticle hemodynamic force [31]. It is possible, via order-of-magnitude analyses, to identify the dominant physical forces that will affect the transport process and those forces will be used in COMSOL to study the nanoparticle movement and trajectories.

3.4 Virtual Reality simulation

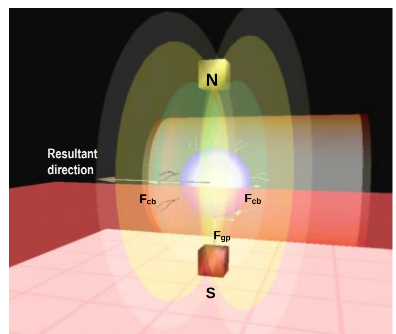
The VR representation defines nanoparticles visualization as 3D models and shows forces applied to the model represented with their vectors that have direction and magnitude**.** A special behavior model is created to control the process of inputting the data from the analytical/numerical models and steer the particle movement using peripherals such as Super glove, Polhemus Fastrak and Phantom force feedback device [32]. These devices allow a person without any experience in VR simulation or magnetic device control to steer the particle in a natural way and at the same time, provides force feedback to feel the forces inside the VR simulation thus enabling him precise steering.

The VR system consists of Windows-7 workstation, using specialized VR (EON 6.0) software. The 3D stereo capabilities and immersions sense will be achieved with stereoscopic Head Mounted Display or 3D stereo wall projection system. For better interaction and precision control, 3D position sensors and power VR gloves can be used that facilitate control and easy movement of the nanoparticles in natural way [33, 34]. In addition, for superior control of the moving nanoparticles a force feedback device, which allows haptic experience, like Phantom (Sensable technologies, MA) have been studied. Acting/reacting forces, collision, and particle motion can easily be controlled using this system. In addition, real time feedback to the user allows natural interaction and control to move the nanoparticles. Precomputed and precompiled constants for specific initial condition, for force equation will be fed to VR simulation to speed up the computation for the math-intensive part of the simulation. All real time calculation for forces, reactions, particle and object moving trajectories will be implemented within the VR simulation using specialized modules. Our previous experience with nanoparticle assembly simulation showed that fast computing and visualization can be achieved with existing technology [35-37]. One of our recent efforts in creating the VR simulation of nanoparticle transport, in the form of video clippings is uploaded into this research group website [38].

In the past, the time delay, or the lag between the user action in the real world and the corresponding update of the state of the object in the virtual world, was a major concern because it leads to unrealistic visualization. With the advances in computation power and using special graphical processors this problem is solved in general case [36]. When nanoparticles with complex computation of dynamic forces are considered, this requires magnitude of computational power many times higher than just for visualization. Our goal is to develop real-time VR simulations that include detailed scale models for Van der Waal's, capillary, contact deformation, electrostatic and magnetic forces, see Figure 4. The magnetic force drives the particle in desired direction in opposition to the forces acting in contact and non-contact regimes. The magnitude and direction of the magnetic force can be easily controlled by the user thus enabling dynamic changes of the position of the particle through control of the velocity. Although some of the forces are relatively small, their direction and magnitude can have influence in determining the reaction force and direction of the movement and hence they are considered within the mathematical model. The magnetic forces are visualized as the driving vectors on the particles together with the rest of the other force vectors.



a



b

Figure 4. VR simulation of forces and magnetic field tracking of the nanoparticles moving

(a) along the sidewall and (b) midstream within the blood vessel

The output of VR simulation provides force feedback to control and steer the particle. Actual position of the nanoparticles is thus continuously monitored and updated on the VR simulation set-up. This algorithm will repeat to update the new position, forces and direction, so most recent values will be used for next iteration of the calculation.

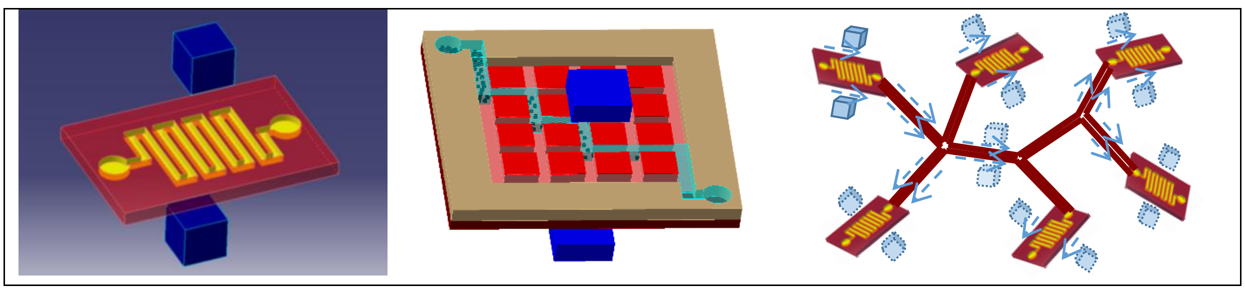
An advantage of the VR steering system is its capability to automate the steering process, without human intervention. Once the desired position is reached the final tracking path can be input as a control algorithm for automated delivery. As a result the particle delivery to a preset position can be achieved automatically with increased or decreased speed if necessary. We will use VR simulation not only to control the system but also as a tool to test and optimize each component of nanoparticles delivery system.

The outcome of this research is the determination of the forces experienced by the multifunctional nanoparticles in different in vivo situations and estimation of the required magnetic force for their steering and retention. Our earlier experience in estimating the forces for nanoassembly were found to be close to the experimental values obtained using an atomic force microscope (AFM) [21, 36]. We has extensive experience in handling virtual reality simulation in a number of different project and we do not anticipate major blockade in overcoming these complication [21, 22, 30, 36]. For example, in 3D space there is no sense of the forces and weight so to represent these force vectors a VR hands simulating assembly and lifting of parts was created. The force vectors are represented by changing the distance between VR hands and its 3D ghost image of the hands. Thus user was able to judge the force required for proper assembly [30].

3.5 In-vitro biochip validation of modeled parameters and VR simulation

Next step is establishing the validity of the developed model and calculated forces experienced by nanoparticle. This will set a platform for *in-vitro* testing of nanoparticle motion, visualization and control.

Biochips with different microfluidic channel network system (shown in Figure 7a, b & c), with dimensions similar to that of arteries, will be fabricated in the laboratory for *in-vitro* testing. A simple microfluidic chip (Figure.5a) with a single channel will be tested applying a known magnetic field. Microfluidic chip with channel network shown in Figure 7b, will provide an idea



a b c

Figure 7 (a, b & c). *In-vitro* biochip designs to be considered for the validation of simulation results.

of particle steering in a more systematic network. To understand the methodology of navigating or steering the particles through the *in-vivo* arteries, a more complex network microfluidic channel system (Figure7c) will be fabricated and tested. This testing will be carried out using axial flow micropumps [39] that are employed to drive biofluids (such as blood) and defined cellular media (with cells), under physiological pressure within the channel network. It may be noted that the movement of a single magnet and its position along the channels is depicted by the arrows. The last configuration helps in understanding the difficulties in transporting the nanoparticles through a complex network of channels having different routing possibilities. These experiments will be performed using a commercially available electromagnetic system and will be interfaced with the VR system, which provides the control strategies for steering the nanoparticle in a given direction. The simulated velocity profile obtained from numerical simulation will be compared with that observed in the actual system. Geometry and dimensions of the microfluidic channels for different materials (mainly biocompatible materials e.g., polyurethane, polycarbonate-urethane, polyimide, and polyethylene) will be investigated. From initial *in-vitro* experiments using microfluidic channels, we found that it requires flux densities of around 4000 gauss. These are in agreement with the values reported in the literature with some clinical trials [19, 32, 40]. Microfluidic channels have been fabricated using excimer laser micromachining and lithography techniques. We has previous experience in making bio-chips for different applications. Nanoparticles with and without surfactant coating and surface functionalization (and with attachment of 'a' drug structure to mimic the actual application) will be tested in these *in-vitro* tests. These investigations are focused to find the effect of 'hydrophobicity' and 'hydrophilicity' characters of the particles (using different surface functionalization layers) on the forces and the velocity of nanoparticles

While testing the biochip, we will be comparing the particle velocity profile and instantaneous positions with simulated values through VR simulation. This helps in evaluating the accuracy of our models and any corrections required for the magnetic force. This kind of comparison will also help in calibrating the VR simulation with real time values.

3.5 Discussions

The difficulty in visualizing and tracking the nanoparticles in vivo is one of the major impediments in clinical applications of the nano-platform based drug delivery system. Magnetic Resonance Imaging (MRI) has so far been used as the primary tool for this purpose, which is expensive and not easily accessible. Hence, we will formulate an 'open architecture' framework and create a VR tool that is capable of driving the nanoparticles in a continuously changing process by analyzing and inputting the forces dynamically. The proposed research facilitates the application of different control algorithms, ultimately providing a novel and real time visualization and tracking of these particles. The availability of such tracking and steering tools will enable the realization of relatively simpler drug delivery techniques. In summary, this proposal seeks to engineer a new generation simulation tool using the known biological and physical inputs for targeted drug delivery systems

4. Conclusions

This project will help in establishing the scientific foundations required for developing reliable and precise drug delivery systems with real time monitoring capabilities. The proposed study and prediction of the nanoparticle movement can facilitate the realization of localized drug delivery mechanisms, capable of reducing systemic toxicity and increasing efficacy of a therapeutic regimen by elevating the local concentration of the drug at the target site.

References

[1] Reddy GR, Bhojani MS, McConville P, Moody J, Moffat BA, Hall DE, Kim G, Koo YE, Woolliscroft MJ, Sugai JV, et al. (2006). Vascular targeted nanoparticles for imaging and treatment of brain tumors. Clin Cancer Res 12, 6677-6686.

[2] Cho K, Wang X, Nie S, Chen ZG, and Shin DM (2008). Therapeutic nanoparticles for drug delivery in cancer. Clin Cancer Res 14, 1310-1316.

[3] Bhojani MS, Reddy GR, Koo Y, Philbert M, Kopelman R, Rehemtulla A, and Ross BD (2007). Multifunctional nanoparticles for targeted imaging and therapy, in Cancer Nanotechnology. Ed: Hari Singh Nalwa and Thomas J.Wester American Scientific Publishers, 81-89.

[4] Silva GA (2007). Nanotechnology approaches for drug and small molecule delivery across the blood brain barrier. Surgical neurology 67, 113-116.

[5] Koo YE, Reddy GR, Bhojani M, Schneider R, Philbert MA, Rehemtulla A, Ross BD, and Kopelman R (2006). Brain cancer diagnosis and therapy with nanoplatforms. Advanced drug delivery reviews 58, 1556-1577.

[6] Hwang CW, Wu D, and Edelman ER (2001). Physiological transport forces govern drug distribution for stent-based delivery. Circulation 104, 600-605.

[7] Karanjakar A (2009). CFD in the design of drug delivery systems Innovations in pharmateutical technology, 74-78.

[8] Proshin AP, and Solodyannikov YV (2005). Mathematical Modeling of Blood Circulation System and Its Practical Application. Automation And Remote Control 67.

[9] Lubbe AS, Alexiou C, and Bergemann C (2001). Clinical applications of magnetic drug targeting. The Journal of surgical research 95, 200-206.

[10] Dobson J (2006). Magnetic nanoparticles for drug delivery. Magnetic Nanoparticles for Drug Delivery 67, 55-60.

[11] Dames P, Gleich B, Flemmer A, Hajek K, Seidl N, Wiekhorst F, Eberbeck D, Bittmann I, Bergemann C, Weyh T, et al. (2007). Targeted delivery of magnetic aerosol droplets to the lung. Nature nanotechnology 2, 495-499.

[12] Berry CC, Wells S, Charles S, and Curtis AS (2003). Dextran and albumin derivatised iron oxide nanoparticles: influence on fibroblasts in vitro. Biomaterials 24, 4551-4557.

[13] Chatterjee J, Haik Y, and Chen C-J (2003). Size dependent magnetic properties of iron oxide nanoparticles. Journal of Magnetism and Magnetic Materials 257, 113-118.

[14] Kim T, and Shima M (2007). Reduced magnetization in magnetic oxide nanoparticles. Journal Of Applied Physics 101, 09M516.

[15] Brightman MW (1965). The distribution within the brain of ferritin injected into cerebrospinal fluid compartments. II. Parenchymal distribution. Am J Anat. 117, 193-219.

[16] Flynn ER, and Bryant HC (2005). A biomagnetic system for in vivo cancer imaging. Physics in medicine and biology 50, 1273-1293.

[17] Kalambur VS, Han B, Hammer BE, Shield TW, and Bischof JC (2005). In vitro characterization of movement, heating and visualization of magnetic nanoparticles for biomedical applications. Nanotechnology 16, 1221-1233.

[18] Tang Q, Zhang Y, Chen L, Yan F, and Wang R (2005). Protein delivery with Nanoscale precision. Nanotechnology 16, 1062-1068.

[19] Pankhurst QA, Connolly J, Jones SK, and Dobson. J (2003). Applications of magnetic nanoparticles in biomedicine. Journal of Physics D: Applied Physics 36, R167-R181.

[20] Dinh TN (2003). Magnetic targeted drug delivery. Master of Science, Thesis, Virginia Polytechnic Institute and State University.

[21] Ikonomov PG, and Ghantasala M (2006). Virtual Reality approach for nanoparticles tracking using simulated forces. Proceedings of the NSTI Nanotechnology Conference-2006, Boston, MA, USA 7-11th May, 1, 693-696.

[22] Ikonomov PG, Ghantasala M, Dwivedi S, and Lyons DW (2006). Simulation of Forces for Virtual Nanoassembly. Journal of Advanced Manufacturing Systems 9, 63-70.

[23] Kim D-H, Park J, Kim B, and Kim K (2002). Modeling and Simulation of Nanorobotic Manipulation with an AFM probe. International Conference on control and automation systems ICCAS 2002.

[24] Perez JM, Simeone FJ, Saeki Y, Josephson L, and Weissleder R (2003). Viral-induced self-assembly of magnetic nanoparticles allows the detection of viral particles in biological media. Journal of the American Chemical Society 125, 10192-10193.

[25] Sitti M, and Hashimoto H (2000). Controlled Pushing of Nanoparticles: Modeling and Experiments. IEEE/ASME TRANSACTIONS ON MECHATRONICS 5, 199-211.

[26] Sitti M, and Hashimoto H (2003). Teleoperated Touch Feedback From the Surfaces at the Nanoscale:Modeling and Experiments. IEEE/ASME TRANSACTIONS ON MECHATRONICS 8, 287-298.

[27] Cummings LJ, Richardson G, and Hazelwood L (2000). Drug delivery by magnetic microspheres Proc. Group. Mathematics in Medicine Study (Nottingham, UK: University of Nottingham) 69-83.

[28] Meyer CR, Boes JL, Kim B, Bland PH, Zasadny KR, Kison PV, Koral K, Frey KA, and Wahl RL (1997). Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations. Medical image analysis 1, 195-206.

[29] Sugii Y, Nishio S, and Okamoto K (2002). Measurement of a velocity field in microvessels using a high resolution PIV technique. Annals of the New York Academy of Sciences 972, 331-336.

[30] Ikonomov PG, and Milkova ED (2004). Virtual Assembly/Disassembly System Using Natural Human Interaction and Control, . Virtual and Augmented Reality Applications in Manufacturing, Springer-Verlag London Ch.7, 111-125.

[31] Prices AR, Secomb TW, and Gaehtgens P (1996). Biophysical aspects of blood flow in the microvasculature. Cardiovasc Res 32, 654-667.

[32] Pulfer SK, and Gallo JM (1997). Targeting magnetic microsheres to brain tumors. Proceedings of the Scientific and Clinical Applications of Magnetic Carriers, 445–455,.

[33] Kode S (2007). Virtual Simulation and Control for Nanoassembly Using Optical Tweezers AFM and Magnetic Forces. M.Sc. thesis, University of Lousiana at Lafayette.

[34] Kode S, Ikonomov P, and Dwivedi SN (2007). Simulation of Virtual Nanoparticle Tracking Using Calculated Forces for an Atomic Force Microscope and Optical Tweezers. Proceedings of the ASEE-NE conference West Virginia, 30-31 March 2007.

[35] Bhavaraju K (2008). An Interactive Virtual Reality Simulation for Nanoparticle Manipulation and Nano-Assembly Using Optical tweezers. M.Sc. thesis, University of Lousiana at Lafayette, .

[36] Bhavaraju K, Ikonomov P, Dwivedi S, and Choudhury A (2008). An Interactive Virtual Reality Simulation for Nanoparticle Manipulation and Nano-Assembly Using Optical Tweezers. Proceedings of the ASEE 2008 Conference, Pittsburg, PA. S2464.

[37] Paudel A, and Ikonomov P (2008). AFM in Nano-Manipulation: Virtual Modeling and Simulation. Proceedings of International Conference of Agile Manufacturing ICAM-2008, July 16-18, , Kalmazoo, MI, 43-48.

[38] Ikonomov PG, Magnetic nanoparticles drug delivery http://www.ceas.wmich.edu/nanoparticle_drug_delivery/

[39] Shumakov VI, Tolpekin VE, Khaustov AI, Romanov OV, and Melemuka IV (1994). Axial-Flow Micropumped System For Assisted Circulation. Biomedical Engineering 28, 231-233.

[40] Grief AD, and Richardson G (2005). Mathematical modelling of magnetically targeted drug delivery. Journal of Magnetism and Magnetic Materials 293, 455-463.

BIPV: Integration of Photovoltaic with the Construction to achieve the concept of Agility

Gaurav Agarwal, Abhishek Agarwal, Shubham Agarwal
Student, B.tech. 2nd year, Faculty of Engineering, Dayalbagh Educational Institute, Agra,, INDIA
agarwalg777@gmail.com

Abstract: Building Integrated Photovoltaics (BIPV) refers to photovoltaic cells and modules which can be integrated into the building envelope as part of the building structure, and therefore can replace conventional building materials, rather than being installed afterwards. Together with the evolvement of the integration of modules in the architecture, the new BIPV products are able to fully replace some building components - the construction parts of the building envelope. BIPV is a multifunctional technology. Besides being a source of electricity, several other purposes can be achieved, such as weather protection, thermal insulation, noise protection or modulation of daylight. BIPV systems represent a powerful and versatile tool for achieving the ever increasing demand for zero energy and zero emission buildings of the near future. In this respect BIPVs offer an aesthetical, economical and technical solution to integrate solar cells harvesting solar radiation to produce electricity within the climate envelopes of buildings with agility.

Introduction

Currently, the world is using fossil fuel at an alarming rate that not only will strain the sources in the near future, but will result in a great amount of pollution as well. The power industry emissions were 10.9 gigatonnes of carbon dioxide equivalents ($GtCO_2e$) per year in 2005, i.e. 24% of global Greenhouse Gas (GHG) emissions, and this is expected to increase to 18.7 $GtCO_2e$ per year in 2030[1]. ‘‘Carbon dioxide equivalent is the unit for emissions that, for a given mixture and amount of greenhouse gas, represents the amount of CO_2 that would have the same global warming potential (GWP) when measured over a specified timescale (generally 100 years)’’ [1].

Of all the renewable energy resources currently available, solar energy is the most abundant, inexhaustible and clean one [2]. In one day, the radiation from the sun on the earth gives about 10,000 times more energy than the daily use for mankind [3]. The challenge is to collect this available energy at a reasonable cost.

As the world’s demand and focus on renewable and non-polluting energy, together with energy efficiency, are ever increasing, zero energy and zero emission buildings are rapidly drawing attention. Zero energy building, need to harvest energy from its surroundings, where energy from the sun is one of the obvious choices. Building integrated photovoltaic (BIPV) systems, where solar cells are integrated within the climate envelopes of buildings and utilizing solar radiation to produce electricity may represent a powerful and versatile tool for reaching these goals with respect to aesthetical, economical and technical solutions [3].

Definition of BIPV

Building integrated photovoltaics (BIPVs) are photovoltaic materials that replace conventional building materials in parts of the building envelopes, such as the roofs or facades. Furthermore, ‘‘BIPV are considered as a functional part of the building structure, and they are architecturally integrated into the building’s design’’. BIPVs have a great advantage when compared to non-integrated systems because there is neither need for allocation of land nor facilitation of the PV system. BIPVs are considered as one of four key factors essential for future success of PV. The on-site electricity producing PV modules can reduce the total building material costs and achieve significant savings in terms of the mounting costs, especially since BIPVs do not require additional assembly components such as brackets and rails etc. The BIPV system generates electricity out of sunlight, with no pollution. All these advantages have caused a worldwide growing interest in BIPV products.

BIPV as a multifunctional building element

BIPV is a multifunctional technology. Besides of being a source of electricity, several other purposes can be achieved, such as weather protection, thermal insulation, noise protection or modulation of daylight (Figure 1). For example, roof integrated systems integrate PV modules into roof tiles; facade integrated systems act as a rain screen; and semitransparent installations can allow for some of the light to enter for day lighting or viewing.

All this has been made possible by the rapid progress of the PV technology which evolved from rigid, standardized and thick solar panels into a variety of solar modules available in rigid and flexible format, opaque or semitransparent, mate or in different colors, providing today’s designers with a rich toolbox with which to expand traditional architecture and transform buildings into energy producing constructions

The following is a non-exhaustive list shown in Fig. 1 of the functions that PV modules can perform beside the fundamental production of electricity:

- Noise protection (up to 25 db sound dumping is possible)
- Thermal insulation (heating as well as cooling), improving the efficiency of cells by cooling through rear ventilation, isolation function is possible as well.
- Visual cover/ refraction (one-way mirroring visual cover)
- Electromagnetic shielding (can be used as Faraday cage but also as repeating antennas)
- Aesthetic quality (integration in buildings as a design element)
- Safety (safety glass function is possible)
- Weatherproof (waterproof and windproof façade or roof of a building)
- Sun protection/ shadowing (degree of shadowing is eligible through positioning and degree of transparency)

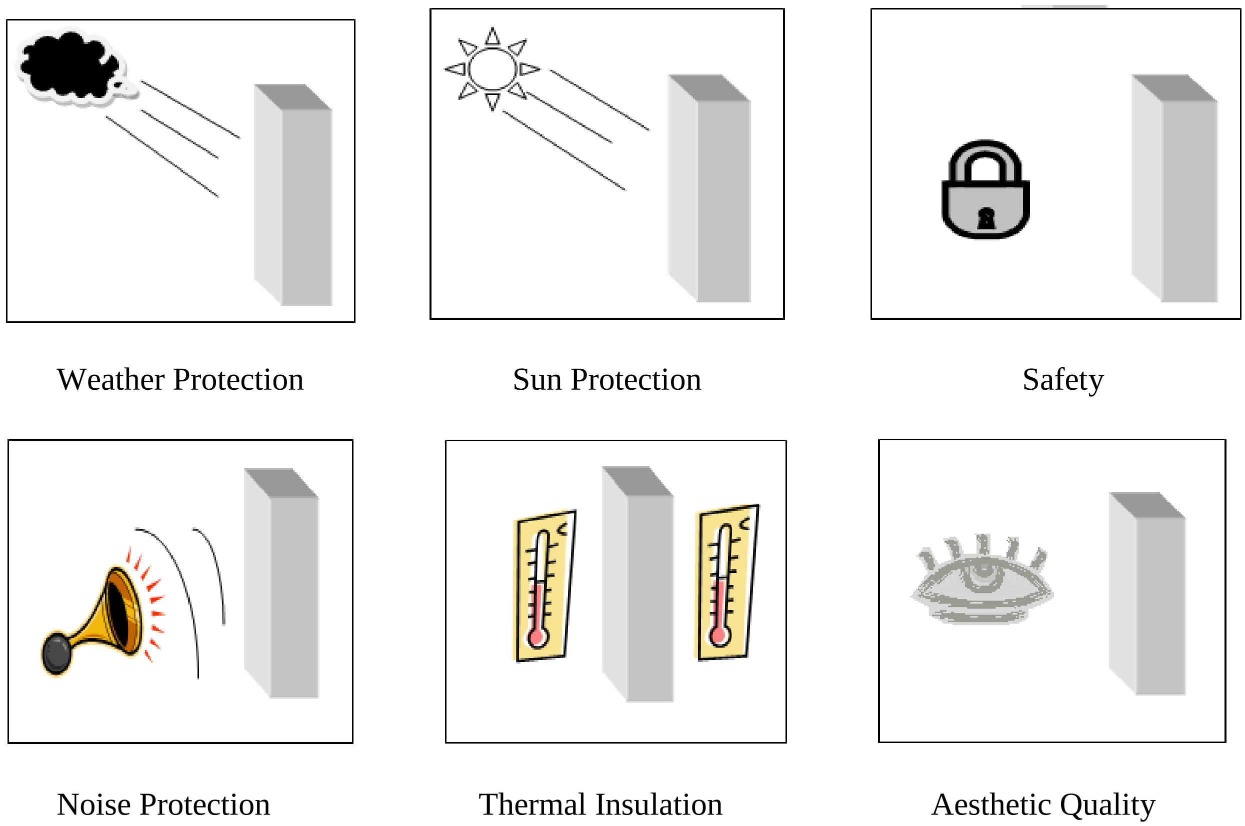


Figure 1: Multiple functions of PV modules[4]

Different levels of PV building integration

PV modules can be applied into the building envelope in several forms depending on their level of integration and on the functionalities they can perform. The large variety and different characteristics of the available BIPV products makes it possible for them to fully replace many of the building components, mainly in façades and roofs. The Building envelope guarantees a border between the controlled inner building environment and the outer climate. Moreover it is bound to be waterproof and fulfill the priority of ensuring the comfortable inner-climate with the possibly least energy expense. Thus, facades and roofs takes over a regulation and control functions in relation to the daylight, ventilation, energy, safety, demarcation and privacy protection, etc. When the Photovoltaic Modules are to be integrated in the building envelope, all these elements need to be considered during the design phase in order to obtain the most suitable product. Additionally, other requirements (from the aesthetics as well as from the pure construction point of view) are to be met. Those are for instance:

- Color, appearance, size
- Weather-tightness
- Wind, rain and snow load
- Resistance and maintenance
- Safety in the construction and utilization phases (fire, electricity and mechanical safety)
- Costs
- Weight and materials used

Test and standards of BIPV

The evaluation of the different BIPV products may involve the following property parameters:

- Solar cell efficiency $\eta = P_{max}/(E)(A)$, where P_{max} is the maximum power point in W or Watt-peak (Wp), E is the input light irradiance in W/m^2 and A is the surface area of the solar cell in m^2
- Open circuit potential or voltage, U_{OC}
- Short circuit electrical current, I_{SC}
- Maximum power point, $P_{max}=(UI)_{max}$
- Fill factor FF is given by $FF= P_{max} /(U_{OC}I_{SC})=(UI)_{max}/(U_{OC}I_{SC})$
- Band gap, Eg
- Quantum yield, Φ=no. of photo-electrons/no. of photons

Classification of BIPV products and their applications

The range of BIPV products is very wide, and they may be categorized in several different ways. Within this work the categorization is mainly performed based on the product descriptions from the manufacturers. BIPV products can be classified in five main categories:

- Standard in-roof systems
- Semitransparent systems (glass/glass modules)
- Cladding systems
- Solar tiles and shingles
- Flexible laminates

Each of the above mentioned categories has a specific market to be addressed, although they all involve the same types of technologies. Within those categories different types of applications are addressed, the main being:

- Pitched roofs
- Flat roofs
- External Building Walls
- (Semi-) transparent facades
- Skylights
- Shading systems

Table 1: Overview of BIPV product and their fields of application

Type of Application / Type of product	Pitched Roofs	Flat roof	External Building Walls	Semi transparent Facades	Skylights	Shading System
Standard in-roof systems	✓					
Semitransparent systems (glass/glass modules)				✓	✓	✓
Cladding systems			✓			
Tiles & shingles	✓					
Flexible laminates		✓				

1. Standard in-roof Systems

Standard in-roof systems are the simplest and most common approach to BIPV taken by the crystalline silicon PV industry by just modifying existing panel designs and mounting systems to make them thinner, more uniform, and flush-mountable on top of existing roofing or siding. This approach has benefits from the manufacturer's perspective because minimal retooling and redesign is needed.

Table 2: Literature data for some of in-roof products [5].

Manufacturer	Product	η (%)	U_{OC} (V)	I_{SC} (A)	P_{MAX} (W)	FF	Area (mm x mm)	P_{MAX}/area (W/m²)
Uni-Solar	PVL-68 PVL-144		23.1 46.2	5.1 5.3	68/module	0.58	2849 x 394 5486 x394	60.6 66.6
Hauptsitz	Sun Power 220 Solar Panel	17.7	48.6	5.7	144/module		1559 x798	
Isofoton	ISF-240	14.5	37.1	8.4	240	0.77	1667 x 994	144.8

✓ Pitched Roofs

The main application for this type of solutions is pitched roofs as they can easily be fixed on top of them. By using standard PV modules combined with a special frame and gutters a fully integrated system can be made. In the Fig. 2 the system is not placed above the existing roofing material but forms an integral part of the roof (BIPV). The PV system can be completely roof filling or it can be installed for instance partly between conventional roof tiles. These solutions require an impermeable roof covering that has to end in the eave

Figure 2: Example of an in-roof PV system.

Semi-transparent Solutions (glass/glass modules)

Although glass/glass based modules are currently still not enough transparent to let in enough sunlight where it is the most needed, they are integrated where only some sunlight penetration is required, being added mainly for aesthetical reasons rather than structural. They are also commonly used to provide sun/wind protection to building surfaces and interiors. The market sectors where these types of BIPV are commonly used are skylights, semi-transparent facades, curtain walls and shading structures (canopies, atrium roofing). The amount of light desired to go through the designed structures can be customized by dimensioning and adjusting the number and spacing of cells in the case of crystalline silicon technology or by modifying the manufacturing process in the case of thin-film. In both cases the more transparent the module, the lower the energy efficiency.

Table 3: Literature data for some semi-transparent products [5].

Manufacturer	**Product**	**η** (%)	**U_{OC}** (V)	**I_{SC}** (A)	**P_{MAX}** (W)	**FF**	**Area** (mm x mm)	**P_{MAX}/area** (W/m²)
Abakus Solar AG	Peak in P210-60		36.50	7.70			2000 x 1066	
Vidursolar	FV VS16 C36 P120		21.6	7.63			1600 x 720	
Glaswerke Arnold GmbH & Co KG	Voltarlux-ASI-T-Mono 4-fach		93	1.97	100/ module	0.55	2358 x 1027	41.3
Schott Solar	ASI THRU-1-L ASI THRU-4-IO	6 6	111 111	0.55 2.22	42 190	0.79 0.77	1122 x 690 1122 x 2619	62.0 64.7
Sapa Building System	Amorphous Silicon thin film Poly crystalline Mono crystalline high efficient	5/cell 16 /cell 22 /cell			32/cell 1.46-3.85 /cell 2.9-3.11 /cell		576 x976/cell 156 x156/cell 125 x125/cell	50 120 155

✓ Skylights

These structures are usually one of the most interesting places to apply PV. They combine the advantage of light diffusion in the building while providing an unobstructed surface for the installation of PV modules. In this type of application, PV elements provide both electricity and light to the building.

Figure 3: Greenhouse in Munich[6].

✓ (Semi-) transparent façades

Glass PV laminates can be applied to windows providing a semi-transparent façade. This can be done with opaque crystalline silicon cells with an adjustable spacing or thin film photovoltaic material that is partially transparent. As the transparency can be varied it can be adjusted to the specific visual and sun shade needs. Integrated into a glass facade it offers limited shading and production of electricity. Applications are typically large atria and façades. As it is a glass product it may give all such functions that glass have, such as insulation (double glazing insulated windows) and fire protection. The cells in glass modules give a playful effect of shading patterns in the space lying behind or underneath.

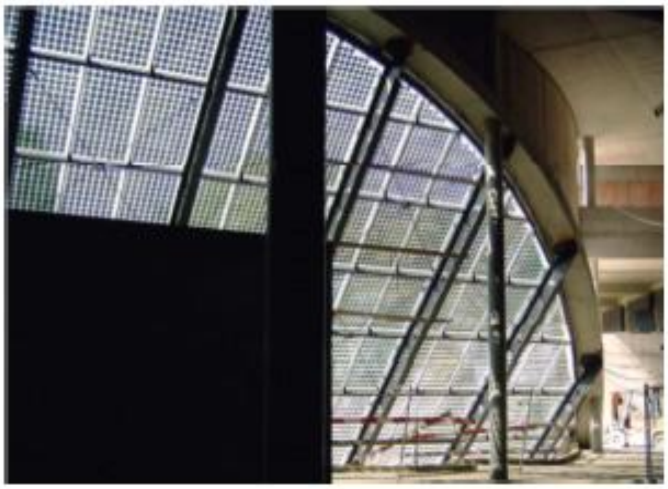


Figure 4: BIPV transparent facade Aalst, Belgium[6]

Figure 5: Semi-transparent facade[7]

✓ Shading systems

PV modules of different shapes can be used as shading elements or as part of an overhead glazing structure. Since many buildings already provide some sort of structure to shade windows, the use of PV shades should not involve any additional load for the building structure. The exploitation of synergy effects reduces the total costs of such installations and creates added value to the PV as well as to the building and its shading system. PV shading systems may also use one-way trackers to tilt the PV array for maximum power while providing a variable degree of shading.

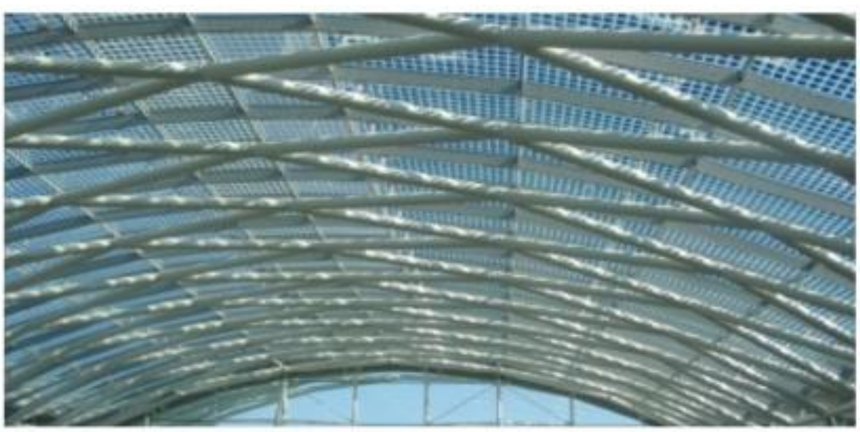


Figure 6: Council's 'Customer First Centre', Hackney,UK.

Cladding Systems

These solar modules can be integrated into the building walls like a conventional cladding element used in facades such as granite or marble. Usually a void is created between the cladding element and the walls offering temperature insulation for the building and improved efficiency for the modules because of the ventilation.

Figure 7: Cladding systems[6]

Solar Tiles and Shingles

The currently available products of this category include i) tiles, designed to interlace with conventional roofing tiles; ii) larger tiles that serve as entire roof wall portions themselves; and iii) thin, flush-mounted panels that overlay conventional roofing or siding. Some tile products may resemble curved ceramic tiles and will not be as area effective due to the curved surface area, but may be more aesthetically pleasing. Some examples of BIPV tile products are given in Table 2, with two of them depicted in Fig. 4.

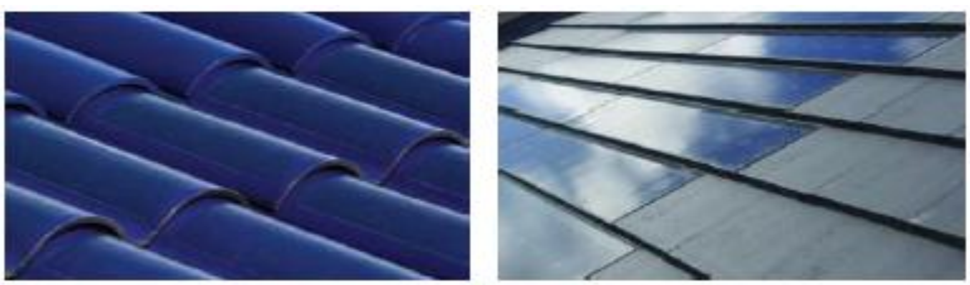


Figure 8: BIPV tile products from SRS Energy (left)[14] and Solar Century (right) [15]

Table 4: Literature data for some BIPV tile and shingles products [5].

Manufacturer	Product	η (%)	U_{OC} (V)	I_{SC} (A)	P_{MAX} (W)	FF	Area (mm x mm)	P_{MAX}/area (W/m^2)
Solardachstein	STEP design		23.15	2.4	1.36/cell	0.76	8 units 100 x100	136
SRS Energy	Sole Powertile		6.3	4.6	15.75/module	0.54	868 x457.2	39.7
Lumeta	Solar Flat Tile		7.4	5.2	28/module	0.73	432 x 905	71.6
Solar Century	C21e Tile	20/cell	12	5.5	52/module	0.78	1220 x 420	101.5

Flexible Laminates

BIPV foil products are lightweight and flexible, which are beneficial with respect to easy installation and prevailing weight constraints for roofs. The PV cells are often made from thin-film cells to maintain the flexibility in the foil and the efficiency regarding high temperatures for use on non-ventilated roof solutions. In most cases the flexible product comes encapsulated and only has to be attached to the existing flat or curved roof. In few cases it supplies a nonencapsulated functional PV laminate on a metal or plastic carrier foil.

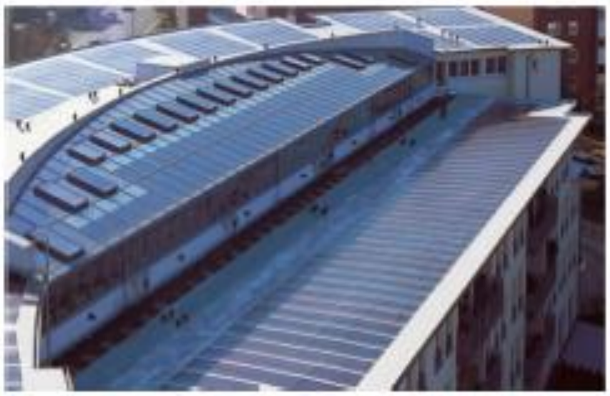


Figure 9a: Membrane Evalon Solar[8]

Figure 9b :BIOSOL thin film plate[9]

Table 5: Literature data for of the foil products [5].

Manufacturer	Product	η (%)	U_{OC} (V)	I_{SC} (A)	P_{MAX} (W)	FF	Area (mm x mm)	P_{MAX}/area (W/m²)
Alwitra GmbH & Co.	Evalon V Solar 40S		138.6	5.1	408/module	0.58	1550 x 6000	42.9
	Evalon V Solar 136		46.2	5.1	136/module	0.58	1050 x 3360	38.5

Comparison

The tile products are more likely to be used on tiled roofs, i.e. residential houses. Due to the easy retrofitting with these products, this market is large. The other products can be used on most structures either together with traditional roofing material or covering the entire roof. The PV foil has a very wide range of usage due to the flexibility, but the efficiency is low, thus the applied area must be relatively large in order to achieve an output comparable with the other products. The modules and the solar cell glazing products can be used on both roofs and facades achieving esthetically pleasing results. This also facilitates using the areas with the highest levels of solar irradiance on geographically challenging locations. However it is not justified to make a comparison between these products, due to different areas of application as well as different demand for effect, costs and available area.

CONCLUSIONS

A building needs heating in winter, cooling in summer and electricity. Eventually, all these three energy will be generated by BIPV. Solar heating facades and roofs covered with thermal solar collectors will provide heating for utilizable water and for support of heating when required; whereas BIPV systems will produce the electricity output required to the building including fresh air needed in summer. In general, a PV system can contribute significant added value to the building in terms of value and image, because all energy-related features of a building, e.g. energy consumption characteristics, play an increasing role in the value determination of a building. BIPV products are capable of creating solar facades, skylights, windows, roofs, walls and multiply the design options of architects and act as multifunctional architectural elements providing many services beyond electricity generation. These may give a great aesthetical appearance in addition to provide weather tightness, solar shading and natural lighting.

It is expected that, continued research and development within both PV and BIPV materials and technologies will yield better and better BIPV solutions in the years to come, e.g. with respect to increased solar cell efficiency, reduced production costs , long-term durability versus climate exposure and improved building integration. These are agile are stepping stones for sustainable future.

References

[1] McKinsey & Company, Pathways to a Low-Carbon Economy. Version 2 of the Global Greenhouse Gas Abatement Cost Curve, McKinsey & Company, 2009.

[2] C. Peng, Y. Huang, Z. Wu, Building-integrated photovoltaics (BIPV) in architectural design in China, Energy and Buildings 43 (2011) 3592–3598.

[3] Swiss BiPV Competence Centre, Energy from the Sun, /http://www.bipv.ch/ index.php?option=com_content&view=article&id=213&Itemid=3&lang=enS, 2010a (accessed November 18, 2011).

[4] Crassard F, Rode J. The evolution of building integrated photovoltaics (BIPV) in the German and French technological innovation systems for solar cells. Master of Science Thesis in Management and Economics of Innovation. Chalmers University of Technology: Go¨ teborg, 2007.

[5] B.P. Jelle, C. Breivik and H.D. Røkenes, Building integrated photovoltaic products: A state-of-the-art review and future research opportunities, *Solar Energy Materials and Solar Cells*, doi: 10.1016/j.solmat.2011.12.016, in press, 2012.

[6] See at the URL: www.design-buildsolar.com

[7] See at the URL: www.schott.com/architecture

[8] See at the URL: www.alwitra.com

[9] See at the URL: www,centrosolar.de

[10] DuPont, <http://www2.dupont.com/Photovoltaique_Integre/fr_FR/assets/downloads/ DuPont_BIPV_flyer3.pdf>, 2011, (accessed December 6, 2011).
[11] Schott Solar AG, ASI® Glass photovoltaic modules, <http://www.schottsolar.com/de/>, 2011, (accessed February 7, 2011).
[12] B.P. Jelle, A. Hynd, A. Gustavsen, D. Arasteh, H. Goudey and R. Hart, Fenestration of today and tomorrow: A state-of-the-art review and future research opportunities, *Solar Energy Materials and Solar Cells* 96 (2012) 1–28.
[13] Alwitra GmbH & Co. using amorphous silicon cells from Uni-Solar, <http://www.cythelia.fr/images/file/membranes/Brochure_evalon-solar_en.pdf>, 2011, (accessed December 6, 2011).
[14] SRS Energy, <http://www.srsenergy.com/maint/files/SPT16%20Technical%20Specifications%20090310.pdf>, 2010, (accessed October 5, 2010).
[15] Solar Century, <http://www.solarcentury.co.uk/installers-and-roofers/products/solar-tiles/c21e-solar-roof-tiles/>, 2011, (accessed December 14, 2011).
[16] Creaton AG, <http://www.creaton.de/en/productrange/roof-accessories/photovoltaics/>, 2011, (accessed December 6, 2011).
[17] Rheinzink, <http://www.rheinzink.com/en/products/roof-systems/roof-covering-systems/quick-stepr-the-rheinzink-steppedroof/>,
2011, (accessed December 6, 2011).
[18] Sapa Building System, <http://www.sapagroup.com/Companies/Sapa%20Building%20System%20AB/Pictures/brochures/Solar_BIPV_low.pdf>, 2011,
(accessed December 6, 2011).
[19] Uni-Solar (left), <http://www.uni-solar.com/products/commercial-products/pvl/>, 2011, (accessed December 6, 2011).
[20] Hauptsitz, <http://www.schweizer metallbau.ch/fileadmin/user_upload/00_Produkte/80 _Sonnenenergie- Systeme/pdf_f/pdf_e/Sunpower_black_e pdf>, 2011, (accessed October 5, 2010).

Manufacturing and Electroplating of Nanoengineered Polymers

Brandon J. Voelker[1], Muralidhar K. Ghantasala[1], Paul V. Engelmann[2] and Jeff Wheeler[3]

[1] Department of Mechanical and Aerospace Engineering, [2]Industrial and Manufacturing Engineering Department, Western Michigan University,
1903, West Michigan Avenue, Kalamazoo, MI-49008
[3] exactheat LLC. 1045 Harts Lake Road, Battle Creek, MI 49037

A polyethylene based polymer with metallic nanoparticles was recently developed as a nanoengineered polymer. The motivation for this project is the proposed use of these novel nanoengineered polymeric materials in many industrial and commercial applications. This necessitated the production of sheets using the nanoengineered polymer and electroplating selected areas using different metallic layers. The nanoengineered polymer resin provided by SABIC as beads of~1 mm dia. was formed into sheets using compression molding. This research exploited selective electroplating methods to deposit nickel and copper layers of different thicknesses using pulsed electroplating techniques. These studies showed that higher duty cycle helped to improve the plated film properties. The quality of the electroplated nickel and copper films were examined using Scanning Electron Microscopy and Optical Microscopy methods, while the mechanical properties (e.g. Young's Modulus, Hardness) of these films were analyzed using Nanoindentation. This paper describes the details of the manufacturing methods used for making the sheets of these novel nanoengineered polymers and further discusses the results of our studies on the electroplated copper and nickel films.

1. Introduction

The development of a nanoengineered polymer has opened up great potential for many new commercial and industrial applications. The polyethylene based material with metallic inclusions benefits from a wide variety of forming and manufacturing possibilities. The polymer resin ships from the manufacturer as 1 mm diameter polymeric beads. All subsequent forming operations were completed at Western Michigan University's Plastics Lab. Several forming methods and parameters were examined to determine the most robust processing technique to produce the required substrates, without altering the semi-conductive nature of the polymer.

Industrial and commercial usage of this polymer benefits from the addition of an electrically conductive metallic coating. Successful application of a metallic coating lowers the resistance to negligible levels in select patterns atop the polymer substrate. Previous research into the electroplating of insulator and semi-conductive polymers has been met with limited success.

One method involves an electrode in contact with the substrate to be plated. This technique depends on the outward propagation of the deposited material originating at the contact point of the electrode with the substrate [1]. However, this method yields a slow rate of deposition and great effort must be used to control deposition to avoid the tendency of the film thickness to taper off as the distance from the electrode contact point increases.

Proceedings of the 2014 ICAM International Conference on Advanced and Agile Manufacturing. Held at Oakland University, Rochester, MI 48309, USA, Copyright © 2014, ISPE USA and ISAM USA

Research conducted by Ghantasala and Sood aimed to electroplate on a selectively resistive (2–20 ohm-cm) silicon wafers without the need for a base film [2]. Their methodology consisted of ion implantation on the face of the silicon to induce selective seeding. In this case 19 keV Pd+ ions are implanted at various dosage levels. Electroplating was preformed to deposit permalloy on the substrate. This would demonstrate that continuous film occurred only in the regions where ion implantation was preformed [2]. The absence of any permalloy deposits in the regions without ion implantation indicate the importance of ions on the substrate surface and their purpose of inducing nucleation sites (eventually evolving into continuous plating). This clearly shows the importance of the nucleation process to occur on a semi-conductive substrate surface and the metallic ions on the surface.

There are numerous factors that alter the deposition and properties of an electroplated material. This research aims to focus on three important factors (material, current density, and duty cycle) with the goal of optimization given a set of parameters. Optimization is defined in this context as an electroplated coating with superior mechanical and electrical properties.

2. Experimental Procedure

2.1 Manufacturing of Nanoengineered Polymer Test Plaques

Elevated temperature extrusion was chosen as the primary method of forming for the samples, due to the ease of manufacturing and also the various output geometries that can be achieved. It was crucial during the forming stage that the parameters were selected such that the dispersion of the metallic inclusions in the polymer (that contribute to its semi-conductive nature) were not adversely effected [3][4].

The samples were produced in two stages. In the first stage the polymeric beads were feed into a 3.175 cm Killion Extruder with a coat-hanger type sheet die shown in Figure 1. The specimen was extruded as a continuous sheet which is 13 centimeters. Every 15 centimeters a cut was made to separate the sample into discrete units.

Figure 1: 3.175 cm Klein Extruder with a Coat-Hanger type Sheet Die

The sample sheet is then transferred to the compression stage, in an effort to provide a consistent surface finish. The sample plaques were formed in a P-20 Steel and Ampco 940 SiNiCrCu compression mold, which were placed into a 222,000 Newton Drake Compression Molding Press. During this compression process the sample is subjected to 133,500 Newtons of force for 60 seconds. The result is a sheet with a thickness of .7 mm. It is observed after the compression stage the ductility and resistance to fracture (compared to non-compression molded samples) had increased.

The surface finish of the sheet, post-compression molding was determined sufficient. However it was desired to explore other manufacturing possibilities in an effort to develop a robust process. Thus a rolling operation was examined to determine its validity as a forming operation, as well as its effect on the surface finish. As such the samples were fed from the extruder directly through a Killion Three-Roll Mill, Figure 2.

Figure 2: Extrudate being fed into the 3-Roll Mill

This process utilizing rolling compression of the sample greatly reduces the sample fabrication time as a continuous sheet is produced without the need for individual, discrete, compression stages. The resulting sheet samples are found to have a similar surface finish and thickness as those samples that were compression molded. To prepare the samples for electroplating 50 mm diameter circles were cut from the polymer sheets to create substrate wafers that could then fit within the wafer electroplating fixture. The samples were cleaned with acetone to remove any contamination from the forming process.

2.2 Electroplating of Nanoengineered Polymer

The electroplating was performed in a clean room environment. As it were desired to only deposit film on one side of the substrate only on a selected area, a special fixture was designed that allows for complete immersion of the substrate while exposing a selected portion of the substrate on the front side (where electroplating is required) to the electrolyte solution. The electrolyte solution was contained in a 1000 mL Pyrex beaker located on a hot plate, as shown in the set-up in Figure 3. A hot plate was used to maintain proper electrolyte temperature as monitored by a thermometer. Also presented in the figure is the anode (nickel or copper plates), as well as the sample jig that encases the nanoengineered polymer substrate.

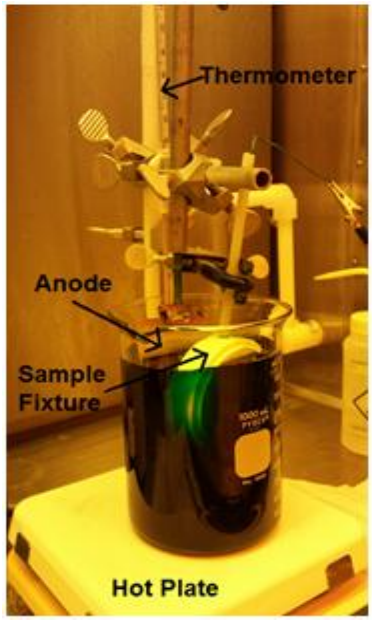


Figure 3: Electroplating Setup

The current source used for electroplating was a Dynatronix MicroStar Pulse Precise Series DuPR10-.5-1.5XR Pulse Power Supply, capable of providing 1.5 A peak current (.5 A continuous) at a maximum of 10 V output. A number of electroplated nickel and copper samples were prepared covering the range of parameters presented below;

- Film Material: Nickel, Copper
- Current Density (mA/cm^2): 5, 10, 15
- Duty Cycle (%): 25, 50 , 75, 100

Current density is calculated based on the area to be plated. The primary effect of current density is the rate at which the material is deposited onto the cathode. Thus for a higher current density, a given amount of deposit is achieved in a shorter period of time. The increase in current density is brought on by an increase in the average current (given a constant plating area). However, this can result in adverse effects on the microstructure of the films.

The duty cycle is the ratio of the ON Pulse time to the ON + OFF Pulse time, with 100% duty cycle corresponding to Direct Current (DC) plating. Pulse plating covers the range of duty cycles from greater than 0% to less than 100%. The result was a periodic interruption of current flow. This interruption allowed for a dispersion of a negatively charged layer that forms around the cathode while current was applied. This layer can inhibit ion mobility, thus decreasing the process efficiency [5]. Early work by W. Kleinekathofer indicated pulse plating (as compared to DC) has a finer micro-structure and a reduced porosity [6]. One of the requirements of pulse plating is the need for higher peak current to maintain a constant cathode current density. For low duty cycles this can produce a large spike in the current, which can lead to “burning” of the deposited material [7]. For example if a 100 mA average current is desired with a 10% duty cycle operating at 1 Hz; the result would be a 1 A current flow for .1 sec followed by a .9 sec pause with no current flow.

2.3 Characterization:

The electroplated copper and nickel films were examined in a Scanning Electron Microscope (Philips XL30-FEG, USA). This system is capable of providing a resolution of 3.5 nm at its maximum voltage of 30 kV. Further, instrumented indentation was employed to measure the mechanical properties of the deposited films using a CSM Nano-Indentation Tester having relevant software.

3. Results and Discussion

3.1 Manufacturing of Nanoengineered Polymer

One of the most crucial aspects of the polymer sheet forming process is to preserve the nature and distribution of the metallic nanoparticles within the host polymer matrix. As the relative distribution of the metallic nanoparticles can greatly affect the electrical conductivity of the polymer, process optimization was performed during this step. Results from the initial forming operations determined that the density of the inclusions within the polymer had decreased after extrusion. This yielded sheets with higher resistance. To remedy this situation it was determined the screen pack within the extruder (which aids in the melting process by increasing the back pressure) had to be removed, as it was interfering with the metallic inclusions. Removal of the pack filter was successful in restoring the extrudates resistance to pre-forming levels.

The manufacturing of the polymer samples into substrates can be achieved through a variety of methods, each with their own set of optimization parameters and results. The elevated temperature extrusion chosen in this analysis is no different. One of the key parameters concerning this forming method is the temperature at which the polymeric beads were melted. Initial tests were conducted at a melt temperature of 200^{o}C. At this temperature the extrudate featured “tearing” at the surface of the polymer, indicating of a melting temperature that is too low. As such the melt temperature was increased to 215 and eventually to 225^{o}C to eliminate the “tearing” effect of the polymer.

Following the extrusion, the compression molding process utilized to form the plaques was optimized to produce an acceptable surface finish. This included a minimum compression time of 60 sec to allow for the polymer to cool. Additionally the compressive pressure was adjusted between 30,000 and 44,000 psi in which it was determined the optimal value is 30,000 psi. This process proved sufficient; however the requirement of the sheets to be cut into discrete units for the press increases the manufacturing time and limits large scale viability.

The rolling operation chosen to replace the press allows for a greater rate of production, as a continuous sheet could be generated with no interruptions to the process flow. Optimization of this operation has been carefully controlled to ensure the material output rate from the extruder matched the speed at which the rollers accept the material. Otherwise stretching and ripping can occur resulting in failure of the material. The rate of the rollers is adjusted to match the extrudate velocity. The extrudate velocity is a function, among other parameters, of the die pressure as well as temperature. It was determined at a die pressure of 22 MPa and temperature of 245^oC the rollers operated at their maximum rotational velocity of 75.1 RPM. To decrease the rotation velocity of the rollers the die pressure is decreased to 11-12.5 MPa at a temperature of 240^oC which resulted in a roller speed of 65 RPM, a stable value.

3.2 Electroplating of Nanoengineered Polymer

SEM images of the copper and nickel electroplated samples prepared at different duty cycles 25%, 50%, 75% and 100% were compared as in Figure 4. To provide for an accurate comparison the current density was kept constant at 10 mA/cm^2. All these samples were electroplated for duration of 90 min each.

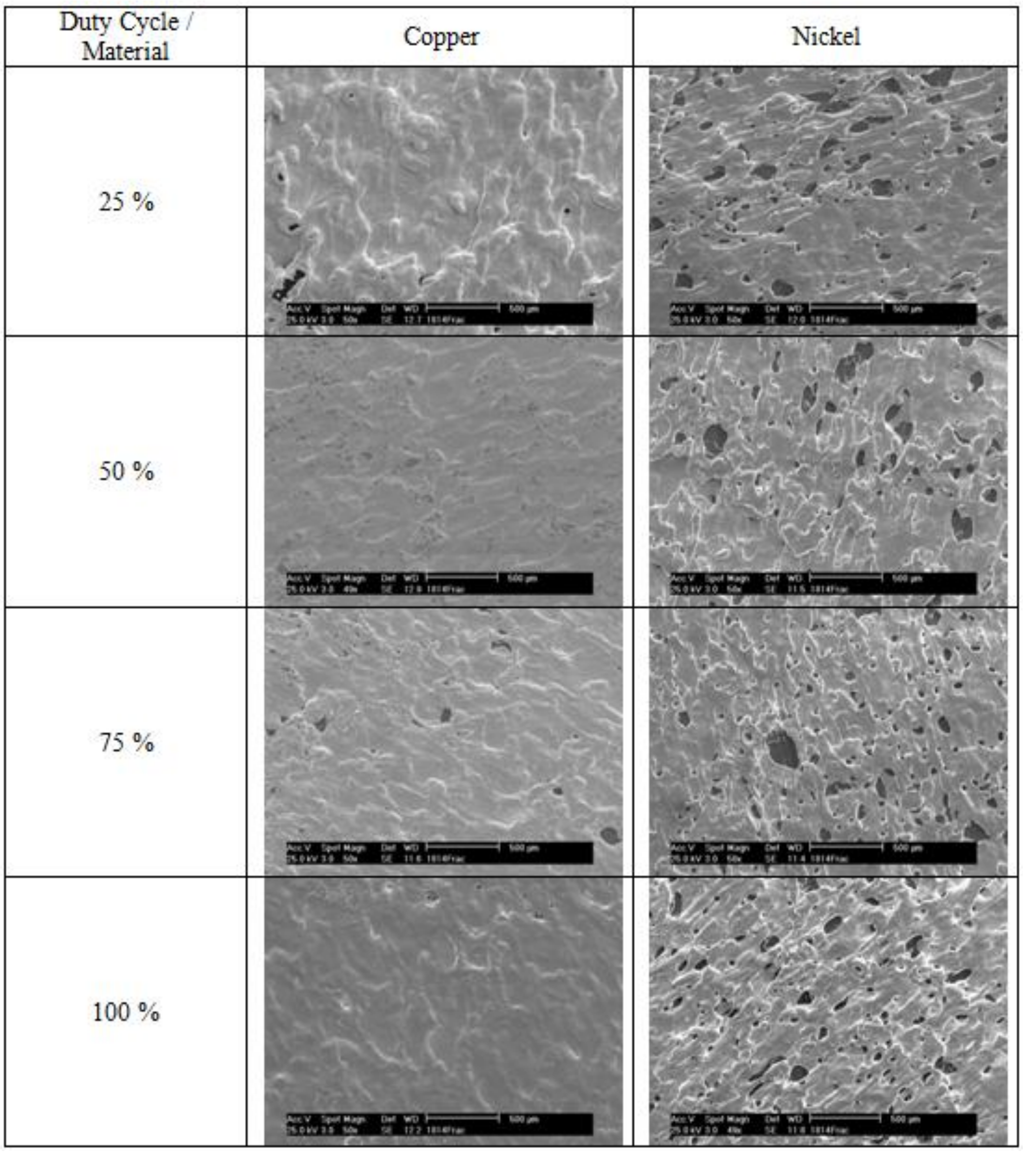


Figure 4: SEM pictures of copper and nickel films deposited at duty cycles 25%, 50%, 75% and 100% respectively (50x)

From the SEM images presented in Figure 4 it is clearly seen that the nickel films appear relatively porous compared to the copper films at all duty cycles. Throughout every duty cycle the copper film presents a greater level of homogeneity with fewer discrepancies. It can also be observed in both cases, the porosity of the deposited films decreased with increasing duty cycle.

In addition to the film topography, the microstructure and size of the individual grains in the deposited Nickel and Copper films were also determined from the SEM analysis, in order to understand the nucleation process. Figure 5 presents the SEM images of these samples at a higher magnification level of 5000X.

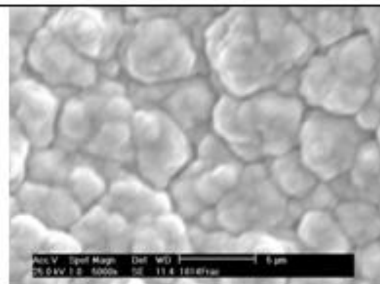
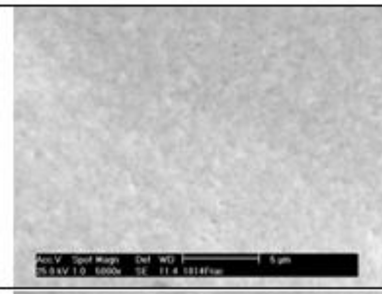
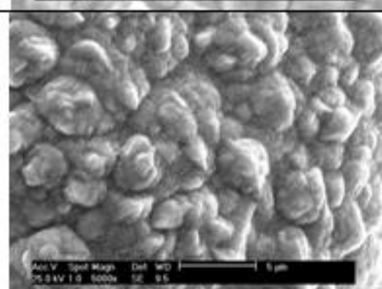
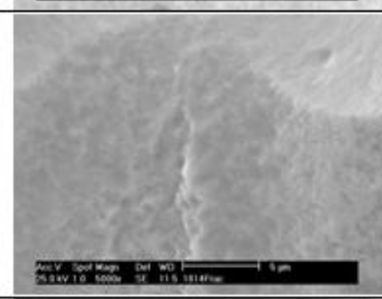
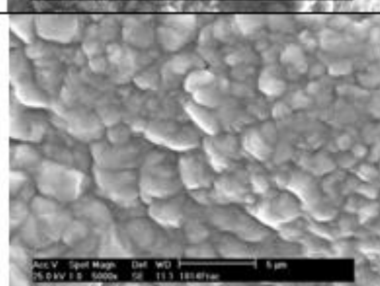
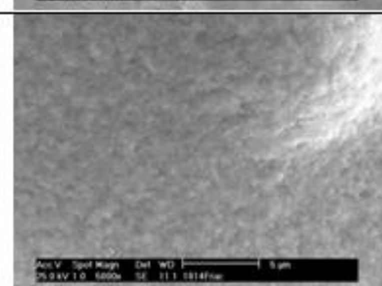
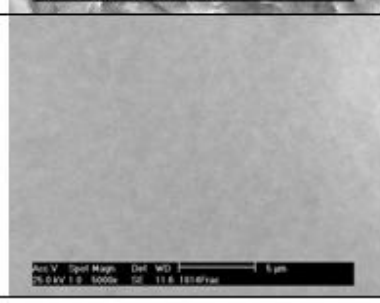
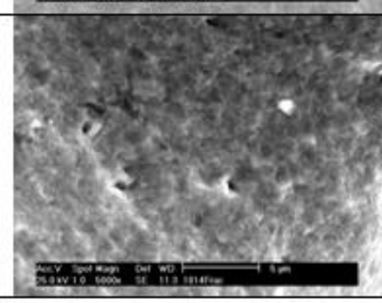
Duty Cycle / Material	Copper	Nickel
25 %		
50 %		
75 %		
100 %		

Figure 5: SEM pictures of copper and nickel films deposited at duty cycles 25%, 50%, 75% and 100% respectively (5000x)

These SEM pictures in Figure 5 clearly showing a large amount of variation in the microstructure and associated grain size between the copper and nickel films. While the copper films showed a distinct granularity, the nickel films appear smooth and quite reflective topography. At a 25% duty cycle the copper film displays an average grain size of the order 3 to 5 µm. The copper films deposited at a duty cycle of 50% exhibited an average grain size of 1 to 3 µm. As the duty cycle further increased to 75%, this trend continues with the grain size decreased to .5 to 2 µm range. Interestingly, relatively smooth sub-µm grains were observed in copper films electroplated using DC constant current (100% duty cycle). In contrast, Nickel films deposited using pulse plating at different duty cycles appears smooth without any granular features. However, the DC plated (100% duty cycle) nickel films seem relatively rough compared to others.

As microstructural variation with grains of different sizes known to influence the physical and mechanical properties of thin films, these changes will be correlated with the film properties. Smaller grain sizes yield significant advantages in terms of wear resistance [8]. Additionally the decrease in grain size leads to an increase in the density of grain boundaries which in turn improves the fracture mechanics of the material.

These films were analyzed using nanoindentation technique also for Young's Modulus and Hardness properties. Correlating the nanoindentation results to the SEM data shows that the relatively larger grain structures of the Copper deposits are related to a higher Young's Modulus and lower Hardness when compared to the Nickel deposits. The Nickel deposits with the smaller, featureless, grains correspond to a relatively higher hardness. Consistent with mechanical theory in which smaller grains (which yields an increase in grain boundary density) have improved mechanical properties such as wear properties, fracture resistance, as well as the observed Hardness. This trend continues within the Copper samples wherein it is found an increase in duty cycle (and thus a decrease in grain size) results in a higher Hardness value.

Examination of these plots indicates that the data conforms well to a linear trend line. It shows that Young's Modulus and hardness are both linear functions of current density as well as duty cycle. This trend appears as a positive slope for both copper and nickel. An increase in duty cycle and/or current density correlates to a direct increase in the samples Young's Modulus and hardness, an important determination in the characterization of the thin metallic films. Additionally, copper exhibits a higher Young's Modulus when compared to nickel for each duty cycle/current density combinations. This was contrary to the bulk properties relationship of the two materials (bulk copper Young's Modulus = 110 GPa, bulk nickel Young's Modulus = 200 GPa)[9], indicative of a scaling factor that alters the materials mechanical properties at the micro-scale. From the SEM results the grain size is roughly 3-5 microns. Thus the grain size of bulk copper can be assumed to be in this range. The bulk value for the Young's Modulus of Nickel was not observed during the course of this research. The nickel films in general exhibits a higher hardness when compared to the copper films, consistent with the relationship of the bulk material mechanical properties.

Conclusions

The manufacturing of the polymer resin from beads into test plaques was demonstrated in this paper using a combination of forming methods. The chosen process for this research involved elevated temperature extrusion coupled initially with compression molding, later replaced by a rolling operation. It was determined both post extrusion operations resulted in a similar surface finish of the sample; however the increase in throughput of the rolling operation was a clear advantage. Optimization efforts of the manufacturing process resulted in samples with unaffected electrical conductance properties in a choice of geometries. Copper and Nickel electroplated samples were prepared at different duty cycles and current densities. The analysis of the copper and nickel samples showed that copper films were much more granular with the grain size varying in the range 0.5 µm to 5 µm, whereas Nickel films were relatively smooth. It was determined that within a given material, mechanical properties rely on the plating current density and duty cycle. An increase in either factor results in improvement of mechanical properties. With respect to the individual material results, copper was found to have a higher Young's Modulus (for equivalent deposition parameters) while nickel exhibited a higher Hardness (for equivalent deposition parameters).

Bibliography

[1] V. Fleury, W. Watters, L. Allam, and T. Devers, "Rapid electroplating of insulators."

[2] M. Ghantasala and D. Sood, "Novel method for electroplating on silicon without the need of a continuous-plating base film," in *Far East and Pacific Rim Symposium on Smart Materials, Structures, and MEMS*. International Society for Optics and Photonics, 1997, pp. 52–58.

[3] M. Müller, B. Krause, B. Kretzschmar, and P. Pötschke, "Influence of feeding conditions in twin screw extrusion of pp/mwcnt composites on electrical and mechanical properties," *Composites Science and Technology*, vol. 71, no. 13, pp. 1535–1542, 2011.

[4] B. Krause, P. Pötschke, and L. Häußler, "Influence of small scale melt mixing conditions on electrical resistivity of carbon nanotube-polyamide composites," *Composites Science and Technology*, vol. 69, no. 10, pp. 1505–1515, 2009.

[5] M. Chandrasekar and M. Pushpavanam, "Pulse and pulse reverse plating conceptual, advantages and applications," *Electrochimica Acta*, vol. 53, no. 8, pp. 3313–3322, 2008.

[6] W. Kleinekathöfer and C. Raub, "Die abscheidung von nickel mit pulsierendem strom," *Surface Technology*, vol. 7, no. 1, pp. 23–34, 1978.

[7] A. Sonin, "Jet impingement systems for electroplating applications: mass transfer correlations," *Journal of The Electrochemical Society*, vol. 130, no. 7, pp. 1501–1505, 1983.

[8] D. Jeong, F. Gonzalez, G. Palumbo, K. Aust, and U. Erb, "The effect of grain size on the wear properties of electrodeposited nanocrystalline nickel coatings," *Scripta Materialia*, vol. 44, no. 3, pp. 493–499, 2001.

[9] William D. Callister, Jr., Materials Science and Engineering: An introduction, John Wiley & Sons, USA, 7th edition, p.137, 2007

Intuitive 3D-Vision Based Wand for Robot Tool Path Teaching

Khalid Mirza[1], Sai Prasanna, Michael Truitt and Hudhaifa Jasim **Ying Wang[2]**

[1]Department of Electrical and Computer Engineering
Oakland University
Rochester, Michigan 48309
Email: mirza@oakland.edu

[2]Department of Electrical & Computer Engineering
Duke University
Durham, NC 27708

Abstract

This paper presents a new intuitive method for robot tool path teaching using vision based teaching wand. Manual teaching of robots for frequently changing applications is an issue when using traditional robot teach pendants. For someone with minimal robot programming skills, using a robot teach pendant to teach an application requires a steep learning curve and is usually a slow process. This is one reason why some industries who could potentially use industrial robots do not use them because they have applications that change frequently, they lack in-house robot programming skills, or they cannot afford expensive external resources to assist in programming robot applications. Our approach is a quick and intuitive way to control the robot tool position in the 3D space without sacrificing accuracy. Just as a joystick is a very natural way to control motion in 2D, our method extends this approach to the three dimensions for robot position control. Using an illuminated teaching wand and a single IR camera, we track the offsets in 3D to generate velocity commands for the robot. A flexible velocity calibration scheme is implemented that does not restricts the user to start the teaching wand at any specific location in camera view to generate robot motion. Transformation of human hand motion into smooth robot motion is achieved by removal of teaching wand jitter and additional convenience features like axis constrained motion help realize robot teach pendant like control.

Keywords — Human-robot interaction, Intuitive robot teaching, Vision-based tracking.

1 Introduction

Industrial robot has proven itself to be a key component for advanced manufacturing. Not only it can dramatically improve the quality of products being manufactured, it also has the potential to increase the production throughput speeds and improve the overall financial savings over the long term. Yet many of the small to medium industries, who could potentially use industrial robots to their advantage, are still hesitant to employ robots in their production lines. The reasons can vary from the high initial investment for setting up a robot based manufacturing cell to the specialized skills needed to program and run the robots. Nevertheless, one of the main reasons that robots are not as extensively used is due to the fact that it is not easy to teach robot applications.

Programming an industrial robot application can be done manually by directly using the robot or off-line in a simulation package using CAD models. Off-line programming (OLP) for robots is similar to how CAD systems are used to generate the programs for CNC machines. Using CAD models for the robots, fixtures and the different components, a program structure is built up which is similar to manual teach programming. Intelligent tools use CAD data to generate the sequences of position and process information. The advantage of this method is the significant saving of time and money specially when designing complex work cells. Other benefits include the ability to analyze the behavior of the work cell before actually implementing it.

Even though OLP method is widely used for programming robots for large and complicated work cells, it has its limitation due to the fact that complete geometrical description of robot working area and CAD model for the workpiece must be known in advance. Various calibrations also need to be performed to correlate the off-line programming to the actual implementation in the work cell, and any deviation of the CAD model from the real world will additionally require the robot path to be corrected on the plant floor. These limitations pose problems for small to medium sized industries who have to deal with batches of workpieces that are manufactured in low volumes and change quite frequently. It usually becomes cost prohibitive to use offline programming to handle such manufacturing jobs employing OLP.

Manual robot teaching methods traditionally use a teach pendant [1] [2] [3] as human robot interaction device. However, using a teach pendant can be awkward, inefficient and time-consuming to generate robot programs on the manufacturing floor. A robot needs to be manually jogged to all the target positions, and position and orientation of the robot end effector are then recorded to program the path for robot motion. Controlling an industrial robot with six degrees of freedom is not an intuitive task when using a teach pendant that typically can jog the robot in only single direction at a time and the programmer may have to switch between different coordinate systems to effectively move the robot end effector to the desired target position. The manual jogging of robot is also unattractive from the ergonomic point of view [4]. It is an acquired skill that comes with long experience and only then manual teaching can be performed with some efficiency to program robot tasks.

To overcome the shortcomings of OLP, some methods have been proposed to enhance it by using Augmented Reality (AR) that overlays computer generated graphics onto the real world images so that both can be perceived at the same time. The advantage of this approach is that the operator can visualize the robot trajectories in the real environment and robot teaching methods do not require workpiece to be modeled in CAD or its localization in the workcell. AR has been used in a variety of ways for industrial robot teaching also, including using marker-based optical tracking systems [5] [6] [7]. This approach suffers from having significant error when determining the depth information of the marker and can also stop working when the marker gets occluded. Other approaches include the use of laser range finder to get better depth accuracy [8].

Different types of manual teaching devices have also been a topic of research to facilitate ease of programming robots. Researchers attached accelerometers to human arms to let one arm control the position and the other control the orientation of the robotic arm [9] and others recorded various joint movements using motion sensors which were then applied for robot arm control [10]. These approaches solved the intuitive control of robot arm problem but were limited in positioning and orienting the robot end effector accurately. Another teaching system using a teach pen with optical markers that were tracked by a motion capture system was proposed to improve the positioning accuracy [11]. However, this accuracy was dependent on the position estimation of the pen tip and the experiments that were performed required the motion capture platform to be separate from the robot platform.

Joysticks and haptic devices have also been used to control the robots. A force/torque sensor attached to the robot end effector was used with a joystick to control the forces in a polishing application was presented in [12]. A 6 DOF force-feedback joystick was used to accurately control the robot and conduct delicate tasks such as surgery in [13]. However, this application was limited to small workspaces and cannot be used in industrial environments. Another technique based on Wii Remote Controller was used to capture human hand gestures for controlling an industrial robot in [14]. These methods addressed the ease of manually teaching a robot but were limited in one way or another for use in an industrial environment.

A mechanical joystick can be a very intuitive way to control a mechanical device in two dimensions. There are numerous applications that use joysticks in such a way that does not require too much effort for one to become skilled. However, for controlling or positioning a device in three dimensions, the intuitive nature of typical mechanical joysticks falls apart, as it becomes very hard to manually control simultaneous motion in all three dimensions both smoothly and accurately.

The vision based teaching wand approach presented in this paper demonstrates an innovative and robust way to control the position of the robot tool tip. This method works in the same way as a 2D joystick controller but extends the intuitive control in three dimensions. A single IR camera tracks the position of a teaching wand and derives its position in three dimensions. However, the wand does not need to be moved to any specific location in the field of view of the camera to start robot motion. Instead, an operating workzone is automatically created in 3D around any starting position of the teaching wand. Moving the teaching wand in any direction then generates the corresponding motion offsets for the robot that mirrors the wand movement. While moving the wand, the user is monitoring the tool tip to accurately position it where it is required for teaching the point. Just like using a mechanical joystick, the velocity of the robot motion is controlled by how far the teaching wand is moved from its starting position. This results in an intuitive way to to quickly and precisely control the velocity of the tool tip. The method mimics the robot teach pendant like control to manually teach positions but gives the added advantage to move along any 3D vector in world coordinates and have full control over the velocity of the motion while moving.

2 System Overview

2.1 System Setup

The teaching system setup consists of a teaching wand and an IR camera connected to a vision PC as shown in Figure 1. The position of the teaching wand is tracked using the IR camera and the image data is communicated to the vision PC for processing. The motion enable and mode information from the teaching wand is also sent to the vision PC which then computes the 3D motion offsets in Robot World Coordinates. The offset data is provided to the robot controller to control the corresponding velocity of the tool tip in Robot World Coordinates.

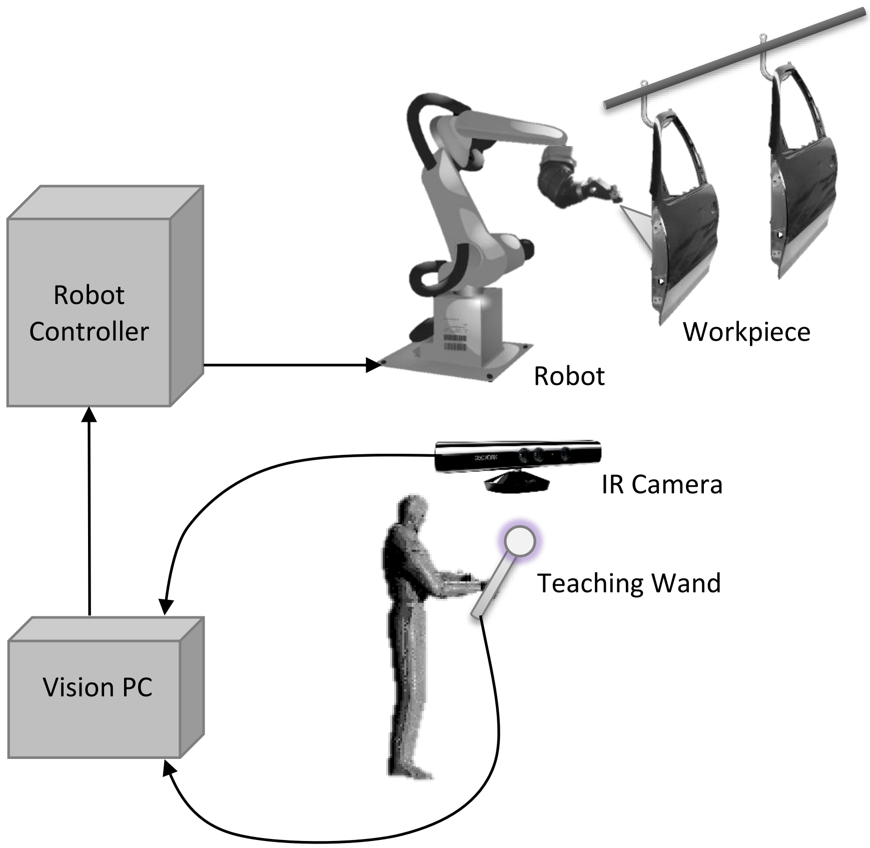


Figure 1: Teaching system setup

2.2 Teaching Wand Operation

2.2.1 Position tracking

The teaching wand has an IR Illuminated Sphere (IRIS) that can be seen by the IR camera as a brightly lit circle from any direction. The vision processing algorithm derives the centroid of the IRIS circle as the position of the teaching wand in Camera Spatial Coordinates. The x and y position of the IRIS in Camera Spatial Coordinates maps directly to its 2D image and is derived from the x and y in Image Pixel Coordinates. The radius of the IRIS circle is known and its relative size seen in the Image Pixel Coordinates at any distance from the IR camera can be calibrated to correspond to its z position in Camera Spatial Coordinates. Hence, the 3D (x, y, z) position of the teaching wand can be tracked in the Camera Spatial Coordinates with just using a single IR camera. Aligning the Camera Spatial Coordinates with the Robot World Coordinates makes the robot motion to directly correspond to the teaching wand movements, hence resulting in a very intuitive way to control the robot tool tip position.

The precision of position tracking depends on the image resolution and its quality based on the lighting conditions. If the teaching wand was being used to directly generate position commands for the robot tool tip, then the entire system needed to be highly accurate, including tracking of the *z*-depth of IRIS and calibration of the IR camera coordinates to the robot coordinate system. However, using the teaching wand as a 3D visual joystick to generate velocity offsets for the robot tool tip, the IR camera only needs to be approximately calibrated with the Robot World Coordinates. The positional accuracy of the teaching wand comes from the operator monitoring the placement of the robot tool tip just as in the case if a robot teach pendant was being used.

2.2.2 Self-Calibration

In order for the teaching wand operation to be intuitive, it must be able to control robot motion at any start position within the IR camera field of view and then deriving distance offsets relative to that start point. This relinquishes the need to locate the teaching wand at one fixed start point in the 3D space for generating robot motion offsets. This is not only a useful feature but also a quite necessary one as it will otherwise be extremely difficult for an operator to precisely position the teaching wand at a fixed start point in 3D space without any guides.

Moving the teaching wand in the view of the IR camera and pressing a button to enable motion sets the start point at that location in 3D space as shown in Figure 2. The start point locates the origin of the motion offset frame which is aligned with the Camera Spatial Coordinates. Any subsequent motion of the teaching wand now generates the motion offset in that direction relative to the start point.

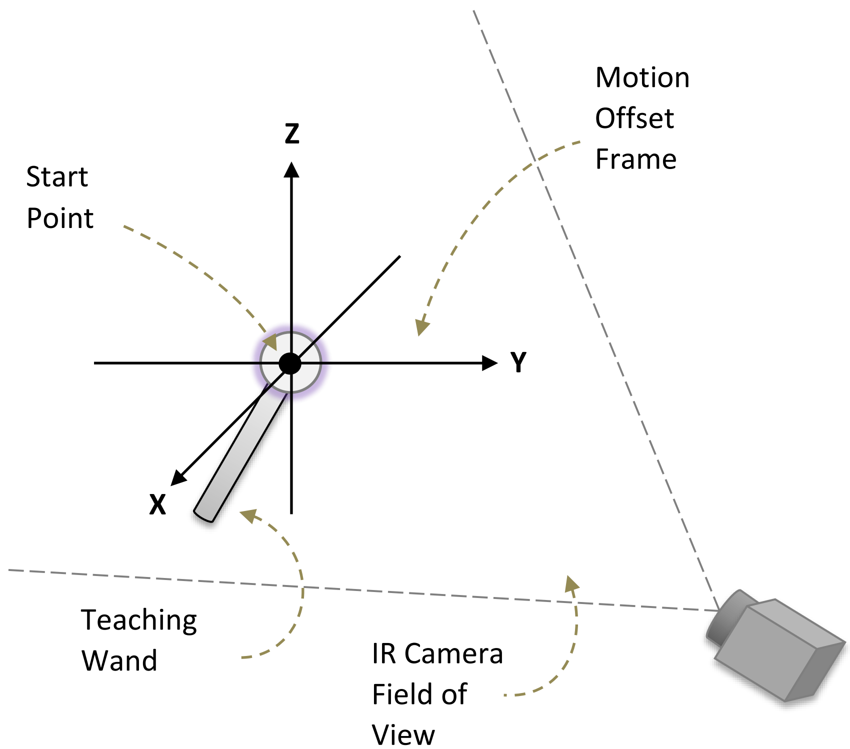


Figure 2: Start point and motion offset frame

2.2.3 Deadzone

After enabling the robot motion, trying to hold the teaching wand steady at the start position can be a challenge and it will have an inherent jitter associated with its tracked position. This slight variations in position should not cause any robot motion and hence need to be suppressed. Figure 3 shows a small deadzone allocated around the start point. Any teaching wand motion within that region does not result in motion offsets. Only offsetting the teaching wand from the start position in the direction of intended robot motion will cause it to move outside the deadzone that in turn will generate a smooth robot motion offset.

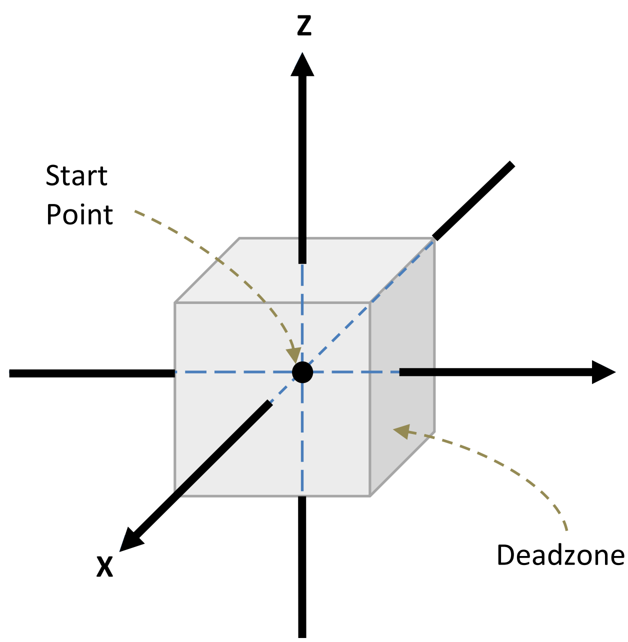


Figure 3: Start point and deadzone

2.2.4 Pure Axis Motion Zones

One of the convenient features available on a robot teach pendant is the ability to jog the robot along any one of the (x, y, z) axes in the Robot World Coordinates. Attempting such a pure axis motion using the visual teaching wand will require extraordinary steady hands. To facilitate this type of motion, pure axis motion zones are created along the Camera Spatial Coordinates as shown in Figure 4. As long as the teaching wand stays within the pure axis zone, motion offsets are generated only along that particular axis ignoring any jitter along the other two axes. Since the Camera Spatial Coordinates are aligned with the Robot World Coordinates, this results in a robot motion along any one of the (x, y, z) axes in the Robot World Coordinates.

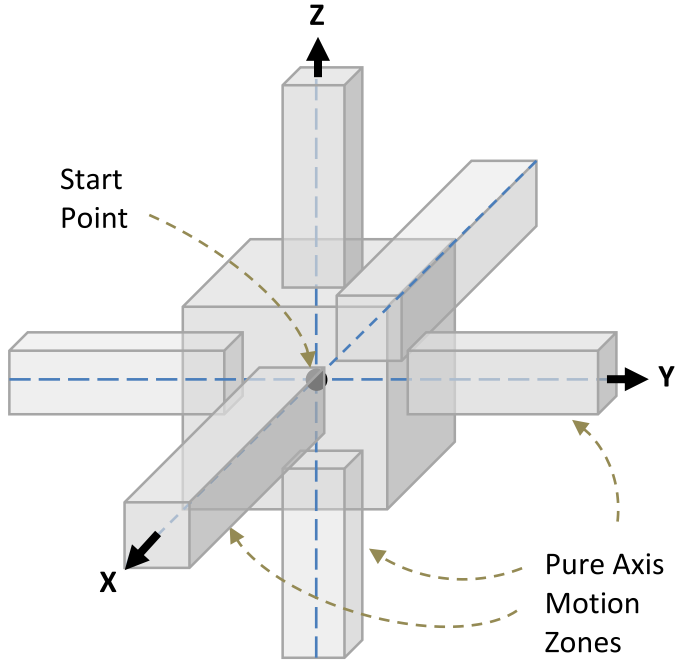


Figure 4: Start point and pure axis motion zones

2.2.5 Motion Offset Generation

Combining the effects of the deadzone, pure axis motion zones and the robot motion enable, the flowchart for a safe way to generate the robot motion offsets is shown in Figure 5.

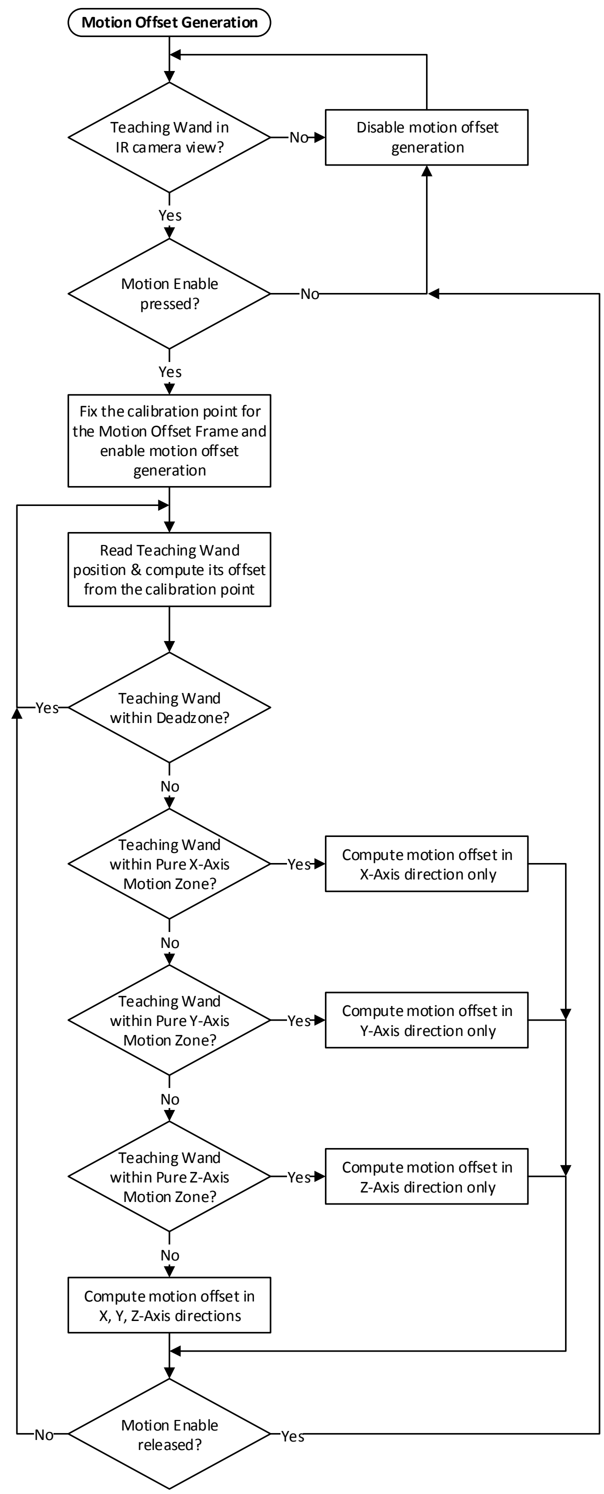


Figure 5: Motion offset generation flowchart

2.3 Hardware

2.3.1 Teaching Wand

The teaching wand is the user interface to control the tool tip position of the robot. Its operation is similar to a visual joystick that is used to generate velocity offsets in three dimensions. The main components of the teaching wand are:

a. IR Illuminated Sphere (IRIS)

Infrared (IR) light is the electromagnetic radiation with wavelengths longer than those of visible light. As a result, infrared light is used in industrial, scientific, and medical applications to illuminate and track objects without interference from visible light. IR camera can also be used to capture an image in darkness as long as an IR source is provided. An IR camera which is sensitive between the spectrum of 3-5μm and 8-14μm wavelengths will works well in most of the lighting conditions.

On the teaching wand, the IR Illuminated Sphere (IRIS) consists of IR LEDs enclosed in a translucent sphere that uniformly distributes the illumination of the LEDs. The result is that images returned from the IR camera shows IRIS as a perfect circle from any view angle. This is important, since the diameter of the captured circle is used to measure the depth position of the teaching wand. The size of IRIS also dictates how large of an operating work zone can be realized for the teaching wand. It is desirable that the IRIS covers a large enough area in the image for the vision algorithms to accurately derive its centroid location.

b. Inertial Measurement Unit (IMU)

An inertial measurement unit is utilized to measure the orientation of the teaching wand relative to the Camera Spatial Coordinates. This tilt in the teaching wand orientation will be used to control the orientation of the robot tool tip. A microcontroller interprets the output of the sensor into yaw, pitch, and roll Euler angles. The use of Euler angles fits the application nicely, since they can be used to represent the orientation relative to a reference or frame.

c. Microcontroller

An Arduino Micro is used as a microcontroller to interface the peripherals attached to the teaching wand and the PC which is running the vision algorithms. The microcontroller reads the inputs, processes the data and eventually sends messages to the vision PC over the serial ports.

d. Control Buttons

There are several buttons on the teaching wand that control its operation:

- Power – to turn on/off the IR LEDs
- Motion Enable – for self-calibration and enable generation of motion offsets
- Mode – to switch between position and orientation control

- Emergency Stop – safety for the handheld teaching device to disable robot motion

Some of the prototypes for the teaching wand are shown in Figure 6.

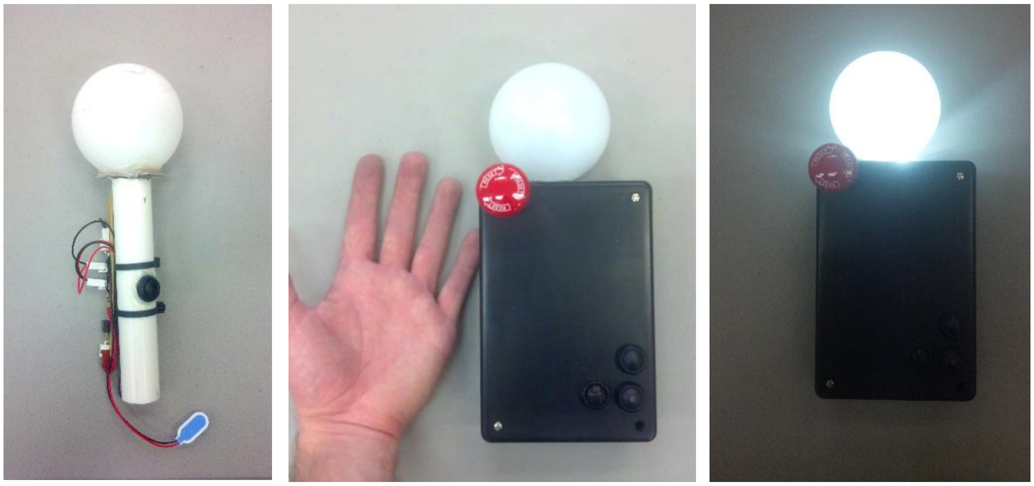


Figure 6: Teaching wand prototypes

2.3.2 IR Camera

Microsoft Kinect for Xbox 360 is a readily available device with multiple vision sensors that has excellent capabilities for its price. Kinect not only includes an RGB sensor but also a depth sensor that works in the IR range, which can be used for tracking the IRIS on the teaching wand. The main specifications for the Kinect IR sensor are given in Table 1.

Property	*Specifications*
Sensor type	CMOS (1600×1200 resolution)
Field of View	58° Horizontal, 40° Vertical, 70° Diagonal
Spatial x/y resolution (@2m distance from sensor)	3mm
Operating range	0.8m – 3.5m

Table 1: Specifications for the Kinect depth sensor

Since the Kinect IR sensor is utilized to capture the position of the IRIS instead of using the RGB sensor, lighting is not a major point of concern. This is very useful in factory environments where control on ambient lighting is always an issue. On the other hand, such industrial manufacturing environments typically may not have strong IR sources. Since the IRIS emits its own IR light, it is presumed that the image seen by the IR sensor does not contain any other sources of strong IR light or objects that are radiating high levels of heat. This will ensure that the vision algorithm tracks the IRIS properly and is not confused by other IR sources.

3 System Implementation

The main goal of the new teaching method presented in this paper is to implement generation of smooth robot velocity offsets corresponding to IRIS movement in the Camera Spatial Coordinates. We are not interested in the absolute location of IRIS in Robot World Coordinates but rather we need to measure the relative movement of IRIS from the start point when motion was enabled. The implementation of this method is detailed in the following sections.

3.1 Image Processing

Most existing object tracking methods use visible light intensity and/or color to track objects [15]. With these methods there are inherent problems with the interference of ambient light source in the tracking environment. IR light source and cameras are a good way to eliminate ambient light interference issues. The IR image captured by the Kinect IR sensor is an RGB image with IR content. As we are only interested in locating the position of the IRIS, we can ignore other properties like its color. RGB images are computationally expensive for any real-time computing or target tracking based applications. Hence, the image should first be converted into grayscale before any further processing. RGB to grayscale is done using the method mentioned in [16] [17] as given below:

$$I_{gray} = \left(0.2989 * I_r + 0.5870 * I_g + 0.1140 * I_b\right) \tag{1}$$

where I_{gray} is the grayscale image, and $(I_r\ , I_g\ , I_b)$ are the RGB components of the image.

3.2 Image Filtering

Even though Kinect IR sensor is suitable for most of the lighting conditions, it suffers from an inherent noise issue due to low signal-to-noise ratios. It is always desirable to perform noise reduction before processing the image to derive position information. The ideal filtering technique for our purpose would reduce noise while preserving the edges so minimum distortion is introduced in the circular shape of the IRIS image. A standard median filter [18] is chosen as it is best suited to reduce the noise in the IR image [19] while preserving the edges.

3.3 Image Thresholding

Even though we are using an IR sensor, the presence of ambient light can have negative effects on detecting and tracking the IRIS. In order to avoid any false positive detection of the IRIS, another level of thresholding is introduced. The filtered image is converted into a binary image based on a threshold value t_{bw} [17]. Since the IR light source will be much brighter than other light sources, a 70% fixed threshold was chosen for the grayscale image to reject all non IR light sources present in the image. The illumination level is calculated as follows:

$$t_{bw} = 0.70 * 255 \tag{2}$$

$$I_{bw} = \begin{cases} 0 & \text{for } I_{gray} \le t_{bw} \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

where I_{bw} is the resulting binary image.

Most of the speckle and high frequency noises are filtered out by the above filtering methods. However, there might still be some leftover high frequency noise content that will need to be further removed. This can be accomplished by using a morphological filter [20]. The binary image is first eroded as given by:

$$A \ominus B = \bigcap_{b \in B} A_{-b} \tag{4}$$

where A is binary image and B is the erosion mask (disk or square). This is then followed by a dilation operation on the image as given by:

$$A \oplus B = \bigcup_{b \in B} A_b \tag{5}$$

3.4 Object Detection

A number of techniques have been proposed for object detection from 2D images. With the noise filtered out from the image, we chose to detect IRIS by using a size invariant circular Hough transform [21] [22]. Hough transform based methods for object detection work by allowing image features to vote for the location of the object. Circular Hough transform returns the radius of the IRIS circle r^p and its centroid $\boldsymbol{c^p} = (c_x^p, c_y^p)$ in pixel coordinates.

3.5 Operating Range

As IRIS tracking is an essential part of our operation, useful information cannot be extracted from the image if IRIS is too close or too far off from the camera. Hence, the operating range for the teaching work zone needs to be calculated for optimal use.

If IRIS is too close to the camera, its image will fill most of the view frame. Any movement of IRIS may then cause the detected image to leave the camera view, which in turn results in a failure to generate motion offsets. The size of the IRIS relative to the image size in pixels can be used to determine the closest operating range of the teaching wand. For the closest operating range, a clearance equal to the diameter of the IRIS in pixels was empirically selected to provide sufficient clearance around IRIS to detect its movement within the camera view. The resulting maximum diameter D_{max} of IRIS in pixels that is allowed for operation can then be calculated as:

$$D_{max} = \frac{\min(I_{\text{height}}, I_{width})}{3} \tag{6}$$

where I_{height} and I_{width} are the height and width of the image in pixels.

If IRIS is too far away from the camera, its pixel size in the image becomes too small for it to have enough resolution to accurately determine its diameter for computing the z-depth. The minimum diameter of IRIS in view was empirically selected to be equal to 10% of the image size before it is considered too small for accurate z-depth offset generation. The minimum diameter D_{min} of IRIS in pixels can then be calculated as:

$$D_{min} = 10\% \frac{\min(I_{\text{height}}, I_{width})}{3} \tag{7}$$

This above analysis implies that diameter of the detected IRIS in pixels should be between (D_{max}, D_{min}) for best operation.

3.6 Jitter Filtering

Tracking a hand held teaching wand in 3D will always have some jitter content associated with it, depending on how steady or unsteady is the operator hand motion. Generating motion offsets with jitter in it will result in an undesirable jittery motion of the tool tip position. Hence, the jitter content in the motion offsets needs to be filtered out in order to produce a smooth robot motion. Since the jitter is mainly high frequency content in the motion data, a weighted moving average filter [18] is implemented to filter it out:

$$J_{x,y,z} = \frac{1}{n} \sum_{k=0}^{n} w_i * \begin{bmatrix} c_x^p \\ c_y^p \\ r^p \end{bmatrix} \tag{8}$$

where w_i is the weight applied to the filter $(w_i = 1)$.

3.7 Z-Depth Calculation

In order for the teaching wand to be intuitive to use, robot velocity offsets need to be mapped directly to IRIS movement in 3D Camera Spatial Coordinates. A change in IRIS position in spatial (x, y) coordinates produces a corresponding change in the detected position of the IRIS in the (x, y) pixel coordinates. This is a linear change, since a change in the Camera Spatial Coordinates is proportional to the corresponding change in the image pixel coordinates.

Similarly, a change in IRIS position along the z axis in the Camera Spatial Coordinates produces a corresponding change in the size of the radius of the detected IRIS in pixel coordinates. This change is non-linear and z depth cannot be linearly derived from the change in radius size in pixels as we did for the (x, y) axes. Instead, the relationship is derived by experimentally measuring and mapping the size of the IRIS radius r^p (pixels) in image to its position S_z in Camera Spatial Coordinates. Table 2 shows measured distance between the IRIS and the IR sensor and the detected IRIS radius in pixels:

S_z	r^p
5	77.995
10	41
15	28
20	21.0638
25	17

30	14.57
35	12
40	10.7107
45	9.2795
50	9
55	8
60	6.6835

Table 2: Mapping between spatial z-depth of the IRIS to its radius in pixels

A two term exponential curve fitting technique is used to fit the measured data. The general equation of a two term exponential is given by:

$$S_z = ae^{bJ_z} + ce^{dJ_z} \qquad (9)$$

Where the coefficients a, b, c, d need to be determined by curve fitting. The computed curve fitting coefficients are given in Table 3 and the result is shown in Figure 7.

Coefficient	*Value (with 95% confidence bounds)*
a	$146\ \ (133.2, 158.9)$
b	$-0.2142\ \ (-0.2404, -0.1881)$
c	$31.8\ \ (27.12, 36.48)$
d	$-0.02641\ \ (-0.03027, -0.02255)$

Table 3: Curve fitting coefficients computation results

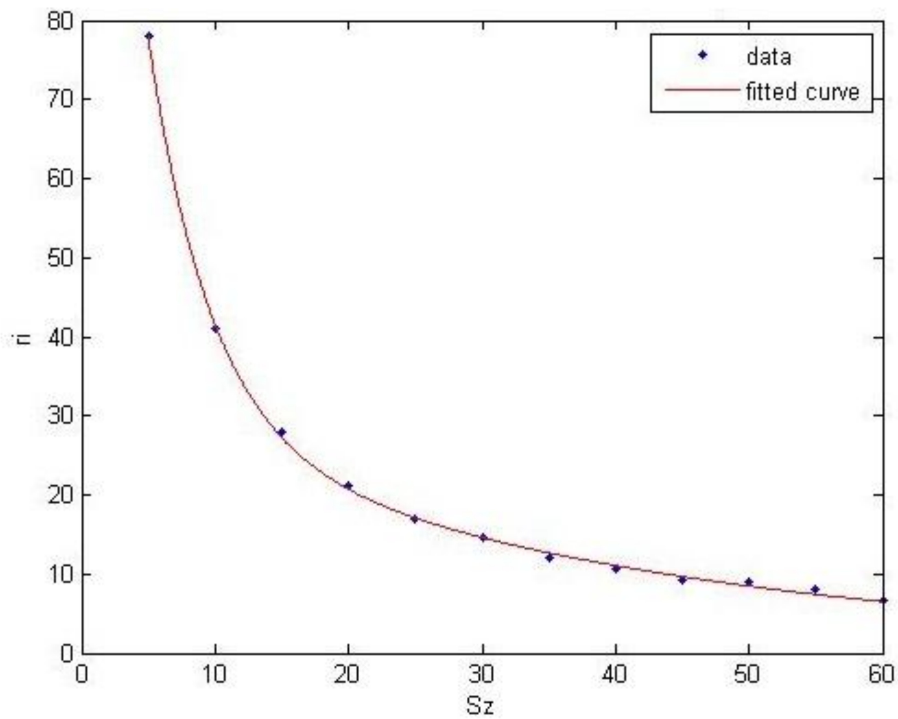


Figure 7: IRIS distance from IR camera vs. IRIS radius in pixels

3.8 Spatial Coordinate Conversion

So far we have the IRIS offsets from the start position in the (x, y) directions in the image pixel coordinates and the IRIS offset in the z direction is in the Camera Spatial Coordinates. Before the (x, y, z) offsets can be used for generating velocity commands, they need to be converted to and normalized in the Camera Spatial Coordinates, otherwise the scaling of motion in the (x, y) directions will vary based on the z-depth S_z of the IRIS. The filtered (x, y) location of the IRIS in pixel coordinates is given by $\left(J_x, J_y\right)$. Using the perspective transform, the corresponding IRIS location in Camera Spatial Coordinates $\left(S_x, S_y\right)$ can be computed as:

$$S_x = S_z * \frac{J_x}{f} \quad \text{and} \quad S_y = S_z * \frac{J_y}{f} \tag{10}$$

where f is the focal length of the Kinect IR sensor.

3.9 Offset Generation

Ease of setup is a major concern when trying to use the teaching wand in an industrial environment. User should be able to pick up the teaching wand and start programming the robot just as easily as when using a robot teach pendant. This implies that we should be able to start the teaching process once the wand is inside the camera frame, irrespective of its distance from the center of the frame. This method performs a self-calibration when robot motion is enabled by setting the initial detected location of the wand as the start point, where the origin of the spatial coordinate system for relative measurements is located. Hence, the initial detected location **O** of IRIS in the Camera Spatial Coordinates, which is now also the reference start point for relative measurements, is given by:

$$\mathbf{O} = \begin{bmatrix} S_x \\ S_y \\ S_z \end{bmatrix} \tag{11}$$

Moving the teaching wand away from the start position starts the generation of distance offsets. The distance **D** of the new position of the wand from its start point in Camera Spatial Coordinates is given by:

$$\mathbf{D} = \begin{bmatrix} D_x \\ D_y \\ D_z \end{bmatrix} = \mathbf{S} - \mathbf{O} \tag{12}$$

where **S** is the new position of the IRIS in the Camera Spatial Coordinates.

3.10 Deadzone, Pure Axis Motion and Boundary Condition Generation

Deadzone is an operational zone around the start point within the workspace of the teaching wand where motion offsets are suppressed even when there is a slight wand movement. As hand held motions are prone to jitter, any jitter around the start point of the teaching wand in Camera Spatial Coordinates should not cause any corresponding robot motion.

Additionally, traditional robot teach pendants provide manual motion keys that can be used to jog the robot just along a selected axis in Robot World Coordinates. This pure axis motion can be very useful when manually jogging the robot while teaching applications. For the teaching wand, pure axis motion computation is comprised of generating motion along any single axis in Camera Spatial Coordinates, which is aligned with the Robot World Coordinates, while rejecting any small deviations along the other two axes.

Both the deadzone and pure axis motion zones can be implemented by creating an operational zone along each of the (x, y, z) axes in Camera Spatial Coordinates, as shown by pure axis motion zones in Figure 4. A distance offset within these zones allow motion only along axis in the direction of the zone while suppressing generation of motion offsets along the other two axes.

This behavior can be implemented by using a simple band reject filter [23] as long as the deviations along the other two axes remain within a certain bound. Once these deviations cross the boundary conditions for the pure axis motion zone, the distance offsets can then be calculated from the moving average filtered value as described in the previous sections. Based on the ideal band reject filter, the filtered distance offset **H** in the Camera Spatial Coordinates that implements the operating zones is given by:

$$H_x = \begin{cases} 0 & \text{if} \quad -w \le D_x \le w \\ D_x \pm w & \text{otherwise} \end{cases} \tag{13}$$

$$H_y = \begin{cases} 0 & \text{if} \quad -w \le D_y \le w \\ D_y \pm w & \text{otherwise} \end{cases} \tag{14}$$

$$H_z = \begin{cases} 0 & \text{if} \quad -w \le D_z \le w \\ D_z \pm w & \text{otherwise} \end{cases} \tag{15}$$

$$\mathbf{H} = \begin{bmatrix} H_x \\ H_y \\ H_z \end{bmatrix} \tag{16}$$

where w is the width of the band reject filter measured in Camera Spatial Coordinates and it defines the bounding box for the pure axis motion zones.

3.11 Normalized Velocity Offsets

Robot motion is generated using the smoothed velocity offset which are created by the filtered distance offset **H** in the Camera Spatial Coordinates. Since we are using the teaching wand as a visual joystick, the distance offsets directly maps into the velocity offset in the Camera Spatial Coordinates. With the Camera Spatial Coordinates aligned with the Robot World Coordinates, these velocity offsets can be used to move the robot tool tip relative to the Robot World Coordinates. However, before the velocity offsets can be sent to the robot controller for generating motion, they need to be normalized. To constrain the velocity offsets between (v_{max}, v_{min}) when filtered distance offsets are bounded between (H_{max}, H_{min}), a simple normalization can be performed to obtain the required velocity offset using:

$$\mathbf{v} = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} = \frac{(\mathbf{H} - H_{min})}{(H_{max} - H_{min})}(v_{max} - v_{min}) + v_{min} \quad (17)$$

These velocity offsets can now be provided to the robot controller to jog the robot.

4 Experimental Results

In the experimental setup, the teaching wand was moved within its operational zone and was tracked by the Kinect IR sensor. The vision PC used MATLAB [24] to communicate with the Kinect sensor and run vision processing algorithms to derive the location of IRIS. The tracked data was recorder and processed to derive the corresponding velocity offsets for controlling the robot.

Experiment 1 demonstrates the results for a pure motion along x axis. Using a width of $(w = \pm 1.6)$ for the pure axis motion zones, the teaching wand was moved along x axis while the deviations in the y and z axis remained within the bounding box for the pure axis motion zone. Figure 8 shows the tracked distance offsets of IRIS and the computed smooth normalized velocity offsets after filtering. Notice that no velocity offsets are present in the y and z axis directions, resulting in a pure motion along x axis.

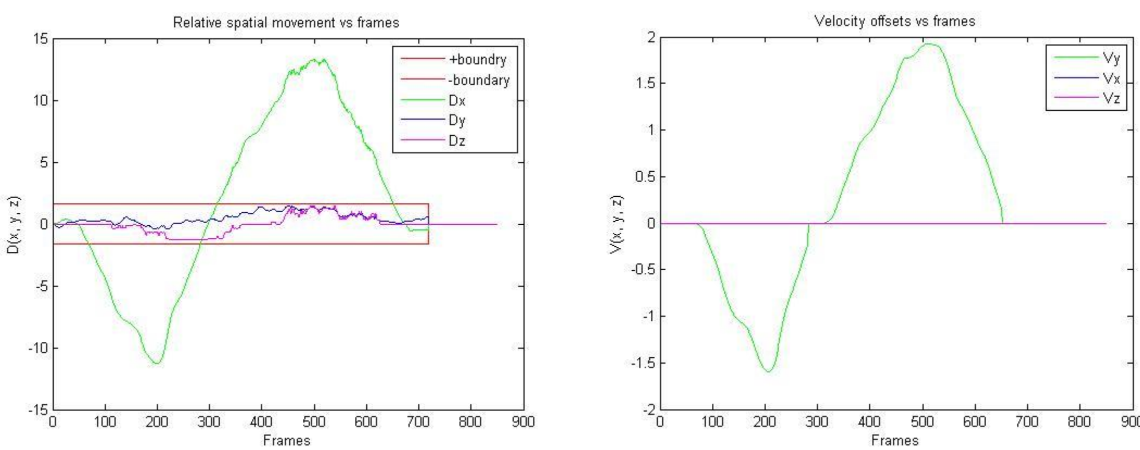


Figure 8: Distance and velocity offsets for pure axis motion

Experiment 2 uses the same tracked data from the last one but reduces the width for the pure axis motion zones to $(w = \pm 1.2)$. As shown in Figure 9, some of the tracked distance offsets of IRIS in the y and z axs directions are now outside the bounding box for the pure axis motion zone. The corresponding computed normalized velocity offsets now has some components in the y and z axis directions. This demonstrates the need for properly setting the width w of the pure axis motion zones to completely filter out any small deviations when trying to manually move the teaching wand along only one of the axis.

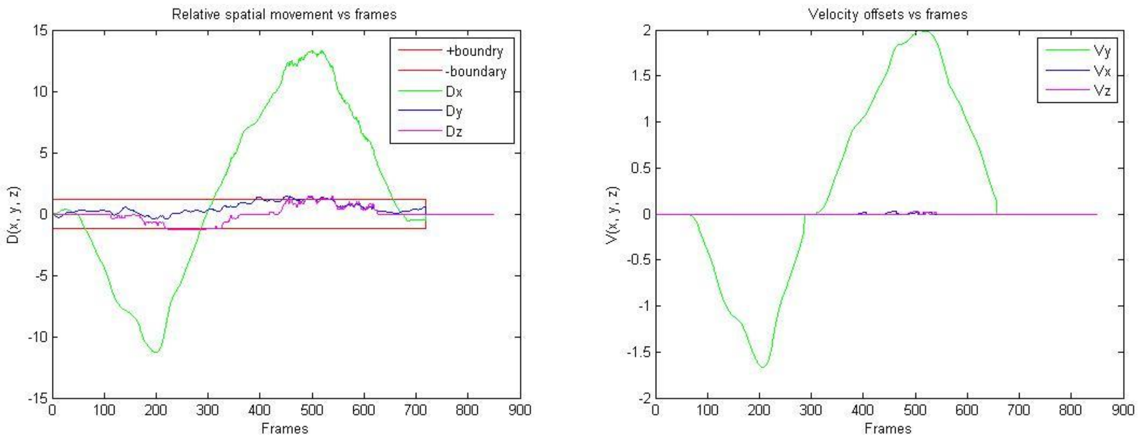


Figure 9: Distance and velocity offsets for pure axis motion with smaller bounding box

Experiment 3 demonstrates the results for a general motion of the teaching wand in 3D space. Using a width of $(w = \pm 1.6)$ for the pure axis motion zones, the teaching wand was moved in the operating zone in all directions. Figure 10 shows the tracked distance offsets of IRIS and the computed smooth normalized velocity offsets after filtering. Notice that any motion within the bounding box did not result in generation of the corresponding velocity offsets.

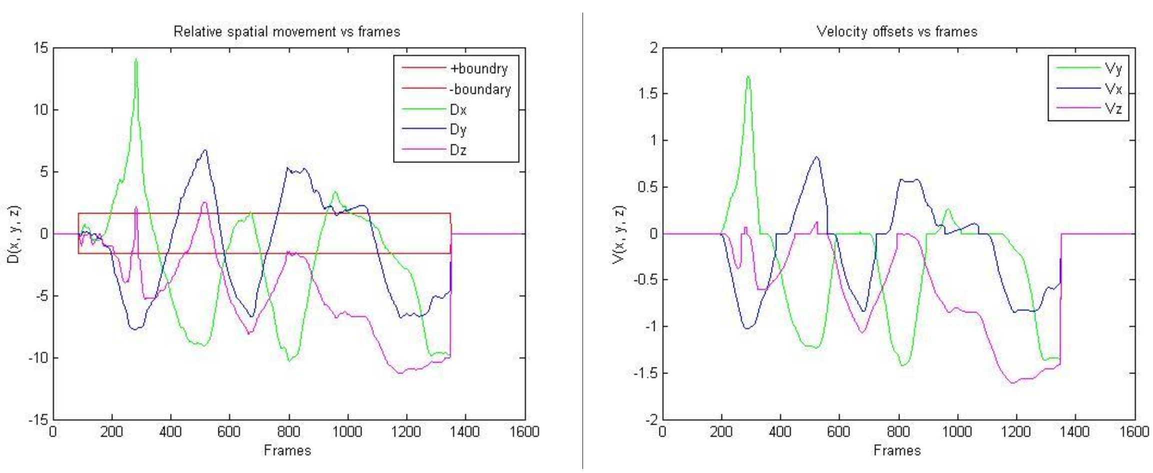


Figure 10: Distance and velocity offsets for motion in all three directions

5 Conclusions and Future Work

This paper presented an intuitive and robust method for robot tool path teaching using a visual teaching wand. Using a single IR camera, the IR illuminated source on the teaching wand was tracked in the 3D space. The use of an IR source helped eliminate the effects of other light sources in the environment resulting in a reliable tracking of its 3D location. This is an important feature for this teaching method to be considered for use in a typical industrial environment. The handheld teaching wand had jitter associated with it that needed to be removed in order to achieve a smooth robot motion. The idea of a deadzone and bounding boxes about the axes was presented to not only help remove jitter but also to implement convenience features like axis constrained motion that assists the handheld teaching wand realize some of the useful robot teach pendant jog controls. Additionally, the distance offsets were generated relative to a start point that is dynamically created in 3D space when the motion enable button is pressed on the teaching wand. This eliminates the need to accurately position the teaching wand in 3D space and robot motion can be started by the teaching wand from any location as long as it is within the operating view of the camera. The use of distance offsets from the start point to map directly into robot velocity offsets helped realize a joystick like control in 3D space that can be a very natural way to control the positioning of the robot tool tip. Aligning the robot world axes with the camera spatial axes makes the robot motion control mirror the teaching wand movements in the 3D space. The teaching wand only needs to be moved in the direction of the required tool tip motion without worrying about precisely locating the teaching wand in the camera view. These features allow the operator to mainly concentrate on the accuracy of the tool tip positioning while moving the handheld teaching wand in a natural way in the 3D space.

Currently, the threshold values and bounding box sizes used in the vision processing algorithms are preset based on a typical setup and operation. These values can possibly be automatically generated based on an initial self-calibration of the system, resulting in optimal performance in different lighting environments and when used by different operators. Hough transform was used on each frame to locate the teaching wand which is compute intensive and can potentially slow down tracking. Other methods need to be investigated that work with only a subset of image frames and can reliably track the target in order to reduce the computation requirement for this teaching method. Advanced filtering methods also need to be investigated that generate smooth velocity offsets for the robot that result in the most precise and intuitive control of the tool tip position.

The ongoing research is looking at the use of an inertial measurement unit on the teaching wand to measure its orientation. This information can then be used to generate velocity offsets for orientation control of the robot tool tip. Hence, a full 6 degree of freedom control of the robot tool tip can then be realized making the teaching wand an intuitive way to teach manual applications on a robot.

Bibliography

[1] ABB. (2014, May) IRC5 Teach pendant. [Online]. http://new.abb.com/products/robotics/controllers/irc5/irc5-data

[2] KUKA. (2014, May) smartPAD. [Online]. http://www.kuka-robotics.com/en/products/controllers/smartPAD/start.htm

[3] FANUC. (2014, May) iPendant. [Online]. http://www.fanucamerica.com/file-repository/DataSheets/Vision-Software-Controls-Options/iPendant.pdf

[4] Bjorn Hein, Martin Hensel, and Heinz Worn, "Intuitive and Model-based On-line Programming of Industrial Robots: A Modular On-line Programming Environment," in *2003 IEEE International Conference on Robotics and Automation*, Pasadena, CA, 2008, pp. 3952 - 3957.

[5] Rainer Bischoff and Arif Kazi, "Perspectives on Augmented Reality Based Human-Robot Interaction with Industrial Robots," in *Proceeding of IEEE/RSJ International Conference on Intelligent Robots and Systems*, Sendai, Japan, 2004, pp. 3226-3231.

[6] T. Pettersen, J. Pretlove, C. Skourup, T. Engedal, and T. Løkstad, "Augmented Reality for Programming Industrial Robots," in *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, 2003, pp. 319-320.

[7] H. C. Fang, S. K. Ong, and A. Y. C. Nee, "Robot Programming using Augmented Reality," in *International Conference on CyberWorlds*, 2009, pp. 13-20.

[8] Chuen Leong Ng, Teck Chew Ng, Thi Anh Ngoc Nguyen, Guilin Yang, and Wenjie Chen, "Intuitive Robot Tool Path Teaching Using Laser and Camera in Augmented Reality Environment," in *2010 11th Int. Conf. Control, Automation, Robotics and Vision*, Singapore, 2010, pp. 114-119.

[9] P. Neto, J. N. Pires, and A. P. Moreira, "Accelerometer-based control of an industrial robotic arm," in *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication*, Toyama, Japan, 2009, pp. 1192–1197.

[10] J. Aleotti, A. Skoglund, and T. Duckett, "Position teaching of a robot arm by demonstration with a wearable input device," in *International Conference on Intelligent Manipulation and Grasping*, Genoa, Italy, 2004, pp. 459–464.

[11] Hsien-I Lin and Yu-Hsiang Lin, "A Novel Teaching System for Industrial Robots," *sensors*, vol. 14, no. 4, pp. 6012-6031, March 2014.

[12] F. Nagata et al., "Joystick teaching system for polishing robots using fuzzy compliance control," in *Proceedings of IEEEInternational Symposium on Computational Intelligence in Robotics and Automation*, Alberta, Canada, 2001, pp. 362–367.

[13] Y. H. Kim, W. M. Park, Rha, and K. H. Rha, "Laboratory-level telesurgery with industrial robots and haptic devices communicating via the internet," *International Journal of Precision Engineering and Manufacturing*, vol. 10, no. 2, pp. 25-29, 2009.

[14] Pedro Neto, J. Norberto Pires, and A. Paulo Moreira, "High-level programming and control for industrial robotics: using a hand-held accelerometer-based input device for gesture and posture recognition," *Industrial Robot: An International Journal*, vol. 37, no. 2, pp. 137-147, 2010.

[15] Qing Wang, Feng Chen, Wenli Xu, and Ming-Hsuan Yang, "An experimental comparison of online object-tracking algorithms," in *Wavelets and Sparsity*, vol. 8138, San Diego, 2011, p. 11.

[16] Ming Cui, Jiuxiang Hu, Anshuman Razdan, and Peter Wonk, *Color to Gray Conversion Using ISOMAP*., 2009.

[17] Otsu N., "A Threshold Selection Method from Gray-Level Histograms," in *IEEE Transactions on Systems, Man, and Cybernetics Vol. 9, No. 1*., 1979, pp. 62-66.

[18] Jae S Lim and Cliffs Englewood,. NJ: Prentice Hall, 1990, pp. 469-476.

[19] Chih-Wei Kuo, Chih-Chin Lai, Ming-Dar Tsai, Yuan-Chang Chang, Chih-Lung Lin, "A novel approach to fast noise reduction of infrared image," *Elsevier*, vol. 54, pp. 1-9, 2011.

[20] Tengku Mohd Tengku Sembok, Z.A. Bakar, and N Jamil, "Noise removal and enhancement of binary images using morphological operations," in *IEEE*, 2008.

[21] D.J. Kerbyson T.J Atherton, "Size invariant circle detection," *Image and Vision Computing*, vol. 17, no. 11, pp. 795-803, 1999.

[22] J. Princen, J. Illingworth, and J. Kittler. H.K Yuen, "Comparative study of Hough transform methods for circle finding," *Image and Vision Computing*, vol. 8, no. 1, pp. 71–77, 1990.

[23] Dietrich Schlichthärle, *Digital Filters Basics and Design*. Berlin Heidelberg: Springer, 2011.

[24] MathWorks. (2014) MATLAB. [Online]. http://www.mathworks.com/products/matlab/

Ruled Based Approach for Scheduling Flow-shop and Job-shop Problems

Mohammad Komaki, Shaya Sheikh, Behnam Malakooti

Case Western Reserve University

Systems Engineering

Email: komakighorban@gmail.com

Abstract

In this study, we investigate the scheduling of n-jobs on m-machines flow shop and job shop. These scheduling problems are NP-hard and finding the good solutions in reasonable computational time is the interest of the many researches. In this study, we proposed a novel graphical method to schedule two jobs on m machines job shop where the optimality of the final solution is guaranteed. Also, another heuristic algorithm called Rule based Approach to schedule n jobs on m machines which is a constructive heuristic algorithm and can be applied to flow shop and job shop scheduling problems.
Key words: Manufacturing System; Scheduling; flow shop; job shop; Ordinal and Cardinal Head-Tail Methods;

1. Introduction

The main purpose of scheduling problems is finding the optimal (or near optimal) sequence of tasks (or jobs) on a set of limited machines (processors) to utilize efficiently the available resources. Effective schedules can result in customer satisfaction and improved utilization of resources. Scheduling problems can also be classified as flow shop or job shop. In Flow shop, n jobs are processed by m different processors in the same order of processors, but in job-shop, each job may be processed by a different sequence of processors. The classes of sequencing problems are n×1 (where n jobs are processed on one processor), n×2, n×3, 2×m, and n×m.

In this paper, we focus on the scheduling of the n-jobs on the m-machines which either are flow shop or job shop scheduling problems. It is shown that most of the variation of these problems in general cases are NP-Complete problems [2] implying that finding algorithm which can obtain the solution in polynomial computational time is almost impossible.

Johnson [11] studied the minimizing the makespan, the completion time of the last job on the system, of two and three-stage flow shop problem and proposed a heuristic algorithm which can find the optimal solution. Since then a lot of researches have been carry out to tackle the flow shop scheduling problems. These studies can be grouped into exact methods such as Branch and Bound (B&B), heuristic algorithms and metaheuristic algorithms. For example, Ignall and Schrage [3] and McMahon and Burton [5] applied B&B to flow shop scheduling problem. Campbell et al. [1], Gupta [10] and Nawaz et al. [6] proposed heuristic algorithm to solve flow shop scheduling problem. A comprehensive review of proposed heuristic algorithms can be found in Ruize and Maroto [9]. Also, recently Malakooti [8] has provided comprehensive review of single and multi-

objective scheduling problems.
The reminder of the paper is organized as following. Next section discusses the algorithms to schedule n jobs by two processors (flow shop and job shop) and section 3 extends the discussed algorithms in section 2 to three and m processors. Section 4 discusses the graphical method to schedule 2 jobs on m machines job shop which the solution is guaranteed to be optimal. Section 5 proposes another heuristic algorithm called Rule based or Head-Tail method which is a constructive algorithm and can be applied to job shop and flow shop scheduling problems.

2. SEQUENCING n JOBS BY 2 PROCESSORS

In this section scheduling of n jobs on 2 processors are discussed. Section 2.1 discusses the well-known Johnson's rule for scheduling n jobs on 2 machines and next section extends this method for job shop.

2.1 Flow Shop of n×2 (Johnson's Rule)

For sequencing n jobs by two processors, Johnson's algorithm provides an optimal solution for minimizing the total time to complete all jobs (which is defined as Makespan, M). The algorithm is based on the shortest processing time (SPT) rule when there are two processors, A and B. Label the first processor as A and the second processor as B, and denote A→B as the order of the processors. Denote

A_i = Processing time on processor A (the first processor) for job i

B_i = Processing time on processor B (the second processor) for job i

Johnson's Algorithm is as follows.

1. List the processing times of all jobs, i = 1, 2, …, n, on processors A and B by $\{A_1, A_2, \ldots, A_n; B_1, B_2, \ldots, B_n\}$. Consider n unassigned locations that must be assigned to the n jobs such that only one job is assigned to each location and vice versa.
2. Choose the job that has the shortest processing time in the processing times list, and label it as job j.
   a) If the shortest processing time is associated with processor A, place job j in the first available position on the job locations.
   b) If the shortest processing time is associated with processor B, place job j in the last available position on the job locations.
3. Remove both A_j and B_j from the list.
4. Repeat steps 2 and 3 until all jobs have been placed in the sequence.

2.2 Job-Shop of n×2

In the job-shop problem, each job may have a different sequence of processors. In the case of two processors, some jobs must be processed first on Processor A and then on Processor B, A→ B, but other jobs must be processed first on processor B and then A, B→A. Although this problem is more complex than the n×2 flow shop problem covered in the last section, it can be solved by an extension of Johnson's algorithm.
Consider Johnson's algorithm covered in the last section. Use the following procedure instead of step 2 of Johnson's algorithm.
Choose job j associated with the shortest processing time from the list.

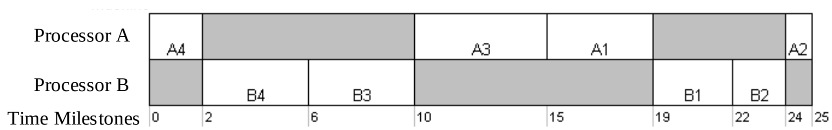
a) If the shortest processing time is associated with processor A, and
i) if job j has sequence A→B, then place job j in the first available space.
ii) if job j has sequence B→A, then place job j in the last available space.
b) If the shortest processing time is associated with processor B, and
i) if job j has sequence A→B, then place job j in the last available space.
ii) if job j has sequence B→A, then place job j in the first available space.

Example 2.1: Job-shop Sequencing of 4 Jobs by 2 Processors

Consider the following problem.

Job i	1	2	3	4
Processor A processing time	4	1	5	2
Processor B processing time	3	2	4	4
Processor Sequence	A→B	B→A	B→A	A→B

The ordered list of processing times is {A_2, A_4, B_2, B_1, A_1, B_3, B_4, A_3}.
1. The shortest time in the list is A_2, and job 2 has sequence B→A. Place job 2 at the end of the sequence. Remove A_2 and B_2 from the list, the list becomes {A_4, B_1, A_1, B_3, B_4, A_3}.
2. The shortest time in the list is now A_4, and job 4 has sequence A→B. Place job 4 in the first available space. Remove A_4 and B_4 from the list result, the list becomes {B_1, A_1, B_3, A_3}.
3. The next shortest time in the list is B_1, and job 1 has sequence A→B. Place job 1 in the last available space. Removing both A_1 and B_1 from the list results in {B_3, A_3}.
4. Only job 3 remains, so it is placed in the remaining space for the sequence. The solution is 4→ 3→ 1→ 2 which has the following Gantt chart with a makespan of 25 days.



3. SEQUENCING n JOBS BY m PROCESSORS

3.1 Algorithm for n×3

The problem of sequencing n jobs by 3 processors is similar to the problem of sequencing n jobs by 2 processors. All jobs are processed in the same order on three processors, A, B, and C; first on A, then B, and finally on C. An optimal solution can be found by an extension of Johnson's algorithm if certain conditions are met. If the conditions are not met, a different method, such as the branch and bound algorithm, can be used to obtain the optimal solution.

Consider three processors labeled as A, B, and C, where:

A_i = Processing time on processor A (first processor) for job i

B_i = Processing time on processor B (second processor) for job i

C_i = Processing time on processor C (third processor) for job i

Consider Min. A = Min{A_1, A_2, …, A_n}, Max. B = Max{B_1, B_2, …, B_n}, and Min. C = Min{C_1, C_2, …, C_n}. Johnson's algorithm will render an optimal solution if either of the following two conditions holds:

a) Min. A $\geq$ Max. B

or b) Min. C $\geq$ Max. B

That is, for all jobs, the minimum processing time on either processor A or processor C must be greater than or equal to the maximum processing time of processor B. If the above condition is not met, optimality cannot be guaranteed.

The following steps can be used to find the solution to n×3 problems.

1. Create two dummy processors and label them as processor AB and processor BC.
    - The processing time of job i on processor AB is equal to the sum of the processing time for job i on processors A and B.
    - The processing time of job i on processor BC is equal to the sum of the processing time for job i on processors B and C.
2. Now apply Johnson's algorithm to the n jobs on the dummy processors AB and BC.
3. Use the job sequence order obtained in Step 2 to make a solution for the three processors, A, B, and C (This can be accomplished either by using a Gantt chart or a table).

It should be noted that in most applications the above optimality conditions (a or b) may not hold, and the above method will render a heuristic solution.

Example 3.1: Sequencing 4 Jobs by 3 Processors

The following table shows the processing times (in days) for four jobs on three processors. The processing order is A→B→C. Find the best sequence and verify whether the solution is optimal.

Job i	1	2	3	4	Min.	Max.
Processor A	10	6	7	10	6	NA
Processor B	3	1	5	6	NA	6
Processor C	8	8	9	9	8	NA

First, check the conditions for which the Johnson's algorithm can render an optimal solution.

Min. A = Min{10, 6, 7, 10} = 6; Max. B = Max{3, 1, 5, 6} = 6; Min. C = Min{8, 8, 9, 9} = 8

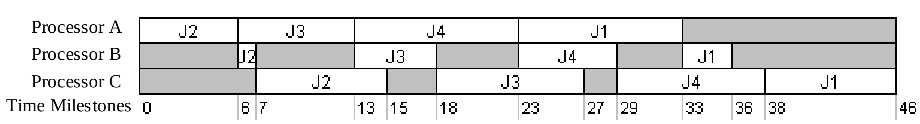
Since Min. A ≥ Max. B, an optimal solution can be found. Now, calculate the processing times for the dummy processors AB and BC. The processing time AB_i is the sum of processing times A_i and B_i. The time BC_i is the sum of B_i and C_i. This is presented in the following table.

Job i	1	2	3	4
Processing Time AB_i	13	7	12	16
Processing Time BC_i	11	9	14	15

Now apply Johnson's algorithm to the above table. The list of processing times is {13, 7, 12, 16; 11, 9, 14, 15}. The shortest time in the list is AB_2 = 7. Place job 2 at the beginning of the sequence. Removing both AB_2 and BC_2 from the list, it becomes {13, 12, 16; 11, 14, 15}. The shortest time in the list is now BC_1 = 11. Place job 1 at the end of the sequence. Removing both AB_1 and BC_1 from the list, it becomes {12, 16; 14, 15}. The shortest time in the list is now AB_3 = 12. Place job 3 in the first available space, which is the second space. Only job 4 remains, so it is placed in the remaining space in the sequence. In summary the solution is:

Iteration	Job Assignment
1	2 _ _ _
2	2 _ _ 1
3	2 3 _ 1
4	2 3 4 1

Therefore, the optimal sequence is 2→3→4→1. Now show the Gantt chart for the sequence on the three processors A, B, and C. The makespan is 46.



3.2. n×2 Composite Approach for Solving n×m Flow Shop (Campbell Method)

In the previous section, we show that in some cases it may be possible to find the optimal solution for n×3 problems. For n×m where m is more than 3, there is no easy way to find the optimal solution. The method often used to solve this problem is based on the results of the method in the previous section where the n×3 was reduced to a n×2 problem so that Johnson's rule can be applied to it. Such a method was developed by Campbell (1970) which uses the n×2 Johnson's algorithm as a subroutine. The general approach is to reduce the n×m problem to a series of n×2 problems and then solve each problem by Johnson's algorithm in a manner similar to how n×3 problems were solved in the last section. The pseudo-processors for the algorithm can be generated using the rules provided for the n×3 method. The idea is to generate a set of composite pseudo processors by incrementally combining different processors. A set of n×2 problems can be generated, each having different composite processors. Each of these composite processor problems can then be solved, and the best alternative can be chosen from the given set of solutions.

The process for creating composite pseudo-processors is as follows:

1. Set k = 1.
2. Consider the first k processors versus the last k processors. Find the summation of job times of the first k processors versus the summation of job times of the last k processors. Solve the generated n×2 problem by Johnson's rule. Record the solution and the makespan.
3. Set k equal to k+1 and repeat step 2 until k = m-1.
4. Choose the sequence that has the minimum makespan.

Example 3.2: Flow Shop of Four Jobs with Four Processors Using n×2

Consider the problem of sequencing four jobs by four processors (A, B, C, and D) as presented in the following table. All jobs are processed in the order (A→B→C→D).

Job i	A	B	C	D	Total
1	4	1	3	2	10
2	1	6	5	3	15
3	5	2	4	1	12
4	2	4	1	5	12
Total	12	13	13	11	49

The total processing times on processors A, B, C, and D are 12, 13, 13, and 11, respectively. Therefore, it can be observed that the worst (feasible) scheduling would have a makespan of 49 minutes, doing jobs sequentially one at a time. The ideal scheduling would have a makespan of 15, but usually such an ideal solution is infeasible.

For k=1, consider just the first and last processors' processing times, i.e. only consider processors A and D. Now solve this 4×2 problem. Using Johnson's algorithm, the sequence is 2→4→1→3. For this sequence the makespan is 23, calculated by using a table or a Gantt chart as discussed earlier.

For k=2, consider the first two and the last two processors. Create two pseudo-processors, AB and CD. For AB add up the processing times of A and B. For CD add up the processing times of C and D. The resulting problem is shown below. Applying Johnson's algorithm gives a sequence of 1→4→2→3 with a makespan of 26.

Next, for k=3, create pseudo-processors for the first three processors, ABC, and the last three processors, BCD. Applying Johnson's algorithm gives a sequence of 4→2→3→1 with a makespan of 26.

For k=1		
Job i	A	D
1	4	2
2	1	3
3	5	1
4	2	5
2→4→1→3, f_1 = 23		

For k=2		
Job i	AB	CD
1	5	5
2	7	8
3	7	5
4	6	6
1→4→2→3, f_1 = 26		

For k=3		
Job i	ABC	BCD
1	8	6
2	12	14
3	11	7
4	7	10
4→2→3→1, f_1 = 26		

Now, choose the best solution. The best sequence is 2→4→1→3 (associated with the k=1 solution) with a makespan of 23.

4. JOB SHOP OF 2 JOBS BY m PROCESSORS

In the previous sections, the sequencing of n jobs by one, two, three, and m processors was introduced. It was shown that optimality can be guaranteed only for n×2 problems. In this section, we discuss an algorithm for processing two jobs on m processors where each job may have a different processing sequence. For example, consider a job shop problem where two jobs are processed by five different operations: sawing, drilling, turning, tapping, and milling. Each job is processed by a different order of these processors.

The objective is to find a solution that minimizes the makespan. A graphical method is used to determine the sequence with the shortest processing time. This algorithm is proposed by Akers [12] and the solution to this algorithm is guaranteed to be optimal. See Figure 4.1 for an example.

1. On the **x**-axis, depict the processing times of job 1 on the M processors sequentially.
2. On the **y**-axis, depict the processing times of job 2 on the M processors sequentially.
3. Shade the areas where the same processor is on both the **x**-axis and the **y**-axis. This represents both jobs being processed by the same processor, which is a conflict.
4. By trial and error, try to find the best path from the origin to the upper right corner of the graph while not passing through the shaded areas. The line may travel either horizontally, vertically, or diagonally (45°). The best path will move diagonally for the longest portion.
5. The solution can be interpreted as follows:
    - A diagonal line (45°) indicates that both jobs are being processed at the same time.
    - A horizontal line indicates that only job 1 is being processed while job 2 is idle.
    - A vertical line indicates that job 2 is being processed while job 1 is idle.

By maximizing the portion of the path that is diagonal, both jobs will be processed at the same time as much as possible, and therefore the makespan will be minimized. A shaded area mean that both jobs cannot be processed simultaneously on the same processor, so it is an infeasible area to schedule.

Example 4.1: Sequencing 2 Jobs 4 Processors

Consider the following processing times for two jobs by four processors. Job 1 has processor sequence C→A→D→B and job 2 has processor sequence B→A→C→D. Find the best job sequece for this problem and show the start and finish of each job on each processor.

Processor	A	B	C	D	Total	Sequence
Job 1	3	1	2	6	12	C→A→D→B

Job 2	2	3	2	3	10	B→A→C→D

Job 1 has a total flow time of 12 and is depicted as the x-axis. Job 2 has a total flow time of 10 and is depicted as the y-axis. The rectangle size is 12×10. See Figure 4.1.

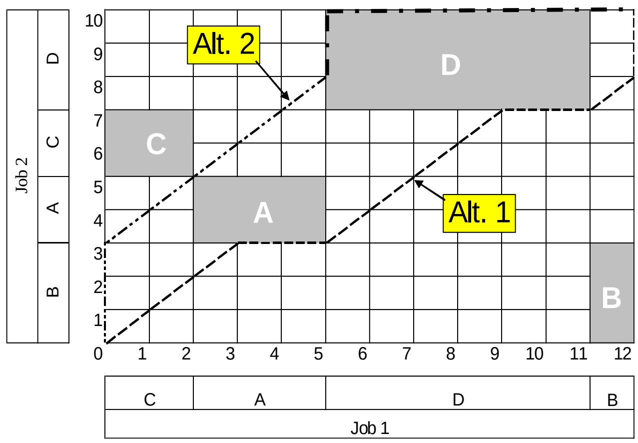


Figure 4.1: Sequencing two jobs by four processors, A, B, C, and D. Alternative 1 has a makespan of 14 and Alternative 2 has a makespan of 17.

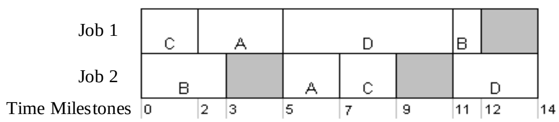
The blocked areas, A, B, C, and D, indicate infeasible areas. Now, find a line that connects (0, 0) to (12, 10). Such a line is a solution to the problem. However, note that the line that directly connects (0, 0) to (12, 10) passes through the A and D shaded regions, and therefore it is not a feasible solution. Now, try to find the line with the longest diagonal line portion that connects (0, 0) to (12, 10) without going through any of the shaded areas of A, B, C, and D. For the purpose of illustration, we present two alternatives, 1 and 2, that connect the two points, (0, 0) and (12, 10).

Alternative 1 starts from the origin and goes diagonally towards the blocked area of processor A. After reaching the blocked area for processor A at point (3, 3), the only feasible way is to move horizontally until it passes the shaded area of A. Recall that this horizontal line indicates that only Job 1 is being processed on processor A while Job 2 is idle, awaiting the same processor to become available. This occurs at point (5, 3). At (5, 3) the line runs diagonally until it reaches the shaded area of processor D at point (9, 7). Once again the line must run horizontally until it passes the shaded area of Processor D. At point (11, 7), the line runs diagonally again. At point (12, 8), the line must run vertically along the edge of the graph until it reaches the corner point. Processors B and C are not in conflict in Alternative 1. The length of the diagonal portion of the line can be measured by adding the number of blocks along the path that are crossed by the diagonal line. Alternative 1's line has a diagonal length of eight.

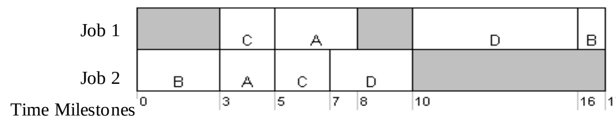
Now, consider Alternative 2. The path runs vertically from point (0, 0) to point (0, 3), and then it runs diagonally until it reaches the shaded area of processor D at point (5, 8). It then runs vertically along the shaded area of processor D up to point (5, 10). Lastly, it runs horizontally starting at point (5, 10) until it reaches the corner point (12, 10). This alternative has only one conflict, in using processor D. Its diagonal length is five.

Alternative 1 has a diagonal length of eight. Alternative 2 has a diagonal length of five. This means that Alternative 1 is better than Alternative 2. By trial and error, one can verify that alternative 1 has the longest feasible diagonal portion, and therefore its solution is optimal. The Gantt charts for Alternatives 1 and 2 are shown below. The shaded areas represent idle times. For Alternative 1, the makespan is 14 and for alternative 2, the makespan is 17.

Schedule for Alternative 1



Schedule for Alternative 2



5. Sequencing of n×m: Head-Tail Approach

5.1 Overview of Head-Tail for Solving n×m

There are two types of n×m sequencing problems: a) flow shop, where the sequence of processors is the same for all jobs, and b) job shop, where the sequence of processors may be different for different jobs. The n×m job shop problem is much more difficult to solve than the n×m flow shop problem. In this chapter, two methods for solving flow shop problems are discussed. The first method is based on the n×2 Shortest Processing Time which is covered in Section 2. The second method is based on the 2×m graphical method which is covered in Section 4. Section 4 covers job shop problems and also uses the 2×m graphical method as its basis. The problem of sequencing n jobs by m processors is a combinatorial problem; therefore finding the optimal solution for large problems is computationally infeasible. In this section, we develop a new heuristic for solving n×m sequencing problems. The approach is based on the Uni-directional Head-Tail method. To solve the n×m sequencing problem, first convert the n×m problem into a series of 2×m problems; then solve each 2×m problem by the graphical method covered in Section 4. Then generate an n×n flow matrix which will be used by the Head-Tail method.

Uni-direction Head-Tail Methods for Solving n×m: Four methods are presented in the Cellular Layout and Networks Chapter for solving Uni-directional flow problems. They are (in order of complexity) Ordinal, Cardinal, Simultaneous, and Integer Programming. When the final solution is obtained, the Pair-wise Exchange method can be used to attempt to find an improved solution. All four methods can be used to solve sequencing problems. The Ordinal method requires much less computational time than the Cardinal method, but the solution may not be as good as the solution obtained from Cardinal method. In this chapter, we only show the application of the Cardinal uni-directional head-tail method.

2×m Composite Formulation of n×m Problems: Consider job i versus job j. Solve this problem by the 2×m graphical method, record the solution, and identify which job is completed first. Suppose that job i is completed first. Now solve the 2×m problem such that job j is completed first, and job i is completed second. Record both these solutions and their makespans. These solutions correspond to (i, j) and (j, i) in the flow matrix, respectively. Note that it is not important which job is started first, but it is important to know which job is completed first.

Steps of Head-Tail Methods: By using a set of simple rules, this method progressively builds the solution by connecting pairs of jobs. The following notations are used for connecting pairs of sequences of jobs.

s_p = (i → j) is a candidate pair sequence where job i is completed before job j.
s_q = (→ → →…) is a partially constructed sequence solution so far.

1. Consider the ranking of all pairs of jobs (i, j). Rank pairs based on their makespans from the shortest to the longest makespan.
2. Consider the highest ranked unassigned pair, s_p = (i , j), such that
   a. Both i and j are not in s_q
   b. Either i or j is in s_q.
3. Use relevant Head-Tail Rules to combine s_p and s_q.
4. Repeat steps 2 and 3 until all jobs are assigned to s_q.
5. Try to improve the obtained final solution by using the Pair-Wise Exchange method.

Uni-Directional Cardinal Rule (Conflict Resolution): If there is a common job, i, in s_p = (i→j) and s_q = (…→ i →…), then consider inserting j (the uncommon job) between each pair of jobs in in s_q. Choose the alternative that has the best objective function value (e.g. the minimum makespan). For example, if s_p = (2→5) and s_q = (1 → 2 → 4), the uncommon job is 5. The common job is 2. Consider all possible locations that job 5 can be inserted in sequence s_q. They are: (5 → 1 → 2 → 4), (1 → 5 → 2 → 4), (1 → 2 →5 → 4), and (1 → 2 → 4 → 5). Choose the sequence that has the lowest makespan as the new s_q.

5.2 Flow-Shop of n×m: Cardinal Head-Tail

For flow-shop problems, the sequence of processors is the same for all jobs; therefore, the Gantt-Chart solution will generally provide the optimal solution for 2×m problems. Therefore, we suggest using the Gantt Chart directly to find the makespan for a given 2×m problems, i.e. it is not necessary to use the 2×m graphical method. The steps of Cardinal Head-Tail method are shown by solving the following example.

Example 5.1: Solving a 4×4 Flow Shop Problem by Cardinal Head-Tail

Consider Example 3.2, the flow shop problem of four jobs by four processors. Solve this problem by the Cardinal Head-Tail method.

2×m Composite Formulation of n×m for Flow-Shop: First generate the n×n makespan matrix by considering each pair of jobs. For example, consider job 2 versus job 3. Solve this problem by 2×m graphical method (or the Gantt chart). The makespan is 17 when job 2 is scheduled before job 3. Now find the best solution when job 3 is scheduled to be finished before job 2. The makespan is 21 for this solution.

The makespans for all combinations of each pair of jobs are shown in the matrix below.

Job/Job	1	2	3	4
1	0	19	16	16
2	17	0	17	20
3	16	21	0	17
4	14	20	14	0

Cardinal Head-Tail Methods for Solving n×m Flow Shop Problems: Now, apply the Cardinal head-tail method to the matrix. First, order the pairs in ascending order.

Rank	1	2	3	4	5	6	7	8	9	10	11	12
w^o Pair	4→1	4→3	1→3	1→4	3→1	2→1	2→3	3→4	1→2	2→4	4→2	3→2
w_{ij}	14	14	16	16	16	17	17	17	19	20	20	21

1. Consider the first ranked pair of jobs, s_p = (4→1); add s_p to s_q; s_q = (4→1).
2. Consider s_p = (4→3); apply Cardinal Rule 2 and find the best location for job 3. There are three alternatives to be considered as presented below.

Sequence	3→4→1	4→3→1	4→1→3
Makespan	19	18*	18

   4→3→1 and 4→1→3 have the same makespan. Alternative s_q= (4→3→1) is selected arbitrarily.
3. Consider s_p = (1→3), but ignore it due to its conflict with (3→1) in s_q = (4→3→1). Ignore (1→4) and (3→1) as 1, 4 and 3 are already assigned in s_q = (4→3→1).
4. Consider s_p = (2→1); as compared to s_1= (4→3→1), there are four possible location for job 2 as presented below.

Sequence	2→4→3→1	4→2→3→1	4→3→2→1	4→3→1→2
Makespan	23*	26	25	26

2→4→3→1 has the lowest makespan, 23.
A summary of the method is provided below. The makespan is 23.

Iteration	Used Rank	s_p	Action Taken	Resulting sequence, s_q	Candidates
1	1	4→1	Add s_p as a new sequence	(4→1)	2,3
2	2	4→3	Insert station 3 before station 1	(4→3→1)	2
3	6	2→1	Insert station 2 before station 4	(2→4→3→1)	-

5.3 Job-Shop of n×m: Cardinal Head-Tail

As stated before, in job-shop problems, different products may have different machine sequences. As a result, job-shop problems are far more difficult to solve than flow-shop problems. The same Cardinal Head-Tail method presented in the previous section can be used to solve this problem. Note that it is necessary to generate an n×n flow matrix by solving 2×m problems by the graphical method (of Section 5) for each pair of jobs (i,j).

Example 5.1: Solving a 4×4 Job Shop Problem by the Cardinal Head-Tail Method

Consider a job-shop problem of processing four jobs by four processors as shown in the following table. Solve this problem by the Cardinal Head-Tail method.

Job i	A	B	C	D	Processor Sequence
1	4	1	3	2	D→C→B→A
2	1	6	5	3	A→B→C→D
3	5	2	4	1	B→A→C→D
4	2	4	1	5	C→A→B→D

2×m Composite Formulation of n×m Job Shop Problem: For example, consider job 2 versus job 3. Use the 2×m graphical method (of Section 5) to solve this problem. The solution (details not shown) indicates that the makespan for completing job 2 before job 3 is 16, and the makespan for completing job 3 before job 2 is 18. The makespans and their ranking for all pairs are presented below.

Makespan Matrix of Pairs of Jobs				
Job/Job	1	2	3	4

Rankings of Pairs of Jobs				
Job/Job	1	2	3	4

1	0	12	10	12
2	15	0	16	19
3	12	18	0	13
4	12	16	12	0

1	-	2	1	2
2	8	-	9	12
3	2	11	-	7
4	2	9	2	-

The steps of the Cardinal Head-Tail method are presented below.

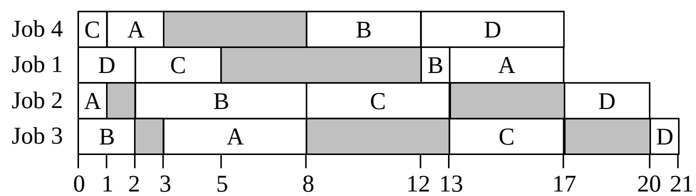
1. Consider the first pair of jobs, $s_p = (1\rightarrow3)$; set $s_q = (1\rightarrow3)$.
2. Consider $s_p = (4\rightarrow1)$; find the best location for job 4. The best solution is $s_q = (4\rightarrow1\rightarrow3)$.

Sequence	4→1→3	1→4→3	1→3→4
Makespan	13	20	20

3. Consider $s_p = (1\rightarrow2)$; this is in conflict with s_q. Find the best location for job 2 in the sequence $(4\rightarrow1\rightarrow3)$.

Sequence	2→4→1→3	4→2→1→3	4→1→2→3	4→1→3→2
Makespan	21*	21*	21*	21*

Each sequence has a makespan of 21. Therefore, each of these solutions is optimal. For the purpose of illustration, the sequence $(4\rightarrow1\rightarrow2\rightarrow3)$ is selected. Since all jobs are assigned, the final sequence is $(4\rightarrow1\rightarrow2\rightarrow3)$. The makespan is 21. The Gantt chart for this solution is presented below.



Iteration	Used Rank	s_p	Action Taken	Resulting sequence, s_q	Candidates
1	1	1→3	Add s_p as a new sequence	(1→3)	2, 4
2	2	4→1	Insert job 4 before job 1	(4→1→3)	2
3	2	1→2	Insert job 2 before job 3	(4→1→2→3)	-

Conclusion

In this paper, several novel heuristic algorithms to schedule n jobs on m machines, flow shop and job shop, are proposed. First of all, the basic algorithm to schedule n jobs on two machines flow shop is discussed and then this method is extended to job shop. Then, the algorithms to schedule n jobs on three machines are discussed. A novel graphical algorithm to schedule two jobs on m machines job shop is proposed which the optimality of the solution is guaranteed. Finally, a constructive heuristic algorithm which is based on the decomposing of the problem into 2×m problem and uses thee earlier discussed method is presented.

As direction of the future study, one can test the efficiency of the proposed methods in this paper in the assembly scheduling, for example see [13].

References:

[1] Campbell, H. G., R. A. Dudek, and M. L. Smith, "A Heuristic Algorithm for The n Job, m Machine Sequencing Problem." Management Science, Vol. 16, No. 10, 1970, pp. 630-637.

[2] Garey, M. R., Johnson, D. S., & Sethi, R. (1976). The complexity of flowshop and jobshop scheduling. Mathematics of operations research, 1(2), 117-129.

[3] Ignall, E., & Schrage, L. (1965). Application of the branch and bound technique to some flow-shop scheduling problems. Operations research, 13(3), 400-412
[4] Mahdavi, I., Komaki, G. M., & Kayvanfar, V. (2011, September). Aggregate hybrid flowshop scheduling with assembly operations. In Industrial Engineering and Engineering Management (IE&EM), 2011 IEEE 18Th International Conference on (pp. 663-667). IEEE.
[5] McMahon, G. B., & Burton, P. G. (1967). Flow-shop scheduling with the branch-and-bound method. Operations Research, 15(3), 473-481.
[6] Nawaz, M., Enscore Jr, E. E., & Ham, I. (1983). A heuristic algorithm for the m-machine, n-job flow-shop sequencing problem. Omega, 11(1), 91-95.
[7] Taillard ED, Benchmarks for basic scheduling problems, European journal of operation research, 64:1, 278-285,1993.
[8] Malakooti, B. (2013). Operations and Production Systems with Multiple Objectives. John Wiley & Sons.
[9] Ruiz, R., & Maroto, C. (2005). A comprehensive review and evaluation of permutation flowshop heuristics. European Journal of Operational Research,165(2), 479-494.
[10] Gupta, J. N. (1972). Heuristic algorithms for multistage flowshop scheduling problem. AIIE Transactions, 4(1), 11-18.
[11] Johnson, S. M. " Optimal Two and Three Stage Production Schedules With Setup Times Included." Naval Research Logistics Quarterly, Vol. 1, 1954, pp. 61-68.
[12] Akers, S. B. "A Graphical Approach to Production Scheduling Problems." *Operations Research,* Vol. 4, 1956, pp. 244-45.
[13] Mahdavi, I., Komaki, G. M., & Kayvanfar, V. (2011, September). Aggregate hybrid flowshop scheduling with assembly operations. In *Industrial Engineering and Engineering Management (IE&EM), 2011 IEEE 18Th International Conference on* (pp. 663-667). IEEE.

Multi-Objective Scheduling Using Rule Based Approach

Mohammad Komaki, Shaya Sheikh, Behnam Malakooti

Case Western Reserve University

Systems Engineering

Email: komakighorban@gmail.com

Abstract

Scheduling is finding the best possible assignment of jobs to the available resources in such way that the resulting solution should minimize the objective function. In practice, enterprises may have several conflicting criteria to evaluate the resulting solutions. In this study, multi-criteria scheduling problem of n- jobs on m-machines has been investigated. For this problem, we proposed a simple constructive heuristic algorithm which uses certain types of rules. The proposed method is efficient and it needs much less computation effort comparing to the conventional scheduling methods.
Key words: Scheduling; Production planning; flow shop and job shop; Multi-Criteria Scheduling, Rule-based Method

1. INTRODUCTION

Scheduling is to find the optimal order of processing different tasks on a set of different processors while effectively utilizing the limited available resources. Effective schedules result not only in improved utilization of resources but also customer satisfaction. The most important objectives of scheduling are listed below; these objectives are usually in conflict, and therefore Multi Criteria Decision Making (MCDM) approaches can be used to find a satisfactory solution.

1. Minimize the total time to complete all jobs (to maximize productivity)
2. Maximize meeting due dates (to maximize customer satisfaction)
3. Maximize utilization of resources (to minimize operational cost)
4. Minimize the average number of jobs waiting in the system (to minimize queues)

In scheduling problems, the job arrival time can be either static or dynamic. In static problems, all jobs are known and are ready to be processed at the beginning of the scheduling period. In dynamic problems, jobs may arrive intermittently throughout the system operations, and the sequence of jobs may change as new jobs arrive. Sequencing problems can also be classified as flow shop or job shop. In Flow shop, n jobs are processed by m different processors in the same order of processors. But in job-shop, each job may be processed by a different sequence of processors. The classes of sequencing problems are n×1 (where n jobs are processed on one processor), n×2, n×3, 2×m, and n×m. Optimal solutions for n×1, n×2, and 2×m problems can be obtained. For n×m problems, heuristics are used.

Multi-Objective scheduling problems have been studied by many researchers such as Dileepan and Sen [4], Fry et al. [5], Wassenbove and Baker [14], and Haral et al. [6] (single machine), Allouche

et al [1] and Rajendran [10] (flow shop), Malakooti et al. [11-12] (multiprocessor), Sheikh [13] (Hybrid Flow shop), Manikas and Chang [15] and Ramesh and Cary [16] (job shop).
In multi-objective scheduling problems specially in bi-objective problems, generally there are two approaches; first one is that one of the criteria is considered as objective function and the rest are considered as constraints, the second approach is that all criteria are considered jointly as objective function, see Dileepan and Sen [4]. In some of them, they have presented exact method, some other suggested heuristic/ metaheuristic algorithms, for example, Arroyo and Armentano [2] suggested Genetic algorithm for multi-objective flow shop, Daniels and Chambers [3] studied bi-objective flow shop and proposed constructive heuristic algorithm. Comprehensive reviews of the multi-objective scheduling can be found in Nagara et al. [7], Hoogeveen [8], Kindt and Billaut [9], and Malakooti [17].
In this paper, we address the multi-objective scheduling problems of n jobs on single machine, flow shop, and jobs shop. Also, we propose a new algorithm called Rule-based method. This algorithm is constructive and uses simple rules to find the sequence of jobs.
The rest of the paper is organized as following. Next section presents the preliminaries and Section 3 is devoted to scheduling on n jobs on 2 processors and Section 4 extends the discussed methods in Section 3 to the m processors. Section 5 covers multi-criteria scheduling of 2 jobs on m machines, and in Section 6 Rule-based Method is proposed. Finally, Section 7 is conclusion and future work.

2. PRELIMINARIES

2.1 Sequencing Performance Measurements

The sequencing problem can be stated as determining a sequence, S, such that one or several sequencing objectives are optimized. The following sequencing criteria can be used to measure the performance of static sequencing problems. It is assumed that the start times for all jobs are set at zero. Therefore, all due dates are defined with respect to the starting time, zero.
Flow Time: The flow time, $F_{i,S}$, for a given job i in sequence S, is its total waiting time, $W_{i,S}$, plus its processing time, t_i.

$$F_{i,S} = W_{i,S} + t_i \quad (2.1)$$

where $W_{i,S}$ is the waiting time for job i in sequence S and t_i is the processing time for job i.
Average Flow Time: The average flow time is the average of the flow times of all n jobs in a given sequence, S.

$$\bar{F}_S = \frac{1}{n}\sum_{i=1}^{n} F_{i,S} \quad (2.2)$$

where n is the number of jobs to be processed. Different sequencing solutions may have different average flow times and different total flow times.
Average Lateness: Lateness for a job is the difference between the completion time of the job and its due date. A job will have positive lateness if it is completed after its due date; it will have a negative lateness if it is completed before its due date. The total of absolute values of latenesses are used to calculate the average lateness. The average lateness of a sequence is always a positive number. The average lateness for a given sequence, S, is:

$$\bar{L}_S = \frac{1}{n}\sum_{i=1}^{n}\left|L_{i,S}\right| \quad \text{or} \quad \bar{L}_S = \frac{1}{n}\sum_{i=1}^{n}\left|\left(F_{i,S} - d_i\right)\right| \tag{2.3}$$

where d_i is the due date for job i and $L_{i,S}$ is lateness of job i in sequence S.

Average Tardiness: Tardiness measures if a job is late with respect to its due date. Jobs completed earlier than their due dates are not tardy, so they are excluded from the calculation. Thus, tardiness is the same as positive lateness. Therefore, tardiness is calculated as the maximum of 0 and lateness. Here "0" signifies that the job was either early or on time. The average tardiness of a given sequence, S, is:

$$\bar{T}_S = \frac{1}{n}\sum_{i=1}^{n}T_{i,S} \quad \text{or} \quad \bar{T}_S = \frac{1}{n}\sum_{i=1}^{n}\mathrm{Max}\left[0, L_{i,S}\right] \quad \text{or} \quad \bar{T}_S = \frac{1}{n}\sum_{i=1}^{n}\mathrm{Max}\left[0, \left(F_{i,S} - d_i\right)\right] \tag{2.4}$$

where $T_{i,S}$ is tardiness of job i in sequence S.

Makespan: Makespan for a given sequence is the total time required to complete all n jobs on given processors. For static systems, the makespan of a sequence, S, is:

$$M_S = F_{n,S} \tag{2.5}$$

where $F_{n,S}$ is the flow time of the last job, n, in the given sequence, S. Note that for n×1 problems, there is only one processor; therefore the processor is utilized all the time, and the makespan is the same as the summation of the processing times, t_i, of all jobs. This is not true for other sequencing problems, i.e., n×2, n×m, and m×2.

Utilization Factor: Utilization refers to how effectively processors are utilized. By maximizing utilization, the amount of idle time in a system is reduced. Utilization for a given sequence, S, is calculated as the total available processing time divided by the total makespan. For m processors:

$$\bar{U}_S = \frac{1}{m^* M_S}\sum_{i=1}^{n}t_i \tag{2.6}$$

For n×1 problems, since there is only one processor, i.e. m=1, the utilization is always 1 regardless of the sequence.

2.2 Bi-Criteria Composite Approach for n×1

Consider the following two objectives:

Minimize average flow time, $f_1 = \bar{F}$

Minimize average tardiness, $f_2 = \bar{T}$

The first objective also minimizes the average waiting time for jobs in the system. The second objective may also minimize the average lateness. For a given additive utility function, the following method can be used to generate the best alternative (see Fry et al. [5]).

$$\text{Minimize } U = w_{SPT}\bar{F} + w_{EDD}\bar{T} \tag{2.7}$$

where w_{SPT} and w_{EDD} are the weights of importance for f_1 and f_2, respectively.

1. Find a sequence, S_{SPT}, which minimizes average flow time. This sequence is obtained by using the shortest processing time (SPT) rule. Now, calculate $\bar{F}$ and $\bar{T}$ for the S_{SPT} sequence.
2. Find a sequence, S_{EDD}, which minimizes average tardiness. Use earliest due date (EDD) rule to find the solution. Calculate $\bar{F}$ and $\bar{T}$ for the S_{EDD} sequence.
3. If the two sequences are the same, i.e. $S_{SPT} = S_{EDD}$, there is only one optimal solution that minimizes both objectives. Stop.
4. Consider the given weights of importance (w_{SPT}, w_{EDD}) for the two objectives, respectively. Apply the following composite MCDM approach to find the sequence of jobs.

Composite MCDM Approach for Sequencing Jobs: This approach assigns a priority index to each job for each objective. The objective weights (w_{SPT}, w_{EDD}) are then used to create a Composite Priority Index for each job.

- First, define the priority increment as 1/(n-1) where n is the number of jobs.
- Assign a job priority index based on the SPT sequence, $P_i(SPT)$, from 0 to 1 to each job. 0 is assigned to the first job in the SPT sequence, and the priority increment, 1/(n-1), is assigned to the second job. Each subsequent job in the sequence is given a priority index that is 1/(n-1) higher than the previous job until the last job in the sequence is assigned a value of 1.
- Assign job priorities for the EDD sequence, $P_i(EDD)$, in a similar manner.
- Use weights of importance for the objectives (w_{SPT}, w_{EDD}) to find the Composite Priorities Index, Z_i, for each job:

$$Z_i = w_{SPT}P_i(SPT) + w_{EDD}P_i(EDD) \quad (2.8)$$

- Sequence the jobs in ascending order of their composite priority index, Z_i.

Example 1: Bi-Criteria Analysis for Sequencing n Jobs by 1 Processor

Consider sequencing four jobs by one processor. Each job has the following processing times and due dates in days. The starting time is zero for all jobs.

Job i	1	2	3	4
Process time t_i	4	9	5	2
Due date d_i	18	10	7	26

Suppose the weights of importance for the two objectives are (0.6, 0.4) for f_1 (average flow time) and f_2 (average tardiness), respectively. The solution using the SPT method to minimize f_1 is:

$S_{SPT} = (4\rightarrow1\rightarrow3\rightarrow2)$ where $f_1 = \bar{F} = 9.75$ and $f_2 = \bar{T} = 3.5$

The solution using the EDD method to minimize f_2 is:

$S_{EDD} = (3\rightarrow2\rightarrow1\rightarrow4)$ where $f_1 = \bar{F} = 14.25$ and $f_2 = \bar{T} = 1$

The above two solutions are not the same; therefore find the Composite Priority Index, Z_i, for each job, and sequence the jobs using this index.

Composite Priority Index: The number of jobs in this problem is four (n = 4). Therefore, the priority increment is 1/(n-1) = 1/(4-1) = 1/3. The Priority Indices for SPT and EDD for each job are presented in the following tables. The Composite Priority Index is then found by combining the two objectives where $Z_i = 0.6P_i(SPT) + 0.4P_i(EDD)$

SPT Sequence				
Job i	4	1	3	2
$P_i(SPT)$	0	1/3	2/3	1

\+

EDD Sequence				
Job i	3	2	1	4
$P_i(EDD)$	0	1/3	2/3	1

⇨

Composite Priority Index				
Job i	1	2	3	4
Z_i	0.47	0.73	0.4	0.4
Rank	3	4	1 or 2	1 or 2

In this example, there are two alternative solutions because the composite priority index for jobs 3 and 4 are the same. The two solutions are $(3\rightarrow4\rightarrow1\rightarrow2)$ and $(4\rightarrow3\rightarrow1\rightarrow2)$. When there are multiple sequence solutions, calculate the average flow time, f_1, and the average tardiness, f_2, for each of the sequences. Choose the sequence that best minimizes the additive utility function $U = w_1f_1 + w_2f_2$.

For example, for sequence $3\rightarrow4\rightarrow1\rightarrow2$, the solution is:

Job i	t_i	W_i	F_i	d_i	$\lvert L_i\rvert$	T_i

Objectives

3	5	0	5	7	2	0
4	2	5	7	26	19	0
1	4	7	11	18	7	0
2	9	11	20	10	10	10
Total	20	23	43		38	10

$f_1 = \bar{F} =$	10.75
$f_2 = \bar{T} =$	2.5

For sequence (4→3→1→2), the solution is the same as the Weighted SPT example solution (see Example 2.1), where $f_1 = \bar{F} = 10.00$ and $f_2 = \bar{T} = 2.5$.
Since the sequence (4→3→1→2) dominates the sequence (3→4→1→2), i.e. it has a shorter average flow time ($\bar{F}$), choose (4→3→1→2) as the best sequence for this bi-criteria problem.

2.3 GENERATING THE EFFICIENT FRONTIER FOR n×1

For small sequencing problems, such as Example 2.1, it is possible to generate all possible sequences and then identify the efficient set of alternatives. The set of efficient alternatives for Example 2.1 is shown in Table 2.1 and Figure 2.1.

Alternative	a_1	a_2	a_3	a_4	a_5
Min. f_1	9.75	10	11.25	13.75	14.25
Min. f_2	3.5	2.5	2	1.5	1.0
Sequence	4→1→3→2	4→3→1→2	4→3→2→1	3→2→4→1	3→2→1→4
Efficient?	Yes	Yes	Yes	Yes	Yes

Table 2.1: Set of efficient alternatives for Example 2.1

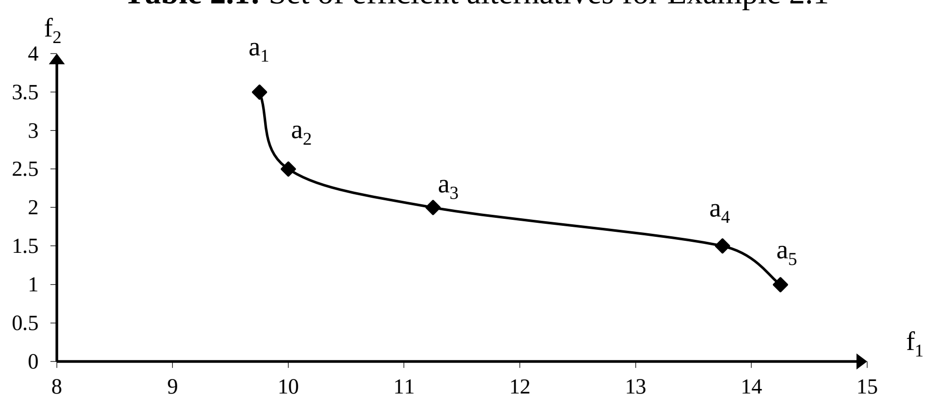


Figure 2.1: Efficient frontier for Example 2.1 where both objectives are minimized

Generating Efficient Alternatives Using Composite MCDM: For larger sequencing problems, generating all possible sequences is not feasible. The Composite MCDM method can be used to generate a set of efficient alternatives. This can be accomplished by varying the weights w_{SPT} and w_{EDD} in the Equation (2.8).
For example, consider five sets of weights (w_{SPT}, w_{EDD}): (0, 1), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25), and (1, 0). Solving five different Composite MCDM sequencing problems may result in five efficient alternatives. If the two sets of adjacent weights generate the same sequence, then consider using some other set of weights that is not between the two adjacent set of weights.

Example 2: Efficient Frontiers for Sequencing n Jobs by 1 Processor

Consider Example 2.1. Generate five efficient points using five different weights. Using weights **w** = (0.5, 0.5) for (w_{SPT}, w_{EDD}), the solution is:

Job i	1	2	3	4
Z_i	0.50	0.67	0.33	0.50
Rank	2	4	1	2

Sequence	3→4→1→2
f_1	10.75
f_2	2.50

For the weights **w** = (0.75, 0.25), the solution is:

Job i	1	2	3	4
Z_i	0.42	0.83	0.50	0.25
Rank	2	4	3	1

Sequence	4→1→3→2
f_1	9.75
f_2	3.50

For the weights **w** = (0.25, 0.75), the MCDM solution is:

Job i	1	2	3	4
Z_i	0.58	0.50	0.17	0.75
Rank	3	2	1	4

Sequence	3→2→1→4
f_1	14.25
f_2	1.00

Table 2.3 shows the five sequences generated by using five sets of weights. The weights (1,0) and (0.75,0.25) generated the same solution. Therefore consider using a different set of weights, e.g. use (0.65, 0.35) which is between (0.75, 0.25) and (0.5, 0.5). This generates a new solution as presented in Table 2.3. Note that only alternatives a_1, a_2, and a_5 are efficient. To generate more efficient alternatives, one can improve upon the generated inefficient points by using the pair-wise exchange method (which is discussed in the Cellular Layouts and Networks Chapter). For example, consider alternative a_7, 3→1→4→2. By exchanging 1 and 2, a new efficient solution a_4, 3→2→4→1, is found.

Weights	(1, 0) or (0.75, 0.25)	(0.65,0.35)	(0.5, 0.5)	(0.4, 0.6)	(0.25, 0.75) or (0, 1)
Alternative	a_1	a_2	a_6	a_7	a_5
Min. f_1	9.75	10	10.75	11.25	14.25
Min. f_2	3.5	2.5	2.5	2.5	1.0
Sequence	4→1→3→2	4→3→1→2	3→4→1→2	3→1→4→2	3→2→1→4
Efficient?	Yes	Yes	No	No	Yes

Table 2.2: Composite MCDM approach for Example 2.3

2.4 Tri-Criteria for n×1

Tri-criteria sequencing problems can be formulated as:

f_1 = Minimize the average flow time ($\bar{F}$), use SPT Rule (minimizes avg. waiting)

f_2 = Minimize the average tardiness ($\bar{T}$), use EDD Rule (customer's priority)

f_3 = Maximize satisfying job priorities, use Job Priority Rule (producer's priority)

f_1 and f_2 are measured by $\bar{F}$ and $\bar{T}$ respectively. In the following, we present how to measure f_3.

Job Priority Rule (PRI): Suppose the producer's priority or goal sequence is S_g. Let a given current sequence solution be labeled as S_c. The PRI objective, f_3, seeks to minimize the difference between the current sequence, S_c, and the goal sequence, S_g. f_3 is measured by the Rectilinear distance between the goal sequence and the current sequence.

$$\text{Minimize } f_3 = \sum_{i=1}^{n} | c_i - g_i | \tag{2.9}$$

Where c_i is the position (rank) of job i in the current sequence, and g_i is the position (rank) of job i in the goal sequence. For example, suppose the current job sequence is (4→1→3→2) and the goal (priority) sequence is (1→2→3→4). f_3 is calculated as presented below.

Job i	1	2	3	4	f_3
Current position, c_i	2	4	3	1	
Goal position, g_i	1	2	3	4	
$\|c_i - g_i\|$	1	2	0	3	6

Note that if the current sequence is the same as the goal sequence, then, $f_3 = 0$; this is the best possible (ideal) solution for job priorities.

The approach for generating efficient alternatives for tri-criteria sequencing problems is similar to the approach of bi-criteria problem. For given weights of importance for the three objectives, f_1, f_2, and f_3, use the composite MCDM approach for all three objectives.

Example 3: Tri-Criteria n×1

Consider Example 2.2. For the third objective, suppose the goal job priority sequence is 1→2→3→4. For the additive utility function, if the weights of importance for the three objectives, f_1, f_2, and f_3, are (0.3, 0.3, 0.4), respectively, find the solution to the MCDM problem. First, find $Z_i = 0.3P_i(SPT) + 0.3P_i(EDD) + 0.4P_i(PRI)$.

For SPT, the sequence is (4→1→3→2) with SPT job priority indices of (0, 1/3, 2/3, 1) for jobs 4, 1, 3, and 2, respectively.

For EDD, the sequence is (3→2→1→4) with EDD job priority indices of (0, 1/3, 2/3, 1) for jobs 3, 2, 1, and 4, respectively.

For PRI, the sequence is (1→2→3→4) with PRI job priority indices of (0, 1/3, 2/3, 1) for jobs 1, 2, 3, and 4, respectively.

Therefore, the composite priority index, Z_i, can be found as follows:

Job i	1	2	3	4
$0.3P_i(SPT)$ $0.3P_i(EDD)$ $0.4P_i(PRI)$	0.3*1/3 + 0.3*2/3 + 0.4*0 =	0.3*1 + 0.3*1/3 + 0.4*1/3 =	0.3*2/3 + 0.3*0 + 0.4*2/3 =	0.3*0 + 0.3*1 + 0.4*1 =
Z_i (sum)	0.30	0.53	0.47	0.70
Rank	1	3	2	4

Therefore, the sequence solution is 1→3→2→4. This solution is presented below where f_3 is measured as |1-1| + |2-3| + |3-2| + |4-4| = 2.

Job i	t_i	W_i	F_i	d_i	$\|L_i\|$	T_i
1	4	0	4	18	14	0
3	5	4	9	7	2	2
2	9	9	18	10	8	8
4	2	18	20	26	6	0
Total	20	31	51		30	10

Objectives	
$f_1 = \bar{F}$ = 51/4 =	12.75
$f_2 = \bar{T}$ = 10/4 =	2.5
f_3 =	2

3. SEQUENCING n JOBS BY 2 PROCESSORS

3.1 Multi-Criteria of n×2

Bi-criteria n×2 sequencing problem is:

Minimize f_1 =Makespan

Minimize f_2 =Total tardiness

The best sequence for f_1 can be obtained by using Johnson's Rule. The best sequence for f_2 can be found using the earliest due date (EDD) rule. The MCDM sequencing of n jobs by 2 processors is similar to the MCDM sequencing of n jobs by 1 processor covered in Section 2.3. For n×2 problems, makespan is used instead of average flow time which was used in n×1 problems, because in n×1 problems makespan is constant but in n×2 problems this value can be optimized. The following examples show the use of composite MCDM approach for sequencing n jobs by 2 processors, and also how to generate a sample of efficient alternatives.

Example 4: Bi-Criteria of Sequencing 4 Jobs by 2 Processors

Consider two processors (A→B) having the following due dates, d_i (in days). Find the best sequence for the bi-criteria problem, where the weights of importance for objectives f_1 and f_2 are (0.6, 0.4), respectively.

Job i	1	2	3	4
Processor A processing time	4	9	5	2
Processor B processing time	3	1	7	6
Due date for job i, d_i	8	14	20	12

Minimum Makespan: The optimal sequence using Johnson's Rule is (4→3→1→2) with a makespan of 21.

Minimum Tardiness: Using the earliest due date rule, the optimal sequence is (1→ 4→2→3).

The above two solutions are not the same; therefore proceed with finding the composite priority index.

Minimum Makespan: (4→3→1→2)

Job i	A_i	B_i	F_i	d_i	$\lvert L_i \rvert$	T_i
4	2	6	8	12	4	0
3	5	7	15	20	5	0
1	4	3	18	8	10	10
2	9	1	21	14	7	7
Objectives			f_1=21			f_2=17

Minimum Tardiness: (1→ 4→2→3)

Job i	A_i	B_i	F_i	d_i	$\lvert L_i \rvert$	T_i
1	4	3	7	8	1	0
4	2	6	13	12	1	1
2	9	1	16	14	2	2
3	5	7	27	20	7	7
Objectives			f_1=27			f_2=10

Composite Priority Index: The number of jobs in this example is four, i.e. n = 4. Therefore priority increment is 1/(n-1) = 1/(4-1) = 1/3.

SPT sequence (Johnson's Rule)				
Job i	4	3	1	2
Priority P_i(SPT)	0	1/3	2/3	1

EDD sequence				
Job i	1	4	2	3
Priority P_i(EDD)	0	1/3	2/3	1

Now find $Z_i = 0.6P_i(SPT) + 0.4P_i(EDD)$. For example, for job 1, $Z_1 = 0.6(2/3) + 0.4(0) = 0.4$.

Job i	1	2	3	4
Z_i	0.4	0.87	0.6	0.13
Rank	2	4	3	1

Therefore, for the bi-criteria problem, (4→1→3→2) is the best sequence. For this sequence, the objective values are:

Job i	A_i	B_i	F_i	d_i	$\lvert L_i \rvert$	T_i
4	2	6	8	12	4	0
1	4	3	11	8	3	3

3	5	7	18	20	2	0
2	9	1	21	14	7	7
Objectives			f_1=21			f_2=10

Since both objectives are minimized, alternative (4→1→3→2) with (f_1 = 21, f_2 =10) dominates the solutions of both SPT (21, 17) and EDD (27, 10).

By using the pair-wise exchange method (as discussed in the Cellular Layout and Networks Chapter), the current alternative (4→1→3→2) may be improved. By exchanging the positions of jobs 4 and 1, sequence (1→4→3→2) can be generated with objective values of f_1 = 21 and f_2 = 8. This solution, (1→4→3→2), dominates the previous solution, (4→1→3→2).

4. MULTI-CRITERIA SEQUENCING n JOBS BY m PROCESSORS

The MCDM of n×m problems is similar to the MCDM of n×2 problems.

Example 5: Bi-Criteria Sequencing Four Jobs by Three Processors

The following table shows the processing times (in days) for four jobs on three processors. The processing order is A→B→C. Find the best sequence and verify whether the solution is optimal.

Job i	1	2	3	4	Min.	Max.
Processor A	10	6	7	10	6	NA
Processor B	3	1	5	6	NA	6
Processor C	8	8	9	9	8	NA
d_i	18	35	22	28		

Minimum Makespan: Using Johnson's algorithm, the optimal sequence for minimizing the makespan is 2→3→4→1. For this sequence, the makespan is 46 and the total tardiness is 43.

Minimum Total Tardiness: According to the earliest due date rule, the sequence is (1→3→4→2). The solution for this sequence is presented below. The makespan is 50 days and the total tardiness is 41 days.

Minimum Makespan: (2→3→4→1)

Job i	F_i	d_i	T_i
2	15	35	0
3	27	22	5
4	38	28	10
1	46	18	28
Objectives	f_1 = 46		f_2 = 43

Minimum Tardiness: (1→3→4→2)

Job i	F_i	d_i	T_i
1	21	18	3
3	31	22	9
4	42	28	14
2	50	35	15
Objectives	f_1 = 50		f_2 = 41

Suppose the weights of importance for the two objectives are (0.5, 0.5), respectively. Using the Composite Priority Index, the following solution is generated:

Job i	1	2	3	4
Z_i	0.50	0.50	0.33	0.67
Rank	2	2	1	4

The best sequence is either 3→1→2→4 or 3→2→1→4. Since sequence 3→1→2→4 has f_1 = 48 and f_2 = 33, but sequence 3→2→1→4 has f_1 = 48 and f_2 = 39, choose sequence 3→1→2→4 as the best solution. This solution dominates (i.e. it is better in terms of both objectives) the EDD solution.

5. MULTI-CRITERIA JOB SHOP OF 2 JOBS BY m PROCESSORS

Consider minimizing makespan as f_1 and minimizing total tardiness as f_2. The MCDM analysis of two jobs by m processors is simple. There are only two possible efficient solutions to this problem, i.e. either job 1 is finished first or job 2 is finished first. To generate the two solutions, first find the solution associated with the longest diagonal line such that job 1 is finished first; then find the solution associated with the longest diagonal line such that job 2 is finished first.

Example 6: MCDM of Sequencing 2 Jobs by m Processors

Consider the following processing times for two jobs by four processors. Job 1 has processor sequence C→A→D→B and job 2 has processor sequence B→A→C→D. Find the best job sequece for this problem and show the start and finish of each job on each processor. Suppose job 1 is due in 15 days and job 2 is due in 10 days.

Processor	A	B	C	D	Total	Sequence
Job 1	3	1	2	6	12	C→A→D→B
Job 2	2	3	2	3	10	B→A→C→D

To minimize f_1, makespan, solve the problem by the 2×m graphical algorithm (see the solution to Example 5.1). The minimum makespan is 14 days. Job 1 is finished first. Now, measure f_2, the total tardiness for this sequence, $f_2 = 4$ (see the following tables).

To minimize f_2, find the optimal path (by trial and error) such that job 2 is finished as close or before its due date, and then job 1 is finished as close or before its due date. This solution is shown in Figure 5.1. For this solution, the total tardiness is $f_2 = 2$ and the makespan is $f_1 = 17$. Both alternatives are efficient with respect to each other. The two alternatives can then be evaluated by the decision maker so that the best alternative can be selected.

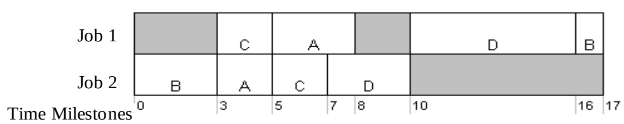


Fig. 5.1. Optimal solution of the problem

Alt 1: Job 1 Finished First			
Job i	F_i	d_i	T_i
1	12	15	0
2	14	10	4
Object:	f_1=14		f_2=4

Alt 2: Job 2 Finished First			
Job i	F_i	d_i	T_i
2	10	10	0
1	17	15	2
Object:	f_1=17		f_2=2

MCDM Alternative		
	Alt 1	Alt 2
Min. f_1	14.00	17.00
Min. f_2	4.00	2.00

6. Sequencing of n×m: Rule-based Approach

6.1 Overview of Rule-based Approach

In this section, we develop a new heuristic for solving n×m sequencing problems. The approach is based on the Rule-Based method or Head-Tail method. To solve the n×m sequencing problem, first convert the n×m problem into a series of 2×m problems; then solve each 2×m problem by the graphical method, see Malakooti [2013]. Then generate an n×n flow matrix which will be used by the Rule-based method.

2×m Composite Formulation of n×m Problems: Consider job i versus job j. Solve this problem by the 2×m graphical method, record the solution, and identify which job is completed first.

Suppose that job i is completed first. Now solve the 2×m problem such that job j is completed first, and job i is completed second. Record both these solutions and their makespans. These solutions correspond to (i, j) and (j, i) in the flow matrix, respectively. Note that it is not important which job is started first, but it is important to know which job is completed first.

Steps of Rule-based Methods: By using a set of simple rules, this method progressively builds the solution by connecting pairs of jobs. The following notations are used for connecting pairs of sequences of jobs.

$s_p = (i \rightarrow j)$ is a candidate pair sequence where job i is completed before job j.

$s_q = (\rightarrow \ \rightarrow \ \rightarrow \ldots)$ is a partially constructed sequence solution so far.

1. Consider the ranking of all pairs of jobs (i, j). Rank pairs based on their makespans from the shortest to the longest makespan.
2. Consider the highest ranked unassigned pair, $s_p = (i , j)$, such that
   a. Both i and j are not in s_q
   b. Either i or j is in s_q.
3. Use relevant Head-Tail Rules to combine s_p and s_q.
4. Repeat steps 2 and 3 until all jobs are assigned to s_q.
5. Try to improve the obtained final solution by using the Pair-Wise Exchange method.

Uni-Directional Cardinal Rule (Conflict Resolution): If there is a common job, i, in $s_p = (i \rightarrow j)$ and $s_q = (\ldots \rightarrow i \rightarrow \ldots)$, then consider inserting j (the uncommon job) between each pair of jobs in in s_q. Choose the alternative that has the best objective function value (e.g. the minimum makespan). For example, if $s_p = (2 \rightarrow 5)$ and $s_q = (1 \rightarrow 2 \rightarrow 4)$, the uncommon job is 5. The common job is 2. Consider all possible locations that job 5 can be inserted in sequence s_q. They are: $(5 \rightarrow 1 \rightarrow 2 \rightarrow 4)$, $(1 \rightarrow 5 \rightarrow 2 \rightarrow 4)$, $(1 \rightarrow 2 \rightarrow 5 \rightarrow 4)$, and $(1 \rightarrow 2 \rightarrow 4 \rightarrow 5)$. Choose the sequence that has the lowest makespan as the new s_q.

6.2 Multi-Criteria of n×m

This method is the same as the multi-criteria n×m method in Section 4. The MCDM approach for n×m is based on the MCDM approach discussed for n×2 MCDM problems. Note that because heuristics are used in solving nxm problems, there is no guarantee that the generated solutions are efficient.

Example 7: MCDM of Flow-Shop Sequencing 4 Jobs by 4 Processors

Consider the problem of sequencing four jobs by four processors (A, B, C, and D) as presented in the following table. All jobs are processed in the order (A→B→C→D).

Job i	A	B	C	D	Total
1	4	1	3	2	10
2	1	6	5	3	15
3	5	2	4	1	12
4	2	4	1	5	12
Total	12	13	13	11	49

Due dates of the jobs are given in the following table. Find the sequencing associated with the weights of importance of (0.5, 0.5) for the two objectives.

Job i	1	2	3	4
d_i	15	25	13	22

Minimum Makespan: The best sequence for minimizing the makespan is 2→4→1→3. For this solution, the makespan is 23 days and total tardiness is 17 days.

Minimum Tardiness: According to the earliest due date rule, the sequence was found to be 3→1→4→2. For this solution, the makespan is 29 days and total tardiness is 5 days.

Job i	F_i	d_i	T_i
2	15	25	0
4	20	22	0
1	22	15	7
3	23	13	10
Objectives	f_1 =23		f_2=17

First Alternative: (2→4→1→3)

Job i	F_i	d_i	T_i
3	12	13	0
1	16	15	1
4	21	22	0
2	29	25	4
Objectives	f_1=29		f_2=5

Second Alternative: (3→1→4→2)

For bi-criteria problems, there may exist a set of efficient alternatives that range between the two extreme alternatives. Now suppose the weights of importance for the two objectives are (0.5, 0.5). The Composite Priority Index is: $Z_i = 0.5P_i(SPT)+0.5P_i(EDD)$.

Job i	1	2	3	4
Z_i	0.33	0.17	0.83	0.67
Rank	2	1	4	3

For the sequence (2→1→4→3), f_1 is 24 and f_2 is 12.

7. Conclusion

Multi-objective scheduling of n jobs on single machine, flow shop, and job shop is addressed in this paper. A new constructive heuristic algorithm called Rule-based method is developed to solve the multi-objective scheduling problems. This algorithm decomposed the problem into the 2×m method which can be solved easily by graphical method. The pair of jobs with the minimum objective function is selected and the next jobs are inserted in all possible positions of the current sequence of jobs until sequence of all jobs are found. This method is effective and its computational effort is much less than conventional methods. This algorithm can be applied to other area such as layout planning.

References:

[1]Allouche, M. A., B. Aouni, J.-M. Martel, T. Loukil, and A. Rebai “Solving Multi-Criteria Scheduling Flow Shop Problem Through Compromise Programming and Satisfaction Functions.” European Journal of Operational Research, Vol. 192, Issue 2, 2009, pp. 460-467.

[2]Arroyo, J. E. C., V. A. Armentano, “Genetic Local Search for Multi-Objective Flowshop Scheduling Problems.” European Journal of Operational Research, Vol. 167, Issue 3, 2005, pp. 717-738

[3]Daniels, R.L., R.J. Chambers, “Multiobjective Flow-Shop Scheduling.” Naval Research Logistics, Vol. 40, 1990, pp. 85–101.

[4]Dileepan, P., T. Sen, "Bicriterion Static Scheduling Research for a Single Machine." Omega, Vol. 16, Issue 1, 1988,

pp. 53-59.
[5]Fry, T.D., R.D. Armstrong, and H. Lewis, "A Framework for Single Machine Multiple Objective Sequencing Research." Omega, Vol. 17, Issue 6, 1989, pp. 595-607.
[6]Haral, U., R. W. Chen, W. G. Ferrell, Jr., and M. B. Kurz, “Multiobjective Single Machine Scheduling with Nontraditional Requirements.” International Journal of Production Economics, Vol. 106, Issue 2, 2007, pp. 574-584.
[7]Nagara, A., J. Haddocka, and S. Heragu “Multiple and Bicriteria Scheduling: A Literature Survey.” European Journal of Operational Research, Vol. 81, Issue 1, 1995, pp. 88-104.
[8]Hoogeveen, H. “Multicriteria Scheduling.” European Journal of Operational Research, Vol. 167, Issue 3, 2005, pp. 592–623.
[9]Kindt, T. V., J.-C. Billaut, Multicriteria Scheduling: Theory, Models and Algorithms, Springer, Berlin, 2002.
[10]Rajendran, C. ”Two-Stage Flow Shop Scheduling Problem with Bicriteria.” Journal of the Operational Research Society, Vol. 43, 1992, pp. 871–884.
[11]Malakooti, B., S. Sheikh, C. Al-Najjar, H. Kim “Multi-Objective Energy Aware Multiprocessor Scheduling Using Bat Intelligence”, Journal of Intelligent Manufacturing, 24.4 (2013): 805-819..
[12]Malakooti, B., H. Kim, S. Sheikh “Bat Intelligence Search with Application to Multi-Objective Multi-Processor Scheduling Optimization”, International Journal of Advanced Manufacturing Technology, Vol. 60, Issue 9, 2012, pp. 1071-1086.
[13]Sheikh, S. “Multi-Objective Flexible Flow Lines with Due Window, Time Lag and Job Rejection” International Journal of Advanced Manufacturing Technology, 2012, DOI: 10.1007/s00170-012-4112-5.
[14]Wassenbove, L.N.V., K. R. Baker, "A Bicriterion Approach to Time/Cost Trade-Offs in Sequencing." European Journal of Operational Research, Vol. 11, 1982, pp. 48-54.
[15] Manikas, A., & Chang, Y. L.. Multi-criteria sequence-dependent job shop scheduling using genetic algorithms. *Computers & Industrial Engineering*, *56*(1) (2009), 179-185.
[16]Ramesh, R., & Cary, J. M. Multicriteria jobshop scheduling. *Computers & industrial engineering*, *17*(1), (1989). 597-602.
[17] Malakooti, Behnam. *Operations and Production Systems with Multiple Objectives*. John Wiley & Sons, 2014.

Rule Based Layout Planning and Its Multiple Objectives

Mohammad Komaki, Shaya Sheikh , Behnam Malakooti

Case Western Reserve University
Systems Engineering
Email: komakighorban@gmail.com

Abstract

In this paper, we present a novel constructive approach for solving facility layout problems considering objectives of total flow and qualitative closeness. This approach, called Head-Tail, utilizes a set of rules to solve single and multiple row layout problems. We apply pair-wise exchange method to improve the partially constructed layouts. The computational efficiency of Head-Tail approach is examined by comparing it with Systematic Layout Planning. The results show that Head-Tail method can find efficient solution with promising quality. A bi-criteria numerical example that minimizes total flow and total qualitative closeness is also presented.

Key words: Rule-based layout planning, multi-objective optimization, Head-Tail Method

1. Introduction

Effective layout designs are necessary in many different businesses such as manufacturing facilities, and service industries. Diminishing physical space availability, high cost of re-layouts, safety issues, and substantial long-term investments are good reasons for getting the design right the first time. Different criteria can be considered in facility layout design including operational costs, installation costs, utilization of space and equipment, work-force management, materials handling, inventory work in process, inventory of incoming materials and inventory of outgoing final products, ease of future expansion, and safety factors. Most of these objectives are difficult to quantify. In this paper, new methods are introduced for solving multiple objective layout problems. Researchers have used many criteria to come up with the optimal layout design. Raman et al. [1] developed three factors of flexibility, productive area utilization and closeness gap to measure the layout effectiveness. Rosenblatt [2] discussed a facility layout problem with two criteria of minimizing costs and maximizing closeness rating. Fortenberry and Cox [3] discussed a multi-criteria facility layout problem that combines quantitative with qualitative criteria and minimizes total weighted work-flow. Malakooti [4,5,6] presented heuristic methods for solving multi-objective facility problems.. Malakooti and Tsurushima [7] proposed an expert system for solving multi-crtieria manufacturing problems. Islier [8] developed a genetic algorithm for solving multi objective facility layout by minimizing transportation load, maximizing compactness of departments and minimizing the difference between requested and available area. Pelinescu and Wang [9] presented a hierarchical multi criteria optimal design of fixture layout. Yang and Kuo [10] proposed a hierarchical AHP and data envelopment analysis approach for solving layout design problem. Chen and Sha [11,12] developed a five-phase heuristic method using paired comparison matrix for solving multi-objective facility layout problem. Sahin and Ramazan [13] and Singh & Singh [14] have used heuristic method for solving multi-criteria layout problems. In this paper, we develop a new rule based constructive approach called Head-Tail method. We use this method for solving bi-criteria layout optimization problems.

2. Head-Tail Layout

Head-Tail approach starts with the pair of departments that have the highest flow as the initial partial layout. More departments are progressively added to the existing partial layout until all departments are assigned. The problem is to find the order of departments that minimizes the total flow as measured in Equation (1). The selection of the location of each department is based on qualitative or quantitative rules and other pertinent information. We call $s_p = (i \leftrightarrow j)$ as a candidate pair of departments and $s_q = (\leftrightarrow \ \leftrightarrow \ \ \leftrightarrow \ldots)$ as a partially constructed sequence.

Generic Steps and Rules of Head-Tail Method

1. Consider the ranking of all pairs (i, j). Set the highest ranked pair as s_q.
2. Consider the highest ranked unassigned pair, $s_p = (i , j)$, such that a. Both i and j are not in s_q b. Either i or j is in s_q.
3. Use relevant Head-Tail Rules to combine s_p and s_q.
4. Repeat steps 2 and 3 until all stations are assigned to s_q.
5. Try to improve upon the obtained solution by use of the Pair-Wise Exchange method.

Temporarily Discarding Unassigned Pairs: If a pair $s_p = (i,j)$ does not have any common element with $s_q = (\ldots\ldots\ldots)$, then $s_p = (i,j)$ is discarded temporarily only in the current iteration, but it is reconsidered in the next iteration.

Permanently Discarding Assigned Pairs: If a pair $s_p = (i, j)$ has two common elements with $s_q = (\ldots\ i.., j\ldots)$, then $s_p = (i,j)$ is discarded permanently (because both i and j are already assigned).

Consider a distance matrix, D, consisting of elements d_{ij} where d_{ij} is the distance between pairs of departments i and j. Also consider a flow matrix, W, where each element (w_{ij}) represents the flow from department i to department j. For every pair of departments (i, j) where $i<j$, create a new w_{ij} equal to (w_{ij} + w_{ji}). These values are used in calculating the total flow. For a problem with n departments, there will be n locations. The distances can be presented by a symmetric flow matrix; that is, the distance from department i to department j, d_{ij}, is equal to the distance from department j to location i, d_{ji}. The total flow for a given sequence solution is:

$$TF = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} d_{ij} \qquad (1)$$

Rules for Single Row Layout

Order all pairs (i, j) in descending order of w_{ij} in a flow Matrix. Label the ordered vectors of w_{ij} by $\mathbf{w}_o$. Consider $s_p = (i \leftrightarrow j)$, $s_q = (\leftrightarrow \ \leftrightarrow \ \ \leftrightarrow \ldots)$, and $\mathbf{w}_o$ = Vector of w_{ij} in descending order of w_{ij}.

Conflict Resolution Rule: If there is a common department in sequences s_p and s_q, i.e. $s_p = (i\leftrightarrow j)$ and $s_q = (\ldots n\leftrightarrow\ i\ \leftrightarrow m\ldots)$, then place the new department, j, next to the common department, i, in s_q on the side that has a higher ranking to its adjacent department, i.e. compare $j\leftrightarrow n$ and $j\leftrightarrow m$ flows and insert j on the side that has higher flow.

Linear Mirror Property: The mirror image of sequence s_q, i.e. its reverse sequence, is equivalent to s_q. Both sequences will have the same total flow. In the following, we show how to solve a single-row layout using Head-Tail.

Example 1: Single-Row Layout

Consider the information given below. Suppose the layout should be in the form of a 1x6 rectangle.

Locations

1	2	3	4	5	6

Distance D

	1	2	3	4	5	6
1	-	1	2	3	4	5
2		-	1	2	3	4
3			-	1	2	3
4				-	1	2
5					-	1
6						-

Flows V

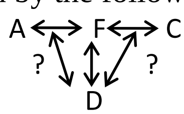
	A	B	C	D	E	F
A	-	10	20	30	10	100
B		-	50	60	10	20
C			-	10	20	80
D				-	50	60
E					-	10
F						-

Ranking of Flows V

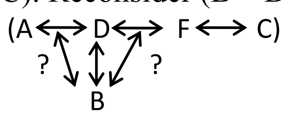
	A	B	C	D	E	F
A	-	11	8	7	11	1
B		-	5	3	11	8
C			-	11	8	2
D				-	5	3
E					-	11
F						-

Figure 1: Initial Data for Example 1

1. s_p = (A↔F); then set s_q = (A↔F)
2. s_p = (C↔F); The mirror image is s_p = (F↔C). Therefore, s_q = (A↔F↔C).
3. Discard (B↔D) temporarily as it does not have a common element with s_q.
4. s_p = (D↔F); there is conflict for location of D, use the Conflict Resolution Rule to find the location for D. This is shown by the following graphical representation.

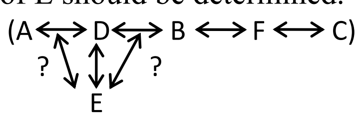


Since the flow of (D,A) = 30 is more than the flow of (D,C) = 10, locate D between A and F; therefore, s_q = (A↔D↔F↔C). Reconsider (B↔D), the location of B should be decided.



The flow of (B, F) = 20 is more than the flow of (B,A) = 10; locate B on the right side of D. Therefore, s_q = (A↔D↔B↔F↔C).

5. Consider (B↔C), discard it because it is a duplicate of elements in the sequence.
6. Consider (D↔E), the location of E should be determined.



The flow of (E,A) = 10 is equal to the flow of (E,B) = 10, therefore E can be put on either side of D; arbitrarily, locate it on the right side of D. Therefore, s_q = (A↔D↔E↔B↔F↔C).

The final layout is (A↔D↔E↔B↔F↔C). Using the above given distances and Equation 1, the total flow of this layout is Total Flow = (10*3) + (20*5) + (30*1) + (10*2) + (100*4) + (50*2) + (60*2) + (10*1) + (20*1) + (10*4) + (20*3) + (80*1) + (50*1) + (60*3) + (10*2) = 1260

Rules for Multi-Row Layout

We can extend the single row Head-Tail method to solve multiple row layout problems. The single row approach expands horizontally. The multi-row layout approach uses all rules of the single row

approach, but allows both horizontal and vertical expansions.
The following rules show how to progressively augment s_q by combining existing $s_q = (\ldots \leftrightarrow i \leftrightarrow \ldots)$, the partially built layout, with a new $s_p = (i \leftrightarrow j)$; i.e. find a location for department j. By use of the following rules, all departments will be assigned and a complete solution will be obtained.

Expanding Phase

Rule 1 (Attaching Sequences, Horizontal and Vertical)
a) Horizontal: If there is one common department in the pair of sequences s_p & s_q, such that the common department is either at the head or the tail of sequence s_p and s_q, then form a new sequence by attaching s_p to s_q by their common department.
For example, given $s_q = (C \leftrightarrow B \leftrightarrow E)$ and $s_p = (A \leftrightarrow C)$, new sequence is $s_q = (A \leftrightarrow C \leftrightarrow B \leftrightarrow E)$.
b) Vertical: A sequence can be attached in the middle of an existing sequence at a 90° angle if the new sequence has one common department with the existing sequence.
For example, adding the sequence (B↔F) to the sequence (A↔B↔C↔D) in Figure 2a results in the sequence in Figure 2b.
Rule 2 (Attaching Sequences, Diagonal): If there is a common department in sequences s_p and s_q, e.g. $s_p = (i \leftrightarrow j)$ and $s_q = (\ldots \leftrightarrow i \leftrightarrow \ldots)$, and Rule 1 fails, i.e. department j cannot be placed next to i in s_q, then insert it at a 45° angle on the side which is available and has the highest ranking priority with respect to its adjacent departments.
Rule 3 (Extending): Keep the existing sequence as open (extended) as possible by stretching all connected lines as straight as possible.
That is, add the new sequence to the old sequence in the form of a straight line if possible. For example, adding (E↔F) to the sequence in Figure 2b results in the sequence in Figure 2c.

Reshaping Phase

Rule 4 (Rotation): To fit the partial constructed layout tree into the given layout locations, if possible, use vertical, horizontal, or diagonal rotation of branches while satisfying closeness preferences. In Example 2, in Figure 5c, E is rotated around D resulting in Figure 5d to allow B to fold down to C in Figure 5e.
Rule 5 (Folding): To fit the partially constructed layout tree into the given layout locations, if possible, bend the branches of the layout tree while satisfying closeness preferences.
For example, suppose the given layout is 2x3 where the given constructed sequence is shown in Figure 2c. Consider folding branch B-F-E. It can be folded to the right or left. Because the ranking of (E,C) is higher than (E,A), fold the branch to the right connecting E to C as shown in Figure 2d. Figure 2d uses the folding rule to fit into the given layout space. To do this, bend A and D upwards and connect them as shown in Figure 2e. This is the final layout which now fits in the 2x3 space.

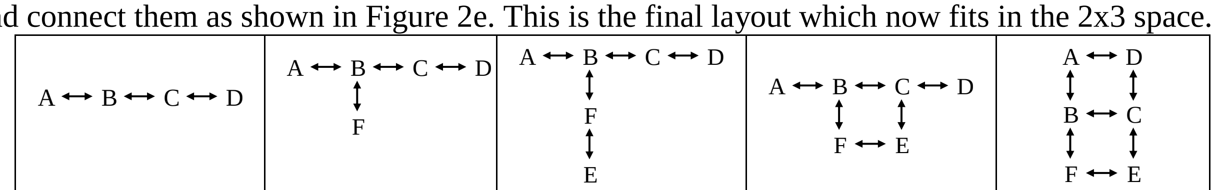


Figure 2a: Expanding	**Figure 2b:** Expanding	**Figure 2c:** Expanding	**Figure 2d:** Folding	**Figure 2e:** Folding

Fitting Phase

Rule 6 (Conflict Resolution for Multi-Row): If the generated solution by using the reshaping rules does not fit into the given layout space, cut low priority (ranking) departments and paste them into the closest empty spaces based on closeness preferences.

Suppose that Figure 3a shows the sequence obtained from Phase I where the layout space is 2x3. In Phase II folding results in Figure 3b. The sequence given in Figure 3b cannot be folded to fit the layout's shape of 2x3. Suppose that the departments (A,B) have the lowest flow of all attached departments in the layout; then A is the best choice to be cut. Apply rule 6 by cutting department A and pasting it to the open space. The result is shown in Figure 3c.

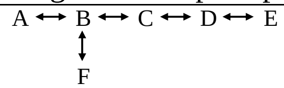
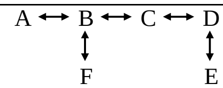
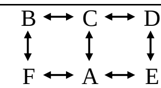
		

Figure 3a: Given Solution	**Figure 3b:**Folding	**Figure 3c:**Cut and Paste A

Example 2: Head-Tail Multi-Row Layout

Consider the information given below. Suppose the layout should be in the form of a 2x3 rectangle.

Locations

1	2	3
4	5	6

Distance D

	1	2	3	4	5	6
1	-	1	2	1	2	3
2		-	1	2	1	2
3			-	3	2	1
4				-	1	2
5					-	1
6						-

Flows V

	A	B	C	D	E	F
A	-	10	20	30	10	100
B		-	50	60	10	20
C			-	10	20	80
D				-	50	60
E					-	10
F						-

Ranking of Flows V

	A	B	C	D	E	F
A	-	11	8	7	11	1
B		-	5	3	11	8
C			-	11	8	2
D				-	5	3
E					-	11
F						-

Figure 4: Initial Data for Example 2

Expanding Steps

1. $s_p = (A \leftrightarrow F)$. $s_q = (A \leftrightarrow F)$ 2. $s_p = (C \leftrightarrow F)$ or $s_p = (F \leftrightarrow C)$. $s_q = (A \leftrightarrow F \leftrightarrow C)$
3. $s_p = (B \leftrightarrow D)$. Ignore s_p temporarily. 4. $s_p = (D \leftrightarrow F)$. s_p can be attached to s_q, see Figure 5a.
5. $s_p = (B \leftrightarrow D)$. B can be attached to D, see Figure 5b.
6. $s_p = (D \leftrightarrow E)$. E can be attached to D. Since (E,C) has a higher flow than (E,A), E is attached to the right side of D. See Figure 5c

All departments are assigned; reshape to fit the layout to the 2x3 block requirement.

Reshaping Steps

7. Consider Figure 5c. (E,C) has a flow of 20 compared to a flow of 10 for (E,A). (B,C) has a flow of 50 compared to a flow of 10 for (B,A). Therefore, the advantage of having B next to C (an improvement of 40) outweighs the advantage from having E next to C (an improvement of 10). Therefore, in Figure 5c, E can be rotated from left to the right side. This rotation will allow B to fold towards (closer) to C instead of A. The result of this rotation rule is shown in Figure 5d.
8. In Figure 5e, B is folded down toward C.

The final solution layout is shown in Figure 5f. The total flow for this solution is Total Flow =

700.

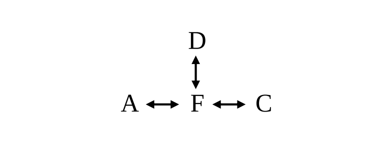


Figure 5a: Expanding

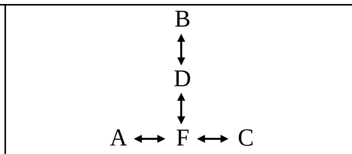


Figure 5b: Expanding

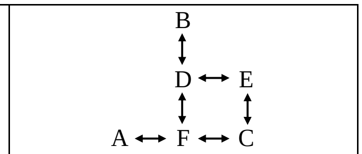


Figure 5c: Expanding

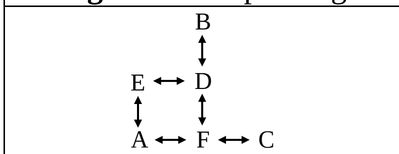


Figure 5d: Rotation of E

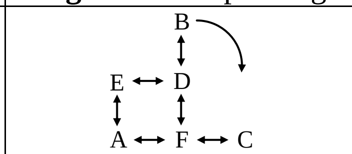


Figure 5e: Folding

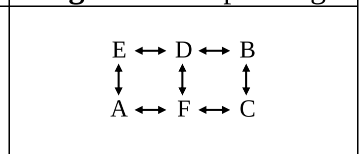


Figure 5f: Solution

Example 3: Comparing Head-Tail Layout and SLP

Consider a facility consisting of seven departments, each with individual space requirements as shown in the table below. Each department can be broken up into units or blocks of 100 square feet. The total allotted area for the new facility is 2400 square feet and must be laid within a rectangular space with a dimension of 4x6 blocks. Each department can fit into a square or a rectangle (e.g. a department can be presented by a 1x6 or 2x3 block shape).

Department	1	2	3	4	5	6	7	Sum
Square feet	600	200	400	200	500	100	400	2400
No. of Blocks	6	2	4	2	5	1	4	24

The matrix below shows the given desired proximity or closeness for each pair of departments. We first use SLP to find a layout solution to this problem for a. 400x600 facility, and b. 200x1200 facility.

	1	2	3	4	5	6	7
1		A	O	A	I	A	A
2			U	E	O	E	X
3				X	U	U	E
4					U	O	O
5						I	I
6							X
7							

Figure 6 show the initial solution where each department is presented by a node, and the desired closeness for each pair of departments is presented by lines that connect the two departments. Figure 7 shows the first iteration of the rearrangement of the relationship diagram to minimize the strain on these bands.

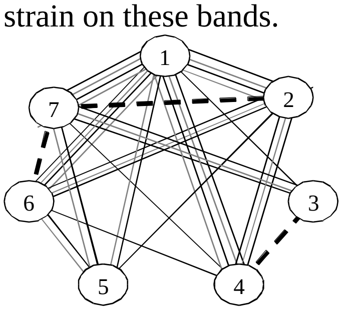


Figure 6: Initial graph to present the problem

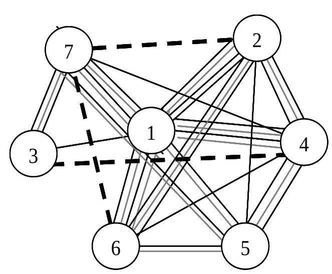


Figure 7: First iteration to improve the total closeness

Figure 8 shows the next iteration. Re-arrangement of the network should be repeated until no more improvement can be achieved. Once the final graphical relationship is achieved, the space requirements for each department are considered. Figure 9 shows the area of each department.

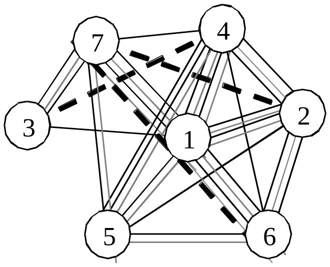


Figure 8: Further rearrangement to improve the total closeness

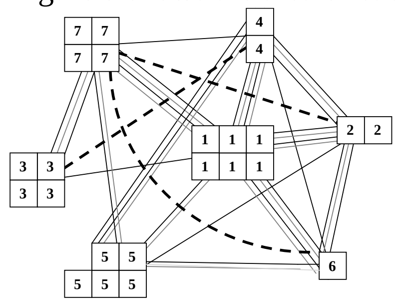


Figure 9: Relationship diagram showing departmental space requirements

Now we can collapse and rotate the spatial relationship diagram shown in Figure 9 to fit into the 4x6 space. This is shown in Figure 10. For the 2x12 block facility, the solution is shown in Figure 11.

2	2	1	1	4	4
5	6	1	1	7	7
5	5	1	1	7	7
5	5	3	3	3	3

Figure 10: Relationship diagram for a 4x6 area

5	5	5	4	2	1	1	1	7	7	3	3
5	5	6	4	2	1	1	1	7	7	3	3

Figure 11: Layout solution for a 2x12 area

The above solution generated by SLP is based on space requirements, availability, closeness preferences, and other relative factors. The A, E, I, O, U, and X ratings can be converted to the numerical values, 5, 4, 3, 2, 1, and 0, respectively. Then the ranking of the pairs of departments will be performed. Note that pairs that tie have same ranking.

Rankings of Pairs (i,j)

	1	2	3	4	5	6	7
1	-	1	11	1	8	1	1
2		-	15	5	11	5	X
3			-	X	15	15	5
4				-	15	11	11
5					-	8	8
6						-	X
7							-

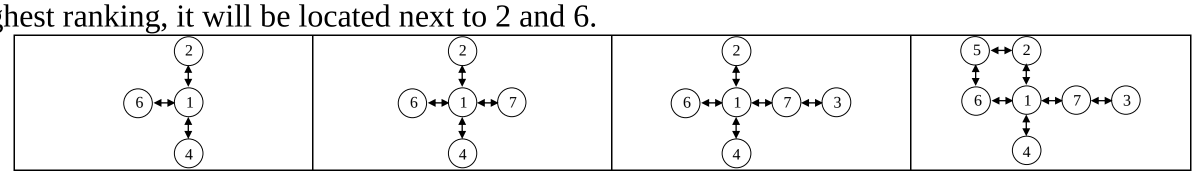
The construction of the layout by the Head-Tail method is as follows.

Expanding Steps

Step 1: $s_p = (1\leftrightarrow2)$. $s_q = (1\leftrightarrow2)$ Step 2: $s_p = (1\leftrightarrow4)$. $s_q = (4\leftrightarrow1\leftrightarrow2)$.

Steps 3, 4, and 5: Departments 6, 7, and 3 are added; see Figure 12.

Step 6: Next consider $s_p = (1\leftrightarrow5)$. Department 5 can be located in four possible locations diagonally from department 1 as shown in Figure 13, Step 6. Because pairs (5,2) and (5,6) have the highest ranking, it will be located next to 2 and 6.



Step 3: Connect 6 to 1	Step 4: Connect 7 to 1	Step 5: Connect 3 to 7	Step 6: Connect 5 to 6 and 2

Figure 12: Graphical Representation of layout tree Steps 3, 4, 5, and 6.

Since all departments are assigned, the expansion phase is complete.

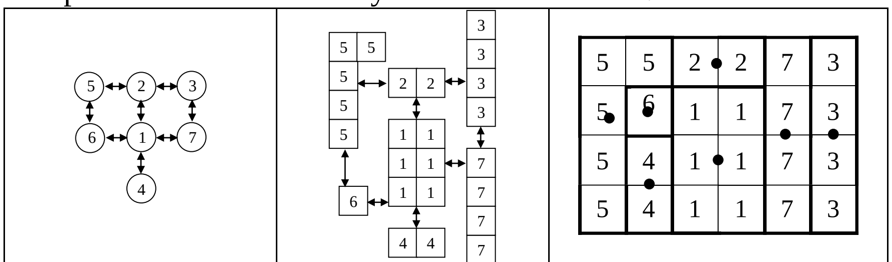
Reshaping Steps

Step 7: Consider the next pair, (4↔6). Departments 4 and 6 cannot be folded. Consider the next pair, (4↔7). Departments 4 and 7 cannot be folded. Consider the next pair, (2↔3). Fold 3 closer to 2 (see Step 7 in Figure 13).

Fitting Steps

Step 8: Present the actual size and shape of the departments.

Step 9: Move the departments to fit in the layout size of 4x6 units.



Step 7: Connect 3 to 2	Step 8: Insert Spaces	Step 9: Final Layout

Figure 13: Graphical Representation of Steps 7, 8, and 9

The final layout is shown in Step 9 of Figure 13. The center of each department is presented by a "●". The distances between each pair of departments are shown in the following table.

	1	2	3	4	5	6	7
1		200	300	200	310	250	200
2			400	400	350	250	300
3				500	510	450	100
4					210	150	400
5						80	410
6							350
7							

The total preferential closeness is: Total Closeness = $\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\text{Closeness}(i,j)*\text{Distance}(i,j)\right)$ = 13,620

Comparison of Head-Tail and SLP Solutions: The above example was solved by SLP method. The layout and the distances for the SLP solution are shown in Figure 14. The total closeness for the SLP solution is 15,500. The Rule-based Planning solution (Example 3) was 13,620 which is considerably better than the solution found by the SLP method.

	1	2	3	4	5	6	7
1		300	300	300	330	150	250
2			600	400	230	150	550
3				400	390	450	250
4					630	450	150
5						180	480
6							400
7							

Distances for the SLP Solution

2	2	1	1	4	4
5	6	1	1	7	7
5	5	1	1	7	7
5	5	3	3	3	3

Figure 14: SLP Layout Solution

3. Bi-Criteria Layout Planning

In bi-criteria facility layout problem, the two objectives are:

Minimize Total Flow $$f_1 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} d_{ij}$$

Minimize Total Closeness $$f_2 = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} p_{ij}$$

For example, suppose that the flow between departments A and B is very high and from the total flow point of view should be located very close to each other. However, from a safety point of view, they should be very far from each other. The above bi-criteria problem can be used to handle such conflicting objectives.

Suppose that an additive utility function can be used to rank multi-criteria layout alternatives. Let w_1 and w_2 be the weights of importance of the two objective functions, respectively. First find the normalized values of the two objective functions. To normalize the flow information, find the largest flow ($d_{ij,max}$) and smallest flow ($d_{ij,min}$), and then use the formula shown below to normalize all flow values presented by d_{ij}: they will be between 0 and 1. Similarly, closeness information can be normalized between 0 and 1 using the following formula where $p_{ij,max}$ is the largest and $p_{ij,min}$ is the smallest p_{ij} values.

$$v_{ij}' = \frac{(w_{ij} - w_{ij,min})}{(w_{ij,max} - w_{ij,min})} \qquad p_{ij}' = \frac{(p_{ij} - p_{ij,min})}{(p_{ij,max} - p_{ij,min})}$$

Then for each pair of departments (i,j), calculate

$$z_{ij} = w_1 d_{ij}' + w_2 p_{ij}'$$

The matrix of z_{ij} values is called the composite matrix, Z. The additive utility function (U) is defined as the weighted sum of the flow and the closeness values using weights of importance of w_1 and w_2 for the flow and closeness, respectively. In order to minimize the additive utility function $U = w_1 f_1' + w_2 f_2'$,

$$\text{Minimize } U = \sum_{i-1}^{n}\sum_{j=1}^{n} d_{ij}(w_1 d_{ij}' + w_2 p_{ij}') = \sum_{i=1}^{n}\sum_{j=1}^{n} d_{ij} z_{ij} \qquad (2)$$

The value of $z_{ij} = w_1 d_{ij}' + w_2 p_{ij}'$ is constant for i, j, and given weights. Therefore, they only need to be calculated once for each pair of departments. Therefore, for the generated Z matrix, the bi-criteria layout problem can be solved by Head-Tail method. The alternative with the smallest U value is the best layout.

Example 4: Bi-Criteria Layout

For a four department facility layout problem presented in Figure 14, find the best bi-criteria layout solution. Suppose that the decision maker's weights of importance for flow and closeness objectives are $w_1 = 0.3$ and $w_2 = 0.7$, respectively. Consider using 0 to 5 numerical ratings for X to A qualitative ratings.

Location

1	2	3	4

Assignment

A	B	C	D

Flow Matrix

	A	B	C	D
A	-	100	40	10
B		-	60	30
C			-	90
D				-

Qualitative Matrix

	A	B	C	D
A	-	X/0	E/4	A/5
B		-	U/1	E/4
C			-	X/0
D				-

Converted Qualitative Matrix

	A	B	C	D
A	-	0	4	5
B		-	1	4
C			-	0
D				-

Figure 15: Initial data for the bi-criteria layout problem

First convert qualitative closeness data to quantitative data as presented in Figure 15. Then normalize flow matrix and closeness matrix values as shown in Figure 16.

Normalized Flow Matrix

$(v_{ij}') = (v_{ij}-10)/(100-10)$

	A	B	C	D
A	-	1	0.333	0
B		-	0.556	0.222
C			-	0.889
D				-

Normalized Qualitative Matrix

$(p_{ij}') = (p_{ij}-0)/(5-0)$

	A	B	C	D
A	-	0	0.8	1
B		-	0.2	0.8
C			-	0
D				-

	A	B	C	D
A	-	0.3	0.66	0.7
B		-	0.307	0.627
C			-	0.267
D				-

Figure 16: Normalized flow and qualitative closeness values

A composite matrix, Z, can be generated by using $z_{ij} = w_1 v_{ij}' + w_2 p_{ij}' = 0.3 v_{ij}' + 0.7 p_{ij}'$.

Solving the Bi-Criteria Problem by the HEAD-TAIL Method: First rank all pairs in descending order of z_{ij} values: they are (A,D), (A,C), (B,D), (B,C), (A,B), and (C,D).

The steps of HEAD-TAIL for assigning departments are:

1) (A↔D); 2) (C↔A↔D); 3) (C↔A↔D↔B)

This is the final solution, and its mirror is (B↔D↔A↔C). The objective values for this layout are: $f_1 = 2*100 + 1*40 + 1*10 + 3*60 + 1*30 + 2*90 = 640$ and $f_2 = 2*0+1*4+1*5+3*1+1*4+2*0 = 16$

No improved solution can be found by applying the Pairwise Exchange Method to the final solution of this example.

Using the additive utility as the performance measure for each facility, the first iteration of the Pair-Wise Exchange method produces the results shown below.

Coefficient	Facility Pairs	Initial distance	Pair-Wise Exchange new distance					
			AB	AC	AD	BC	BD	CD
0.3	AB	1	1	1	2	2	3	1
0.66	AC	2	1	2	1	1	2	3
0.7	AD	3	2	1	3	3	1	2

0.307	BC	1	2	1	1	1	1	2
0.627	BD	2	3	2	1	1	2	1
0.267	CD	1	1	3	2	2	1	1
U = $\sum_i \sum_j d_{ij} z_{ij}$		5.548	1722	**4.682**	4.828	4.828	4.748	1788

The lowest score is 4.682 for swapping A and C.

Improved Layout:

C	B	A	D

The swap is made and the initial layout for the second iteration is as follows:

Coefficient	Facility pairs	Initial distance	Pair-Wise Exchange new distance					
			AB	AC	AD	BC	BD	CD
0.3	AB	1	1	1	2	2	1	1
0.66	AC	2	1	2	3	1	2	1
0.7	AD	1	2	3	1	1	1	2
0.307	BC	1	2	1	1	1	3	2
0.627	BD	2	1	2	1	3	2	1
0.267	CD	3	3	1	2	2	1	3
U= $\sum_i \sum_j d_{ij} z_{ij}$		4.682	**4.402**	5.548	4.748	4.682	4.762	**4.402**

The lowest score is 4.402 for swapping A and B. The new layout is.

Improved Layout:

C	A	B	D

Next iteration is:

Coefficient	Facility pairs	Initial distance	Pair-Wise Exchange new distance					
			AB	AC	AD	BC	BD	CD
0.3	AB	1	1	2	1	1	2	1
0.66	AC	1	2	1	3	1	1	2
0.7	AD	2	1	3	2	2	1	1
0.307	BC	2	1	1	2	2	3	1
0.627	BD	1	2	1	1	3	1	2
0.267	CD	3	3	2	1	1	2	3
U= $\sum_i \sum_j d_{ij} z_{ij}$		4.402	4.682	4.828	1788	1722	**4.042**	4.682

The lowest score is 4.042 for swapping B and D. The new layout is shown below.

Improved Layout:

C	A	D	B

Continuing with the next iteration:

Coefficient	Facility pairs	Initial distance	Pair-Wise Exchange new distance					
			AB	AC	AD	BC	BD	CD
0.3	AB	2	2	3	1	1	1	2
0.66	AC	1	3	1	2	2	1	1
0.7	AD	1	1	2	1	1	2	1
0.307	BC	3	1	2	3	3	2	1
0.627	BD	1	1	1	2	2	1	3
0.267	CD	2	2	1	1	1	3	2
U= $\sum_i \sum_j d_{ij} z_{ij}$		**4.042**	4.748	4.468	4.762	4.762	4.402	4.682

No more improvement is possible, stop. The layout remains with f_1 (total flow) = 640 and f_2 (total closeness) = 32.

Final Layout:	C	A	D	B

Efficient Frontier

By systematically varying weights (w_1, w_2) in Equation (2), a set of efficient bi-criteria alternatives can be generated. To do this, first generate the Z matrix for each given set of weights and then use head-Tail method to solve the problem. Note that for a given set of weights, when PWE method is used, alternatives generated in different iterations of PWE method may also be efficient and should be recorded. That is, each Pairwise Exchange solution in each iteration of PWE method can be efficient. Consider the set of four weights given in the first row of the following table. For each given set of weights, solve the problem by Head-Tail method. The generated solution is recorded under each set of weights. The set of generated four alternatives are all efficient.

(w_1, w_2)	(0.99, 0.01)	(0.6, 0.4)	(0.5, 0.5)	(0.01, 0.99)
Layout	ABCD	BACD	DBAC	CADB
Min. f_1	420	460	580	640
Min. f_2	32	28	20	16
Efficient?	Yes	Yes	Yes	Yes

Conclusion

In SLP, the process of improving the initial graph of the layout becomes very difficult as the number of departments increases. Also the evaluation process in SLP is subjective and depends on the layout designer. A different approach for solving layout problem is by use of Head-Tail Layout method which is a powerful tool for solving both structured and ill-structured layout problems. Head-Tail can solve very large layout problems with minimal computational efforts. Because this method is Head-Tail, it can incorporate complicated constraints while developing the layout solution. Furthermore, it can be used in conjunction with Expert systems approaches to solve layout problems.

References:

[1] Raman D., S. V. Nagalingam, G. C.I. Lin, "Towards measuring the effectiveness of a facilities layout", *Robotics and Computer-Integrated Manufacturing*, doi:10.1016/j.rcim.2007.06.003, 2008.

[2] Rosenblatt, M. J., "The facility layout problem: a multi-goal approach." *International Journal of Production Research*, 17, 1979, pp. 323-332.

[3] Fortenberry, J. C., J. F. Cox, "Multiple criteria approach to the facilities layout problem", *International Journal of Production Research*, Vol. 23, Issue 4, 1985, pp.773-782.

[4] Malakooti, B., "Multi-Objective Facility Layout: A Heuristic Method to Generate All Efficient Alternatives", *International Journal of Production Research*, Vol. 27, No. 7, 1989, pp. 1225-1238.

[5] Malakooti, B., "Computer Aided Facility Layout Selection with Applications to Multiple Criteria Manufacturing Planning Problems." Large Scale Systems: Theory and Applications, Special Issue on Complex Systems Issues in Manufacturing, Vol. 12, 1987, pp. 109-123.

[6] Malakooti, B., A. Tsurushima, "An Expert System Using Priorities for Solving Multiple Criteria Facility Layout Problems." *Inter. Journal of Production Research*, Vol. 27, No. 5, 1989, pp. 793-808.

[7] Yang, Z., B. Malakooti, "A Multiple Criteria Neural Network Approach for Layout of Machine -Part Group Formation in Cellular Manufacturing." *Third IE Research Conference*, 1994, pp. 249-254.

[8] Islier A. A., "A genetic algorithm approach for multiple criteria facility layout design", *International Journal of Production Research*, 1998, Vol. 36, No. 6, pp. 1549-1569.

[9] Pelinescu D. M., M. Y. Wang, "Multi-objective optimal fixture layout design", *Robotics and Computer-Integrated Manufacturing*, Vol. 18, 2002, pp. 365-372.

[10] Yang T., Ch. Kuo, “A hierarchical AHP/DEA methodology for the facilities layout design problem”, *European Journal of Operational Research*, 147, 2003, pp. 128–136.
[11] Chen C.-W., D. Y. Sha, “Heuristic approach for solving the multi-objective facility layout problem”, *International Journal of Production Research*, Vol. 43, No. 21, 1, November 2005, pp. 4493–4507.
[12] Sha, D., C. Chen, "A New Approach to the Multiple Objective Facility Layout Problem", *Integrated Manufacturing Systems*, Vol. 12, No.1, 2001, pp. 59-66.
[13] Şahin, Ramazan. “A Simulated Annealing Algorithm for Solving the Bi-objective Facility Layout Problem.” *Expert Systems with Applications*, Vol. 38, Issue 4, 2011, pp. 4460-4465.
[14] Singh S. P., V. K. Singh, “Three-Level AHP-Based Heuristic Approach for a Multi-Objective Facility Layout Problem.” *International Journal of Production Research*, Vol. 49, Issue 4, 2011, pp. 1105-1125.

A Synopsis of Multiplicative Z Utility Theory for Solving Risk Problems[1]

Behnam Malakooti

Case Western Reserve University

Systems Engineering

Email: bxm4@case.edu

Abstract

This paper provides a synopsis of newly developed approach, Z Utility Theory, for solving risk problems. Descriptive experimental evidence show that decision behavior existing utility theory models such as Expected Utility Theory, Cumulative Prospect Theory, and Rank Dependent Expected Utility, cannot solve risk problems. Z-Utility-Theory normative model is developed in response to above challenge. Z-Utility-Theory generalizes Expected Utility Theory, Prospect Theory, Cumulative Prospect Theory, Mean-Variance approach, Cobb-Douglas production function, and arithmetic and geometric mean functions. Z-Utility-Theory functions have only four parameters and can solve all different decision problems. Malakooti [11] shows that ZUT substantially outperforms all existing risk methods including EUT, CPT, and Rank Dependent Expected Utility. He also shows that ZUT solves all existing decision paradoxes.

Key words: Risk, Decision Making, Z Utility Theory, Stochastic Dominance, Editing rule

1. INTRODUCTION

In this paper, we show how risk can be accurately quantified by utility functions for both normative (rational and prescriptive) and descriptive decision behaviors. The developed approach is called Z Utility Theory. We show that Z Utility Theory (a normative model) overcomes all the obstacles that were faced by Expected Utility Theory (von-Newman-Morgenstern [1]), Prospect-Theory (Kahneman and Tversky [2], Cumulative-Prospect-Theory (Tversky-Kahneman,[3]), and Mean-Variance approach of Markowitz [4].

Historical Perspective: We can divide the history of solving risky decisions into four eras.

I. Prior to St. Petersburg paradox (Expected Value Based Rationality) where Expected Values (EVs) and rules of thumb were used to make decisions. Expected Value (EV) is linear in both probabilities and payoffs.

---

[1] The information contained within is the sole intellectual property of Professor Behnam Malakooti of Case Western Reserve University, Cleveland, Ohio.
Proceedings of the 2014 ICAM, International Conference on Advanced and Agile Manufacturing, Held at Oakland University, Rochester, MI 48309, USA

II. From Bernoulli [5] to von-Newman-Morgenstern [1] (Value-based Rationality) where value functions were used. Value functions assumed to be concave for Risk Averse (and convex for Risk Prone). The culmination of this period is Expected Utility Theory, EUT, which is linear in probabilities.

III. From Allais [6] to Tversky-Kahneman [3] (Subjective Probability Based Rationality) where in addition to value functions, subjective probabilities in place of probabilities were used. The subjective probabilities, Savage [8], were evolved into rank-dependent expected utility methods (e.g. Quiggin [7]). The culmination of this period is Cumulative-Prospect-Theory, CPT, (Tversky-Kahneman [3]) which is linear in subjective probabilities. CPT assumes value function is concave for gains (and convex for losses) and weight function has a reverse S shape. In this period, researchers focused on weakening or removing some of EUT Axioms, e.g. see Machina [9] among many others.

IV. From Markowitz [4] to Present (2013) (Dual/Nonlinear Rationality) where dual functions of risk are developed. Markowitz [4] developed mean-variance analysis by using Expected Value (EV) (presenting mean) and Standard Deviation (presenting risk). However, Mean-variance analysis results in utility functions that violate stochastic dominance. Instead of Standard Deviation, Z Utility Theory uses a new multiplicative function (called Risk Value) to measure risk whose utility function is compliant with stochastic dominance. Z Utility Theory simultaneously uses three different functions to measure risk: value function of payoffs, weight function of probabilities, and multiplicative Risk Value function of both payoffs and probabilities. We will illustrate that each of these three functions can be used independently to explain risk behavior. Based on experimental evidence of this paper, we show that value and weight functions have similar shapes, they are double helix which is a generalization of shapes suggested by Prospect Theory and Cumulative Prospect Theory. ZUT also allows for reference points (e.g. reflecting initial wealth) to be determined independent from gain and loss decomposition which is the main characteristics of PT and CPT. ZUT generalizes all past generations of risk models including EUT, Mean Variance, PT, and CPT. Malakooti [10] shows several applications of Z Theory for solving Multiple Criteria Decision Making (MCDM) and applications to operations and production systems.

2. MULTIPLICATIVE Z UTILITY FUNCTION

2.1. Risk Problem: Consider a single attribute (criterion) decision problem under risk where each alternative is presented by $Y(y,p) = (y_1,p_1;\ y_2,p_2;\ldots;\ y_k,p_k)$, where y_i, is the payoff of state i and p_i is the probability of state i, and the summation of probabilities is one. Therefore each alternatives has a k-tuple vector of payoffs $Y= (y_1,\ y_2,\ldots,y_k)$ and probabilities $P=(p_1,p_2,\ldots,p_k)$. Note that different alternatives may have different number of states, k, and therefore different probabilities. For example, Y_1=(-\$12,0.7;\$20,0.3) has k=2 states and Y_2=(-\$5,0.4;\$15,0.59;-\$21,0.01) has k=3 states. Suppose $y_{min}=\min\ \{y_{i,j}\}$ for all alternatives j=1,2,..,n and all states i=1,2,…,k. Therefore, y_{min} is the minimum payoff considering all possible alternatives. For risk problems, M-ZUT can be written as (2.2) where each alternative may have a different k. So for risk problems we use Expected Value (EV) instead of Linear Value (LV), and Risk Value (RV_M) instead of Deviational Value (DV_M).

$$EV = \sum_{i=1}^{k} p_i y_i , \qquad MV_M = \prod_{i=1}^{k} (y_i - y_{min})^{p_i} + y_{min} , \quad \text{and} \quad RV_M = |EV\text{-}MV_M| \tag{2.1}$$

$$U = EV + z_M * RV_M \qquad -1 < z_M \leq 0 \tag{2.2}$$

Malakooti [11] proves that M-ZUT (2.2) satisfies stochastic dominance. Therefore, Risk problems can be solved by M-ZUT which is nonlinear in probabilities and payoffs. In the following the contour and graph of EV and MV_M is presented.

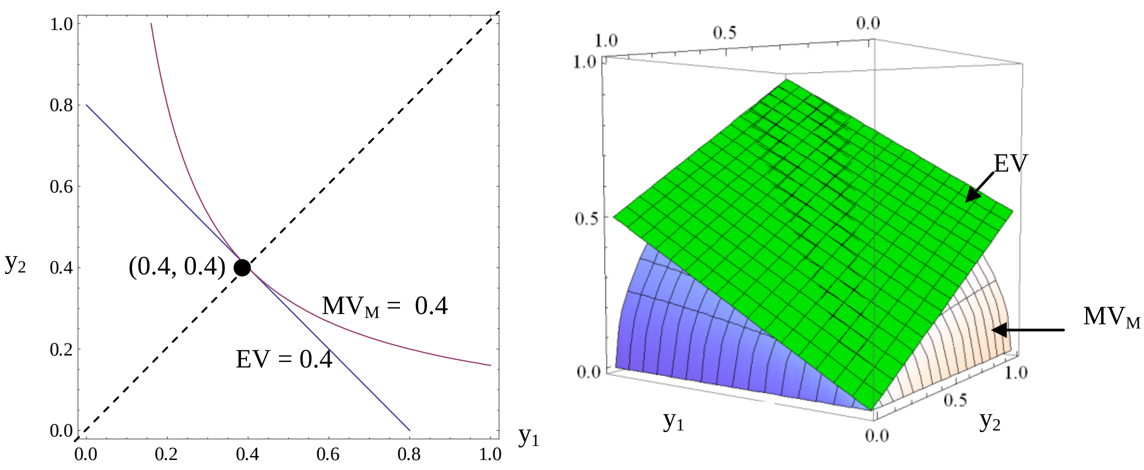


Figure 1: A symmetric concave Multiplicative ZUF for z_M = -1, U=MV_M**:** **a**. Contours of EV and MV_M, **b.** Functions of EV and MV_M

Risky Alternatives: An alternative is Risky if its RV_M>0, otherwise it is Certain when RV_M=0.

Certainty Equivalence is Equal to Utility Value in M-ZUT (CE=U): Certainty Equivalence (CE) of a certain alternative is equal to its Expected Value (EV). But CE of a risky alternative is U. The type of risk attitude is determined by comparing CE and EV. At a given alternative the DM is: Risk Averse (RA) if CE <EV; Risk Prone (RP) if CE>EV, and risk Neutral (RN) if CE=EV.

2.2. Axioms and Theorem of M-ZUT for Risk

Malakooti [11] provides the set of Axioms of Z Theory which include Completeness, Transitivity, Continuity, Monotonicity, and ZUT independence. For more detail see Malakooti [11].

M-ZUT Risk Classifications: For all alternatives:

- z_M < 0 (the DM is Risk Averse) if and only if M-ZUT is concave in payoffs y.
- z_M > 0 (the DM is Risk Prone) if and only if M-ZUT is convex in payoffs y.

- z_M = 0 (the DM is Risk Neutral) if and only if M-ZUT is additive in payoffs y.

Note that when z_M >0, M-ZUT is convex and risk prone, however, M-ZUT when using z_M >0 may not satisfy stochastic dominance. See Supplement for an example of Solving Allais Paradox by M-ZUT.

Expected Utility Theory as a Special Case of M-ZUT: Suppose that v(y) is used instead of y in M-ZUT, Equations (2.1 and 2.2), then M-ZUT uses value functions and it becomes V-M-ZUT.

When z_M=0 in V-M-ZUT then V-M-ZUT=p_1*v(y_1)+p_2*v(y_2)+…+p_k*v(y_k)=Expected Utility Value of EUT of von-Newman and Morgenstern [1]. That is, ZUT generalizes EUT which is linear in probabilities, p. When z_M = 0, the DM is: Risk Averse if and only if v(y) is concave; Risk Prone if and only if v(y) is convex; and Risk Neutral if and only if v(y) is linear.

Cumulative Prospect Theory (and Rank Dependent Expected Utility) as a Special Case of M-ZUT: Because Prospect Theory, Kahneman and Tversky [2], violates stochastic dominance, Tversky and Kahneman [3] developed Cumulative Prospect Theory to remedy this problem. In M-ZUT, Equations (2.1 and 2.2), use v(y) instead of y, and use w(p) instead of p (that is probability weights are used instead of probabilities), then M-ZUT (Equations (2.1) and (2.2)), uses both value and weight functions, when z_M=0 then, (2.2) becomes U=w(p_1)*v(y_1)+ w(p_2)*v(y_2)+…+ w(p_k)*v(y_k) which is the same as the Cumulative Prospect Theory of Tversky and Kahneman [3] and RDEU. That is, ZUT generalizes CPT and RDEU which is linear in weights, w(p).

2.3 Stochastic Dominance Standard Form (SDSF)

Stochastic Dominance Standard Form (SDSF): Consider Y_A and Y_B. Present Y_A as its SD Equivalent Y_A^S; and Y_B as its SD Equivalent Y_B^S such that both Y_A^S and Y_B^S have the same probabilities of states, their payoffs are in non-decreasing order, and the number of states is minimized. Similarly, n different risky alternatives can be presented in SDSF having k number of states. These representations of alternatives can be graphically verified by presenting the c.p.d., G, of alternatives. See Allais Paradox (Table 4.2) in SDSF for two pairs of alternatives and also for four alternatives.

Approach for Generating Alternatives in SDSF (Splitting States): First for each alternative, reorder its states such that its payoffs will be in non-decreasing order. Then split the probability of each state of each alternative such that all alternatives have the same probabilities for all states using the minimum number of states of nature. See Allais Paradox below.

Collapsing States (Combining States): When comparing a sub-set of SDSF alternatives to simplify their presentations to the DM, use the SDSF of the sub-set of alternative. Collapsing will reduce the number of states to the minimum because SDSF has the minimum number of states for any set of alternatives.

SDSF of Allais (1953) Paradox: Allais Paradox consists of four alternatives Y_A=\$1 for sure, Y_B=(\$5,0.10;\$1,0.89;\$0,0.01), Y_C=(\$1,0.11;\$0,0.89), and Y_D=(\$5,0.1;\$0,0.9) where payoffs are in \$1million. In most cases DMs prefer Y_A to Y_B; and Y_D to Y_C. The SDSF for comparing Y_A to Y_B (and also Y_C to Y_D) requires only three states. The SDSF for four alternatives has four states. Note that EV, RV_M, and $RV_{M,C}$ (where C=0.3) of alternatives are independent of SDSF presentations. The CE=U of alternatives for z_M=-0.35, i.e. U= EV-0.35RV_M, is shown.

	Y_A vs. Y_B in SDSF			Y_C vs. Y_D in SDSF		
	P_{AB}^S	Y_A^S	Y_B^S	P_{CD}^S	Y_C^S	Y_D^S
y_1	0.01	1	0	0.89	0	0
y_2	0.89	1	1	0.01	1	0
y_3	0.1	1	5	0.1	1	5
EV	-	1	1.39	-	0.11	0.5
RV_M	-	0	1.39	-	0.11	0.5
CE=U	-	1	0.904	-	0.072	0.325

a. Two pairs in SDSF

	Y_A,Y_B, Y_C, Y_D in SDSF			
P^S	Y_A^S	Y_B^S	Y_C^S	Y_D^S
0.01	1	0	0	0
0.88	1	1	0	0
0.01	1	1	1	0
0.1	1	5	1	5
EV	1	1.39	0.11	0.5
RV_M	0	1.39	0.11	0.5
$RV_{M,C}$	0	0.22	0.06	0.4

b. All alternatives in SDSF

Table 1: SDSF presentation of Allais Paradox in $

CONCLUSION

In this paper we summarized the key points of a new theory call Z Utility Theory which is being developed by Malakooti [11] for decision making under risk. This theory generalizes Expected Utility Theory, Prospect Theory, Cumulative Prospect Theory, Mean-Variance approach, Cobb-Douglas production function, and arithmetic and geometric mean functions. Also, ZUT satisfies the first and second stochastic dominance. Moreover, we introduced a new editing rule which is called Stochastic Dominance Standard Form (SDSF) which can be shown that increases the accuracy of decision making. According to the benchmarking on the well-established dataset, Malakooti [11] shows that ZUT substantially outperforms all existing risk methods including EUT, CPT, and Rank Dependent Expected Utility. He also shows that ZUT solves all existing decision paradoxes.

REFERENCES:

[1] von Neumann J.,O. Morgenstern, Theory of Gaines and Economic Behavior. Princeton: Princeton University Press, 1st edition 1944, 2nd edition 1947.

[2] Kahneman, D., A. Tversky, (K-T) (1979): "Prospect Theory: An Analysis of Decision Making Under Risk." *Econometrica*, 47(2), 263 – 292.

[3] Tversky, A., and Kahneman, D. (T-K) (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and uncertainty*, 5(4), 297-323.

[4]Markowitz, H.M. (1959): "*Portfolio Selection: Efficient Diversification of Investments*," New York: JohnWiley & Sons (Reprinted by London: Yale University Press, 1970; 2nd ed. Oxford, UK: Basil Blackwell, 1991.)

[5] Bernoulli, D. (1738): "Specimen Theoriae Novae de Mensura Sortis," in Commentarii Academiae Scien-tiarum Imperialis Petropolitanae. V. Translated by L. SOMMER (1954) as "Exposition of a New Theory on the Measurement of Risk," Econometrica, 22, 23-36.

[6] Allais, M. (1953): "Le comportement de l'hommerationneldevant le risque: critique des postulatsetaxiomes de l'écoleAméricaine," *Econometrica*, 21(4),503–546

[7] Quiggin, J. (1982): “A theory of anticipated utility,” *Journal of Economic Behavior and Organization*, 3(4), 323–43.
[8] Savage, L. J. (1954): “*The Foundations of Statistics*.” Wiley, New York.
[9]Machina, Mark J., (1982): “Expected utility’ analysis without the independence axiom, *Econometrica*, 50, 277-323.
[10] Malakooti, Behnam. *Operations and Production Systems with Multiple Objectives*. John Wiley & Sons, 2014a.
[11] Malakooti, Behnam. Risk, Ambiguity, Uncertainty, Catastrophic, and Multi-attribute Decision Patterns and Quantifications: Z Theory, Case Western Reserve University, 2014b.
[12] Cobb, C. W., P. H. Douglas, (1928): “A Theory of Production”. *American Economic Review*,18,139–165.
[13] Keeney, R. L., (1974): “Multiplicative Utility Functions”, *Operations Research*, 22(1), 22-34.

Virtual Reality 3D Simulations of Concrete and Asphalt Laboratories

Tyler Bayne[1], Spencer Hoin[1], Dr. Pnina Ari-Gur[1], Dr. Marwa Hassan[2] , Mr. Peter Thannhauser[1], Dr. Roman Rabiej[1], Dr. Pavel Ikonomov[1], Mr. Jeff Johnston[3], and Dr. Dan Litynski[1]

[1]College of Engineering and Applied Sciences, Western Michigan University, Kalamazoo, MI, 49008
[2]Louisiana State University, Baton Rouge, Louisiana
[3]Muskegon Community College, Muskegon, MI
Email: tyler.w.bayne@wmich.edu; pnina.ari-gur@wmich.edu;

Abstract

Virtual reality 3D laboratory experiments are being developed to provide students with a virtual hands-on experience. Students perform these virtual experiments, which simulate the physical ones, prior to doing the physical experiment. In some cases, the physical experiment is unavailable and the virtual one is the only hands-on experience for the students. The concrete and asphalt virtual experiments were developed using the game development engine Unity 3D. They contain realistic 3D models of the physical items which are manipulated with scripts programmed using C# and/or JavaScript. The students view the virtual experiments from a *first person* view. In the concrete compression test the student mixes different aggregates and creates their own concrete cylinder. The strength is then measured using a compression tester. In the asphalt experiment, the student mixes and tests the viscosity of an asphalt mixture. The virtual experiments have been introduced into numerous universities and colleges around the globe. Feedback is being collected to provide specific updates and improve the overall quality of the virtual modules.

Research Objective

The virtual lab project is being developed in order to provide students with a virtual hands-on experience using virtual reality 3D experiments. These virtual experiments simulate the physical ones and allow the students to perform step-by-step tasks enhancing their educational experience.

Figure 1: The safety glasses and mixing table in the concrete experiment

Methods Used

The base for the virtual experiments is the physical lab, performed prior to starting the simulation development. Some of the experiments have been developed using Unity 3D, a game development engine. Unity 3D is used by professional development studios to create games for PC, Xbox, iOS, and other platforms. The Unity 3D scripts are programmed using C# and/or JavaScript. The virtual experiments contain realistic 3D models of the different items around the laboratory room. The 3D models were created using Adobe Photoshop and 3ds Max before being imported into Unity 3D and paired with the scripts.

The perspective of the students performing the laboratory is a *first person* view. This allows true simulation of the physical, hands-on experiment. During the experiment, the mouse is used to control the "head" of the virtual student while the arrow keys are used for moving around the virtual laboratory room. A task is completed by pointing to the object and clicking the mouse.

Concrete and Asphalt Modules

In the concrete compression test, the students mix different aggregates and cement to create their own concrete cylinder. The team developed this experiment first. In the asphalt laboratory, the viscosity is measured after the asphalt is mixed.

Figure 2: The mixing barrel in the concrete experiment

The concrete experiment begins with the students putting on the safety glasses (safety first!), as can be seen in Figure 1. The simulation does not let the students continue until this step is completed. In the next step, the students pour a bucket of water into the mixing barrel to prepare for the mixing of the concrete. The virtual experiment then guides the students through the process of mixing the different aggregates together for the concrete mixture.

At this stage, the students are ready to create a concrete cylinder. As seen in Figure 2, the completed mixture is poured into the mixing barrel, the barrel is spun, and the concrete is poured into the mold. The students then carry the cylinder to the curing chamber in the adjoining room.

One of the benefits of the virtual experiment is that the students do not have to wait for the sample to cure. They can immediately take a cured version of the sample (labeled by its "curing time") and start the compression test. The students test the sample that has cured for seven days by carrying it to the compression tester (Figure 3).

Figure 3: The compression tester and computer in the concrete experiment

After the sample is inside the compression chamber, the students press the green button on the control panel to start the machine. It compresses the cylinder until it breaks and the strength is displayed on the computer.

Proceedings of the 2014 ICAM, International Conference on Advanced and Agile Manufacturing, Held at Oakland University, Rochester, MI 48309, USA

The students then go back to the curing room and repeat the testing with a sample that has cured for 28 days.

Asphalt mixing and viscosity testing virtual experiment has also been developed. This experiment takes place on a table top. This table is shown in Figure 4.

Again, the first step is to put on the safety glasses. The students cannot move to the next step until the safety glasses are worn. Next, the students adjust the temperature settings by clicking on the temperature dial. They can choose between 135 and 165 °C.

Figure 4: The asphalt mixing experiment

The students can then pour the asphalt binder into the chamber situated on the scale. The weight is displayed after the asphalt is poured. The next step is where the students place the asphalt chamber into the hot environmental chamber.

After the spindle is lowered, the students set the machine to 20 RPM and turn the motor on. The spindle starts spinning and measures the viscosity of the asphalt. The students press the "Print" button and the result is displayed on the machine's screen (Figure 5). The viscosity measurement shows different values for each student because it is a randomly generated number within a set range, depending on the temperature selected in the beginning.

Results

The initial implementation of the lab in a teaching environment has shown promising results so far. Further assessment data will be collected and analyzed to gauge the success of the virtual laboratory experiments and to provide feedback for improvement. Also, a scanning electron microscope simulation is being developed where the students choose between different samples and manipulate the microscope to capture a clear image of a specimen. This is a more involved instrument and was slated to start last to build on our experience with developing modules.

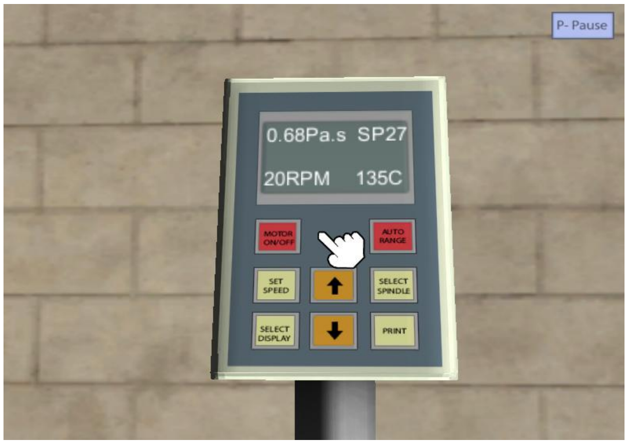


Figure 5: The controls and display in the asphalt experiment

Conclusion

The objective for this research project is to allow students to perform a hands-on simulation of engineering laboratories. The impact on student learning has been positive so far. An electron scanning microscope virtual experiment is currently in development, building on the experience with other modules.

The virtual experiments have been introduced into the curriculum at Western Michigan University, Louisiana State University, Muskegon Community College, as well as numerous other universities and colleges around the globe. Feedback is being collected and used to provide specific updates to the virtual laboratories as well as increasing their overall quality.

Acknowledgment

The generous grants and equipment from Hewlett Packard's Higher Education HP Technology for Teaching Grant Program (2007), Hewlett Packard's 2010 Catalyst Initiative, and the National Science Foundation’s grant number 1140348 made the virtual laboratory project possible.

EDUCATION ENHANCEMENT AND ATTRACTING STUDENTS TO STEM CAREER IN SHIPBUILDING AND MARINE INDUSTRY

Dr. Suren N Dwivedi

Department of Mechanical Engineering and Manufacturing

University of Louisiana at Lafayette

Lafayette, Louisiana

Email id: Suren7483@gmail.com

Varun Kumar Pyata

Department Of Computer Science

University of Louisiana at Lafayette

1.0 Abstract:

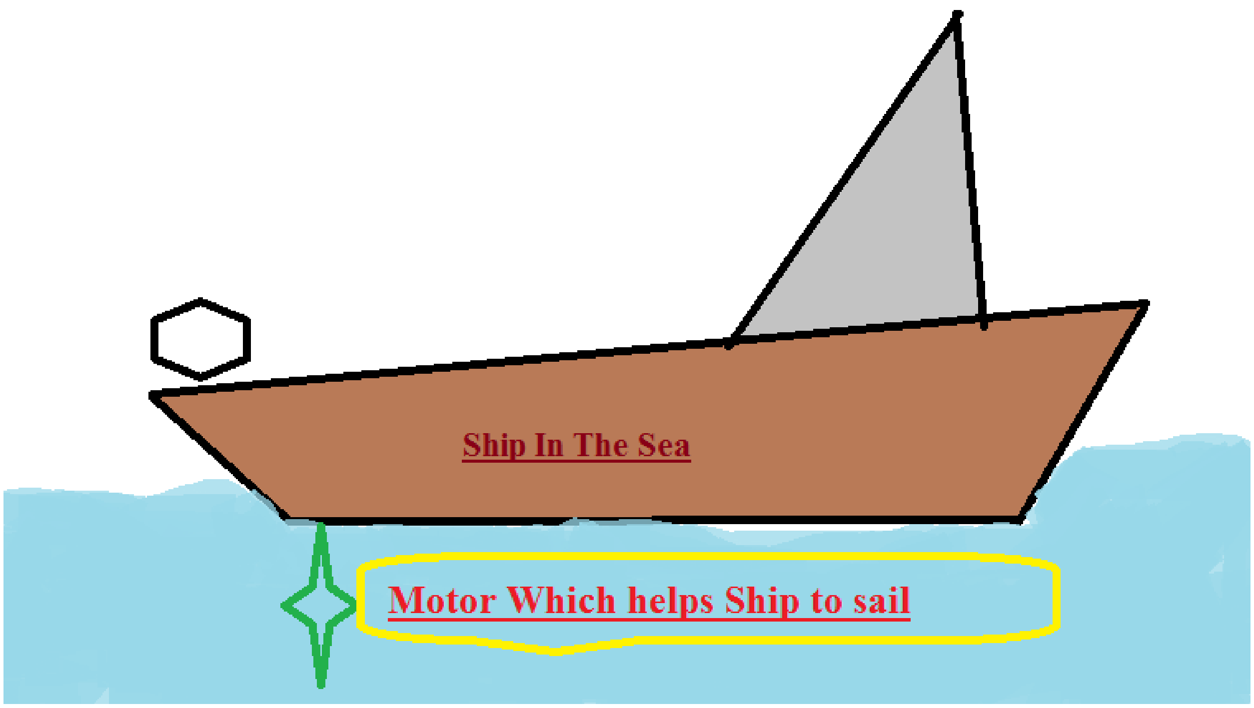
This paper aims to educate technicians in the area of shipbuilding. It is a multifaceted paper involving professional development for the college faculty and secondary school teachers, curriculum and education material enhancement, and career pathways development. This paper will improve and modernize technician education at the college level by preparing faculty and students with industry-relevant instructional modules and by attracting secondary school students to these educational pathways to provide a highly competent, well-educated and advanced workforce for the shipbuilding industry. The paper also increases career awareness about the shipbuilding-affiliated workforce with active learning kits while presenting a positive image of the industry. Its mission is to focus and coordinate member resources on issues, challenges, and opportunities facing the ship industry in Louisiana and across the nation. By way of these entities' long and productive partnership with their collaborators, comprehensive industry needs have been surveyed and validated through literature review. This will familiarize students with technical knowledge about the critical technologies of Virtual Reality, Lean Manufacturing and Rapid Product Development. Such familiarization will equip students with the necessary knowledge and skills, thus helping them to tackle practical problems when they join the industry.

Keywords: Shipbuilding, Marine Engineering, students, teachers, Lean Manufacturing.

2.0 Introduction:

As we all know Shipbuilding is not an easy task. It requires immense planning and perfect execution in order to make the ship sustain any situation. This is why having good knowledge in building a Ship is very essential. The responsibility of Shipbuilding engineer is not only to build and design ships in the shipyards, but also needs to perform maintenance and repair work on a ship making sure that it can withstand and overcome any situation resulting in smooth sailing in the sea at any point of time. There are few universities which offer education on building a ship and steps to maintain it, to give the engineers a broader idea and motivate students to Shipbuilding Engineering is the main idea of this paper.

A person responsible for designing a ship, performing operation and maintenance of ship's engines is called Marine Engineering. A Marine Engineer must be capable of dealing with many obstacles like physical space constraints, extreme weather conditions, deep water, and remote locations. Having this constrains in mind an Engineer has to develop a high quality operating platform which can adapt any circumstances. Ships are very important means of transport for imports and exports so duty of a Marine Engineer is not complete once he builds a ship. In fact the responsibility doubles in maintaining and operating the developed platform time to time. For maintaining the ship and designing it to sail safely from source to destination is the major task. The Engineer should be intelligent to judge the possible design problems and environment problems which they have to overcome. In order to safeguard the ship from all disasters.



3.0 Key Aspects for Students who wish to become a Marine Engineer:

- Huge scope of development as it is an never ending studies
- Huge demand in the market as they are very few marine engineers
- Very interesting field of study
- Should have good knowledge on terminology and functionality of building Ships
- Need to have in-depth knowledge on physics and mathematics
- Should be accurate in measuring the speed, length and distance to be maintained
- Make sure the ship is never out of fuel and is always under safe conditions
- Be prepared to sail for long periods of time when needed
- As Ships are the major means of transportation for exports and imports a Marine engineer will directly be serving a huge contributor for their Nations Pride.

4.0 Steps to follow before building the ship:

4.1 Purpose for building the Vessels

Before starting to build a ship an Engineer should analyze deeply about all functionality and working of the vessels. Should know about the history of the ships, its success rates its failures. The purpose for building the ship and contribution to the advancement by building the ship. It is this gathering of bits and pieces of evidence that will create a defined outline of what the ship should or could have looked like.

4.2 Design the Ship

In order to build the ship with good effect it should have a perfect design. An outline structure should be designed of how a Ship should look like what is the capacity it can withstand, what is the distance it can travel, what should be the time incurred in developing the ship, the cost to be spent in building the ship. This perfect outline of Yacht and the vessels should help the engineer to proceed in building the ship according to the sketched plan.

4.3 Building of the Vessels.

Having the detailed structure towards the construction of ship. An Engineer should now focus on how to build a Ship allowing it to sail in any given conditions. Important factor which an engineer needs to remember is to measure the kinetic energy it needs to possess, the fuel is can hold, measure the time it need to take for sailing from one given point to destination. Starting with the construction of ship very importantly the platform which is the base of the ship should be very strong and according to the length we analyzed during design phase.

4.4 Rigging and Hoisting the Sails

Sailing of the ships are not designed or modeled earlier. It is the engineer's choice to sail in the best possible ways. The sailing process should involve the efficient way of using the fuel, best ways to overcome from disasters. It is the engineer's responsibility to make sure the ship is not over flooded with many round tripping mechanisms. Engineer should make sure the ship sails to the desired destination without any problems before he starts sailing in the sea.

4.5 Shipping

Shipping includes of either exchange of goods from one source to destination or moving passengers from one place to another. There are some ships which do only the specified work for which it is constructed, this is planned during the design phase. If the ship is modeled for exchanging goods then it is designed with much harder platform and bigger capacity. If the vessel is used for transporting passengers then it is built with less motors and increasing speed of time to reach destination.

4.6 Maintenance

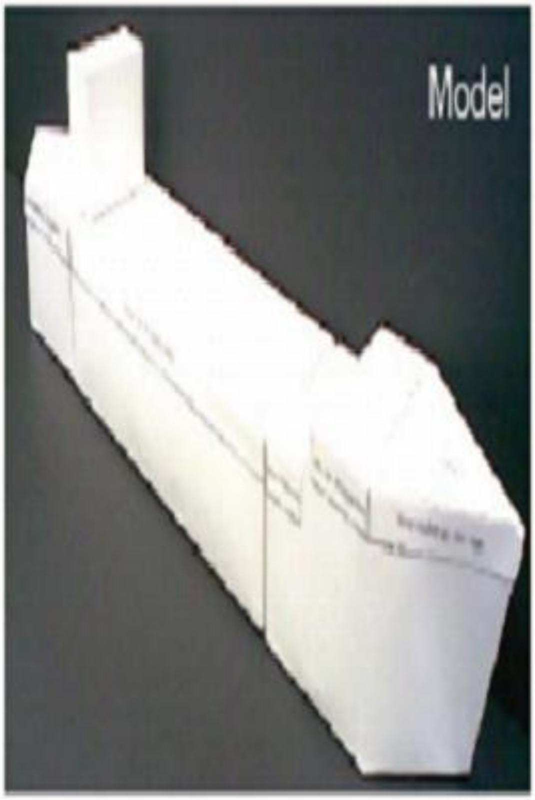
Maintenance is an important phase of engineering. The responsibility of Marine Engineer is not complete only when he builds a vessel. It is very important to maintain it in every phase to make sure the ship is always safe from any disasters. A ship should go on sailing only after the engineers are sure on how to maintain it if there are any problems during the journey. There are many ways which effects and halts the ships from sailing like lack of fuel, environment changes, overloaded ship, fire in the ship, many more an Engineer should have mechanism to overcome every possible disasters ensure the ship and the possession inside the ship are always safe.

5.0 Examples

Example 1

IDEA	Building a Ship
Objectives	Students are placed in different departments to work on their assigned tasks
Responsible Tasks	Students have to calculate the Weight, draft and kinetic energy of the ship
Learning skills	Students learn about the components of a ship, operations within a shipyard, methods of ship construction, design calculations

Material Requirements	Teacher’s manual, Student handout, student book, scissors, rulers, glue sticks, protractors and tape.



Example 2

IDEA	Construct a Submarine
Objectives	Students with enact as employees of PQR yard
Responsible tasks	Identify the number of parts required in building a submarine, construction of a submarine and cost of a ship

Learning skills	Students will learn basics about ship, about submarines and the construction methodologies, How to estimate the cost and how to measure the required number of parts
Material Requirements	Teacher's manual, students handout, students book, cost estimator, ruler

Example 3

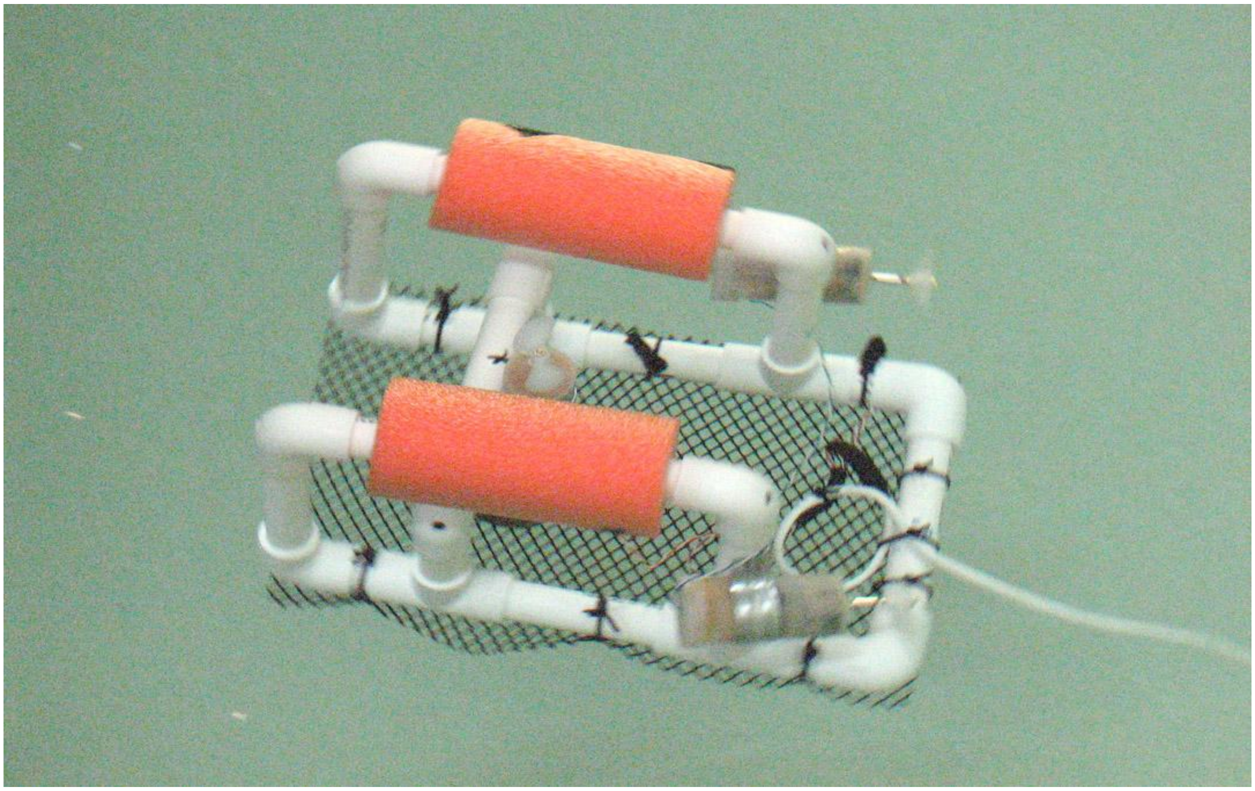
IDEA	Identifying possible causes for disasters
Objectives	Students should be able to estimate the possible causes for disaster and the immediate preventive measures of safeguarding the ship

Responsible Tasks	Students have to imagine as Disaster agencies and analyze what are the factors that cause disasters Should model the design of the ship in such a way it can overcome the disasters
Learning skills	Students learn about the fundamental components of the ship. Importance of a strong design and preventive methods for a ship.
Material Requirements	Teacher's manual, Student handout, student book, Disaster agency manual.

Example 4

IDEA	To build an Sea perch robot

Objectives	Students should be able to understand the underwater robotics program and ocean engineering principles
Responsible Tasks	Students need to build an underwater Remotely Operated Vehicle (ROV) which engineers all the operations done inside the sea.
Learning skills	Students learn about the fundamentals of Robotics, underwater functionalities, resources and live beneath the sea.
Material Requirements	Teacher's manual, Student handout, student book, scissors, glue, Artificial Intelligence material for building the robot.



6.0 Conclusion

In this paper we have presented the various benefits to the students by learning Marine Engineering. In purpose of creating interest for students towards ship building we have taken many examples to show how a ship is build what will be the responsibilities of the teacher and the student. We have illustrated about the challenges and advantages of marine Engineering, what are various aspects an engineer will look into before building the ship and the steps to follow to overcome any unconditional occurrence of disasters. Ship Building is very interesting concept to learn as till today ships are major means of transportation to exchange goods. This field is also best for the people who like traveling many place. In this advanced world, where we need to have immense knowledge towards a subject to prove ourselves eminent in the society Marine Engineering is really challenging and brings huge respect to us as we will be part towards our nations pride.

7.0 References

- http://www.nebhe.org/wp-content/uploads/ETOP-2013-Paper-Student-Reactions-to-PBL-in-Photonics-Technology-Educaton.pdf
- Donnelly, J., Massa, N.M. (2011). Problem Based Learning in Sustainable Technologies: Increasing the STEM Pipeline. Proceedings of HI-TEC, San Francisco, CA.http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- Massa, N.M. (2011). Creating Real-World PBL Challenges in Sustainable Technology to Increase the STEM Pipeline. Proceedings of the National Association for Workforce Improvement (NAWI), Burlington, VT. http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- http://www.ageofsail.net/aospro1b.asp
- Dr. Alok K. Verma P.E., CMfgE Technical POC and PI " Project- Based Learning "
- . Massa NM, Dischino M, Donnelly J, Hanes F. (June, 2010). Problem-Based Learning in Sustainable Technologies: Increasing the STEM Pipeline. Transactions of the

American Society for Engineering Education Annual Conference, Louisville, KY. http://www.nebhe.org/wp-content/uploads/ASEE-2010-Paper.pdf.

- http://bie.org/?gclid=CKOP4PbyiLwCFZRr7AodizwAkw
- https://www.tsl.texas.gov/sites/default/files/public/tslac/ld/projects/trc/2007/manual/craftillos/boat.jpg
- Massa, M., Dischino, M., Donnelly, J., Hanes, F. (Jan., 2009). Problem-Based Learning in STEM Education. Proceedings of the ASTE International Conference, Hartford, CT. http://www.nebhe.org/programs-overview/photon/conference-papers/.
- Donnelly, J. "Problem Based Learning", OPN May 2008, Optical Society of America..
- Morrisville, NC: Lulu Publications. http://stores.lulu.com/photon2.
- Hanes, F., Donnelly, J., Massa, N.M., Audet, R.H. (Oct., 2007). PHOTON Problem Based
- Hanes, F., Donnelly, J., Massa, N.M., Audet, R.H. (Oct., 2007). PHOTON Problem Based Learning (PBL): Using “Real-World” Problems to Develop Optics/Photonics Challenges for Students. Hands-on Science, Bucharest, Romania. http://www.nebhe.org/programs-overview/photon/conference-papers/.
- Hanes, F., Donnelly, J., Massa, N.M. (Oct., 2007). PHOTON2: Optics/Photonics Education Using Web-based Learning. Hands-on Science, Bucharest, Romania. http://www.nebhe.org/programs-overview/photon/conference-papers/..nebhe.org/programs-overview/photon/conference-papers/.
- https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&docid=-19_-EBUAjfHxM&tbnid=KQ1Is38cwwjcqM:&ved=0CAUQjRw&url=http%3A%2F%2Fwww.societyofrobots.com%2Frobotforum%2Findex.php%3Ftopic%3D13178.0&ei=l1XcUueuFqnY2AXQo4DgCA&bvm=bv.59568121,d.b2I&psig=AFQjCNECXXaubVF0qClMPmsam7sYwVtlmg&ust=1390257933745084
- http://www.seaperch.org/what_is

Design and Implementation of an Unmanned Ground Vehicle

Shreyas H, Thirumalesh H S
Department of Electrical and Electronics Engineering, SJCE, Mysore, India
Email: shreyas9693@gmail.com, hsthirumalesh@gmail.com

Abstract

There has been an increase in the rate of human crimes in today's world prompting safety measures to be taken without risking human lives. Robots can help us achieve this goal. In this paper we present the design of an efficient and economical robot that can perform both spying and rescue operations. This robot is an Unmanned Ground Vehicle (UGV) that can monitor enemy attacks and intruders secretly as if it were a first person to map an unknown location.These robots also find application in factories to move parts around on the floor. A UGV is basically a locomotive robot that can be controlled without human presence in the vehicle. This is done via wireless networking. The robot is controlled through a dual tone multi frequency (DTMF) call made to a mobile phone placed on the robot from the operator. The HT9170B DTMF Decoder IC is used to decode the dial tone into binary sequence. The decoded signals are sent to the ATMEGA8 microcontroller which generates appropriate output signals that are predefined in the microcontroller by a program. Since the robot is controlled through mobile network, the range over which it can be controlled is significantly large. The mobile also ensures real time video transfer from the robot's surroundings to the operator's display unit, thereby reducing the risks involved in direct human interaction with hostile environments.

***Index Terms*-DTMF, UGV, video transmission.**

I. Introduction

The deployment of autonomous vehicles is rapidly becoming possible in today's world because of technological advances in networking and in miniaturization of electromechanical systems. Robots can coordinate their actions through communication networks,and can perform challenging tasks such as search and recovery operations, manipulation in hazardous environments, exploration, surveillance, and also environmental monitoring for pollution detection and estimation [1]. An unmanned ground vehicle (UGV) is basically any piece of mechanized equipment that moves across the surface of the ground without an onboard human presence [2]. The potential advantages of employing such robots in espionage are numerous. They can spy on people in ways people can't move and from views humans can't reach.They can perform tasks faster than humans with much more consistency and accuracy. They can also capture moments that are too fast for the human eye.

In this paper, the design and implementation of a UGV which leverages the already existing dual tone multi frequency (DTMF) technology to establish a secure, fast and reliable connection with the operator's display unit is presented.The robot is controlled through a DTMF call made to the mobile phone placed on the robot from the operator. The HT9170B DTMF Decoder IC decodes the dial tune into binary sequence. These decoded signals are sent to the ATMEGA8 microcontroller which generates appropriate output signals that are predefined in the microcontroller by a program. The mobile is also used to transfer video in real time from the robot's surroundings to the operator's display unit.

The rest of this paper is organized as follows: Section II contains the hardware description of the robot, Section III contains the software description. The testing of the robot and results obtained are presented in Section IV and the conclusion is presented in Section V.

II. Hardware Description

The hardware comprises of:
1. The vehicle
2. An ATmega8 microcontroller
3. An Android mobile
4. A DTMF decoder

1. Vehicle :

The vehicle comprises of a fiber body chassis which houses two geared DC motors rated at 3V each. The two wheels are coupled to the shaft of the motor. A picture of the robot is shown in Figure 1. There is an arrangement made to place the android mobile on top of the vehicle. The agility of the vehicle is improved by making it as compact and light as possible. The motors are powered by a 5V regulated supply from a 9V DC rechargeable battery . The CD7805 IC regulates 9V DC to 5V DC [3]. The L293DNE IC is used to drive the motors by providing sufficient current for their operation [4].

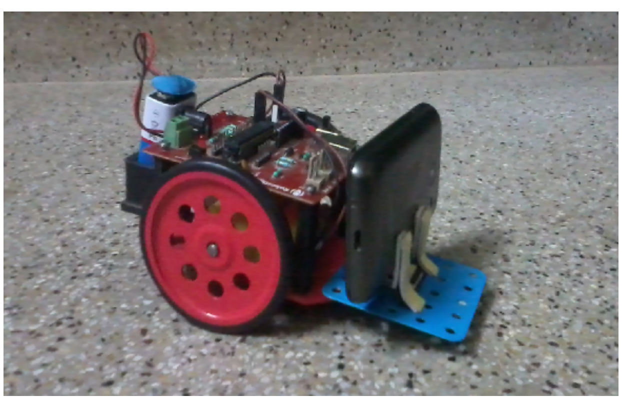


Figure 1: Robot

2. ATmega8 Microcontroller :

The microcontroller used is ATmega8A-PU, an 8-bit microcontroller from Atmel [5]. It controls the robot's overall movement. Based on the program written by the user onto this microcontroller, appropriate control signals are sent to the motors causing the robot to move as per the given command. The Atmega8 microcontroller has three PORTS, namely B , C and D. We utilize PORTB to send control ouput signals to the motors. PORTC receives external signals from the decoder in the form of a binary sequence. An onboard clock of frequency 4 MHz is used to provide clock pulses to the microcontroller.

3. Android Mobile :

The mobile is placed on a panel provided on top of the robot. It serves two purposes:
• It receives DTMF tone from the operator's mobile and sends them to the DTMF decoder.
• It records the video and sends it to the user's laptop using an android application called IP Webcam.
The photo resolution and the video resolution depend upon the quality of the camera in the phone that is placed on the robot.

4. DTMF Decoder :

The decoder module used is HT9170B [6]. The input to this decoder module is the DTMF tone. It converts the DTMF tone to it's corresponding 4-bit binary sequence. As shown in the DTMF frequency matrix in Figure 2, when a key is pressed, a particular set of two frequencies superimpose over one another. For example, when '1' is pressed frequencies of 697 Hz and 1209 Hz superimpose over one another and an audio tone unique to that number is generated. The decoder module recognizes the corresponding key pressed on the mobile on receiving the DTMF tone and produces a 4-bit binary sequence output.

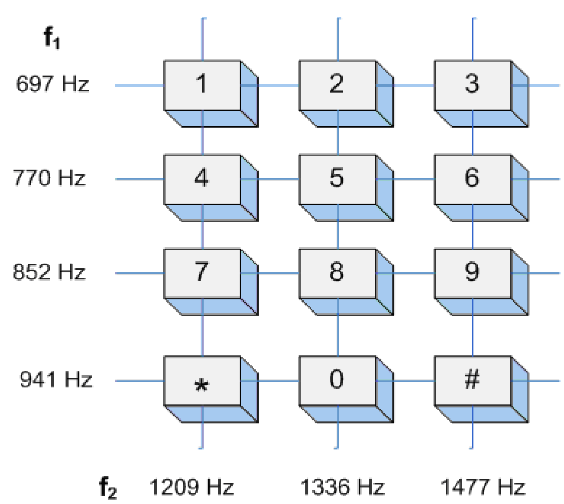


Figure 2: DTMF Frequency Matrix

III. Software Description

Integrated Development Environment (IDE) AVR Studio 4 has been used to write the C-codes and convert them to HEX files which are later burned onto the Atmega8 microcontroller using the AVR programmer - Khazama. An android application called IP Webcam is used for real time video transfer from the mobile phone on the robot to the operator's display unit. In order for this application to work, both the laptop and the mobile phone must be connected to the same network. This application provides the user with numerous modes to view the real time video such as Java and Flash. Options are also provided to enhance the stream quality and zoom in for a closer view of the surroundings. The night vision option can be used in places where there is insufficient light.

This application is first installed on the operator's mobile phone. The IP address provided in the application is entered on the browser of the laptop, thus opening a window which enables the user to view the real time video provided by the mobile phone placed on the robot. Then the user can choose to view the video as he/she pleases. Figure 3 shows a flowchart which represents the interaction between the robot and the software.

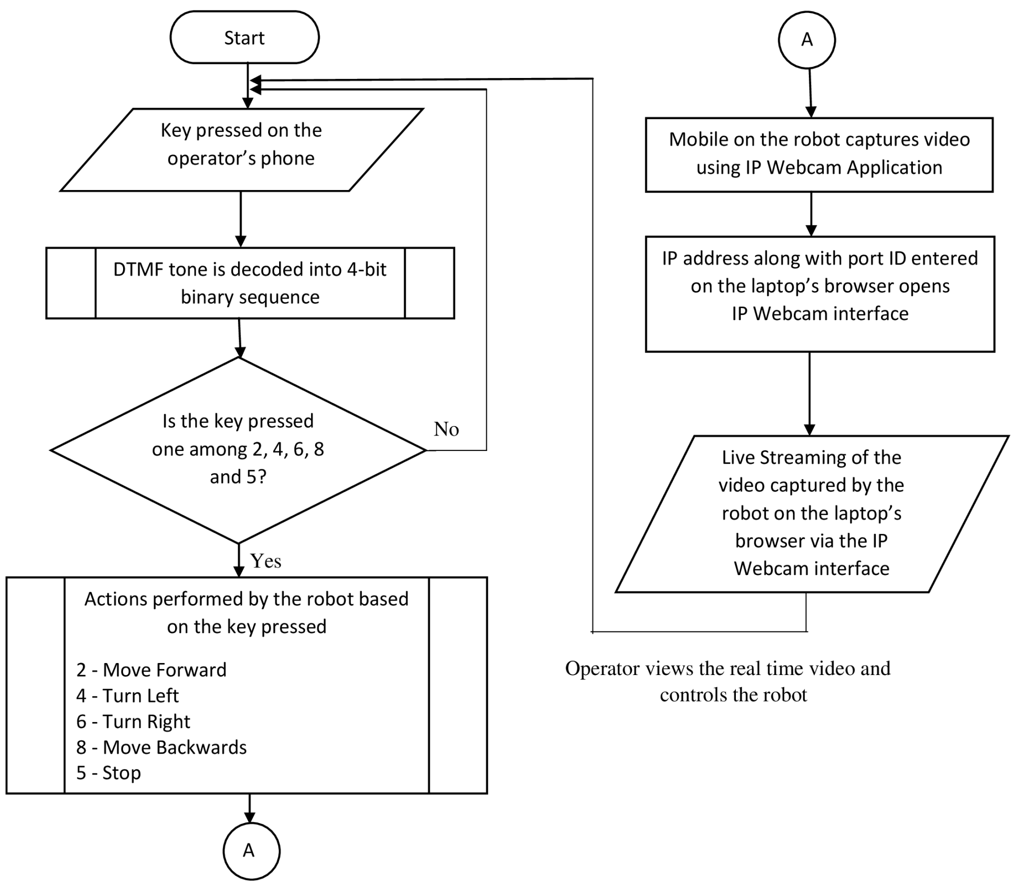


Figure 3: Flowchart - Interaction between the robot and the software

IV. Testing and Results

The C-program is written such that the robot initially waits for the signal sent to it by the operator on pressing a certain key on the mobile phone. The mobile phone and the operator's laptop are connected to the same Wi-Fi network. The operator can watch the live video delivered by the robot on his/her laptop's browser with the help of an android application called IP Webcam. The operator has to open the link http://<local ip address>:port id on a browser. On opening the link, the android webcam server is seen. Then the operator has to select the way in which he/she wants to view video. Figures 4 and 5 represent screenshots taken on the operator's laptop from the real time video delivered by the mobile phone on the robot in the presence of light and absence of light (night vision) respectively.

Figure 4: Screenshot taken in the presence of light

Figure 5: Screenshot taken in the absence of light (night vision)

On pressing a certain key from the operator's mobile, the corresponding DTMF tone is sent to the android mobile. The DTMF tone as heard from the android mobile is sent to the DTMF MT8870 module via a 3.5mm headjack . The table shown in Figure 6 lists the actions performed by the robot on pressing the following keys on the mobile phone.

Key Pressed	4-bit Binary Sequence	Operation performed by the robot
2	0010	Move Forward
4	0100	Turn Left
6	0110	Turn Right
8	1000	Move Backwards
5	0101	Stop

Figure 6: Operations performed by the robot according to the key pressed

The robot can be controlled at any point of time as the video displayed on the laptop's browser is real-time. In order to test the robot, both the laptop and the mobile phone on the robot were connected to a Wi-Fi network and the range of video transmission was found to be roughly 30 meters in radius. The working of the robot is represented in the form of a block diagram in Figure 7.

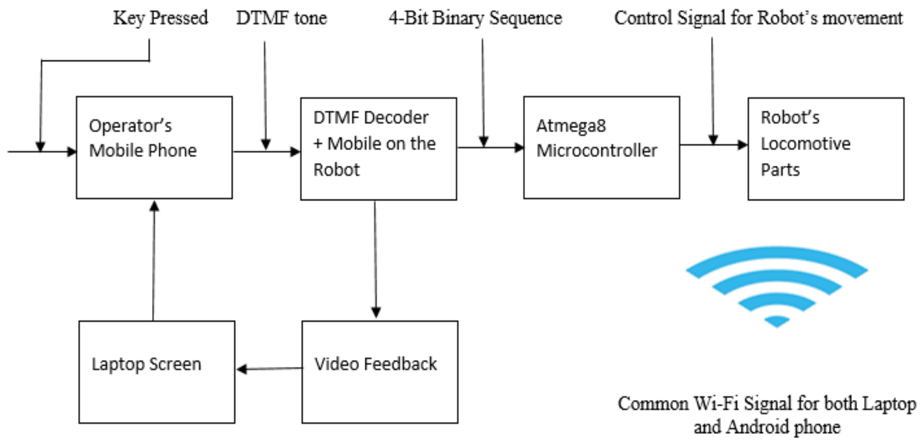


Figure 7: Block Diagram - UGV Control

The android mobile that is on placed on the robot is connected to the HT9170B DTMF decoder via a 3.5mm audio jack. Once a key is pressed on the operator's mobile a unique DTMF tone is generated which is decoded into a binary sequence by the HT9170B DTMF decoder. The microcontroller comprehends the binary sequence and causes the robot to operate accordingly. The mobile is configured to operate in auto-answer mode. Referring to the application circuit of the Decoder IC HT9170B in Figure 8, D0,D1,D2,D3 are it's outputs which are connected to the pins 23(PC0), 24(PC1), 25(PC2), 26(PC3) of the ATmega8 microcontroller shown in Figure 9.

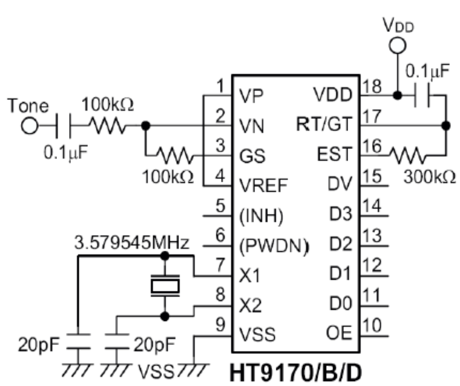


Figure 8: Application Circuit - IC HT9170B

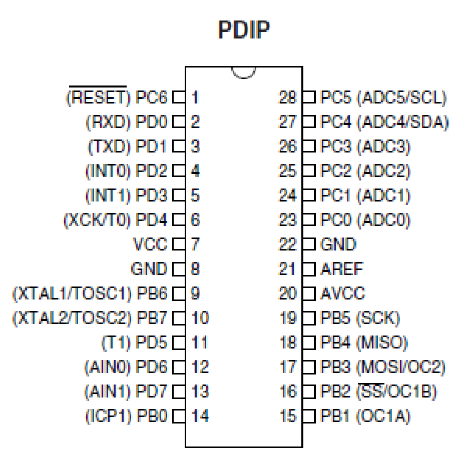


Figure 9: Pinout diagram - ATmega8A-PU

There is a negligible delay in the real time video delivered by the mobile on the robot. This can be adjusted by increasing or decreasing the stream quality. However, sometimes there is a delay of 4 seconds in the audio. This happens due to browser buffering and depends on the browser used. It can be resolved either by using the Flash renderer option provided by IP Webcam or by using a different browser.

V. Conclusion

This paper presents the design and implementation of a DTMF-based UGV. The vehicle was found to be capable of traveling over flat surfaces. It was also observed during test runs that the control signals for the UGV were received from the operator's mobile using DTMF signals as expected. A slight delay was observed in the response time of the UGV due to the execution speed of the algorithm in the microcontroller. The video output from the mobile placed on the robot was observed to be real-time as expected, but with a negligible delay.

Future enhancements/ modifications include: incorporating a wide area network (WAN) so that the range of the real-time video transfer can be extended and improving the mechanics of the robot so that it can travel over rough terrain.

References

[1] Jorge Cortes, Sonia Martinez, Timur Karatas, Francesco Bullo, "Coverage Control for Mobile Sensing Networks", IEEE Transactions on Robotics and Automation, Vol. 20, No. 2, pp. 243-255, April 2004.

[2] Phuoc-Nguyen Nguyen-Huu, Joshua Titus, "Reliability and Failure in Unmanned Ground Vehicle (UGV)", GRRC Technical Report 2009-01, University of Michigan, February 2009.

[3] http://cpc.farnell.com/1/3/7805-regulator

[4] http://www.ti.com/product/l293d

[5] http://www.atmel.in/devices/ATMEGA8A.aspx

[6] http://www.holtek.com/english/docum/comm/9170.htm

Mold Design for Injection Molding Using Additive Manufacturing

Jorge Rodriguez, Charles Crouch, Joseph DeMenter, Brian Guenther, Leah VanEeuwen

Department of Industrial and Manufacturing Engineering
Western Michigan University
Kalamazoo, MI 49008-5336

email: jorge.rodriguez@wmich.edu

Abstract

Prototype injection molding currently can take weeks and thousands of dollars to develop and manufacture. Using Additive Manufacturing to produce injection molds to rapid quote products can drastically reduce those costs and time frame. Functional prototype molds were developed and analyzed with CAD software, then printed on a 3D printer, and tested using a plastic injection molding unit. Calculations for cooling and mold compression were used to develop guidelines for future parts, with measurements taken to validate calculations and dimensions on the molded parts. Upon project completion, a company can rapidly and efficiently produce functional parts for quoting for prospective clients.

Introduction

Prototype plastic injection molds are used to produce short runs of parts to be utilized in quoting customer needs. Prototype injection molds can take weeks and tens of thousands of dollars to produce. The use of additive manufacturing to aid in the process of developing prototype plastic injection molds can reduce the lead time and overall cost of producing functional parts for quoting prospective clients [4, 5, 6].

Injection molding is the process of injecting molten plastic material into a mold cavity with the proper injection parameters to produce a functional part [10]. Plastic injection molds are primarily manufactured using tool steels which can withstand the high temperature environment. It is important to understand the methods and concepts of injection molding when developing an additively manufactured plastic mold that will be used as an insert for plastic injection molding.

Several variables need to be taken into consideration during injection molding, e.g., compression pressure, hold time, and cooling. Compression affects injection molded parts in terms of quality, due to part deformation caused by volumetric shrinkage. Volumetric shrinkage is the reduction of molecule size in a particular plastic when changing between molten and solid states. Compression also affects the proper sealing of the mold halves, without a proper seal flashing is likely to occur. This aspect was considered in our design by using an injection mold that is placed into a steel mold. The compression ratios then needed to be studied and adjusted in the design of the molds.

Additive manufacturing is the process of additively producing a 3D part through a layered process; it has also been previously referred to as rapid prototyping, rapid tooling, rapid technology, and layered manufacturing [1]. Recent developments have improved tolerances, thinner layers, and wider material selection. Although the 3D printing technology, particularly when referring to FDM - Fused Deposition Modeling - is well understood, there is always room for improvement [3]. But numerous technological developments nowadays are in the area of materials used in 3D printing. The overall goal in these developments is to move from the rapid prototype to a functional part.

A 3D printing company that has been a pioneer in developing printing technology and materials is Objet (now part of Stratasys Corp). Objet [7] applies polyjet technology that is similar to a inkjet printer using a print head that instead of printing water-based ink, lays down thin layers of material. The Objet has the ability to print in layers as thin as 0.0063 inches; this is thinner than a human hair. By being able to apply material at such thin layers, a higher resolution mold was able to be made. Access to an Object machine was available for this project [9].

The objective of this project is to perform design and testing for additive manufactured mold design. The designs will be determined based on part design, mold material, cooling channels, injection material, and mold life. Initial design guidelines for some geometry features will be studied. The goal of the project is to determine the feasibility of utilizing additive manufacturing for short runs, as well as quoting of functional injection molded parts.

Methodology

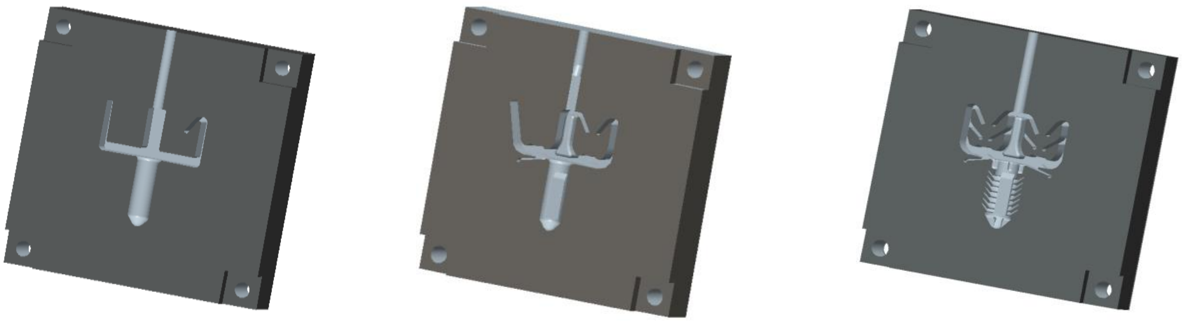
The main steps taken to complete this feasibility study were:

- Research process and material properties
- Create initial/modified designs
- Determine injection parameters
- Perform Moldflow and Finite Element Analysis
- Define cooling channel
- Do testing of printed inserts.

Some of the specific developments are:

Material Properties. For the temperature-resistant materials available for 3D printing using Objet, different material properties were obtain to help understand how the materials were going to react in the injection molding process. Some properties were provided by the printing company, but for some properties the information was not available, and further research provided approximated values. The material being used is VeroGray, with a Modulus of Elasticity of 3,000 MPa (compared to aluminum which is 70,000 MPa); the heat deflection temperature of Vero Gray is at 120°F (compared to aluminum which is 1,200°F). Poisson's Ratio was determined by looking at similar materials (i.e., Fullcure 720), and by performing compression testing.

Mold Design Progression. In designing the mold it was decided to do a progression of designs. With the progression of design , it started with basic features and progressively work our designs into the final part. The reason this was done is to help understanding the effects with a simple geometry and build upon it. This would also help in getting a better understanding of the molding process and deciding injection parameters. Three primary designs were developed for experimentation in this project; the designs went through a progression from basic to more complex (Figure 1).



(a) Initial Design *(b)Modified Design* *(c)Final Design*

Figure 1. Progression in Mold Design

The first design was made up of simplified features from the final part. The first design (Figure 1a) has two fingers that protrude upwards with a curved section to represent one of the 8 curved fingers on the final design. Rectangular and cylindrical features were tied into this design as well to give a simplified representation of the center recessed chamber and the christmas-tree retainer feature located at the bottom of the part. The modified design (Figure 1b) incorporated the rounded outside corners, a recess located where the primary center heat mass, as well as an angled side wall that was used to determine if there were any deflection. This design also added variable wall thickness as well as the tension finger that is located just below the main elbows where the curved fingers can be found. By removing material from the part design this reduced the overall volume of the mold cavity which in turn reduces that amount of heat transfer. The final mold design (Figure 1c) was the final test part. This design included all complex features such as the tension fingers, variable wall thickness, and christmas-tree retainer. On all three mold designs a 1.5° draft angle was used to help with manual ejection of the part from the molds.

Moldflow Analysis. This analysis was done to define the initial injection parameters for testing the designs. The initial parameters that were observed within the Moldflow results were clamping force, temperatures upon injection, volumetric shrinkage, VP switchover, expected cooling times, and melt temperatures. The first analysis with Moldflow was also done to determine what should be expected from our part once injecting the designated material into the additively manufactured injection mold. The final analysis with Moldflow was done using our final injection parameters that were defined through the injection molding tests. The final analysis results verified the experimental results that were being observed, the temperature at mold open was upwards of 447°F which was located in the exact locations were mold degradation occurred.

Finite Element Analysis. An FEA of the mold design was performed to determine the maximum deflection around the part cavity. The maximum deflection would help to determine the amount of material needed to add to the overall height of the mold in order for the molds to completely seal when they were injected. The clamping force is calculated with the part and runner area, the cavity pressure, the separating force, and the factor of safety. The parameters and equations are as follows:

- Projected part and runner area = 5.73 in^2
- Cavity pressure = 20 MPa (2,900 psi) - from Moldflow Analysis
- Separating Force = 8 tons
- Factor of safety = 1.25
- (PA) * (CP) = SF (SF) * (FS) = 10 tons

The FEA ran with 10 tons of clamping force on the front face of the mold, and 20 MPa of injection pressure on every face of the part cavity. The maximum deflection around the part cavity (Figure 2) was 0.15mm. This meant an additional 0.15mm to the overall height of our mold design.

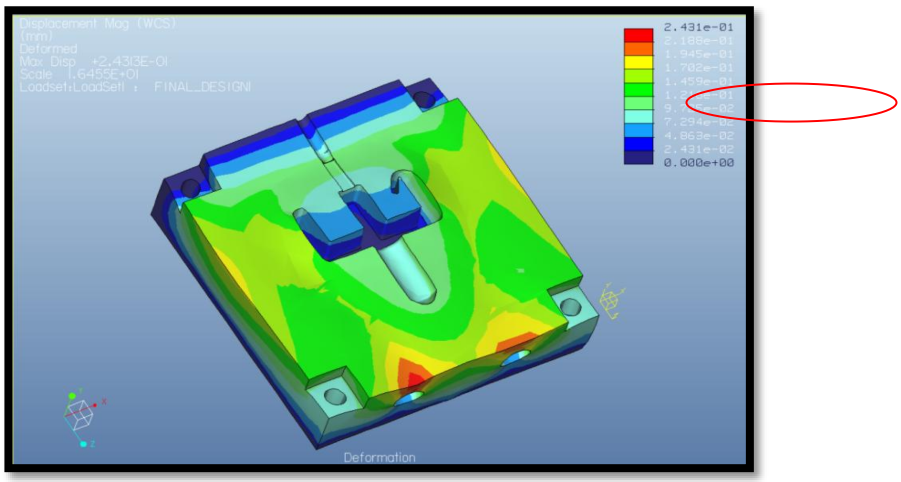


Figure 2. FEA Deflection Results.

After completing the final design an FEA was run again to verify that the maximum deflection around the part cavity did not exceed 0.15mm at the actual injection molding parameters used during testing [2]. For the final design the maximum deflection did not exceed 0.15mm.

Cooling Channel Design. The first step was to do heat transfer calculation based on the thermal properties on VeraGray (approximations) and Nylon and Polypropylene to estimate the amount of heat energy being introduced to the molds. Using the properties of cooling process water, and a static situation to simulate a worst case scenario, the volume the cooling channels needed to displace was calculated. The intention was to define the channels' cross-sectional area. Unfortunately, these simplified calculations did not provide feasible dimensions. Polypropylene offered better results, but still unfeasible for our mold size. A second analysis was done taking into account the cooling pump and dynamic factors, producing a feasible design for Polyprolpylene, a 0.25” circular cross section was specified.

Taking advantage of using printed mold inserts, this first design for cooling channel is a "conformal" design, where the channel follows the profile of the cavity in a 2D path. A second design was used, where the channels follow a 3D path actually going under the center of mass of the part as well as having a rectangular cross section, where higher thermal mass is expected, and therefore mold degradation was expected. In both designs analysis was done to structural integrity was ensured, and special access points were required to clean out the support material used for the channels during the printing process.

Results

Testing of three different designs was performed using a Millacron 100-ton injection molding machine. The testing protocol was to do injection, record parameters, inspect injected part and inserts, until failure of the printed molds. Do part inspection afterwards. The results are as follows:

Test Set 1. This set is for the initial design. Part material was both polypropylene and nylon. The molds tested with polypropylene had cooling channels, and the molds tested with nylon were tested with cooling channels and without cooling channels. The table below shows that the inserts were able to take several shots, that the cooling channels have an effect, but few quality parts are obtained.

	Polypropylene	Nylon	Nylon
Cooling	With	With	Without
# of Mold Halves	5	25	10
# of Shots	16	34	14
Quality Parts	1	2	4

Test Set 2. This set is for the modified design. Polypropylene and Nylon were used. In this test set polypropylene was tested without cooling channels, and nylon was tested with and without cooling channels. The table below shows a smaller number of acceptable parts, with inserts without channels performing better than with channels, for Nylon.

	Polypropylene	Nylon	Nylon
Cooling	Without	With	Without
# of Mold Halves	2	2	16
# of Shots	6	4	23
Quality Parts	2	0	3

Test Set 3. This set is for the final design. Polypropylene and Nylon were used. In this test set there were no cooling channels. the table below shows that even a smaller number of shots were possible, with polypropylene giving the only quality part.

	Polypropylene	Nylon
Cooling	Without	Without
# of Mold Halves	2	6

# of Shots	10	3
Quality Parts	1	0

Discussion

Throughout testing there were two materials that were primarily tested during injection molding. The main material tested was nylon, with nylon a total of 24 mold sets were tested, 11 with cooling and 13 without cooling. The average number of injection shots obtained before failure was 2.42 with cooling, and 2.53 without cooling. This result of having more shots being performed with molds that did not have cooling was due to the cooling lines having to be so close to the mold cavity that it lowered the integrity of the mold.

In testing with polypropylene a total of 6 molds were tested; 2 with cooling and 4 without cooling. The average number of injection shots obtained before failure was 4 with cooling and 5 without cooling. The reason for fewer mold sets being tested is the average mold life with polypropylene is extended due to lower temperature of material melting point. The result of having more shots with molds that did not have cooling was due to the cooling lines being too close to the mold cavity again which weakens the integrity of the material.

The two materials that were used, nylon and polypropylene, exhibited different results from one another during injection. With Nylon the cavity fills easily and transfer more detail from the mold to the part. Due to the higher injection temperature of Nylon, the mold would degrade after only one to two shots, with mold material melting and adhering to part. Nylon also required longer cycle times to allow the mold to cool down to extend mold life. Polypropylene was capable of producing features well but was difficult to get to flow into more detailed areas. This issue was due to the higher injection viscosity of polypropylene which made setting the injection parameters more difficult. Lower cycle times with polypropylene due to the lower injection temperature, and produce more parts per mold with the mold degrading after four or five shots.

Conclusions and Recommendations

Upon completion of the feasibility study on the use of AM techniques to manufacture injection molds it was determined that it is feasible to produce functional parts. There was evidence that AM molds have the potential to supply short production runs and produce functional prototypes. Throughout the testing it was possible to determine that the mold material VeroGrey was able to withstand the clamping and injection pressures required for injection molding, and there is a new material VeroGreen that has higher heat deflection temperature, expected to improve thermal life.

Through the testing it was concluded that cooling had no significant impact on mold life or number of quality parts produced. Due to the low thermal conductivity of the mold material, for cooling to be effective at removing heat from the mold the channel needed to be too close to the part surface, causing a thin wall between the mold cavity and the cooling channel. This thin wall would then crack under the pressures of injection, causing the cooling channels to become filled with injection material, thus destroying the mold.

The following recommendations are made: a) to have exact material properties, this is of vital importance for a project of this nature. In order to have valid prediction of performance there is need to have reliable properties; b) redesign inserts to accommodate cooling channels, the structural integrity after thermal fatigue is believed to be the main cause of insert failure. Cooling channels should help when properly implemented.

Bibliography

1. *3D Printing*. (n.d.). Retrieved 7 14, 2012, from Wikipedia: http://en.wikipedia.org/wiki/Additive_manufacturing
2. Dickens, P. (2000). Failure Mechanism Analsyis of Stereolithographic Injection Mould Tooling. *Rapid Prototyping Journal*, 62-74.
3. DLG. (2012). *Digital Lab Graph*. Retrieved Feb. 12, 2012, from http://factoryoffactories.com/rapidtool.htm
4. Jacobs, P. F. (1996). Recent advances in rapid tooling for Stereolithography. *Proceedings of the national conference on Rapid Prototyping and Tooling Research* (pp. 18-19). Buckinghamshire: Buckinghamshire College.
5. Jacobs, P. F., & Hilton, P. D. (2000). Rapid Tooling: Technologies and Industrial Applications. In P. F. Peter D. Hilton. New York: Marcel Dekker.
6. Nagahanumaiah, Subburaj, & Ravi. (2008). Computer aided rapid tooling process selection and manufacturability evaluation for injection mold development. *Computers in Industry, 59*(2-3), 262-276.
7. Objet 3D Printers. (2012, 6 5). *New Objet Materials: The power behind your 3D Printer*. Retrieved 2012, from http://www.objet.com/Portals/0/docs2/New%20materials%20data%20sheets_low%20res.pdf
8. *PolyJet Materials Data Sheet*. (n.d.). Retrieved 7 8, 2013, from http://www.stratasys.com/materials/polyjet/~/media/29592222B80C489BAC28803DB08C10E5.ashx
9. RayCE Americas.
10. Zheng, T. a. (2011). *Injection Molding: Integration of Theory and Modeling Methods.* Longdon, New York: Springer Heidelberg Dordrecht.

Project Based Learning for STEM (Science, Technology, Engineering, Mathematics) **Education**

Dr. Suren N Dwivedi

Department of Mechanical Engineering and Manufacturing

University of Louisiana at Lafayette

Suren7483@gmail.com

Lafayette, LA

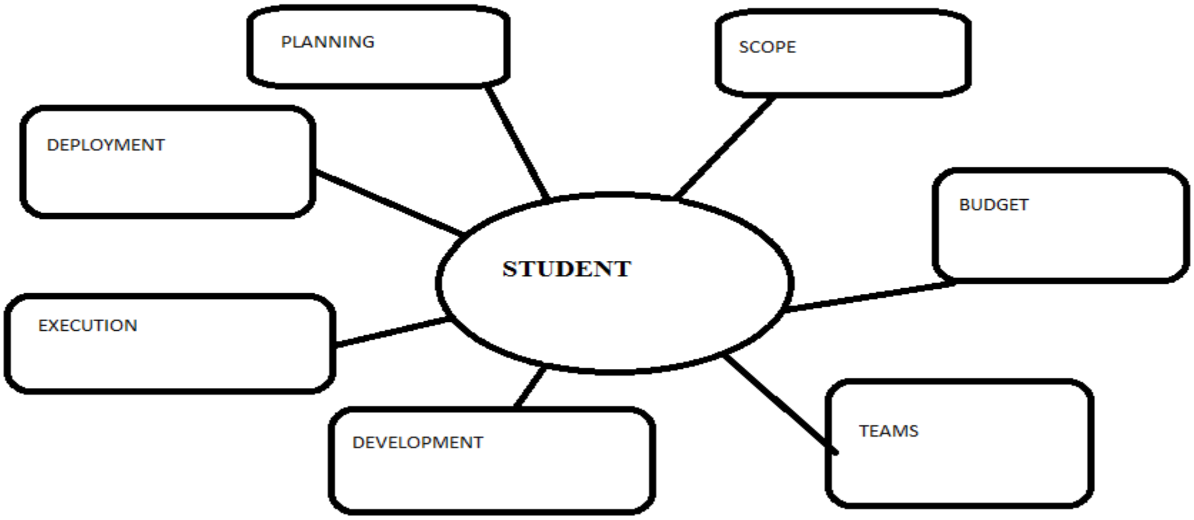
1. Abstract:

Project-based learning (PBL) for STEM Education is an interesting and effective way to learn about science, technology and mathematics. It is a technique used in making concepts understand for completing a goal with greater knowledge. The main purpose of PBL is to **learn the activities by implementing them**. Project-based learning is an approach emphasizing on teaching students by engaging them in investigation of the particular task and making them learn things by practically implementing it. In this way students get significant solutions to complex problems by asking and refining questions, sharing and debating ideas, experimenting and predicting solutions, gathering data and analyzing in teams, drawing conclusions, building leadership qualities, suggesting new proposals and broader methodologies to implement, creating artifacts. This approach will incorporate students with all the professional values needed for their successful career and make them technically flexible of facing any challenges with ease and with great effectiveness. The main aim of PBL is making students enjoy what they do which helps them remember about the lecture for greater depth, builds confidence to face any problem with ease, makes students learnt to work collectively in teams and build collaboration within their peers, makes students follow systematic approach of recognizing and solving a problem, builds competition within themselves for better results. This method brings great help for teachers as well to monitor every work what student does and guide them accordingly to complete their task with good effect ensuring students build their career with strong foundation.

2. Introduction

A project in contemporary world is defined as organized way to achieve a goal. In order to achieve an aim with effectiveness a project needs to have careful planning or proposal, scope, cost incurred, time desired for completion.

In this advanced world where students are engaged to learn many things it is becoming difficult for them to remember about the lectures they hear in class. As the days pass, in vain of learning new things students forget about the lectures taught by the teacher in class a week before. There are two reasons for it 1.No practice about the lecture learnt, 2.No time to implement the lectures learnt practically. This has become a huge concern for teachers, parents, students and the society. Then arise the question how can we make students remember about the lectures they learnt for long time? Is it possible for them to incorporate lessons to greater depth in this fast pacing world? Then the concept called Project Based learning was introduced giving solutions to all this questions. They strongly believe just by watching soccer in television we cannot become a good soccer player it is more important to practice and practically involve in playing the game in order to become a good player. This approach allows students to learn simultaneously by implementing them. In this way it gives student understand the lecture better, since they are doing practically they can any time approach the teacher for doubts during their project. They start taking responsibilities and focus on learning in an enjoyable and effective manner. This approach helps students know what they are doing, what is the need of doing, what is the goal and design the best effective solution. This phenomenon also helps teachers to contact students easily and know their progress better by seeing the work students accomplish in completing every module and task. It is also easy for them to clear the doubts of the students and make them understand concepts with great effect.



3. Advantages of BPL

How PBL helps in effective learning for Students?

- By the outcome of their recent developments and results many researchers have confirmed that Project-Based learning is the best and most enjoyable way to learn
- In this way a Student will actively engage themselves towards a project and incorporates lessons with interest.
- PBL helps a student to think and analyze about a task in a broader and effective way
- In this way students will not only get greater depth of knowledge on a subject but also build collaboration within their mates to achieve a goal.
- PBL helps students in taking responsible tasks at their own and solve problems with confidence
- Since teachers and other expects are also the members of the project it provides greater flexibility of utilizing there advertise and skills for the students.
- Students tend to propose new and different innovations to make their project more creative.
- PBL allows teacher to closely monitor the students work and help them to achieve their task with high quality.
- PBL helps the student to build leadership qualities and learn how to manage things.
- Project Based Learning technique helps students relatively to enter into problem

solving learning journey.

4. How to follow PBL approach?

4.1 Introducing about PBL

Initially teacher gives an overview to students what a Project Based learning is about. For what we are using PBL and make students familiarize about this approach. Teacher has to assign a project to students and guide them through out ensuring students are able to understand the concepts completely. Teacher has to make students understand all the protocols and the procedure needed to follow in order to develop their projects with best effect. Teacher will also mention their role in the project and explain students about their participation in completing the given project in time

4.2 Problem Overview

Next phase is to explain students about the task assigned to them. Teacher has to explain students about the problem, what are their responsibilities and their roles in accomplishing the task. It is the responsibility of teacher to make students understand the problem and create interest in solving them by illustrating with examples.

For example: Task is to build a ship

Teacher has to explain student about the ship and its advantages, about the procedure involved in building the ship, about various factors to consider while progressing the work, various precautionary measures while building a ship, risks involved, and the benefits students get by building this ship.

4.3 Analyzing steps for efficient solution

Teacher has to explain students about the steps to be followed before starting the project for efficient solution. This step is very important because it involves the constructive methods of how to solve a problem. This approach should be very interactive by allowing students to express their views and strategies in achieving their goal with efficiency

From the above example:

After teacher giving complete synopsis about their projects Students can express their various views by which they can achieve his/her aim. Every student has different methods of build the ship depending on the department they are put into. They can freely communicate their various ideas to the teacher about their task and analyze the best approach to complete the task

4.4 Estimating Cost and Time incurred for completing the task:

Teachers will explain students about various materials and equipment's required for completing the task. Depending on the materials required students along with their mentor should start estimating a desired cost that has to be put in for completing the task and time that is needed in completion of the project. The cost and time go in logger heads the more number of days needed for completing the project the more cost we invest.

From the above example:

After analyzing how to proceed further with building the ship. Teacher should explain students about the materials required in building ship and the use of every material used n building. In common for building a paper ship we require bunch of paper sheets, scissors, glue, work book, weighing machine to measure the weight of the ship according to its platform and to measure kinetic energy inside the ship. Students need to estimate the cost for these materials and the time by which they can complete building a ship by using the required material. Students need to take permissions from the teacher for the cost and time invested for the correct advice.

4.5 Designing of the problem:

In this step students will start implementing the plan to achieve their goal. According to the various steps designed earlier according to their responsibilities they start working on the modules assigned to them. Teacher will keep monitoring every students work and guide them accordingly ensuring every student will learn maximum knowledge during the project work

In case a student has a different design to follow in order to achieve his goal at any time he should be free to communicate with the teacher and discuss why he is planning to chance the plan which was implanted before.

From the above example:

Depending on their departments students will start building a ship. They start with the structural part one they measure the kinetic energy and weight the ship can withstand. Basing on these measure the student start building a ship bearing in mind to accomplish the project with success and uniqueness. In this way it builds competition within their peers and every student starts enjoying the work they do in learning about their project how to build a ship.

4.6 Submitting the solution

After completing the project students will submit their respective projects to the teacher for grades. It is easy for the teacher to judge the grade basing on the outcome of the project and how unique the project is being developed. In this way right from the child hood every student starts thinking differently and try to complete their task in a professional manner with maintaining quality.

4.7 Additional Innovations

The problem solving learning teaching will immensely help students to think beyond what is already existing and express their new ideas they found out during developing there project. In this way student will never forget what he/she has learnt during developing the project and also creates an interest in student for future implementing advanced features for the project which is developed already by them.

5 Advantages to the students:

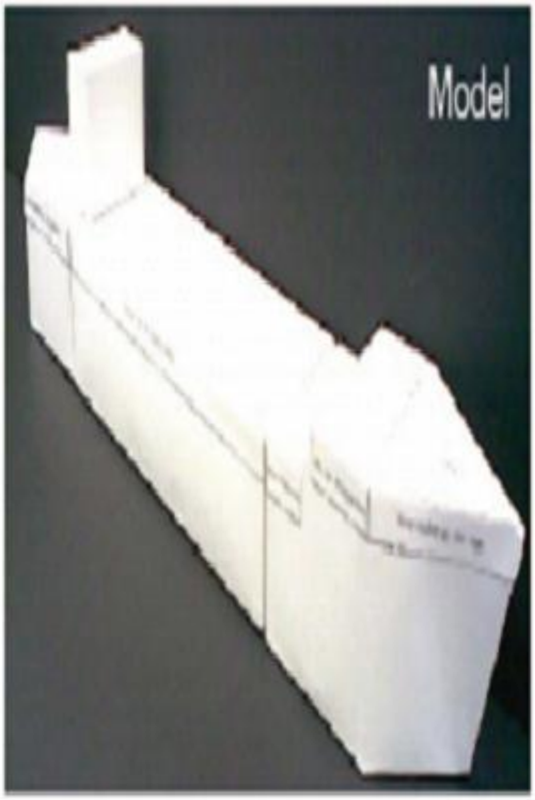
- Intense learning on the subject

- Builds systematical approach in solving a problem
- Builds confidence on how to solve a problem with ease
- Self- learning
- Self- thinking
- Self- analyzing
- Building responsibilities in solving a problem
- Collectively working with the peers
- Builds Leadership qualities
- Competition within the peers
- New idea for innovations

6 Examples

Example 1

IDEA	Building a Ship
Objectives	Students are placed in different departments to work on their assigned tasks
Responsible Tasks	Students have to calculate the Weight, draft and kinetic energy of the ship
Learning skills	Students learn about the components of a ship, operations within a shipyard, methods of ship construction, design calculations
Material Requirements	Teacher's manual, Student handout, student book, scissors, rulers, glue sticks, protractors and tape.



Example 2

IDEA	Construct a Submarine
Objectives	Students with enact as employees of PQR yard
Responsible tasks	Identify the number of parts required in building a submarine, construction of a submarine and cost of a ship
Learning skills	Students will learn basics about ship, about submarines and the construction methodologies, How to estimate the cost and how to measure the required number of parts

Material Requirements	Teacher’s manual, students handout, students book, cost estimator, ruler

Example 3

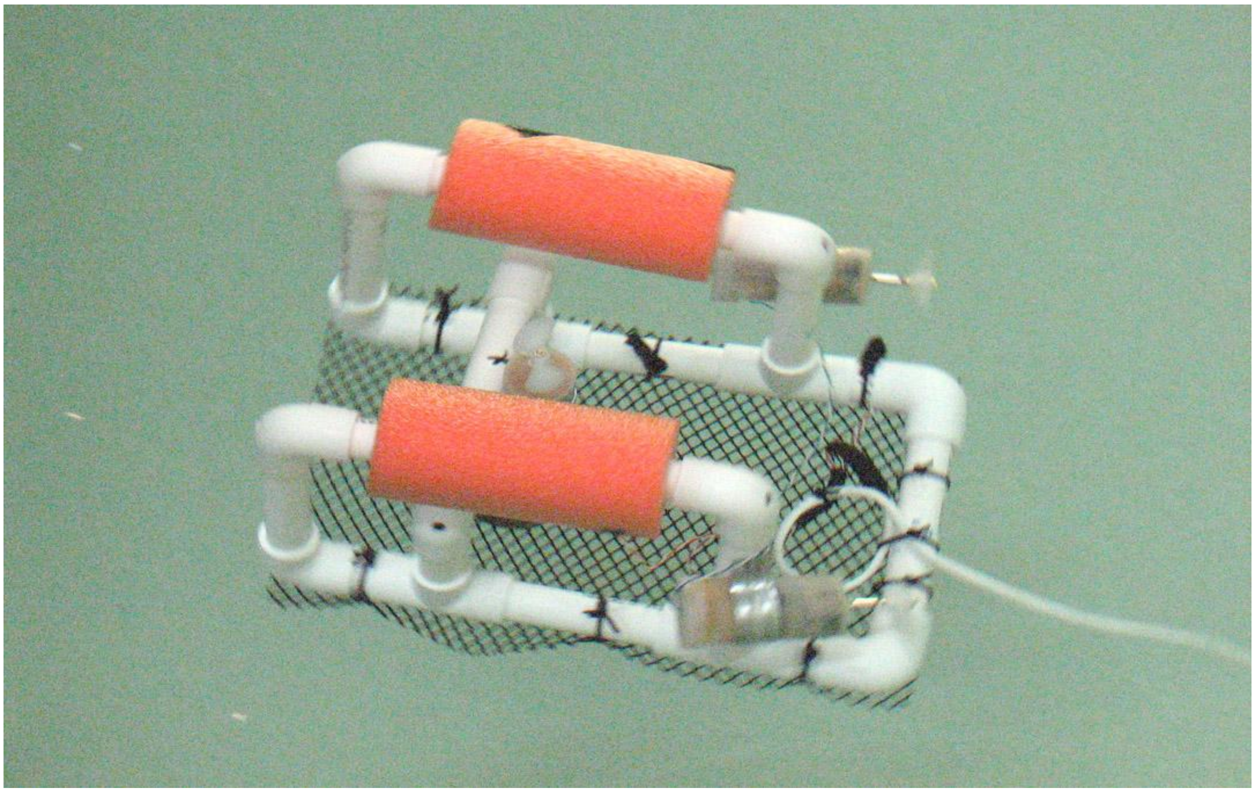
IDEA	Identifying possible causes for disasters
Objectives	Students should be able to estimate the possible causes for disaster and the immediate preventive measures of safeguarding the ship
Responsible Tasks	Students have to imagine as Disaster agencies and analyze what are the factors that cause disasters Should model the design of the ship in such a way it can overcome the disasters

Learning skills	Students learn about the fundamental components of the ship. Importance of a strong design and preventive methods for a ship.
Material Requirements	Teacher's manual, Student handout, student book, Disaster agency manual.

Example 4

IDEA	To build an Sea perch robot
Objectives	Students should be able to understand the underwater robotics program and ocean engineering principles
Responsible Tasks	Students need to build an underwater Remotely Operated Vehicle (ROV) which engineers all the operations done inside the sea.

Learning skills	Students learn about the fundamentals of Robotics, underwater functionalities, resources and live beneath the sea.
Material Requirements	Teacher's manual, Student handout, student book, scissors, glue, Artificial Intelligence material for building the robot.



7 Conclusion

In this paper we have presented how Problem based learning is an effective approach to make students like and enjoy what they are learning. In traditional way students listen to the lecture and practice later without much guidance and participation from the teacher. But by using problem solving learning method student and teachers will actively participate and ensures that the project is developed with intense care and focus. Students share their knowledge with their peers and make them understand the course what they are learning. There will not be any chance for the student to escape from learning the course as they are under continuous supervision of the teacher. Students will learn their work efficiently as they have to submit every work they put in for developing the module assigned to them. This approach builds professionalism in the student's right from child hood and ensure they have right strategies to achieve their goal.

8 Reference

- http://www.nebhe.org/wp-content/uploads/ETOP-2013-Paper-Student-Reactions-to-PBL-in-Photonics-Technology-Educaton.pdf
- Donnelly, J., Massa, N.M. (2011). Problem Based Learning in Sustainable Technologies: Increasing the STEM Pipeline. Proceedings of HI-TEC, San Francisco, CA.http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- Massa, N.M. (2011). Creating Real-World PBL Challenges in Sustainable Technology to Increase the STEM Pipeline. Proceedings of the National Association for Workforce Improvement (NAWI), Burlington, VT. http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- Dr. Alok K. Verma P.E., CMfgE Technical POC and PI " Project- Based Learning "
- Dischino M., DeLaura J.A., Donnelly J., Massa, N.M., Hanes, F. (2011). Increasing the STEM Pipeline through Problem-Based Learning. Proceedings of the Joint International Conference of the International Association of Journals and Conferences and American Society for Engineering Education, Hartford, CT. http://www.nebhe.org/wp-content/uploads/IAJC-ASEE-2011-Paper.pdf.
- Hanes, F., Donnelly, J. (2011). STEM PBL: Increasing the STEM Pipeline Through Problem-Based Learning. Proceedings of SustaiNext Summit East, Philadelphia, PA.

Web.http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.

- Massa, N. M., Dischino, M., Donnelly, J., Hanes, F. (June, 2011). Creating Real-World Problem-based Learning Challenges in Sustainable Technologies to Increase the STEM Pipeline. Proceedings of the American Society of Engineering Education (ASEE) Annual Conference, Vancouver, BC. http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- Donnelly, J., Massa, N. M., Dischino, M., Hanes, F. (Aug. 2010). Implementation Stories from the PHOTON PBL Project. Proceedings of SPIE Annual Optics & Photonics Conference, San Diego, CA. http://www.nebhe.org/programs-overview/professional-curriculum-development/about-stempbl/stem-conference-papers/.
- . Massa NM, Dischino M, Donnelly J, Hanes F. (June, 2010). Problem-Based Learning in Sustainable Technologies: Increasing the STEM Pipeline. Transactions of the American Society for Engineering Education Annual Conference, Louisville, KY. http://www.nebhe.org/wp-content/uploads/ASEE-2010-Paper.pdf.
- http://bie.org/?gclid=CKOP4PbyiLwCFZRr7AodizwAkw
- https://www.tsl.texas.gov/sites/default/files/public/tslac/ld/projects/trc/2007/manual/craftillos/boat.jpg
- Massa, M., Dischino, M., Donnelly, J., Hanes, F. (Jan., 2009). Problem-Based Learning in STEM Education. Proceedings of the ASTE International Conference, Hartford, CT. http://www.nebhe.org/programs-overview/photon/conference-papers/.
- Massa, N.M., Dischino, M., Donnelly, J., Hanes, F. (Aug., 2008). Problem-Based Learning in Photonics Technology Education. Proceedings of Optics and Photonics, Vol. 7061, SPIE, August, 2008. http://spie.org/x648.html?product_id=797135
- Donnelly, J. "Problem Based Learning", OPN May 2008, Optical Society of America..
- Morrisville, NC: Lulu Publications. http://stores.lulu.com/photon2.
- Hanes, F., Donnelly, J., Massa, N.M., Audet, R.H. (Oct., 2007). PHOTON Problem Based

- Hanes, F., Donnelly, J., Massa, N.M., Audet, R.H. (Oct., 2007). PHOTON Problem Based Learning (PBL): Using "Real-World" Problems to Develop Optics/Photonics Challenges for Students. Hands-on Science, Bucharest, Romania. http://www.nebhe.org/programs-overview/photon/conference-papers/.
- Hanes, F., Donnelly, J., Massa, N.M. (Oct., 2007). PHOTON2: Optics/Photonics Education Using Web-based Learning. Hands-on Science, Bucharest, Romania. http://www.nebhe.org/programs-overview/photon/conference-papers/..nebhe.org/programs-overview/photon/conference-papers/.
- https://www.google.com/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&docid=-19_-EBUAjfHxM&tbnid=KQ1Is38cwwjcqM:&ved=0CAUQjRw&url=http%3A%2F%2Fwww.societyofrobots.com%2Frobotforum%2Findex.php%3Ftopic%3D13178.0&ei=l1XcUueuFqnY2AXQo4DgCA&bvm=bv.59568121,d.b2I&psig=AFQjCNECXXaubVF0qClMPmsam7sYwVtlmg&ust=1390257933745084
- http://www.seaperch.org/what_is

