Thesis for the Degree of Master of Engineering

Depth and Disparity Visualization of 3D Camera Images in Stereoscopic Imaging Technology

by

Rohit Ramesh

Department of Telecommunication Engineering

The Graduate School

Pukyong National University

August 2011

Depth and Disparity Visualization of 3D Camera Images in Stereoscopic Imaging Technology

입체영상 기술에서의 3D 카메라 영상의 깊이 및 시차 시각화

Advisor: Prof. Shin-Il Jeong
(Co-Advisor: Prof. Wan-Young Chung)

by
Rohit Ramesh

A thesis submitted in partial fulfillment of the requirements for the degree ofMaster of Engineeringin Department of Telecommunication Engineering, The Graduate School,Pukyong National University

August 2011

Depth and Disparity Visualization of 3D Camera Images in Stereoscopic Imaging Technology

A dissertation

by

Rohit Ramesh

Approved by:

____

Professor Wan-Young Chung,
(Chairman)

____

Professor Moon-Gab Joo,
(Member)

____

Professor Shin-Il Jeong,
(Member)

August 26, 2011

Table of Contents

List of Figures

List of Tables

List of Abbreviations

2D	Two Dimensions
3D	Three Dimensions
S3D	Stereoscopic 3D
CCD	Charged Coupled Device
QSPDE	Quadtree Segmentation of Pixel's Disparity Estimation
MSE	Mean Squared Error
PSNR	Peak Signal Noise Ratio
Normalized CC or NK	Normalized Cross Correlation
MD	Maximum Difference
SD	Standard Deviation

Acknowledgement

I would like to express my sincere gratitude to all the people who made an immense contribution in versatile ways throughout the entire period of research as well as successful completion of this thesis. First and foremost, I am thankful to **Professor Shin-Il Jeong** for giving me an opportunity to pursue master's degree in Pukyong National University, South Korea. Also, I am grateful to **Professor Wan Young Chung** to give his kind and encouraging guidance in USN laboratory. I feel delighted that he introduces this interesting topic for my research work and guided me throughout till the completion of this thesis. I personally like to pay courteous regard to my thesis reviewers Professor Wan Young Chung, Professor Shin-Il Jeong and Professor Moon-Gab Joo for their dedication of time and effort in reviewing this thesis and giving invaluable suggestions and advices in favor of completing the dissertation in an effective and systematic way.

In addition, I would like to thank all my colleagues in Ubiquitous Sensor Network Lab, both graduates and undergraduates, who have been very kind and helpful in assisting on my research as well as sharing some unforgettable delighted moments together. I am thankful especially to my lab mates, Sang-Joong Jung, Heung-Sub Shin and An Seong-Mo for their continuous help and support during the period of the research work. My sincere thanks go to an Iranian friend Mrs. Nasim Arshad who always encourage and guide me for the past 2 years.

I am fortunate to be in close with my younger sister Ms. Nidhi Sinha, who is also present in South Korea to pursue her master's course majoring in biotechnology. She has always been my biggest strength whose presence

incessantly gave me confidence and endurance to work hard and give my best.

Last but not least, I would like to show sincere appreciation to my parents thousands of miles away in India for their endless love, support and encouragement during the whole time of my studies.

Depth and Disparity Visualization of 3D Camera Images in Stereoscopic Imaging Technology

Rohit Ramesh

Department of Telecommunication Engineering, The Graduate School,

Pukyong National University

Abstract

The advancement in the stereo vision technology has provided a modern approach in the process of making 3D movies, animations and graphics. In the world of 3D, stereoscopic imaging technology has been extensively used in many applications ranging from Augmented Reality to Virtual Reality. Also, 3D imaging systems is always a challenging task in order to control the convergence and focus of the pair of cameras so as to obtain a clear stereoscopic (3D) image.

The main goal of this research is to extract more accurate and abundant information related with depth and disparity estimation of stereo image pairs by image processing techniques. The visual perception to view a pair of 2D images simultaneously is a crucial factor to build a 3D image. In this thesis, we present the effect of binocular cues (depth and disparity) between the intensities of two stereo images which has essential relevance for the visualization of a stereoscopic or 3D image. Due to this stereoscopic effect, two images of the same scene taken at slightly different positions from the cameras differ in their horizontal positions providing the depth visualization between two images. The term 'disparity' represents the amount of horizontal shifting between the left camera image and the right camera image. Meanwhile, this dissertation seeks to investigate the formation of a 3D image from two stereo image pairs and obtain other relevant information related to it. Thus, this thesis reviews the study based on the comparison between two stereo image pairs and acquiring the significant analytical and calculative approach in the domain of 3D visualization.

1.0 Introduction

The stereo visualization approach to extract 3D information from 2D images is an emerging and innovative technology in the field of making 3D movies and animations which are attracting many countries and researchers to develop more advanced stereoscopic imaging technology. The basic principle of making a 3D image mainly depends on the alignment of the two digital cameras. When both cameras are placed on a horizontal platform then a "parallel-axis stereoscopic camera rig" is formed. The conceptual background of forming a 3D image originates by looking two images at different projections placed at some distances apart.

These two stereoscopic cameras work in resemblance with the human stereo vision. When we see an object, our visual mechanism, vergence and focus, changes to fix the location of the object [1]. Also, focus of eye lens changes to acquire clear images of the object. In the same fashion, the focus of the camera lenses is adjusted to view the object distinctly. To obtain a clear stereoscopic image, zero vertical shifts between the two cameras is considered.

Binocular stereo vision brings the well known correspondence problem. Until now, many researches has been done to undertake the stereo-matching problem (popularly known as correspondence problem) taking either 'area-based' or 'feature-based' approach. Area based approach mainly includes 'cross-correlation' method while feature based includes 'edge-detection' and 'texture extraction' methods. Even though, viewing from different aspects, feature based algorithm are considered as most pragmatic way to implement stereo vision algorithms [2].

1.1 Motivation

Stereoscopy is a technique capable of recording three-dimensional visual information or creating the illusion of depth in an image. A 3D (three dimensional) or S3D (stereoscopic 3D) film is a motion picture that enhances the illusion of depth perception. It is a visual presentation system that is used to maintain or recreate moving images of the third dimension. Originated from stereoscopic photography, a special motion picture camera system is used to record the images as seen from two perspectives and special polarized 3D glasses are used to provide the illusion of depth when viewing the film. Since, 3-D films were prominently featured first time in the year 1950s in American cinema, later on it became more and more successful throughout in year 2000s around the globe. The unprecedented success of 3D movie "Avatar" in 2009 has brought much more interest in the world of stereo vision technology. Hence, the motivation of this thesis is to gather more adequate and abundant information on depth and disparity visualization of stereo image pairs which can be helpful for further development of stereoscopic imaging technology.

1.2 Challenges

When viewed stereoscopically, it showed that the two images are combined by the brain to produce 3D depth perception. Human stereo vision changes focus of the eye lens to fixate the object. As a result, the clear images of the object are obtained. Also, vergence angle between left and right eyes changes in order to fixate different object distances [3]. However, sometimes a troublesome problem occurs with Virtual Reality when users suffer from motion sickness-like symptoms popularly called as cybersickness [4]. This symptom is mainly occurred when there is difference in the virtual distance comprehended by horizontal disparity and focus, causes eyestrain or headache. Another important reason of cybersickness is the object distance. If object distance of a stereoscopic image is very near to the eye lenses then it brings eyestrain to the viewers. Therefore, in the procedure of forming a stereoscopic still or motion image, it becomes essential to take proper care of the object distance in order to avoid cybersickness to the viewers. Along with this problem, there is another difficulty in clear visualization of the horizontal disparity between the two stereo image pairs when both images are taken at very small shifting of the cameras. Therefore, in those kinds of images, disparity between the two images is very less and hence it becomes an arduous job to distinguish the clear shifting at every corresponding points or edges of the two images.

1.3 Research Aims

By highlighting the challenges occurred, our research aim is to manipulate a linear relation between the object distance and the image distance with the help of geometry of a stereoscopic camera lens system. The diagram of the geometry of camera lens system is shown in Fig. 1-1 [5]. The operation of a stereo camera mainly depends on the vergence and focus of the camera. To reduce the problem of cybersickness, as mentioned before, it is inevitable to know the position of the CCD sensors with respect to the lens because it only determines the distance to an object through the lens equation.

An object is considered which is placed at 'pd' distance from the camera lens in the figure 1-1. 'f' be the focal length and 'id' be the image distance of the object. The equation of the camera lens is represented as:-

$$\frac{1}{f} = \frac{1}{i_d} + \frac{1}{p_d} \quad (1)$$

Finally, on manipulating the lens equation, the linear relationship between the object distance and the image distance with respect to focal length of the camera lens is represented as:-

$$i_d = \frac{p_d \times f}{p_d - f} \quad (2)$$

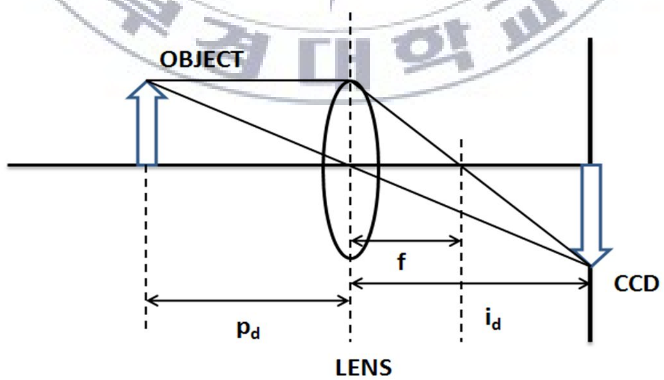


Figure 1-1 Geometry of a Stereoscopic Camera Lens System.

Taking f1=25mm, f2=35mm, and f3=50mm respectively, we can obtain the corresponding graph in MATLAB 7.4.0. as shown in figure 1-2.

Observing this graph, it is concluded that the most probable focal length of lens to be used for getting stereoscopic images is between 25mm to 35mm. By using focal length of lens above 50 mm can easily cause the problem of getting abrupt 3D images which will cause the problem of cybersickness among the viewers.

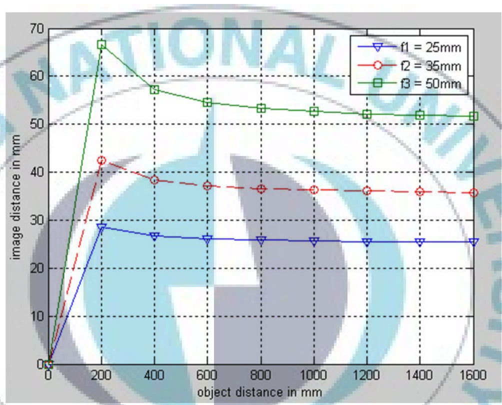


Figure 1-2 Graph between Object Distance and Image Distance.

Also, main goal of the research is to overcome the visualization problem of horizontal disparity between the two images when there is very less disparity between them.

1.4 Contribution

Through the research works, lots of efforts had been contributed to overcome the challenges faced during building the basic concepts of forming an anaglyph image taking two stereo image pairs. The main idea of this thesis is to use the two stereo image pairs to form an anaglyph image and further obtain more information by image processing techniques. In contrast to the information available on several research papers, journals and also on internet, we are concentrating on obtaining the useful selective information which could be helpful for our research on stereoscopic images. Main objectives of our thesis are as follows:

1. The formation of 3D anaglyph image from two stereo image pairs of the same scene taken at slightly different positions.

2. Manipulation of linear relation between object distance and image distance in order to acknowledge the cause of cybersickness and select suitable range of using focal length of the lens for obtaining better stereoscopic images.

3. To reduce the presence of unavoidable environmental conditions like noise from a combined image (imbricating left camera image over right camera image to obtain a single synthesized image) using suitable filters.

4. Visualization of depth between two images when combined together to form a single image in three dimensional (3D) view.

5. Comparing the existing texture extraction and feature extraction methods with our efforts of using the same methods in different perspective for better visualization of stereo images.

6. Propose a new technique for visualizing stereo image pairs which have less disparity among them in order to overcome well known correspondence problem.

7. Distinguishing the difference between the formation of a 3D image and a single synthesized image obtained by 'Image Fusion' method.

8. Comparing the two original stereo image pairs with several distorted 3D images in order to obtain image quality measurements.

1.5 Chapter Organization and Summary

The structure of this dissertation is organized as follows. Chapter 2 covers the literature review and related works which inspired the idea of forming a 3D image. Chapter 3 describes about the depth and disparity maps obtained by imbricating left camera image over right camera image. Chapter 4 reviews the comparison between texture extraction algorithm and feature extraction algorithm. Also, the difference between forming a 3D image and single synthesized image obtained by image fusion method is discussed. In Chapter 5, original stereo images are compared with their distorted 3D images to obtain image quality measurements between them. The experimental results are discussed on the performance of building 3D images in Chapter 6. Lastly, Chapter 7 summarizes the discussions and works that have been presented throughout the whole thesis.

2.0 Literature Reviews and Related Works

This chapter gives a brief introduction on background of the stereoscopic imaging technology which leads to the creation of a 3D anaglyph image. The formation of a 3D image is the basic step for obtaining further information on perception of depth and binocular disparity in this research. The rest of chapter shows discussions and related works on classifications of depth perception as well as detailed description on horizontal disparity between stereo image pairs.

2.1 Background of Stereoscopic Imaging Technology

Stereoscopy word is originated from Greek where stereos mean “Solid” and skopeo means “to look” or “to see”. “Stereoscopic Imaging Technology” refers to a technique for creating or enhancing the illusion of depth in an image by presenting two images separately to the left and right eye of the viewer. Both of these 2D images are then combined in the brain to give the perception of 3D depth. The 3D image can be seen by a viewer using special polarized glasses. These polarized glasses help to combine separate images from two source images and also filter both images seperately.

The basic principle behind obtaining two separate 2D images is by mounting two digital cameras on a horizontal platform called as ‘stereoscopic camera rig base’ system and then taking pictures of the same scene at the same time. A tripod is used in order to give support to both cameras with rig base and a 3D monitor to view both combined images as a single 3D image. The alignment of both cameras is crucial factor for obtaining clear stereoscopic image. The two cameras need to be kept aligned within close tolerances in all three axes. The two images must be identical except for the essential slight image shift in the horizontal plane (x axis), which gives stereoscopic parallax (disparity). There must be no difference between the images in the vertical plane (y axis) as the images obtained by different vertical alignment of two cameras create distortion and hence cannot be fully corrected during post processing of images. Therefore, zero vertical shifts between the two cameras are considered. Also, there is a need to control the zoom, focus, interaxial (distance between the cameras) as well as the convergence point of the two optical axes to obtain clear stereoscopic images.

2.2 Formation of 3D Anaglyph Images

Anaglyph images were the earliest method of presenting theatrical 3-D which requires less specialized hardware. It is one of the most significant terms associated with stereoscopy. Anaglyph images are used to provide a stereoscopic 3D effect, when viewed with glasses where the two lenses are different (usually chromatically opposite) colors, such as red and cyan. Thus the images are made up of two color layers, superimposed, but compensate with respect to each other to produce a depth effect. The polarized glasses with colored filters in each eye separate the appropriate images by canceling the filter color out and provide better view of the stereoscopic image. The visual cortex of the brain fuses the two images of different colors into perception of a three dimensional (3D) scene or composition.

In historical methods using camera filters, on film, two images from the perspective of the left and right eyes were projected or printed together as a single image, one side through a red filter and the other side through a contrasting color such as blue or green or mixed cyan. Nowadays, an image processing computer program is used mostly to simulate the effect of using color filters, using as a source image a pair of either color or monochrome images. In recent simple practice, the left eye image is filtered to remove blue & green. The right eye image is filtered to remove red.

Figure 2-1 3D Anaglyph Images.

2.3 Depth Perception in 3D Image

Depth perception is the visual ability of human vision to perceive the world in three dimensions (3D) and the distance of an object. Our minds are able to perceive the world in three dimensions primarily because we have binocular vision. Depth perception arises from a variety of depth cues. These are typically classified into monocular cues that require the input from just one eye and binocular cues that require input from both eyes. Size of object play key role in monocular cues. The objects which are far make smaller visual angles than near objects. Binocular cues include stereopsis, yielding depth from binocular vision.

2.3.1 Monocular Cues

Monocular cues provide depth information when viewing a scene with one eye. When an observer moves, the apparent relative motion of several stationary objects against a background suggests about their relative distance, this is popularly known as motion parallax. If information about the direction and velocity of movement is known, motion parallax can provide absolute depth information. This effect can be seen clearly while travelling in a train. Nearby things pass quickly, while far off objects appear stationary.

2.3.2 Binocular Cues

It provides depth information when viewing a scene with both eyes. By using two images of the same scene obtained from slightly different angles, it is possible to view the distance to an object with a high degree of accuracy. If an object is far away, the disparity of that image falling on both retinas will be small. If the object is close or near, the disparity will be large. It is the effect of stereopsis that creates illusion to people into thinking that they perceive depth when viewing 3D movies and stereoscopic photos.

Binocular depth cues are based on the simple fact that a person's eyes are located in different places. One cue, binocular disparity, refers to the fact that different optical images are produced on the retinas of both eyes when viewing an object. By processing information about the degree of disparity between the images it receives, the brain produces the impression of a single object that has depth in addition to height and width. Another term called as ‘Convergence’ which is basically a binocular cue for distance or depth perception. By the method of stereopsis, the two eye balls focus on the same object and in doing so they converge to each other. The convergence stretches the extraocular muscles and the kinesthetic sensations from these extraocular muscles help in depth and distance perception. The angle of convergence is smaller when the eye is fixating on far away objects. Convergence is effective for distances less than 10 meters.

2.4 Binocular Horizontal Disparity

The horizontal difference between the viewpoints of two eyes to see an object of the same scene at slightly different positions is called “Binocular Horizontal Disparity”. It is the form of depth perception mostly used by the human brain, and is the most easily manipulated for perception tasks. In stereopsis, the brain uses binocular disparity to extract depth information from the two-dimensional retinal images. Usually, the human eyes are horizontally separated by about 50-75 mm (interpupillary distance) depending on each individual. Thus, each eye has a slightly different view of the images. This can be easily observed when alternately closing one eye while looking at a vertical edge. The binocular disparity can be observed from apparent horizontal shift of the vertical edge between both views.

In stereo vision technology, binocular disparity refers to the difference in coordinates of similar features within two stereo images. It is calculated from stereo image pairs taken from a set of digital stereo cameras. The horizontal distance between these two cameras, called as the baseline, can affect the disparity of a specific point on their respective image plane. As the baseline increases, the disparity increases due to the greater angle needed to align the sight on the point. However, in stereo vision, binocular disparity is referenced as coordinate differences of the point between the right and left images instead of a visual angle. The units of images are usually measured in pixels.

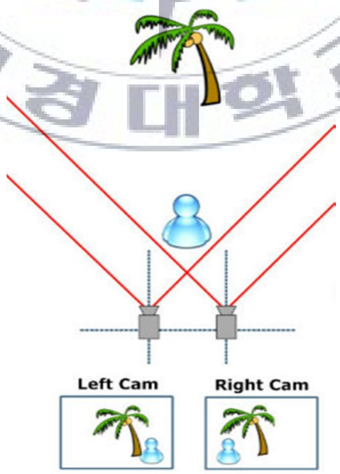


Figure 2-2 Stereo Image Pairs of the Same Scene.

The specifications of digital cameras to be used for our research work are described in the table below.

Table 2-1 The Specifications of Digital Cameras used in the 'Rig' Model.

DIGITAL CAMERA PROPERTIES	SPECIFICATIONS
Model	SONY-PMW EX3
Focal length	½ inch
CMOS Sensor Size	½ inch
Pixel Resolution	1920*1080(HQ Mode)
Signal to Noise Area	54 dB
Wide angle view	5.8 mm
Shutter Angle	180 degree

The visualization of horizontal disparity can be used in further extraction of information from stereo image pairs. Horizontal disparity is most useful is for depth and distance calculation. Disparity and distance from the cameras are negatively correlated. As the distance from the cameras increases, the disparity decreases. This allows finally creating depth perception in stereo images.

Figure 2-3 Left Camera Color Image and Right Camera Color Image.

2.5 Related Works

Context awareness refers to the idea about basic principle, formation and working of stereo image pairs in form of 3D image. All relevant information related with creation of stereoscopic (3D) images describes about the perception of depth being visualized by human vision and work out in detail comprising the importance of binocular horizontal disparity visualization of 3D images. These factors play crucial role in building a clear stereoscopic image and helps to avoid distortions which is mainly responsible for creating cybersickess among the viewers. Below are the summarization of some of the researches conducted by other researchers on stereoscopic imaging technology consisting of depth perception, monocular cues, binocular cues and horizontal disparity of 3D images.

(1) Clark et al. [6] describes a stereo vision approach to depth perception in which a set of programs was developed. (a) Position and orientation of cameras in 3-space has been considered. (b) Using two

pictures, loci of point pairs are found. (c) Each point pair is computed in 3-space. (d) Resultant depth information is extracted.

(2) K. C. Kwon et al. [7] proposed a binocular stereoscopic camera vergence control method for getting horizontal disparity information. The key-object in left and right images are extracted by labeling the centre area of image and then the disparity information is calculated.

(3) R. Srinivasan et al. [8] describes about the effect of monocular cues like shading, texture, occluding contours and binocular cues like disparity on stereo image pairs. The basic approach is to integrate the factors like monocular and binocular cues to build a strong shape of stereo image.

(4) Pascal et al. [9] proposed a simple and parallel technique to achieve the dense depth maps. Both feature based and area based approach is considered by correlation followed by interpolation techniques.

2.6 Chapter Summary

This chapter discusses the prior works in formation of a 3D image by the superposition of two images producing the depth effect. These two images are taken of the same scene at slightly different positions. The depth information provides two categorization namely monocular cues and binocular cues for perception of depth by looking from one eye and both eyes respectively. There is a direct relationship between depth and disparity to see an object considering monocular and binocular cues. The difference between two eyes is an important factor to view two objects at slightly different positions in order to create a disparity called as binocular horizontal disparity. Thus, our thesis work suggests the studies of formation of an anaglyph image providing the depth information as well as it shows horizontal disparity between two images.

3.0 Depth and Disparity Maps of Stereoscopic Image Pairs

In this chapter, stereoscopic views estimating the depth information between two images as well as the horizontal disparity map of stereo images is discussed. In general, the 2D to 3D conversion from two images is assigned to generate depth map information from 2D images. To obtain depth map image, the basic hardware component used is two digital cameras placed on a horizontal platform called as 'rig' for obtaining the stereoscopic image pairs. To generate the stereo image pairs, the depth information of the objects inside the scene has to be figured out. Finally, the left eye and right eye images give to the user a 3D visualization when combined together.

3.1 Depth Map

A depth map is defined as a two-dimensional array where the x and y distance information corresponds to the rows and columns of the array in both the stereo image pairs, and the corresponding depth measurement (z values) are stored in the array's elements (pixels). In order to obtain the depth information, a pre-processing of an image is applied to extract the relevant information from the two input images. Pre-processing of an object basically includes the classification of an image. In this process, detecting regions in an image (like objects which are close to the screen and objects that are far away from the screen) is the foremost important step for creating depth map of the scene. The position of objects placed near the screen is called as ‘negative parallax’, objects placed at the screen are called as ‘zero parallax’ and objects situated behind the screen are termed as ‘positive parallax’. The precise control of placing the objects in positive parallax and negative parallax tends to create the required depth between two images and hence a crucial factor for obtaining a clear 3D image. Geometrically, this parallax values depends on the inter-ocular difference and baseline distance between the two cameras.

3.1.1 3D Visualization of Depth Map

In 3D computer graphics, a depth map is an image which contains the information relating to the distances of the surfaces of scene object. This is possible by imbricating the left camera image over the right camera image. Figure 3-1 represents left camera image and right camera image in gray. These two images are overlapped together to form a depth map. Figure 3-2 shows the illusion of depth map obtained from two gray images. Originally these images are color obtained from left camera and right camera respectively and then it is converted into gray by image processing techniques. It is observed that the darker shades means object is closer and lighter shades means object is farther away. Hence, a depth is estimated between them. Also, the distance of depth between these two images can be observed in three dimensional views as shown in figure 3-3.

Figure 3-1 Left Camera Gray Image and Right Camera Gray Image.

Figure 3-2 Depth Map of both Gray Stereo Image Pairs.

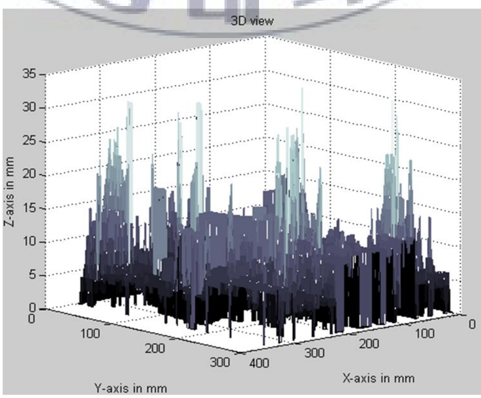


Figure 3-3 Depth between two images in 3D view.

3.2 Effects of Disparity and Depth on 3D Images

‘Convergence’ control as well as ‘Focus’ control plays a significant role for creating a clear 3D image. This mainly depends on the alignment of digital cameras on stereoscopic camera rig system. After placing the two digital cameras on the rig model, horizontal shifting is performed by moving the rig in different horizontal positions. The effect of this shifting is observed on the formation of different 3D images. By taking proper measurements of the shifting between two digital cameras, finally one position is obtained where a clear 3D image can be observed. Along with the disparity of 3D image, the position of objects (considered for this experiment) are also placed at different distances from the stereo camera rig system in order to visualize the depth.

3.2.1 Movement of Stereoscopic Camera Rig

In this part, several experiments are performed by shifting the camera rig system in horizontal directions and its effect on 3D image is observed. The 3D camera rig system used for our experiment is shown in figure 3-4.

Figure 3-4 Stereoscopic (3D) Camera Rig.

At first, the two digital cameras mounted on the rig base as shown in figure 3-5 are placed at extreme end positions and their shifting from the centre is measured as 140mm.

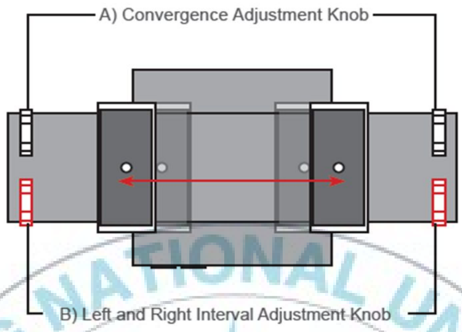


Figure 3-5 Horizontal Motion of Stereoscopic Camera Rig Base.

The resultant 3D image obtained from the cameras when placed at extreme end positions are obtained in figure 3-6.

Figure 3-6 3D Image when Stereo Cameras are at 140mm from the centre.

From this figure, it can be analyzed that depth visualization of different objects placed at different locations cannot able to form a clear 3D image and also brings eye pain while looking the stereo image with the polarized glass for long time.

Now, the two cameras placed on camera rig base are slowly moved to each other. At one particular position at about 100mm from the centre, it is observed that relatively better 3D image is obtained as compare to the image obtained in figure 3-5. The position of different objects is same as before. Depth visualization in this case is also slightly improved as shown in figure 3-7.

Figure 3-7 3D Image when Stereo Cameras are at 100mm from the centre.

From this experiment, it can be analyzed that if the movement of the camera rig base is decreased then at some particular point a clear 3D image with proper depth visualization can be obtained.

On continuing the experiment, it is observed that when both the cameras are at a distance of 70mm from the centre and depth between all objects are considered same then a clear 3D image is obtained as shown in figure 3-8.

The distance between some objects (3 water bottles of the same size are taken for this experiment) from the camera rig base are measured as 1.35m, 1.77m and 2.25m respectively. The objects at different distances create depth which is now easily visualized as compare to previous images.

Figure 3-8 3D Image when Stereo Cameras are at 70mm from the centre.

3.2.2 Depth Visualization of Objects at Different Locations

Some of the factors like 'negative parallax', 'zero parallax' and 'positive parallax' as shown in figure 3-9 are essential for creating depth between the different objects in a 3D image. 'Zero parallax' is a point where the object is placed on plane of the screen.

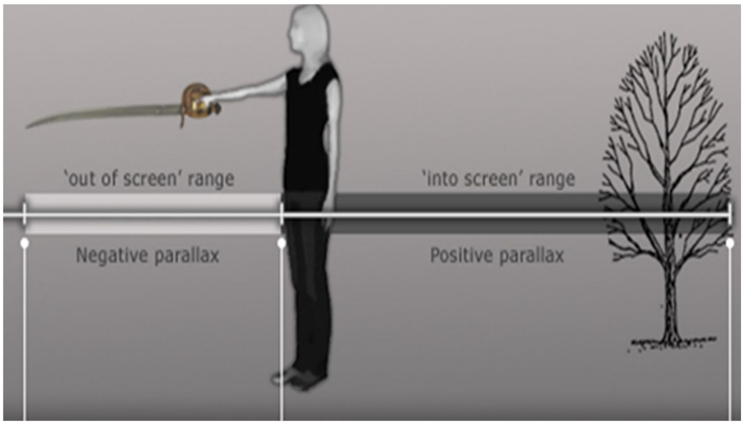


Figure 3-9 Factors for Creating Depth in a 3D Image.

In figure 3-9, the position where a girl is standing is called as 'zero parallax' point. The objects placed in front of this screen that is from girl's hand till end of the sword are called as 'negative parallax'. This position is mainly used in making of 3D movies or animations where all objects gives an illusion effect to appear out of the screen and they are visualized as it is very close to our eyes. The regions behind the screen (from girl's position to end of the tree) are called as 'positive parallax'.

On considering the different parallax effect, an experiment is performed in which measurements between object very close to the camera as well as objects at farthest point from the camera are observed. The distance between the 'man sitting on chair' and camera rig system is 1.22m as shown in figure 3-10. This is called as 'negative parallax'. This region has significant importance for creating an illusion effect in making of 3D movies and animations. If the figure 3-10 is visualized on 3D screen by wearing 3D polarized glasses then 'hand of the man' will seems to be very close to our eyes. This visual perception also helps to visualize depth between different objects in a 3D scene.

Figure 3-10 3D Image considering Parallax Effect.

In figure 3-10, the distance between man and computer placed behind him is 2.42m. This region is called as 'positive parallax'.

3.3 Disparity Map

The term 'disparity' used to describe the positions of the corresponding objects in an image seen by the left and right eyes. The disparity map is a map from which the depth at every pixel can be understood clearly. The shifting or disparity of each pixel of right camera image as compared to left camera image or vice-versa can be computed from estimated depth map value. Disparity map in the form of grayscale image can be generated by using simple image processing in MATLAB 7.4.0.

3.3.1 Median Filtered Disparity Map

In the process of taking stereoscopic images, sometimes it becomes tedious job to avoid all environmental conditions which can affect the quality of the images. In one case, if the images of some objects are taken inside a closed room then the possibilities of distortion in an image can be reduced to some extent but not completely. Here, one such distorted disparity map caused by some environmental condition like noise as shown in figure 3-11 is considered in the final image. By image processing technique, the distorted image of the disparity map caused by noise factor is tried to reduce by using a filter called as median filter.

Figure 3-11 Disparity Map containing Noise Factor.

The median filter is basically a nonlinear digital filter which is often used to remove noise. Such noise reduction technique is a pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is widely used in stereo image processing because it also preserves edges while removing noise. Median filtering is one kind of smoothing technique same as linear Gaussian filtering. Usually, all filters used for smoothening purpose are effective at removing noise in smooth patches or smooth regions of an image, but adversely affect edges. In the process of reducing the noise in an image, it is also important to preserve the edges because it plays a crucial factor in the visual appearance of images.

The median filter is better than using Gaussian filter at removing noise while preserving edges ranging from small to moderate levels of noises. However, its performance is not that much better than Gaussian filter for high levels of noises, whereas, for small levels of noise like for example salt and pepper noise, it is particularly effective. Because of this, median filtering is widely used in stereo image processing. Figure 3-12 shows the median filtered disparity map after removing noise from the image.

Figure 3-12 Median Filtered Disparity Map.

3.4 Chapter Summary

In this chapter we had discussed about the depth and disparity maps of stereoscopic image pairs. Some experiments are performed by moving the stereo camera rig at different positions in order to visualize the depth between some objects. Also, a 3D image having parallax effects are considered which has significant role in creating the illusion effect in the making of 3D movies and animations. Some factors like noise affecting the disparity map are observed and also the importance of using median filter to remove the noise in comparison to Gaussian filter is mentioned. This prior works had discussed about the pre-processing step of the visualization of images before entering into the next phase of the image processing methods.

4.0 Horizontal Disparity Visualization

This chapter illustrates about the current research works performed in order to visualize the horizontal disparity between two stereo image pairs. Several researchers had already made contribution in detecting the difference between stereo images and acquired the information related with depth and disparity visualization. This section distinctly elaborates the difference between the current work and all the existing work on disparity visualization of stereoscopic images. Some existing algorithms has been studied deeply and on further modification being performed in those algorithms, a new approach of visualizing the horizontal disparity in more robust way is proposed and implemented. Lastly, this chapter discusses about the comparison between all methods and techniques to show the authenticity of the work at end of each section. Finally, this chapter ends with the brief chapter summary.

4.1 Texture Extraction Algorithm

It can be used in the area of stereo image processing which involves using algorithms to detect various desired portions or shapes (features) of a digitized image. One of the important characteristics of texture extraction method is edge detection. Edge detection is a fundamental tool in image processing, computer vision and pattern recognition, particularly in the areas of feature detection and feature extraction, which aim at identifying points in a digital image or pointing discontinuities. The result of applying an edge detector to stereo image pairs may lead to a set of connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in an image. Sometimes, applying an edge detection algorithm to an image may significantly reduce the amount of data to be processed and may therefore filter out information that may be regarded as less relevant, while preserving the important structural properties of an image. Edge detection is a fundamental tool used in most image processing applications to obtain information from the frames of an image as a precursor step to feature extraction and object segmentation. This process detects outlines of an object and boundaries between objects and the background in the image. There are many filters to perform edge detection.

4.1.1 Edge Detection by 'Gabor' Filter

The differences in the position of two images is foremost the basic step for 3D reconstruction of the images from two-dimensional projections. This horizontal shift between left and right images taken from stereoscopic cameras creates a horizontal disparity. Usually, the disparity estimation is classified into either 'area-based' or 'feature-based' methods. Considering all the previous disparity estimation techniques, mostly area-based methods are used. According to this method, at first a rectangular area of one image is chosen and by using correlation method, matching is done between the selected rectangular area and all the corresponding positions of the other image. The best matching correlation found is then used to calculate the disparity [10]. Disparity is computed when the filter is totally inside the image area. For this purpose, 'Gabor' filter is regarded as a useful filter for binocular disparity estimation [11]. While considering ‘feature-based’ method also, ‘Gabor’ filters is regarded as most appropriate for texture extraction, representation as well as discrimination between two images [12]. One of the different approaches for disparity estimation is by combining phase based and energy based algorithm using ‘Gabor’ filter by Fourier transform method [13]. In our research work, texture extraction method using ‘Gabor’ filter involves following steps:

(1) The left camera color image and right camera color image is converted into grayscale.

(2) ‘Gabor’ filter is applied on both grayscale images.

(3) The horizontal disparity is visualized between two stereo image pairs as shown in figure 4-1 (back portion of head is taken in both images).

Figure 4-1 Disparity Visualization Using ‘Gabor’ Filter.

4.1.2 Edge Detection by ‘Canny’ Filter

The main motive of using ‘Canny’ edge detection is to derive an ‘optimal’ operator for detecting the edges and finding their accurate localization. The accuracy of detecting the edges is found to be inversely proportional to their localization. The more accurate the detector is, the less accurate the localization [14]. Canny et al. [15] proposed a computational theory of optimal edge detection algorithm which originally implies good detection, better localization of the edges and also achieved to get minimal response of the edges.

Feature based correlation method using ‘Canny’ filter is used almost similar to steps of ‘Gabor’ filter. The major differences between both filters are:

(1) In case of ‘Canny’ filter, direct conversion of the color stereo image pairs into color edge detection without converting them into grayscale is possible [16].

(2) By the use of this filter, it is easy to visualize all the corresponding edges or points to estimate the clear disparity as compared to edge detection by using ‘Gabor’ filter.

In figure 4-2, the left camera color image and right camera color image is directly converted into color edges by using ‘Canny’ filter. Here, back portion of vessel in both the images is taken in order to show the horizontal disparity between the stereo image pairs.

Figure 4-2 Disparity Visualization Using ‘Canny’ Filter.

4.1.3 Comparison between ‘Gabor’ and ‘Canny’ Filters

One major advantage of using ‘Canny’ filter for estimating the disparity is that it greatly reduces the computation time. The reason is that the conversion of the color images which is being obtained by stereoscopic cameras into grayscale images is not required. Another advantage is that efficiency as well as clarity for recognizing the edges by ‘Canny’ filter is increased as compared to detecting edges through ‘Gabor’ filter.

4.2 Stereo Matching Algorithm

The correspondence problem in stereo matching can be broadly classified into two categories, the intensity-based matching and the feature-based matching. In intensity based, the matching process is applied directly to the intensity of the two images, while in feature based, the features are first extracted from the images and then the matching process is performed. Edge elements, corners and line segments are those features which are useful related with the change of view of the images, and also they have been widely used in binocular stereo vision work.

For the visualization of horizontal disparity between two stereo image pairs, the process of comparing left camera image and right camera image is popularly known as 'Stereo Matching'. This method is used in the field of stereo vision for many years and it has large contribution in generating depth and disparity maps.

4.2.1 Correlation Method

Image correlation is a kind of method that employs tracking and adjustment of images for accurate 2D and 3D measurements of changes in images. It is one of the popular approaches for estimating the horizontal disparity. Usually, by correlation method, epipolar lines or constraints are considered on both images to visualize the horizontal disparity. Epipolar constraints play a crucial role in stereo matching techniques. A pair of stereo images is epipolar or raster scan aligned as the first step in location of image feature points. The image point in the right image which gives the maximum correlation value is chosen as the corresponding point for the associated point in the left image [17]. In parallel-axis stereoscopic camera, these lines are always considered as parallel lines. It is the effect of epipolar constraint that disparity estimation in horizontal directions is more viable in comparison to the vertical directions [18]. Usually, by correlation method the disparity is estimated taking the windowed regions of left and right images instead of taking whole image [19]. The windowed region in right image is taken on epipolar line constraint and then the disparity is visualized with the windowed region in left image as shown in figure 4-3. At this moment, one limitation arises that the size of the window must be enough large for adequate stereo-matching. For some images, correlation based stereo matching can be easily visualized but for few other images the limitation of using correlation method brings drawback for proper disparity visualization.

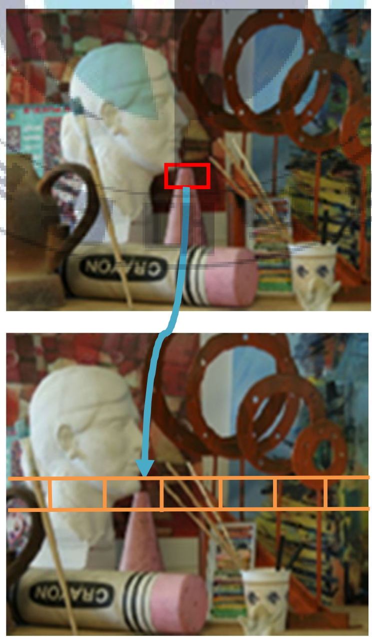


Figure 4-3 Disparity Visualization Using Standard Correlation Method.

4.2.2 QSPDE (Proposed Method)

Image segmentation is always a challenging task in image processing, pattern recognition as well as in 3D reconstruction. In stereo vision, image segmentation plays a key role in partitioning a stereo image pairs into multiple segments. The main purpose is to partition a digital image into smaller segments (set of pixels) which are much easier to analyze. Image segmentation is the process of assigning pixels of an image to regions which have common properties [20]. Normally, based on these classifications, the two digital images or stereo image pairs are compared. Therefore, iteration process can also be performed easily by dividing an image into M*M blocks using quad-tree decomposition [21]. Quadtree image segmentation is basically used for describing the disparity map of fixed block size image [22].

After deep analysis of all existing methods, it is observed that for some stereo images, correlation based stereo-matching is used for disparity visualization but in case if the horizontal shift between the two images is not clearly visible at every corresponding points or edges, then it brings the limitation of using correlation method. In figure 4-4, the two images at the top are from left camera and two images at bottom are from right camera. They have least disparity between them which are enough capable to make 3D images but horizontal shifting between left camera and right camera images are not clearly visible.

Hence, to estimate the horizontal disparity between any stereo image pairs in more robust way, a method is proposed which is named as "Quadtree Segmentation of Pixels Disparity Estimation (QSPDE)" [23].

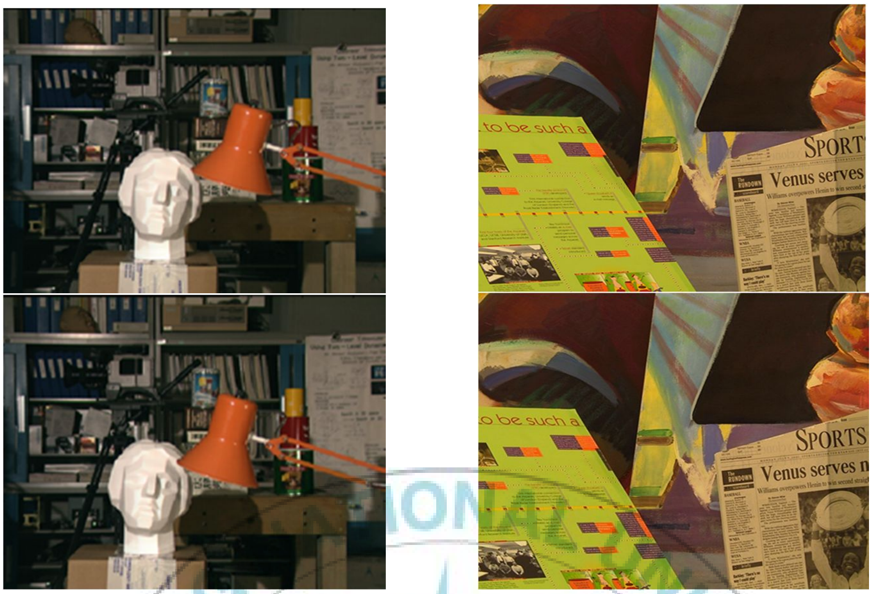


Figure 4-4 Least Disparity Stereo Image Pairs.

Using QSPDE method, both left camera image and right camera image is divided into four equal segments. Each segment is further divided four another smaller size segments or better called as pixels in similar way as shown in figure 4-5.

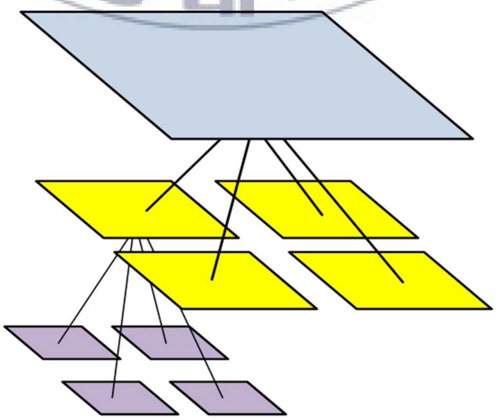


Figure 4-5 Quadtree Segmentation of an Image.

Now, using proposed QSPDE method, it is free to choose any one pixel of the left camera image and compare it with the corresponding pixel of right camera image in order to estimate the disparity.

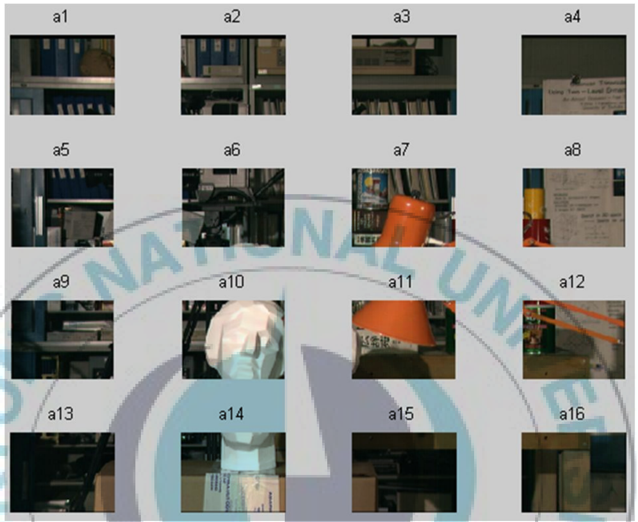


Figure 4-6 Quadtree Segmentation of Left Camera Image.

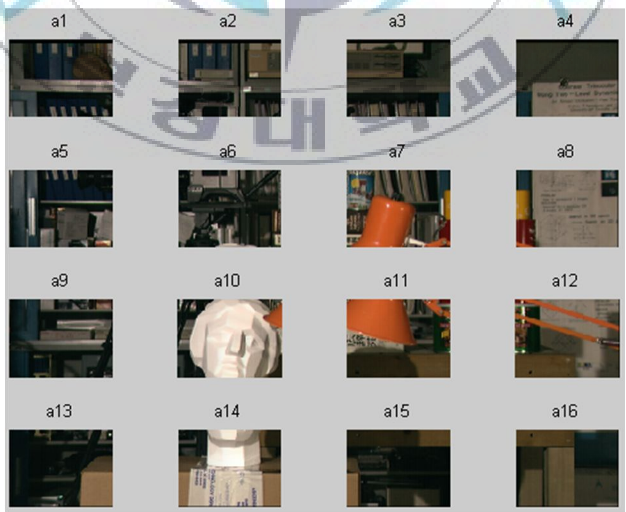


Figure 4-7 Quadtree Segmentation of Right Camera Image.

Figure 4-6 and 4-7 shows the quadtree segmentation of left camera and right camera respectively obtained by using MATLAB 7.4.0. Now, let us choose some pixels like ‘a7’, ‘a10’ and ‘a14’ of left camera image and compare it with same pixels of right camera image respectively. It is observed that the visualization of the horizontal disparity between both images becomes easier in comparison to the feature based stereo-matching methods.

4.2.3 Comparison between Correlation and QSPDE Methods

QSPDE method is basically based on the concept of correlation method of stereo-matching but not limited to choose any fixed particular window or block size of image and make comparison with the same window or block size in other image as similar in correlation method . QSPDE method is more useful to visualize the horizontal disparity of any corresponding segments or pixels than disparity visualization by correlation method and also the limitation of using a fixed size window and comparing with the other image containing the epipolar lines or constraints is removed. Table 4-1 shows the comparison between texture extraction, correlation and QSPDE methods. Their merits and demerits highlight the authenticity of using QSPDE method more than the other existing methods.

Table 4-1 Comparison between All Three Methods

Points	Correlation Method	Texture Extraction	QSPDE (Proposed Method)
Merits	Clear disparity for few stereo pairs	Disparity of particular area by edge detection	Easy disparity visualization for any stereo pairs
Demerits	Limitation is to choose a particular area or region	Difficulty in disparity estimation for any areas or regions	No demerits

4.3 Image Fusion Algorithm

In image fusion technology, the images which combine together are referred as 'source' images and the resultant combined image is known as 'fused' image [24]. There are many techniques which describes suitable image fusion processes. Usually, some of those techniques are based on 'Statistics', 'Wavelet', 'Scene Content' and 'Image Overlay' [25].

Among all four techniques, image overlay is considered as one of the method whose operational characteristics totally depend on the quality of stereoscopic image pairs. Also, the fused image and all relevant information contained in the input images should be maintained without affecting the quality of the images. The fusion process of the two images should not introduce any artifacts or incompatibility, which can make any trouble, give false or misleading information to the human observer, or in any succeeding image processing steps [26].

The goal of the research is that with the help of image fusion technique, a new image can be created which is more suitable for visualizing the horizontal disparity in comparison to the feature extraction using edge detection methods. The new fused image as shown in figure 4-8, provides more comprehensive information regarding visualization of the disparity rather than taking both source images separately.

Figure 4-8 Single Fused Image.

4.3.1 Uni-Directional Horizontal Disparity

Feature based stereo matching method using edge detection is taken into account so as to compare between the ways of visualizing the disparity either taking both images separately using 'Canny' filter or perceive horizontal disparity from a single fused image. There is one limitation which arises by using 'Canny' filter is that it is inevitable to take both stereo image pairs together in order to visualize the disparity. In addition, either left camera image is considered first to compare with right camera image or vice-versa, the disparity can be visualized in one direction at one time. This method of visualizing the horizontal disparity is called as 'Uni-Directional' as shown in figure 4-9.

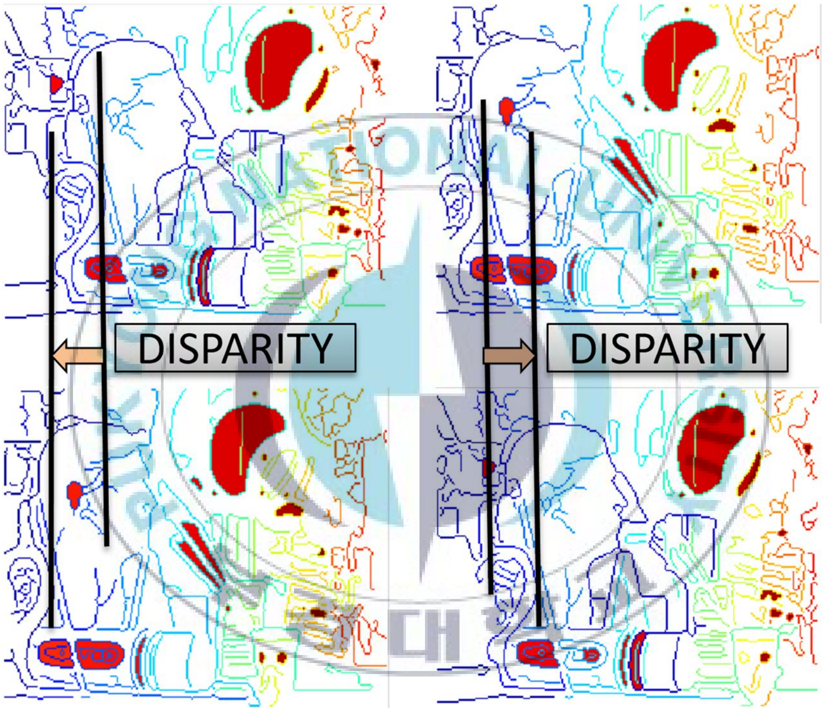


Figure 4-9 Uni-Directional Horizontal Disparity Visualization.

By 'Image Fusion' method also, the disparity between two images can be visualized in two opposite directions individually, but the difference between the previous and the current method is that in case of image-fusion, a single fused image is enough to estimate the disparity in two directions without taking both images separately.

4.3.2 Bi-Directional Horizontal Disparity

In an image fusion technology, the basic foremost step is to adjust spatial alignment of the source images and then to overlap the two images to form a single fused image. This method has several advantages. Firstly, it continues to provide the horizontal disparity on a single synthesized image rather than using two separate source images [27]. Secondly, by increasing or decreasing the threshold intensity of the overlapped images, horizontal disparity can be visualized in two opposite directions taking any one image as reference and then comparing with another on the same fused image. Also, a single threshold is much effective for edge detection method based on image fusion [28]. In figure 4-10, it is observed that on single synthesized image obtained from 'Image' Fusion' method, the horizontal disparity can be visualized in both directions.

Finally, the fused image greatly helps to estimate the disparity with much less computation time as compared to visualization of horizontal disparity by edge detection method using 'Canny' filter. Consequently, it can be easily noticed that by using image fusion method, the two source images can also be obtained on the same fused image by adjusting the threshold intensity of the two stereo image pairs. This process of visualizing the horizontal disparity on single image is termed as 'Bi-Directional'.

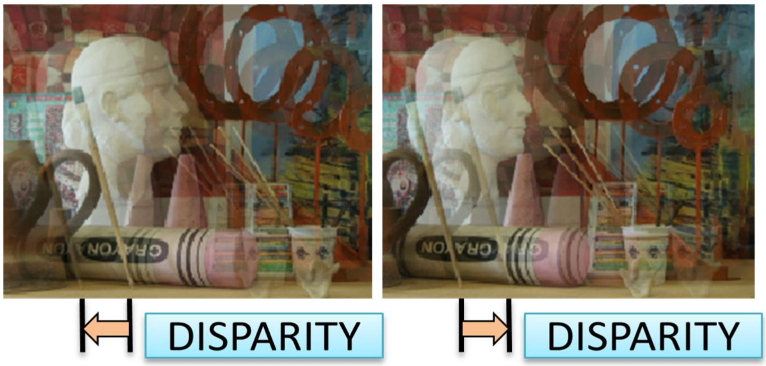


Figure 4-10 Bi-Directional Horizontal Disparity on Single Fused Image.

If either of left camera image or right camera image is taken as reference, then by decreasing the threshold intensity of other image, the horizontal disparity can be visualized. Thereafter, if threshold intensity is more reduced then that image gets hidden and only the reference image is visible in the picture. So, both left camera image as well as right camera image can also be obtained separately by adjusting the proper threshold intensity on single fused image.

One of the major advantages of using image fusion method is that both stereo image pairs need not required using any filter for distinguishing the horizontal disparity between them. By using image fusion technology, the color images obtained from left camera and right camera are overlapped to form a color fused image. Also, it reduces the long process for visualizing the horizontal disparity by firstly converting the original images into some another form with the help of any filters. As a result, image fusion technique is considered as more robust way for visualization of the disparity with the help of single synthesized image. One important point to be taken into consideration is that the fused image is the result of overlying left camera image over right camera image. So, fused image is different from a 3D image. In the process of making 3D image, left camera image is imbricated on right camera image to see the result with the help of 3D polarised glasses.

4.3.3 Grayscale Representation

The threshold value as used in image fusion method can also be used for preserving a large number of features for edge detection of black and white images [29]. This is helpful in visualizing the horizontal disparity between stereo image pairs after converting color images into gray images. Image binarization is one of the technique in image processing which is used to binarize a grayscale image such that some pixels of an image will represent as 'foreground' in gray while other pixels represent 'background' region to be converted into white. This method is used in stereoscopic imaging technology for visualizing the horizontal disparity by converting the color stereo images into grayscale images. The foremost step is to combine left camera image and right camera image into a single synthesized image based on image fusion method. Now, image binarization concept is used to convert right camera color image into grayscale image which will be higher than the fixed threshold and left camera color image is converted into white image which is below than the fixed threshold on a single fused image as shown in figure 4-11 [30]. Now, horizontal disparity between the two images can be easily visualized for any least disparity stereo image pairs. The objective of image fusion is to combine two stereo image pairs in order to generate a single fused image that contains a more accurate description of the scene than any of the individual image for better understanding of the human visual system [31].

Figure 4-11 Grayscale Representation for Disparity Visualization.

4.4 Chapter Summary

This chapter illustrates the core of our research work comprising of texture extraction algorithm, stereo matching algorithm and image fusion algorithm which aim to get better horizontal disparity visualization of stereo image pairs. A new method for visualizing the horizontal disparity is proposed which is called as QSPDE method. This proposed method gives a new pinnacle in the arena of all existing stereo matching techniques. At the end of each section comparison between existing method and current research work is mentioned to get clear idea of this chapter.

5.0 Image Quality Measurements

'Image Quality' is the general term used to refer the degree of visibility of relevant information in an image. Stereo image quality is a characteristic of an image that measures the perceived stereo image degradation when compared with an original image. Imaging systems may introduce some amounts of distortion or artifacts in the images considering some of the unavoidable environmental conditions.

In this part of the research, we have two original images taken from left camera and right camera respectively. Some of the noise factors like 'salt & pepper' and 'Gaussian' are taken into consideration as the distorted stereoscopic image pairs. These distorted images are compared with original left camera image and right camera image seperately to finally get the image quality measurements like MSE, PSNR, Normalized CC and Maximum Difference.

5.1 MSE (Mean Square Error)

MSE measures the average of the squares of the "errors." The error is the amount by which the value entailed by the estimator differs from the quantity to be estimated. The distorted stereo image pairs affected by 'Gaussian' noise with different standard deviations and also affected by 'salt & pepper' noise is compared with the original stereo image pairs to obtain MSE as shown in equation (1).

$$MSE = \frac{1}{MN} \sum_{j=1}^{M} \sum_{k=1}^{N} \left(x_{j,k} - x'_{j,k}\right)^2 \tag{1}$$

where,

M, N = Height and width of an image

$x_{j,k}$ = Original image with pixel coordinates j, k

$x'_{j,k}$ = Distorted image with pixel coordinates j, k

5.2 PSNR (Peak Signal Noise Ratio)

The term peak signal-to-noise ratio, abbreviated as PSNR, determines the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of the stereo images. Also, the PSNR is used as a measure of quality of 3D reconstruction of images. The signal in this case is the original image, and the noise is the error which affects both stereo image pairs. It is most easily defined via the mean squared error (MSE) as shown in equation (2).

$$PSNR = 10 \log \frac{(2^n - 1)^2}{MSE} = 10 \log \frac{255^2}{MSE} \tag{2}$$

In our research work, the two stereo images have horizontal shifting between the two images. So, due to disparity, both stereo image pairs have some calculative MSE as well as PSNR values. But, in case if we consider two images to be identical, then MSE will be zero and PSNR will be undefined.

5.3 Normalized Cross Correlation

The correlation between two signals (cross correlation) is a significant approach to feature detection of an image [32].

Normalized cross correlation is one of the methods which in general, used for template matching. It is a process which is used for finding incidences of a pattern or object within an image.

For stereo image-processing applications in which the texture extraction of an image can vary due to some environmental conditions, the images need to be normalized first. This is shown in equation (3) where the summation of the product of original image and distorted image is being divided by the square of the summations of two original images to finally get normalized cross correlation which is represented as NK.

$$NK = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N} x_{j,k} \times x'_{j,k}}{\sum_{j=1}^{M}\sum_{k=1}^{N} x^2{}_{j,k}} \quad (3)$$

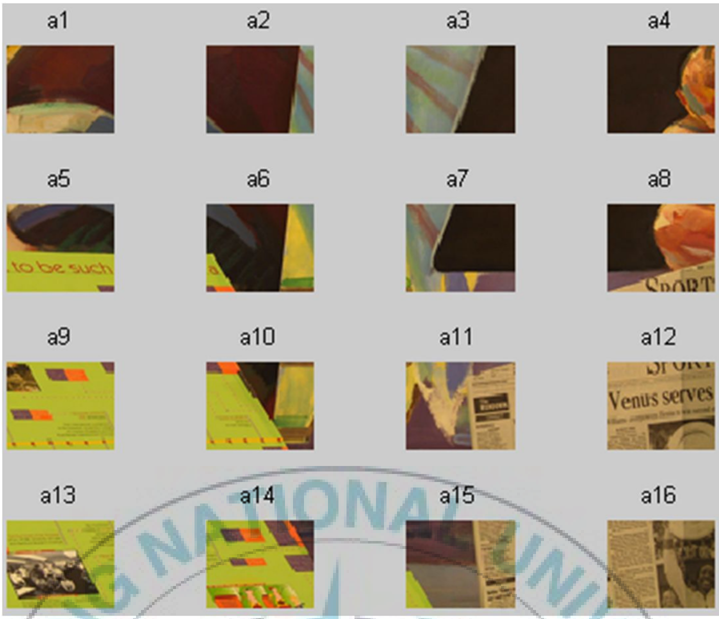
5.4 Maximum Difference

It represents the difference between the original stereo image and distorted image. Here, it can be analyzed that the horizontal disparity between the two images causes the difference between the two stereo image pairs when compared with the distorted images. Larger disparity brings more difference between original and distorted image and lesser disparity has less differences. From equation (4), it is observed that the modulus of difference between original image and the distorted image is calculated to get the maximum difference which is represented as MD.

$$MD = Max\left(\left|x_{j,k} - x'_{j,k}\right|\right) \quad (4)$$

5.5 Pixel's Disparity Estimation

The difference in the amount of shifting between the left camera image and right camera image is crucial factor in the visualization of stereo image pairs. Many research works describes about the visualization of horizontal disparity between two images by various methods like texture extraction, correlation, stereo matching, etc. Mostly in all methods, the visualization can be easily possible by either taking fixed window size in both images, or to divide an image into smaller segments based on quadtree segmentation method. So, for obtaining the calculative result on the amount of shifting, the pixel's disparity estimation method is robust way in order to estimate the horizontal disparity. One pixel in an image is chosen and compared with the pixel of another image of the same object or scene and thus horizontal disparity can be observed. For example, in QSPDE if an image is considered with least disparity as shown in figure 5-1, then the amount of horizontal shifting can be calculated by subtracting pixel's coordinates of left camera image with pixel's coordinates of right camera image. Finally, the modulus of the result is taken to estimate the pixel's disparity as shown in table 5-1.



(a)

(b)

Figure 5-1 Quadtree Segmentation of Least Disparity Image: (a) Left Camera Image (b) Right Camera Image.

In figure 5-1, it is free to choose any pixel between 'a1' to 'a16' and compare it with same pixel in right camera image and easily visualize the disparity. Let us consider coordinates of 'a8', 'a10' and 'a12'pixels in left camera image (a) and subtract with coordinates of same pixel's in right camera image (b). Hence the pixel's disparity estimation between two images is shown in table 5-1.

Table 5-1 Pixel's Disparity Estimation between Stereo Image Pairs

Pixel's Number	Left Camera Image(L)	Right Camera Image(R)	Pixel's Disparity |L-R|
'a8'	363	375	12
'a10'	169	181	12
'a12'	332	344	12

5.6 Chapter Summary

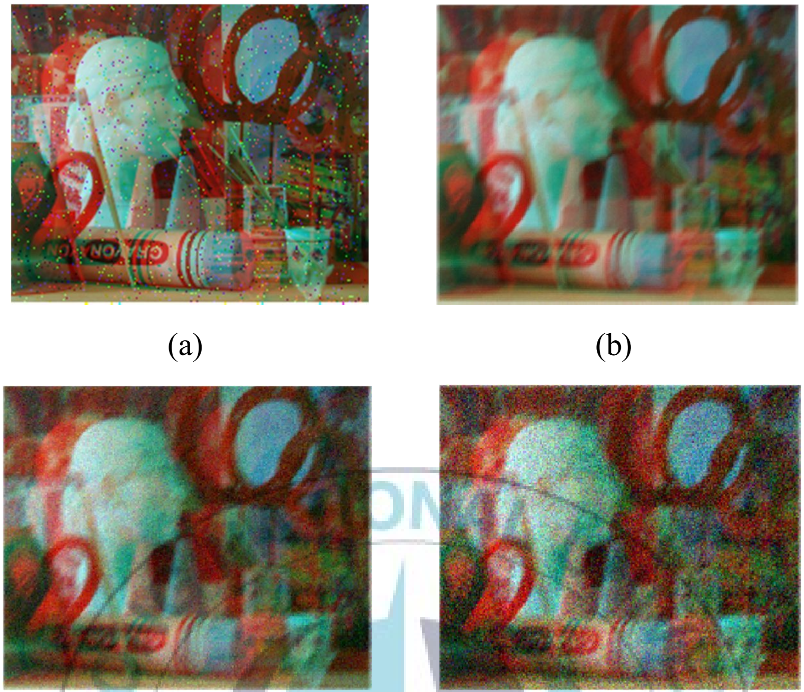
The conclusion of this chapter is that in all existing methods of image quality measurements, only one original image have been used to compare with the distorted image. The difference in our research work is that two stereo image pairs are considered as the original images and both images seperately compared with the distorted images in order to obtain the image quality measurements. Therefore, the analysis graphs of the experiments are presented in the next chapter.

6.0 Experimental Results & Discussions

In this chapter, few experiments are conducted to verify the performance of our thesis work in terms of analyzing the stereoscopic image pairs and making their comparison with the distorted images. Also, the maximum difference between left camera image with the distorted images and difference between right camera images with the same distorted images are obtained. Finally, the results will be shown in details in the form of analysis graphs.

6.1 Performance of Stereoscopic Image Quality Measurements

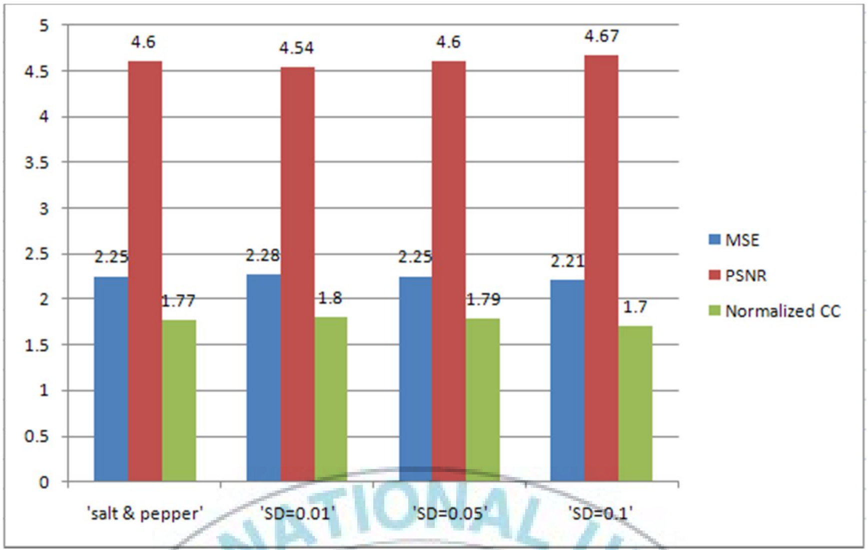
Firstly, the stereoscopic image pairs are combined together to form a 3D anaglyph image. Then, the distorted form of 3D image is obtained taking some environmental conditions like ‘salt & pepper noise’ and ‘Gaussian noise’ by image processing in MATLAB 7.4.0 as shown in figure 6-1. The two original images from left camera and right camera are compared with distorted images seperately to finally obtain the analysis graph of image quality measurements. Different horizontal disparity images of larger shifting as well as smaller shifting between the image pairs are considered to perform the experiment and get the corresponding quality measurement graphs.



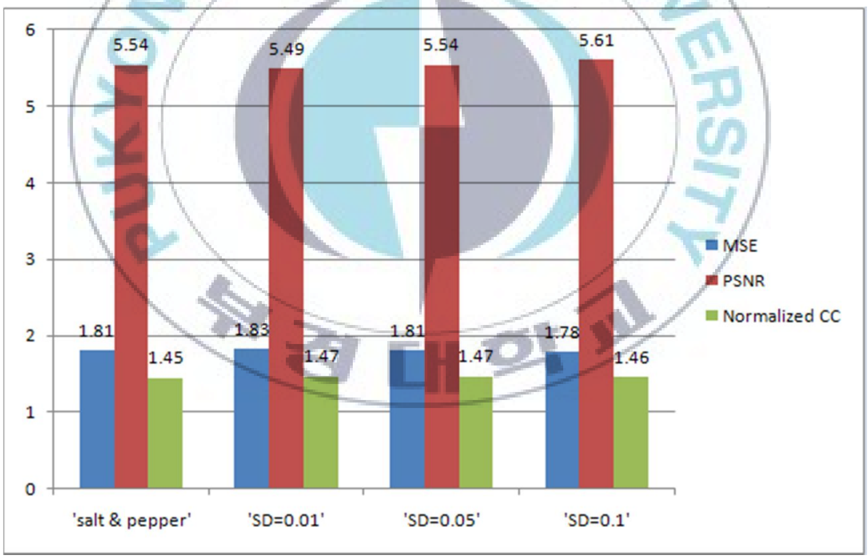
(a) (b) (c) (d)

Figure 6-1 Distorted 3D Images: (a) Salt & Pepper Noise. (b), (c) and (d) Gaussian Noise of SD=0.01, 0.05 and 0.1 respectively.

Now, all these distorted images are compared with two original left camera and right camera image seperately to obtain image quality measurements. Finally, their analysis graph is shown in figure 6-2, where SD= Standard Deviation values of ‘Gaussian Noise’.



(a)



(b)

Figure 6-2 Image Quality Measurements: (a) Vessel Left Image. (b) Vessel Right Image.

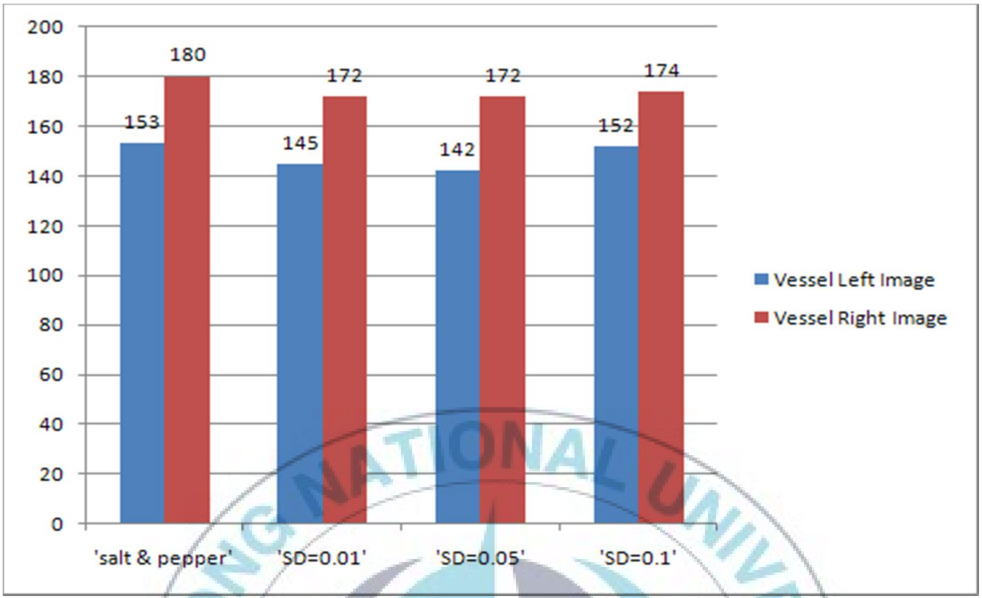
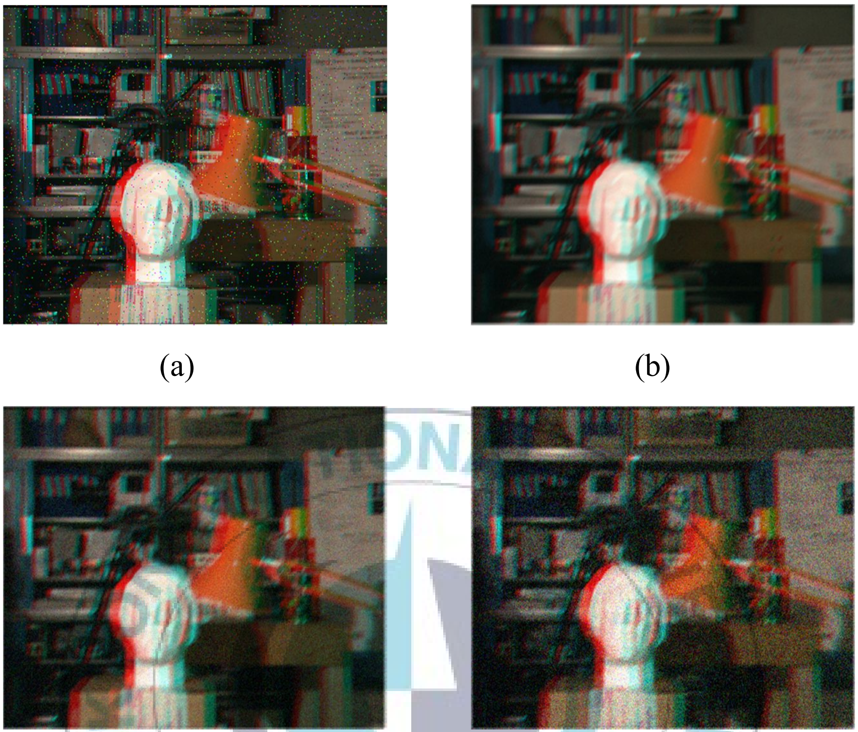


Figure 6-3 Maximum Difference between Vessel's Left and Right Images.

Here, figure 6-3 shows the maximum difference between left camera image and right camera image on comparing with all distorted images.

All graphs were showing the image quality measurements of vessel's image pairs which are having greater disparity. Now, another stereo pairs of 'Tsukuba' images which have lesser disparity as compared to vessel's images are taken for evaluating the different image quality measurements.

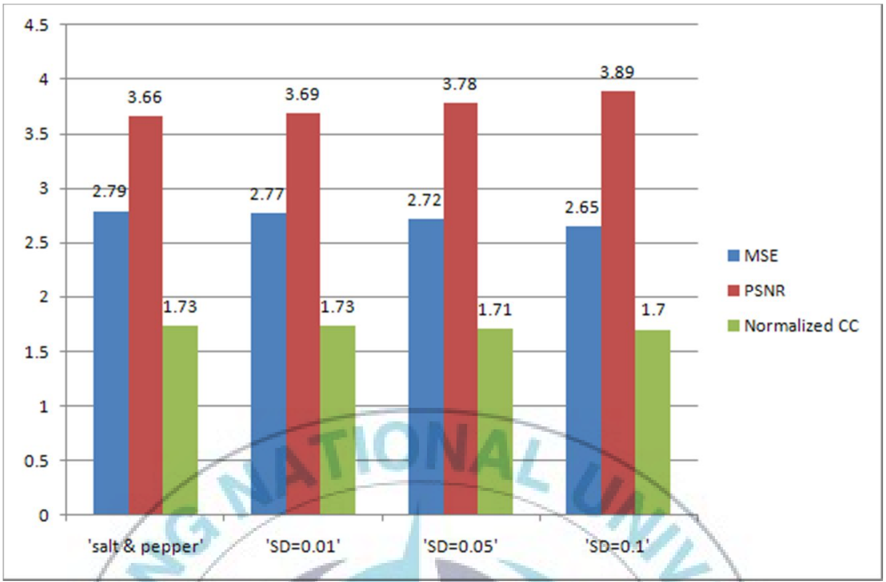
Figure 6-4 shows the 'Salt & pepper noise' and 'Gaussian noise' with different standard deviation values affecting the left and right of 'Tsukuba' images.



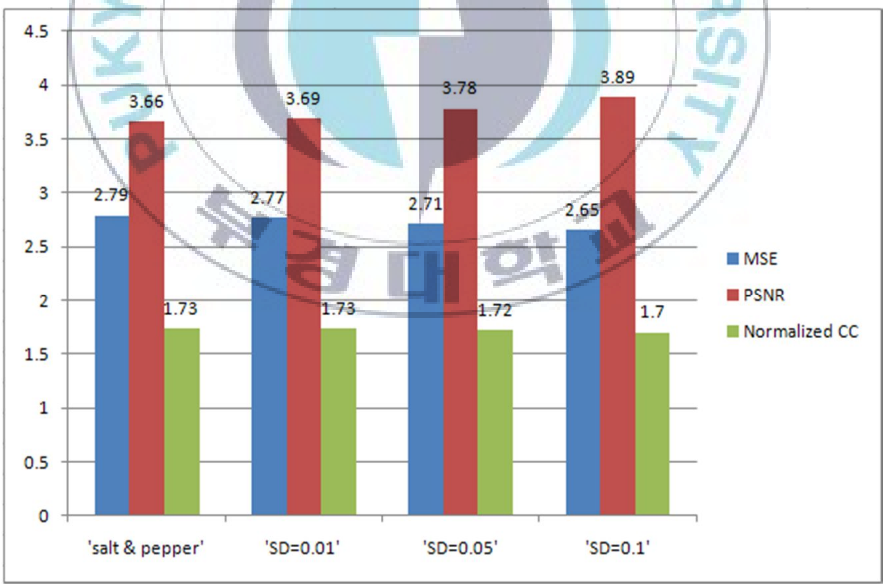
(a) (b)

(c) (d)

Figure 6-4 Distorted 'Tsukuba' Images: (a) Salt & Pepper Noise. (b), (c) and (d) Gaussian Noise of SD = 0.01, 0.05 and 0.1 respectively.



(a)



(b)

Figure 6-5 Image Quality Measurements: (a) 'Tsukuba' Left Image (b) 'Tsukuba' Right Image.

From figure 6-5, it is clear that since the disparity between 'Tsukuba' left image and 'Tsukuba' right image is very less so all image quality measurements values are almost similar in both the analysis graphs.

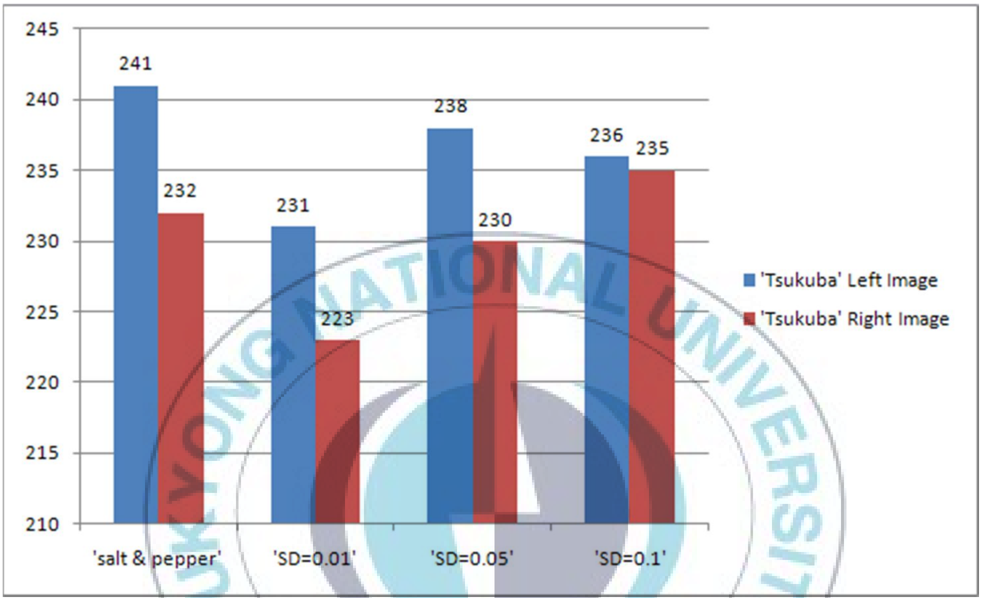


Figure 6-6 Maximum Difference between 'Tsukuba' Left and Right Images.

6.2 Chapter Summary

The effect of distortions in the process of making comparison with the original stereoscopic image pairs is discussed in this chapter. Different horizontal disparity images are considered to perform the experiment by image processing techniques and finally the stereoscopic image quality measurements are obtained with the analysis graphs.

7.0 Conclusion and Future Works

The thesis studied the stereoscopic imaging technology for depth and disparity visualization of 3D camera images. The concept behind formation of a 3D image is discussed by considering two digital cameras mounted on a movable horizontal platform called as 'Stereoscopic Camera Rig'. Later, depth and disparity between two stereo image pairs is visualized and several algorithms related with it are discussed.

7.1 Overall Conclusions

In most of the 3D imaging applications, edge detection method act as a preprocessing stage for feature extraction. The extraction of features such as edges and curves from an image is useful for many purposes like texture analysis, 3D reconstruction of surfaces, image segmentation and stereo-matching of the images.

In this research, depth and horizontal disparity estimation between the two images is presented with the brief description of stereoscopic camera rig model. Stereoscopic camera rig system is an important hardware component in order to develop a clear 3D image. Disparity estimation is being performed by using 'Gabor' filter through texture extraction method. The disparity visualization technique by using 'Gabor' filter brings few drawbacks like for each corresponding edges or regions we cannot able to estimate the disparity between the images effectively. This problem is being sorted out by using 'Canny' edge detection method.

Further it is being analyzed that the visualization of horizontal disparity using ‘Gabor’ and ‘Canny’ filters plays crucial role for the determination of the geometrical differences between two stereo image pairs of the same scene taken at slightly different positions. Sometimes, the use of ‘Gabor’ and ‘Canny’ filters has got some limitations in visualization when it is applied on stereo image pairs of minimum disparity. It is observed that the stereo images which have least disparity are difficult to estimate the clear horizontal shifting between the two images when ‘Gabor’ and ‘Canny’ filters are applied on those images.

Till now, correlation based stereo-matching method has been used in which a selected area of an image is taken in both the images to distinguish the disparity. Since, correlation method is very time consuming operation and is only valid for few stereo images, so it is inevitable to simplify the visualization algorithm without affecting the quality of the stereo images. For this purpose, segment based stereo-matching are widely used. A popular approach is to partition the image into smaller segments and then block matching algorithm is performed to visualize the disparity.

In our research work, a novel method for estimating the disparity is proposed by dividing an image of any size into smaller pixels. This method is named as “Quadtree Segmentation of Pixels Disparity Estimation (QSPDE)”. QSPDE is relevant to any stereo image pairs and better way for distinguishing the horizontal disparity with much more clarity and higher effectiveness.

In addition, many researches on the multi-source image fusion are going on in stereoscopic imaging technology. The method of combining two separate images into a single fused image improves the level of performance in visual perception as well as provides ample information related to the visualization of horizontal disparity. Our present research brings a novel method for visualizing the horizontal disparity in bi-directional ways between two stereo images in form of single fused image and also reduces the computation time in comparison to distinguishing the disparity between both images seperately.

7.2 Future Direction

3D technology is an emerging field in the domain of stereoscopic imaging used in the formation of 3D movies, animations and graphics. Based on the contribution of this thesis, there remains an immense option for future works in this area. At present, the main purpose in this project is to develop an automatic control module of the distance between two cameras in same or opposite directions on horizontal camera rig system.

References

[1] G. Lee, N. Hur and C. Ahn, “Parallel-Axis Stereoscopic Camera with Vergence Control and Multiplexing Functions”, Stereoscopic Displays and Virtual Reality Systems, X, 5006, 2003.

[2] R. A. Lane and N. A. Thacker, “Tutorial: Overview of Stereo Matching Research”, [online] Available: www.tina-vision.net. Tina Memo No. 1994-001.

[3] D. Coombs and C. Brown, “Cooperative Gaze Holding in Binocular Vision”, IEEE Control System, pp. 24-33, 1991.

[4] J. J. L. JR., “A Discussion Cybersickness in Virtual Environments”, SIGCHI Bulletin, Vol. 32, no.1, pp. 47-56, January 2000.

[5] S. Y. Park and N. Lee, “Stereoscopic Imaging Camera with Simultaneous Vergence and Focus Control”, Society of Photo-Optical Instrumentation Engineers, 43(12), pp. 3130-3137, 2004.

[6] C. Thomson, “Depth Perception in Stereo Computer Vision”, a report submitted to Computer Science Department, Stanford University, Research Report Number STAN-CS-75-521, 1975.

[7] K. C. Kwon, Y. T. Lim and N. Kim, “Vergence Control of Binocular Stereoscopic Camera Using Disparity Information”, Journal of the Optical Society of Korea, Vol. 13, no. 3, pp. 379-385, September 2009.

[8] R. Srinivasan, M. Shridhar and M. Ahmadi, “Computing Surface Information from Monocular and Binocular Cues for Vision Applications”, Proceedings of the 27th IEEE Conference on Decision and Control, Vol. 2, pp. 1085-1089, 1988.

[9] P. Fua, "A Parallel Stereo Algorithm that produces Dense Depth Maps and Preserve Image Features", Inria Research Labs. France. Research Report Number 1369, 1991.

[10] D. Bakin, "Stereoscopy on Complex Scenes and Edge Detection", [online]Available: http://www.okob.net/texts/mydocuments/stereoscopy.

[11] M. H. Ouali, D. Ziou and C. Laurgeau, "Dense Disparity Estimation Using Gabor Filters and Image Derivatives", 2nd International Conference on 3D Digital Imaging and Modeling, pp. 483-489, 1999.

[12] V. S. N. Prasad and J. Domke, "Gabor Filter Visualization", [online] Available: http://www.cs.umd.edu/class/spring2005/cmsc838S/assignment-projects/gabor-filter-visualization/report.pdf.

[13] X. Huang and E. Dubois, "Disparity Estimation for the Intermediate View Interpolation of Stereoscopic Images", IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 2, pp. 881-884, 2005.

[14] M. Ali and D. Clausi, "Using the Canny Edge Detector for Feature Extraction and Enhancement of Remote Sensing Symposium (IGARSS), pp. 2298-2300, 2001.

[15] J. Canny, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-8, no. 6, November 1986.

[16] R. Ramesh, H. S. Shin, S. I. Jeong and W. Y. Chung, "A New Depth and Disparity Visualization Algorithm for Stereoscopic Camera Rig",

International Journal of KIMICS, Vol. 8, no. 6, pp. 645-650, December 2010.

[17] J. E. Dudgeon and V. Lakshminarayanan, “Smoothning Algorithm for 3D Surface Rendering from Stereo Images”, Proceedings of Twenty-Fifth Southeastern Symposium on System Theory, pp. 524-528, 1993.

[18] C. W. Lin, E. Y. Fei and Y. C. Chen, “Hierarchical Disparity Estimation Using Spatial Correlation”, IEEE Transactions on Consumer Electronics, Vol. 152, pp. 639-647, August 1998.

[19] H. Yan and J. G. Liu, “Robust Phase Correlation Based Sub-pixel Disparity Estimation”, 4th SEAS DTC Technical Conference-Edinburgh, 2009.

[20] M. A. Wani and B. G. Batchelor, “Edge-region-based Segmentation of Range Images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, pp. 314-319, 1994.

[21] Y. K. Kim, S. Y. Kim and Y. S. Ho, “Stereo Matching Using Global and Local Segmentation”, ITC-CSCC, pp. 1225-1226, 2007.

[22] D. R. Clewer, L. J. Luo, C. N. Canagarajah, D. R. Bull and M. H. Barton, “Efficient Multiview Image Compression Using Quadtree Disparity Estimation”, The 2001 IEEE International Symposium on Circuits and Systems, Vol. 5, pp. 295-298, 2001.

[23] R. Ramesh, H. S. Shin, S. I. Jeong and W. Y. Chung, “A Novel Horizontal Disparity Estimation Algorithm Using Stereoscopic Camera Rig”, International Journal of KIMICS, Vol. 9, no. 1, pp. 83-88, February 2011.

[24] J. Kim, "Fast texture transfer through the use of wavelet-based image fusion", Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition, Vol. 1, pp. 270-275, November 2007.

[25] K. P. Ramesh, S. Gupta and E. P. Blasch, "Image fusion experiment for information content", 10th International Conference on Information Fusion, pp. 1-8, July 2007.

[26] S. G. Nikolov, "2-D image fusion by multiscale edge graph combination", Proceedings of the Third International Conference on Information Fusion, pp. 16-22, July 2000.

[27] Y. Wan, Y. Cao and L. Dun, "A single threshold edge detection algorithm based on multiscale fusion", 2006 International Conference on Computational Intelligence and Security, Vol. 2, pp. 1710-1715, November 2006.

[28] R. Ramesh, H. S. Shin, S. I. Jeong and W. Y. Chung, "Bi-Directional Horizontal Disparity Visualization Using Stereoscopic Camera Rig", Proceedings of the 1st International Symposium on Automotive and Convergence Engineering, pp. 25-28, January 2011.

[29] J. A. M. Hernandez, F. G. Castaneda, J. A. M. Cadenas, "A method for edge detection in gray level images, based on cellular neural network", 52nd IEEE International Midwest Symposium on Circuits and Systems, pp. 730-733, 2009.

[30] R. Ramesh, S. M. An, S. I. Jeong and W. Y. Chung, "Grayscale Representation of Fused Stereo Image Pairs for Efficient Disparity

Visualization", Proceedings of the IEEK Summer Conference, pp. 702-705, June 2011.

[31] L. Jing, P. Quan, T. Yang, Y. M. Cheng, "Color based grayscale fused image enhancement algorithm for video surveillance", Third International Conference on Image and Graphics, pp. 47-50, 2004.

[32] R. O. Duda, P. E. Hart, Pattern Classification and Scene Analysis, New York: Wiley, 1973.

List of Publications

Journals

[1] **Rohit Ramesh**, Heung-Sub Shin, Shin-Il Jeong and Wan-Young Chung, “A New Depth and Disparity Visualization Algorithms for Stereoscopic Camera Rig,” International Journal of KIMICS, vol. 8, no. 6, pp. 645-650, December 2010.

[2] **Rohit Ramesh,** Heung-Sub Shin, Shin-Il Jeong and Wan-Young Chung, “A Novel Horizontal Disparity Estimation Algorithm Using Stereoscopic Camera Rig,” International Journal of KIMICS, vol. 9, no. 1, pp. 83-88, February 2011.

Conferences

[3] **Rohit Ramesh**, An Seong-Mo, Shin-Il Jeong and Wan-Young Chung, “Grayscale Representation of Fused Stereo Image Pairs for Efficient Disparity Visualization,” Proceedings of the IEEK Summer Conference, Jeju, Korea, pp. 702-705, 24 June-2011.

[4] **Rohit Ramesh**, Heung-Sub Shin, Shin-Il Jeong and Wan-Young Chung, “Bi-Directional Horizontal Disparity Visualization Using Stereoscopic Camera Rig,” Proceedings of the 1st International Symposium on Automotive and Convergence Engineering, Ho Chi Minh City (Vietnam), pp. 25-28, 19 January-2011.

[5] **Rohit Ramesh**, Heung-Sub Shin, Shin-Il Jeong and Wan-Young Chung, “A Novel Stereo-Matching Algorithm Using Horizontal Moving Stereoscopic Camera Rig,” Proceedings of KISPS Fall

Conference at Suwon, Korea, vol. 11, no. 2, pp. 349-352, 27 November 2010.

[6] **Rohit Ramesh**, Heung-Sub Shin, Shin-Il Jeong and Wan-Young Chung, "Depth Perception Using a Parallel-Axis Stereoscopic Camera Rig," Proceedings of the Korea Institute of Maritime & Communication Sciences, Republic of Korea Naval Academy, vol. 14, no. 2, pp. 147-148, 29 October, 2010.

[7] Heung-Sub Shin, **Rohit Ramesh**, and Wan-Young Chung, "The Research of Stereoscopic Camera Rig System for Implementation of 3D Stereoscopic Image," Proceedings of the Korea Institute of Maritime & Communication Sciences, Republic of Korea Naval Academy, vol. 14, no. 2, pp. 156-157, 29 October, 2010.

Awards

"An Excellent Presentation Paper Award" in Proceedings of the Korea Institute of Signal Processing and Systems (KISPS), Suwon, Korea, 27 November 2010.