DEVELOPMENT OF POSITION FEEDBACK SENSOR BASED ON VISION USING NEURAL NETWORK

YUSHAZAZIAH MOHD YUNOS

Thesis submitted in fulfilment of the requirements for the degree of Master of Science

December 2009

ACKNOWLEDGEMENT

Alhamdulillah, a big praise to Allah who makes all this is possible for me. Putting my thesis together has been quite of a challenge for me, but I came to realized that this is another journey that I have to go through. There are some people who made this journey a great experience of my life.

I would like to express my deep and sincere gratitude to my supervisor, Assoc. Prof. Dr Zahurin Samad for his detailed and constructive ideas and comments through out my master study. His guidance and encouragement helped me in all the time of research and writing of this thesis.

Sincere thanks to Mr. Faizul Hawary and Ms. Nurulasikin Mohd Suhadis from School of Aerospace Engineering, for their willingness on spending some time to share some thought and knowledge on my research project.

I warmly thank School of Mechanical Engineering and its staff for all the help and facilities provided to me especially to Mr. Norijas, Mr. Abdul Hamid, Mr. Jamari, Mr. Hashim and Mr. Ali Shahbana.

I would also like to gratefully acknowledge the support of all my great friends that I met along this journey. They helped me immensely by giving me encouragement and friendship. It may not enough to put all the big names in this small space.

Special thanks to the greatest gift that I have in my life – my parents, family and friends. Their love and understanding has supported me spiritually. This thesis is dedicated special to my late sister, Yussahila which memory with her is always in my heart.

The financial support from University Kuala Lumpur is gratefully acknowledged.

TABLES OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

r_k	intensity level
n_k	number of pixel
G	number of intensity scale
m	row of image matrix array
n	column of image matrix array
$\bar{A}$	mean intensity value for image A
$\bar{B}$	mean intensity value for image B
r	correlation coefficient
x_j	input of neuron
y_j	output of neuron
w_{ij}	weight connection between neuron
E	sum squared error
y_{jc}	actual output
d_{jc}	desired output
Dw	weight change
p	number of input neuron
a	number of image in one set of image condition
b	number of input condition
Ex	position error
P	image position number
$Incx$	incremental movement
K_p	proportional gain
K_i	integral gain
K_d	derivative gain

K_{cr} critical gain

P_{cr} critical period

LIST OF ABBREVIATIONS

FMS	Flexible Manufacturing System
CCD	Charged couple device
BPNN	Backpropagation neural network
CNC	Computer numerical control
C	capacitance
LVDT	Linear variable differential transformer
LED	Light emitted device
DOF	Degree of freedom
CMOS	Complimentary metal oxide semiconductor
2D	Two dimensional
3D	Three dimensional
RBF	Radial basis function
PGA	Pneumatic group actuator
DFC	Direction flow change
USB	Universal serial
Pc	Personal computer
RGB	Red, green, blue
PCB	Printed circuit board
HSI	Hue, saturation, intensity
GRNN	General Regression neural network
PCA	Principal Component analysis
SONFIN	Self constructing neural fuzzy inference network
ARTMAP	Adaptive resonance theory MAP
ARTMMAP	Adaptive resonance theory mixture MAP

RT	Regression tree
S1	Number of neuron in single hidden layer
S2	Number of neuron in second hidden layer
PID	Proportional, Integral, Derivative
SSE	Sum squared error
trainlm	Levernberg-Marquadt algorithm
traingdx	Gradient descent with adaptive learning rate and momentum algorithm
trainrp	Resilient backpropagation algorithm
trainscg	Scaled conjugate gradient algorithm

PEMBANGUNAN PENDERIA SUAP BALIK KEDUDUKAN BERASASKAN PENGLIHATAN MENGGUNAKAN RANGKAIAN NEURAL

ABSTRAK

Pengukuran untuk kedudukan lurus adalah asas dalam pelbagai proses industri terutamanya sebagai penderia suapbalik kedudukan. Enkoder optik adalah antara penderia kedudukan yang banyak digunakan. Bagaimanapun, beberapa kekurangan didapati dalam penggunaannya. Kerja yang dicadangkan ini mengambil pendekatan dengan membina penderia kedudukan menggunakan pengelasan imej. Imej telah dikelaskan menggunakan rangkaian neural. Imej masukan yang disuap kepada pengelas ini adalah imej skala kelabu yang diambil daripada persekitaran sekeliling (pandangan atas) untuk mewakili kedudukan. Sifat-sifat di dalam imej tersebut kurang jelas dan menyebabkan pengekstrakan sifat sukar dilakukan untuk mengekstrak data statistik. Imej-imej telah diskalakecilkan dan dimasukkan ke dalam rangkaian sebagai satu vektor. Beberapa siri imej telah diambil pada kedudukan yang pelbagai di mana setiap siri mempunyai jarak antara imej yang berlainan. Apabila jarak antara imej ini lebih hampir antara satu dengan yang lain, kesan tindanan antara imej adalah besar. Empat siri imej telah dikaji untuk melihat kesan bilangan imej masukan terhadap rangkaian neural. Kesan kecerahan keadaan persekitaran yang berlainan juga telah dikaji untuk membolehkan sistem ini berfungsi pada keadaan persekitaran yang berbagai-bagai. Struktur rangkaian neural yang terbaik untuk menyelesaikan masalah pengelasan imej ini telah dikaji. Kadar pengecaman telah dikira untuk menunjukkan prestasi sistem. Daripada keputusan yang diperolehi, didapati bahawa rangkaian neural ini berjaya mengelaskan kedudukan imej lebih daripada 80 peratus. Rangkaian neural yang telah dilatih kemudiannya diuji pada pergerakan satu paksi robot gantri sebagai pengesan suapbalik kedudukan yang menggunakan penglihatan. Robot gantri bergerak kepada kedudukan diingini dengan

mengambil perbezaaan antara kedudukan sebenar dan kedudukan diingini. Ralat kedudukan ini kemudian dibetulkan menggunakan gandaan kawalan. Daripada eksperimen yang telah dijalankan, didapati kawalan perkadaran 10 berjaya mencapai objektif. Daripada keputusan yang diperolehi, ia menunjukkan sistem suapbalik penglihatan berkeupayaan untuk digunakan di dalam aplikasi pergerakan lurus satu paksi. Ini boleh disimpulkan bahawa pengelasan imej menggunakan imej kedudukan boleh diaplikasikan sebagai penderia suapbalik kedudukan.

DEVELOPMENT OF POSITION FEEDBACK SENSOR BASED ON VISION USING NEURAL NETWORK

ABSTRACT

Measurement of linear position is fundamental in many industrial processes especially as positioning feedback. Optical encoder is one of the frequently used position sensor. Unfortunately, there are some disadvantages in the usage of this type of encoder. This proposed work presents an approach to build a position sensor using image classification. The image was classified by using supervised backpropagation neural network. Input image that was fed to the classifier is a grayscale image of the surrounding environment (up view) which was captured to represent the position. The features in the image are not clear and thus feature extraction is difficult to be performed to extract statistical data. The image was rescaled and fed into the network as one vector. Series of images were captured at various positions with each series having different distance to each other. As the interval between images becomes closer, there is more overlapping between images. Four series of images interval have been studied to see the effect of different numbers of input element towards the network. The effect of different brightness has been studied to make the system robust enough to be used in variation of environment condition. The best structure of network to solve this kind of image classification problem has been studied. To show the system performance, recognition rates was calculated. The result shows that, the network successfully classified the position image up to 80% accuracy. The trained network was then tested in one axis movement of gantry robot as position feedback sensor using vision. The gantry robot moves to the desired position by taking the difference between the actual position and the desired one. This error position has been corrected by using a gain control. From the experimental works that had been done, it shows that proportional gain of 10 is successfully achieved the objective.

The result obtained shows that the vision feedback system is capable for use in the application of one axis translational movement. It can be concluded that image classification of position image can be applied as a position feedback sensor.

CHAPTER 1
INTRODUCTION

1.1 Background of Research

Positioning plays an important task in manufacturing (Tseng and Liu, 2003) whether it is used in machine tools, coordinate measuring machine, pick and place machine in circuit board assembly, semiconductor manufacturing equipment and others (Alejandre and Artes, 2007). The positioning sensor is also used in movable elements of robot, peripheral device and measuring apparatus (Dumbravescu and Schiaua, 2000). Position, means an object's coordinates (linear or angular) with respect to a selected reference (Fraden, 2004). There are many sensors designed to measure position. It ranges from many methods such as mechanical, capacitance, resistance, magnetic and optical. The output produced from these methods can be analogue or digital.

Optical encoders are precision measuring device frequently used in high accuracy machine (Alejandre and Artes, 2006). An optical based encoder produces a digital pulse and this characteristic makes it as one of favoured position feedback sensor. Digital device is preferred to analogue as there is no needs to have an analogue to digital converter which can add cost to the system (Matthes *et al.*, 1994). Furthermore, digital position information allows speed calculation by numerical calculation and this increases its compatibility with most control system (Sente and Buyse, 1995). Unfortunately, as the encoder has higher resolution, the price also increases (Park *et al.*, 2001; Pereira *et al.*, 2007).

Encoder can be in three forms – tachometer, incremental and absolute (Doebelin, 1996). Tachometer has only one track and goes in the forward direction only. The reverse direction can give the same position as the forward direction and thus it was only applied in speed feedback. Incremental encoder has two tracks in which the lag between the tracks is ¼ cycle. Sometimes it can be three tracks where the additional one is a reference index. Meanwhile, the absolute encoder utilizes multiple tracks which are read parallel to produce a binary or gray code reading of angular position of the shaft.

There are two types of optical encoder namely rotary and linear scale encoder. Rotary encoder attaches to the motor spindle shaft and gives indirect measurement of the shaft rotation. Linear encoder gives direct measurement as it is attached to the movement track. Optical encoder basically consists of a LED, a scale or disk and a photo-sensor. LED will light up a moving scale or the rotating disk and the light changed will be sensed by a photo-sensor. Photo-sensor will produce the digital form of signal which is the square wave. This wave will be converted to pulse wave in terms of voltage by using electric circuit. The coded signal will be read by the microprocessor as a position measurement.

Several researchers have improved characteristic of optical encoder such as by changing the conventional grating scale purposely to improve the resolution and accuracy of the encoder. El-roy and Friedland (1995) proposed a mechanism of wide-carriage-dot-matrix in linear encoder, Fu and Shyu (2001) replaced the grating by using a virtual gray level pattern, Kang and Park (2004) proposed an encoder with chiaroscuro plate and Lee and Zhoa (2004) presented a non-contact microscopic-surface based optical sensor.

In the last 15 years, an increasing use of linear encoders has occurred (Alejandre and Artes, 2006), as a more accurate alternative to rotary encoder. Figure 1.1 shows the example of an optical linear encoder. Most linear optical encoder in the market used an incremental measurement. Their incremental nature require re-homing when the power is turned off and then on. Typically linear encoders are constructed with a stationary scale while the read head move along the scale. As the read head move, power and signal cable which is mounted to the read head cause a problem in cable management. Since the optical scale is made of glass, it is somewhat fragile and susceptible to shock. Another possibilities that limit the use of the linear encoder is the reader can be easily blocked by contaminates which results in mis-positioning (Matthes *et al.*, 1994). Alejandre and Artes (2007) added that linear encoder particularly suffers from misalignment, lack of support rigidity, vibration and temperature change. Affected by the disadvantages of linear optical encoder, a method of reading linear positioning measurement by using a vision system is proposed in this study.

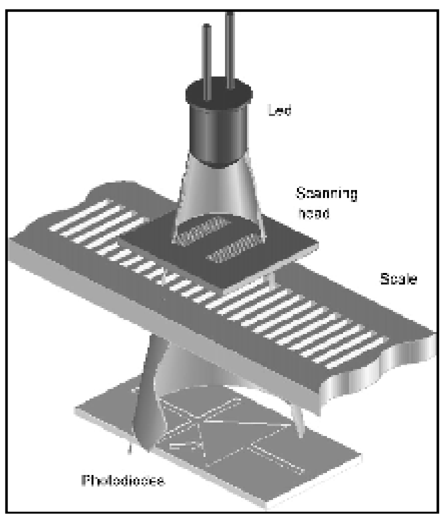

Figure 1.1: Optical linear encoder (Alejandre & Artes, 2006)

The advantage of vision system is its ability to automatically interpret image in high degree of reliability, repeatability and speed (Jain and Dorai, 1996). One of the successful application of vision system is as feedback in control system. Chroust *et al.* (2000) stated that the goal of vision based control system is to track the motion of an object. One of the difficulties in using vision as feedback sensor compared to conventional motion control is the time latency in acquiring the image and processing the image to get the target information. Several researchers reported vision as positioning feedback sensor which give accuracy in assembly process of foam barrier (Gao *et al*., 2002), positioning of robot manipulator (Shen *et al*., 2001; Larsen and Ferrier, 2004), motion control of pneumatic actuator (Hirai *et al*., 2002), visual servoing of path trajectory (Park and Lee, 2003), inspecting the position control accuracy for motor control system (Zhenzhong *et al*., 2001), visually guided a microlathe tool tip (Ojima *et al*., 2007) and movement of autonomous vehicle navigation (Wu and Tsai, 2009).

1.2 Problem Statement

Noble (1995) reported that most industrial vision research focused on the products (inspecting) rather than improving manufacturing process (process understanding and control). A decade later, Golnabi and Asadpour (2007) stated that vision was applied in many application of manufacturing including automated visual inspection, process control parts identification and mostly play an important role in robotic guidance and control. Author of both papers agreed that the restrictions in vision application are the choices of image processing algorithm and the implementation of the system.

Technically, machine vision often operates in constrained and controlled environments in which there is control over illumination. This research will manipulate this constraint by capturing the environment view in uncontrolled condition as a training input to the neural network. Some of the problem faced by vision system are non-uniform illumination, poor contrast, shadows and highlights (uneven lighting), occlusions, sensor noise and background clutter (Jain and Dorai, 1996). This problem was reflected onto the environment view captured by the camera. Petrovic *et al.* (2002) has proposed a robust vision system in angle measurement with uncontrolled lightning condition and it was shown that the system was effected by noise in camera that produced a poor contrast, non adequate lighting and 'parasite' edge that do not belong to the tracked object in the image captured.

In an application of machine vision, the camera will focus onto the object. If the object is moving, the object may go out from camera's field of view and cannot be measured (Daniilidis *et al.,* 1998). If the field of view increased, the accuracy of the image will be decreased. This can be solved by changing the zooming of the camera but by using a simple camera with a fix lens, this is not achievable and it will totally depend on the algorithm of vision system.

In this thesis, the camera will focus only on the environment and the image of the environment will be used to determine position. Generally, in machine vision application, a camera can be placed facing many views and the image can be captured from several environment views which are down, side and up view. Down view gives a view of workpiece meanwhile the up view gives a view of ceiling. The side view gives the view of the laboratory environment. The down view has variation in colours of workpiece and the side view has a variation in objects inside the image. Unfortunately, the down and side view are easily changed and distracted by the movement of workpiece and human entering the laboratory. This problem will affect the network trained earlier. The up view will not be affected by the problem but it doesn’t have much variation and not much difference can be seen from each image. This lead to difficulty in extracting statistical data from the image since the features is not clear. It needs a method to recognize it intelligently. A work by Wu and Tsai (2009) for autonomous vehicle navigation used omnidirectional camera facing upward to reduce the possibly effect from nearby objects and humans. In their work, a circular landmark placed on the ceiling is used as a focusing object.

Colour image contains more detailed information (Chiou *et al*., 2008) and data. As the colour image consists of more than one colour, it requires a big storage space (McColl and Martin, 1989) and thus need more time in image processing. A grayscale image is an image with shades of gray. In this work, the environment view will be fed into the neural network in grayscale mode to avoid the big amount of memory and bigger image size. The grayscale image is a suitable form of image to be used as it is enough to give information for the neural network to differentiate the position image. From the review done so far, there is no research on vision as a position feedback sensor which used an environment view as a measurement scale reported before.

1.3 Research Objective

The proposed research intention is to apply vision in positioning motion control which can adapt with the changing environment. The main objective is to apply vision system as a position feedback sensor. The sub objectives are;

a) To develop image processing algorithm without feature extraction by using neural network to avoid a complex image processing algorithm.
b) To adopt a vision system as direct position measurement sensor.
c) To develop a control algorithm using vision as a positioning feedback sensor.
d) To evaluate the performance of the position control system.

1.4 Research Scope

There are several limitations for the proposed research. This closed loop positioning control system will focus only on one axis linear movement. The system will be applied to Gantry Robot placed at FMS (Flexible Manufacturing System) Laboratory, School of Mechanical Engineering. The gantry robot moves by using a stepper motor and receive the position input in pulses.

The vision system will use a camera with CMOS sensor as image acquisition device. The size of captured image is set at 480 x 640 pixels. The captured image will be used as an input to the classifier which is backpropagation neural network (BPNN), one of the supervised neural networks.

In a control movement of mechanical system, factors such as vibration, friction and inertia will affect the motor performances. In this thesis, the mechanical inaccuracies inside the system will not be taken into consideration. The controller will not be designed to eliminate all the factors above.

1.5 Research Approach

The research was started by doing a study on relevant topics. The development in position feedback sensor was studied. Recent research in vision as positioning feedback sensor was taken care seriously. The application of image classification by neural network was focused.

From the objective set above, there were several tasks to achieve. First, vision algorithm was built by using BPNN to recognize the position along a linear axis. The input is image without feature extraction and the output is position of the image. Several images with different condition were gathered before used to train the network. This training was done offline before being applied to the system.

After the network had successfully classified the position based on the image, it was applied to the gantry robot. Here, the network acts as a database which learned to recognize all positions according to the respective images.

Further, vision sensor was implemented as a position feedback. The feedback sensor will be based on the comparison between the image captured and the network trained. Thus, the algorithm for closed loop position control was built to complete the system. At the end, several testing was run and the data were used to evaluate the performance of the positioning system.

1.6 Thesis Outline

This thesis is arranged in accordance to the objectives and approach as mentioned above. Chapter 2 provides a literature review of related subject. Here, the enhancement on previous research and past knowledge will be explained. The topic discussed include vision as positioning feedback sensor, neural network on image classification and other development of sensor feedback which getting other researcher interest.

Chapter 3 gives description of the proposed research methodology. The method on choosing a parameter for BPNN will be discussed in more detailed. The discussion also includes the algorithm of vision system and close loop control system applied to the gantry robot.

Chapter 4 discusses the result of the best structure of neural network and performance gained from the positioning control system. This includes a reasonable justification to support the results. The thesis closed with summary, conclusion and contribution of the overall research in Chapter 5. There are also recommendations for the improvement for further research.

CHAPTER 2
LITERATURE REVIEW

2.1 Introduction

This chapter presents an overview of the topic covered within the thesis title. Three major topics are focused here which are positioning sensor, vision as positioning feedback sensor and neural network in image classification. Concept of positioning sensor and its application are briefly discussed. After that, the discussion on vision as positioning feedback sensor is presented. These are followed by the review of application of neural network in image processing. The discussions are focused on the application of BPNN. These include the comparison between BPNN with other types of classifier and the improvement done by other researchers toward BPNN. Recent studies and researches on above areas are reviewed thoroughly.

2.2 Positioning Sensor

Precision position motions are commonly found in industrial applications in manufacturing (Tseng and Liu, 2003). The motion systems are used in machine tools, coordinate measuring machine, pick and place machine in circuit board assembly, semiconductor manufacturing equipment (Alejandre and Artes, 2007), movable elements of robot, peripheral device and measuring apparatus (Dumbravescu and Schiaua, 2000). It is also crucial on industrial application such as material transfer, packaging, assembly and electrical wiring (Gan and Cheung, 2003).

To study precisely the application of positioning feedback sensor, an example of a linear stage of *x-y* worktable motion system has been considered by Fischer and Tomizuka (1996). The worktable linear motion system is used commonly in CNC machine center and wire bonding machine. Commonly, linear positioning table is driven by servomotor through a lead screw. The worktable position is estimated by positioning sensor, eg. rotary encoder which is attached to the servomotor shaft. The system is controlled by control algorithm in microprocessor to implement advanced control algorithms and multitasking applications involving: data acquisition process, complex trajectory generation algorithms, supervision tasks execution, identification process, and compensation for the errors in real-time (Kastaneda and Okazaki, 1998). The errors produced from the movement of table such as friction, backlash and inertia resulted in a difference between the position of the motor and the corresponding position of the table. Figure 2.1 shows the principle scheme in closed loop positioning sensor.

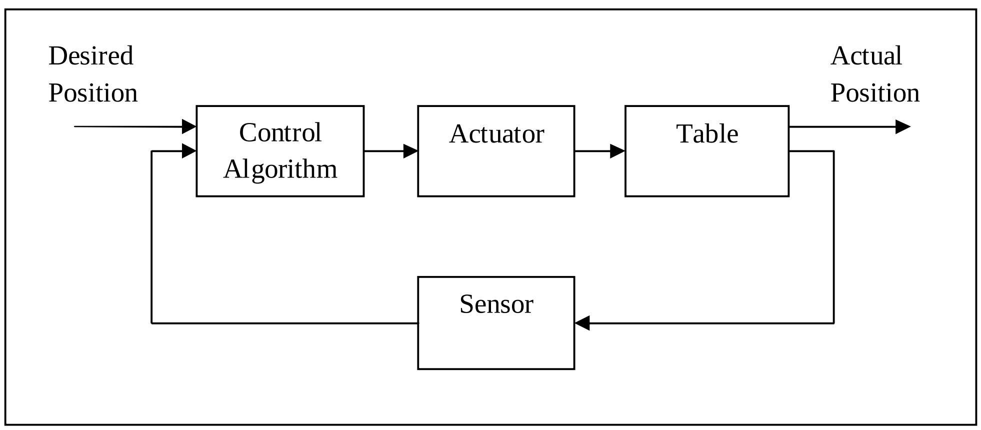

Figure 2.1: Principle scheme in closed loop in positioning worktable
(Fischer and Tomizuka, 1996)

Positioning sensor can be magnetic, resistive, capacitive, inductive and optical. Magnetic sensors rely on electromagnetic fields, and the magnetic properties of materials, in the operation of their basic sensing elements. Capacitive sensor is based on changes in capacitance in response to physical variations. The basic sensing element of the sensor consists of two simple electrodes with capacitance *C*. Resistive sensor or potentiometer as they commonly called are relatively simple devices in which a sliding contact or wiper moves over a resistive element. Inductive sensor or commonly known as linear variable differential transformer (LVDT) utilizes the change in inductance of a secondary coil relative to a primary coil to discern position. Optical encoder is the sensor that used the electric pulse produced by the movement of scale when the light from light emitted device (LED) projected through the scale and detected by the photo-detector. Another popular sensor is interferometer which uses energy in the form of light or sound wavelength. The transmitted wave interacts with the reflected wave and produces the resultant amplitude (Matthes *et al*., 1994; Webster, 2000).

2.2.1 Previous Study on Positioning Sensor

Previous study shows that there are improvements in conventional positioning feedback sensor by inventing a novel sensor or modifying the existing one. Nearly fourteen years back, El-Roy and Friedland (1995) proposed a new linear encoder with 1000 lines per inch grid to increase the accuracy of linear incremental encoder. The encoder reading is based on wide-dot-matrix-catriage of printer mechanism. It used a friction compensation controller to avoid hang off and stick slip problem. The author claimed that the sensor can achieve submicron accuracy. The disadvantage is that the output is analogue output, thus it needs to be equipped with hardware counter and preceding signal conditioning.

Fu and Shyu (2001) took the same objective with the above research by using a virtual gray level pattern produced by binary optic mask. The gray level pattern is claimed better than the doublet grating of conventional linear encoder in terms of high precision manufacture, small period and easy duplication. The sensor consists of a laser diode, lens, cylindrical lens, linear array detector, columnar lens array and detector array unit which makes it complex in assembly structure.

In the application of aerospace which required a simple and non powered sensor, Rice *et al.* (2001) from the Air Force Research Laboratory, USA presented a passive rotary and linear incremental encoder which uses a fiber optic cable. The advantages of this fiber optic sensor are immunity to electromagnetic interference, a long lead remote positioning, reduced size and increase tolerance in harsh environment. The sensor basically consists of a pulse source and a detector connected by a coupler. The coupler convert the signal to the encoder which has a fiber optic time delay network which divides the pulse in power and time among fibers to illuminate a reflective gray code mask. There are serial digital bit streams of the gray code value produced on the mask. The sensor has losses in light reflection of optical network and the single lead fiber structure is needed to reconfigure in order to solve the problem.

Kang and Park (2004) suggested a linear encoder with chiaroscuro plate to solve the on-off switching problem in switched reluctance motor drive. Chiaroscuro is a penetrable round plate with a linearly changed intensity as shown in Figure 2.2.. It was placed between a transmitter and a receiver of phototransistor. When the motor rotated, the plate rotated and the transmitter light penetrated through the plate. The change in light intensity was captured by the receiver. It has its own simple logic circuit thus it is not affected by sampling problem in microprocessor. This low cost sensor (due to simple structure) produced a high resolution in angle control. In order to determine the backward or forward direction, the sensor required two phototransistors.

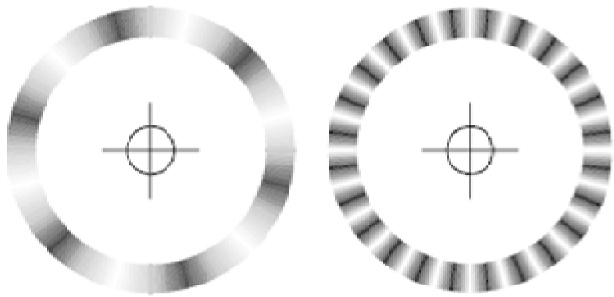

Figure 2.2: Chiaroscuro plate (Kang and Park, 2004)

Yen *et al*. (2004) took the advantage of low cost and high accuracy of compact disk driver by proposing a new optical pick up head table. The laser/photodiode act as a sensor and the carrier platform was mounted on a pickup head to construct an *x*-*y* table. The table applied adaptive controller to counter the influence of coupling effect and platform mass variation. It is proposed to carry an object such as biological specimens or aligned optical fiber. There is a limitation in *x* direction movement because the working movement confined within a small range. The sensor has a problem in oscillation and overshoot was obvious, showing a poor quality of feedback signal.

The noncontact optical sensor for 3-DOF (degree of freedom) measurement for planar and spherical motion is proposed by Lee and Zhou (2004) based on the application of microscopic-surface-based optical sensor. The sensor consists of a LED which illuminates the surface and then the lens collects the reflected light before the image is captured by photo-detector. A unit of sensor is capable to produce the translational motion value in *x* and *y* axis. It needs a dual sensor to get the rotation movement measurement.

In linear stage, for multiple DOF measurement, the system needs a bulky structure in combination of single position sensor. The solving method is by using a planar stage with interferometer sensor. The problem in interferometer in planar stage is sensitivity to environment condition. To overcome this problem, Gao *et al.* (2005) proposed a high precision dual mode surface encoder without bulky part that can measure multiple DOF. The encoder is composed of a slope-sensor unit and an angle grid. The slope-sensor unit consists of a slope sensor and scanning beam slope sensor. It can provide two measurements which are high speed or multiple DOF. The sensor has a complexity in its fabrication due to small and compact angle unit design.

Recently, Treshanchez et al. (2009) used the image acquisition capabilities of an optical mouse sensor to build an absolute rotary encoder. The coded binary image of position printed on a white paper is captured by CMOS camera and produced two dimensional image arrays information. The numerical value obtained is processed to obtain the absolute angular position with minimum resolution. One of the disadvantages of their proposed work is any motion during the initial image reading can caused the sensor to restart again.

2.3 Vision Feedback in Control System

Golnabi and Asadpour (2007) divided the application of machine vision in industry into four categories – automated visual inspection, process control, parts identification and robotic guidance and control. Malamas *et al.* (2003) classified the industrial vision all in one category - quality inspection but in different potential features which are dimensional, surface, structural and operational. Despite this difference, the primary goal in most vision system is to improve productivity and quality in manufacturing process (Graves, 1998). The advantages of machine vision are noncontact measurement sensor, real time data, high reliability and flexibility, a low cost and high efficiency in performance. However, it depends much on decision making problem and imaging environment (Jain and Dorai, 1996). A typical industrial vision system shown by Malamas *et al.* (2003) stated several standard requirements, an image processing software, a computer for processing the acquired image, a camera which in most cases is in fixed position and appropriately illuminated and properly arranged scene.

Visions act as a feedback in control system. In theory, the basic idea in continuous product or process improvement using machine vision is to monitor the critical parameter of the development. It will determine the principal product or process variations. This knowledge is applied further to develop corrective or control process. It will then be incorporated into the manufacturing cycle. The feedback system continuously improved overall product or process quality over time (Nobel, 1995). This statement reflects the usage of vision system in a wide range of applications. One of the applications of vision as sensor feedback is in positioning control system.

2.3.1 Vision Feedback in Positioning System

Zhenzong *et al.* (2001) studied the possibility of vision as an alternative to conventional position accuracy device such as optical encoder, hall sensor and resolver. The system used a stripe of structured light to map two dimensional (2D) image coordinates to three dimensional (3D) object coordinates in 3D movable platform. Radial basis function (RBF) neural network was used as a mapping model of this inspection system. The testing accuracy achieved is 0.081 mm. A much higher inspection accuracy system can be obtained by increasing the number of training sample into RBF network and using a smaller movement step.

One interesting study by Ojima *et al*. (2007) used vision as positioning feedback sensor in the machining process of microlathe. The movement of tool was controlled by visual feedback. The position of tool and workpiece was captured by CCD camera and compared with preregistered template. The image captured defines the 2D positional information (pixel coordinate) for the objects in view to refer to. Corresponding to the relative position of tool and workpiece, the tool path and speed are calculated and converted into appropriate pulse train. From the experimental linear and circular path controlled, the error is within ±5 pixels (30 μm). This small scale experimental setup is not affected much by environment changing thus it did not require any complex image processing method. If this method is to be applied in conventional worktable, much consideration has to be taken in image processing technique.

In vision control feedback application, the positioning can be a position based or image based (Wells *et al.,* 1996; Park and Lee, 2003). In position based method, the visual data is used to compute the relative position of object to the camera. Meanwhile in image based, the information from observed object is used directly as feedback in the system.

The example of position base method is shown by Wells *et al.* (1996). They positioned the end effectors with wrist mounted camera. The study used feedforward neural network to learn the nonlinear relationship between the object's features variations observed from the camera with the change in robot pose. The camera captured *x* and *y* coordinate of four points that represent the object shape features which are taken as inputs to the network. The network presents the output as six Cartesian coordinate (position and orientation) of the camera seen from the reference image (reference position). The experiment has to be done on controlled contrast and illumination environment by enclosing the work area with a black curtain.

In application of vision system, the camera can be mounted either at end effector or fixed at some location in the environment. Larsen and Ferrier (2004) used different mounting method compared to Wells *et al.* (1996) by placing the camera facing downward to the three fingers of flexible robot manipulator. The camera captured the deflection movement of the fingers by detecting the centroid location and major axes of the fingers using moment based technique. The system is controlled using neural network scheme. Attention should be taken on the effect of noise on image and error towards the control scheme. This system was proposed for micro level application.

Hirai *et al.* (2002) used similar camera mounting in pneumatic group actuator (PGA) motion control. PGA is an actuator composed of multiple single motion tubes connected with moving plate. Two circular markers were attached on the moving plate. Two dimensional position of each marker was computed using labelling and image gravity computation from an image captured by a CCD camera and it defined the position and orientation of the plate. The movement was controlled until it reaches the final desired position of the end plate by several iteration of movement. Smooth movement can be achieved by increasing the stiffness of PGA.

Vision feedback can be used together with motion control in application of assembly process. Bright and Jennings (1998) used machine vision on controlling a two axis motion of dispensing polyurethane machine. The conveyor transported the mould under a CCD camera. Two dimensional image coordinate extracted from mould profile define the x and y axis paths and the pulse signal was sent to the motor to control the movement of the mould along the conveyor. The mould profile additionally gave signal to mixing machine in selection of shot size and dispensing rate.

Theoretically in the application of pattern recognition, the objects captured by camera, compared with reference template and the information on respective features are extracted. In the study of position tracking of foam barrier, Gao *et al.* (2002) used pattern recognition to detect the position and orientation deviation of the foam. The deviation is obtained by applying direction flow change (DFC) concept. From the algorithm developed, the position error in two axes, x and y and the rotating angle can be computed. The tracking of foam barrier cannot be done by existing sensor as the material deformation can affect the position indication.

Most of the studies in vision have been researched and tested in simple laboratory environment (Wells *et al.*, 1996). One of the examples is by using a ball and beam experiment. Petrovic *et al.* (2002) used a ball and beam experiment to show the robustness of vision system in non ideal lighting condition and with background effect and using a low cost USB web camera. The image processing method developed consists of a low and high level processing. In low level, the image was converted to edge vector and further, the vector which represents the beam angle was computed in high level processing. Even the control of the ball is possible; the system still has an effect from the noise in camera which produce a worse contrast, non adequate lightning and ‘parasite’ edge that do not belong to the tracked object in the image captured.

The simple experiment was also applied by Park and Lee (2003) in positioning control of ball trajectory on the plate. The objective of the experiment is to change the angle of plate so that the ball can track its trajectory on the plate. The camera captured the image of the ball on the plate and computed the center position by using a fast algorithm combined with Euler position estimator. It used a difference plates and ball materials to show the robustness of the system. This system can be applied in handling a balancing problem of robotic arm. The effect of inertia in hollow ball and damping in rubber ball decreased the accuracy in the system.

2.4 BPNN in Image Processing

BPNN is the most popular in all of neural network applications (Li and Park, 2008). It is the most widely used neural network (Delogu *et al.,* 2008; Nassar and Ammar, 2007; Feitosa *et al*., 2000). BPNN was proposed by Rumerlhart and Parker independently in 1980s (Mokhtarzade and Zoej, 2007). Neural network are being successfully applied in many application such as business, medicine, geology and physics to solve problems of prediction, classification and control. They are able to tackle a variety of image classification tasks in industrial vision environment (Malamas *et al*., 2003). The discussions here focus on BPNN in image processing application in a recent research.

2.4.1 Structure of BPNN

The structure of BPNN consists of an input layer which is connected to output layer by hidden layers. BPNN is a supervised neural network in which the input of the network is trained to reach the target output.

In feed forward action of BPNN, the input data will be trained to reach the target output. Along the way, the input will be weighted in hidden layer and differentiate by activation function to be paired to its target output. The hidden layer which can have one or more layers is the connecting layer between input and output layer. Each neuron has a connection weight to the following layer's neurons. It will be sum together with the bias which has a value of 1. The input from the neuron is given by;

$$x_j = \sum y_i w_{ij} \tag{3.3}$$

where y_i is the output from the neuron and w_{ij} is the weight connection between the neuron (Rumelhart *et al.*, 1986). The choice of numbers of hidden layers and neuron inside each layer are the constraint in network structure.

The activation function determined the output of the neurons. The standard nonlinear activation function for BPNN is log-sigmoid transfer function (Rumelhart *et al.*, 1986). The log-sigmoid will limit the output between 0 and 1 and squashing the output value, y_j;

$$y_j = \frac{1}{1+e^{-x_j}} \tag{3.4}$$

The log-sigmoid transfer function is commonly used in multilayer networks that are trained using the backpropagation algorithm, in part because this function is differentiable (Hagan and Demuth, 1999).

After the feedforward action, the output of the network is compared to the target output. The squared error of the difference between trained output and target output are propagated back to the network. The sum squared error (SSE), E is defined as

$$E = \frac{1}{2}\sum_c \sum_j \left(y_{jc} - d_{jc}\right)^2 \tag{3.5}$$

where y_{jc} is the actual output and d_{jc} is desired output (Rumelhart et al., 1986).

Based on the error received, connection weights are then updated. This process repeats, shown in a number of iteration until the network converges to a state that the error are acceptably low. Another stopping method is by setting the number of iteration process. The network will stop training as it reaches this error value or iteration value. Stopping method is important as it can eliminate the possibility of overtraining condition (Mokhtarzade and Zoej, 2007).

The weight change, Δw in input and hidden layers is guided by gradient descent algorithm following the delta rule by Windrow Hoff (Rumelhart *et al.*, 1986);

$$\Delta w \propto -\frac{\partial E_s}{\partial w} \tag{3.6}$$

In gradient descent algorithm, the weight was adjusted in the negative direction of gradient. It was found that the performance function is decreasing rapidly but it not produced fastest convergence.

2.4.2 Input and Output of BPNN

Input to the network is the feature that was extracted from the image. Meanwhile the target output defined a group, class or specific target feature.

The binary input data was applied by several researcher in text categorization (Yu *et al.*, 2008), document classification (Li and Park, 2008), handwritten numeral recognition (Ping and Lihui, 2002), character recognition (Kamruzzaman, 2001) and printed mathematical symbol (Shamsudin *et al.*, 1999). The target output of the network is the defined character, number or symbol. To achieve higher recognition rate or classification rate, it needs an image processing algorithm which can extract the character or number features that resemble the target output. Binary input simplified the recognition process but, in some application it is not suitable to use as it generally decreases the recognition performance of the neural network (Egmont-Peterson *et al.*, 2002).

Barelli *et al.* (2008) used a grayscale image in detection of defect in surface of mechanical seals. They proposed that the variation of pixel brightness value in image can determine the surface characteristics of the seal. If the surface is good, the pixel value will be homogenous and if there are faults present, the pixel brightness will be enhanced. In other application of grayscale image, the researcher extracted the textural features of the image such as homogeity, energy, contrast, variance of color image (Morquin *et al.,* 2004) , the distance value between element and neighbor pixel and three grayscale level of color features, R (red), G (green) and B (blue) components (Mokhtarzade and Zoej, 2007). Morquin *et al.*, (2004) used respective features in separation process of soil and clod. Mokhtarzade and Zoej, (2007) applied the features in detecting road from background image taken from satellite. Huang (2007) has listed a mathematical formulation for several texture features. The combination of the features is one of the consideration factors in designing the network structure.

Chiou *et al.* (2008) believed that color image gives more information compared to grayscale image. They fed a color image in separating gold and non-gold plating in inspection of PCB. They agreed that color image can easily distinguish any anomalies in image. One problem occurred is the effect of illumination. As a solution, they normalized the color features of the image. In contrast, Shiau *et al.* (2000) used color image over the disadvantages of binary image. They applied the HSI (hue, saturation and intensity) image in classification of web, neps and thrash of textile. Web, neps and thrash are similar to each other in shape which causes difficulties when distinguishing them in binary image processing. They have to deal with the problem in brightness and contrast of the image. To avoid this disturbance, they used a scanner and applied a homogenous brightness and contrast.

The discussion above is about the input features fed into the network. The numbers of input basically depend on the features extracted from the image. Meanwhile, the number of output neurons depends on the number of the output feature. The value of output neuron can be represented in binary numbers. A node of output target was used by Morquin *et al.* (2004) in differentiating an onion from soil clod and Arbach *et al.* (2003) in mammographic mass classification. The output of the network will only give one value either 0 or 1. Zoorofi *et al.* (2001) gave two binary outputs to classify contamination in IC circuit inspection. More combination of binary numbers can be made as shown by Wong *et al.* (2008) to represent five classes of stitching defect and Han *et al.* (2003) to classify seven conditions in wetland image. Mokhtarzade and Zoej (2007) gave the output of the network in a form of 2D matrix image with a same size of input image.

The input and output neuron was connected by hidden layers. In hidden layers, the input was weighted to reach the target output. Nevertheless, the choice of number of hidden layer is crucial in network structure. More hidden layers can avoid the burden storage in each layer (Han *et al.,* 2003) and possibility to increase the accuracy (Mokhtarzade and Zoej, 2007) but it will take more computational time. In the other view, increasing the number of layers increases the complexity of a network topology and there is no guarantee exists that the model's performance will be improved (Shiau *et al.*, 2000). However, determining the exact neural network topology is somewhat subjective and depends on the problem to be solved (Zoorofi *et al.*, 2001).