

Modeling 3-dimensional image from a 2-dimensional image using OpenGL

By

Shuhaida Mohamed Shuhidan

Dissertation submitted in partial fulfillment of

The requirements for the
Bachelor of Technology (Hons)
(Information Technology)

JUNE 2004

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

t
T
385
S 562
2004

1. Computer graphics
2. thesis

CERTIFICATION OF ORIGINALITY

This is to verify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.



(SHUHaida MOHAMED SHUHIDAN)

ABSTRACT

The main objective of the project is to study on the process of transferring two-dimensional (2D) image into a three-dimensional (3D) object by using OpenGL. Other than that, this project also covers on the research that involves jaggedness of a 3D image. From the research made, the outcome product will be 99% free from jaggedness. This kind of practice will maximize the performance of the product in terms of the efficiency. The project is put into practice in order to decrease the cost in producing a three-dimensional image from its corresponding two-dimensional image. At present, the commonly used tools to model a three-dimensional image are quite costly and expensive. Moreover, these tools such as digitized scanner are quite bulky and immobile. The project scope of study will only be within the Computer Graphic and Computer Vision area. A single two-dimensional image does not contain depth value and this will certainly make the image lack of realism. In contrast, the depth value needs to be added in the process of modeling a three-dimensional image from a two-dimensional image. This project includes reading the depth value of a two-dimensional image. These values will be manipulated and assigned as the depth value of the three-dimensional image. Then, each point identified and connected by a triangle. The combination of the triangles will produce a smooth surface of a three-dimensional image. This process will be repeated to another image in order to optimize the smoothness of the three-dimensional image. Lastly, factors that influenced the smoothness of the created three-dimensional model will be identified.

ACKNOWLEDGEMENT

Alhamdulillah, with my greatest gratitude to the Almighty Allah for his gracious blessings throughout this project was undertaken. My deepest gratitude to my parents for their love and supports. To my supervisor, Dr Abas Md Said, thank you so much for the guidance and support throughout completing the projects. There were times when I feel down as his word of wisdom had really encourage me to give more effort to the projects. Im feeling very lucky to have a supervisor that is so highly motivated like him.

I would also like to express my sincere appreciation to all my special friends Aida Fazlina, Goh Kim Nee, Haslina, Mastura, Ng Yin Ping and also Nirwan Bakri for all their contributions towards the project.

I really hope that all the efforts had been given and also with the help of such nice people, the project will benefit the society.

TABLE OF CONTENT

Content	Page Number
CERTIFICATION OF APPROVAL	i
CERTIFICATION OF ORIGINALITY	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
TABLE OF CONTENT	v
LIST OF FIGURES	vii
CHAPTER 1	
INTRODUCTION	1
1.1 Background of study	1
1.2 Problem Statement	2
1.3 Objectives and Scope of work	2
CHAPTER 2	
LITERATURE REVIEW AND THEORY	3
2.1 2-dimensional and 3-dimensional graphic	3
2.2 Bitmap(BMP)	4
2.3 Gray Value and Color Images	5
2.3 Depth Analysis	6
2.4 Human Facial	7
CHAPTER 3	
METHODOLOGY AND PROCEDURE	9
3.1 Plan	10
3.2 Analyze	10
3.3 Code	11
3.4 Testing and Evaluation	12

	3.5	Result	12
	3.6	Tool	13
CHAPTER 4		RESULT AND DISCUSSION	14
	4.1	Software For Project	14
	4.2	Convert RGB To Grayscale Image	15
	4.3	Convert 2-Dimensional Image To 3-Dimensional Image	16
	4.4	Image Acquisition	20
	4.5	Output	21
CHAPTER 5		CONCLUSION	22
	5.1	Recommendations for future work	22
REFERENCES		23

LIST OF FIGURES

- Figure 2.1 Pixel of bitmap image
- Figure 3.1 Flow of Project
- Figure 4.1 Ball placement
- Figure 4.2 Original Image of a ping-pong ball in control lighting
(64x64 pixel) enlarged by 300% scale
- Figure 4.3 3-dimensional ball from a controlled lighting
- Figure 4.4 Original Image of a ping-pong ball in non-control lighting
(64x64 pixel) enlarged by 300% scale
- Figure 4.5 3-dimensional ball from a non-controlled lighting

CHAPTER 1

INTRODUCTION

1.0 INTRODUCTION

Chapter 1 discussed overview of the project, which consists of background study, problem statement, objectives and scope of study. A brief explanation of 2-dimensional and 3-dimensional images will be reviewed. The study will focus more to model a 3D human face from a single 2D image.

1.1 Background Study

The 3-dimensional (3D) image is different than a 2-dimensional (2D) image as the 3D image has the depth of the image. On the other hand, the 2D image only has the width and the height of the image. An example of a 2D image is a piece of paper on the desk with a drawing on it and the example of 3D image is a can of soda on the desk. The height, width and depth can be measured easily as it can be seen by normal human eyes. The study of this project focuses on modeling human face.

The main objective of this study is to model a 3D image from a 2D image. A 2D image may be modeled to a 3D image by adding the depth value of the 2D image. The depth will be calculated as based on the pixel value of the 2D image. Calculating the depth is the key success to the project.

Some coding will be written in order to create a program to model the 3D image using Microsoft Visual C++ 6.0 OpenGL. Microsoft Visual C++ 6.0 is chosen because it is the most productive C++ tool for the highest-performance development for Windows and therefore it is the best tool to model a 3D image. On the other hand, OpenGL is used with computer hardware that is designed and optimized for the display and manipulation of 3D

graphics. Combination of Microsoft Visual C++ 6.0 and OpenGL will surely produce the best result.

1.2 Problem Statement

-Modeling a 3D image had been expensive all the way as it needs high technology equipments as 3D digitizer, scanner and etc.

-Bad lighting may give effect to the process of modeling a 3D image.

-Human face contains complex structures. The depth of one point to the next point may differ a lot. How to increase the smoothness of the surface of the human face?

1.3 Objectives and Scope of Study

The main objective of the project is to develop a program that able to create a 3D model from a 2D image using Microsoft Visual C++ 6.0 OpenGL. Other than that the study also include;

- To find a possible way to model 3D image without using expensive tools.
- To study the lighting that may effect good and smooth 3D image
- To increase the smoothness of 3D human face produced

The scope of the study will be within the Computer Graphic and Computer Vision area.

The duration of the project is from January 2004 to May 2004. In the early stages, few simple objects will be tested. Towards the end, human face will be focused to be modeled in order to fulfill the project's objective.

CHAPTER 2

LITERATURE REVIEW AND THEORY

2.0 LITERATURE REVIEW AND THEORY

Chapter 2 will focus on the description of the relevant findings of the project. Two-dimensional and three-dimensional graphics will be defined and compared. The image format used for this project is bitmap. A brief explanation about bitmaps will be discussed. Then, the depth analysis that determines the z-value will be explained. The study will then go in depth on the human face, which is going to be modeled in the experiment of the project.

2.1 Two-dimensional and three-dimensional graphics

The term three-dimensional means that an object being described or displayed has three dimensions of measurement. It is the width, height, and depth. But, as for the two-dimensional, the image may only have the width and height. As an example, a painting is inherently a two-dimensional object because it is nothing more than canvas with painting applied. Basically, three-dimensional computer graphics are actually two-dimensional images on a computer screen plus the illusion of depth that is called the third dimension.

A three-dimensional image has many advantages over the two-dimensional image. According to A.D. Marshall et al (1993), limited information can be deduced about the physical shape and size of an object in a scene from a two-dimensional image. However, with the three-dimensional image, the geometry of the scene is represented in terms of three-dimensional coordinates based on the axes x, y, and z. Thus, the measurements of the size of an object can be straightforwardly computed from its three-dimensional coordinates. It is easier to interpret a three-dimensional image as it contains adequate information of the height, width, and depth of an image.

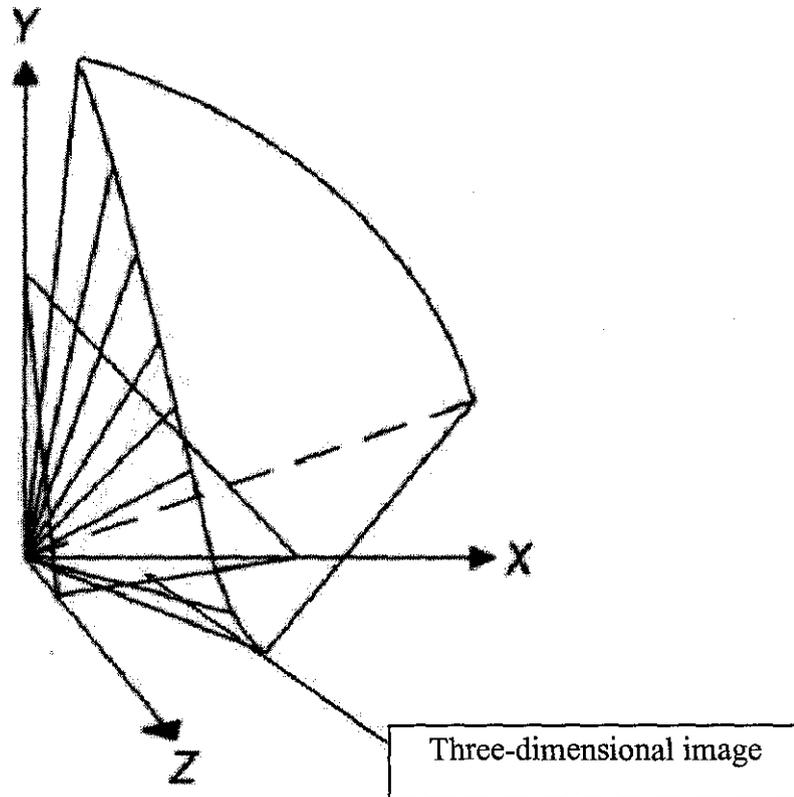


Figure 2.1

Therefore it is obvious that three-dimensional image is better than the two-dimensional as it increase the realism of the image. The depth of an image will clearly be shown in the three-dimensional as compared to a two-dimensional image that does not have the depth value the image.

A single two-dimensional image can be easily converted to a three-dimensional image by adding the depth value of the image. Some algorithm may be needed in order to produce the three-dimensional image from a two-dimensional image.

2.2 Bitmap (BMP)

A bitmap is an array of pixels values that describes an image, in human terms, a complete picture. A bitmap can be of various bit depth and resolution. It is a representation, consisting of rows and columns of dots, of a graphics image in computer

memory. The Windows Bitmap file format is the standard file format used by Microsoft Windows. Bitmap files can contain either 2 (black and white), 16 256 or 16.7 million colors. Most Windows Bitmap files are not compressed. It is possible to save 16 and 256 color images in compressed format but some applications (notably Windows Paintbrush) are not able to read the compressed files.

The default of file extension for Windows Bitmap files is “.BMP”. Occasionally, you may see bitmap files with the extension of “.DIB”. The easiest 256-color bitmap file format is Windows' BMP, where eight bits represents one pixel. Windows' BMP files are device-independent bitmap files most frequently used in Windows system. The bitmap format is based on the RGB color model. BMP is a raw file format. It does not compressed the original image.

According to Ralf Steinmetz et al (2002)

The bitmap format defines a header and a data region. The header region (BITMAPINFO) contains information about size, color depth, color table and compression method. The data region contains the value of each pixel in a line. Lines are flush extended to a value divisible by 32 and padded with zero values.

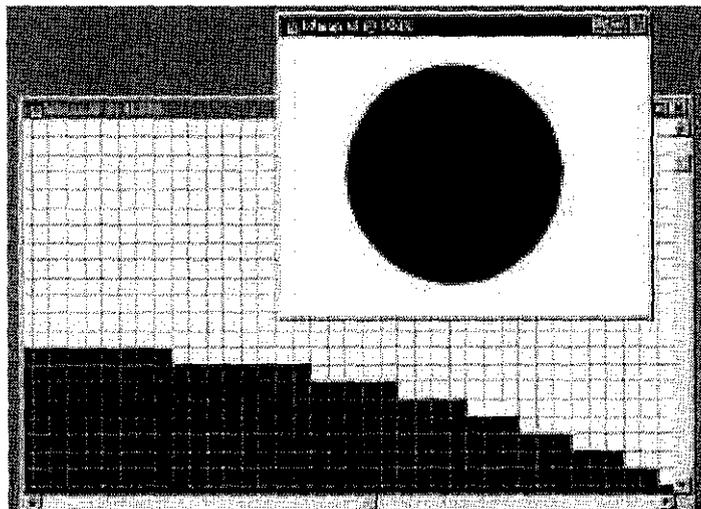


Figure 2.1

Each pixel of a bitmap contains a value. Referring to Figure 2.0, when zooming in on a simple bitmap, it will show its essential nature as a grid of pixel values. Each pixel's value may vary. As for this project, the bitmap file's image will then be converted to a grayscale value image and all the pixel's value will be acquired.

The value of each dot (whether it is filled in or not) is stored in one or more bits of data. For simple monochrome images, one bit is sufficient to represent each dot, but for colors and shades of gray, each dot requires more than one bit of data. The more bits used to represent a dot, the more colors and shades of gray that can be represented.

Bit-mapped graphics are often referred to as raster graphics. The other method for representing images is known as vector graphics or object-oriented graphics. With vector graphics, images are represented as mathematical formulas that define all the shapes in the image. Vector graphics are more flexible than bit-mapped graphics because they look the same even when you scale them to differ the sizes. In contrast, bit-mapped graphics become ragged when you shrunk or enlarge them.

2.3 Gray Value and Color Images

It is stated that the default value for the gray value is known as $G_{max} = 255$. So the gray scale value p , may vary from $0 \leq p \leq 255$. The value 0 will be the darkest black and the 255 value is a white color. Whatever value fall between the values will depend on the gray scale value, the more the value, the lighter of the color will be.

Reinhard Klette et al (1998) says, an image is defined for image points, which area assumed to be in real plane or which can be elements of a discrete set of points. The image E always assume a uniquely determined image value $E(p) = E(x,y)$ in each image point $p = (x,y)$. A digital image is distinguished by a discrete image points and value. The coordinates x and y of the image points (x,y) are assumed to be integers. Their values intervals are assumed to be $1 \leq x \leq M$ and $1 \leq y \leq N$ whereby the M and N values mark the image resolution. The value $A = M \times N$ identify the image size (p.12)

The main purpose to convert the image from RGB color image (bitmap file image) to a grayscale image is to acquire the depth value whereby;

Depth value of an image = grayscale value of each pixel

This will be discussed more on the next sub-topic

2.4 Depth Analysis

A two-dimensional image is lack of its depth value compared to a three-dimensional image. As for this project it is important to do the depth analysis to model a three-dimensional object from a two-dimensional image. The z value of a three-dimensional object is the depth of an object. As described earlier, a two-dimensional image doesn't have the depth information of the image. The value of the light intensity of a particular two-dimensional image from each pixel will determine the depth value of the real object.

According to A.D. Marshall et al (1993), a depth map is the simplest and most convenient way of representing and storing the depth measurement taken from a scene. A depth map is two-dimensional arrays where the x and y distance information correspond to the rows and columns of the array as in an ordinary image (two-dimensional image), and the corresponding depth readings (z values) are stored the array's elements (pixels). Essentially, the depth map is similar to grayscale image except that the z information replaces the intensity information. That is the reason the two-dimensional image of this project needed to be converted to grayscale image.

z-axis information = intensity information

The depth map of a bitmap may look like the figure 2.2

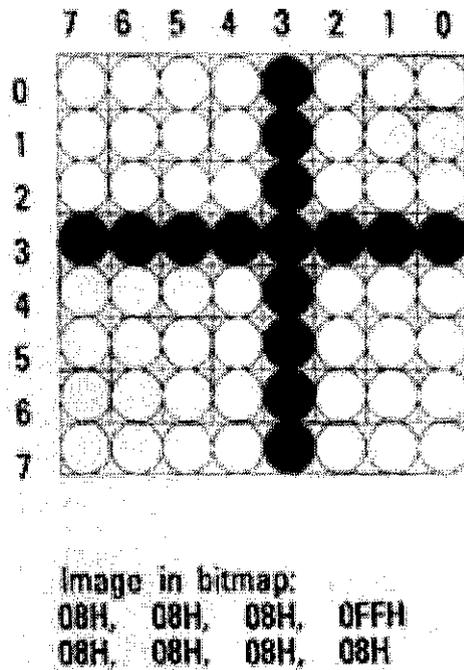


Figure 2.2

As for human facial, the depth value from a pixel to next one may differ. This is because human facial has a complex surface. The depth value may be calculated using some algorithm. It is important to get accurate depth value as to ensure that the three-dimensional model produced is similar to the actual object when display. In other words, is to maximize the smoothness of the two-dimensional image.

Reinhard Klette et al (1998) says that objects may or may not have edges. An object edge is a discontinuous change of orientations in surface points. For example sphere object doesn't have object edges, therefore has exactly one object face. The whole of the object's surface points make up the object surface. This also applies to human face, whereby as for human face the surface have no edges. There is no exactly one-object face. Hence, the problems of shape reconstruction are often oriented towards the reconstruction of object faces (surface reconstruction), which only can be calculated approximately (not necessarily true to a specific scale) for certain application. On the other hand, these problems can be characterized by and analysis of depth values or the distance to individual points on the object surfaces (range data analysis).

2.4 Human Facial

Human facial is a complex surface and also contain information of a particular person. That is the reason of choosing a human face as the main focus of the study.

According to Hyo Jong Lee et al, “A human face has been studied for a long time in several areas because the human face has important information, such as sex, age and emotion.” Thus, the important of human face image does lead to the selection of human face to be modeled mainly for the project. Furthermore, “important face features are eyes, nose and lips.” These are the points that should be focused on a human face image. As it may determine the main points that will shape the three-dimensional model.

Today, human face recognition is quite a popular study. Lots of method and approach is used to get the three-dimensional shape from a two-dimensional image. As for this project, some code is written in order to read a bitmap file type of two-dimensional image then to display it in a window. There are also other methods to create a three-dimensional image like as explained briefly below.

According to Baback Moghaddam et al (2003), “the two most critical factors limiting the performance of automatic face recognition systems are pose and illumination”.

Therefore, Baback Moghaddam had come out with the method of Shape from Silhouette (SFS) techniques to reconstruct three-dimensional shape from multiple silhouettes images of an object. The reconstructed three-dimensional shape is called visual hull, which is a maximal approximation of the object consistent with the object’s silhouettes.

CHAPTER 3

METHODOLOGY AND PROCEDURE

3.0 METHODOLOGY AND PROCEDURE

This chapter contains detail description of methodology and procedure used in order to complete this project. It includes the methods and procedures to come out with findings in the research area together with the product. It is illustrate in the Figure 3.0.

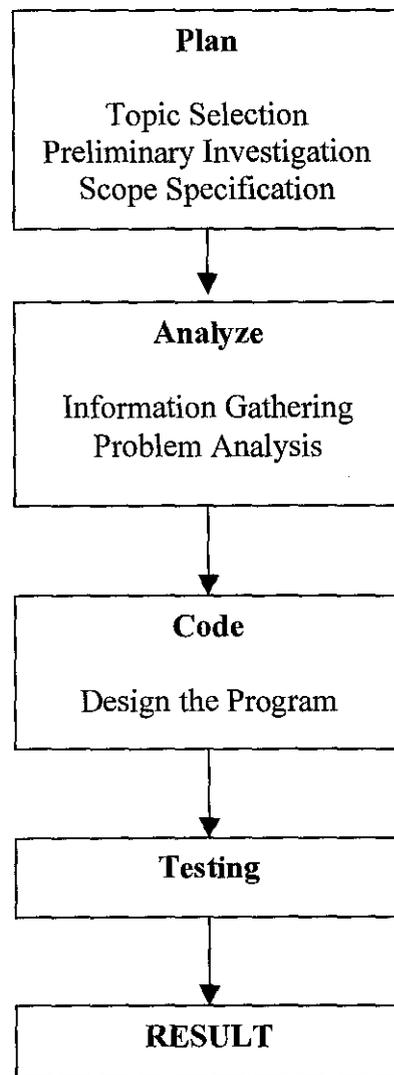


Figure 3.0

3.1 Plan

The first step is to plan the project. Planning is the most crucial part for a project. This phase is the base of a project to grow and to get the required result.

A careful planning regarding on steps on how to complete the project is outlined. The project is break into smaller task.

1. Information gathering
2. Design the output referring to the required requirements and based on the information gathered.
3. Test the code written with different images in order to achieve the objective that is to minimize jaggedness of a three-dimensional image produced
4. Evaluate the application, whether it is usable or not by users.

An outline of the plan created through the Gantt chart using Microsoft Project 2000. The time frame for each task is marked. Every important date in the schedule such as report's submission is highlighted. This outline will then be referred form time to time as to ensure the smoothness of the project. The purpose of the planning phase is to identify clearly the scope of project or the problem by performing the preliminary investigation.

3.2 Analyze

The Analysis Phase is where the project lifecycle begins. During The Analysis Phase is the high-level Project Definition is break down into the more detailed requirements. The Analysis Phase is also the part of the project where to identify the overall direction that the project will smooth the flow of the project.

Gathering information and requirements is the main attraction of the Analysis Phase. The process of gathering information and requirements is done by reading literature reviews. These literature reviews may be gathered through journals, reference books and the Internet to further understand the development of the project.

The literature review is divided by few sub topics, that are;

1. Two-dimensional and three-dimensional graphic
2. Bitmap (BMP)
3. Gray Value and Color Images
4. Depth Analysis
5. Human Facial

All relevant information regarding these sub-topics is gathered as much as possible to help to have clearer understanding towards the project. Some of the information also provides ideas and guidance towards completing the project.

3.3 Code

After the requirements and adequate information is gathered, it is now the time to write the code of the project. This is actually the design phase of the project. During this phase, it will focus on developing a working application according to the scope specification and also the problem analysis based on the information gathered. This is to ensure that the objective of the project is achieved. This project used Microsoft Visual Basic 6.0 and Microsoft Visual C++ software to build the application also use OpenGL. There are two working application expected to be completed at the end of this phase, as stated below.

3.3.1 Convert RGB To Grayscale Image

This application is to ensure the RGB image (color image) will be converted to Grayscale image. The Grayscale value of the image will then be stored to notepad (*.txt) file.

3.3.2 Convert two-dimensional Image To three-dimensional Image

This application is to ensure the two-dimensional Image will be converted to three-dimensional Image. The Grayscale value stored in notepad (*.txt) file will be called and be assigned as the depth value of the three-dimensional image produced.

A logical solution is defined according to the problem statement stated during the earlier phase. This logical solution is then turned into a physical solution by writing the code using related software.

3.3.3 Convert RGB To Grayscale Image

- Microsoft Visual Basic 6.0

3.3.4 Convert two-dimensional Image To three-dimensional Image

- Microsoft Visual C++ 6.0

It is important to ensure that the code written is error free.

3.4 Testing

The code written is tested using some images. Some experiments are done to ensure that the application will produce a smooth three-dimensional image from a two-dimensional image.

The experiments cover that setting for lighting to capture the images, the positioning the image and the image condition to control the lighting. The images will be captured using high-resolution digital camera under different condition as to experiment it and get the factors that may affect the smoothness of the three-dimensional model.

The testing is done by capturing a ping-pong ball using few lighting condition based on try and error to get a good setting for the lighting. It is to ensure that the ball have balance lighting towards every angle of the ball. Once a setting has been established, the setting of lighting was then used to take the images of human face. Few images will be taken as to ensure the best images are captured. A human face image without a controlled lighting is also captured as to compare it with controlled lighting one.

3.5 Tool

The software will be used mainly are;

1. Microsoft Visual C++ 6.0 with OpenGL
2. Microsoft Visual Basic
3. Adobe Photoshop 7.0

Other tools needed are;

1. A set of computer
2. Digital Camera, Fujifilm FinePix A101
 - Simulated display image
 - CCD sensor 1/2.7-inch CCD with 1.3 million square pixels
 - Image file size 1280 ' 960 / 640 ' 480 pixels, 24-bit color
 - File format JPEG, AVI
 - Storage media SmartMedia™ Cards (3.3V, 2 – 128MB)
3. Ping-pong ball

CHAPTER 4

RESULTS AND DISCUSSION

4.0 RESULTS AND DISCUSSION

Modeling a 3-dimensional image from single 2-dimensional image is a possible thing to do. In the end of the project, a window displaying the 3-dimensional image will be displayed and this output will be compared to the 2-dimensional image captured earlier. Some factors will be controlled. The experiment will be done in the laboratory under the controlled environment.

4.1 Software For Project

Visual C++ is chosen as the main platform used to implement the project, as it is the most productive C++ tools for the highest performance development for Windows. Specifically Microsoft Visual C++ brings a new level of productivity to C++, without scarifying the flexibility, performance and control. Microsoft Visual C++ 6.0 has produced a new level of productivity with new features that significantly reduce development time. Developers will spend less time building applications, less time coding, less time compiling, and less time debugging, while enjoying greater component reuse. This will increase speed and generate the code faster than before.

OpenGL is defined as the “software interface to graphic hardware”. It is a 3-dimensional graphics and modeling library that is extremely portable and very fast. It is not a language like the C++ or any programming language. It provides some prepackaged functionality. It has the library that provides various functionality as to ease the developer. In this project, OpenGL is chosen as it is used to optimize for the display and manipulation of the 3-dimensional graphic.

4.2 Image Acquisition

The technique to acquire the 2-dimensional image is also important in order to produce a smooth 3-dimensional image. The key to produce a quality photograph is the lighting. A guideline is referred in order to produce good lighting. The experiment is done indoor. As for that, artificial lighting is used.

4.2.1 Lighting

Artificial light is any light source other than natural light. Artificial light sources include those with a continuous spectrum such as candles, normal electric light bulbs (tungsten lighting), special photographic light bulbs (photoflood bulbs), as well as discontinuous light sources such as fluorescent tubes.

Most of the continuous light sources have more red and less blue than daylight, giving a warmer light. In colour photography you can either alter the colour by using filters or use film designed for the particular artificial light.

There are few methods used to get a photograph for this project

4.2.1 Method 1

Artificial lighting used here is table lamp. Two table lamps are required as to produce a balanced lighting toward the ping-pong ball (the experimented object). One of the table lamp will be placed at the right side of the ping-pong ball, not too near until got highlight to the ping-pong ball, but not too far until cannot see the cast shadow of the ping-pong ball. The other lamp will be placed at left side of the ball. It is to ensure that the light does not directly shine at the ball. Direct it to the wall at the back of the ball. So the light can bounce on the ball as well.

Place the ball around the circle area in the Figure 4.0. It is off-center, which will produce a good quality photograph.

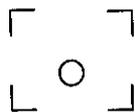
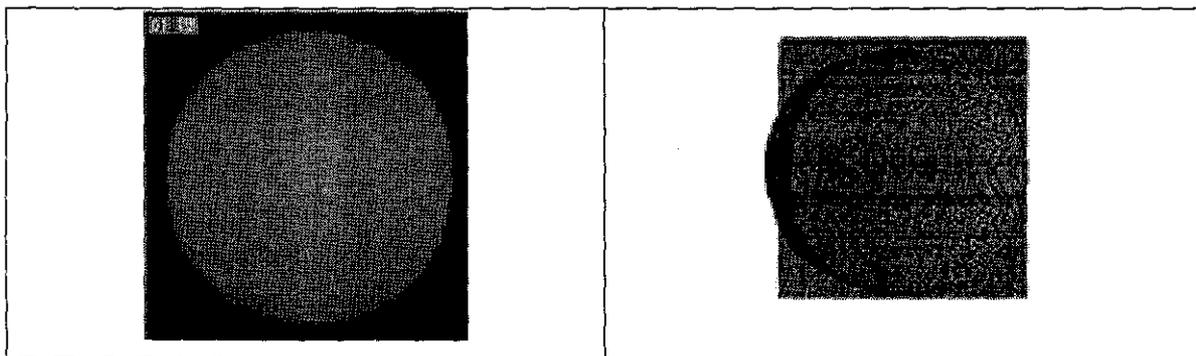
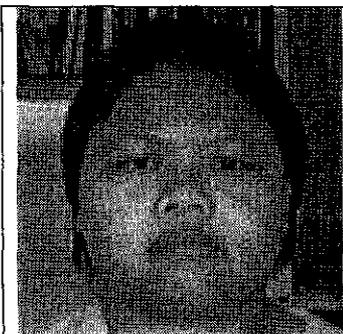


Figure 4.1

4.2.1.1 The output

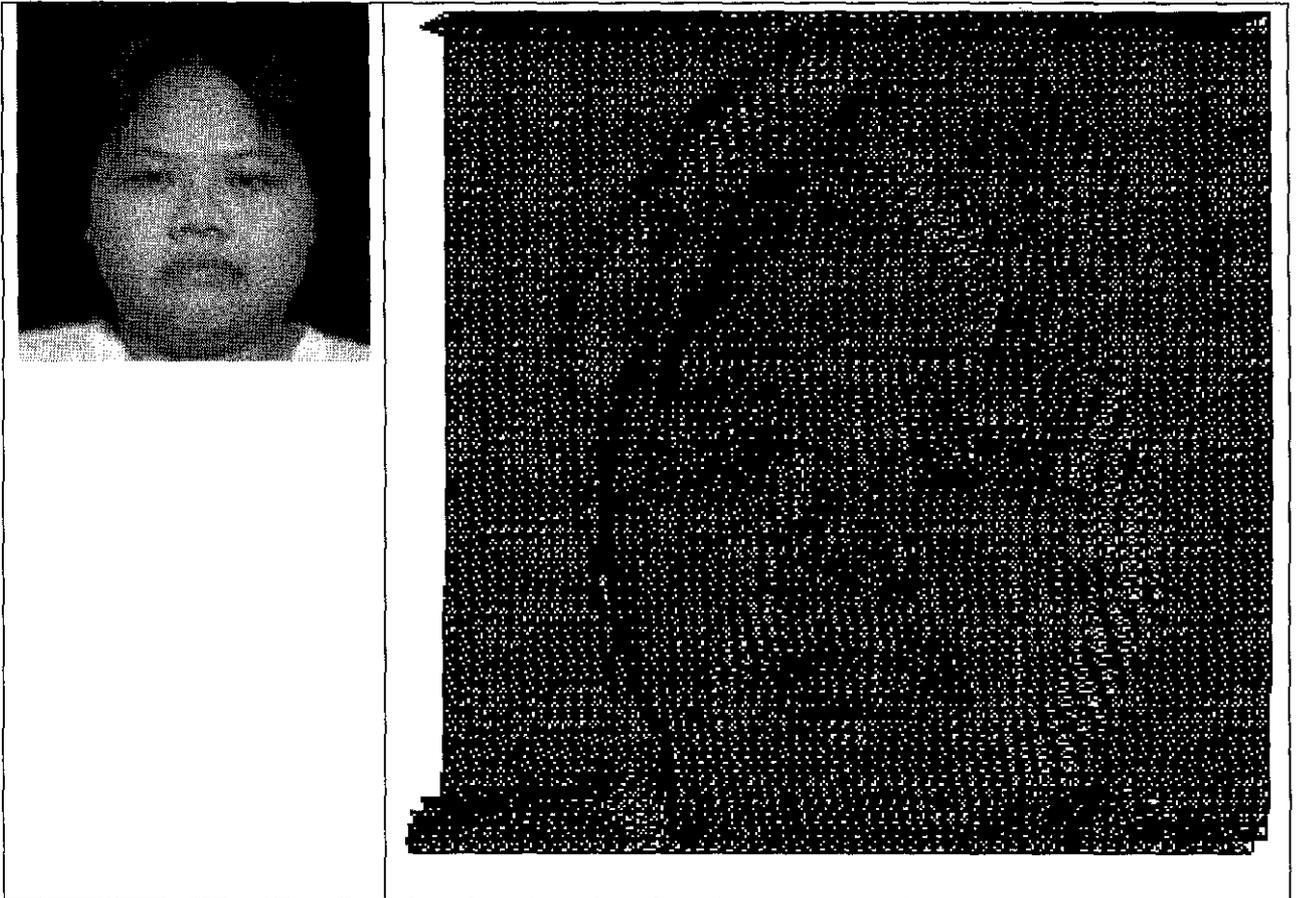


Then the image was further proceed with human face

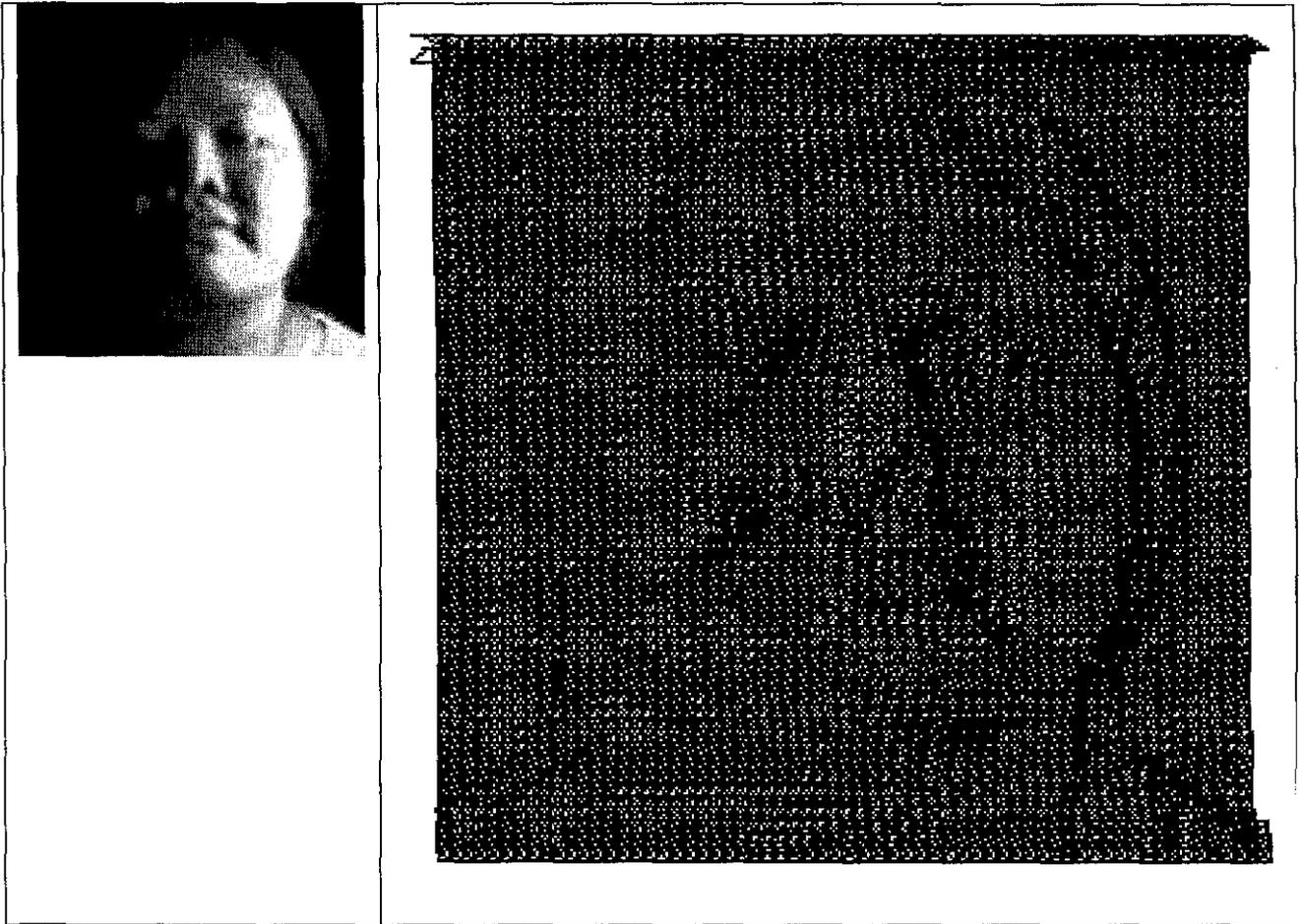


4.2.2 Method 2

4.2.2.1 Source of light (table lamp) placed about 60cm off the ground directly in front of the camera. The background of the image is black as to avoid from light reflection

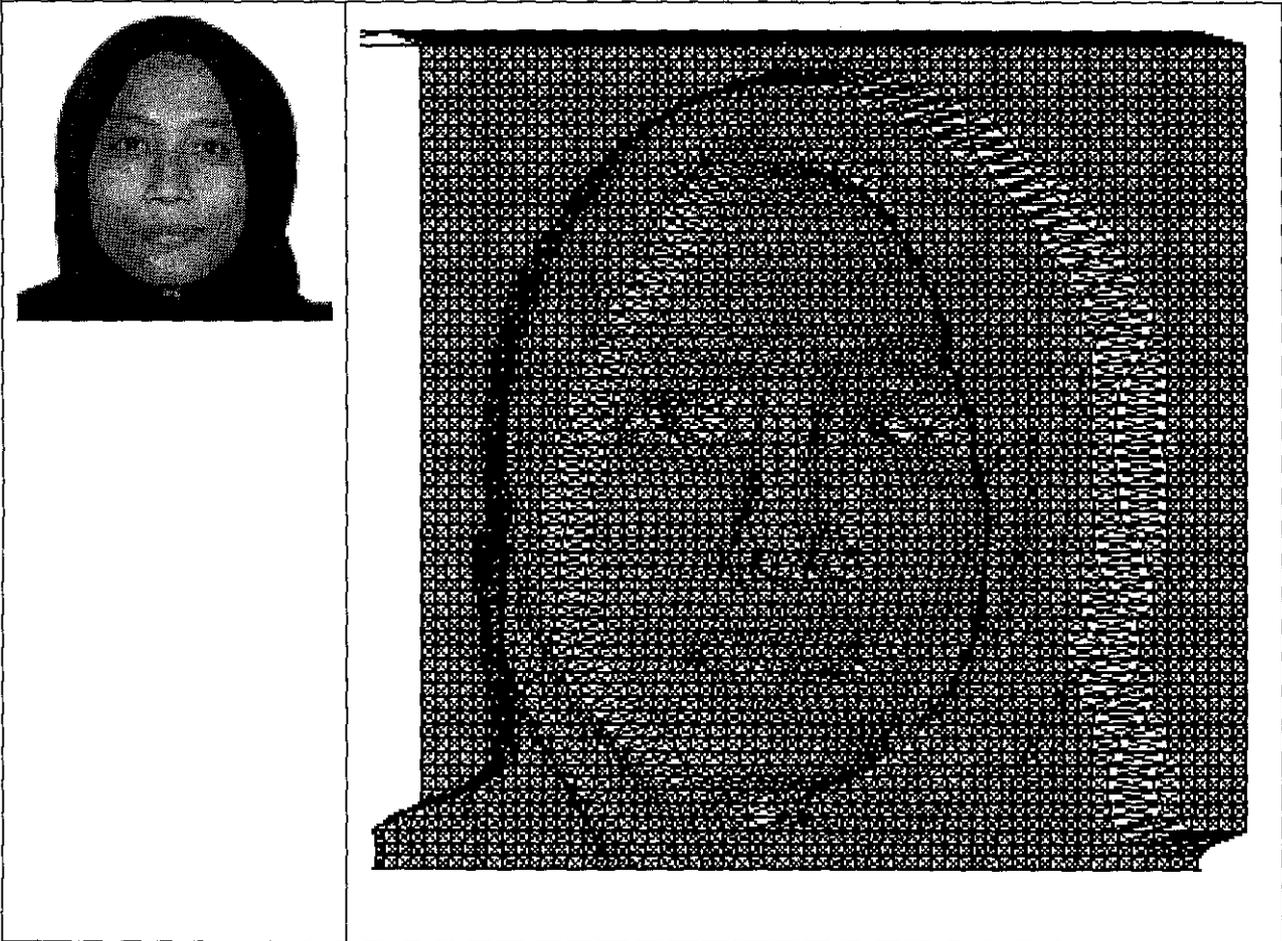


4.2.2.2 A source of light (table lamp) is coming from the side of face



Method 2 of the first experiment whereby the lighting involves a source of light (table lamp) placed about 60cm off the ground directly in front of the camera. The background of the image is black as to avoid from light reflection is chose as t produce the less jagged of three-dimension of human face.

4.2.2.3 Test on other image



CHAPTER 5

CONCLUSION

5.0 CONCLUSION

The depth analysis is very important to this project. The depth of a 2-dimensional image will determine the z-axis of the 3-dimensional model. The depth analysis data can be calculated and stored in a file. This file is yet to be referred in order to acquire the z-axis value to model the 3-dimensional model.

Some factors will be controlled in order to get a smooth 3-dimensional model. This will later be experimented by capturing 2-dimensional images with different conditions.

The end of this project can demonstrate a working product that will model a 3-dimensional object from a 2-dimensional image using OpenGL.

5.1 RECOMMENDATIONS FOR FUTURE WORK

The feature that can be embedded for future work is to add the texture, in the 3-dimensional image (human face) produced. But, there's a problem that may arise. Not everybody's skin color is the same. Some may have dark skin colors and some may have bright. The brightness may have differed value. A standard skin color may be set to standardize the output.

Other possible improvements of the program can concentrate on more sophisticated visual details and bump mapping, on OpenGL support and on more adaptive triangulation of surfaces to speed up rendering.

REFERENCE

- Reinhard Klette, Karsten Schluns and Andreas Koschan 1998, *Computer Vision: Three Dimensional Data from Image*, Springer
- Mashall Martin 1993, *Computer Vision, Model and Inspection*, World Scientific
- Rafael C. Gonzales and Richard E. Woods 1993, *Digital Image Processing*, Addison Wesley
- Katsuchi Ikeuchi and Yoichi Sato 2001, *Modeling From Reality*, Kluwer Academic Publisher
- Shelly, Cashman, Rosenblatt 2001, *System Analysis and Design Fourth Edition*, Course Technology
- Baback Moghadam, Jinho Lee, Hanspeter Pfister and Raghu Machiraju A. 2003, *Model-Based 3D Face Capture with Shape From Silhouettes*, The Ohio State University.
- Ralf Steinmetz and Klara Nahrstedt 2002, *Media Coding and Contents Processing*, IMSC Press Multimedia Series