UNIVERSITY *of the*
WESTERN CAPE

Department of Computer Science
University of Western Cape
Computer Science Honours
Project: Anaglyph Video

By:
Jihaad Pienaar
Supervisors:
Mehrdad Ghaziasgar and James Connan

Glossary

Depth Map	An image or image channel that contains information relating to the distance of the surfaces of scene objects from a viewpoint.
Stereo Pair	Contains two views of a scene side by side, ex. The human eye uses stereo pairs, as the left eye produces a left image and the right eye produces a right image.
OpenCV	A library of programming functions for the use in real time computer vision. [8]
Stereoscopic 3D	Is the illusion of giving an individual the perception of depth in using what would commonly be seen as traditional two-dimensional, or “flat”, visual media.
Graphical user interface (GUI)	Is a human-computer interface (i.e., a way for humans to interact with computers) that uses windows, icons and menus and which can be manipulated by a mouse (and often to a limited extent by a keyboard as well) [5].

Acknowledgements

I would like to thank my parents for their motivation and support as well as providing me with the opportunity to study at the University of the Western Cape. I would also like to thank the National Research Foundation for providing the funding for my honours studies. A big thank you to my supervisor, Mr Mehrdad Ghaziasgar for providing me with the support to complete the project, finally thank you to the rest of the Computer Science department for assisting me throughout my studies.

Contents

Chapter 1

Introduction

Stereoscopic 3D is used in a variety of fields. It started out as a way to make a more immersive entertainment standard, but it has grown and is now used in science, medical research, marketing as well as many other fields.

Medical diagnostics in 3D are helping with cancer and other disease detection and diagnosis, as well as X-ray and MRI analysis. [10]

The entertainment industry also finds great value in stereoscopic 3D, as it is used in studios, theme parks, museums, etc. The stereoscopic 3D provides the viewers with a more immersive and exciting experience.

The average American over the age of 2 spends more than 34 hours watching movies in a week [1]. Television and movies are major forms of entertainment in peoples daily lives. It is because of this, there are millions of dollars being spent in improving these entertainment industries. As such movie and TV studios are spending millions on making their products into better experiences, a way in which this can be achieved is through anaglyph 3D.

The aim of this project is to create an application, in which users can take their videos or movies and create a 3D version of it, making their videos a much more entertaining experience. The user would be able to load their videos into the application and simply convert it to 3D. There are different ways of achieving the 3D effect, this application would use Stereoscopic 3D, also known as Anaglyph 3D [11]

Chapter 2

User Requirements

This chapter focuses on viewing the problem from the users perspective. The solution is based on the information acquired from the user.

2.1 Users view of the problem

The users require an application with an easy to use and friendly graphical user interface (GUI). The application should be compatible with numerous video formats and the conversion process should happen in a timely manner.

The hardware and software requirements for the application are:

(a) Anaglyph 3D viewing glasses.

(b) Windows 7 or Ubuntu operating system.

(c) Open source computer vision library. (OpenCV)

(d) Visual studio 2010.

2.2 Description of the problem

Users who want to have a greater immersive experience when watching movies are forced to pay to watch them at the cinema. This system provides users with a free and simple way to take videos and convert them to a 3D equivalent which can be enjoyed at the users leisure. The system should be simple to navigate and easy to add the videos to be converted.

2.3 Expectations of the software

The software should be simple to install, once installed, users should be presented with a friendly and simple to navigate GUI. There should be an option to add a video from the users hard-drive, this option should be easy to find and labelled effectively. Once the movie is selected, the user should be presented with an option to convert the video. Finally, the conversion should be a swift process and the user should be presented with an option to play the 3D video output.

Chapter 3

Requirements analysis

3.1 The designers interpretation of the problem

To trick the human eye into believing it is seeing depth in 2D/flat images, the anaglyph 3D effect can be used. The effect also involves a red and cyan filter applied to 3D glasses. The original video would be split into its individual frames of images, each of these images would need to be split into a stereoscopic pair, that is, a “left” and “right” image.

To achieve the stereoscopic pair of images a depth map would need to be generated for each of the original images. The original image and the depth map can then be used to generate a stereoscopic pair which can be achieved through a technique called parallaxing. The colour of the left image would then be converted to red and the colour of the right image would be converted to cyan. The two images would be displaced by some amount (depending on the amount of depth we want). Finally, we would re-combine our images to create our 3D anaglyph image. [2] A possible problem would be adjusting the displacement in order to give us the best and most immersive 3D effect.

3.1.1 Development tools

(a) Anaglyph 3D viewing glasses

(b) Open source computer vision library (OpenCV)

(c) Windows 7 or Ubuntu operating system.

(d) Visual studio 2010

(e) Python 2.7.2

(f) The 3D rendering algorithms

3.2 The complete analysis of the problem

(a) User selects a video to convert.

(b) Using OpenCV, the video can be manipulated into its separate frames.

(c) A depth map is generated for each original image.

(d) Using parallaxing, the original image and the depth map are used together to generate a left and right image.

(e) Each original image is then split to create a left and right image

(f) Using Dubois algorithm the left and right image is processed to create a red left image and a cyan right image.

(g) The two images are then combined to create a single 3D image.

(h) The process is applied to each frame of the video and once completed the frames are then put back together to recreate the video.

(i) Using the red/cyan 3D glasses the video can be viewed in 3D.

3.3 Current solutions

Current solutions require users to purchase the application. There are rarely free solutions that provide users with a simple, quick and easy way to convert videos into 3D.
Some of the paid for solutions are:

- Leawo Video Converter - 2D to 3D video converter which costs $29.95 [7]
- DVDFab 2D to 3D Converter - Another paid for 2D to 3D conversion suite. [4]

3.4 Suggested Solution

Since most of the current solutions have to be paid for, the solution suggested will be a free to use application which provides users with an easy way to produce quality 3D videos. The applications will be able to run on Microsoft Windows as well as Linux operating systems.

Chapter 4

User Interface Specification

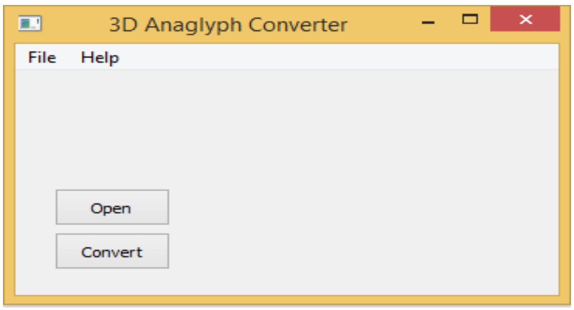
This chapter explores the user interface of the system, how the user will interact with it, and how it will look.

4.1 Description of the user interface

The system will use a graphical user interface (GUI), this means that the user will interact with the system using graphics/images instead of using text based commands.

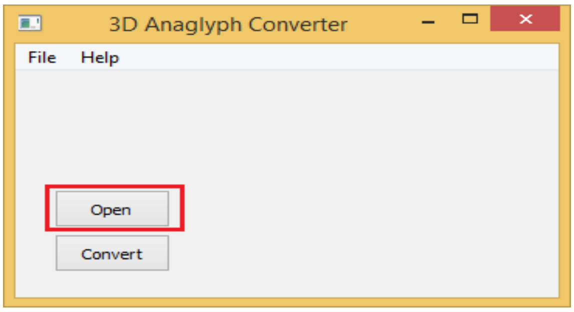
4.1.1 Startup

Upon starting up the application, the user will be produced with a simple window; this window will contain 2 menus and two buttons.



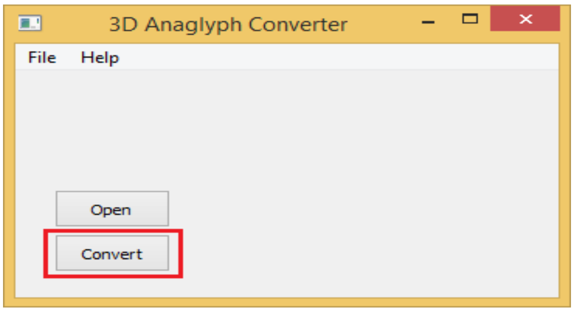
4.1.2 Loading a video

The user can load a video in 1 of two ways: either by selecting the option from the file menu or clicking the button, which will allow the user to select the video they wish to convert.



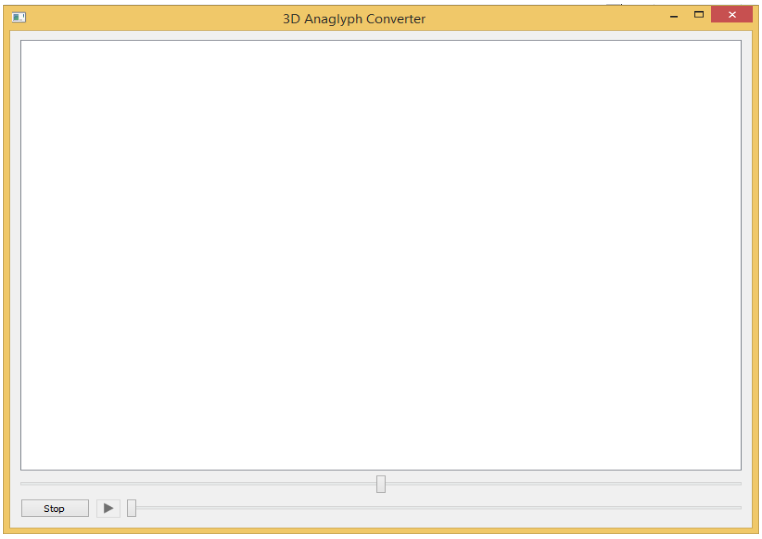
4.1.3 Converting a video

Once a video has been successfully selected the user will be able to convert the video using the convert button, alternatively it can be selected from the file menu.



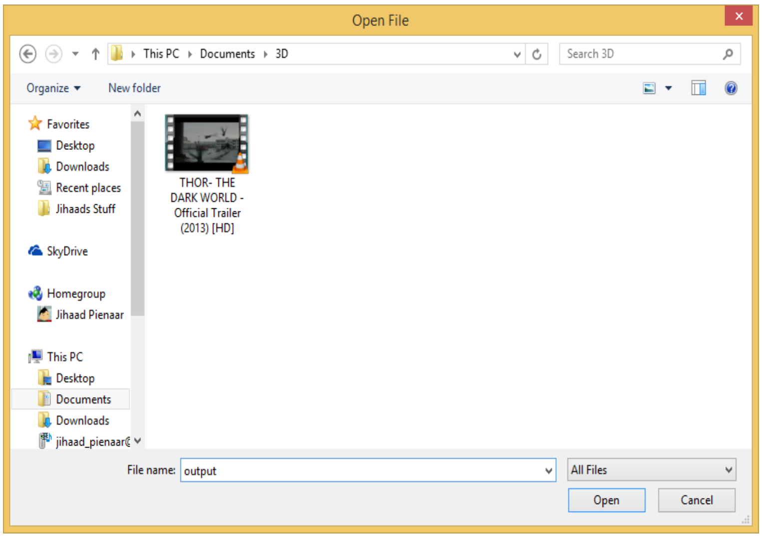
4.1.4 Displaying the video

Once the video has been successfully converted the output will be displayed for the user to view.



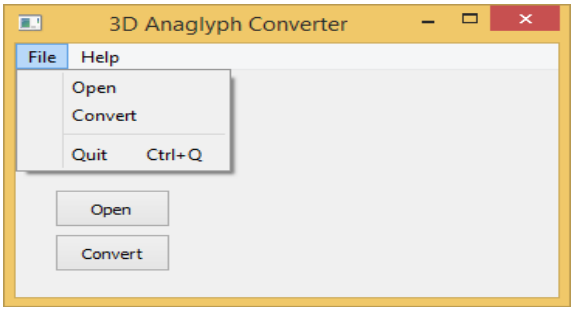
4.1.5 Saving the video

Once the conversion process has completed the user will automatically be prompted to save the video.



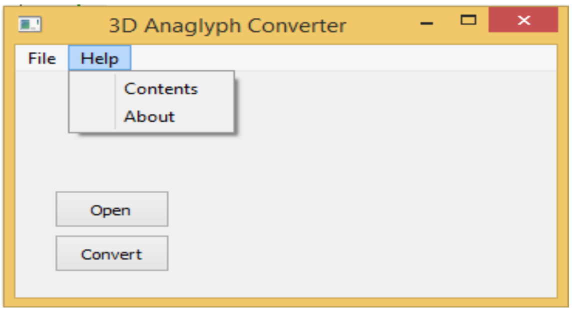
4.1.6 Closing the application

It will also have a quit option allowing the user to easily quit the application.



4.1.7 Help

If there are any problems or difficulty in using the application the user can select the help menu bar.



Chapter 5

High level design

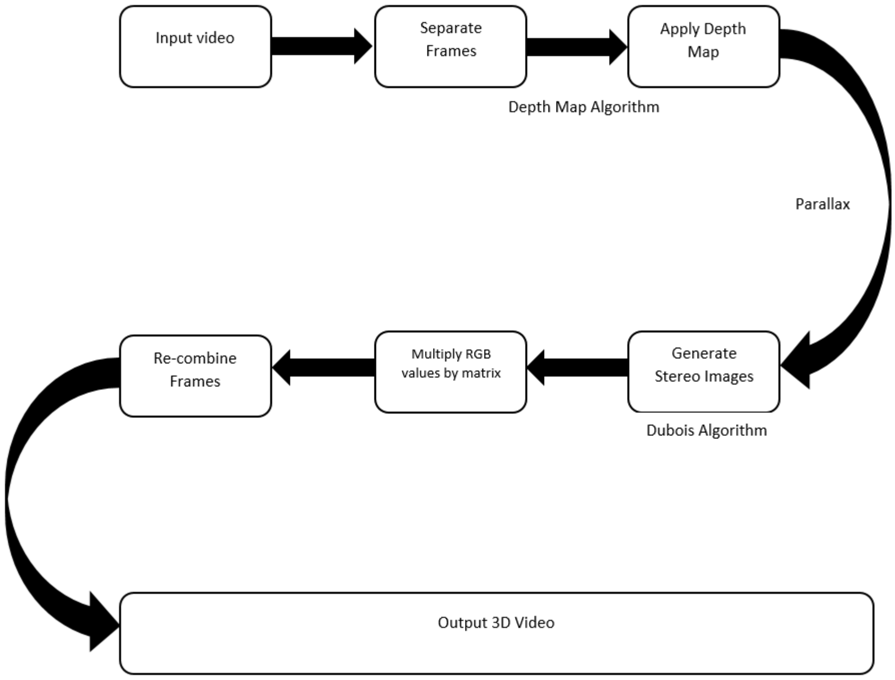
In this chapter, we will look at the high level design view of the problem applied. Since the programming language used was C/C++, an object orientated analysis will not be applied and a class diagram will not be included.

5.1 High level data dictionary

The following contains objects and their descriptions, these objects were used during the development of this research project.

Object	**Description**
OpenCV	OpenCV is an open source computer vision library. It provides functions and data structures that allow for the processing of images.
Stereo Pair	Contains two views of a scene side by side, example. The human eye uses stereo pairs, as the left eye produces a left image and the right eye produces a right image.
Depth Map	An image that contains information relating to the distance between objects, darker shades of grey are usually used to represent object far away, where lighter shades are used for nearer object.
BGR colour space	BGR colour space uses the colours blue green and red, with blue holding the most significant bit and red holding the least significant bit.
RGB colour space	RGB colour space uses the colours red, green and blue with red holding the most significant bit and blue holding the least significant bit.
Least square approximation	An approximation method used to find solutions to equations with more equations than unknowns.
Displacement	The shifting of an initial point to some final point.
Parallax	Difference in the apparent position of an object viewed along two different lines of sight, and is measured by the angle or semi-angle of inclination between those two lines.

5.2 Relationships between objects



5.3 Full solution components

Chapter 6

Low level design

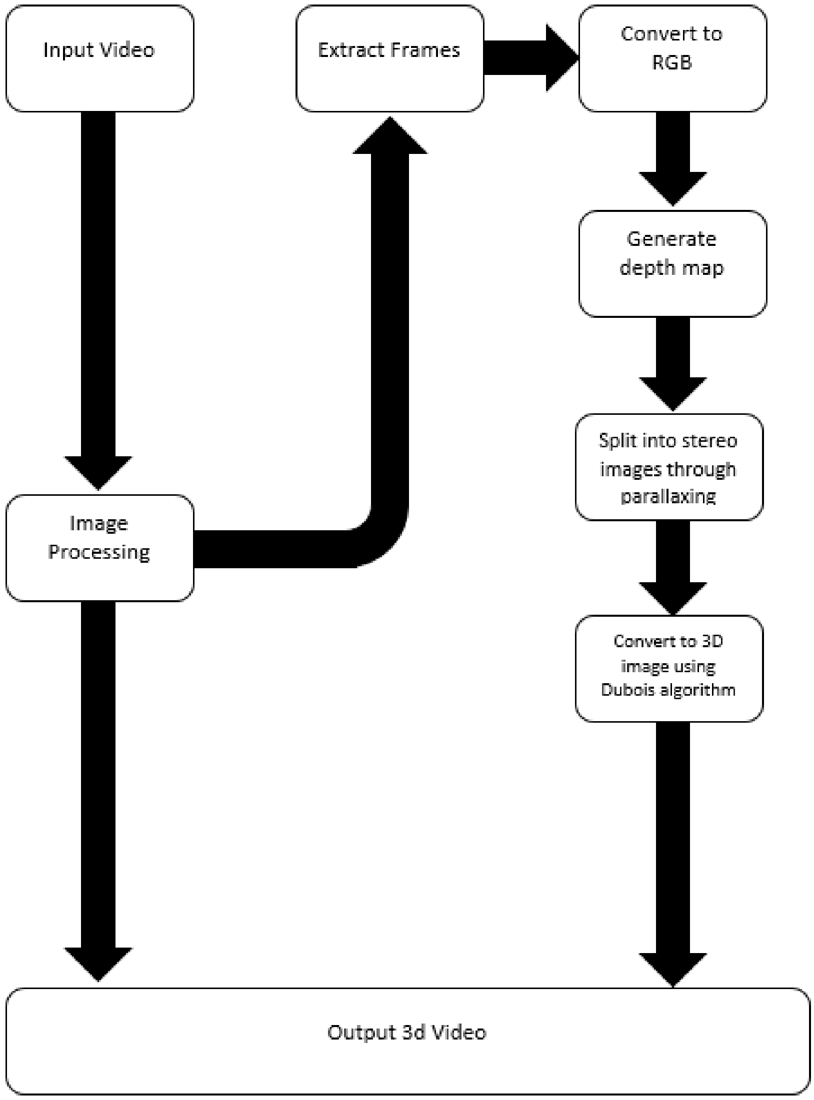
In this section a low level design will be applied to the problem giving details about data types and functions.

6.1 low level data dictionary

The following table contains the system components and concepts that were used.

Class	**Attributes**
Matrix Multiply	MatrixMultiply(LImage[][],RImage[][]) - multiplies the left image RGB values by some pre-calculated left matrix and the right image RGB value with some pre-calculated right matrix.
Depth	Depth (image) - Using the original image, the fourth order moment is calculated for each pixal. The maximum fourth order moment is then used to create a high order statistics map (HOS map).
ImageFill	ImageFill(HOS) - fills the holes in an HOS map to create a depth map.
Displace	Displace(Image, DepthMap) - gets us the amount to displace the original image using the formula: D=30.0*((I-127.5)/(127.5)); where D is the displacement value and I is the pixel value of the depth map.
Parallax	Parallax(Image, DepthMap) - used to create a left and right view of a single image by removing x amount of pixels from the left to create the right image and the same x pixels from the right to create a left image.
Convert to RGB	cvCvtColor(cloneFrame, img_hsv, CV_BGR2RGB) - is used to change the colour space of the image from BGR to RGB.

6.2 Low level design



6.3 Detailed methodology

In this section a detailed analysis of the methodology will be examined, the way each component interacts in order to make the system feasible will be explored.

6.3.1 Video input

The video to be converted is selected by the user. The conversion process will be achieved through a combination of techniques, firstly a depth map is generated for each frame. The depth map and the original image is used to create a pair of stereo images through parallaxing, finally, the Dubois algorithm is used with the stereo images as inputs to create a 3D image.

6.3.2 Convert from BGR to RGB

The frames are captured in the BGR (Blue, Green, Red) colour space. The image manipulation for the various algorithms is best done in RGB (Red, Green, Blue) colour space. The frames would need to be converted before it can be manipulated.

6.3.3 Foreground Extraction

This algorithm will be used to generate a depth map for each frame of the video. The depth map will only contain 2 layers, a foreground and background, this will be used in conjunction with parallaxing to produce stereo images. [9]

The fourth order moment is calculated for the original images, which provides an outline to which section of the image is highly focused and should be the foreground. Finally, the foreground regions are highlighted to produce the depth map. [9]

6.3.4 Parallax

The depth map and the original images will be used to produce a pair of stereo images. The parallax value is calculated for each of the pixels of the original image and the new “left” and “right” images are generated by shifting the original images by parallax/2 to the left and parallax/2 to the right. [6]

6.3.5 Dubois Algorithm

This algorithm devised by Eric Dubois uses least square approximation in order to produce an optimized 3D image. What makes this algorithm superior to other conventional algorithms is that it takes into account the spectral characteristics of the devices; the colour filters in the anaglyph glasses and the cone receptors in the human eye. Based on this characterization the algorithm is produced. [3]

The algorithm uses a pair of stereo inputs and multiplies the RGB values of these inputs by a predefined 3X3 matrix. The final step is to filter the left image with red and the right image with cyan and combine these images to produce the new 3D image. [3]

6.3.6 Re-assembling the Video

In both algorithms, each frame will have some operation done on it, once completed each frame will be re-assembled in order to produce the desired 3D video.

Chapter 7

Code Documentation

The code will be fully documented with the use of comments at all the methods and statements. The code will be available through the form of a CD which will be included.

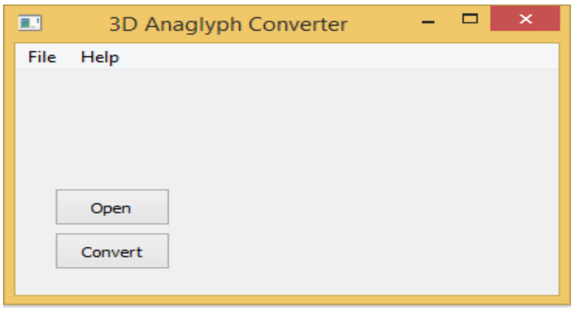
Chapter 8

User Manual

The following figures demonstrate the GUI of the system.

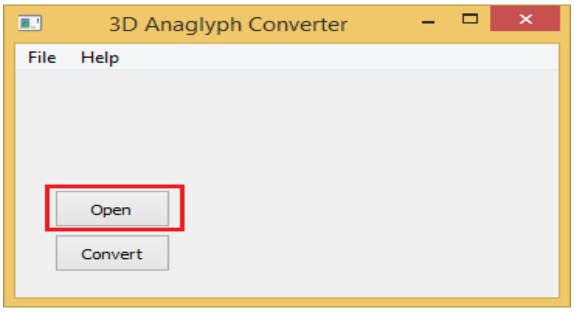
8.1 Starting The System

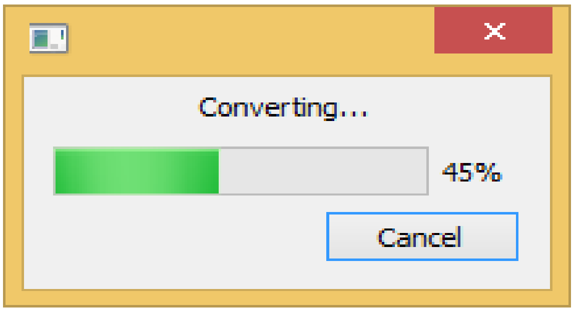
The first window the user would see when starting the application.



8.2 Selecting The Video

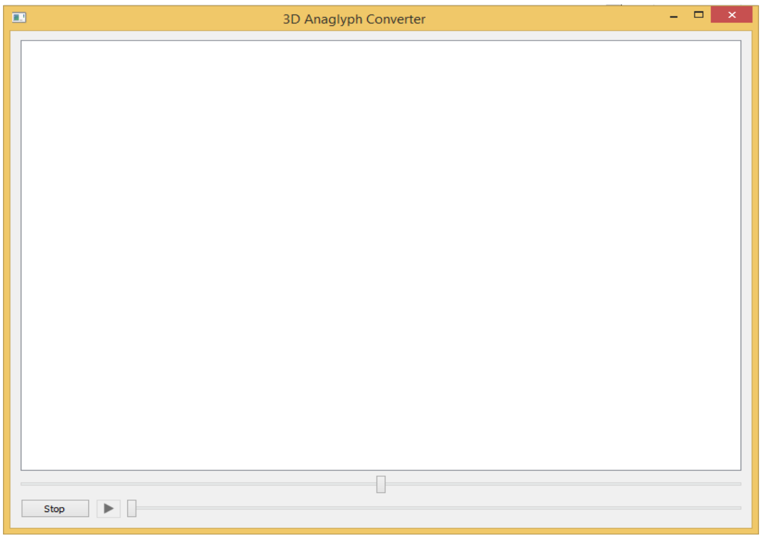
To load the video the user would select the "open" option, to convert the video the user would select the "convert" option. While converting a loading bar should appear.





8.3 Viewing The Output

The output of the conversion process would then be displayed in a new window.



Chapter 9

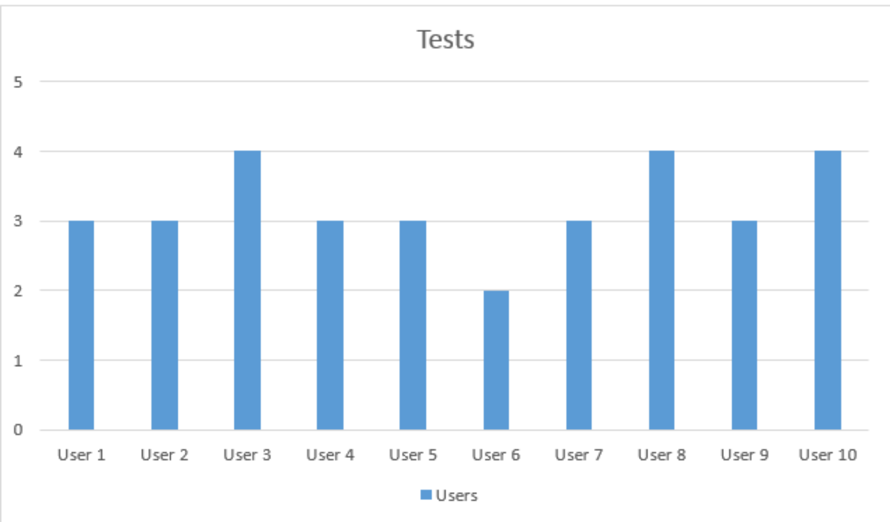
Testing and Results

9.1 Depth Map Test

The accuracy of the depth map generation was tested. The test video used contained about 2320 frames, the frames were manually examined and it was observed that about 77% of the frames were successfully generated.

9.2 3D Effect Test

The strength of the 3D effects were tested through user acceptance testing. 10 interviews were conducted and the results are displayed below:



In the figure above, 60% of the users gave a rating of 3. This indicates that there is room for improvement, such as, increasing the amount of depth layers.

Chapter 10

Conclusion

As defined in chapter 2, the users require a system which is easy to navigate and use. The application produced provides a simple method of converting a 2D video into 3D. The interaction between the users and the system is by means of a GUI which means it would be interactive and easy to grasp. The problem stated is that there are no free methods for users to enjoy a 3D experience, the proposed solution provides a free and easy to use method for users to create 3D videos.

Bibliography

[1] DailyNews. Americans spend 34 hours a week watching TV, according to Nielsen numbers. http://www.nydailynews.com/entertainment/tv-movies/americans-spend-34-hours-week-watching-tv-nielsen-numbers-article-1.1162285, 2013.

[2] Michael Doneus and Klaus Hanke. Anaglyph images still a good way to look at 3d-objects? page 2, 1997.

[3] Eric Dubois. A projection method to generate anaglyph stereo images. pages 1–4, 2001.

[4] FAB. DVD FAB. http://www.dvdfab.com//, 2013.

[5] GUI. GUI. http://www.linfo.org/gui.html, 2013.

[6] Jaeseung Koa, Manbae Kimb, and Changick Kim. Depth-map estimation in a 2d single-view image. *SPIE*, 6696:1–4, 2007.

[7] Leawo. LeawoVideoConverter. http://www.leawo.com/hd-video-converter/, 2013.

[8] OpenCV. OpenCV. http://code.opencv.org/projects/opencv/wiki, 2013.

[9] Jungwoo Park, Jaeho Lee, and Jenq-Neng Hwang. Fast extraction of objects of intrest from images with low depth of field. *ETRI*, 29(3):353–362, 2007.

[10] RealID. Professional Markets. http://www.reald.com/content/markets.aspx, 2013.

[11] Wikipedia. Anaglyph 3D. http://en.wikipedia.org/wiki/Anaglyph_3D, 2013.