

CHAPTER 2

LITERATURE REVIEW

2.3 General Description

2.3.1 The Camera Positioning

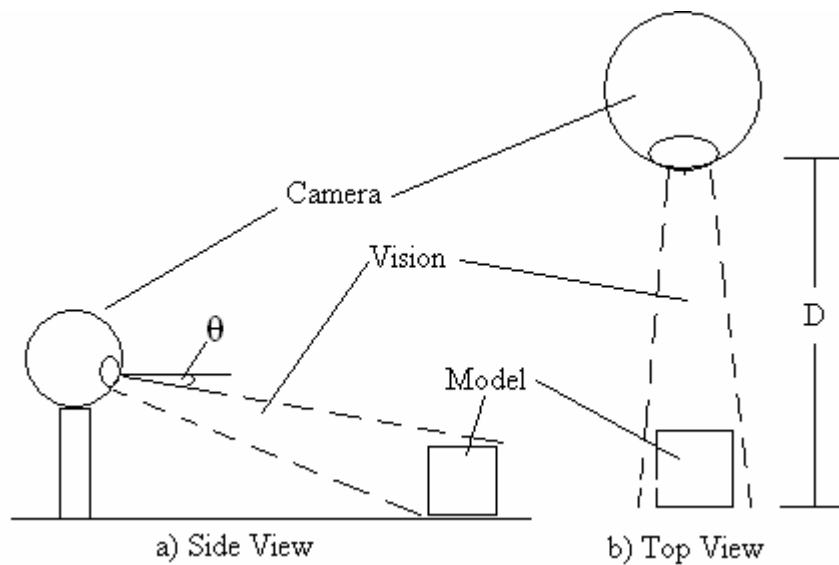


Figure 2.1 Sample of the procedure on how to capture the images

The systems generally work as shown in the Figure 2.1. There is digital camera that represent as an image captured. This system firstly tests on when the model shape placed in the view area. In this picture, the symbols represent:

- i. D = Distance between camera and model shape
- ii. θ = Angel view of the camera to the model shape

2.3.2 The Camera Specification

Developing a camera to suit the project is very much detail. However, this camera is not suitable to use outdoor as it is made from a web camera. It can get a good result if it is used indoor but only bright light or dark scene if use outdoor. The camera specification in Table 2.1 is use as a prototype in terms to obtain all the images. This camera is use as a prototype in terms to obtain all the images. Distance between camera and model is 9 cm and the angel view of the camera to the model shape is about 20° .

Table 2.1 The Specification of the Prototype Camera

Definition: 300k/350k/400k/450k pixels(640x480)	Frame Rate: 320x240 up to 30 frame/sec(CIF)
CMOS chip type: Color CMOS image sensor	640x480 up to 15 frame/sec(VGA)
Hi-resolving power: 640x480	Sensor size: 4.86x3.64mm ²
Video Format: 24-bit RGB	Dynamic range: <72dB
Interface: USB	Focus Range: 5cm-infinite
S/N Ratio: <48dB	Built-in image compression
Automatic white balance	Dynamic image E-MAIL
Automatic color compensated	Manual Focus
Shutter Key	

2.3.3 The Model of Shape

Model shapes are made from hard paper maroon in color. Color chosen is because of the contrast between the light background and the model itself. There are 4 shape model include in this experiment which are rectangle, cylinder, pyramid and cubic.

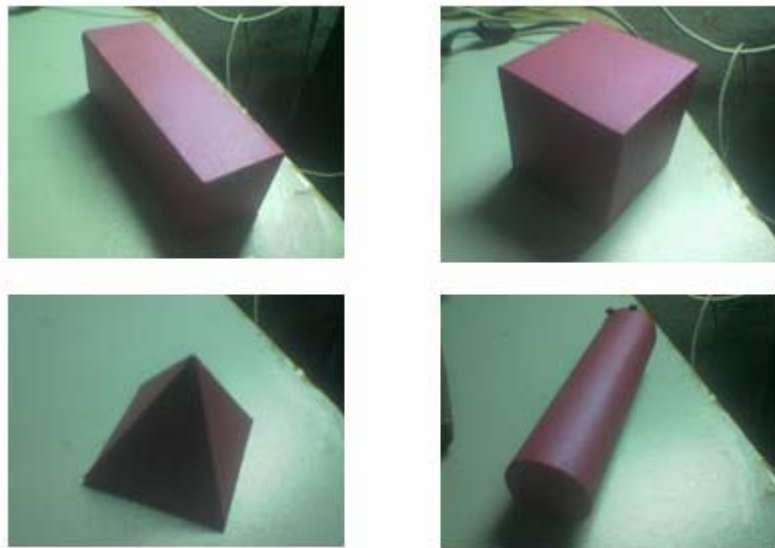


Figure 2.2 Sample of the models

2.4 Research of various method

There are several problem occurs during conducting of this project. The system can be read in various systems. In MATLAB, the system for image acquisition and image processing can have a lot of different step that can be consider finding the appropriate distance. As for this project title, 'Shape Recognition through Vision Systems' noticed that there are several ways to conduct a distance measurement. Listed next is the method that been research that can be applying using MATLAB software.

2.4.1 Autonomous Land Vehicle In A Neural Network(ALVINN)

This is a system that used to shown that neural techniques hold much promise for the field of autonomous road following. Using simple color image preprocessing to create a grayscale input image and a 3 layer neural network architecture consisting of 960 input units, 4 hidden units, and 100 output units, ALVINN can quickly learn, using back-propagation, the correct mapping from input image to output sight direction. This system is using the Hough Transform to get output of the entire database.

The Hough transform is a good system if the user needs to detect the distance accurately where it will be used to isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses and many more. In applications, Hough transform cannot be use as a simple analytic description of features. Due to the computational complexity of the generalized Hough algorithm, the systems restrict the main focus of this discussion to the classical Hough transform. The main advantage of the Hough transform technique is that it is tolerant of gaps in feature boundary descriptions and is relatively unaffected by image noise.

2.2.2: The Image processing using Canny edge Detector

This is the program that used to do the vision system program. This program, will be use the basic image processing from taking the images, read them then process it until the edge detection. This images is then be recognize by using the Singular Value Decomposition (SVD) to get a sum of digits that represent a pixel in the images that already converted to binary. This program is good to get an instance result start from taking the picture until get the SVD value. It will then read from the database and trained using neural

network and check if there is a possible similarity to the database. If the sum of value is almost the same as the database, it will then identify that the vision for that particular image have the value of a certain distance.

2.3 Problem Identification

The ALVINN system is not appropriate because the image will be process in a long time and this will be the problem in vehicle guidance system. The driver of the car will not wait for an hour just to get the result then he can move. Therefore, ALVINN system is not suitable for this method because of the time taken to execute the shape recognition program as in the titles.

The second is image preprocessing using canny edge detector need the SVD value as an important asset to get the similarity in the system. However, if these projects want to have a good system, the database will need a perfect picture where the foreground and the background of the picture are obvious. There should not be any unequal lighting, reflection, unwanted shadow and very less noise for the database. This is where that found out the problem because of a large amount of noise, an unequal lighting condition and shadow in the real time while taking the images in the indoor. In contrary, while taking outside however show worse result as the bright sun in day time or too dark during dusk to night time makes the image spoiled.

To solve the problem, by get a large amount of images and compare for the best one. This is to get select the best image through observation of all and process it to get a perfect output. Actually, this process can be accurately taken if the model images are obtained in a studio vision room where all the lighting must be equal so that even the reflection will be less that it can be ignored after doing the filtering.

2.4 MATLAB for Image Processing

It is a general rule that any evidence that is to be presented in court must be open to scrutiny and be testable. In the case where processed digital images are presented in court and have to establish two things, first is the authenticity of the original images and the digitized versions of these images and the second is the validity of any processing performed on the images for enhancement or other purposes.

Some algorithms are quite complex and the results may be sensitive to subtleties in the implementation. An image that has been extensively processed using proprietary software may well be challenged in court. Refuting that challenge will be difficult if one does not have access to the source code.

Another area of difficulty is recording all the steps used in the image enhancement process. Some applications such as Photoshop provide an extensive 'history' recording process. However, this sort of facility is designed primarily with the aim of being able to undo processes, or revert back to an earlier step in the sequence of processes. The history recording may not allow access to particular parameter values used in any individual step.

2.4.1 The Advantages of MATLAB

2.4.1.1 Recording of the processing used

MATLAB is a general purpose programming language. When it is used to process images one generally writes function files, or script files to perform the operations. These files form a formal record of the processing used and ensures that the final results can be tested and *replicated* by others should the need arise.

2.4.1.2 Access to Implementation Details

MATLAB provides many functions for image processing and other tasks. Most of these functions are written in the MATLAB language and are publicly readable as plain text files. Thus the implementation details of these functions are accessible and open to scrutiny. The defense can examine the processing used in complete detail, and any challenges raised can be responded to in an informed way by the prosecution. This makes MATLAB very different from applications.

It should be noted that some MATLAB functions cannot be viewed. These are generally lower level functions that are computationally expensive and are hence provided as 'builtin' functions running as native code. These functions are heavily used and tested and can be relied on with considerable confidence.

2.4.1.3 Numerical Accuracy

Another advantage of MATLAB is that it allows one to ensure maximal numerical precision in the final result.

In general, image files store data to 8 bit precision. This corresponds to a range of integer values from 0-255. A pixel in a color image may be represented by three 8 bit numbers, each representing the red, green and blue components as an integer value between 0 and 255. Typically this is ample precision for representing normal images.

However as soon as one reads this image data into memory and starts to process it it is very easy to generate values that lie outside the range 0-255. For example, to double the contrast of an image one multiplies the intensity values by 2. An image value of 200 will become 400 and numerical overflow will result. How this is dealt with will vary between image processing programs. Some may truncate the results to an integer in the range 0-255;

others may perform the mathematical operations in floating point arithmetic and then rescale the final results to an integer in the range 0-255.

It is here that numerical precision, and hence image fidelity, may be lost. Some image processing algorithms result in some pixel values with very large magnitudes (positive or negative). Typically these large values occur at points in the image where intensity discontinuities occur, the edges of the image are common sources of this problem. When this image with widely varying values is rescaled to integers in the range 0-255 much of this range may be used just to represent the few pixels with the large values. The bulk of the image data may then have to be represented within a small range of integer values, say from 0-50. Clearly this represents a considerable loss of image information. If another process is then applied to this image the problems can then accumulate. Trying to establish the extent of this problem, if any, is hard if one is using proprietary software.

Being a general programming language it is possible to have complete control of the precision with which one represents data in MATLAB. An image can be read into memory and the data cast into double precision floating point values. All image processing steps can then be performed in double precision floating point arithmetic, and at no intermediate stage does one need to rescale the results to integers in the range 0-255. Only at the final point when the image is to be displayed and/or written to file does it need to be rescaled. Here one can use histogram truncation to eliminate extreme pixel values so that the bulk of the image data is properly represented.

2.4.1.4 Advanced Algorithm

MATLAB is a scientific programming language and provides strong mathematical and numerical support for the implementation of advanced algorithms. It is for this reason that MATLAB is widely used by the image processing and computer vision community.

New algorithms are very likely to be implemented first in MATLAB; indeed they may only be available in MATLAB.

2.5 MATLAB vs. Others Software

MATLAB is always a good way to test algorithms and processing before getting into the actual implementation in C. It features a lot of pre-built functions which makes it much easier to implement very fast your algorithm and test it in a visual. It's a high level signal processing language. Once the algorithm works correctly under MATLAB user can port it easily to C without worries.

The others software also needs very strong in the mathematical solution algorithm and expert much more in that particular software used.

2.6 Neural Network

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. Train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements.

Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation is shown in fig(2). There, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are used, in this *supervised learning* (training method studied in more detail on following chapter), to train a network.

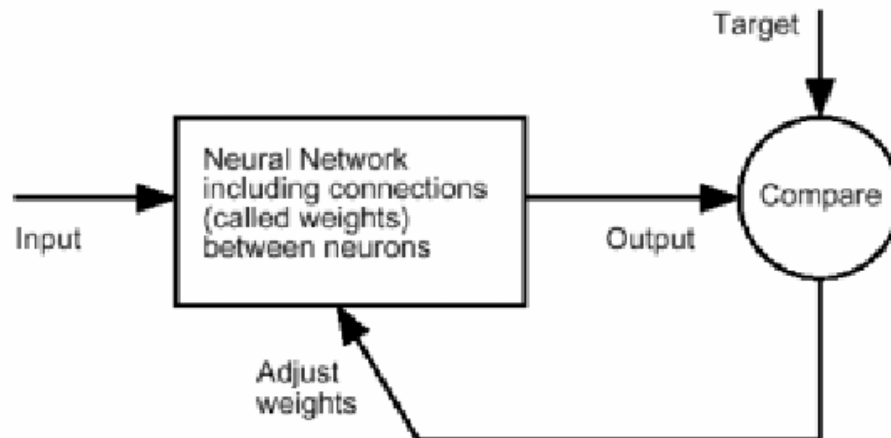


Figure 2.3 Neural Network block diagram

Neural networks have been trained to perform complex functions in various fields of application including pattern recognition, identification, classification, speech, vision and control systems.

Today neural networks can be trained to solve problems that are difficult for conventional computers or human beings. The supervised training methods are commonly used, but other networks can be obtained from *unsupervised training* techniques or from direct *design* methods. Unsupervised networks can be used, for instance, to identify groups of data. Certain kinds of linear networks and Hopfield networks are designed directly. In summary, there are a variety of kinds of design and learning techniques that enrich the choices that a user can make.

The field of neural networks has a history of some five decades but has found solid application only in the past fifteen years, and the field is still developing rapidly. Thus, it is distinctly different from the fields of control systems or optimization where the terminology, basic mathematics, and design procedures have been firmly established and applied for many years.