Durham E-Theses

Metrics for Stereoscopic Image Compression

GORLEY, PAUL,WARD

How to cite:

GORLEY, PAUL,WARD (2012) *Metrics for Stereoscopic Image Compression* , Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/3471/

Metrics for Stereoscopic Image Compression

Paul W. Gorley

A Thesis presented for the degree of

Doctor of Philosophy

Dedicated to

My Father, John Gorley

and in memory of

My Mother, Susan Gorley

Metrics for Stereoscopic Image Compression

Paul W. Gorley

Submitted for the degree of Doctor of Philosophy

2012

Abstract

Metrics for automatically predicting the compression settings for stereoscopic images, to minimize file size, while still maintaining an acceptable level of image quality are investigated. This research evaluates whether symmetric or asymmetric compression produces a better quality of stereoscopic image.

Initially, how Peak Signal to Noise Ratio (PSNR) measures the quality of varyingly compressed stereoscopic image pairs was investigated. Two trials with human subjects, following the ITU-R BT.500-11 Double Stimulus Continuous Quality Scale (DSCQS) were undertaken to measure the quality of symmetric and asymmetric stereoscopic image compression. Computational models of the Human Visual System (HVS) were then investigated and a new stereoscopic image quality metric designed and implemented. The metric point matches regions of high spatial frequency between the left and right views of the stereo pair and accounts for HVS sensitivity to contrast and luminance changes in these regions.

The PSNR results show that symmetric, as opposed to asymmetric stereo image compression, produces significantly better results. The human factors trial suggested that in general, symmetric compression of stereoscopic images should be used.

The new metric, Stereo Band Limited Contrast, has been demonstrated as a better predictor of human image quality preference than PSNR and can be used to predict a perceptual threshold level for stereoscopic image compression. The threshold is the maximum compression that can be applied without the perceived image quality being altered.

Overall, it is concluded that, symmetric, as opposed to asymmetric stereo image encoding, should be used for stereoscopic image compression. As PSNR measures of image quality are correctly criticized for correlating poorly with perceived visual quality, the new HVS based metric was developed. This metric produces a useful threshold to provide a practical starting point to decide the level of compression to use.

Declaration

The work in this thesis is based on research carried out in the Innovative Computing Group, the School of Engineering and Computing Sciences, University of Durham, UK. No part of this thesis has been submitted elsewhere for any other degree or qualification and it all my own work unless referenced to the contrary in the text.

Acknowledgements

I would like to thank all those who supported this work.

First and foremost I would like to thank my Parents, John and the late Susan for their continual support throughout my studies.

In particular thanks to Dr Nicolas Holliman, for supervising my Ph.D. for the endless hours of his time and for giving me the opportunity to under take this research.

I wish to thank my examiners Professor Nishan Canagarajah and Professor Julie Harris, for their time in reading the thesis.

Thanks also to Kodak Corp. for the loan of their 3D Stereo Imaging display equipment and technical discussions regarding these systems. To Professor Alyssa Goodman of the Harvard Initiative in Innovative Computing and the COMPLETE project for the data used to generate the Perseus image, also to Richard Stevens at the National Physical Laboratory for the test objects used in the Mannequin image.

Additionally I would like to thank both the School of Engineering and Computing Sciences, Durham University and the Faculty of Science at Durham University for support of the Durham Visualization Laboratory.

Publications

The research has been documented in part in the following publications,

"Stereoscopic image quality metrics and compression", P.W. Gorley, N.S. Holliman, Stereoscopic Displays and Virtual Reality Systems XIX, Proceedings of SPIE-IS&T Electronic Imaging, SPIE Vol.6803, January 2008.

"Investigating symmetric and asymmetric stereoscopic compression using the PSNR image quality metric", P.W. Gorley and N.S. Holliman, invited paper in Proceedings 44th Annual Conference on Information Sciences and Systems, IEEE Information Theory Society, Princeton USA, March 2010.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In this chapter, the motivation for the compression of stereo images is detailed and then a brief description of the definition of the problem, including an overview of the relevant background topics is given. The objectives, criteria for success and the main contributions of the research are then outlined. Afterwards, the thesis contribution is outlined and the work summarised including defining the symmetric/asymmetric encoding problem and the formulation of a framework for investigating these encoding methods further.

An experiment using human subjects is performed to try and find a stereoscopic image quality threshold and to validate this metric for symmetric compression. A second human based experiment is performed to accurately evaluate symmetric and asymmetric stereo compression techniques. The results of this experiment are compared back to the initial Peak Signal to Noise Ratio (PSNR) trials. The results highlighted problems with the metrics used to investigate these issues and therefore a novel stereoscopic image quality metric based on contrast and luminance changes at matched points within a stereo pair is proposed.

At the end of the chapter a thesis overview and summary is provided.

1.1   Motivation and Problem Definition

Large developments have been made to the efficient representation of visual data over the last few years. Coding methods for visual data, such as images and video have progressed to such an extent that 2D communication systems are available commercially. These systems are based on the Joint Photographic Experts Group (JPEG) [57] and Moving Picture Experts Group (MPEG) [83] standards and have made real-time video conferencing pos-

sible, thus reducing the need for expensive travel. The development of new standards such as MPEG 4 and MPEG 7 have brought different multimedia communication areas together whilst meeting the demands for interaction.

History suggests that the greatest developments in image and video technologies have nearly always occurred by the addition of extra sensations. In the late 1800s, monochrome video was developed which brought an extra sensation and realism to still photographs. Monochrome video quality was later improved by the addition of colour. Most recently the production of Digital and High Definition television has brought enhanced quality and realism. However, even these high resolution, large screen systems, still have limitations when displaying real and natural scenes.

As the human visual system reacts more strongly to 3D than to 2D images [50, 81], adding depth to a visual system is a good way of bringing an extra sensation to the viewer. Three dimensions are perceived by humans from a series of different cues. These include perspective, occlusion and shading. However, for 3D images or video to appear realistic, these cues alone are not enough. The use of stereoscopic methods to produce depth provides an additional cue when viewing 3D images. Stereoscopic displays produce an image for each eye simultaneously. The images are slightly offset and the method is based on the fact that humans perceive 3D by viewing a slightly different image with each eye.

The result of incorporating stereo 3D into visual systems will be a doubling of the data size and thus a likely increase in data for transmission. According to Woo W. [156], the problem of increased data can be solved by,

- Increasing the Channel Bandwidth;

- Improving Channel Utilisation with an efficient protocol; and/or

- Reducing the source itself with efficient compression techniques.

The third of these solutions, to increase the efficiency of compression techniques has received a large amount of consideration in recent years.

This research will initially look at possible approaches to compressing stereo images. The research will investigate the question of how much compression to apply and what approach to adopt when applying compression to these image pairs. The symmetric approach to compressing a stereo image is when both left and right views of the image pair

are compressed with equal amounts. Asymmetric compression is when unequal amounts of compression are applied to the left and right views.

The compression of stereo image pairs, symmetrically and asymmetrically will be compared for a fixed file size. To investigate whether symmetric or asymmetric compression produces a better quality of stereo image from an uncompressed original, currently available monoscopic image quality metrics, such as PSNR, will be used to assess these images. A comparison between symmetrically compressed pairs and those produced from image and depth map encoding will also be undertaken using PSNR. However, results from initial experiments and previous work have shown the PSNR metric has significant problems when evaluating stereoscopic images.

To further investigate symmetric and asymmetric compression, objective human based experiments will be performed. From the results of these experiments, conclusions will be drawn to which compression approach, symmetric or asymmetric produces a better quality of stereoscopic image for a set file size. The results of these experiments will be compared back to the initial PSNR trial.

A new stereo image quality metric will be developed. In line with human visual system properties, this new metric will take into account changes in luminance and contrast within the stereo image pair. The new metric will be used to assess symmetric stereo image quality and to find a stereoscopic image compression threshold, such that stereoscopic images compressed below this threshold will be of an acceptable quality. The results from the human based experiments will be used to evaluate the new metric and confirm this acceptable stereoscopic image quality threshold.

1.2   Stereoscopic 3D

Stereoscopy is defined as any technique that is able to create an illusion of depth within an image or any method that is able to record three-dimensional visual information. To create this illusion of depth within an image or video, a slightly different image is presented to each eye, Figure 1.1. Images seen on the majority of 3D Displays use this method. Stereoscopy was invented in 1840, by Sir Charles Wheatstone. Traditional stereoscopic photographs created a 3D illusion using a pair of 2D images. Depth perception in the brain is created by providing the viewers' eyes with two different images. Each image represents two perspectives of the same object. The difference in the two images is similar to that received in natural binocular vision.
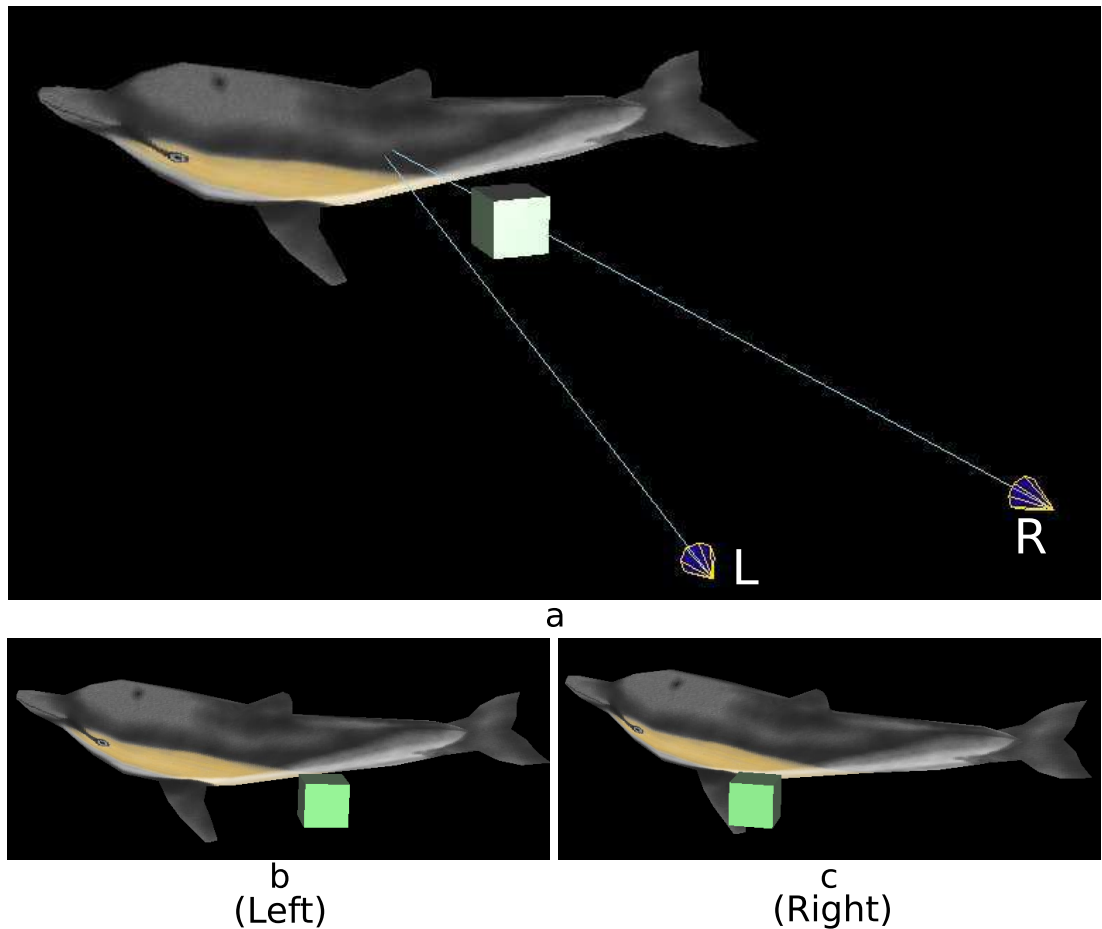
Figure 1.1: Stereo Vision: The two eyes converge on the object of attention (a), The cube is shifted to the right in left eye's image (b), The cube is shifted to the left in the right eye's image (c), [49].

To try and avoid eye strain and distortion, objects within the image that are at infinite distance should be orientated straight ahead, so that the viewers' eyes are neither crossed nor diverging. If no object at infinite distance is contained within the picture, there should be less disparity between the two images.

One of the remaining problems with most stereo displays is focus/fixation mis-match. In normal 2D viewing the screen distance always matches the fixation distance. When using a stereoscopic display, the screen distance is the optical source for all the objects in the screen and the ocular focus distance only matches the fixation distance when the objects are in line with the screen plane. This focus/fixation mismatch is illustrated in Figure 1.2. Chapter 2, Section 2 describes Stereoscopy in more detail.
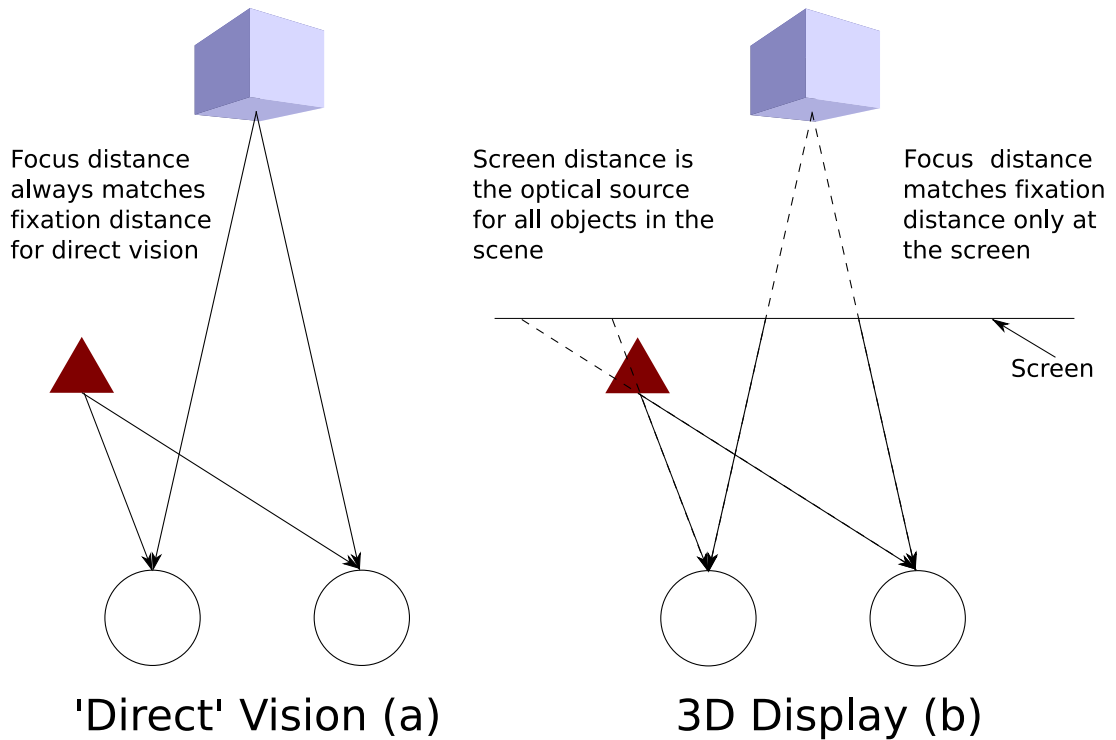
Figure 1.2: Focus/Fixation mis-match: Focus distance matches fixation distance for direct vision (a), Focus distance matches fixation distance only at the 3D display screen (b).

1.3   Stereoscopic Image Compression

There are many different image compression algorithms. These algorithms, primarily designed for monoscopic compression, can be applied to stereoscopic images. Stereoscopic compression can be achieved by either compression across both views simultaneously, or independently compressing the left and right views separately [31].

Compression algorithms can be divided into two distinct categories. They are either lossless or lossy. Lossless algorithms do not change the content of the file, so if you compress and decompress a file it will not be changed. Lossy algorithms, such as JPEG, are able to achieve a better compression ratio by selectively getting rid of some of the information within the file. Due to the better compression ratio, in this research more interest is shown in lossy algorithms, such as JPEG. The question of how much compression to apply to an stereo image pair before the resulting image becomes noticeably lower from the original is investigated. Until now this problem has remained unresolved. Chapter 3 describes Stereoscopic Image Compression in more detail.

1.4 HVS Models and Metrics for Assessing Stereo Image Quality

There are many different image comparison metrics. Wilson et al. [152], classifies these metrics into three different types, however often these different types of metric can overlap. The metric classifications are,

1. Human Perception,

2. Objective measure based on theoretical models, and

3. Subjective measures based on mathematically defined models of the Human Visual System.

Human perception is where a group of participants are asked to view the images and select which image they feel is the best quality from a selection. Objective measure based metrics use mathematical models and numerical value representations of images. The standard way to quantitatively measure image quality is to use the objective measure Peak Signal to Noise Ratio (PSNR). However, these types of metrics are often thought not to correlate well with subjective judgement of image quality. The third type, subjective based image metrics use established models of the human visual system, to assess image quality.

Research into the Human Visual System has produced mathematical models of how humans see the world around them. Figure 1.3 shows how the eye connects to the brain to form the Human Visual System. Stereoscopic images and videos contain large quantities of data. One way to achieve good compression of these images is to discard some of the data. Knowledge of the Human Visual System can allow this data loss to be done selectively so that only data to which the Human Visual System is not sensitive, is discarded. The Human Visual System is most sensitive to changes in contrast and luminance at areas of high spatial frequency. Therefore investigations of the Human Visual System allow metrics to assess the image degradation in these areas. Chapter 4 describes the HVS models and image comparison metrics in more detail.
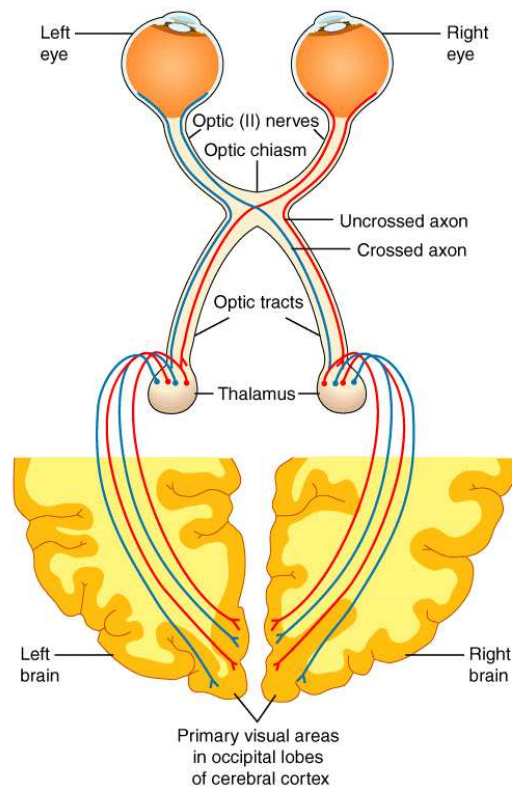
Figure 1.3: Human visual system, interprets the information from visible light to build a representation of the world surrounding the body [151].

1.5   Objectives of the Research

Research Area:

Stereoscopic 3D Displays, Stereoscopic Image Quality, Image Comparison Metrics, Human Perception and Stereoscopic Image Compression Techniques.

Key Objectives:

1. Research existing image and video compression algorithms, evaluate these algorithms and decide which compression techniques to use for the compression of stereo images.

2. Research current metrics for image and video compression, choose an image comparison metric, such as PSNR, that is suitable for comparing compressed stereo images.

3. Design and run a preliminary experiment to investigate symmetric and asymmetric

image compression. Evaluate the results of this experiment and conclude whether the metric is suitable for assessing stereo image compression.

4. Design and run a human based trial to produce a subjective threshold of acceptable stereo image quality and compare this threshold to those predicted using PSNR and the new HVS based metric.

5. Design and run a human based trial to evaluate the symmetric and asymmetric stereo image compression techniques and compare the result to those from Pixel based metrics and the new HVS based metric.

6. Investigate human visual system (HVS) properties and then develop a new (HVS) based metric. This metric should accurately predict the preferred amount of compression to be applied to a particular stereo image and produce a threshold of acceptable image quality for a particular stereo image.

1.6 Criteria for Success

The criteria for the completed research are outlined in this section. The Research will be evaluated against these criteria to determine its success at meeting its objectives. The criteria will be discussed and evaluated in the final chapter of the thesis.

The criteria for success are as follows:

1. A successful evaluation of symmetric and asymmetric compression approaches using existing image quality comparison metrics, such as Peak Signal to Noise Ratio and Sum of Absolute Differences.

2. The successful design and running of a subjective human based trial to evaluate symmetric image compression. This trial should produce an acceptable image quality threshold for symmetric images.

3. The successful design and running of a second human trial to compare the asymmetric compression method with the symmetric method, using human subjects.

4. A successful evaluation of whether current monoscopic image comparison metrics are suitable for stereo image evaluation and whether a more suitable metric can be developed to provide the ability to accurately assess the quality of stereo images.

5. A successful investigation of human visual system (HVS) properties leading to the development of a new HVS based metric. This metric needs to accurately predict the preferred amount of compression to be applied to a particular stereo image and produce a threshold of acceptable image quality for a particular stereo image. If this is achieved, then the metric can be classed as successful.

1.7 Main Contribution and Summary of Work

JPEG compression of the left and right components of a stereo image pair is a way to save valuable bandwidth when transmitting stereoscopic images [115]. In order to assess the effect of compression techniques on perceived image quality, 2D objective measures such as the Peak Signal to Noise Ratio (PSNR) or the Mean Squared Error (MSE) are often used. These measures indicate the difference between the coded and original image.

The transmission and storage of stereoscopic image material involves a large amount of data. Therefore, a substantial research effort is focused on understanding digital image compression (such as JPEG or MPEG coding) to obtain savings in both bandwidth and storage capacity. The same compression techniques used in two dimensional image material can also be applied separately on the left and right view of a stereoscopic image pair. Image compression may compromise perceived image quality however, through a reduction of detail and the introduction of compression artefacts such as blur, blockiness, or ringing.

Based on Binocular Suppression Theory, it is assumed that the binocular percept of a stereo image pair is dominated by the high quality component [67]. Thus, theoretically, when one image of the stereo pair is compressed such that a high quality is maintained, the other view can be coded more heavily without introducing visible artefacts in the binocular percept. The mixed resolution concept was introduced by Perkins [96]. Mixed resolution coding assumes that the binocular percept is not affected when one view is of high quality and the other view of lower quality.

However, according to Seuntiens et al. [115], subjective tests show asymmetric mixed resolution encoding of a JPEG stereo image pair results in image quality somewhat below the average of the two symmetrically encoded image pairs for the right and left view.

This thesis details the initial experiments performed to investigate these two conflicting approaches further. The problems with metrics like PSNR are detailed. To further investigate the symmetric and asymmetric compression problem human based experiments

were performed. The results of these experiments concluded that for the three test images, symmetric compression should be used as it produced better image quality results for a set file size.

A new human visual system (HVS) metric is developed. This new metric takes into account contrast and luminance changes at point matched regions of high spatial frequency. The metric is able to accurately predict stereo image quality.

The accuracy of this metric, is assessed using the results from the symmetric subjective human based trial. In this trial, participants were asked to view and assess the quality of symmetrically compressed stereoscopic image pairs. From this, a threshold of acceptable image quality was obtained for each image and compared to that from the new HVS based metric. The results showed that the metric conservatively estimated the stereoscopic image quality threshold calculated from the results from the human trial. From this, the conclusion was drawn that the new HVS metric was able to accurately predict the image quality threshold for the three test images.

1.8 Thesis Overview

The following sections are contained in the remainder of the thesis.

Chapter 2 - The Study of Stereoscopic 3D

This chapter introduces the history of stereoscopic 3D and then gives details of the background topics of stereoscopic depth perception, hardware and image quality.

Chapter 3 - Stereoscopic Compression

This chapter gives details of current compression formats, as well as existing monoscopic compression methods. It the looks at methods for compressing stereoscopic images.

Chapter 4 - Human Visual System (HVS) Models and Metrics for Assessing Stereoscopic Image Quality

This chapter details the literature relating to the modelling of the human visual system (HVS) and metrics, including those based on the HVS, for assessing stereoscopic image quality.

Chapter 5 - Investigating Symmetric and Asymmetric Encoding using PSNR

This chapter details the initial experiments performed to investigate two conflicting theories. Based on theories of binocular suppression, mixed resolution coding assumes that the binocular percept is not affected when one view is of high quality and the other view of lower quality. Subjective tests show, asymmetric mixed resolution encoding of a JPEG stereo image pair results in image quality somewhat below the average of the two symmetrically encoded image pairs for the right and left view. Work from this chapter is published in, Gorley and Holliman, Investigating symmetric and asymmetric stereoscopic compression using the PSNR image quality metric [44].

Chapter 6 - Evaluating Symmetric and Asymmetric Stereoscopic Compression

This chapter details a trial to measure asymmetric stereoscopic image encoding quality with human subjects, using the Double Stimulus Continuous Quality Scale (DSCQS) from the ITU-R BT.500-11 recommendation. These results are compared to those collected from the human-based symmetric stereoscopic image quality experiment, to evaluate which compression method, symmetric or asymmetric produces the better results.

Chapter 7 - Development of New Stereo Band Limited Contrast (SBLC) Metric

In this chapter we consider computational models of the human visual system (HVS) and describe the design and implementation of a new stereoscopic image quality metric. This point matches regions of high spatial frequency between the left and right views of the stereo pair and accounts for HVS sensitivity to contrast and luminance changes in regions of high spatial frequency, based on Michelson's Formula and Peli's Band Limited Contrast Algorithm. To establish a baseline for comparing our new metric with PSNR we ran a trial measuring stereoscopic image encoding quality with human subjects, using the Double Stimulus Continuous Quality Scale (DSCQS) from the ITU-R BT.500-11 recommendation. Work from this chapter is published in, Gorley and Holliman, Stereoscopic Image Quality Metrics and Compression [43].

Chapter 8 - Conclusions and Future Work

A summary of the work and possible extensions of the research are addressed in Chapter 8. The criteria for success is also revisited and discussed in this chapter.

1.9 Summary

This chapter has presented a brief introduction to the research area, motivation and problems, as well as an overview of Stereoscopic 3D, Image Compression, the Human Visual System and Image Quality Metrics. The chapter also presents objectives for the research, criteria for success and outlines the thesis contents.

Chapter 2

The Study of Stereoscopic 3D

2.1 Introduction

This chapter will provide details in the background areas of Stereoscopic Depth Perception, Hardware and Image Quality. We perceive the world around us in three dimensions. Recent advances in stereoscopic content production and hardware, have allowed for the viewing of three dimensions on computers, the television and at the cinema.

We will initially look at the history of stereoscopy. Humans have only relatively recently explained the binocular depth sense, stereopsis. In 1838, Wheatstone explained that there is a unique depth sense from retinal disparity [150]. Stereoscopic depth perception and the various stereoscopic storage formats and hardware will then be covered. We will then detail the factors for consideration when producing stereoscopic images of high quality.

2.2 History of the Study of Stereoscopic 3D

Stereopsis has been investigated for over two thousand years. Studies began sometime around 300BC when Euclid made observations on binocular vision. Additional studies on the subject of stereopsis have been conducted by others throughout history. Despite this curiosity in the field, it has been argued by Wade that practical technology, for example, the stereoscope, was not invented earlier "because the phenomenon of stereopsis based on disparity had not been adequately described" [135].

For instance, during the middle of the second century AD, Ptolemy studied the horopter, diplopia and correspondence. He had all the data required to construct a theory of depth perception through disparity detection, but did not [28]. In the early 1600's, almost 1500

years later, Descartes and Kepler linked the sensation of binocular convergence with depth perception. However, their work stopped short of a full understanding of the stereoscopic effect.

In 1838 Charles Wheatstone wrote a paper investigating Stereoscopy and Binocular Vision [150]. He used mirrored apparatus to create three dimensional photographs in the form of what is believed to be the first "stereoscope". This apparatus was improved and in the last few years of the 1800s started to appear in still cameras. In the 1960's, film technology had advanced sufficiently so that the head mounted display was produced, see Figure 2.1. The early work on these head mounted displays took place at the Bell Helicopter Company [22], with the intension of making it possible for pilots to see in the dark and land at night in rough terrain. The display was created so that the subject would look through a servo-controlled infrared camera. In 1966 the display was upgraded, replacing the camera outputs by computer images to create the first virtual reality system.

Figure 2.1: The Head Mounted Display, user views outputs from servo-controlled infrared camera to enable the subject to see in the dark [129]

Ivan Sutherland demonstrated the influence of this technology in an example in 1965 [128]. A camera was mounted on the roof of a building, with its field of view focused on two people playing catch. A viewer inside the building, who followed the motion of the ball, moving the camera by using head movements, wore the head mounted display. Unexpectedly, the ball was thrown towards the camera mounted on the roof, and the viewer who was inside the building, ducked. When the camera was set to observe the horizon, the viewer saw a panoramic skyline. Then the camera looked down to reveal that it

was "standing" on a plank extended off the roof of the building, the viewer panicked! This experiment demonstrated that a human could become totally immersed in a remote environment through the eyes of a camera.

The stereoscopic effect was researched to a much greater depth when the technology of displays advanced. The United States military researched it, in the hope of being able to use it to train pilots. From this, the technology has been put to many other uses, in such fields as medicine and entertainment.

2.3   Stereoscopic Depth Perception

When viewing the real world, humans perceive the depth of an object in many different ways. Each eye, left and right, perceives a different image and because our eyes are slightly separated from each other, the perceived images are slightly different. Figure 2.2, is a pair of left and right images, making a stereo pair.
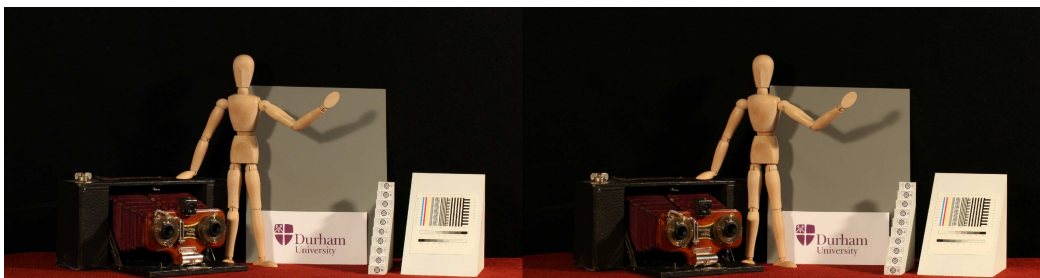
Figure 2.2: Stereo Image Pair, we see two 2D images, one in each eye to perceive 3D.

The images are slightly different in the same way we see slightly different things in each eye. The horizontal pixel difference in these two images is known as the stereoscopic binocular disparity. This is the horizontal difference in the image location seen by the left and right eyes.

The brain is presented with these two different, yet almost identical images, from which it is able to perceive a single 3D image. This perception of the 3D image is known as the stereoscopic affect and there are various ways it can be recreated. For example, the use of colour Anaglyph, where spectacles are used to filter different colours out of the image, producing a different image for each eye or the use of polarisation, where both the light from the images and the glasses are polarized such that the correct image is seen in each eye.

2.3.1   Other Depth Cues

Depth cues are concerned with the perception of depth in one image only. However, the more images that are available, the greater the information relating to the depth of an object. There are a number of other important depth cues which are divided into two categories, displacement, and optic flow, which are used in the animation of virtual reality scenes [98]. Displacement uses objects known as landmarks to enable the user to perceive depth, by assuming that the user knows the depth of these landmarks. As this method requires the user to know the depth of lots of objects, displacement does not allow the results of movement to be anticipated.

It is believed from research that optic flow is the method used by humans to perceive depth and was first identified using bees. Optic flow is the apparent visual motion that you experience as you move through the world. For example, if you were sitting in a moving car or a train, trees, ground and buildings would appear to move backwards. This motion is optic flow [14]. Consider Figure 2.3, the question of whether the ball is moving up and to the right (from 1 to 5), or the observer is moving down and to the left is not knowable, yet is very important information. If a patterned background was present it would be possible to tell the direction in which the ball or the observer is moving.
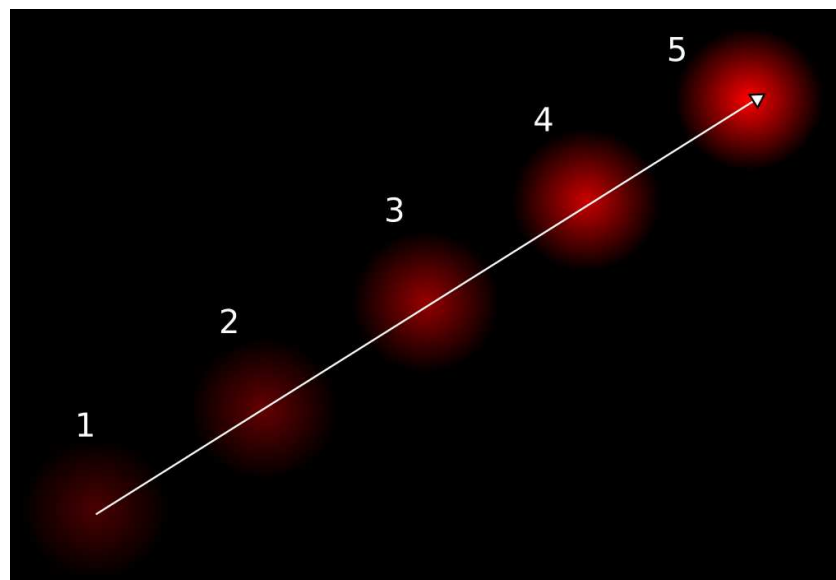
Figure 2.3: Optic flow, the question of whether the ball is moving from 1 to 5 or the observer from 5 to 1 is not knowable. [14]

There are many other methods of depth perception, including; Motion Parallax, Occlusion, Shadows, Linear Perspective, Texture Gradient, Size Gradient, Luminance and Depth of Focus. These are explained below and described in more detail by Robin Morgan [80].

Motion Parallax

Motion parallax is a form of optic flow and is a depth cue that results from our movement. As we move, objects that are closer to us move further across our field of view than objects that are in the distance. Through this it is possible to perceive the depth of objects.

Occlusion

The closest opaque object blocks further objects from view. It is probably the strongest depth cue.

Shadows

These provide information about the height of the object above the plane, Figure 2.4. Shadows are very important when objects are in motion.
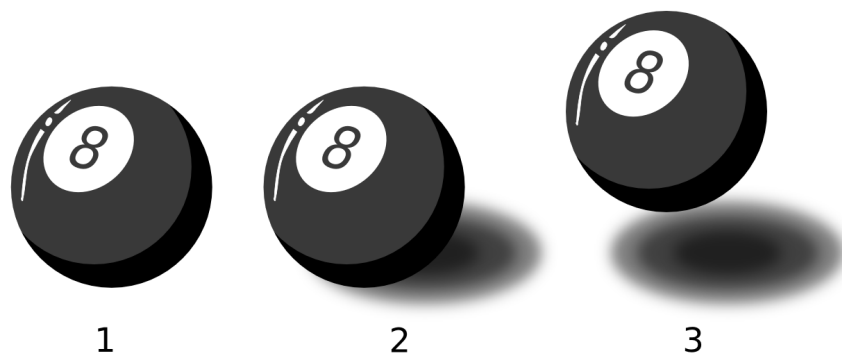
Figure 2.4: Example use of Cast Shadows, helps the observer perceive the height of an object above a plane.

Linear Perspective

This depth cue is based on the principle that parallel lines appear to converge to a single vanishing point. This is used in technical drawing to convey depth. If two lines are seen

as converging to the same point, they are perceived to be heading into the distance. This can be seen in the example below, Figure 2.5, where the rail tracks appear to converge to a single point.

Figure 2.5: Example of Linear Perspective, Rail tracks disappearing into the distance [14]

Texture Gradient

The texture appears to become denser with distance from the viewer. In the example below, the further away from the viewer, the smaller the blocks appear (Figure 2.6).

Size Gradient

Objects which are further away appear smaller than objects which are closer. In the example below, the people in the background appear to be much smaller (Figure 2.7).

Luminance

This is often also referred to as shading and presents information about the position of an object relative to the light source by the way it is shaded.

Figure 2.6: Example of Texture Gradient, The blocks in the driveway appear to become denser with distance from the viewer.

Depth of Focus

This is the principle of accommodation. Objects of interest are brought into sharp focus and the remainder of the image becomes blurred (Figure 2.8). Getting depth of focus correct in stereo image generation is very computationally expensive and requires knowledge of where the user is looking [42].

2.4   3D Stereoscopic Hardware

There are many methods of producing a stereoscopic effect using hardware. In 1833 Wheatstone produced a mirror stereoscope (Figure 2.9), this device was made up of mirrors arranged so that they focused two images into one stereoscopic 3D representation [70].

In more recent years the technology has advanced and there are now many different ways to view stereoscopic 3D images. These technologies will be detailed independently and include,

- Anaglyph

- Shutter Glasses

- Polarised Glasses

- Auto-stereoscopic Displays

Figure 2.7: Example of size gradient, the people further away appear smaller than those in the foreground.
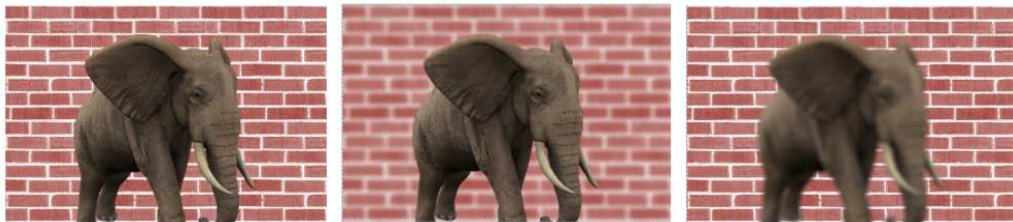
Figure 2.8: Example of depth of focus, objects of interest are brought into sharp focus and the remainder of the image becomes blurred.

The more modern stereoscopic three dimensional hardware, such as stereo displays, provides users with a binocular image. These types of display project a different view of a 3D scene to each eye. The differences between the two views are interpreted by the brain as depth. Objects then appear in depth to be behind or in front of the screen. Good quality displays provide a convincing three dimensional effect for people with normal binocular vision. However, further research is required of the tools and standards to make best use of these displays.

2.4.1 Anaglyph

To view an anaglyph stereoscopic 3D image the viewer must wear glasses that contain two differently coloured lenses. Red/cyan glasses are often used, but red/green and blue/yellow are also available. Figure 2.9, is a anaglyph image of a lion statue, the two images of the
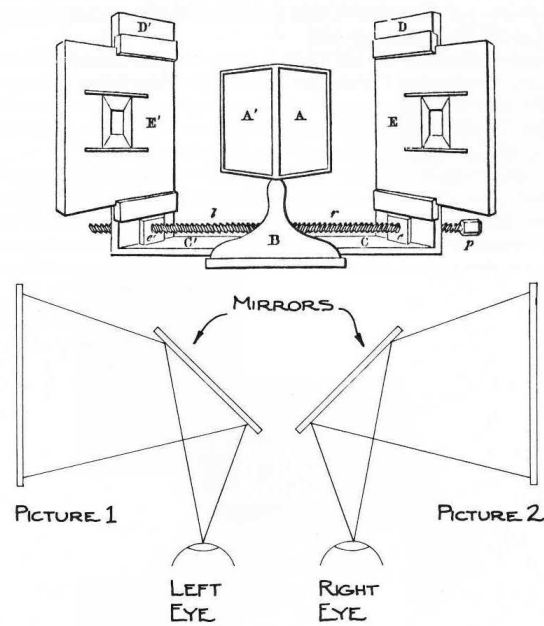
Figure 2.9: Wheatstone's Mirror Stereoscope. The head is brought up to mirrors A and A', and pictures E and E' are viewed stereoscopically. [150]

stereo pair are shown in the two respective colours. Anaglyph does however have certain limitations, for example the images are not seen in true colour as the glasses filter out certain colours.

The following quote by Alexander Klein of Stereoscopy.com, explains how the anaglyph technology works,

> "When you look through the red lens, only red light is allowed through. The eye that is covered with the red lens will see the red image. By the same token, the green lens only allows green light through, so the eye that is covered with the green lens will see the green image. In an anaglyph, when a given color filter stops the other colors, it is called subtractive filtration. Because the red and green images are slightly offset, each eye sees a slightly different view of the picture. This disparity simulates the distance between our two eyes, which provides two views of the same scene, therefore providing us with the perception of depth, or binocular stereopsis." [63]

Figure 2.10: Anaglyph 3-D photo of Edward Kemeys's lion statue outside the Art Institute of Chicago, in Illinois. [108]

2.4.2 Shutter Glasses

The lenses of shutter glasses are darkened one after the other to allow light to reach one eye then the other. The glasses are synchronised to the display or projector so that when the image for the left eye is being shown the right eye is darkened. This is then followed by the left eye lens darkening when the image for the right eye is displayed, Figure 2.11. This means that the correct eye only every receives the correct view. The switch between the two images and the switching of which lens is darkened is done very rapidly, ideally at 60 times a second. This produces a flicker free 3D image. This method has become recently more popular with the reduction in the weight of the shutter glasses and introduction of high quality 3 DLP projectors running at 120Hz.

2.4.3 Polarised Glasses

Un-polarised normal light travels in electromagnetic waves. These waves vibrate in many different axes including the horizontal and vertical. Polarised stereoscopic systems project light with different polarisations for each stereoscopic pair. For example, light might be transmitted in the horizontal axis for the left eye image and the vertical axis for the right eye image. Polarised glasses are then worn so that the viewer only sees the image with
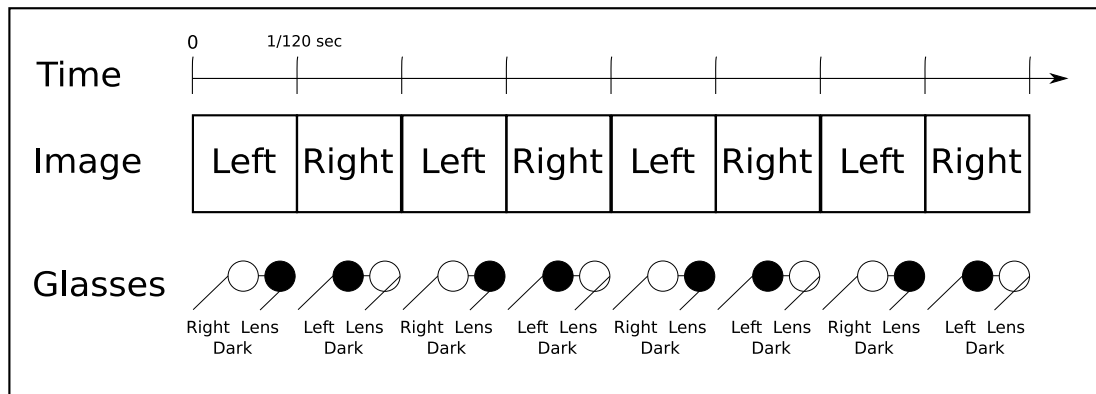
Figure 2.11: Shutter Glasses - Lenses darken independently in synchronisation with the screen or projector's refresh rate

the correct polarisation to pass through the filters on the glasses. Figure 2.12, shows this in more detail [121].

2.4.4 Auto-stereoscopic 3D Display Designs

In order that the viewer is not required to wear glasses or other devices, Auto-stereoscopic three dimensional displays automatically direct the left and right views to the appropriate eye. The three most significant design approaches that accomplish this effect are detailed below [47].

Two-view Twin LCD Auto-stereoscopic Displays

High quality auto-stereoscopic displays have been built successfully by using two LCD elements. The image from one of these elements is directed to the left eye, the image from the other directed to the right. These images are produced in two viewing boxes in space (Figure 2.13). Several display designs have adopted this method [32,33,45]. For the Image to be displayed in Stereo 3D, the viewer should sit in a central position and will have a maximum of 20 to 30 mm of movement around this position before the 3D effect is lost.

Displays of this nature are usually low cost and have a high resolution per view. Often it is possible to switch the screen between 2D and 3D; allowing the display to be used as a standard monitor when the 3D effect is not needed.
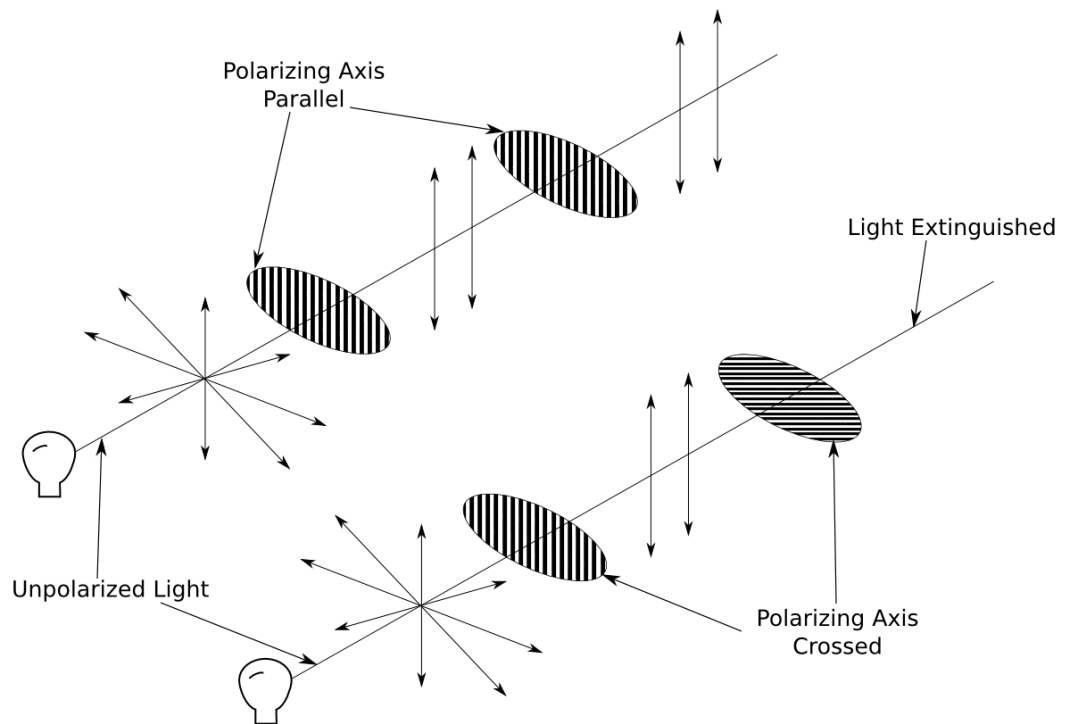
Figure 2.12: Polarised Glasses - Image with the correct polarisation to pass through the filters on the glasses [121]

Multi-view Auto-stereoscopic Displays

Multi-view displays present more than two views at the same time in multiple viewing boxes, for example, there may be eight, nine or even twelve views (Figure 2.14). The user sees two of these views at the same time. The views are presented so that a valid stereo pair is formed and displayed to each of the viewers. This gives a large viewing area and enables multiple viewers to see the effect at the same time.

These displays often have a lot lower resolution per view in comparison to the Twin-view displays, because the underlying display is split into multiple views. They are, however, relatively cheap to produce. Switching between 2D and 3D is rarely a feature of Multi-view displays.

Tracking Two-view Auto-stereoscopic Displays

Tracking two-view screens aim to offer the higher resolution of two-view displays along with the wide viewing freedom of multi-view screens (Figure 2.15). This is done by generating
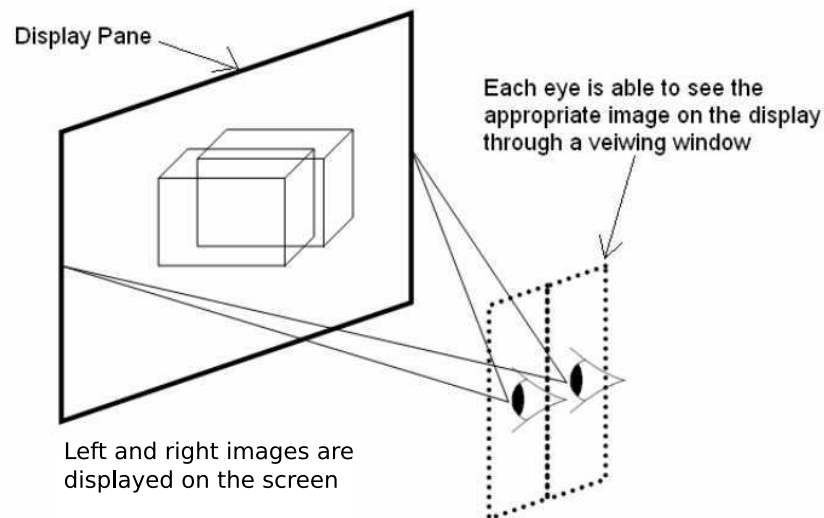
Figure 2.13: Two View Twin LCD Auto-stereoscopic Display [47]

only two views and getting these views to follow the user's head position. These displays have to detect where the viewer is and have a mechanism for rotating the viewing box towards them. Displays of this nature generally have a high resolution per view and a large view area. They are usually expensive due to the high cost of the head tracking and viewing box steering mechanism.

2.4.5   Comparing Stereoscopic 3D Displays

The following categories are useful in comparing the performance of the broad range of display types available and the appropriateness of different displays for a particular job. The following specific technical features of displays should be considered [48].

Total Display Resolution

Stereoscopic displays are designed to display one view to each eye. The total resolution for a display is the sum of all the pixels in all the views and determines the required bandwidth and computational effort necessary to display the images.

Resolution per View

The resolution per view is a very important variable when comparing the human perception of 3D displays. When using a stereo 3D display there is still a need for a high resolution.
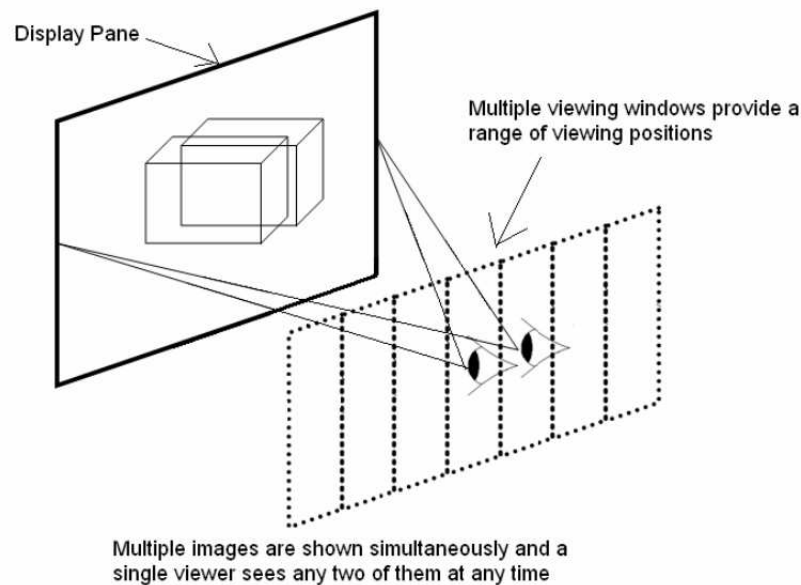
Figure 2.14: Multi-View Auto-stereoscopic Display [47]

This is because the display pixels are split between the number of views. Images on a 3D display can often appear to look better than on a 2D display with an equal total screen resolution. This is because the information received from the two views is integrated, by the human brain, into a single image.

Stereoscopic Resolution

Stereoscopic resolution is the number of depth voxels (intervals) within the range of +/- 100mm. It can be calculated for each of the displays by finding the screen disparity generated in this range. The total of these values is then divided by the stereoscopic pixel width of the display. The stereoscopic resolution of displays with a lower resolution per view is also low. The human eye is able to perceive at least 240 voxels over a range of +/- 100mm. A generic single LCD twin-view display, with a resolution of 1280(h) x 1024(v) is only able to produce 31 voxels with this range [48].

Viewing Freedom

The viewing freedom is the distance over which a viewer is able to move without losing the 3D effect. When using the screens for desktop applications the viewing freedom is less important. The viewers usually position themselves in the centre of the screen. Viewing
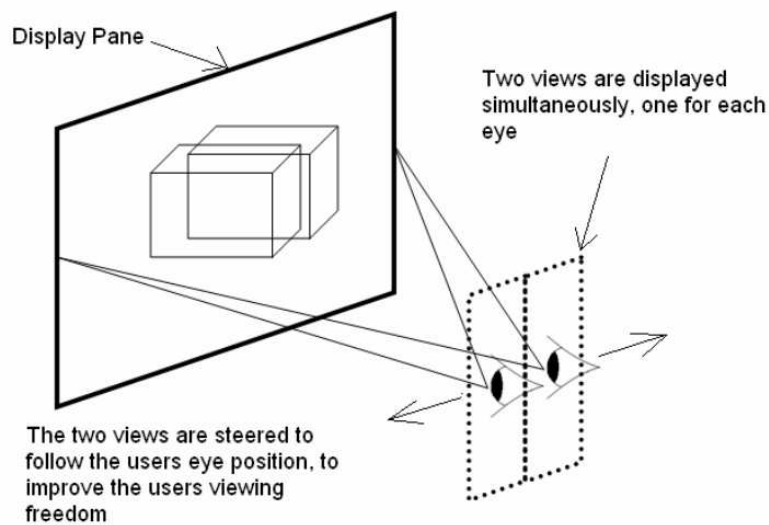
Figure 2.15: Tracking View Auto-stereoscopic Display [47]

freedom becomes of greater importance in applications such as public information displays where users are passing by a display rather than sitting down to use it.

Crosstalk

Inter-view crosstalk is often known as ghosting. It is a key performance characteristic for 3D displays. It is created when light leaks between the left and right viewing windows. In an ideal world the crosstalk would be zero, however in practise it is often a lot more. It is difficult to know the precise crosstalk of a display because the majority of manufacturers do not quote it for their products. For applications where depth judgement is critical, the use of a display with a high crosstalk is unlikely to be acceptable.

Brightness and Contrast

Often the brightness and contrast of a 3D display is not comparable to an equivalent 2D display. In parallax barrier displays, the brightness is reduced as light is blocked by pixels. With displays with two barriers, the rear barrier can be mirror coated on the side facing the illuminator to recirculate light.

2.5 Stereoscopic Image Quality

Badly designed stereoscopic images can often cause user discomfort, such as headaches, during viewing [66]. Therefore, when producing stereoscopic images, there are a number of guidelines that should be followed in order to try and prevent this discomfort.

2.5.1 Camera Geometry

The camera and axes geometry is an important factor that should be considered in stereoscopic image capture. For comfortable stereoscopic images to be produced it is important to get the camera geometry correct.

Camera geometry is the separation and orientation of the two cameras used to take the stereo images. The first is used to take still stereo image pairs and is Parallel Axis geometry, the second is used to take image sequences and is Converging Axis geometry. Different data compression algorithms are required depending on the camera geometry.

Parallel Axis Geometry

The two camera image planes are arranged in a co-planer and collinear fashion. Apart from the different offset position, this gives them identical optical characteristics to enable them to acquire a stereo image pair. Figure 2.16 illustrates the set up.

Here F is the focal length. This is the distance between the image planes and the location of the cameras. $P(x, y, z)$ is a point in the scene, which is projected onto the image planes at PL and PR. If the images are superimposed they do not coincide. The difference between the two images is the disparity.

Converging Axis Geometry

Setting the cameras up in converging axis geometry gives the following properties [87].

- It enables the fields of view to overlap fully in the scene volume of interest;

- It is easily calibrated, such that the origin of the global co-ordinate system is positioned where the optical axes of the two cameras converge (see Figure 2.17).

- The disparity will have both horizontal and vertical components and neither of them is normally zero.

Figure 2.17 illustrates the camera geometry set up.
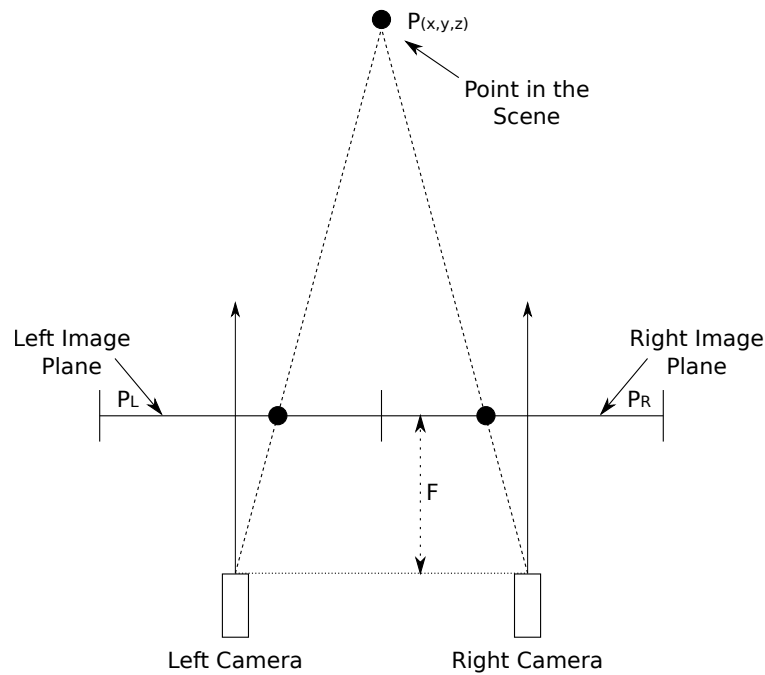
Figure 2.16: Parallel axis geometry: Cameras are arranged in a co-planer and collinear fashion. F is the Focal Length, P(x,y,z) is a point in the scene and PL and PR are the image planes [87]

2.5.2   Accommodation/convergence Relationship

The levels of convergence and accommodation are equal when perceiving objects in the real world. However, when viewing stereo images on the screen, the eyes will remain focused on the actual screen plane, but will converge on the perceived depth of the image. This perceived depth will vary for different portions of the image.

Lipton states, *"Thus there is a breakdown of the habitually tied-together responses of two separate mechanisms"* [71].

This divergence from normal viewing can result in a feeling of discomfort. Jones indicates that the limits to the perceived depth of an object, in order for a viewer to experience an image comfortably, are set at approximately 50mm out of the screen, and 60mm into the screen for a viewing distance of 700mm [56]. These figures represent a conservative lower bound.

2.5.3   Screen Surround

The stereo window is the plane in which the display screen lies. Lipton defined the edges of this stereo window as the screen surround [71]. Understanding the issues associated
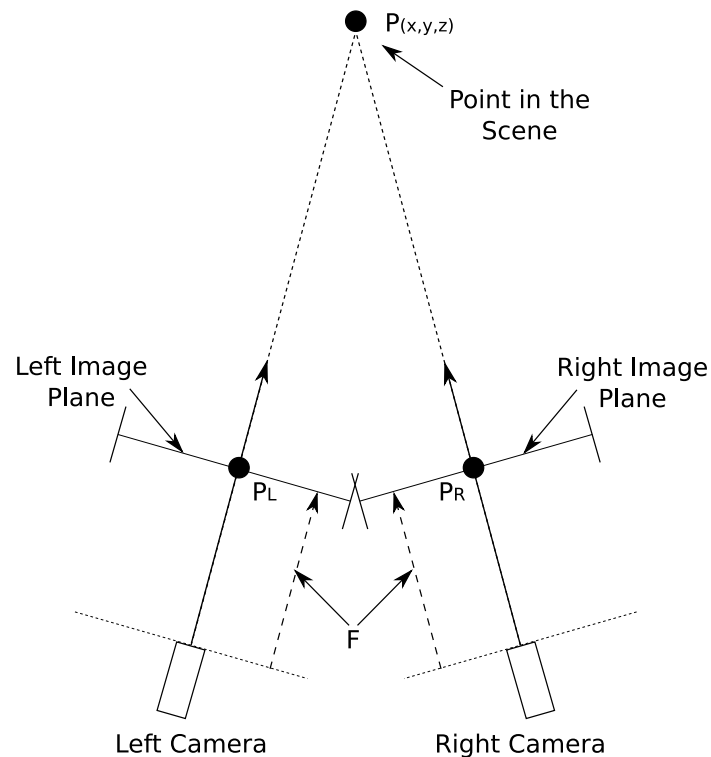
Figure 2.17: Converging Axis Geometry: Cameras are arranged in a toe-in fashion. F is the Focal Length, P(x,y,z) is a point in the scene and PL and PR are the image planes [87]

with screen surround is critical for stereo content production. Stereo images are either seen as projecting into the viewer space or within the stereo window. Images that are projected out into the viewer space can cause discomfort. This can be especially apparent if the image moves towards the edge of the screen.

Bourke suggests that *"Objects in the negative parallax region (in front of the screen) will present conflicting cues to the visual system if they are cut by the border of the screen, a region that is clearly at zero parallax"* [19].

Cues relating to the stereoscopic depth of an object enable the viewer to tell whether an image is positioned in front of or behind the stereo window. However, as sections of images are cut by the edge of the screen, the images may appear to be positioned along the screen plane.

2.5.4 Crosstalk

When using stereoscopic displays, crosstalk is the resulting effect of each eye seeing an image of the unwanted perspective view [71]. Ideally in a stereoscopic system each eye should only see its assigned image. Perfect stereoscopic images free from crosstalk can

be created using a stereoscope. This is because each image is optically separate from the other. Unfortunately stereoscopic displays can be imperfect and often crosstalk can be seen. Pommeray states that when using stereoscopic displays, there are two main opportunities for crosstalk to occur. These are departures from the ideal shutter in the eyewear and CRT phosphor afterglow [100]. Stereographics state that they have produced electro-optical shutters that are so good that no unwanted image will be perceived in the incorrect eye from incomplete shutter occlusion [71]. They state that any crosstalk viewed is from CRT phosphor afterglow.

Crosstalk for various technologies is categorised by Bos [18]. He also states that the percentage of crosstalk varies depending on screen position, with about 7% at the top of the screen due to the active eyewear still settling, about 4.5% in the middle of the screen and up to about 15% by the bottom of the screen as more afterglow is present. The following crosstalk rules of thumb are detailed [69],

- Lipton, Stereographics Corp. suggests reducing the disparity to reduce the double images that give rise to the stereo sensation. Smaller double images mean less crosstalk [71].

- Bos suggests that colour selection helps too. White lines on a blue background produce little crosstalk. Red produces less crosstalk than do blue and green, which have zinc sulphide, which causes afterglow in the standard P-22 phosphor set [18].

2.6   Summary

This chapter has presented the background of stereoscopic depth perception, hardware, display methods and image quality. Further research should be undertaken in stereoscopic image quality and compression, investigating different compression approaches along with the affect they have on stereo image quality. Methods of evaluation stereoscopic images should be investigated, the research should take into account current image quality metrics as well human visual system properties.

Chapter 3

Stereoscopic Image Compression

3.1 Introduction

This chapter will provide coverage of the main research area of stereoscopic image compression. The image compression formats of GIF (Graphics Interchange Format), PNG (Portable Network Graphics) and JPEG (Joint Photographic Experts Group) will be detailed. Current monoscopic and stereoscopic compression techniques are investigated, with particluar interest in the conflicting results of investigations in symmetric and asymmetric stereoscopic compression methods.

3.2 Image Compression Formats

This section gives an overview of the various techniques that are currently used in image compression. It is by no means a complete overview of all image compression methods. Compression algorithms can be divided into two distinct categories: they are either lossless or lossy. Lossless algorithms do not change the content of a file. If you compress and then decompress a file, it has not changed. Lossy algorithms achieve better compression ratios by selectively removing some of the information within the file. Such algorithms are best when used to compress photographic image files [11]. Greater savings in file size are achieved using lossy compression. However, it is necessary to calculate how much compression can be applied before unacceptable degradation occurs.

The following three image compression formats will be examined in detail,

- GIF (Graphics Interchange Format)

- PNG (Portable Network Graphics)

- JPEG (Joint Photographic Experts Group)

The GIF compression method was chosen as it is a compression method best used with vector graphics, PNG was chosen as it is a lossless compression method and JPEG was chosen as it is a commonly used lossy compression method. For these compression formats to be useful for stereoscopic image compression, the most appropriate method of applying the compression to the image pair, must also be considered.

3.2.1 GIF (Graphics Interchange Format)

The GIF format was produced in 1987 by CompuServe. It is a compressed format which was designed to reduce the amount of time it takes to transfer images over a network. GIF is a bitmap image format that uses a maximum of 256 colours. The format can either be used for pictures or animations. For animations each frame is also limited to 256 colours. GIF is now widely used on the World Wide Web.

In 1989-1990, CompuServe created a revised version of the GIF format, the 89a [52]. This new version provided support for multiple images in a stream and interlacing. It also enabled storage of application specific data. In the early 1990's when the Internet started to gain popularity, GIF became one of two formats for website based images. The other was X BitMap (XBM), used for black and white images. JPEG images did not start to be used until the introduction of true colour images.

GIF Technical Details

Unlike JPEG, the GIF file format uses lossless compression. Assuming that only 256 colours are used, this enables the size of the image to be reduced without dropping the quality of the image [97]. This enables the original image to be retrieved by uncompressing the reduced image. Due to the 256 colour restriction, the GIF format is unsuitable for photographs. GIF compression is ideal for images that contain sharp transitions, such as those in diagrams.

The colours are stored in a palette or table which allocates actual colour values to a colour number. When GIF was created, very few people had the ability to view images in more than 256 colours; this is why the 256 colour limitation seemed acceptable. Also for simple graphics, such as cartoons and drawings, 256 colours is usually adequate. To enable binary transparency, one of the colours in the palette can be set as transparent.

Alternatives

PNG was created with the objective to both replace and improve upon the GIF format. The PNG format overcomes the GIF patent problem and also provides more features and greater compression than GIF. The biggest disadvantage of PNG compared to GIF is that animation is not supported. For more information on the PNG format please refer to Section 3.2.2.

GIS File Format

The GIS image format is based on the GIF standard. It includes a special embedded GIS tag. This tag is known as the stereoscopic descriptor and is used by a viewer to determine how a stereoscopic image is stored. If a GIS file does not contain a GIS tag, then the image is simply a GIF file stored with a GIS extension. The specification states that files not containing a tag have to be stored in "Frame Doubling" stereoscopic 3D format by default [122]. To make use of other 3D options the GIS tag must be present.

The main advantage that GIS has over JPS is that when using a GIF viewer to look at a GIS file, the file will appear to be an animated GIF that alternates between the left and right views stored in the image. This gives the viewer a sense of 3D perspective.

The GIS file format uses lossless compression and also supports all the stereoscopic 3D format options available in the JPS file format. Unlike the JPS and PNS formats which support up to 16,777,216 colours, the GIS format, like GIF, is limited to only 256 colours.

3.2.2 PNG (Portable Network Graphics)

PNG was designed and created with the intention to both replace and improve upon the GIF format. It was defined in 1995-96 to overcome problems with GIF copyright issues (GIF uses patented LZW compression [148]) and thus a patent licence is not required to use the PNG file format. PNG provides a transferable format for storing, transmitting and displaying bitmapped (raster) images. It can compress and decompress images without loss of quality. PNG supports both colour and grey-scale images. To enable transparent images to be created using PNG, the format has an optional alpha channel, allowing the image to have multiple levels of opacity [3].

According to Roelofs [105], PNG is particularly well suited for use in two areas, as an intermediate editing format for repeatedly saving and restoring images and as a final display format for the World Wide Web. It doesn't compete with JPEG in terms

of compression and is best used with computer-generated images. To try and ensure that PNG was a better format than GIF the following features were included in the specification [110],

- Includes black and white, paletted and full colour formats.

- From 1 to 16 bits per colour.

- Transparency channel and alpha maps for overlapping images.

- Gamma indication for portability over different formats.

- Chroma, text and date chunks to store additional information.

PNG Technical Details

A PNG file's first 8 bytes contain the file signature. This is followed by a series of chunks. The signature contains the letters PNG and two new lines [2]. Using chunks within the PNG file format enables the PNG format to be extended and also to ensure its compatibility with older versions. Animation is not supported by the PNG file format.

The colour within PNG images can be made up in two ways. The images are either made up using image channels or they make use of palette indexed colour at a depth of 8 pixels per channel. PNG supports gamma correction. This is required because different computers, and especially their displays, interpret colours in different ways [39].

The PNG file format offers a variety of transparency alternatives. Single pixels can be declared as transparent within true colour or greyscale images. Alpha channels can be added to palette entries within images.

PNG uses a method of compression known as deflation. It is a non patented lossless compression. The compression is done by predicting the colour of each pixel based on the colour of the previous pixel. The predicted colour is subtracted from the actual colour for each pixel.

PNS File Format

The PNS file format is basically the same as the PNG file format. It stores stereoscopic 3D images in the "Cross Eyed" format. This is like the side by side format but has the right image first. Unlike the JPS format there is no special tag present. This format makes

use of the PNG loss-less compression technique and is therefore good for storing original images.

3.2.3   Comparison of Compression Formats

PNG compared with GIF

For the majority of images, PNG is able to achieve a greater compression than GIF. PNG also has a much wider range of colour depths and choice of transparency options than GIF; these options include alpha-channel transparency. The increased range of colour depths allows for both smoother fades and greater colour precision. The main advantage of GIF is that it supports animation while PNG does not (see Section 3.2.2, PNG Technical Details).

PNG compared with JPEG

PNG is not able to produce smaller file sizes than JPEG for photographic and photo-like images. This is because JPEG uses a lossy compression method especially designed for photographic image data. A large increase in file size (5-10 times) is seen when using PNG instead of JPEG for such images [2].

When images contain objects that include sharp transitions, text or line art, then PNG is the better file format choice. This is because sharp objects are in the high frequency domain and removed by JPEG's discrete cosine transform. If an image contains both sharp and photographic objects it is necessary to choose between the larger but sharper PNG and the smaller JPEG that contains artefacts around sharp objects.

3.2.4   JPEG (Joint Photographic Experts Group)

JPEG is an industry standard for compressing images. This format provides compression where you lose sharpness from the original [11]. The compression is provided by dividing the image into tiny, 8x8, pixel blocks, which are halved over and over until an adequate ratio is achieved.

JPEG itself specifies only how an image is transformed into a stream of bytes, but not how those bytes are encapsulated in any particular storage medium. A further standard, created by the Independent JPEG Group, called JFIF (JPEG File Interchange Format) specifies how to produce a file suitable for computer storage and transmission (such as over the Internet) from a JPEG stream. In common usage, when one speaks of a "JPEG file" one generally means a JFIF file.

JPEG is the most common format used for storing and transmitting photographs on the World Wide Web. It is not as well suited for line drawings and other textual or iconic graphics because its compression method does not perform well on these types of image.

A DCT-based (discrete cosine transform) method is specified for "lossy" compression and a predictive method for lossless compression [136].

Lossy Encoding

Detailed here is a common encoding method for the JPEG standard. This encoding method is lossy in terms of data compression. This compression makes use of the discrete cosine transform. It uses the transform to convert the images and then compresses the resulting coefficients. The level of compression used can be varied, which gives some control over the final image quality and size. The amount of compression varies depending on the image being compressed. Images that do not contain many high frequency details can be compressed the most.

The transform is applied to blocks of 8x8 pixels. This size is a trade-off between image quality, computational complexity and speed of compression. Two methods are used in compression; the coefficients are quantised, and are then Huffman or arithmetically compressed. The quantising of the coefficients is the lossy part of the sequence, where high frequency information is discarded.

According to Anderson, the encoding of an image using the discrete cosine transform can be split up in to nine steps [11].

1. Convert non-greyscale images into YCbCr components.

2. Down sample the blue difference and red difference chroma components (CbCr).

3. Group pixels into 8x8 blocks for processing.

4. DCT each pixel block.

5. Un-wrap the coefficients.

6. Scale each coefficient by a 'quantization' factor.

7. Eliminate near-zero coefficients.

8. Huffman encode data.

9. Add header information and quantization factors.

Lossless Encoding

The lossless encoding scheme is rarely used in practice. It makes use of a simple prediction and differencing method. It is possible to get 2:1 compression on images that contain between 2 and 16 bits per pixel. However, if the image has colour components, that is RGB values, then each component has to be encoded separately.

Images are encoded by scanning sequentially left to right, top to bottom. Current pixel values are predicted from previous pixel values. Encoding is done using either Huffman or arithmetic methods on the difference between the current value and the predicted value.
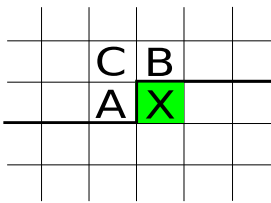
Pixel Neighbourhood X = current pixel		Encoder Type	Predicition Scheme
		0	No Prediction
		1	A
		2	B
C **B**		3	C
A **X**		4	A + B - C
		5	A + ((B - C)/2)
		6	B + ((A - C)/2)
		7	(A + B)/2

Figure 3.1: Lossless Mode prediction schemes, the predictor combines up to three neighboring samples at A, B, and C in order to produce a prediction of the sample value at the position labeled by X. Any one of the predictors shown in the table below can be used to estimate the sample located at X [11].

There are eight prediction types and associated methods available to the encoder, Figure 3.1. Type 0 is used for the hierarchical progressive mode and type 1 is always used for the very first scan line. Type 2 is used to find the first pixel in a new row.

JPS Image Format

The JPS image format is based on the JPEG standard. It includes a special embedded JPS tag. This tag is known as the stereoscopic descriptor and is used by a viewer to determine how a stereoscopic image is stored. If a JPS file does not contain a JPS tag then the image is simply a JPEG file stored with a JPS extension. Files not containing a tag are stored in cross-eyed stereo 3D format by default.

Many popular stereo formats are supported by the JPS standard. The formats include, Interleaved (for use with shutter glasses), anaglyph for use with "red and green" coloured

three dimensional glasses and both over under and side by side formats for parallel and cross-eyed viewing.

The JPS format uses lossy JPEG compression with a high compression ratio. Siragusa and Swift define and provide more details relating to the JPS image format [122].

JPEG Multi-Picture Format

JPEG Multi-Picture Format (MPO, extension .mpo) is a JPEG-based format that has been designed for multi-view images. It contains a series of JPEG files concatenated together. There are special EXIF fields that describe its purpose. It is used by the Fujifilm FinePix Real 3D W1 camera [21].

3.3 Monoscopic Image Compression Techniques

The following three techniques will be considered,

- Wavelet Transform Compression

- Fractal Image Compression

- Vector Quantization

3.3.1 Wavelet Transform Compression

Wavelet transform is a type of data compression that is mainly used for reducing the file size of images. It is sometimes used for video compression and audio compression. It is one of the most effective types of image compression. The aim is to reduce the amount of image data stored within a file to as little as possible. This results in a certain loss of quality. Therefore like JPEG, Wavelet transform is a lossy compression technique.

JPEG 2000

JPEG 2000 is one of the most familiar formats using the wavelet transform method of compression. Like JPEG it was designed by Joint Photographic Experts Group and was produced with the intention of superseding the original discrete cosine transform based method. It has improved compression performance, as well as better editability and scalability. It is able to work at high compression ratios without the images becoming "blocky and blurry".

JPEG 2000 applies a form of transform coding to compress images. It supports very low and very high compression rates including lossless compression. One of JPEG 2000 main strengths is its ability to handle a very large range of effective bit rates [7].

Legal Problems

The majority of present software available does not support JPEG2000. This is due to the possible danger of patents on the compression method. JPEG2000 is itself not licence free, however, contributing companies and organisations have agreed that the core coding system can be obtained free of charge. The official JPEG homepage states that,

"It has always been a strong goal of the JPEG committee that its standards should be implementable in their baseline form without payment of royalties and license fees. The up and coming JPEG 2000 standard has been prepared along these lines, and agreement reached with over 20 large organizations holding many patents in this area to allow use of their intellectual property in connection with the standard without payment of license fees or royalties." [5]

However, the JPEG committee has also noted that undeclared and obscure submarine patents may still present a hazard:

"It is of course still possible that other organizations or individuals may claim intellectual property rights that affect implementation of the standard, and any implementers are urged to carry out their own searches and investigations in this area." [5]

Comparison with PNG

JPEG2000 provides support for lossless encoding. It was not designed however, to replace current lossless image formats. The PNG format still provides better compression for images that contain many pixels of the same colour. The PNG file format contains a greater range of compression features than the JPEG2000 format.

Assuming that there are no further changes to the PNG standard, it is expected that it will be used more heavily for compressing diagram type images and that the JPEG2000 standard will be used to compress more picture based images. Marcellin et al. in their technical overview of JPEG 2000 suggest that in the future, PNG and JPEG2000 will be used in a similar manner to the way that GIF and JPEG were used in the 1990's [74].

3.3.2 Vector Quantization

Vector quantization is based on the principle of block based coding. It is a fixed length algorithm for lossy data compression. Initially, vector quantization was seen as a challenging problem, as it required multidimensional integration [41]. In 1980 an algorithm based on a training sequence was proposed by Linde, Buzo and Gray [68]. The use of this training sequence removed the need for multidimensional integration.

Image compression is a common application of vector quantization. The idea behind it is simply to approximate regions within an image. The idea is similar to rounding off in mathematics. A one dimensional vector quantization example is shown below (Figure 3.2).
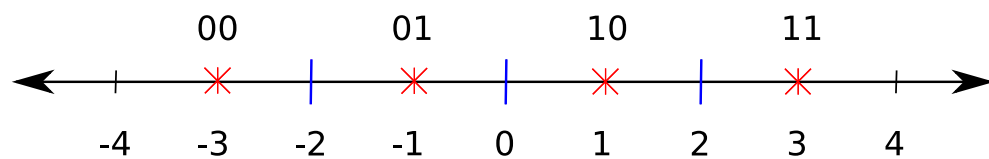
Figure 3.2: One Dimensional Vector Quantization, all numbers less than -2 are approximated by -3, numbers between -2 and 0 are approximated by -1, 0 and 2 by 1 and above 2 by 3 [41]

In this example all numbers less than -2 are approximated by -3, numbers between -2 and 0 are approximated by -1, 0 and 2 by 1 and above 2 by 3. The approximate values can then be uniquely shown by two bits. It is therefore a one dimensional, two bit vector quantization.

A two dimensional vector Quantization is explained below (Figure 3.3). Two dimensional, vector quantization is used for images.

Here pairs in each region are approximated by the position of the red star associated with that region. In this example there are 16 regions and therefore 16 stars. 4 bits are required to uniquely identify each region. It is therefore a two dimensional, four bit, vector quantization.

3.3.3 Fractal Image Compression

Fractal Image compression makes use of fractals for lossy compression. This compression method was designed to be used for images of natural scenes. Within an image, certain parts resemble other parts of the image. Fractal compression relies on this similarity. Fractal compression was invented in 1988 by Arnaud Jacquin [106]. More recently Michael
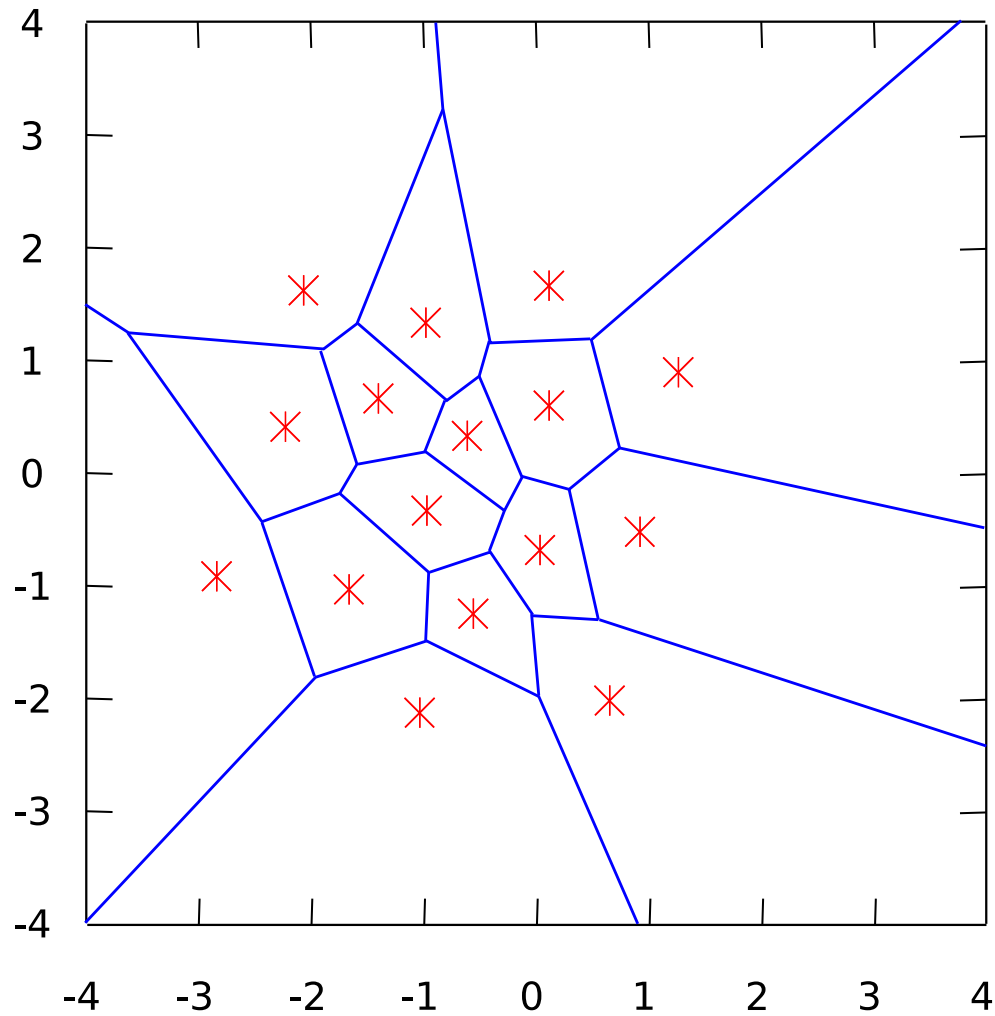
Figure 3.3: Two Dimensional Vector Quantization, pairs in each region are approximated by the position of the red star associated with that region [41]

Barnsley [13] and Alan Sloan [54] have researched the compression method and have taken out patents.

Fractal image compression is based on the principle that images often contain several similar objects. These similar objects or fractals need only be stored once enabling compression within an image. An example to illustrate fractals within an image is the Sierpinski Triangle (Figure 3.4). The triangle can be copied onto itself [132].

3.4   Stereoscopic Compression Approaches

Compression plays a significant role in the storage and transmission of digital video data and images. The main methods of compression are designed to remove the temporal and
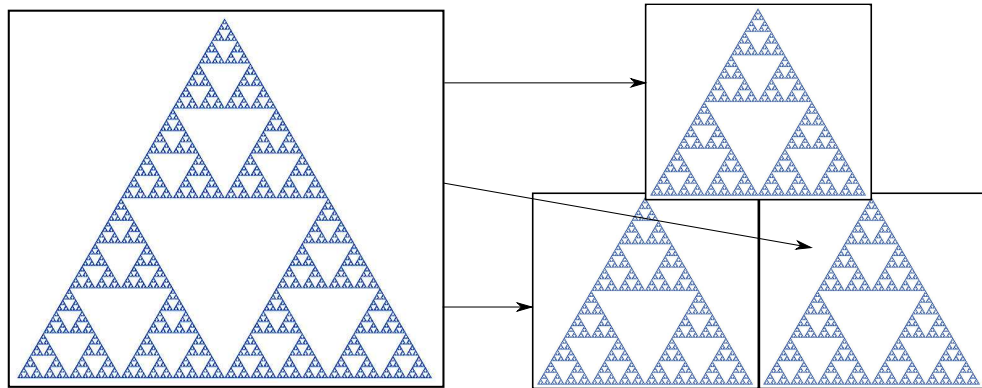
Figure 3.4: Fractal image compression: Sierpinski Triangle [134].

spatial redundancies that are present in both images and video sequences [62]. Three dimensional stereo images are produced by simulating the way human eyes observe objects through two horizontally separated perspectives. This means that two images are necessary for the production of one 3D image. These images are referred to as left frame and right frame. To enable these frames to be transmitted and a 3D image produced at the receiving end, twice the bandwidth is needed in comparison to transmitting a monocular image. Therefore, some kind of data compression of these two frames is necessary. According to Naemura et al. [85] it is well known that there is a high correlation between the left and right frames. Good quality stereoscopic image compression should make use of this correlation [120].

3.4.1 Block Based Stereo Image Compression

The Block Based stereo image compression technique takes into account the correlation between the left and right parts of a stereo image. This is done by coding one image using conventional motion-compensated Discrete Cosine Transform methods such as MPEG-2 and the other image is predicted from present and previous images. The biggest advantage of this technique is that the conventionally coded image can be decoded using a normal MPEG-2 decoder and can then be displayed on a conventional 2D television screen. These advantages are critical at the moment, though over the next few years, digital terrestrial television will mature to enable the viewing of 3D images and video. It is believed that the cost of the extra equipment (such as LCD glasses) and bandwidth will reduce significantly and may well be within reach of many customers.

The Digital Imaging Group, University of Glamorgan, have developed a block based

compression technique that they believe has an easy development method and low cost implementation. The basic method is detailed in Digital Image and Data Compression, Block Based Stereo Image Compression [87]. They found for general purposes, DCT compression of the reference image followed by block-based compression of the disparity map gave optimal results. They conclude that the approach yields the best compromise between quality and compression rate.

Multi-resolution Based Block Matching

In 1995 a technique for multi-resolution Block Based Matching was proposed by Sethuraman et al. [113]. It was believed that only one image of a stereo pair needed to be of a high resolution to enable depth to be perceived satisfactorily. Sethuraman et al. made use of this visual property in the proposed technique, to enable image sequences to undergo greater compression. Using the technique, only one image sequence is coded independently using motion compensated prediction, whilst the sequence for the other eye is generated from the coded image sequence using disparity compensation, estimation and motion compensation of previous images in the sequence.

The right and left image sequences are decomposed on a frame by frame basis. The decomposition is done in a series of different resolution levels and is performed using a 2D wavelet transform. The first stream is entropy and run length encoded before transmission. The second stream is created using hierarchical block based motion compensation and estimation. The initial estimations are coarse, but are refined further at each stage. The estimation method assumes that the camera axis is both parallel and aligned in the horizontal and that the disparity is scalar. If the blocks to be estimated only appear once in the image, temporal prediction is used to generate the approximated frame, as disparity estimation is not possible.

3.4.2   Object Based Stereo Image Compression

For true disparity between left and right frames to be taken into account, object based stereo image compression would be more efficient than the block based techniques proposed in Section 2.6.2. Video compression methods, such as MPEG-4, are proof of this. However, the complexity level and difficulty of encoding the shape of the objects will also increase.

To overcome this problem, the Digital Imaging Group, Glamorgan [87] have proposed a hybrid scheme with characteristics of both block-based and object-based coding meth-

ods to achieve data compression for stereo image pairs. They detail that their proposed algorithm comprises the following operations [87]:

1. Extract objects by contour analysis for both frames, and then match the objects to find those areas which are similar in terms of their shapes.

2. Enclose the matching object pairs by object bounding rectangles, in a similar spirit to MPEG-4. A parity based contour filling algorithm is further used to identify the interior and exterior pixels of the object, in which the shapes of the objects are identified by binary object planes. These planes are only used as an aid in the coding process rather than being encoded.

3. Encode the objects in terms of three groups:

   - Objects that are enclosed by closed contours;
   - Objects enclosed by contours that terminate at the image frame boundaries;
   - Unidentified areas that are treated as the background.

3.4.3 Inter-frame Disparity/Depth coding

2D Motion Compensation

Object and Block based stereo image compression exploits the correlation between the left and right image streams. If however, there is a large degree of motion within an image sequence, the above techniques become inefficient. For these motion intense image sequences, a better approach is to use object or block based compression methods in conjunction with motion compensation. According to Ebroul Izquierdo M. [53] and Tzovaras D. et al. [133] when using a compression method such as MPEG-2, which already makes use of motion compensation, the same motion vectors can be used for motion in the depth map or disparity field. However, this 2D Motion compensation method only works when the object's depth in relation to the camera remains constant.

3D Motion Compensation

3D motion compensation has the advantage over 2D compensation in that is it able to cope with depth within an image sequence. It is able to cope with more complex motions, including rotational motion and objects moving towards and away from the camera. These extra advantages do, however, come at the cost of increased computational complexity.

Calculation of 3D motion compensation is able to be determined by the presence of the extra view for disparity and depth estimation.

3.4.4 Residual Image Coding

Disparity estimation is one of the main focuses of stereo image coding. This technique takes into account the redundancy in the stereo image pair. Making use of the residual image for compression has been investigated by Frajka et al. [38]. Residual image coding aims to make stereo image compression more efficient, by improving the coding of the residual image. The method individually encodes the blocks of the residual image that correspond to the blocks used within the disparity estimation process. This enables the correlation properties of the residual image to be used to improve the compression. Occlusion within the images is handled by not estimating the blocks that have large distortions.

3.4.5 Multi-resolution Auto-stereoscopic Image Compression

Large amounts of storage space and large bandwidths for transmission are required by multi-view stereoscopic images. If the image display and the source are close, then it may be possible to put up with high bandwidths, however, compression of the information is essential for transmission over long distances. Large amounts of work have been done, and techniques developed, for the compression of 2D images, however, less has been done on the coding of stereoscopic images. It is possible to apply the current 2D compression techniques to the individual views within a stereo pair; however, this does not take into account the high correlation between the two images.

Shah and Dodgson present more details on the issues relating to multi-resolution auto-stereoscopic image compression [116]. Four different multi-view image compression techniques were investigated. The four techniques were: differential pulse code modulation (DPCM), disparity estimation, three-dimensional discrete cosine transform (3D-DCT) and principal component analysis (PCA).

They concluded that the lossless methods, DPCM and disparity estimation, showed that it was possible to encode multi-view images at an acceptable compression rate. The results from the DPCM showed that, for a given pixel, it was the nearby pixels within a view that provide a better predictor than the corresponding pixel values in adjacent views.

They proved this by investigating disparity estimation using search space measures. They observed that, although the search space measures performed well, the overhead of

encoding the shifts, reduced the advantages of the method and therefore a novel method of encoding was developed. This method collected similar scan-lines, which were then disparity estimated, rather than estimating the disparity for each scan-line separately.

Investigations were also carried out into lossy compression. They found that the JPEG-like encoding technique, 3D-DCT, showed that multi-view images can be coded efficiently. An investigation into the performance of the 3D-DCT system for a number of different multi-view images using different quantisation methods was carried out. They found that a set of quantisation volumes proved to be effective in reproducing good quality images, as well as a similar or better compression rate when compared to the 2D JPEG coding scheme. The second experiment used the PCA coding scheme. They concluded from the two experiments that it was possible to code new images with respect to blocks but not with respect to views.

3.4.6   Symmetric and Asymmetric Compression

In recent years there has been an increased interest in three-dimensional displays, television systems and virtual environments [103]. With this increased interest, the processing of stereoscopic images, image sequences and video has become more important. As stereoscopic images require an increase in the bandwidth compared to monoscopic images, the compression of stereo images is one of the most important factors that must be evaluated to enable the extensive use of three-dimensional systems.

The space required to store a stereoscopic image is normally twice the amount need to store a monoscopic one. Similarly, for stereoscopic video, twice the bandwidth is required. One method of reducing the storage space or required bandwidth is to compress the images. Existing research shows that both symmetric and asymmetric compression methods are able to significantly reduce the size of the stereoscopic content. However, there is an inconsistency within current research, with different papers suggesting that both symmetric and asymmetric compression methods produce better results. In this section, the existing research is acknowledged and reviewed and conclusions drawn as to the state of current publications.

Definition of Symmetric and Asymmetric Compression Methods

The same techniques that are used for the compression of monoscopic two-dimensional (2D) images can be applied to a stereoscopic three-dimensional (3D) image. The technique

is often applied independently to the left and right views of the stereo pair.

Symmetric compression of a stereoscopic image pair is when the same amount of compression is applied to the left and right image within the pair.

If a different amount of compression is applied to each image within the pair, then the stereoscopic image is said to be compressed asymmetrically. In basic terms, asymmetric compression is when a high quality image is presented to one of the eyes and a low quality image to the other. Examples of symmetrically and asymmetrically compressed images are shown in Figure 3.5 and Figure 3.6. The uncompressed original is shown in Figure 3.7.
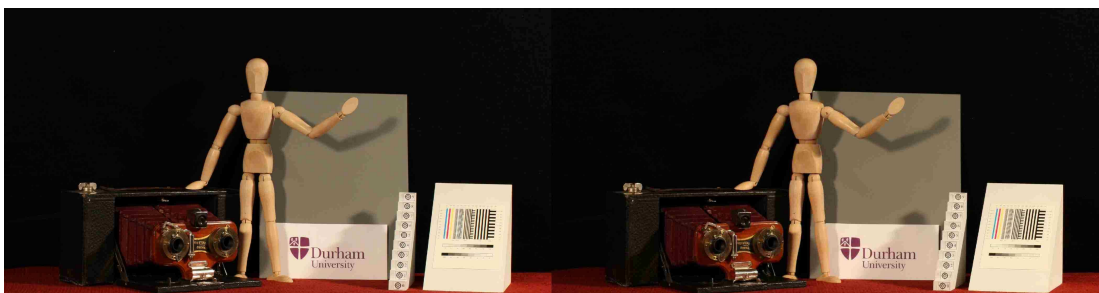
Figure 3.5: Symmetrically compressed stereo image pair. The same amount of compression is applied to both images.
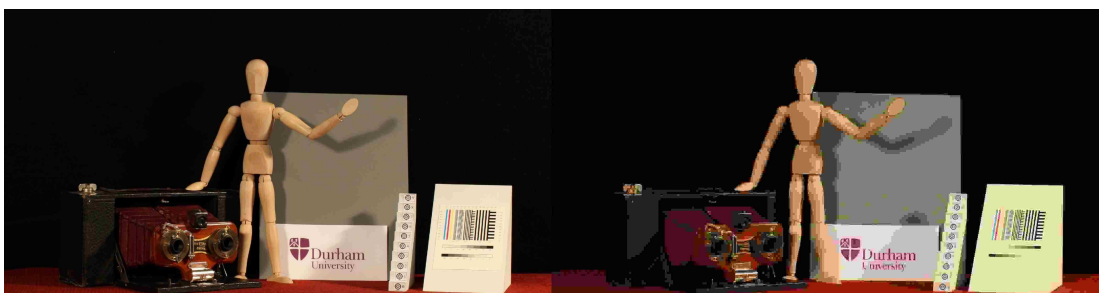
Figure 3.6: Asymmetrically compressed stereo image pair. A different amount of compression is applied to each image in the pair independently.

Suppression Theory of Binocular Vision

If two similar images are presented to each eye independently, the result is a single percept of the scene. This observation is known as binocular fusion [61, 64, 84]. If the eyes are presented with distinctly different images, then instead of these images being combined they will be in competition with each other. This competition is known as binocular

Figure 3.7: Uncompressed DVL Test Scene, there is no compression applied to either the left or right views.

rivalry [61,64,84]. This means that the percept within the image will be unstable, shifting between the patterns of each eye. This will mean that certain sections in one of the images within the pair will dominate sections in the other image and vice versa. This results in a shift in dominance in each eye. Information in the less dominant image will be suppressed, whilst the information in the more dominant image will be visible [58].

This suppression theory is often used as the basis for validating the following asymmetric stereo compression scheme: One image of the stereo pair is used to contain details of the scene, whilst the other image within the pair is simply used to contain the disparity information. This method enables the second image in the pair to become less dominant and be highly compressed, without affecting the depth information in the overall compressed stereo image [51]. This compression method has been used in several different ways by different people. A multi-resolution approach is adopted by Serthuraman et al. [113], whilst Dinstein et al [31], use a disparity compensation technique. In both these papers, the disparity is calculated between the stereo images. However, Dinstein et al. [51] and Perkins [96] also provided methods where the disparity between the images is not computed. In these approaches a Discrete Cosine Transform (DCT) based compression method is applied to one of the images within the pair. Reynolds and Kenyon, asymmetrically compress the images using wavelet compression based on the theory of binocular suppression [103].

Existing Symmetric and Asymmetric Comparison Experiments

Symmetric and Asymmetric Compression of Images

The transmission and storage of stereo images requires a nominal doubling of the file size. Research has been undertaken to discover if additional savings can be made by

exploiting properties of the human visual system and especially to asymmetric compression encoding of stereo pairs [125]. JPEG compression of the left and right views of a stereo image pair is a method to save valuable bandwidth and disk space when transmitting and storing stereoscopic images. Seuntiens et al. [115] investigates the effect of varying both the Symmetric / Asymmetric JPEG encoding and camera base distance on the overall stereo image quality. They investigate how perceived depth, sharpness and eye strain vary for 48 different conditions. The stereoscopic depth of the image was varied by altering the camera base distance and the JPEG compression ratio was also changed. All the experimental conditions they use are shown in Table 3.1.

Table 3.1: Compression ratio and camera base distance conditions

Camera-Base Distance = 0, 8 and 12cm			
Image Compression Ratios (L,R)			
(Orig,Orig)	(Orig,1:30)	(Orig,1:40)	(Orig,1:60)
(1:30,Orig)	(1:30,1:30)	(1:30,1:40)	(1:30,1:60)
1:40,Orig)	(1:40,1:30)	(1:40,1:40)	(1:40,1:60)
1:60,Orig)	(1:60,1:30)	(1:60,1:40)	(1:60,1:60)

The results of the asymmetric and symmetric coding showed that how perceived image quality varies with file size is complicated. They found cases where the image quality of a symmetric coded pair was higher than that of an asymmetric coded pair, even when the averaged bit rate for the symmetric pair is lower, than for the asymmetric pair. From this it can be concluded that symmetric image compression out performs asymmetric compression in these cases, but still more investigation is required.

Siegel et al [119] and Sethuraman et al [113] investigate compression methods for left and right image pairs and the correlation between the two images. They find that presenting a good quality image to one eye and a poorer quality image to the other, results in a better perception of depth, than just presenting one good image alone. They also observe that a binocular stereoscopic image pair with one blurred image and one sharp image is able to successfully stimulate appropriate depth perception. They investigate and find that if one member of the image pair is compressed without loss or nearly without loss and that the other image in the pair is heavy compressed, the high and low resolution views will be perceived both accurately and comfortably. When using natural images, Dinstein et al [51], Pastoor [93] and Yano & Yuyama [158] showed that it is possible to significantly reduce the compression of one image in a stereo pair without affecting the

overall impression of sharpness.

Both Meegan et al [76] and Stelmach et al [124], performed experiments showing that when left and right views of an image are combined, the higher quality view masks coding artifacts in the lower quality view. They showed that properties of binocular fusion make it possible to exploit asymmetric coding to improve compression efficiency.

Mixed resolution of random dot stereo-grams [58] showed that binocular fusion can occur if either the low or high frequency spectrum are identical, however, those frequency components that are not identical cause binocular rivalry. Asymmetric encoding of images has been assessed as a bandwidth reduction technique [126]. The best method of compression varied, depending on the type of degradation, with asymmetric compression being used for low-pass filtered images but for quantization, symmetric compression should be used. They concluded that, depending on the type of image being compressed, asymmetric compression can be used to reducing storage and transmission bandwidth for some stereoscopic images.

Perkins performed experiments to compare mixed resolution data compression of stereo pairs [96]. Images were encoded asymmetrically with one high resolution and one low resolution image making up each stereo pair. He found that the eyes can easily fuse mixed resolution stereo pairs and perceive depth in them and that the stereo images produced were perceived as closer to the high quality image than the low quality image. Images were subjected to a high pass filter to blur them. It was concluded that, for the levels of blur applied, asymmetrically encoding one image with a high compression and one with a low compression produced, better results for a set file size than encoding one image heavily and the other not at all.

A large number of subjective studies are undertaken by Seuntiens, [114] to investigate asymmetric compression, crosstalk and depth perception. In the asymmetric experiments, the findings show that the observers rate the images with a global score based on the lowest quality image within the pair and therefore symmetric compression produces better results. The same experiments are repeated, after applying asymmetric blur to the images, by Meegan et al. [76]. Meegan finds that applying blur, results in a better quality asymmetrically compressed image. Benoit et al. suggest that depending on how the stereo pair has been distorted, an asymmetric compression method applied to the image may theoretically produce better results [9]. Tests to identify the impact of eye dominance, reveal no effect by Seuntiens [114] and Kalva et al. [59] when evaluating stereoscopic images.

Shneor et al. [118] analysed eye dominance when viewing asymmetric coded images and showed that displaying the high quality component to the dominant eye improved the image quality score.

Seuntiens et al [115] undertake further investigations into how asymmetric JPEG compression and camera base distance affects image quality, perceived depth, sharpness and eye-strain. Investigations were performed using two stereoscopic still scenes and three different camera base distances (0, 8, 12 cm) and four different compression ratios (original, 1:30,1:40 and 1:60) where used. The JPEG compression was applied independently to the left and right images within the pair, resulting in a 4 x 4 matrix of all possible symmetric and asymmetric combinations. Again the DSCQS method of ITU-R BT recommendation 500 [1] was followed and observers were asked to assess image quality, sharpness, depth and eye strain. The results show that JPEG compression has a negative effect on image quality and perceived sharpness and eye-strain. No effect is shown for perceived depth. Also the results of eye-strain and sharpness correlated well with image quality. Increasing the camera base distance resulted in increased depth being perceived, however it also resulted in more eye strain being reported.

How to code images in line with factors of the Human Visual System is investigated by Dinstein et al [31]. They state that, when presenting a sharp image to one eye and a blurred image to the other, people should be able to perceive a world that is both sharp and three dimensional. Both Dinstein et al [31] and Siegel et al [120] conclude that further work is needed as the question of what to compress, one eye's view heavily (asymmetric compression), or both with equal amounts (symmetric compression) still remains unanswered.

Symmetric and Asymmetric Compression of Video

Recent studies using the ITU recommendation 601 have reported minimal loss of perceived video quality for resolution mismatched stereo video inputs [126]. However, many factors affect stereoscopic quality, including the contents of what is being compressed and the method used for resolution reduction. For example, the quality threshold for cinema production is different to that for a normal desktop based stereo viewing environment [107].

Fehn et al extended a multi-view coding method based on the MPEG 4 standard [78], to take into account mixed resolution video [35]. Their approach is motivated by the suppression theory of binocular vision [58] which they detail and indicate that the

depth effect and perceived sharpness of a stereo pair that is asymmetrically encoded is dominated by the higher quality component of the pair [77, 93, 125]. They hypothesized that, using asymmetric compression they would be able to achieve very high compression gains without reducing the overall visual quality of the three-dimensional video. The results showed that the use of asymmetric encoding is a suitable method to reduce the overhead needed for transmitting the additional right view stream. The reduction resulted in a joint bit rate slightly larger than the amount required to transmit a single monoscopic video. However, they did not validate their experiments using human subjects.

Stelmach & Tam [124] compared asymmetrically compressed stereo sequences, with 26 subjects, using the ITU recommendation 500 [1]. They used three different stereo sequences for the assessment, with the addition of a fourth sequence for training participants. Each of the sequences used were 10 seconds in duration and were stored in the ITU-R 601 format at 60Hz interlaced with a resolution of 720 x 480. The videos were processed independently of each other using MPEG-2 with standard Main Level settings. The bit rates used were 6,3,2 and 1 Mbits/s and were combined into the following symmetric and asymmetric stereo sequences (left:right), 6:6, 6:3, 6:2, 6:1, 3:3, 3:2, 3:1 and Uncompressed:Uncompressed. The left stream was always displayed at the higher bit-rate. Their results showed that subjective quality of the asymmetric sequence fell approximately mid way between the quality of the left and right views. The results were as predicted and consistent with known properties of binocular vision. They went on to investigate asymmetric compression of videos further, with human subject trials, again following the DSCQS method from the ITU recommendations [1, 126]. They looked at the asymmetric compression method as a way of reducing overall bandwidth and storage of stereoscopic videos. The results showed that although asymmetric compression was a promising technique for reducing these transmission and storage overheads, the results depended on the type of degradation being applied to the image. For example, for video being degraded using quantization, symmetric compression works far better than for images degraded using low pass filtering.

The design and investigation of a three-dimensional stereo video system is undertaken by Kalva et al. [60]. They state that the characteristics of the Human Visual System can be exploited to enable the individual views to be compressed at different quantities without it affecting the quality of the final three-dimensional video. They examine the bounds of asymmetric compression and its relationship to eye dominance and find that asymmetric

encoding may be able to produce better results depending on eye dominance, however they conclude that although the possibilities of dominance and suppression mechanisms during binocular fusion exist, they are not yet well understood and should be investigated further. [112]

The results from the above research into symmetric and asymmetric compression is conflicting. The reason for this maybe that some of the images where blurred or filtered before the compression methods were applied or that the experiments investigated images at high levels of compression, beyond what is likely to be used in practise.

3.5   Summary

This chapter has presented image compression formats, monoscopic and stereoscopic image compression approaches and details of current symmetric and asymmetric stereoscopic compression experiments.

JPEG is the most commonly used image compression format. There is conflicting evidence in the research as to whether symmetric or asymmetric stereoscopic image and video compression produces the better results. More work needs to be undertaken to investigate these two compression approaches further.

Chapter 4

Human Visual System (HVS) Models and Metrics for Assessing Stereoscopic Image Quality

4.1 Introduction

This section gives an overview of the literature relating to the modelling of the human visual system (HVS) and how humans perceive visual information. It provides background on human visual system anatomy and physiology, human visual system models and current metrics based on HVS models, that are used to assess image quality.

4.1.1 Motivation

One of the goals of the work is to evaluate compression of stereo images, including the methods of compression and how much compression can be applied without noticeable reduction in image quality. In order to evaluate the stereo image quality, some type of assessment measure is required. For an accurate assessment of image quality, one needs to study and understand the way the human visual system (HVS) works and the way in which visual information is combined and depth perceived. Although the fields of Anatomy, Physiology and Psychology have helped us develop an understanding of stereoscopic perception, models of the human visual system are still scarce and simplistic. For the most part, metrics are used to estimate the quality of the images for each eye separately and then these two estimations are combined in a metric to give an overall estimation of

quality.

4.1.2 Outline

The section begins with an overview of human vision, including the anatomy of the eye, optical nerve and areas of the brain that are used to enable us to see. The area of visual perception and existing models of the human visual system are then examined. The mathematics behind the models are explained in this section, including Luminance, Band Limited Contrast, Colour and Temporal and Spatial Masking. The final area details current visual quality metrics, with descriptions and evaluations of the most popular metrics, including human visual system metrics and how, if possible, they have been applied to evaluating the quality of stereoscopic images.

4.2 Human Vision

Binocular vision provides humans with depth perception. This is produced by the difference in the points of the two images that incident on the retina and is known as stereopsis, which helps to provide precise information about the depth of objects. Human vision also uses other cues to interpret depth perception.

4.2.1 The Eye

The eye contains the lens, pupil and retina. A small image is focused onto the retina using the lens. The eye is able to adjust to different lighting conditions by using the iris. The image produced on the retina of the eye is not what is perceived. Instead a percept is computed by the brain, based on the sensory image information picked up by the retina. It is during this computation that the image on the retina is inverted. Figure 4.1 illustrates the anatomy of the eye.

The sclera is the outermost layer of the eye. It is tough and is used to maintain the shape of the eye. The cornea is clear and is the front part of the sclera. All light that enters the eye passes through the cornea. The extra-ocular muscles are used to move the eye. The second layer of the eye contains blood vessels and is called the choroid or uveal tract. It contains the two structures, the iris and the muscles of the ciliary body for controlling the size of the lens.

The light sensing part of the eye, the retina, is the inner most part of the eye. It
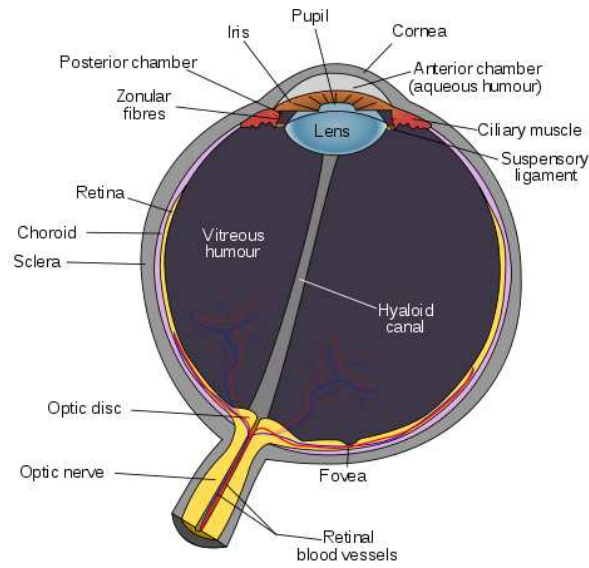
Figure 4.1: Anatomy of the human eye [104].

contains rods, for low light vision and cones for colour vision and detail. The centre of the retina contains the fovea. Here there are only cones and is where vision is sharpest. The eye's blind spot is where the optic nerve, containing the retinal arteries, joins the eye.

The Visual Angle Defined

The visual angle is an important concept in relation to the eye. Visual angles are defined in degrees, minutes and seconds of arc. A visual angle of approximately one degree can be approximated as a 1-cm wide object viewed at 57cm. The following equation is used to calculate the visual angle θ which is measured from the optical centre of the eye [145].

$$\tan\frac{\theta}{2} = \frac{h}{2d} \tag{4.1}$$

or

$$\theta = 2\arctan\frac{h}{2d} \tag{4.2}$$

Where h is the height of the object and d is the distance from the object to the optical centre of the eye (Figure 4.2).
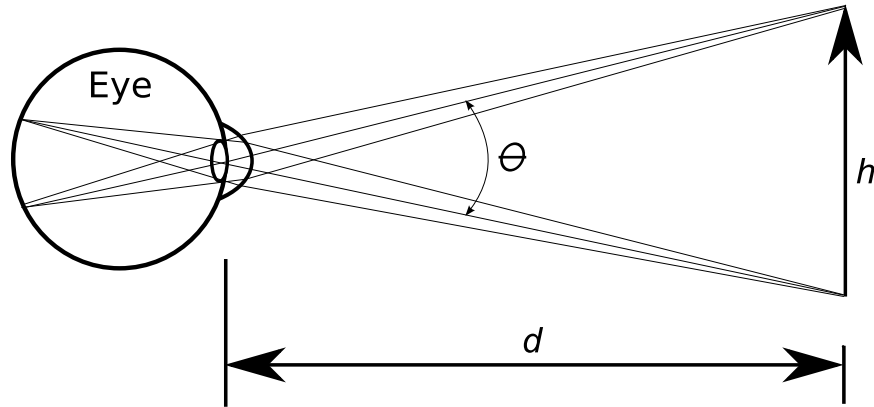
Figure 4.2: Visual Angle: The diagram shows an observer's eye looking at a frontal extent (the vertical arrow) that has a linear size h, located in the distance d [145].

The Lens

The lens of the eye has two parts, the curved surface of the cornea and the crystalline lens. The nodal point is the optical centre of the lens and is approximately 17mm from the retina. Usually the distance from the cornea is used as a measure for the eye. Where optics are concerned, it is better to measure the distance to the nodal point, instead of the centre of the eye. Imaging properties of a simple lens can be described using the following equation [137].

$$\frac{1}{f} = \frac{1}{d} + \frac{1}{r} \tag{4.3}$$

Where f is the focal length of the lens (i.e. the distance of the focus from the lens), d is the distance from the object to the optical centre of the eye and r is the distance to the formed image (Figure 4.3).

Chromatic Aberration

Humans focus different wavelengths of light at different depths within the eye. This means that the human eye is not corrected for Chromatic Aberration. Long wavelength red light is refracted less than short wavelength blue light.

According to Podoleanu [99], the eye's chromatic aberration can give rise to strong illusory depth effects. An example of this is when both blue and red text is displayed on a black background. For about 60% of observers the red text appears closer. 30% see the blue text as closer and the remaining 10% see both the red and blue text in the same
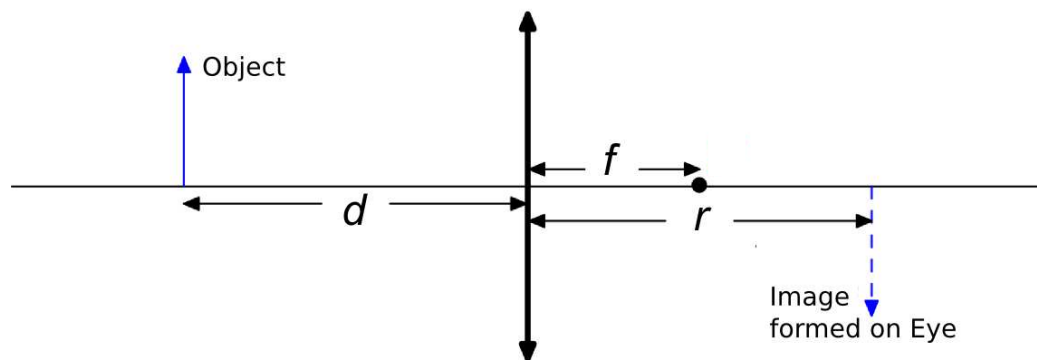
Figure 4.3: The Lens: d is the distance of the object from the lens; r is the distance of the image from the lens and f is the focal length.

plane.

Receptors

The eye contains two types of photoreceptor cells, these are rods and cones. Cones are used to detect normal levels of light including colour, whereas rods are used to detect low light levels. The input from rods is spread over a large area with many rods adding to each fibre optic nerve signal. The cones within the eye are located in a small area in the centre of the retina called the fovea. The cones are closely packed together and it is here that the vision is most sharp.

Acuities
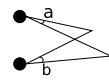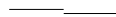
The eye's ability to see fine detail is measured by visual acuity. It is a quantitative measure of the ability to see an image that is in focus at a certain distance. Normal visual acuity is often referred to as 20:20 vision. The standard definition is the ability to resolve a spatial pattern separated by a visual angle of one minute of arc [24].

Five of the most basic acuities are summarised in Table 4.1. Ware has more details relating to the eye, in *Information visualization, perception for design*, Chapter 2 [145].

4.3   Important Human Visual System Characteristics

Due to metric based image comparison problems, research into the operation of the human visual system (HVS) has resulted in objective quality metrics, for 2D image comparison,

Table 4.1: Summary of the Basic Acuities of the Eye [145]

Point Acuity (1 minute of arc): The ability to resolve two distinct point targets.	
Grating acuity (1-2 minutes of arc): The ability to distinguish a pattern of bright and dark bars from a uniform grey patch.	
Letter Acuity (5 minutes of arc): The ability to resolve letters. The Snellen eye chart is a standard way of measuring this ability. 20/20 vision means that a 5-minute letter target can be seen 90% of the time.	
Stereo Acuity (10 seconds of arc): The ability to resolve objects in depth. The acuity is measured as the difference between two angles (a and b) for a just detectable depth difference.	
Vernier acuity (10 seconds of arc): The ability to see if two line segments are collinear.	

based on HVS characteristics. [29, 89, 149]

Some of the most important human visual system properties include:

- Sensitivity to contrast changes rather than just luminance changes. [149]

- Varying sensitivity to artefacts and errors at different spatial frequencies. This can be modelled by Band Limited Contrast, which estimates the visibility threshold for stimulus at different spatial frequencies. [147]

- Higher level perceptual factors, such as attention, eye movements and how different

types of coding artefacts are unacceptable. [90]

- Stereoscopic rivalry between the left and right images. [139]

- Masking, this refers to reduced inability to detect a stimuli on a spatially or tempo-rally complex background. [79]

4.4 Current Models of the Human Visual System

Current models of the human visual system can be split into two different categories. These are neuro-biological models and models based on properties of human vision. Low level processes of the optic nerve and the eye are estimated and then a neuro-biological model is produced. These models tend to be very complex and are therefore not very useful in real-world applications [153]. Aspects of human vision relevant to picture quality are predicted by psychophysical models. These model the human visual system sensitivity to light, also known as *luminance masking or lightness non-linearity* [111]. They are often built upon psychophysical experiments, implemented as a series of processes, as shown by Boev et al. in Figure 4.4 [4].

The response of the HVS depends on local variations of luminance (i.e. relative luminance and contrast), and it varies very little in relation to the absolute luminance value for an image. This property of the HVS is known as *Webster-Fechner Law* [111]. Also, the calculated perception depends on the image content and the colour and luminance model used, which make HVS modelling very complex. [94, 95, 154]
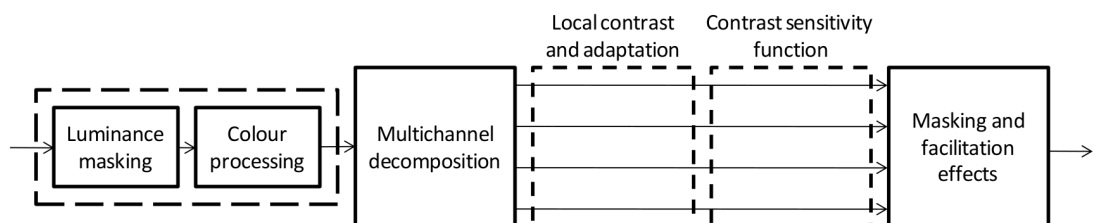
Figure 4.4: Block diagram of typical psychophysical HVS model [4].

4.4.1 Luminance and Relative Luminance

Luminance is the measure of apparent intensity, it is the measure of the intensity of the luminosity of light travelling in a certain direction. In developing models, the human visual

system sensitivity to luminance is measured for different wavelengths using subjective tests, where subjects match light intensities from different light sources. Wyszecki and Stiles, define luminance with the CIE function $V(\lambda)$. It is the ratio of radiance between two wavelengths, which have the same apparent intensity when using direct comparison [157]. The CIE function $V(\lambda)$ is used in the following equation (4.4) to predict if two spectral power distributions $P_{1\lambda}$ and $P_{2\lambda}$ would match:

$$\int_{\lambda} P_{1\lambda} V(\lambda) d\lambda = \int_{\lambda} P_{2\lambda} V(\lambda) d\lambda \tag{4.4}$$

Relative luminance is based on the photometric definition of luminance above, however, with the output value for luminance normalized to a reference white [101]. As with the photometric definition, it is based on the luminous flux density in relation to a particular direction. This is radiant flux density weighted by the luminosity function $\overline{y}(\lambda)$ of the CIE Standard Observer. For red, green and blue colour spaces that use the ITU-R BT.709 primaries, relative luminance, I, can be calculated from the linear red, green and blue components [82].

$$I = 0.2126R + 0.7152G + 0.0722B \; \text{(Red, Green and Blue Values)} \tag{4.5}$$

The formula takes into account the luminosity function. It is weighted so that the green light contributes the most to the intensity perceived by humans, and the blue light the least. It is how this relative luminance varies in regions of high spatial frequency within an image that can be used to evaluate image quality. The overall relative mean luminance of the whole image can calculated using the following equation,

$$L = \frac{1}{255 * mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I(x,y) \tag{4.6}$$

where m and n are the height and width of the image and (x, y) are the normalized RGB values.

4.4.2 Michelson Formula and Band Limited Contrast

One of the most important perceptual attributes of an image is apparent or perceived contrast. In recent years techniques to enable modification and manipulation of contrast within an image have been developed. Peli proposes a definition of local band limited contrast for images, based upon the Michelson formula and relative luminance [94]. The Michelson formula is:

$$Michelson's\ Formula\ (C_{(x,y)}) = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \tag{4.7}$$

where L_{max} is the maximum relative luminance value for the image and L_{min} is the minimum relative luminance value for the image. The band limited contrast for each point can then be calculated using the following formula,

$$Band\ Limited\ Contrast\ (BLC_{(x,y)})\ per\ Pixel = \frac{C_{(x,y)}}{L_{(x,y)}} \tag{4.8}$$

Where $C_{(x,y)}$ is the Michelson Formula per pixel (x, y) and $L_{(x,y)}$ is the Relative Luminance per pixel (x, y).

How band limted contrast varies across an pair is shown if Figure 4.5.

Figure 4.5: Band Limited Contrast varying across an image. [94]

Osberger et al. produce a human visual system based metric based upon the Peli local band limited contrast algorithm, that produces a perceptual image quality rating and contrast distortion map for a monoscopic 2D image [91]. They conclude that their algorithm based metric provides accurate results in assessing monoscopic image quality. Tolhurst and Tadmor investigate band-limited contrast in natural images. They investigate the amount of change in Michelson's Contrast needed to detect changes in colour amplitude [131].

4.4.3 Colour

Colour is a visual perceptual property. Colour categories and physical specifications of colour can be associated with objects and materials. However, colour is a psychophysical property of the human visual system as opposed to a physical property of an object or material [15]. The human perception of colour derives from the spectrum of light, that interacts in the human eye with the spectral sensitivity of the three light cone receptor types of the eye.

When trying to produce accurate models of the human visual system, it is critical to have an understanding of the way humans perceive colour. The way humans perceive colour is studied by *colour-matching experiments*. Brainard and Freeman asked observers to adjust the intensities of primary lights to match the colour appearance of a test light [20]. Details of asymmetric colour matching experiments are described by Wandell [138]. In these experiments, a subject studies the colour of an object under a light source and then must select an object that looks the same colour from under a second light source.

4.4.4 Temporal and Spatial Masking

The human visual system has different levels of sensitivity to different spatial frequencies. Daly has presented a human visual system model that takes into account temporal and spatial masking [29]. A visual threshold map across the image is produced to try and predict the visibility of artefacts within the image by calculation. Fairchild et al. have also produced a model which also includes temporal and spatial masking, however this model does not predict threshold differences, but tries to predict image quality well above the visual threshold of acceptability [34]. A temporally-aware model for video is proposed by Watson, which tries to predicts the differences in video quality threshold values [146]. They try to measure the observer's internal perceptual scale for visual impairment. This is done by asking observers to rate which of two videos is more impaired.

4.5 Current Metrics for Assessing Visual Quality of Images

4.5.1 Pixel Division

Pixel Division is the simplest image comparison metric. A third image is produced from two input images by dividing the pixel values in the first image by the corresponding values in the second image. Pixel division can be done on a single image where the Red, Green

and Blue values of each pixel within the image are divided by a constant. The resulting images can then be visually compared [75].

Equation for Pixel Division

A single pass over the image is used to extract the numerical representations of each of the images. The images are then compared using the following equation [30],

$$Q(i,j) = \frac{P_1(i,j)}{P_2(i,j)}$$
(4.9)

Where P_1 and P_2 are the input images and Q is the output image. i and j are the indices of the pixel.

If the division is to be done on a single image, then the equation is modified and is displayed below,

$$Q(i,j) = \frac{P_1(i,j)}{C}$$
(4.10)

where C is a specific constant.

For colour images where the pixel values are vectors, the individual red, blue and green components are simply divided separately to produce the output.

Usage

Pixel division is best used to detect change within images. Pixel division gives a fractional change between pixels in two images. An example of using pixel division to detect change between two images is shown below (Figure 4.6).

Figure 4.6: Example of Pixel Division [37].

Image A and B are of the same scene except two objects have been moved slightly. Dividing image A by image B using a floating point pixel type gives the resulting image C. Pixels that have remained constant between the two images have a value of 1 whereas the pixels that have been changed have a value somewhere between 0 and 1. This means only the objects that are changed are noticeable within the new image.

4.5.2 Mean Square Error and Peak Signal to Noise Ratio

The peak signal to noise ratio (PSNR) is an engineering term and is the ratio of the maximum possible power of a signal to the power of the corrupting noise that affects the quality of this signal. Because of the wide range of this signal, PSNR is usually expressed as a logarithmic decibel (dB) scale [102].

One of the main uses of PSNR is to measure the quality of image compression. The Mean Square Error (MSE) is calculated for two images. One of the images is usually an approximation or compressed image of the first. MSE is defined as:

$$MSE = \frac{1}{mn} \sum_{x=0}^{m-l} \sum_{y=0}^{n-l} \| I(x,y) - I'(x,y) \|^2 \tag{4.11}$$

where $I(x,y)$ is the pixel values for the original image; $I'(x,y)$ is the compressed version, and m and n are the dimensions of the images. The Peak Signal to Noise Ratio (PSNR) is then defined as:

$$PSNR = 10 \cdot log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \tag{4.12}$$

where MAX_I is the maximum pixel value of the image.

A low value of MSE means that there is a lesser error and as PSNR has an inverse relationship with MSE, this means that a high value of PSNR equates to a lesser error. The higher the PSNR value of a compression scheme the better it is. Cho et al. used PSNR to compare two methods of stereoscopic image compression [26]. Frajka et al. state that for comparison of stereo images using PSNR, it should be calculated using the average of the MSE of the reconstructed left and right images [38].

4.5.3 Structural Similarity Index

The similarity between two images can be measured using the structural similarity (SSIM) index. The index is used to compare one image against another that is regarded as perfect

quality. Surface Luminance of an object within an image can be calculated by adding together the illumination and reflectance of that object. Structures within a scene of an image are independent of illumination. Therefore, when investigating structures within an image, it is necessary to remove any influence illumination may have. Structural information is made up of attributes that represent the structural objects within the scene and this information is not dependant on the luminance or contrast of the image.
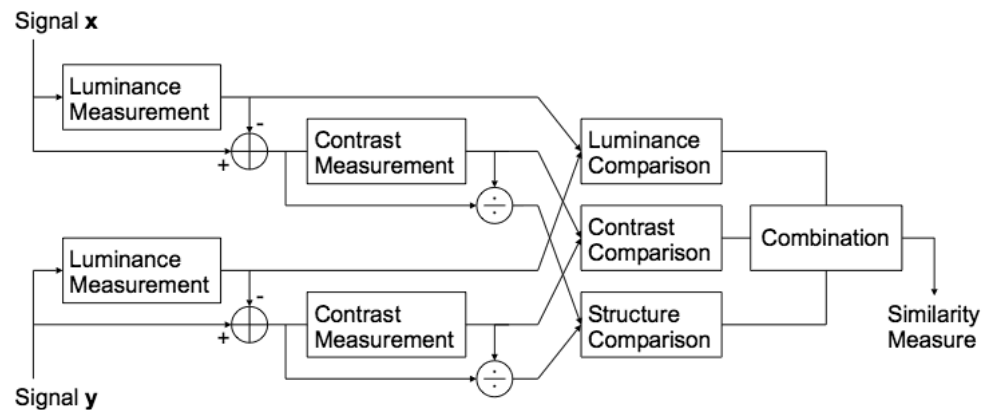
Figure 4.7: Diagram of the Structural Similarity Measurement System [141].

The system diagram for structural similarity is shown in Figure 4.7. The inputs x and y are two aligned none negative signals and one of the signals has perfect quality. From this it is possible to measure the quality of the second signal in comparison to the first. The comparison measurements are split into three different tasks for luminance, contrast and structure. Each is calculated separately and then combined to produce the similarity measure. The whole structural similarity process is explained in detail by Wang et al. [141].

Extensions and improvements to the Structural Similarity Index are detailed by Wang and Simoncelli [140] and by Wang et al. [142,143]. Implementations are also available from Wang et al. [144]. The Structural Similarity Index Metric (SSIM) has benefits over Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) as it is measured locally, rather than computed over the entire image. The SSIM is compared for an 8x8 window, which moves one pixel at a time across the image. The mean SSIM is then the arithmetic average of the local scores.

4.5.4   Colour Histograms

Colour Histograms is an image comparison method that is used by many applications. They are an efficient method of image comparison and are insensitive to small changes in a camera viewpoint. They are often used in multimedia applications and examples of their use include querying an image database [88] and scene break detection [12]. Their popularity stems from the following factors [92].

- The overall algorithm design is detailed in Digital Image and Data Compression, Object Based Stereo Image Compression [55]. Colour Histograms are computationally trivial.

- Small changes in camera viewpoint tend not to affect colour histograms.

- Different objects often have distinctive colour histograms.

However, as colour histograms lack spatial information, it is possible for two completely different images to have very similar histograms. For example the images below have similar colour histograms (Figure 4.8). Despite their different appearances there is approximately the same quantity of red in each image. The red within the first image is well scattered whereas the red in the second image is all contained on the golfer's shirt.

Figure 4.8: Two images with similar Colour Histograms [92].

Swain and Ballard, detail how to identify objects using colour histograms [130] and Stricker and Swain, analyse information capacity and sensitivity using colour histograms [127].

4.5.5 Colour Coherence Vectors

Colour coherence vectors (CCV) consist of two histograms and are an approach for comparing images that incorporate spatial information. This measurement classifies pixels into coherent and incoherent regions and histograms are produced for each region [109]. Incoherent pixels are not part of a continuous region. Colour coherence vectors represent each colour within an image. The measure of the coherence within the images allows finer distinctions to be made than when using simple colour histograms [25].

Computing Colour Coherence Vectors

When computing a colour coherence vector, the first few stages are very similar to calculating a colour histogram. Average pixel values are used to replace pixel values in a small local area. The colour space is then discretized so that only a distinct number of colours are in the image. These pixels are then classified within a given colour bucket as either incoherent or coherent. Pixel groups are computed by determining connected components. Two pixels are connected, if one pixel is one of the eight neighbours of the other and is within the discrete colour bucket. As shown by Patrick Winston and Berthold Horn, only a single image pass is required to calculate connected components [155]. Pixels only belong to one connected component and are classified, coherent or incoherent, depending on whether they are larger or smaller than a certain size (τ).

An Example of Computing Colour Coherence Vectors

For the purpose of this example, $\tau = 5$ and it will be assumed that each of the three colour components have the same value. For example, a pixel with (R, G, B) value $(20, 20, 20)$ will have a single value of 20. It will be assumed that after the image has been slightly blurred some of the pixel intensities will have varied, Table 4.2. The colour space is then discretized into three buckets, so that the values 10-19 are in bucket 1, 20-29 bucket 2 and 30-39 bucket 3, Table 4.3. After they are discretized the values are then replaced with their corresponding bucket value, Table 4.4.

Next the connected components are calculated. Each group of connected components is assigned a letter, Table 4.5. A table is created of the letters associated with each of the discretized components, Table 4.6.

The pixels can then be classified as coherent and incoherent. Regions A and B are coherent as they have greater than τ pixels. The regions C and D are incoherent. Therefore

the Colour Coherence Vector can be calculated as, $\langle (15,4), (16,0), (0,1) \rangle$, Table 4.7.

Table 4.2: Computing Colour Coherence Vectors: Blurred Pixel Intensity Values

16	15	22	21	10	22
17	14	20	13	21	24
16	17	23	38	17	23
14	25	14	22	25	15
18	17	11	12	22	27
19	15	12	20	21	24

Table 4.3: Computing Colour Coherence Vectors: Discretized Buckets

Bucket	Colour Value
1	10 to 19
2	20 to 29
3	30 to 39

Table 4.4: Computing Colour Coherence Vectors: Discretized Bucket Values

1	1	2	2	1	2
1	1	2	1	2	2
1	1	2	3	1	2
1	2	1	2	2	1
1	1	1	1	2	2
1	1	1	2	2	2

Table 4.5: Computing Colour Coherence Vectors: Letter Assignment

A	A	B	B	C	B
A	A	B	C	B	B
A	A	B	D	C	B
A	B	A	B	B	C
A	A	A	A	B	B
A	A	A	B	B	B

Colour Coherence Vectors are a new method for comparing colour images that combines Colour Histograms with spatial information. It is very useful for searching for images in which a subset of an image appears.

Table 4.6: Computing Colour Coherence Vectors: Associated letters for the Discretized Components

Label	A	B	C	D
Colour	1	2	1	3
Size	15	16	4	1

Table 4.7: Computing Colour Coherence Vectors: Incoherent/Coherent Pixel Classification

Colour	1	2	3
α	15	16	0
β	4	0	1

Or

$$\langle (15, 4), (16, 0), (0, 1) \rangle$$

4.6 Current Monoscopic HVS Based Comparison Metrics

The pixel based metrics described in Section 4.5 are limited in their accuracy when assessing image quality. Due to this, much research has taken place into the development of more advanced HVS image quality assessment techniques. Both Daly [29] and Lubin [73] have produced metrics based on the early stages of human vision and these have been able to determine the presence of compression errors within different areas of an image. However, as these models produce varying thresholds for separate areas within the image, they are not able to accurately predict overall image quality [40]. This is because the human visual system is less sensitive to the peripheral areas of images. [65]

Osberger et al, [90] have developed a HVS based metric that weights areas of a monoscopic image to take into account factors known to influence visual attention. The metric uses Peli's Band Limited Contrast Algorithm [94] and shows a large improvement compared to PSNR, when compared to subjective human opinion.

Although large amounts of work have been undertaken on assessing image quality using human visual system metrics, very little work has been done on metrics for assessing stereoscopic image quality.

4.7   Stereoscopic Image Quality Metrics

There are currently only a small number of metrics for evaluating the quality of stereoscopic images. In this section we will detail some of the recent papers and metrics in the literature. Most of these existing stereo metrics are based upon current 2D image quality metrics, with modifications to account for stereoscopic 3D information within the images.

Lu et al., propose a metric for evaluating the quality of 3D asymmetric encoding of video. [72] Their metric is based upon a spatial frequency domain model and is a weighted sum of the differences between images and spatial frequencies between the degraded and original images.

A metric is proposed by Shao et al. to evaluate depth map image rendering of video and depth video. [117] It measures the colour and sharpness of edge distortion within the videos. For the distortion in colour measure, they compare the luminance difference between the rendered image and the reference. To calculate the sharpness of the distorted edge, they evaluate the proportion of the remaining edge to the original edge.

A framework for a full reference stereoscopic image quality metric is proposed by Boev et al. [17] This metric, called 3DQ, is composed of two components, a stereoscopic quality component and a monoscopic quality component. Perceived degradation of depth cues are assessed by the stereo component whilst the mono component looks at distortions caused by blur, noise and contrast changes. A block diagram of the metric is given in Figure 4.9.
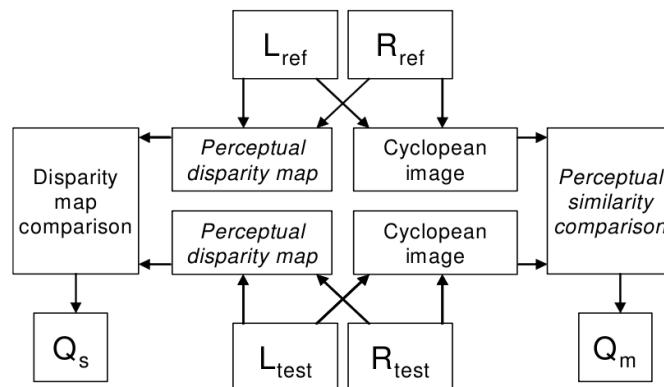
Figure 4.9: Block Diagram of the 3DQ Metric, Stereoscopic quality (Qs) and Monoscopic quality (Qm) for a stereo-pair [17].

Benoit et al. [9], propose tests using the ITU recommendations [1] and an extension

of a 2D metric involving measurement of disparity map distortion. It makes use of the existing SSIM metrics [23, 144].

Hewage et al., [46] undertake a comparison of subjective tests and 2D metrics, for video and 3D video created using depth maps. The results show that the metric correlates well with subjective viewer opinion, for perception of image quality and depth.

The work details a subjective evaluation of the effect of compression artefacts on perceived image quality, including overall image quality and depth perception. They conclude that whilst subjective image quality tests should remain the 'gold standard', use of a good objective assessment metric would provide an acceptable compromise for 3D image quality assessment.

4.8 Summary

This chapter details the background research into human vision, the eye, visual system characteristics and the models used to assess image quality. Existing image quality metrics are presented including, 2D human visual system based metrics, as well as developments is stereoscopic 3D metrics.

Existing metrics such as Mean Square Error and Peak Signal to Noise Ratio are currently most commonly used for the assessment of both monoscopic and stereoscopic image quality. However, new human visual system (HVS) based image quality metrics have been developed and have shown to perform well on monoscopic 2D images. Some work has been undertaken into HVS based stereoscopic image quality metrics, but a good quality 3D stereoscopic image quality metric that can compete with the results from image quality trial is still lacking.

Chapter 5

Investigating Symmetric and Asymmetric Compression Methods using Objective Image Quality Metrics

5.1 Introduction

The compression of the left and right components of a stereo image pair is a way to save valuable bandwidth when transmitting stereoscopic images [115]

In order to assess the effect of compression techniques on perceived image quality, objective measures such as Peak Signal to Noise Ratio (PSNR) are often used. The storage of stereoscopic images involves a large amount of data. Therefore, a substantial research effort is focused on understanding compression (such as JPEG and JPEG 2000) to obtain savings in storage capacity [115]. The same compression techniques used in 2D image material can be applied separately to the left and right views of a stereoscopic image pair. Image compression may compromise perceived image quality however, through a reduction of detail and the introduction of compression artefacts such as blur, blockiness, or ringing.

The experiments, detailed in this chapter will compare the two different stereo pair compression methods, symmetric and asymmetric encoding. Symmetric encoding of the image pair is when the left and right images are compressed by equal amounts, resulting in equal degradation. Asymmetric encoding is when the compression and therefore degra-

dation of the left and right images is unequal [76]. JPEG and JPEG 2000 will be used for the two encoding methods. The methods will be assessed using Peak Signal to Noise Ratio (PSNR) and the Sum of Absolute Differences (SAD) image quality metrics.

5.2   Motivation

Compression of the left and right views of a stereo image pair is a method to save valuable bandwidth and disk space when transmitting and storing stereoscopic images. Based on theories of binocular suppression, it is assumed that the binocular percept of a stereo image pair is dominated by the high quality component [67]. Thus, theoretically, when one image of the stereo pair is compressed such that a high quality is maintained, the other view can be coded more heavily without introducing visible artefacts in the binocular percept. The asymmetric compression concept was introduced by Perkins [96]. Asymmetric coding assumes that the binocular percept is not affected when one view is of high quality and the other view of lower quality. An example of asymmetric compression applied to the test image *Mannequin*, is shown in Figure 5.1.

Figure 5.1: Asymmetric Compression: *Mannequin*.

According to Seuntiens et al. [115], subjective tests show asymmetric compression of a JPEG stereo image pair, with the same file size, results in image quality somewhat below the average of the two symmetrically encoded image pairs for the right and left view. A large number of subjective studies, to investigate symmetric and asymmetric compression have been undertaken [113, 114, 125]. Further work is needed as the question of what to compress, one eye's view heavily (asymmetric compression), or both with equal amounts (symmetric compression) still remains unanswered [31, 120].

5.3   Hypothesis

Due to the high bandwidth requirements of stereoscopic 3D versus 2D, which of the two compression techniques, symmetric or asymmetric, produces a better level of stereoscopic image quality, for a fixed file size, is of great interest. There is conflicting evidence in the current literature, as to which of these compression methods produces a better result.

Due to this conflict, a null hypothesis is proposed, stating that both compression methods, symmetric and asymmetric, will produce the same image quality for a given file size. An investigation of the two methods will be undertaken using both JPEG and JPEG 2000 compression formats. The results from this investigation will be used to evaluate the hypothesis.

5.4   Experimental Method

5.4.1   Test Images

A sample of stereoscopic images was chosen. The following types of images were used, Photo-realistic computer generated: *Masha*; Photograph: *Mannequin* and Non-photo-realistic computer generated: *Perseus*.

The original images were in an uncompressed two-view side by side format. One view of each stereo pair is shown in Figures 5.2, 5.3 and 5.4.

Figure 5.2: *Masha.*   Figure 5.3: *Mannequin.*   Figure 5.4: *Perseus.*
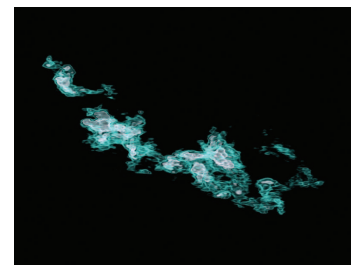
5.4.2   Image Compression Formats: JPEG and JPEG 2000

The compression methods JPEG and JPEG 2000 were chosen for this experiment. Both compression techniques, including the reason for their use, are detailed below.

JPEG

The JPEG compression method, a standardized image compression mechanism, was used in this experiment. It is designed to be used to compress either full-colour or grey-scale images of natural real world scenes. The Independent JPEG Group (IJG) scale was used, with a lower compression value resulting in a better quality compressed image.

JPEG 2000

JPEG 2000 is a wavelet-based image compression standard [6]. JPEG 2000 can operate at higher compression ratios without generating the characteristic 'blocky' artefacts of the original DCT-based JPEG standard. It also allows more sophisticated progressive downloads of images on websites. Meegan et al. state that blur but not blockiness is an acceptable form of monocular degradation when images are compressed asymmetrically [76]. Compression using JPEG 2000 produces blurring artefacts rather than blockiness, so to investigate this difference it was decided to repeat the experiments using this compression method. JPEG2000 was chosen as opposed to other wavelet based compression methods, as like JPEG, it provides compression settings from 1-99.

5.4.3 Stereoscopic Compression Methods: Symmetric and Asymmetric Compression

A program was produced to calculate the total file sizes for symmetrically and asymmetrically compressed image pairs for all possible JPEG and JPEG 2000 compression settings. A series of asymmetric pairs was produced for each symmetric pair, whilst maintaining a constant file size between the pairs. Images were compressed using full colour (24-bits). The design of the tool is shown in Table 5.1.

The asymmetric compression chosen was the one with the greatest difference in compression between the left and right views for each file size. For example, although compressing *Perseus*, with the left and right IJG compression values (40,40), (31,62) and (25,98) all result in the same file size (57Kb), the asymmetric compression of (25,98) was chosen as it has the greatest difference in the left and right compressions. The compressed stereo pairs were then compared back to the original using PSNR and SAD.

Table 5.1: Design of Asymmetric/Symmetric Compression Tool

		Left	Image	**A**	**B**	**C**	..	**CT**	**CU**
Right	Compression	Setting	1	2	3	..	98	99
Image	Setting	Size	X1	X2	X3	..	X98	X99
1	1	Y1	**X1 Y1**	X2 Y1	X3 Y1	..	X98 Y1	X99 Y1
2	2	Y2	X1 Y2	**X2 Y2**	X3 Y2	..	X98 Y2	X99 Y2
3	3	Y3	X1 Y3	X2 Y3	**X3 Y3**	..	X98 Y3	X99 Y3
:	:	:	:	:	:		:	:
98	98	Y98	X1 Y98	X2 Y98	X3 Y98	..	**X98 Y98**	X99 Y98
99	99	Y99	X1 Y99	X2 Y99	X3 Y99	..	X98 Y99	**X99 Y99**

5.4.4   Image Comparison Methods

Peak Signal to Noise Ratio (PSNR)

A tool was coded to compare stereo image pairs. It reads images from separate files, comparing them by calculating the Peak Signal to Noise Ratio (PSNR) of the differences in the two images [102]. PSNR is calculated from a comparison based on the difference in pixels, it is an objective image quality metric. First the Mean Square Error (MSE) is calculated for two images. One of the images is usually an approximation or compressed image of the first. MSE is defined as:

$$MSE = \frac{1}{mn} \sum_{x=0}^{m-l} \sum_{y=0}^{n-l} \parallel I(x,y) - I'(x,y) \parallel^2 \tag{5.1}$$

where $I(x,y)$ is the pixel values for the original image; $I'(x,y)$ is the compressed version, and m and n are the dimensions of the images. The Peak Signal to Noise Ratio (PSNR) is then defined as:

$$PSNR = 10 \cdot log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \tag{5.2}$$

where MAX_I is the maximum pixel value of the image. PSNR is applied across each stereo image pair as a whole.

The efficiency of compression is typically measured using the peak signal to noise ratio, or PSNR. The higher the PSNR value of a compression scheme the better. Using PSNR to compare images to their original uncompressed equivalent takes into account the quality retained by the compression method, as well as how much the image was compressed. It

is also the only generally accepted objective quality metric and is routinely used in JPEG and MPEG-4 comparisons [8].

For these reasons, it was decided to use PSNR as the image comparison metric in this experiment.

Sum of Absolute Differences (SAD)

The Sum of Absolute Differences (SAD) is a widely used simple image and video quality metric. It is mainly used in block matching in adjacent frames in video. It looks at the absolute value of the difference in each pixel in the two frames. The differences are then summed to create a simple metric.

The equation for Sum of Absolute Differences (SAD) is,

$$SAD = \frac{\sum_{x=0}^{m-1}\sum_{y=0}^{n-1} \parallel I(x,y) - I'(x,y) \parallel}{m \times n \times 255} \times 100\% \text{ (For an 8 bit image)} \qquad (5.3)$$

In these experiments the metric will be used to evaluate the difference between the compressed image and the uncompressed original.

5.5   Results and Analysis

5.5.1   JPEG Symmetric and Asymmetric Compression using PSNR

The PSNR results are presented from the symmetric and asymmetric JPEG compression experiments. The three test images were compressed both symmetrically and asymmetrically and a constant file size maintained between each image pair compared. As both symmetric and asymmetric JPEG image compression increased, both the file size and overall PSNR image quality reduced.

The compressed images were compared back to the originals using the coding PSNR measurement tool. The PSNR results for all three JPEG compressed images, *Masha*, *Mannequin* and *Perseus*, are displayed in Table A.1.

Figures 5.5, 5.6 and 5.7 show that symmetric JPEG compression out performed the asymmetric compression in nearly every case. The only exception to this was for *Masha*, when the image was compressed to the most extreme levels. Also there is only a very small difference in the two compression methods for this case. For the symmetric image

the left and right views are both compressed with a JPEG compression value of 99, whilst the asymmetric image is compressed with 98 and 99. Both images have the same file size.

The results also show that in the majority of cases, the greater the difference in the left and right view compression amounts, the worse the PSNR image quality value and thus the worse the quality of the compressed image in comparison to the original.
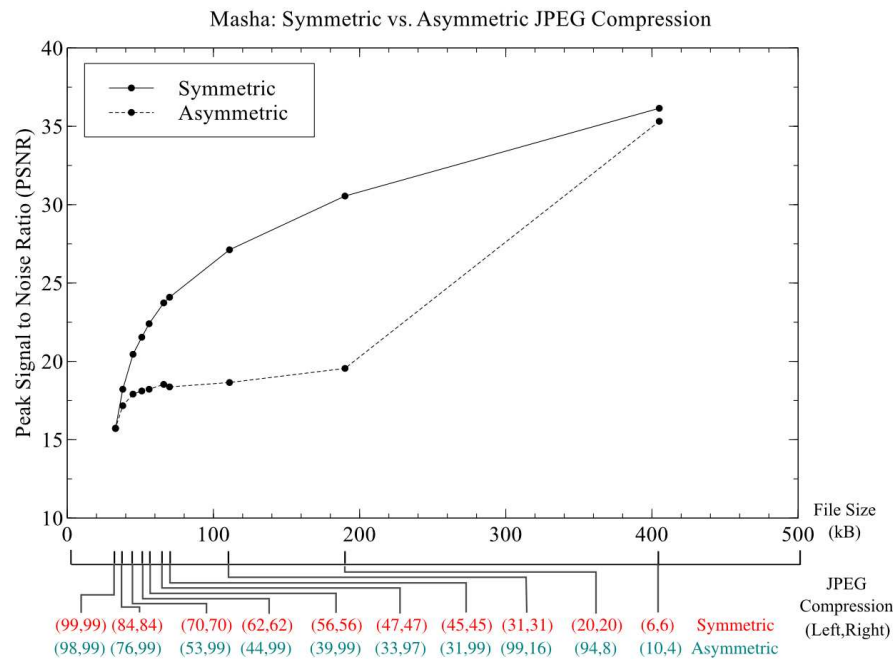
Figure 5.5: Symmetric vs. Asymmetric JPEG Image Compression, for the test images *Masha*.

5.5.2 JPEG 2000 Symmetric and Asymmetric Compression using PSNR

As with JPEG compression, again the images were compressed symmetrically and asymmetrically, with constant files sizes being maintained between each of the pairs. Similarly, file size and overall image quality reduced with increasing JPEG 2000 compression. PSNR was again used for the comparison and the results for the JPEG 2000 compressed images, *Masha*, *Mannequin* and *Perseus*, are shown in Table A.2.

Figures 5.8, 5.9 and 5.10 show that symmetric JPEG 2000 compression out performed the most extreme asymmetric compression in every case. There were only small differences in quality, for *Masha*, when extreme levels of compression were applied. However, at these extreme compression levels, there was only a small difference in the compression applied

to the left and right views within the pair.

The results again suggest that the greater the difference in the compression between the left and right images within the pair, the worse the overall PSNR image quality value and the greater the difference in quality between the compressed image and the original.

5.5.3   JPEG Symmetric and Asymmetric Compression using SAD

This section details the results from Sum of Absolute Differences assessment of the symmetric and asymmetric compression applied to the test image *Masha*. In every case there was a greater difference for the asymmetrically compressed image pairs. The results showed that the more compression that was applied to the image, the greater the difference between the two methods. As with the PSNR experiments the results suggest that the symmetric compression method should be used for stereoscopic image compression. The Sum of Absolute Differences results for the test image *Masha* are shown in Table A.3 and displayed in Figure 5.11.

The results show that as the file size decreases, and the compression increases, the sum of absolute differences increases dramatically for asymmetric compression. The differences in SAD results for the symmetrically compressed image are much lower but SAD still increases with decreasing file size. This suggest that there are much fewer absolute differences between the original and symmetrically compressed image.

5.5.4   Evaluation Issues and Limitations

Although PSNR is one of the most commonly used measurements of the quality of reconstruction in image compression, probably due to the simplicity of its computation, it has some limitations and does not provided an ideal measure for stereoscopic image quality. PSNR is a measure of how close a likeness the compressed image is to the original. It is a pixel based comparison method and does not consider whether or not the resulting compressed image produces a stereoscopic effect or the quality of any effect produced.

JPEG 2000 was used throughout the experiments within this chapter to compare images produced with blurring artefacts as opposed to those blocking artefacts produced when compressing with JPEG. The images compressed using JPEG 2000 produced better results that those compressed using JPEG, when using PSNR. However, JPEG 2000 has one large limitation when compressing images; there are still issues of compatibility relating

to its use in most browsers. Many browsers simply do not support JPEG 2000.

The experimental results showed that PSNR has some limitations and does not provided an ideal measure for stereoscopic image quality. The experiment with the *Masha* test image was therefore repeated using the new metric, Sum of Absolute Differences. The results from this repeated experiment, showed that SAD has similar limitations to PSNR when used to assess stereoscopic images and therefore is not an ideal measure for stereoscopic image quality.

5.6 Conclusions

This section aims to summarise the work accomplished and the findings of the experiments performed in this section, by drawing conclusions from the examination of the results. The section finishes by looking into necessary further work.

One of the main aims for this work was to investigate and compare two possible compression methods to try and reduce the required file size and bandwidth for the storage and transmission of stereoscopic images.

This chapter looks at two compression methods, symmetric and asymmetric encoding, and concludes that symmetric compression produced the better results. The uncompressed images used are ground truths and the aim was to compress the images so that they are as close a match to the original as possible. It is concluded that for both, JPEG and JPEG 2000 stereoscopic image encoding, symmetric compression, as opposed to asymmetric compression, should be used for all image types. The same conclusion is drawn when using either of the two image comparison methods, PSNR and SAD.

5.7 Future Work

Further investigation should be undertaken in two areas, a subjective evaluation of the symmetric and asymmetric compression methods and the development of a new image quality metric that has a better correlation with subjective human opinion than existing metrics such as PSNR.

Subjective image quality studies should be undertaken to evaluate the two different compression methods, symmetric and asymmetric compression. File sizes in the asymmetric trial should remain equal to those in the symmetric trial to allow a direct comparison between the methods. Like existing stereoscopic image quality studies, the trials should

be based on the Double Stimulus Continuous Quality Scale (DSCQS) from the ITU-R BT.500-11 recommendations [1].

Given that experiments in existing literature have shown that both PSNR and SAD do not seem to correlate well with human subjective opinion [51, 96], it is felt that work should be done on the development of a new Human Visual System (HVS) based image quality metric [91]. This metric should take into account changes between the left and right views of the stereo pair and how these changes affect the overall stereoscopic image quality. Any new metric developed should be evaluated against PSNR. For it to be classed as a success, it should be a better predictor of human image quality preference than PSNR and ideally be able to predict the maximum amount of compression that can be applied to a stereoscopic image pair before unacceptable degradation occurs.
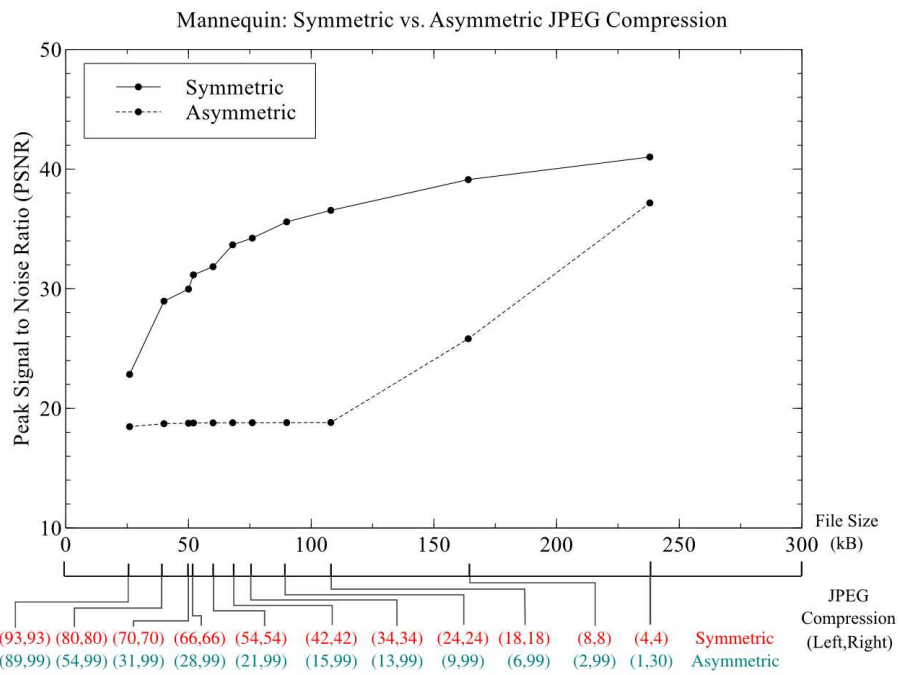
Figure 5.6: Symmetric vs. Asymmetric JPEG Image Compression, for the test images *Mannequin*.
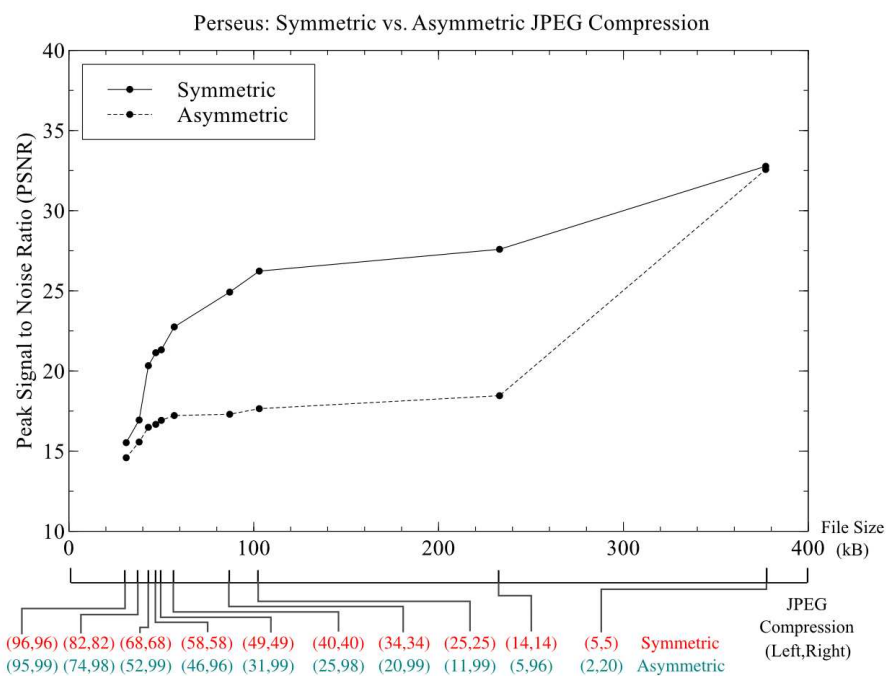
Figure 5.7: Symmetric vs. Asymmetric JPEG Image Compression, for the test images *Perseus*.

Masha: Symmetric vs. Asymmetric JPEG2000 Compression
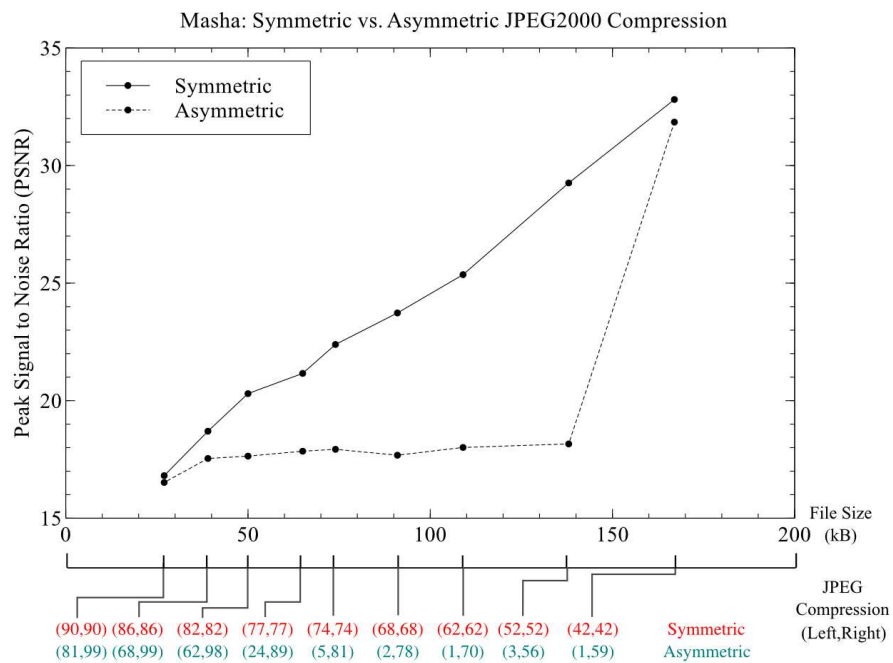
Figure 5.8: Symmetric vs. Asymmetric JPEG 2000 Image Compression, for the test images *Masha*.

Mannequin: Symmetric vs. Asymmetric JPEG2000 Compression
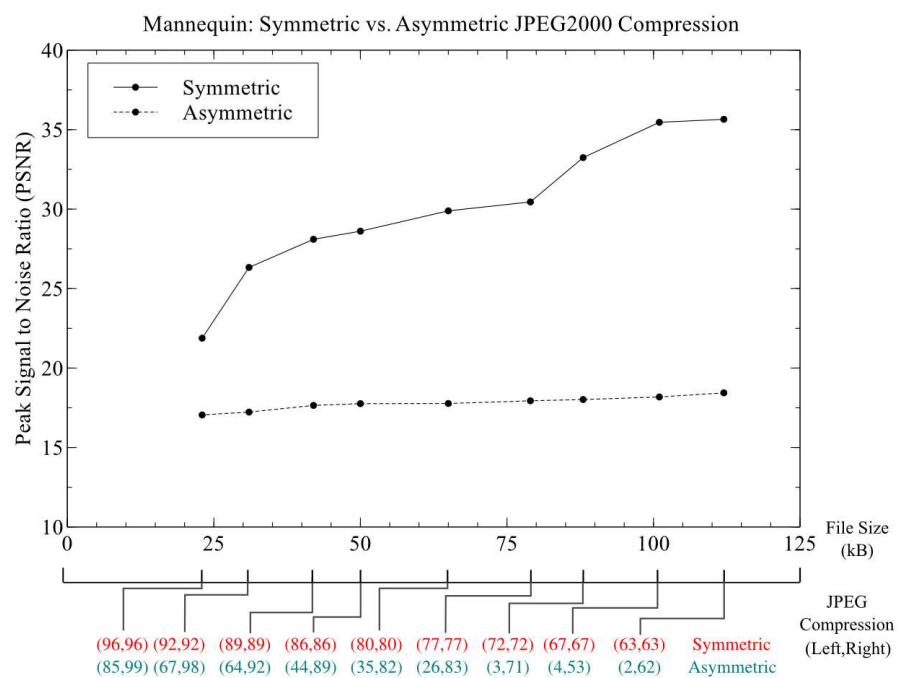
Figure 5.9: Symmetric vs. Asymmetric JPEG 2000 Image Compression, for the test images *Mannequin*.
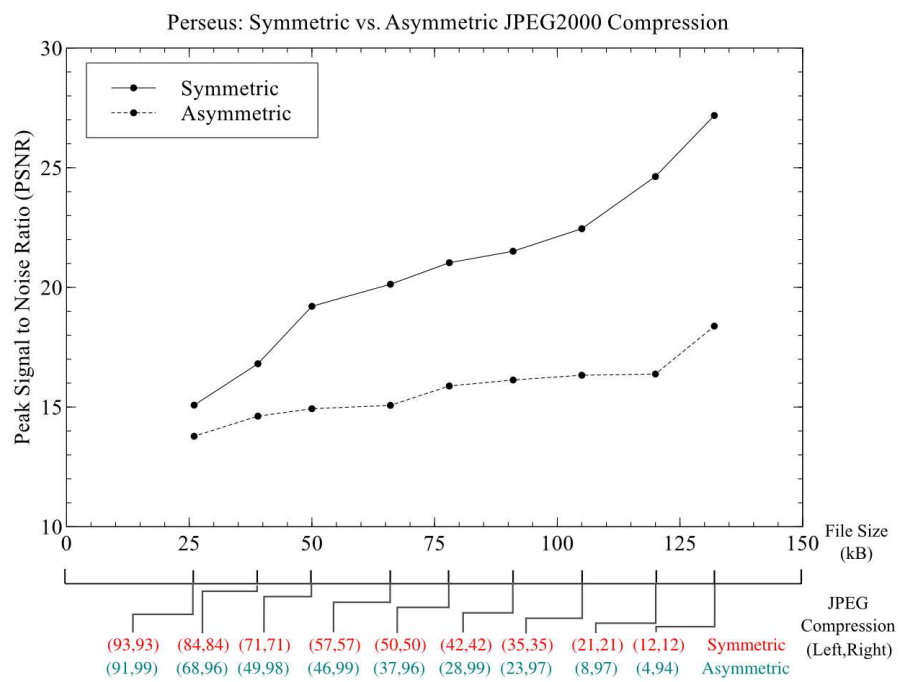
Figure 5.10: Symmetric vs. Asymmetric JPEG 2000 Image Compression, for the test images *Perseus*.
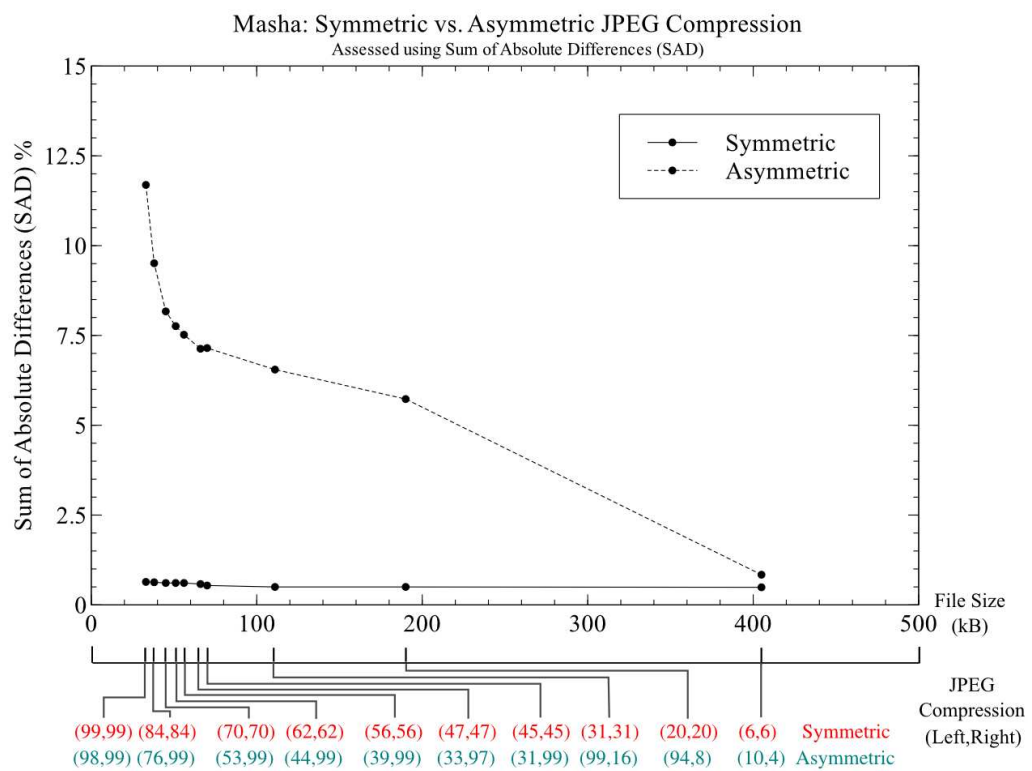
Figure 5.11: SAD Evaluation of Symmetric and Asymmetric Compression Methods for *Masha*. The symmetrically compressed image has much lower differences in SAD, but SAD still increases with decreasing file size.

Chapter 6

Subjective Evaluation of Symmetric and Asymmetric Stereoscopic Image Compression

6.1   Introduction

In this chapter, symmetric and asymmetric stereoscopic image compression techniques are investigated using subjective quality measures. In experiments, detailed in Chapter 5, how Peak Signal to Noise Ratio (PSNR) measures the quality of varyingly coded stereoscopic image pairs was evaluated. The results suggest that, when measuring image quality using PSNR, symmetric, as opposed to asymmetric stereo image compression, will produce significantly better results.

As PSNR measures of image quality are widely criticized for having low correlation with human perception [10, 86], this chapter evaluates symmetric and asymmetric stereoscopic image compression. Subjective trials using the Double Stimulus Continuous Quality Scale (DSCQS) from the ITU-R BT.500-11 recommendation are designed and run. The observers were asked to assess the image quality and the trials separately measured symmetric and asymmetric stereoscopic image encoding quality. Consistent file sizes were maintained between the symmetric and asymmetrically compressed image pairs.

From the results of the symmetric subjective experiment, a stereoscopic image quality threshold is calculated. This provides a recommendation as to the amount of JPEG compression that could be applied across each stereo image before noticeable degradation

is perceived.

The results from the symmetric and asymmetric trials showed that, as expected, increasing the overall compression level resulted in a reduced image quality. For the asymmetric trial it was also noted that increasing the difference in the compression between the left and right views, resulted in a reduced image quality. It is concluded that in general, for stereoscopic image compression, using JPEG, a symmetric as opposed to asymmetric compression approach across the left and right images of the pair should be used.

6.2 Motivation

Additional space and bandwidth is required for storage and transmission of stereoscopic images. The space required to store a stereoscopic image is normally twice the amount need to store a monoscopic one. One method of reducing the storage space is to compress the images. Therefore, substantial research effort has been focussed on digital image compression, using, for example, JPEG [115] to obtain bandwidth and storage capacity savings. However, the question of how much compression to apply and the approach to adopt in applying the compression to the image pair remains unresolved.

Traditionally, to assess the effect of different compression methods on perceived image quality, 2D objective measures such as Peak Signal to Noise Ratio (PSNR) or Mean Squared Error (MSE) are used. In earlier investigations, experiments were run to evaluate the quality of symmetric and asymmetric coding using PSNR, Chapter 5.

The results strongly suggested that symmetric coding should be used for compressing stereo image pairs, however, the PSNR metric does not correlate well with subjective judgement of image quality [86].

Existing research shows that both symmetric and asymmetric compression methods are able to significantly reduce the size of the stereoscopic content. However, there is an inconsistency within current research, with different papers suggesting that both symmetric and asymmetric compression methods produce better results [31, 120]. Therefore, further investigation into these two compression methods will be undertaken, using subjective human trials.

6.3 Hypothesis

The prediction is that, for both symmetric and asymmetric compression, the subjective quality of the compressed images from the subjective human based trial will decrease with increased JPEG compression.

It is expected that an image quality threshold will be calculated from the symmetric trial. The prediction is that symmetric compression will out perform asymmetric compression for image compressions below the image quality threshold calculated from the results of this first trial. To evaluate this prediction the results of, this second asymmetric human quality preference trial will be compared with the findings of the first symmetric human trial. File sizes are kept equal in both trials.

6.4 Experimental Method

The experimental method followed was based on the ITU-R recommendation 500. Two trials will be undertaken, the first to assess the quality of symmetrically compressed images and calculate an image quality threshold for these images. The second trial will evaluate asymmetrically compressed images in comparison to the symmetric equivalent. File sizes of the compressed images will be kept constant in both trials to allow a comparison of the two compression methods.

The trials assessed the quality of both symmetrically and asymmetrically compressed stereoscopic images, relative to uncompressed originals. The Double-Stimulus Continuous Quality-Scale (DSCQS) method for stereoscopic image evaluation was followed.

This (DSCQS) method was cyclic, in that the assessors viewed a pair of 3D stereo pictures of the same image, one compressed, and the other the uncompressed original and were asked to assess the quality of both. In line with the ITU-R recommendations, each session lasted no longer than half an hour.

In the first set of subjective human trials, the assessors were presented with a series of symmetrically compressed stereo image pairs displayed in a random order, with compression amounts covering all required combinations. The order the images were displayed was random and different for each subject. The left and right images within the stereo pair were compressed by equal amounts. Following these first sessions, the mean scores for each test condition and picture were calculated.

In the second set of subjective human trials, the assessors were presented with a series

of asymmetrically compressed stereo image pairs, again these were in a random order and were displayed randomly for each subject. In these images the left and right views are compressed with varying amounts of JPEG compression. To enable accurate comparison of these compression approaches, the file sizes were chosen to correspond with those in the earlier symmetric human trial.

The results from this asymmetric human trial were compared to the findings of the symmetric trial to evaluate which method produces the more preferable results.

6.4.1   Equipment and Viewing Conditions

A full resolution auto-stereoscopic Kodak Stereo Imaging [27] display was used for viewing the images. This allows three dimensional images to be viewed without special glasses, displays a full resolution view to each eye and has zero cross-talk.

It provides double the number of pixels of an equivalent 2D monitor, has a 45 by 36 degree field-of-view and a resolution of 1280x1024 pixels, creating a virtual image from the two displays while the user views the final image in three dimensions through two 32 mm apertures. Although special glasses are not required, the user must sit in a particular place to see the 3D effect. A 17 inch IBM LCD monitor is used for the image quality scoring screen, consistent for both trials.

The two displays were driven independently, but using the same type of graphic card (nVidia Quadro FX family) and same software driver (nVidia ForceWare Release 80). The experiment was conducted in a darkened room, with constant minimal light levels and with equipment arranged as shown in Figure 6.1.

The equipment and viewing conditions were consistent in both the symmetric and asymmetric human preference trials.

6.4.2   Test Images

The experiment was performed with three stereoscopic test image pairs, *Masha*, *Mannequin* and *Perseus*, Figures 6.3, 6.4 and 6.5. These images were the same as those used in the objective trial detailed in Chapter 5. These image types correspond to computer generated photorealistic, a stereoscopic photograph and computer generated non-photorealistic. In both trials *Masha* was shown 45 times, *Mannequin* and *Perseus* 43 times each in-line with ITU-R Recommendations.

For the symmetric trial, each was assessed with levels of compression 5 to 100 in steps

Figure 6.1: The environment used, reflections of objects or lights behind participants were eliminated.

of 5 in comparison to the uncompressed original. In the asymmetric trial, the test images were compressed asymmetrically with varying differences between the left and right views.

A series of asymmetric pairs were created for file sizes corresponding to the symmetric pair. For example, the asymmetric (Left,Right) compressions, (100,8) & (80,11) & (60,14) & (41,18) & (9,100) & (11,80) & (14,61) & (18,40) for $Masha$ at a file size of 139kB, correspond to symmetrically compressing the image at (25,25).

The images were compressed such that file sizes were consistent to those in the first symmetric trial. This was to enable the two compression methods, symmetric and asymmetric to be accurately compared. Tables 6.1, 6.2 and 6.3 contain all the asymmetric compression settings and the equivalent file sizes and symmetric compression amounts used for the three images.

Table 6.1: $Masha$: Asymmetric Compression Settings.

File	Symmetric	Left Compression				Right Compression			
Size	Comp	Constant				Constant			
(kB)	(JPEG)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)
139	(25,25)	(100,8)	(80,11)	(60,14)	(41,18)	(9,100)	(11,80)	(14,61)	(18,40)
126	(30,30)	(100,11)	(80,14)	(60,18)	(41,23)	(10,100)	(14,80)	(18,61)	(23,40)
94	(50,50)	(100,18)	(80,26)	(60,38)	(41,55)	(19,100)	(27,80)	(37,61)	(57,40)
76	(65,65)	(100,26)	(80,44)	(60,66)	(41,82)	(28,100)	(44,80)	(64,61)	(83,40)
62	(75,75)	(100,40)	(80,67)	(60,86)	(41,98)	(41,100)	(67,80)	(84,61)	(97,40)
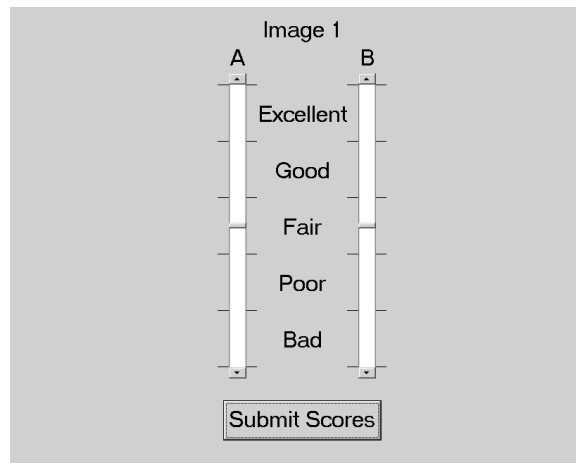
Figure 6.2: Scoring scale used in the results collection. The underlying 0-100 values are not shown.

Figure 6.3: *Masha.*

Figure 6.4: *Mannequin.*
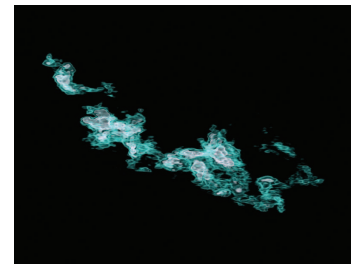
Figure 6.5: *Perseus.*

6.4.3 Participants

For the symmetric trial a total of 20 candidates (16 male, 4 female) were recruited within the Durham University population. Ages varied from 18 to 54 with a mean of 23 years.

In the asymmetric trial, a total of 20 candidates (15 male, 5 female) were recruited within the Durham University population. Ages varied from 18 to 28 with a mean of 22 years.

In both trials all participants were non-expert, in that they were not directly concerned with image quality in their normal work, and were not experienced image assessors. Participants were not aware of the purpose of the experiment, or that one of the images was uncompressed. They received a nominal payment of five pounds.

Table 6.2: *Mannequin*: Asymmetric Compression Settings.

File Size	Symmetric Comp	Left Compression Constant				Right Compression Constant			
(kB)	(JPEG)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)
185	(45,45)	(100,17)	(91,23)	(79,29)	(70,31)	(16,100)	(17,96)	(26,79)	(31,71)
179	(50,50)	(100,19)	(91,25)	(79,31)	(70,32)	(17,100)	(18,96)	(28,79)	(34,71)
111	(75,75)	(100,39)	(91,58)	(79,75)	(70,84)	(42,100)	(45,96)	(76,79)	(84,71)
98	(50,80)	(100,57)	(91,68)	(79,84)	(70,90)	(53,100)	(57,96)	(84,79)	(90,71)
80	(85,85)	(100,69)	(91,80)	(79,92)	(70,97)	(69,100)	(71,96)	(92,79)	(97,71)

Table 6.3: *Perseus*: Asymmetric Compression Settings.

File Size	Symmetric Comp	Left Compression Constant				Right Compression Constant			
(kB)	(JPEG)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)	(L,R)
62	(40,40)	(100,24)	(85,31)	(70,42)	(56,57)	(24,100)	(30,86)	(41,71)	(54,57)
58	(50,50)	(100,28)	(85,37)	(70,53)	(56,68)	(28,100)	(36,86)	(51,71)	(66,57)
54	(60,60)	(100,33)	(85,46)	(70,64)	(56,77)	(33,100)	(45,86)	(63,71)	(77,57)
49	(70,70)	(100,43)	(85,60)	(70,76)	(56,88)	(43,100)	(59,86)	(75,71)	(88,57)
44	(80,80)	(100,55)	(85,73)	(70,88)	(56,100)	(75,100)	(72,86)	(87,71)	(100,57)

6.4.4 Protocol

Only volunteers meeting the minimum criteria of acuity of 20:30 vision, stereo-acuity at 40 sec-arc and passing the colour vision test, were used in the experiments. Prior to the start of the experiments, candidates received instructions and completed a practice trial of the Kodak Stereo Display. This contained five sets of stereo images viewed and rated in the same ways as for those in the trial, but these practise trials were not included in the experimental analysis.

The participants then completed the 131 experimental trials in individual sessions. Participants were advised not to switch between the two images more that 3-4 times, however no restriction was enforced. Participants were requested to be as accurate as possible in judging image quality but not to spend too long on each image set, although no time limit was imposed. An additional display showed image quality ranking sliders and participants were asked to record image quality results, for each pair shown, using this screen. Answers could not be changed and once image quality scores were recorded and submitted, the next image comparison screens were displayed.

After the experiments, all participants were debriefed and given a chance to ask ques-

tions. The three vision tests and the practise trial took about 15 minutes and the experiment lasted half an hour including small breaks after each image type.

6.4.5   Grading Scale

In each trial the images were rated on a sliding scale of, Excellent, Good, Fair, Poor, Bad. Figure 6.2 shows the grading scale. Participants were asked to assess the overall picture quality of each stereo pair by marking on a vertical scale. Ten centimetre vertical scales were displayed on the screen in pairs to accommodate the double presentation of each test picture. These provided a continuous rating system to avoid quantising errors, but were divided into five equal lengths corresponding to the normal ITU-R five-point quality scale. The associated terms categorising the different levels were the same as those normally used in the ITU-R recommendation 500; but here they were included for general guidance. Results were recorded by moving a vertical slider to the desired position along the scale. The 0 to 100 values were not shown to the observer. The absolute difference score between the compressed and uncompressed original image was calculated in each case.

6.5   Results and Analysis

6.5.1   Symmetric Human Trial

The results for the first subjective human trial, using symmetric images, are presented. Initially each image file size (kB) was calculated for the JPEG compression settings, 5-95 in intervals of 5 and maximum compression.

The mean score and standard deviation error were calculated for the results from the human trial. The 95% confidence interval was calculated from the results, to provide a reliable maximum and minimum grade series for each of the images. Data for each of the images was then subjected to a t-test, comparing the minimum and maximum compression applied to each of the images, to check that the compression had a statistically significant effect on perceived image quality. Data collected from the subjective human trial was subjected to one-way Analysis of Variance (ANOVA), with compression as the within-subject independent variable and score as the dependant variable. The ANOVA was performed to evaluate whether there is evidence that the means for each compression differ significantly.

In terms of investigating an image quality threshold, which of the means are different needs to be evaluated and therefore the data was subjected to a Tukey Multiple Comparison Test [16], comparing the difference between each pair of means with adjustment for multiple testing. From this, a threshold for each image was established.

The image quality threshold is calculated as the point where the compressed images become statistically different from the uncompressed original. Images above this quality threshold are perceived as statistically the same as the uncompressed original.

The results from the trial are considered in terms of the difference between the ranking scores for two images shown in each case. The greater the difference in the rating the worse the compressed image was perceived in relation to the uncompressed original. Unsurprisingly, overall the results showed perceived image quality reduced with compression for all three images.

Masha

Figure 6.6 shows the mean score and standard deviation from Table C.1 for each compression setting for the test image $Masha$. It was important to confirm the prediction that the perceived image quality from the subjective human based trial decreased with increased JPEG compression. The results of the t-test between maximum and minimum compression did show that there was a 0% probability that they were statistically the same. The results from the one-way ANOVA revealed that there was a significant effect of compression on perceived image quality (F value = 85.094 and p value = 0.000). If no difference in perceived image quality compared with compression had been found, this may have indicated problems with participants' viewing position during the trial or a display problem such as an optical or mechanical misalignment.

The results from the Tukey multiple comparison test showed that the scores up to and including JPEG compression setting 30 were statistically the same. Therefore, the image quality threshold for $Masha$ is JPEG compression level 30.

The Tukey comparison returned a value of 0.996 probability that the image compressed at 30 was statistically the same as the uncompressed original. Compressing up to and including this setting results in images that are statistically perceived to be the same. For $Masha$, compressing to this level gives a 64% reduction in the stereo image file size.

Figure 6.7 shows the mean score and maximum and minimum grade series calculated from the 95% confidence interval plotted against image file size (kB). Marked on the graph

is the calculated image compression threshold, where statistically, the subjects perceived the images become different from the original.
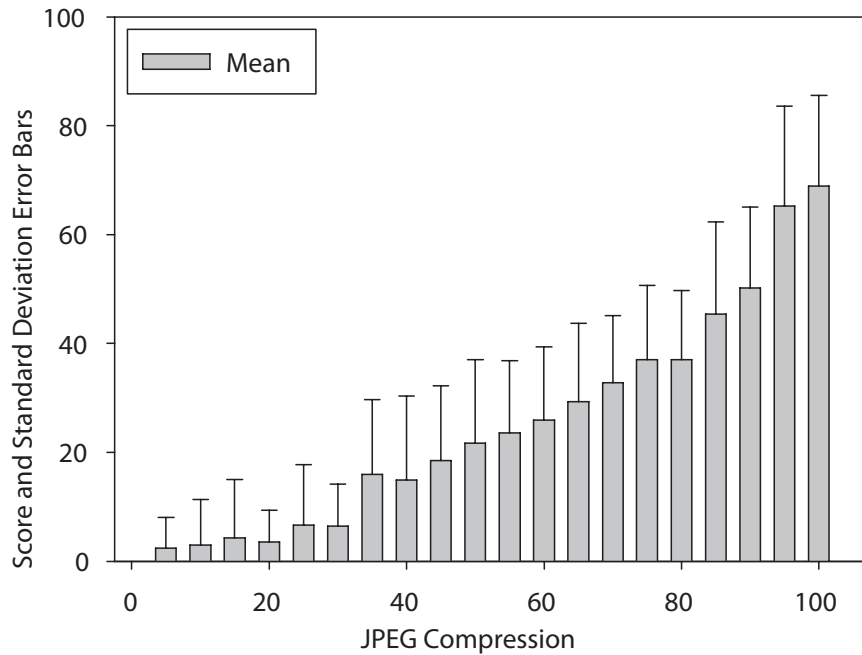
Figure 6.6: *Masha*: Mean Perceived Absolute Difference Score & Standard Deviation Error Bars.

Mannequin

Figure 6.8 shows the mean score and standard deviation from Table C.2 for each compression setting for the test image *Mannequin*. Again the t-test between maximum and minimum compression was used to confirm the prediction for *Mannequin*, that perceived image quality from the subjective human based trial decreased with increased JPEG compression, with the results showing that, there was a 0% probability that they were statistically the same. Results from the one-way ANOVA revealed that there was a significant effect of compression on perceived image quality (F value $= 68.788$ and p value $= 0.000$).

The results from the Tukey multiple comparison test showed that the scores up to and including JPEG compression setting 50 were statistically the same and returned a value of 0.210 probability that the image compressed at 50 was statistically the same as the uncompressed original. Compressing up to and including this image quality threshold results in images that are statistically perceived to be the same. For *Mannequin*, compressing to this level gives a 85% reduction in the stereo image file size.
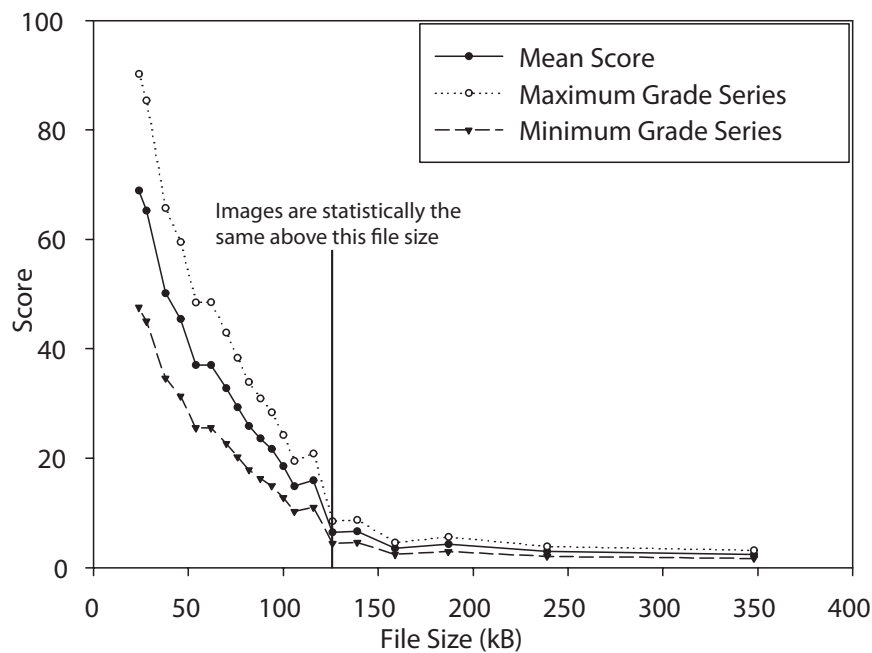
Figure 6.7: *Masha*: Perceived Mean & Maximum and Minimum Grade Series Calculated from the 95% Confidence Interval.

Figure 6.9 shows the mean score and maximum and minimum grade series calculated from the 95% confidence interval plotted against image file size (kB). Marked on the graph is the calculated image compression threshold, where statistically, the subjects perceived the images become different from the original.

Perseus

Figure 6.10 shows the mean score and standard deviation from Table C.3 for each compression setting for the test image *Perseus*. To confirm the prediction that again, the perceived image quality from the subjective human based trial decreased with increased JPEG compression, a t-test between maximum and minimum compression was used and again, showed there was a 0% probability that they were statistically the same. Results from the one-way ANOVA revealed that there was a significant effect of compression on perceived image quality (F value = 57.827 and p value = 0.000).

The results from the Tukey multiple comparison test showed that the scores up to and including JPEG compression setting 60, were statistically the same. The Tukey comparison returned a value of 0.158 probability that the image compressed at 60 was
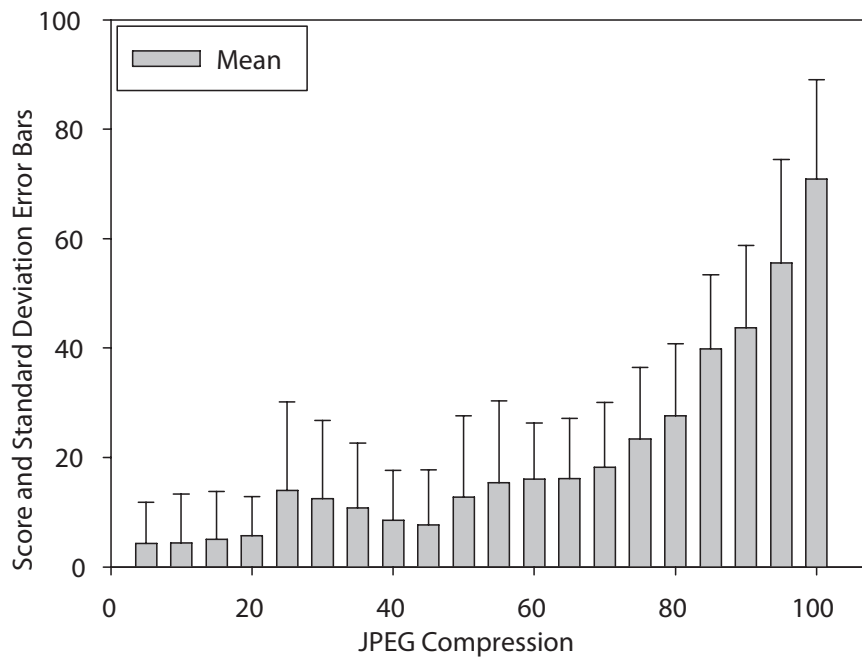
Figure 6.8: *Mannequin*: Mean Perceived Absolute Difference Score & Standard Deviation Error Bars.

statistically the same as the uncompressed original. Compressing up to and including this setting results in images that are statistically perceived to be the same. For *Perseus*, compressing to this level gives a 62% reduction in the stereo image file size.

Figure 6.11 shows the mean score and maximum and minimum grade series calculated from the 95% confidence interval plotted against image file size (kB). Marked on the graph is the calculated image compression threshold, where statistically the subjects perceived the images become different from the original.

6.5.2 Asymmetric Human Trial

The results from the investigation of asymmetric compression using a subjective human trial are presented. Initially appropriate asymmetric compression amounts, equivalent to those used in the first subjective trial for symmetric compression were calculated. Each of the three images were then compressed asymmetrically to produce asymmetric images of all the relevant files sizes. In each case symmetric compression was compared to a series of varying compressed asymmetric image pairs, all with a constant file size. The file sizes were consistent with those used in the symmetric trial.

The mean score and standard deviation were then calculated for each of the results
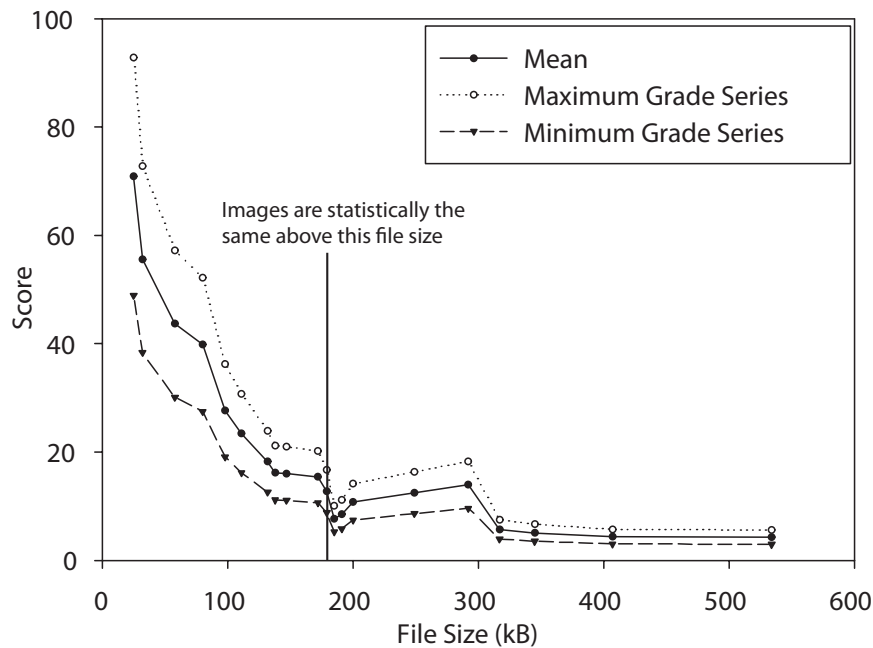
Figure 6.9: *Mannequin*: Perceived Mean & Maximum and Minimum Grade Series Calculated from the 95% Confidence Interval.

from the human trial. The 95% confidence interval was then calculated, this provided a reliable maximum and minimum grade series for each of the images and equivalent file sizes. A t-test was performed between the results of the symmetric experiment and those of the asymmetric compression that produced the best perceived image quality. The t-tests were performed to evaluate if there was any significant difference between the perceived quality of the symmetric compressed pair and the best asymmetric equivalent.

Data collected from the subjective human trial was subjected to one-way Analysis of Variance (ANOVA), with the Asymmetric Difference between the left and right views as the within-subject independent variable and score as the dependant variable. The ANOVA was performed to evaluate whether there is evidence that the means for each compression differ significantly and whether the overall trend suggested an increase or decease in image quality with an increasing symmetric difference.

The results from the trial are considered in terms of the difference between the ranking scores for two images shown in each case. The greater the difference in the rating, the worse the compressed image was perceived in relation to the uncompressed original. The results of both symmetric and asymmetric compression of two-view stereo images are compared. For asymmetric compression, the general trend was, the greater the difference
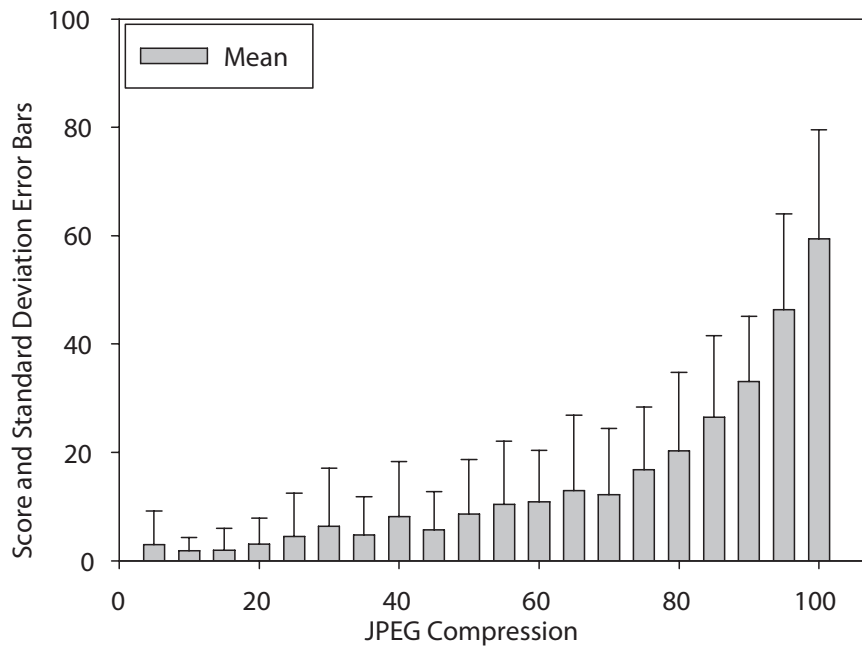
Figure 6.10: *Perseus*: Mean Perceived Absolute Difference Score & Standard Deviation Error Bars.

in the amount of compression applied to the left and right view, the poorer the perceived quality. For each of the three images the results from the asymmetric trial confirmed the prediction that the perceived image quality from the subjective human based trial decreased with an increase in asymmetric image compression. Tables B.4, B.6 and B.8 contain the results from the asymmetric trial.

Masha

Figure 6.12 shows the mean score from Table B.4, for each of the file sizes and symmetric and asymmetric compression settings for the test image, *Masha*. It was important to confirm the prediction that, as with symmetric compression, the perceived image quality from the subjective human based trial decreased with increased asymmetric JPEG compression. The results of the one-way ANOVA's showed there was a significant effect in that increasing the difference in compression of the left and right images within a stereo pair resulted in a decrease in the overall perceived image quality (F value = 489.37, p value = 0.0061). If it had been found that there was no difference in perceived image quality for varying amounts of asymmetric compression, this may have indicated that there was a problem with the participants' viewing position within the trial or a display problem such
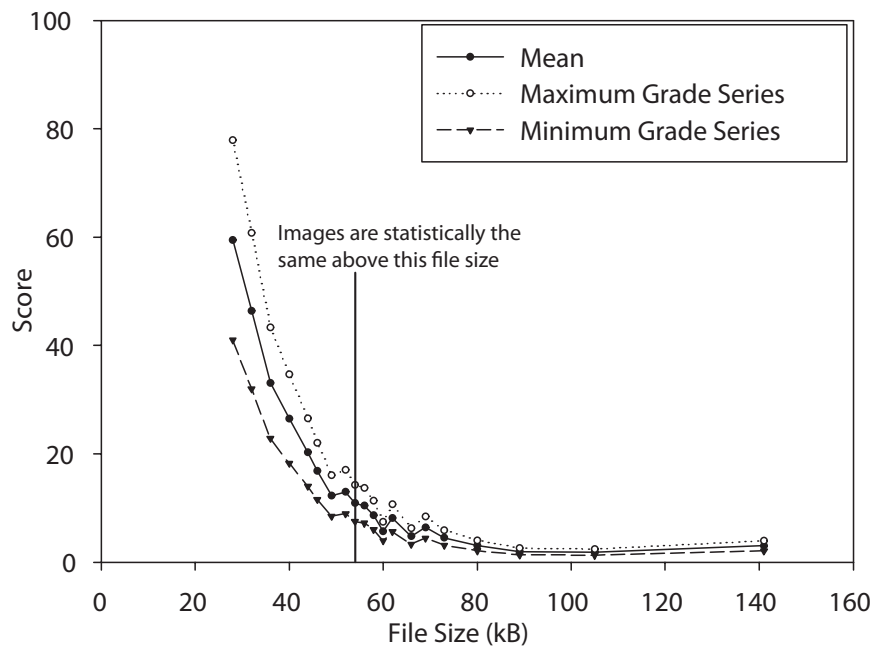
Figure 6.11: *Perseus*: Perceived Mean & Maximum and Minimum Grade Series Calculated from the 95% Confidence Interval.

as an optical or mechanical misalignment.

For the test image $Masha$, the results showed that applying small amounts of difference in the compression between the left and right views could lead to a increase in perceived image quality. However, there was no statistical difference between the symmetric and asymmetric results in these cases.

To assess the statistical significance of the differences between the best asymmetric compression and the corresponding symmetric compression, t-tests were performed. The results of these t-tests are displayed in Table B.5 and in Graph 6.13. The results show that there was only significant difference for images compressed above the symmetric threshold found in the first human trial. Thus, when compressing images up to the image quality threshold, symmetric compression produced better results. Above this threshold, at high levels of degradation, there were instances where applying a small difference in compression between the left and right images could be beneficial. Although it is unlikely, that images with these compression levels and degradation would be used in practice.

The results from the subjective human trial showed that applying large amounts of asymmetric difference in the compression of the stereo pair $Masha$ produced poor quality stereoscopic images in comparison to the symmetrically compressed image of equivalent

file size.
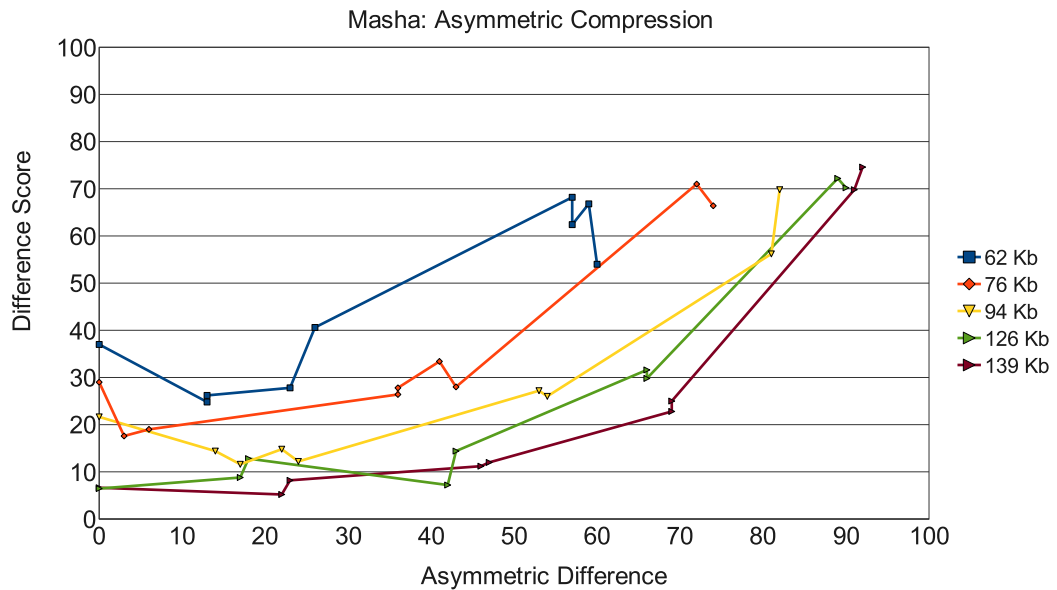
Figure 6.12: *Masha*: Asymmetric Compression Results. The difference score between the asymmetrically compressed images and the uncompressed original is plotted against the asymmetric difference amount between the left and right views.

Mannequin

Figure 6.14 shows the mean score from Table B.6, for each of the file sizes and symmetric and asymmetric compression settings for the test image, *Mannequin*. Again to confirm the prediction that, as with symmetric compression, for *Mannequin*, the perceived image quality from the subjective human based trial decreased with increased asymmetric JPEG compression, a one-way ANOVA was performed. The results of this one-way ANOVA showed there was a significant effect in that increasing the difference in compression of the left and right images within a stereo pair resulted in a decrease in the overall perceived image quality (F value $= 489.37$, p value $= 0.0061$).

Again, as with *Masha*, the results showed that applying small amounts of difference in the compression between the left and right views of the test image *Mannequin* could lead to an increase in perceived image quality. Although, as with the test image *Masha*, there was no statistical difference between the symmetric and asymmetric results in these cases.
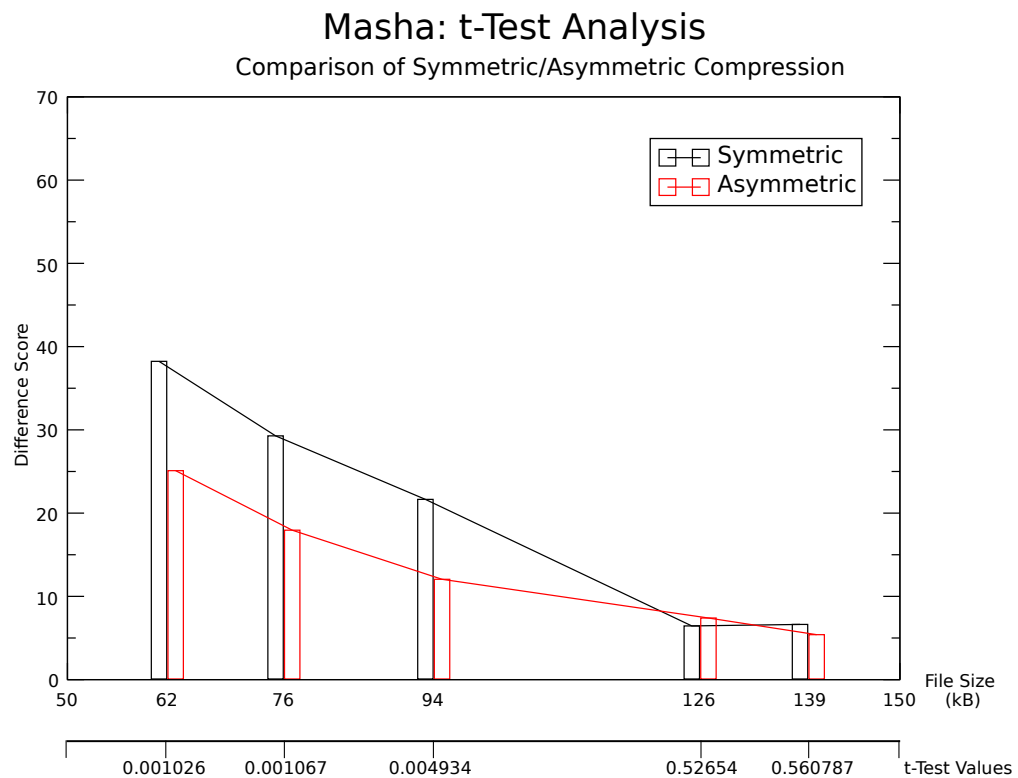
Figure 6.13: *Masha*: Asymmetric t-Test Results.

t-tests were performed to assess the statistical significance of the differences between the best asymmetric compression and the corresponding symmetric compression. The results of these t-tests are displayed in Table B.7 and show that there was only significant statistical difference for images compressed above the symmetric threshold found in the first human trial. As with *Masha*, compressing images up to the image quality threshold, symmetric compression produced better results. Above this threshold, although there were instances where applying a small difference in compression between the left and right images could be beneficial, in practice there are unlikely to be used due to the high levels of degradation.

Applying large amounts of asymmetric difference in the compression of the stereo pair *Mannequin* again produced poor quality stereoscopic images in comparison to symmetrically compressed images of equivalent file size. In general, the larger the difference in left and right compression, the worse the image was perceived.
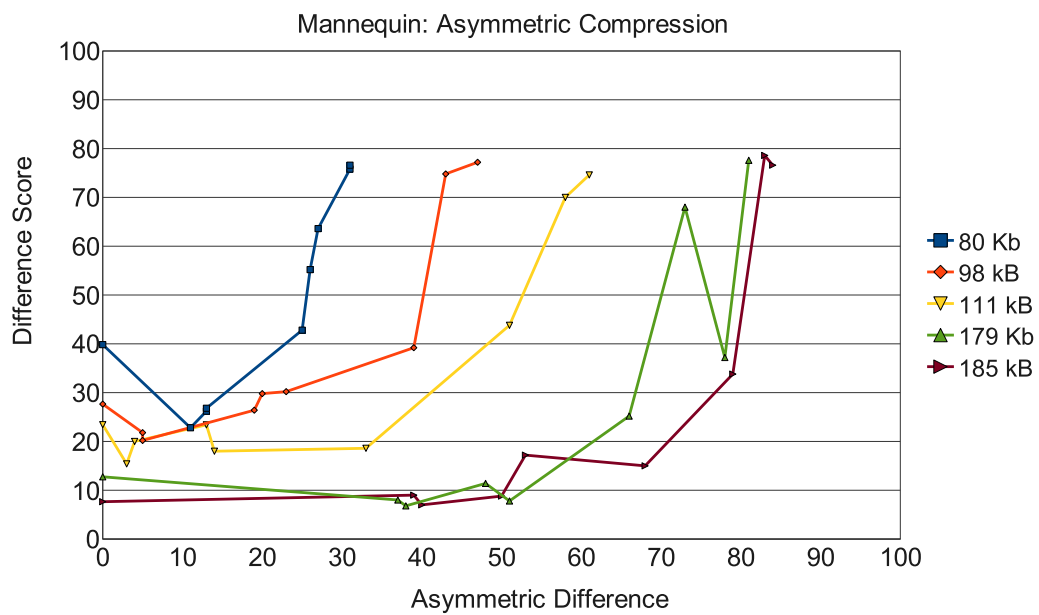
Figure 6.14: *Mannequin*: Asymmetric Compression Results. The difference score between the asymmetrically compressed images and the uncompressed original is plotted against the asymmetric difference amount between the left and right views.

Perseus

Figure 6.15 shows the mean score from Table B.8, for each of the file sizes and symmetric and asymmetric compression settings for the test image, *Perseus*. To confirm the prediction that, as with symmetric compression for *Perseus*, the perceived image quality from the subjective human based trial decreased with increased asymmetric JPEG compression a one-way ANOVA was performed. The results of this one-way ANOVA showed there was a significant effect in that increasing the difference in compression of the left and right images within a stereo pair resulted in an decrease in the overall perceived image quality (F value = 489.37, p value = 0.0061).

Again the results showed that applying small amounts of difference in the compression between the left and right views of the test image *Perseus* could lead to an increase in perceived image quality. Again, there was no statistical difference between the symmetric and asymmetric results in these cases.

This was consistent across all three test images. Again t-tests were performed to assess the statistical significance of the differences between the best asymmetric compression and the corresponding symmetric compression, for consistent file sizes. The results of these

t-tests, on the *Perseus* image, are displayed in Table B.9 and show that there was only significant difference for images compressed above the symmetric threshold found in the first human trial. Above this threshold, at high levels of degradation, there were again instances where applying a small difference in compression between the left and right images could be beneficial. Although, as with the other test images, it is unlikely that these compression levels would be used in practice.

Applying large amounts of asymmetric difference in the compression of the stereo pair *Perseus* again produced poor quality stereoscopic images in comparison to symmetrically compressed images of equivalent file size. As with the other two test images, the larger the asymmetric difference in left and right compression, in general, the worse the image was perceived.
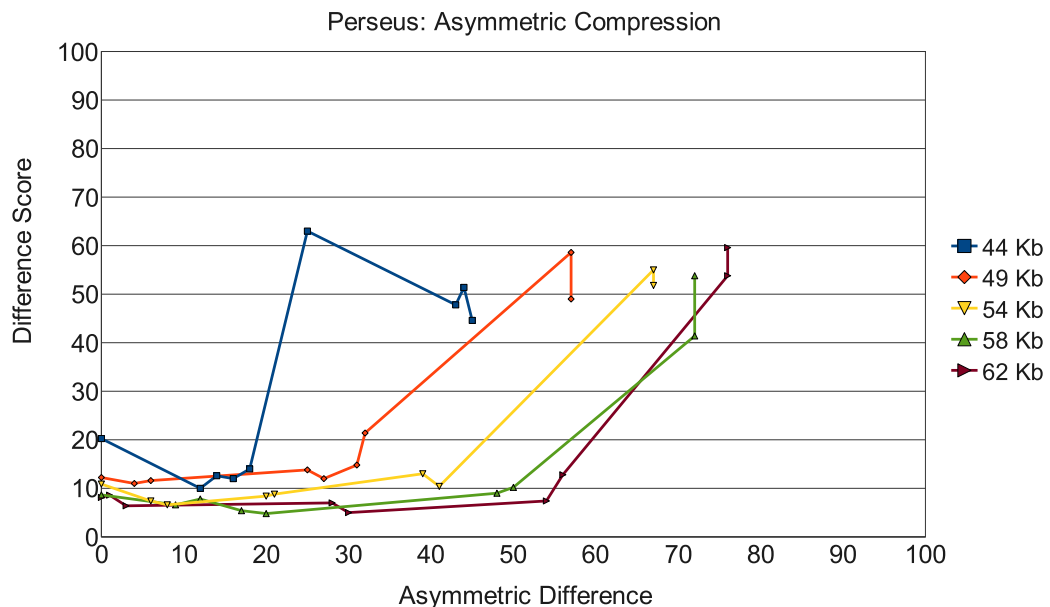
Figure 6.15: *Perseus*: Asymmetric Compression Results. The difference score between the asymmetrically compressed images and the uncompressed original is plotted against the asymmetric difference amount between the left and right views.

6.6   General Discussion

The analysis of the t-test and the ANOVA data from the symmetric trial revealed that, as expected, the results between the uncompressed image and the most compressed image showed that there was an 100% probability that the images were statistically different. It

can therefore be concluded that the results from the symmetric human based experiment have shown, as predicted, symmetric JPEG compression of the stereoscopic images has a detrimental effect on perceived image quality. For each of the three images it was possible to establish a baseline of human stereoscopic image quality preference.

The analysis of the t-test and ANOVA results from the asymmetric trial showed that as expected increasing the overall asymmetric compression level resulted in a reduced image quality and that there was a 100% probability that the most asymmetrically compressed image was statistically different to the uncompressed original. It can therefore be concluded that the results from the asymmetric subjective human trial have shown that for the three test images, increasing asymmetric compression of stereoscopic images has a detrimental effect on perceived image quality. For each of the three test images it was possible to establish which of the compression methods, symmetric or asymmetric produced images of a better perceived quality.

For the test image *Masha*, the results suggest that in general the symmetric compression approach should be used. Where there is a statistical difference between the two compression methods symmetric compression outperformed asymmetric.

Comparing the asymmetrically compressed stereo image pair with the best perceived image quality result, with the corresponding symmetric image showed that applying small amounts of asymmetric compression may result in a greater perceived image quality for *Masha*. However, the t-tests performed between these results showed that there was no statistically significant difference between the asymmetric and symmetric results, unless the image was compressed beyond the symmetric image quality threshold. There was a statistical difference for the images with greater compression, however as these compressions are above the image quality threshold, the images have very high levels of degradation and therefore are unlikely to be used in practice.

In the case of *Mannequin*, again the results suggest that the symmetric compression approach should be used. Symmetric compression again outperformed asymmetric, where there is a statistical difference between the two compression methods.

Applying small amounts of asymmetric compression, again may result in a greater perceived image quality. However, as with *Masha*, the results of the t-test between the symmetrically compressed image and the best perceived asymmetric equivalent, showed no statistical difference between the two compressed images, for compressions below the symmetric image quality threshold. There was statistically significant differences for the

images compressed above the threshold, but again the images have very high levels of degradation and therefore are unlikely to be used in practice.

For *Perseus*, the results suggest that the symmetric compression approach should be applied. Again, where there is a statistical difference between the two compression methods, symmetric compression outperformed asymmetric.

Applying small amounts of asymmetric compression to *Perseus*, again may result in a greater perceived image quality, but as with the other two images, there was no statistical difference for images below the symmetric image quality threshold. Again, there was a statistical difference for the images with greater compression, however as these compressions are above the image quality threshold, again the images have very high levels of degradation and therefore are unlikely to be used in practice.

Overall the results suggest, that for the three test images, *Masha*, *Mannequin* and *Perseus*, perceived image quality decreases with an increase in the asymmetric difference amount between the compression of the left and right images. From this it is concluded that the results show that overall, increasing the difference between the views has a decremental affect on image quality. The results showed that in general, for stereoscopic image compression using JPEG, a symmetric as opposed to asymmetric compression approach across the left and right images of the pair should be used.

6.7   Conclusion

How the quality of stereoscopic images varies with different compression methods has been investigated. The experimental methodology presented in this chapter aimed to and succeeded in generating statistically robust results.

In particular, the human trials assessing the compressed images, provided data that allowed for a detailed understanding of benefits and drawbacks for compressing images symmetrically and asymmetrically. A number of observations have been made and are as follows.

It is concluded that increasing the symmetric compression level results in a reduced perceived image quality of symmetrically compressed images. The symmetric human trial was able to produce a threshold for acceptable image quality. From the results collected it was concluded that ideally, this threshold level should not be exceeded if an acceptable quality of the stereo image is to be maintained.

The hypothesis was that the asymmetric human trial would show that compressing

images symmetrically would result in a better perceived quality of stereo image. It is concluded from the results that in general this is the case and the symmetric, as opposed to asymmetric compression, should be used for stereo image compression.

However, the results of this trial showed that there was a small number of instances were asymmetric seemed to outperform symmetric compression. Analysis showed that overall, increasing the difference in the left and right views of the stereo image reduced perceived image quality for a specific file size. The results also showed that applying a small difference in levels of left and right compression may result in an increase in the perceived image quality. However, these differences in image quality were only statistically significant for images sizes above the calculated symmetric image quality threshold and were images with greater than practical amounts of compression applied and therefore not likely to be used in practice.

Overall, it is concluded that, for stereoscopic image compression using JPEG, a symmetric, as opposed to asymmetric compression approach across the left and right images of the pair, should be used.

6.8 Future Work

The results of the symmetric and asymmetric subjective human trials, suggest that overall symmetric compression produces images with a better perceived quality. The results at high levels of compression are much closer than suggested by PSNR. Further work should still be undertaken in the development of a new Human Visual System (HVS) based image quality metric. This metric should then be evaluated against both PSNR and the subjective image quality results.

Chapter 7

Development and Evaluation of the New Stereo Band Limited Contrast Metric

7.1 Introduction

How Peak Signal to Noise Ratio (PSNR) measures the quality of varyingly coded stereoscopic image pairs, has already been investigated in Chapter 5. However, PSNR measures of image quality are widely criticized for correlating poorly with perceived visual quality.

In Chapter 6, symmetric and asymmetric compression approaches were evaluated using a subjective trial. This trial measured stereoscopic image encoding quality, using human subjects. It used the Double Stimulus Continuous Quality Scale (DSCQS) from the ITU-R BT.500-11 recommendation. From the symmetric trial a image quality threshold was calculated for each of the test images.

In this chapter the development and evaluation of a new stereoscopic image quality metric that can be used to rank the quality of compressed images and guide the choice of compression is detailed.

Computational models of the Human Visual System (HVS) have been considered and here, a new stereoscopic image quality metric, Stereo Band Limited Contrast (SBLC), is developed. The new metric uses point matches between the left and right views in order to account for HVS sensitivity to contrast and luminance changes in regions of high spatial frequency. The metric is based on Michelson's Formula and Peli's Band Limited Contrast

Algorithm.

The SBLC metric will be compared with PSNR, using the image quality thresholds established in the symmetric trial detailed in Chapter 6. The results suggest that SBLC is a better predictor of human image quality preference than PSNR and could be used to predict threshold compression levels for stereoscopic image pairs.

7.2   Motivation

Stereoscopic images require additional storage space and bandwidth for transmission. Therefore, substantial research effort has been focussed on digital image compression, using for example JPEG [115], to obtain bandwidth and storage capacity savings. However, the question of how much compression to apply and the approach to adopt in applying the compression to the image pair remains unresolved [31, 120].

There is currently extensive interest in metrics for automatically predicting the compression settings for stereoscopic images, to minimize file size, but still maintain an acceptable level of image quality.

Traditionally, to assess the effect of these compression methods on perceived image quality, 2D objective measures such as Peak Signal to Noise Ratio (PSNR) or Mean Squared Error (MSE) have been used. A previous chapter details investigations to evaluate the quality of symmetric and asymmetric coding using PSNR. The results strongly suggest that symmetric coding should be used for compressing stereo image pairs, however, it is felt the PSNR metric does not correlate well with subjective judgement of image quality [10, 86]. This is not surprising, as PSNR is essentially a simple pixel based comparison method.

It was therefore felt important to investigate current metrics for evaluating stereoscopic image compression further and develop a new stereoscopic image quality metric.

7.3   The New Stereo Band Limited Contrast (SBLC) Metric

Visual perception research is dominated by the study of contrast sensitivity and changes in luminance. Current metrics are designed for 2D image comparison and are not able to easily predict the quality of 3D images, therefore, a new stereoscopic image quality metric was designed and implemented.

Aspects of human vision relevant to picture quality are predicted by modelling the human visual system's sensitivity to light [111]. The response of the HVS depends on the

image content and the colour and local variations of luminance (i.e. relative luminance and contrast) [94, 95, 154].

The new Stereo Band Limited Contrast(SBLC) metric, accounts for HVS sensitivity to contrast and luminance changes at regions of high spatial frequency. The metric can be used to rank stereoscopic pairs in terms of image quality.

The SBLC metric, uses Sift and the RANSAC algorithm [36] to extract edges, corners and regions of high spatial frequency within the image. Points are matched between the left and right views of the stereo pair. For each of the matched points, the surrounding pixels are considered as shown in Figure 7.1. The matched point and the surround pixels make up the matched region.

Stereo Band Limited Contrast - Region Size
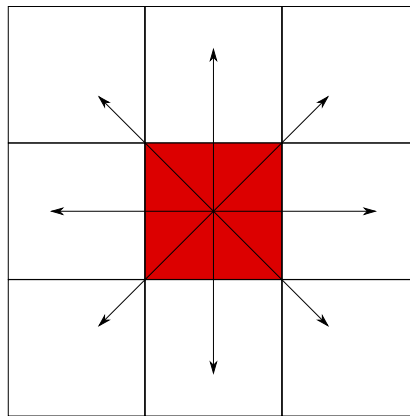
Figure 7.1: For each matched point (shown in Red), the surrounding pixels are considered.

Pixels outside the range of the image are discarded. The Relative Luminance, I, for every matched point in each region is then calculated using:

$$I = 0.2126R + 0.7152G + 0.0722B \quad \text{(Red, Green and Blue Values)} \tag{7.1}$$

Current research suggest that metrics based on the Michelson's Contrast Formula, are better at predicting monoscopic image quality that PSNR [91, 94, 131]. Using Relative Luminance (I), Michelson's Contrast Formula [94] is calculated for both of the corresponding matched regions within the stereoscopic pair and then the average of matched regions in both the left and right views, C, is calculated. This is repeated for all the matched regions. The overall Relative Mean Luminance (L) of the whole image is calculated using the following equation:

$$L = \frac{1}{255 * mn} \sum_{x=0}^{m-1} \sum_{y=0}^{n-1} I(x, y) \tag{7.2}$$

The Stereo Band Limited Contrast (SBLC) is calculated from the mean of the ratio, $C(x)/L$ for every matched point x. The value for the compressed image is then deducted from that of the original to give the Stereo Band Limited Contrast. This was calculated using the following equation, where p is the total number of matched point regions.

$$SBLC = \left(\frac{1}{p} \sum_{x=0}^{p} \frac{C_{Orig}(x)}{L_{Orig}} \right) - \left(\frac{1}{p} \sum_{x=0}^{p} \frac{C_{Comp}(x)}{L_{Comp}} \right) \tag{7.3}$$

7.4 Hypothesis

The prediction is that the image quality rating produced by the metric will decrease with increased JPEG compression. It is expected that the image quality thresholds produced by the SBLC metric will be closer to the subjective results in Chapter 6, compared with the thresholds produced from PSNR, in Chapter 5.

To evaluate this hypothesis, how subjective image quality varies with the full range of JPEG compression was investigated. The new HVS based SBLC metric was evaluated in comparison to PSNR by comparing them to the results collected in the previous subjective human symmetric image quality trial.

7.5 Calculating Image Quality Threshold from SBLC and PSNR

To enable a comparison of SBLC to PSNR and the results for the subjective trial an estimation of an image quality threshold was calculated from the PSNR results and the SBLC results. These image quality thresholds are calculated by estimating the first point of inflection in the curve and thus gives an estimation as to when the compressed image starts to significantly differ from the uncompressed original.

The point of inflection is calculated by estimating the 1$^{\text{st}}$ and 2$^{\text{nd}}$ derivatives from a series of points.

For Figure 7.1, these are estimated in the following way:

Table 7.1: Calculating Point of Inflection

	X	Y	$\left(\frac{dy}{dx}\right)$	$\left(\frac{d^2y}{dx^2}\right)$
1	X_1	Y_1	$\left(\frac{dy}{dx}\right)_1$	$\left(\frac{d^2y}{dx^2}\right)_1$
2	X_2	Y_2	$\left(\frac{dy}{dx}\right)_2$	$\left(\frac{d^2y}{dx^2}\right)_2$
3	X_3	Y_3	$\left(\frac{dy}{dx}\right)_3$	$\left(\frac{d^2y}{dx^2}\right)_3$
4	X_4	Y_4	$\left(\frac{dy}{dx}\right)_4$	$\left(\frac{d^2y}{dx^2}\right)_4$
5	X_5	Y_5	$\left(\frac{dy}{dx}\right)_5$	$\left(\frac{d^2y}{dx^2}\right)_5$

Estimation of $\left(\frac{dy}{dx}\right)_3$ at point (X_3, Y_3) is calculated from $\frac{(Y_4-Y_2)}{(X_4-X_2)}$

The point of inflection is calculated for the estimate of the second derivative at point (X_3, Y_3) using the following equation:

$$\left(\frac{d^2y}{dx^2}\right)_3 = \frac{\left(\frac{dy}{dx}\right)_4 - \left(\frac{dy}{dx}\right)_2}{(X_4 - X_2)} \tag{7.4}$$

7.6 Comparison of SBLC Results to PSNR and Subjective Preference

The results for SBLC are presented. They are compared to those collected for both PSNR and from the symmetric subjective human trial in previous chapters. A baseline of quality for each image was established from the symmetric human preference trial results.

The PSNR and the new SBLC metric values were then calculated for each JPEG compression. In each case a predicted point of inflection in the graph signified the threshold where the quality of the images started to degrade heavily. Therefore, the point of inflection was calculated, for each image, from an estimation of the second derivative. The image quality thresholds were then compared to the baseline established using the subjective human trial.

The results collected using the new SBLC metric, PSNR and from the previous symmetric human preference trial, for each compression setting for the test images, *Masha*, *Mannequin* and *Perseus* are shown in Tables C.1, C.2 and C.3 respectively and displayed in Figures 7.2, 7.3 and 7.4.
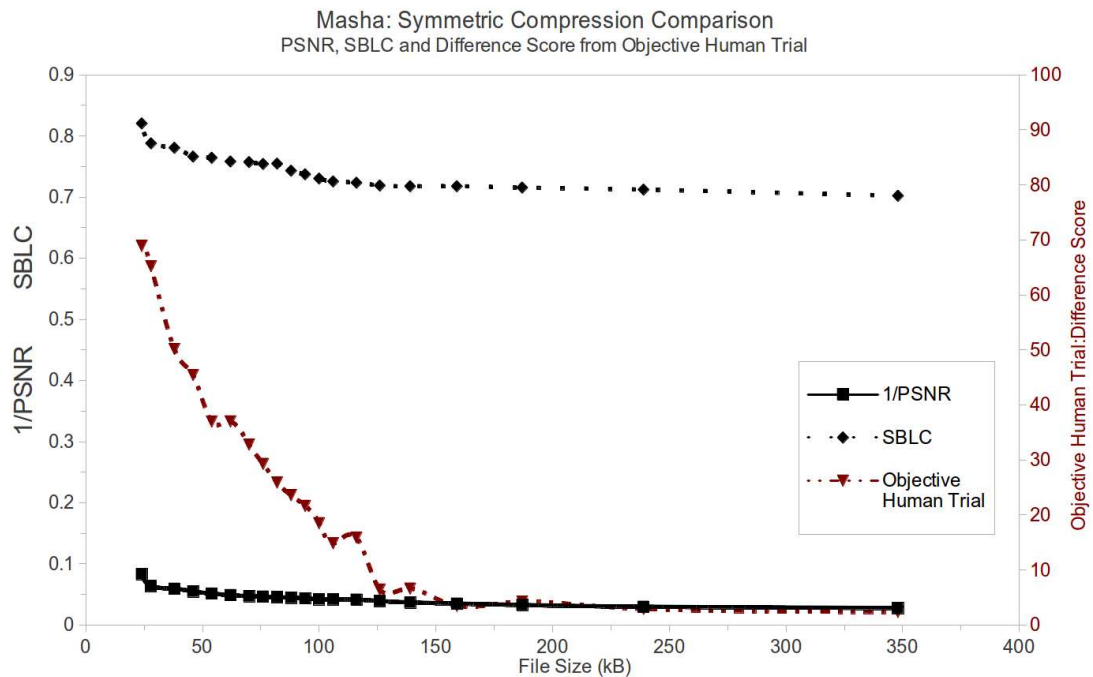
Figure 7.2: *Masha*: PSNR, SBLC and Difference Score from the Human Trial

7.6.1   Symmetric Subjective Human Trial

The results from the Tukey multiple comparison test undertaken in the previous chapter, gave compressing up to JPEG compression level 30 as the image quality threshold for *Masha*. As with the previous chapter, Figure 7.5 again shows the mean score and maximum and minimum grade series calculated in Chapter 6 from the 95% confidence interval plotted against image file size (kB). Marked on the graph is the calculated image compression threshold, where statistically, humans perceive the images become different from the original.

For the image *Mannequin* the Tukey multiple comparison test undertaken, gave compressing up to JPEG compression level 50 as the image quality threshold. Figure 7.6 shows the mean score and maximum and minimum grade series calculated from the 95% confidence interval plotted against image file size (kB). Marked on the graph is the calculated image compression threshold, where statistically, humans perceive the images become different from the original.

The results from the Tukey multiple comparison test undertaken on *Perseus*, gave compressing up to JPEG compression level 60 as the image quality threshold. Figure 7.7 shows the mean score and maximum and minimum grade series calculated from the 95%
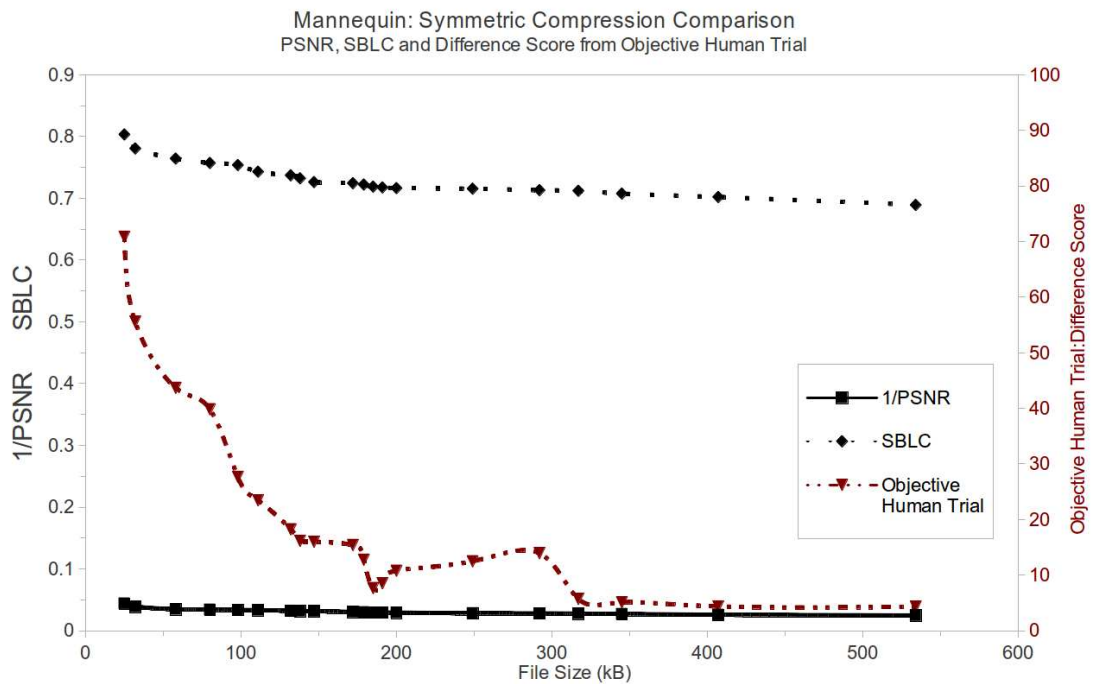
Figure 7.3: *Mannequin*: PSNR, SBLC and Difference Score from the Human Trial

confidence interval plotted against image file size (kB). Again marked on the graph is the calculated image compression threshold, where statistically, humans perceive the images become different from the original.

7.6.2   PSNR

Figure 7.8 shows the 1/PSNR values from Table C.1 for each compression setting for the test image *Masha*. The reciprocal of PNSR is used so that the graphs follow the same trend and can be easily compared with those from the new SBLC metric and human factor trial results. The point of inflection in the results is calculated from an estimation of the second derivative and found to be between JPEG compressions 20-25 (139-159kB),thus giving an estimated image quality threshold of JPEG compression 20 for *Masha*, when using PSNR alone.

The PSNR results for *Mannequin* and *Perseus* are shown in Table C.2 and Table C.3. Figure 7.9 and Figure 7.10 show the 1/PSNR values for each compression setting for these images. Again, points of inflection, were calculated from estimations of the second derivative and gave JPEG compressions of 30-35 (66-69kB) and 15-20 (80-89kB), giving PSNR estimated image quality thresholds of 30 and 15 for *Mannequin* and *Perseus*
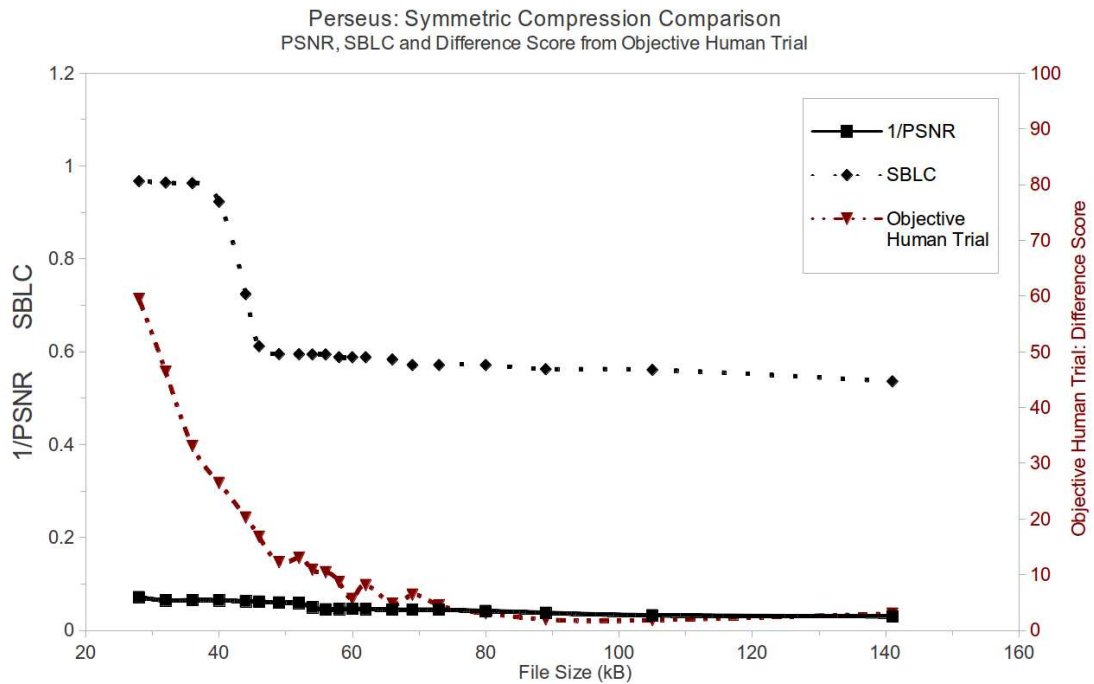
Figure 7.4: *Perseus*: PSNR, SBLC and Difference Score from the Human Trial

respectively.

7.6.3   SBLC

Figure 7.11 shows the New SBLC image metric values from Table C.1 for each JPEG compression setting for the test image *Masha*. The point of inflection again is calculated from an estimation of the second derivative, producing a point of inflection between JPEG compressions 30-35 (116-126kB), giving an estimated image quality threshold of JPEG compression 30 for *Masha* when using SBLC metric.

The SBLC results for *Mannequin* and *Perseus* are shown in Table C.2 and Table C.3. Figure 7.12 and Figure 7.13 show the SBLC values for each compression setting for the test images. Here the points of inflection in the images gave JPEG compressions of 30-35 (200-249kB) and 40-50 (58-62kB), giving SBLC estimated image quality thresholds of 30 and 40 for *Mannequin* and *Perseus* respectively.
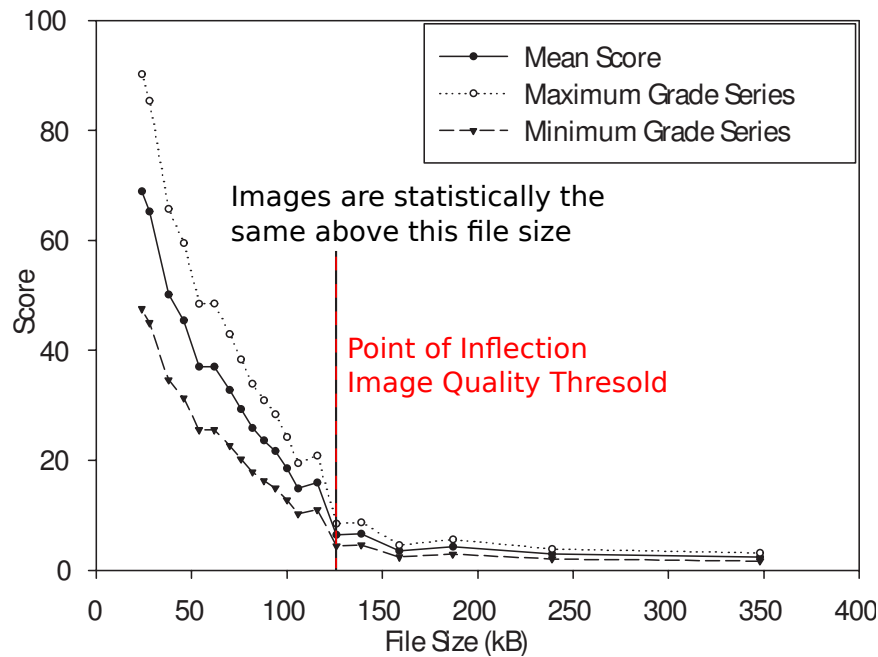
Figure 7.5: *Masha*: Perceived Mean & 95% Confidence

7.7 General Discussion

For *Masha* the new SBLC metric provides a conservative estimation for the amount of compression that can be applied before the perceived image quality threshold, calculated in Chapter 6 is reached. The threshold estimation from the PSNR graph suggests that you can compress further than the established baseline. In this case it is believed that the SBLC metric produces a useful threshold that can be used as a starting point when deciding the required image compression to be used. It allows compression at least as far as the SBLC metric suggests with any noticeable degradation.

In the case of *Mannequin*, the predicted thresholds from PSNR and the new SBLC Metric were the same. As *Mannequin* is a photograph and PSNR works better with photographs [123] it is understandable that PSNR prediction is closest to the SBLC in this case. Both estimations provide an acceptable, conservative, compression level for the image.

For *Perseus* the SBLC metric estimation is much closer to the subjective threshold than that from PSNR. In this case PSNR does not provide an suitable threshold to predict the acceptable image compression baseline. The new SBLC metric again produces an acceptable, conservative, compression level for the image.
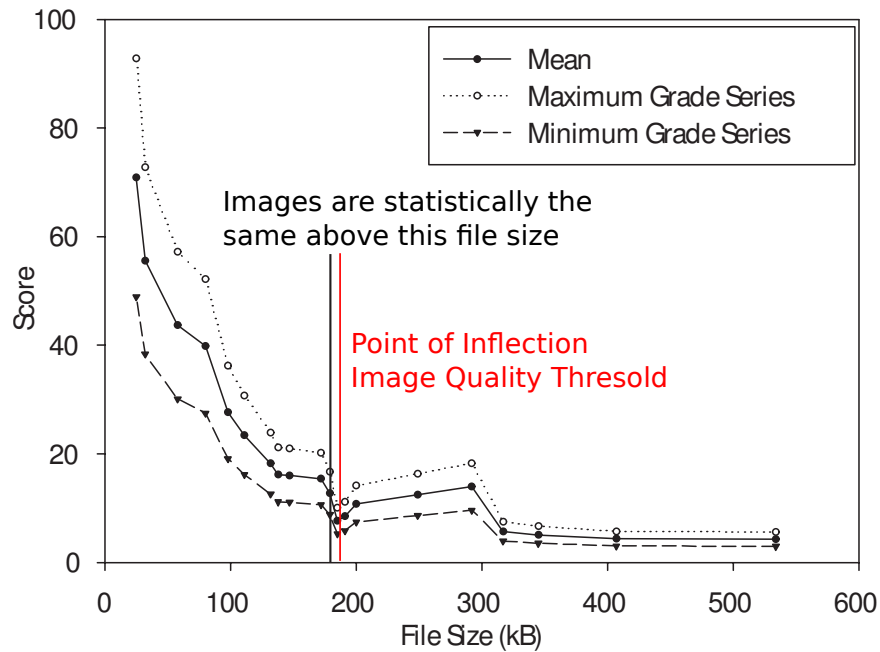
Figure 7.6: *Mannequin*: Perceived Mean & 95% Confidence

Overall, the results suggest that the new metric is a better predictor of human image quality preference than PSNR and could be used to predict a conservative threshold compression level for stereoscopic image pairs, where no perceived degradation would be apparent.

7.8 Conclusion

How the quality of stereoscopic images varies with compression has been investigated.

As expected it has been shown that PSNR, the new stereoscopic image quality metric and the perceived image quality, all reduce with image compression and reduce with a reduction in image quality.

The new SBLC metric produces a conservative estimate of perceived image quality for all three images. It is concluded that SBLC produces a better estimation of the stereoscopic image quality baseline compared to that produced from the PSNR results and that SBLC produces a useful threshold that can be used as a practical starting point when deciding the required image compression to be used.
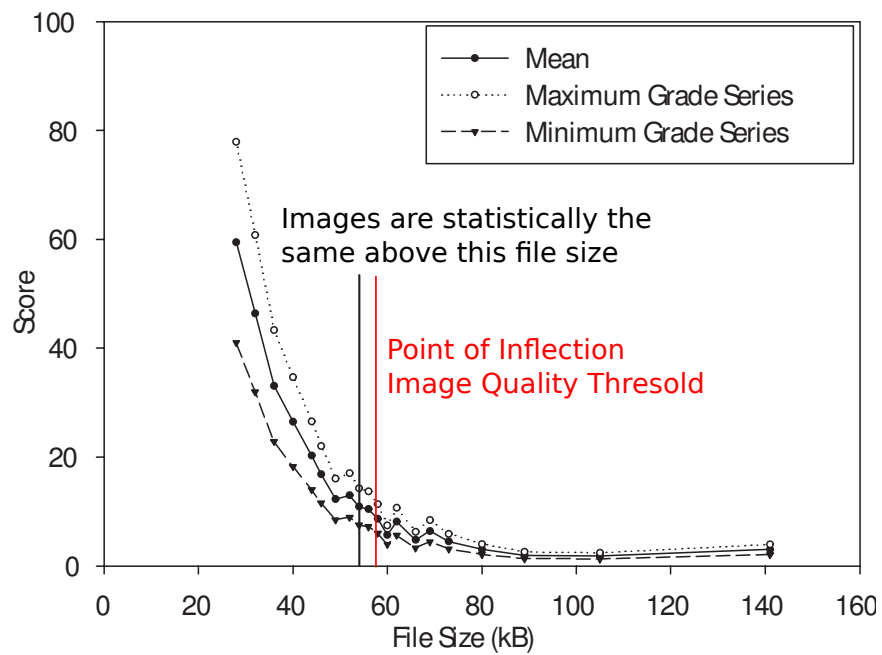
Figure 7.7: *Perseus*: Perceived Mean & 95% Confidence
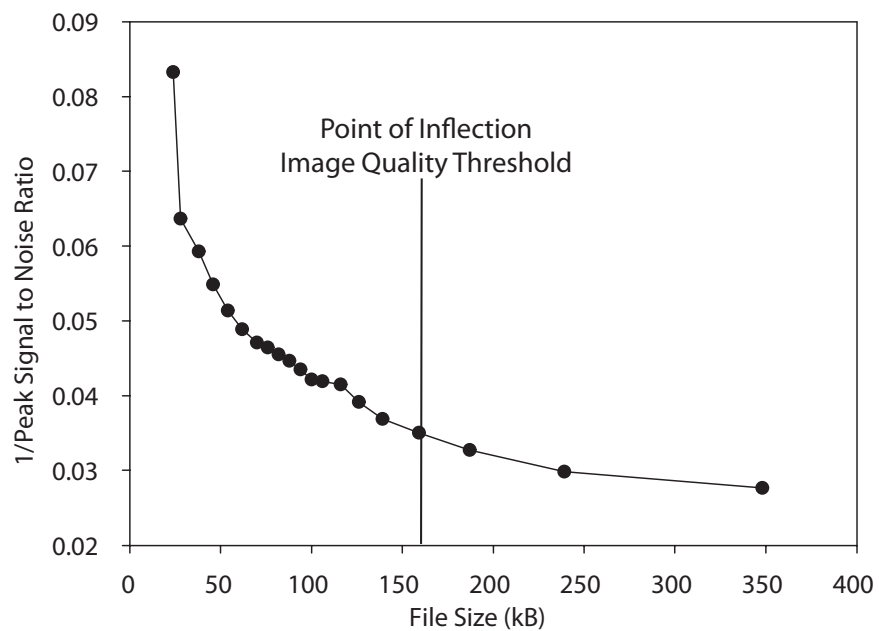
Figure 7.8: *Masha*: 1/PSNR vs. File Size(kB)
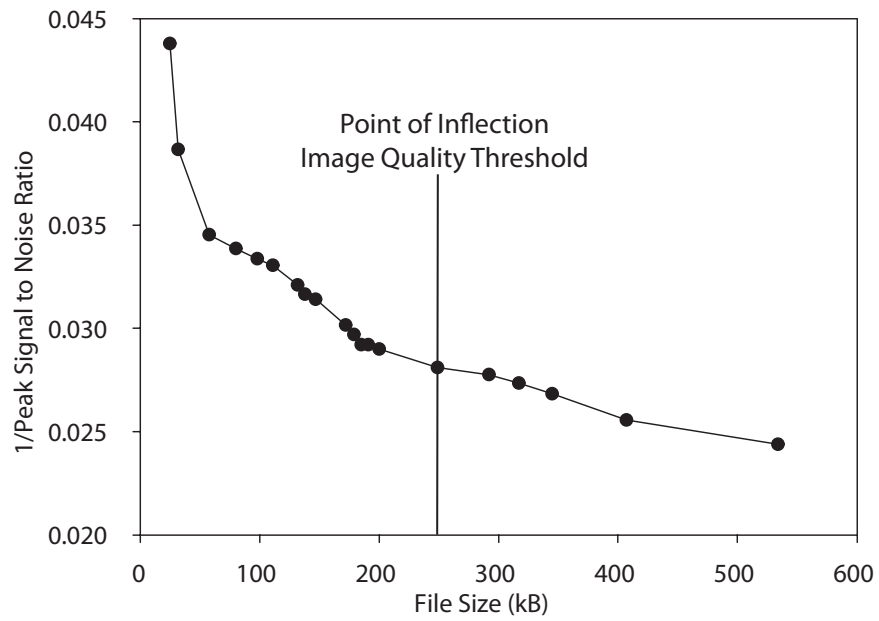
Figure 7.9: *Mannequin*: 1/PSNR vs. File Size(kB)
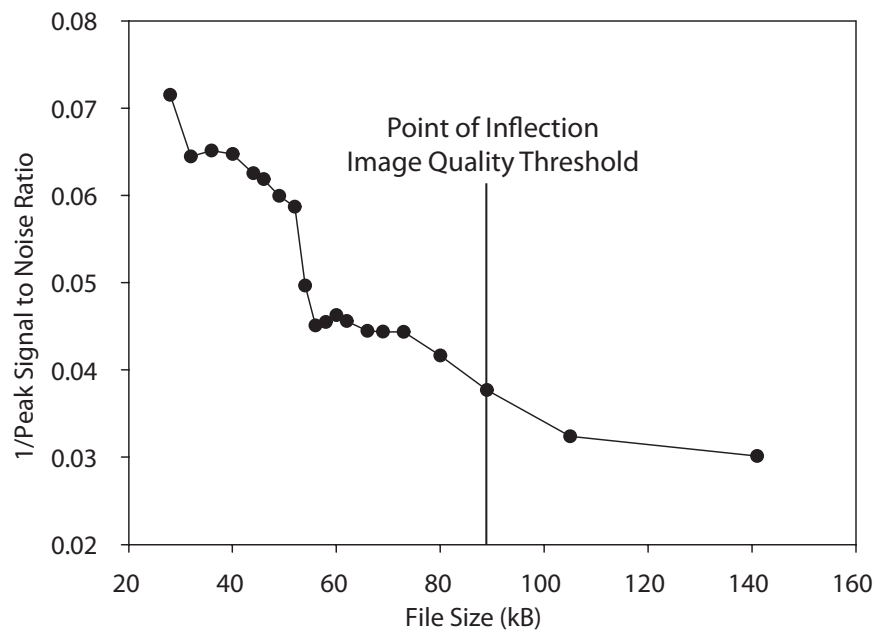
Figure 7.10: *Perseus*: 1/PSNR vs. File Size(kB)
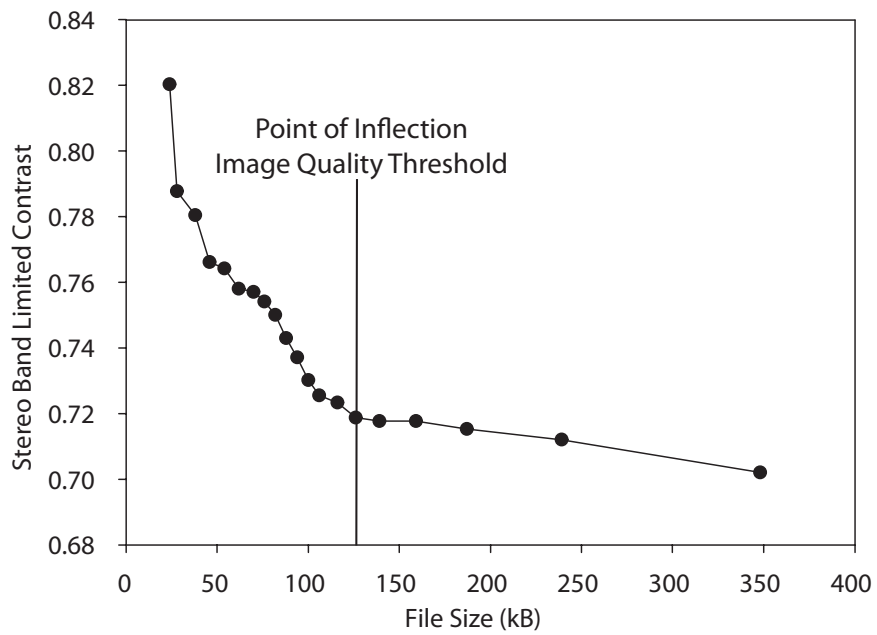
Figure 7.11: *Masha*: SBLC vs. File Size(kB)
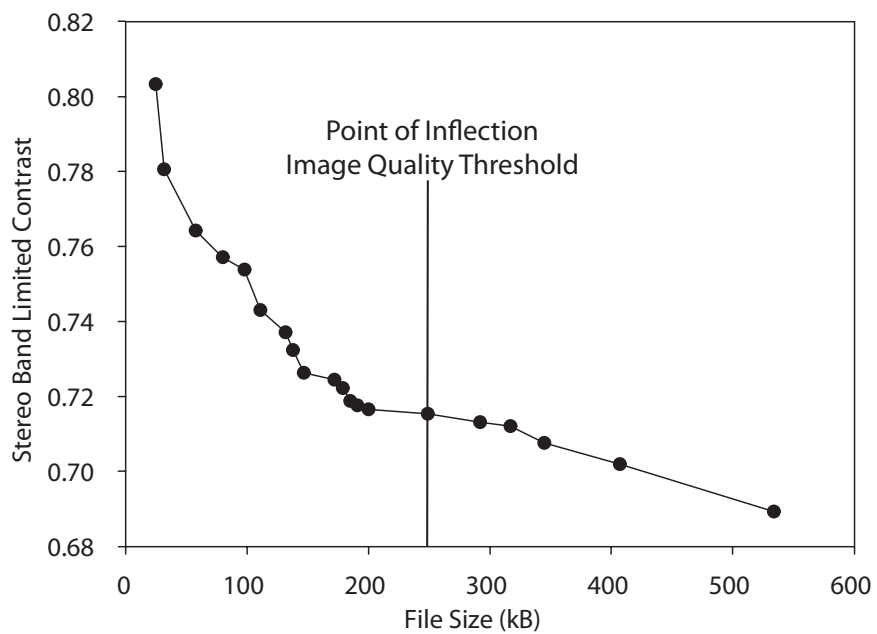
Figure 7.12: *Mannequin*: SBLC vs. File Size(kB)
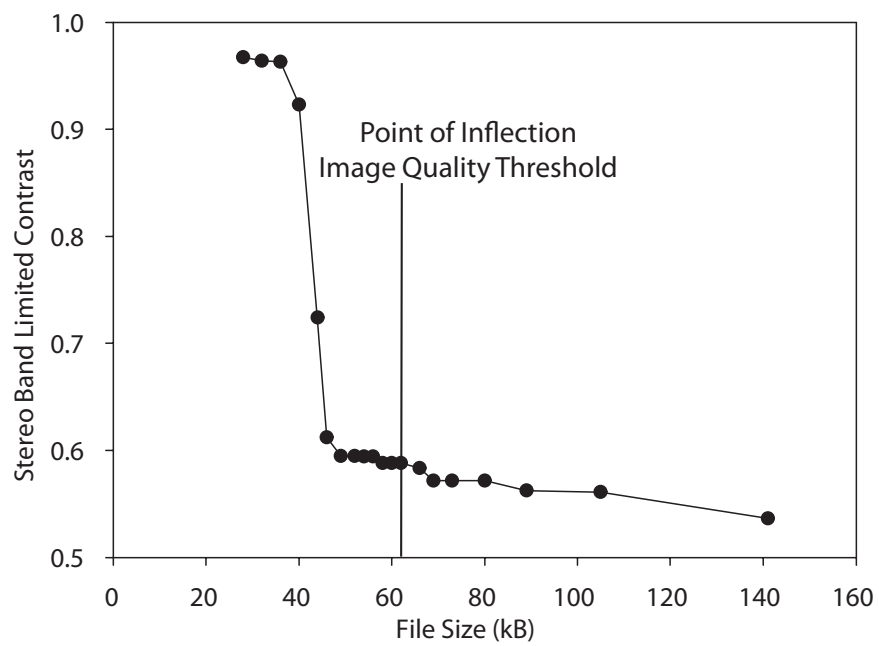
Figure 7.13: *Perseus*: SBLC vs. File Size(kB)

Chapter 8

Conclusions and Further Work

8.1   Introduction

The thesis is, that stereoscopic compression needs new ways to operate, compared to existing methods. This includes methods for applying compression and for measuring it. This chapter draws together results and the evaluation of those results to present conclusions about symmetric and asymmetric stereoscopic compression approaches and the metrics used to evaluate these compression techniques. It describes possible future work that builds upon this research.

The research flow; experiments undertaken and the conclusions from those experiments are shown in Figure 8.1.

8.2   Major Contributions and Conclusions

This section details the major contributions of the research. It describes the results and conclusions from each contribution. Overall conclusions are then drawn.

1. **Symmetric and asymmetric compression evaluation using existing image comparison metrics, PSNR and SAD.**

   The main aim of the initial research undertaken was to investigate and compare two possible compression methods to try and reduce the required file size and bandwidth for the storage and transmission of stereoscopic images.

   The work evaluates the two compression approaches, symmetric and asymmetric encoding, using the objective quality metrics, PSNR and SAD and concludes that
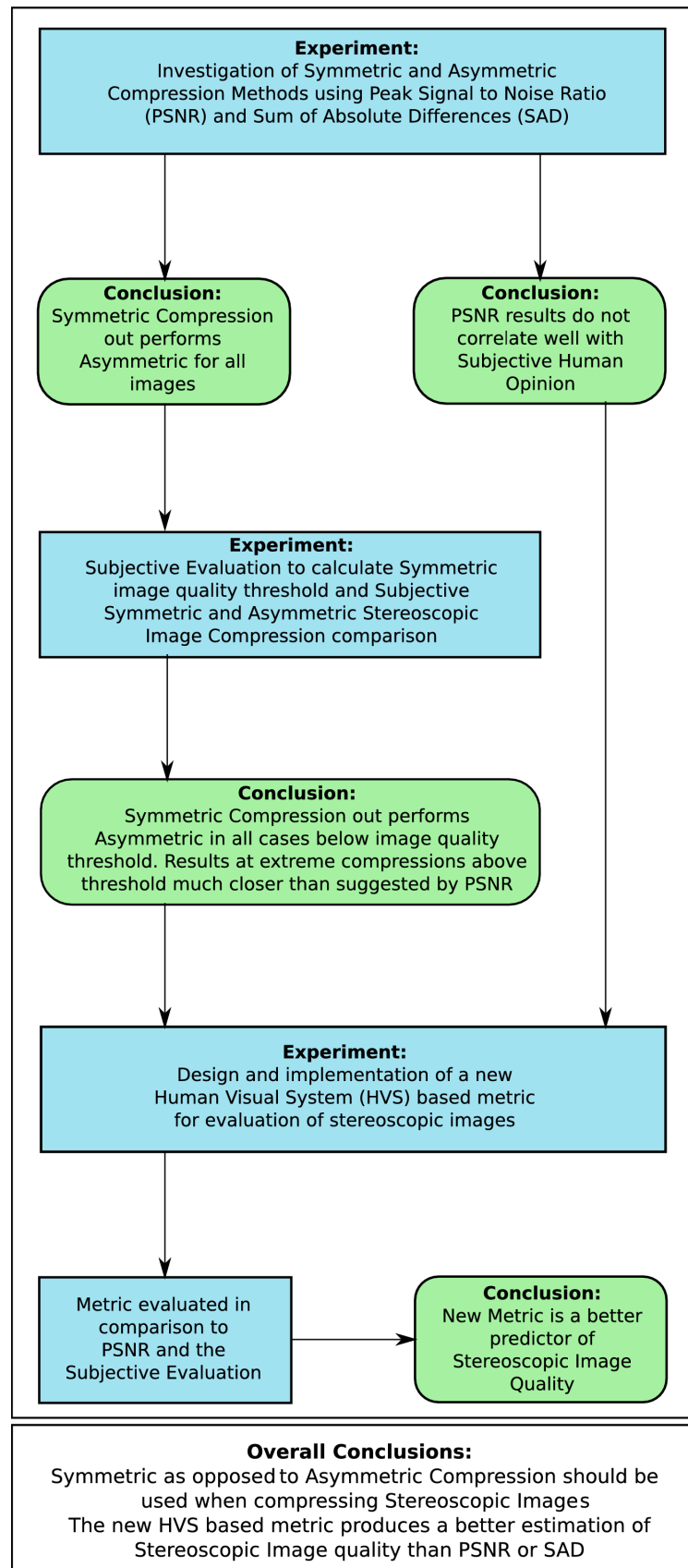
Figure 8.1: Research Flow: Experiments and Conclusions

symmetric compression produced the better results. It is concluded that for JPEG and JPEG 2000 stereoscopic images, the symmetric compression approach, as opposed to asymmetric compression, should be used for all image types. The same conclusion is drawn when using either of the two image comparison methods, PSNR and SAD.

2. **Subjective evaluation of symmetric and asymmetric stereoscopic image compression and the calculation of symmetric image quality thresholds**

How the quality of stereoscopic images varies with different compression methods was investigated. The experimental methodology presented in this work generated statistically robust results.

In particular, the subjective trials assessing the compressed images, provided data that allowed for a detailed understanding of benefits and drawbacks for compressing images symmetrically and asymmetrically.

The following observations were made. It is concluded that increasing the asymmetric compression level results in a reduced perceived image quality of asymmetrically compressed images. The symmetric human trial produced a threshold for acceptable image quality for each of the images. From the results collected it was concluded that ideally this threshold level should not be exceeded if an acceptable quality of the stereo image is to be maintained.

For the asymmetric trial, analysis showed that overall, increasing the asymmetric difference in the left and right views of the stereo image reduced perceived image quality for a specific file size. It is concluded that in general, for stereoscopic image compression, using JPEG, a symmetric as opposed to asymmetric compression approach across the left and right images of the pair should be used.

3. **An investigation of the Human Visual System (HVS) and the design and implementation of a new HVS based metric, comparing the results from current image quality metric, PSNR, and the subjective trial**

A new stereoscopic image quality metric, Stereo Band Limited Contrast has been developed. The new SBLC metric produces a conservative estimate of perceived image quality for all three images. It is concluded that SBLC produces a better estimation of the stereoscopic image quality threshold compared to that produced from the PSNR results and that SBLC produces a useful threshold that can be

employed as a practical starting point when deciding the required image compression to be used. Images can be compressed at least as far as the SBLC metric suggests without any noticeable degradation.

8.2.1 Overall Conclusions

The conclusions of this research are,

- When investigating symmetric and asymmetric compression using current image quality metrics such as PSNR and SAD, it is concluded that symmetric compression out performs asymmetric compression for all the images tested.

- It is concluded, from current research, that PSNR results do not correlate well with subjective human opinion [10, 86].

- When evaluating symmetric and asymmetric compression using subjective human opinion, it is concluded that in general, symmetric compression out performs asymmetric compression.

- It is concluded that the new HVS based stereo image quality metric is a better predictor of stereo image quality compared to the existing metrics, PSNR and SAD. The new metric also produces a conservative estimation of subjective stereo image quality that can be used as a starting point for increasing compression levels.

8.3   Criteria for Success

The success of the research is measured against the criteria detailed in Chapter 1. An analysis of how each criterion is achieved in the research is stated here.

1. *A successful evaluation of symmetric and asymmetric compression approaches using existing image quality comparison metrics, such as Peak Signal to Noise Ratio (PSNR) and Sum of Absolute Differences (SAD).*

   An experiment was undertaken to assess both symmetric and asymmetric compression techniques, using PSNR and SAD. The experiment was successful and showed that symmetric compression out performed asymmetric in all cases.

2. *The successful design and running of a subjective human based trial to evaluate symmetric image compression. This trial should produce an acceptable image quality threshold for symmetric images.*

   A subjective human trial for the assessment of symmetric image quality was designed and undertaken. Symmetrically encoded images with various amounts of compression were compared back to the original uncompressed images. The trial was successful and the results were shown to be statistically robust. An image quality threshold was calculated from the results of the trial.

3. *The successful design and running of a second human trial to compare the asymmetric compression method with the symmetric method, using human subjects.*

   A second subjective human trial was undertaken. The trial compared asymmetrically compressed images back to their original uncompressed equivalents. Asymmetrically compressed images were picked with equivalent file sizes to those compressed in the symmetric trial. This enabled a direct comparison between the two compression approaches to be undertaken and results collected.

4. *A successful evaluation of whether current monoscopic image comparison metrics are suitable for stereo image evaluation and whether a more suitable metric can be developed to provide the ability to accurately assess the quality of stereo images.*

   Current monoscopic image comparison metrics were evaluated in relation to stereoscopic image quality assessment. Popular metrics such as PSNR and SAD were found not to correlate well with subjective human opinion and it was concluded that a new stereo image quality metric, based upon human visual system properties, should be developed.

5. *A successful investigation of human visual system (HVS) properties leading to the development of a new HVS based metric. This metric needs to accurately predict the preferred amount of compression to be applied to a particular stereo image and produce a threshold of acceptable image quality for a particular stereo image. If this is achieved, then the metric can be classed as successful.*

   Following investigation into the human visual system, the new HVS based metric, Stereo Band Limited Contrast was developed. The new SBLC metric is able to produce a conservative estimate of perceived image quality and from this it was

concluded that SBLC produces a better estimation of the stereoscopic image quality baseline, compared to that produced from the PSNR results. SBLC also produces a useful threshold that can be used as a practical starting point when deciding the required image compression to be used.

8.4 Further Work

The final part of this chapter looks at future work that could be undertaken. It will detail the possible further experiments including the hypothesis, likely experimental methods and the expected outcomes of the work.

8.4.1 Subjective evaluation of the compression of stereoscopic and monoscopic images

A series of symmetrically compressed stereoscopic and monoscopic image pairs should be compared to each other to determine for a set file size, whether 2D or 3D images are preferred. Each monoscopic and stereoscopic image pair will be compressed at a series of different compression levels, with a set file size being maintained between the pairs.

Hypothesis

A null hypothesis would be proposed for this experiment. It is expected that for every compression level and its equivalent file size, the quality of both the monoscopic and stereoscopic images will be perceived as being the same.

Experimental Method

The Double Stimulus method from the ITU Recommendations BT 500 would be used. Each one of the twenty participants would be shown a stereoscopic and monoscopic image and asked to rate them on a scale of 1-100. The compression applied to the stereoscopic and monoscopic images will be varied to enable the file size to be kept constant. In line with the ITU-R recommendations, each session should last no longer than half an hour.

The test images, Masha, Mannequin and Perseus will be used to evaluate the quality of compressed stereoscopic and monoscopic images.

Expected Outcome

It is expected that for high quality, low compressed images, the stereoscopic images will be perceived as better quality. As the compression is increased, at some point it is expected that the monoscopic image will become the one with a better perceived image quality. It is thought that this is likely to be the case as the stereoscopic images would be nearly double the file size for any given compression.

8.4.2 Evaluation of asymmetric compression using the Human Visual System (HVS) stereoscopic image quality metric

The metric should be used to assess the quality of the asymmetrically compressed stereoscopic images and evaluated to see if it is able to accurately predict whether symmetric or asymmetric compression should be applied to a stereoscopic image pair.

Hypothesis

It is hypothesised that the metric will predict that symmetric encoding is the better compression method and the results will be in line with those found from the subjective human trial assessing symmetric and asymmetric stereoscopic compression.

Experimental Method

The test images, Masha, Mannequin and Perseus will be used and the same compression setting used in the subjective study of symmetric and asymmetric compression will be applied. The metric will be used to compare the compressed images to the uncompressed originals and the results compared to those collected using the metric for symmetric compression.

Expected Outcome

It is expected that the metric will be able to produce a conservative image quality threshold for asymmetric stereoscopic compression. It is also expected that the metric results will be in-line with those from the trial, in suggesting that symmetric compression is better, compared to asymmetric. The results collected from the metric will be closer to those from the human trial, compared with PSNR.

8.4.3 Large subjective human study of symmetric and asymmetric image quality

Further study into symmetric and asymmetric stereoscopic image compression using a large human factors trial. This trial would allow small differences in asymmetric compression to be compared to the symmetric equivalent. In the subjective symmetric/asymmetric trials, detailed in Chapter 6, there were instances where small amounts of asymmetric compression produced better results than the symmetric approach. This trial will investigate whether these differences are statistically significant.

Hypothesis

The prediction is that, for both symmetric and asymmetric compression, the perceived quality of the compressed images will decrease with increased JPEG compression.

A null hypothesis, that there will be a no statistically significant difference between images compressed with a small asymmetric difference and their symmetric equivalents is proposed.

Experimental Method

The Double stimulus method from the ITU Recommendations BT 500 would again be used. The trial should have around 200 participants. Each one of the participants would be shown an asymmetrically compressed stereoscopic image and the equivalent uncompressed image and asked to rate them on a scale of 1-100. This would then be repeated with a symmetrically compressed image of the same file size and the uncompressed original. This would enable the two methods to be compared.

Around 20 test images will be used to evaluate the quality of compressed stereoscopic images.

Expected Outcome

It is expected that symmetric compression will out perform asymmetric compression for all image compressions and that there will be a statistically significant difference between images compressed with a small asymmetric difference and their symmetric equivalents.

8.5   Summary

In this chapter the conclusions regarding the research undertaken are presented and the direction for future work outlined. There is current inconsistency in the research regarding symmetric and asymmetric compression approaches to stereoscopic image compression. Existing image quality metrics such as Peak Signal to Noise Ratio (PSNR) are widely used in the assessment of image quality, however, the results from these metrics do not correlate well with subjective human opinion.

The research initially investigated the symmetric and asymmetric compression approaches using the image quality metrics, PSNR and Sum of Absolute Differences (SAD). Due to the criticisms of these metrics, symmetric and asymmetric compression was then investigated using a subjective human trial. From the results of these two experiments it was possible to draw the conclusion that symmetric compression should be used. The only exception to this was that sometimes applying very small amounts of asymmetric difference (i.e. tiny amounts of difference in the compression between the left and right views), to the images, resulted in a better image quality for a set file size at very high compressions.

As PSNR does not correlate well with subjective human opinion, human visual system (HVS) based image quality metrics were investigated and a new HVS stereo image metric, Stereo Band Limited Contrast, was developed. The new SBLC metric, accounts for sensitivity to contrast and luminance changes at regions of high spatial frequency. It extracts edges, corners and regions of high spatial frequency within the image and point matches these regions between the left and right views of the stereo pair. The metric has been shown to correlate well with subjective trial results.

Bibliography

[1] *Methodology for the subjective assessment of the quality of television pictures*, Tech. Report Recommedation ITU-R BT.500-11, International Telecommunications Union, 1998.

[2] *Portable network graphics specification*, Tech. Report Version 1.1, 1999.

[3] *Portable Network Graphics (PNG) Specification (Second Edition)*, 2003, `http://www.w3.org/TR/PNG/`.

[4] *Modelling of the stereoscopic hvs*, Tech. report, Mobile 3DTV, 2009.

[5] *Official jpeg 2000 page*, August 2009, `http://www.jpeg.org/jpeg2000/`.

[6] Michael D Adams, *The JPEG-2000 still image compression standard*, ISO/IEC JTC 1/SC 29/WG 1 N2412 [EB/OL], `http://www.ece.uvic.ca/~mdadams/publications/jp`, pp. 2000–2005.

[7] Michael D. Adams, *The jpeg-2000 still image compression standard*, ISO/IEC JTC 1/SC 29/WG 1 N 2412, 2001.

[8] Hazem Munawer Al-Otum, *Qualitative and quantitative image quality assessment of vector quantization, jpeg, and jpeg2000 compressed images*, Journal of Electronic Imaging **12** (2003), no. 3, 511–521.

[9] Alexandre Benoit, and Patrick Le Callet, and Patrizio Campisi, and Romain Cousseau, *Quality Assessment of Stereoscopic Images*, EURASIP Journal on Image and Video Processing **vol. 2008** (2008), 13 pages.

[10] A. Almohammad and G. Ghinea, *Stego image quality and the reliability of psnr*, 2010, pp. 215–220.

[11] C. Anderson, *JPEG: Joint Photographic Experts Group*, November 2005, `http://www.cs.auckland.ac.nz/compsci708s1c/lectures/jpeg_mpeg/jpeg.html`.

[12] Farshid Arman, Arding Hsu, and Ming-Yee Chiu, *Image processing on compressed data for large video databases*, MULTIMEDIA '93: Proceedings of the first ACM international conference on Multimedia (New York, NY, USA), ACM, 1993, pp. 267–272.

[13] Michael F. Barnsley, *Fractal modelling of real world images*, (1988), 219–242.

[14] S. S. Beauchemin and J. L. Barron, *The Computation of Optical Flow*, 1995.

[15] Brent. Berlin and Paul Kay, *Basic color terms : their universality and evolution / brent berlin and paul kay*, University of California Press, Berkeley :, 1969 (English).

[16] J. M. Bland and D. G. Altman, *Multiple significance tests: the bonferroni method.*, British Medical Journal, 310, 170 (1995).

[17] A. Boev, A. Gotchev, K. Egiazarian, A. Aksay, and G. B. Akar, *Towards compound stereo-video quality metric: a specific encoder-based framework*, SSIAI '06: Proceedings of the 2006 IEEE Southwest Symposium on Image Analysis and Interpretation (Washington, DC, USA), IEEE Computer Society, 2006, pp. 218–222.

[18] Philip J. Bos, *Ss-5 Time Sequential Stereoscopic Displays : The Contribution of Phosphor Persistence to the "Ghost" Image Intensity*, (19910728), no. 27, 603–606.

[19] P. Bourke, *Creating stereoscopic images that are easy on the eyes*, February 2003, `http://astronomy.swin.edu.au/~pbourke/stereographics/goodstereo/`.

[20] David H. Brainard and William T. Freeman, *Bayesian color constancy*, J. Opt. Soc. Am. A **14** (1997), no. 7, 1393–1411.

[21] Camera and Imaging Products Association, *Multi-Picture Format - CIPA DC-007-Translation*, February 2009, `www.cipa.jp/english/hyoujunka/kikaku/pdf/DC-X007_E.pdf`.

[22] Wayne Carlson, *A critical history of Computer Graphics and Animation*, February 2009, `http://design.osu.edu/carlson/history/lesson17.html`.

[23] M. Carnec, P. le Callet, and D. Barba, *An image quality assessment method based on perception of structural information*, 2003, pp. III: 185–188.

[24] Nondestructive Testing (NDT) Resource Centre, *Visual acuity of the human eye*, December 2005, `http://www.ndt-ed.org/EducationResources/CommunityCollege /PenetrantTest/Introduction/visualacuity.htm`.

[25] S Chan, K Martinez, P Lewis, C Lahanier, and J Stevenson, *Handling sub-image queries in content-based retrieval of high resolution art images*, in Proc. Int. Cultural Heritage Informatics Meeting, 2001, pp. 157–163.

[26] S. Cho, K. Yun, C. Ahn, and S. I. Lee, *Disparity-compensated stereoscopic video coding using the mac in mpeg-4*, ETRI Journal **349** (2005), no. 3, 326–329.

[27] J.M. Cobb, *Autostereoscopic desktop display: an evolution of technology.*, Stereoscopic Displays and Applications XVI, Proceedings of SPIE, vol. 5664, 2005, pp. 139–149.

[28] Robert A. Crone, *The history of stereoscopy*, Documenta Ophthalmologica **81** (1992), no. 1, 1–16.

[29] Scott Daly, *The visible differences predictor: an algorithm for the assessment of image fidelity*, (1993), 179–206.

[30] E. Davies, *Machine vision: Theory, algorithms and practicalities*, Academic Press, 1990.

[31] I. Dinstein, G. Guy, J. Rabany, J. Tzelgov, and A. Henik, *On stereo image coding*, In the International Conference on Pattern Recognition, 1988, pp. 357–359.

[32] D. Ezra and G. J. Woodgate, *Autostereoscopic Display Apparatus,*, March 1998, US pat. 5,726,800.

[33] D. Ezra, G. J. Woodgate, B. A. Omar, N. S. Holliman, J. Harrold, and L. S. Shapiro, *New autostereoscopic display system*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (S. S. Fisher, J. O. Merritt, and M. T. Bolas, eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 2409, March 1995, pp. 31–40.

[34] Mark D. Fairchild, Garrett M. Johnson, Jiangtao Kuang, and Hiroshi Yamaguchi, *Image appearance modeling and high-dynamic-range image rendering*, APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization (New York, NY, USA), ACM, 2004, pp. 171–171.

[35] C. Fehn, P. Kauff, Sukhee Cho, Hyoungjin Kwon, Namho Hur, and Jinwoong Kim, *Asymmetric coding of stereoscopic video for transmission over t-dmb*, 3DTV Conference, 2007 (2007), 1–4.

[36] M.A. Fischler and R.C. Bolles, *Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography*, **24** (1981), no. 6, 381–395.

[37] R. Fisher, S. Perkins, A. Walker, and E. Wolfart, *Pixel division*, 2003.

[38] A. Frajka and K. Zeger, *Residual image coding for stereo image compression*, Image Processing. 2002. Proceedings. 2002 International Conference on, vol. 2, 2002, pp. 271–220.

[39] Roelofs G, *Png: The definitive guide*, 1999.

[40] G. A. Geri and Y. Y.Zeevi, *Visual assessment of variable-resolution imagery.*, J Opt Soc Am A **12** (1995), no. 10, 2367–75.

[41] Allen Gersho and Robert M Gray, *Vector quantization and signal compression*, Springer, 1991.

[42] Ovidiu Ghita, Paul F. Whelan, and John Mallon A, *A computational approach for depth from defocus*, Tech. report, 2004.

[43] P. W. Gorley and N. S. Holliman, *Stereoscopic image quality metrics and compression*, Stereoscopic Displays and Applications XIX, SPIE Proceedings, vol. 6803, 2008.

[44] P.W. Gorley and N.S. Holliman, *Investigating symmetric and asymmetric stereoscopic compression using the psnr image quality metric*, Information Sciences and Systems (CISS), 2010 44th Annual Conference on, Mar. 2010, pp. 1 –6.

[45] T. Hattori, T. Ishigaki, K. Shimamoto, A. Sawaki, T. Ishiguchi, and H. Kobayashi, *Advanced autostereoscopic display for G-7 pilot project*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (J. O. Merritt, M. T. Bolas, and S. S. Fisher, eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 3639, May 1999, pp. 66–75.

[46] C.T.E.R. Hewage, S.T. Worrall, S. Dogan, and A.M. Kondoz, *Prediction of stereoscopic video quality using objective quality models of 2-d video*, Electronics Letters **44** (2008), no. 16, 963–965.

[47] Nick Holliman, *Practical Technology: Lecturer at the university of durham, explains digital stereoscopic imaging...*, 3D Displays.

[48] Nick Holliman, *3d display systems*, Tech. report, University of Durham, 2002.

[49] Fred Hsu, *Autostereogram tutorial eye: Binolcular vision diagrams*, March 2005, http://en.wikipedia.org/wiki/Stereopsis.

[50] J. Hsu, Z. Pizlo, C. Babbs, D. Chelberg, and E. Delp, *Design of studies to test the effectiveness of stereo imaging, truth of dare: is stereo viewing really better*, 1994.

[51] I, *Dinstein et al. compression of stereo images and the evaluation of its effects on 3-d perception*, SPIE Applications of Digital Image Processing XII, 1153,522-30, 1989.

[52] CompuServe Incorporated, *Graphics Interchange Format*, July 1990, http://www.w3.org/Graphics/GIF/spec-gif89a.

[53] M.E. Izquierdo, *Stereo image analysis for multi-viewpoint telepresence applications*, **11** (1998), no. 3, 231–254.

[54] Arnaud E. Jacquin, *Fractal image coding based on a theory of iterated contractive image transformations*, vol. 1360, SPIE, 1990, pp. 227–239.

[55] Jianmin Jiang, Mike Reddy, Eran A. Edirisinghe, and Peng Zhang, *Digital imaging and data compression: Object-based stereo image compression*, December 2005, http://www.comp.glam.ac.uk/digimaging/object.htm.

[56] Graham R. Jones, Delman Lee, Nicolas S. Holliman, and David Ezra, *Controlling Perceived Depth in Stereoscopic Images*, Proc. SPIE Stereoscopic Displays and Virtual Reality Systems VIII **4297** (2001), 42–53.

[57] JPEG, *Joint photographic experts group*, February 2009, http://www.jpeg.org/.

[58] B Julesz, *Foundations of cyclopean perception*, 1971.

[59] H. Kalva, L. Christodoulou, and B. Furht, *Evaluation of 3dtv service using asymmetric view coding based on mpeg-2*, 3DTV07, 2007, pp. 1–4.

[60] Hari Kalva, Lakis Christodoulou, Liam M. Mayron, Oge Marques, and Borko Furht, *Design and evaluation of a 3d video system based on h.264 view coding*, NOSSDAV '06: Proceedings of the 2006 international workshop on Network and operating systems support for digital audio and video (New York, NY, USA), ACM, 2006, pp. 1–6.

[61] Lloyd Kaufman (ed.), *Sight and mind: An introduction to visual perception*, University of Illinois Press, 1974.

[62] Myung Jun Kim, Yun Gu Lee, and Jong Beom Ra, *A fast multi-resolution block matching algorithm for multiple-frame motion estimation*, IEICE - Trans. Inf. Syst. **E88-D** (2005), no. 12, 2819–2827.

[63] A. Klein, *Stereoscopy.com - anaglyphs*, February 2005, `http://www.stereoscopy.com/faq/anaglyphs.html`.

[64] Leonid L. Kontsevich, *Pairwise comparison technique: a simple solution for depth reconstruction*, J. Opt. Soc. Am. A **10** (1993), no. 6, 1129–1135.

[65] P. Kortum and W. S. Geisler, *Implementation of a foveated image coding system for image bandwidth reduction*, Human Vision and Electronic Imaging, SPIE Proceedings, vol. 2657, 1996, pp. 350–360.

[66] M. Lambooij, M. Fortuin, W. A. Ijsselsteijn, and I. Heynderickx, *Measuring visual discomfort associated with 3d displays*, vol. 7237, SPIE, 2009, p. 72370K.

[67] D. A. Leopold, J. C. Fitzgibbons, and N. K. Logothetis, *The role of attention in binocular rivalry as revealed through optokinetic nystagmus*, Tech. report, Cambridge, MA, USA, 1995.

[68] Y. Linde, A. Buzo, and R. Gray, *An algorithm for vector quantizer design*, Communications, IEEE Transactions on **28** (1980), no. 1, 84–95.

[69] J. S. Lipscomb and W. L. Wooten, *Reducing crosstalk between stereoscopic views*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (S. S. Fisher, J. O. Merritt, and M. T. Bolas, eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 2177, April 1994, pp. 92–96.

[70] L. Lipton, *Foundations of the stereoscopic cinema*, Van Nostrand Reinhold, 1982.

[71] L. Lipton, *Stereographics, Developers Handbook*, Stereographics Corporation, 1997.

[72] F. Lu, H. Wang, X.Y. Ji, and G.H. Er, *Quality assessment of 3d asymmetric view coding using spatial frequency dominance model*, 3DTV09, 2009, pp. 1–4.

[73] J. Lubin, *A visual discrimination model for imaging system design and evaluation*, Vision Models for Target Detection and Recognition, Eli Peli, Editor, World Scientific, New Jersey (1995), 245–283.

[74] Michael W. Marcellin, Michael J. Gormish, Ali Bilgin, and Martin P. Boliek, *An overview of jpeg-2000*, 2000.

[75] A. Marion, *A introduction to image processing*, Chapman and Hall, 1991.

[76] D. V. Meegan, L. B. Stelmach, and W. J. Tam, *Unequal weighting of monocular inputs in binocular combination: Implications for the compression of stereoscopic imagery*, Journal of Experimental Psychology: Applied **7(2)** (2001), 143–153.

[77] L. M. J. Meesters, W. A. Ijsselsteijn, and P. J. H. Seuntiens, *A survey of perceptual evaluations and requirements of three-dimensional tv*, Circuits and Systems for Video Technology, IEEE Transactions on **14** (2004), no. 3, 381–391.

[78] Philipp Merkle, Karsten Müller, Aljoscha Smolic, and Thomas Wiegand.

[79] A. A. Michelson (ed.), *Studies in optics*, U. Chicago Press, Chicago, Ill., 1927.

[80] Robin Morgan, *Depth Perception*, 1st ed., Doubleday, 1982.

[81] T. Motoki, H. Isono, and I. Yuyama, *Present status of three-dimensional television research*, Proceedings of the IEEE **83** (1995), no. 7, 1009–1021.

[82] Motta, *A standard default color space for the internet: srgb, available at http://www.w3.org/graphics/color/srgb.html*, 1996.

[83] MPEG, *Moving Picture Experts Group*, February 2009, `http://www.chiariglione.org/mpeg/`.

[84] Gerald M. Murch, *Review: [untitled]*, The American Journal of Psychology **87** (1974), no. 4, 742–746.

[85] Takeshi NAEMURA, Masahide KANEKO, and Hiroshi HARASHIMA, *Compression and representation of 3-d images (surveys on image processing technologies : Algorithms, sensors and applications)*, IEICE transactions on information and systems **82** (19990325), no. 3, 558–567.

[86] G. Nyman, J. Radun, T. Leisti, J. Oja, H. Ojanen, J. L. Olives, T. Vuori, and J. Hkkinen, *What do users really perceive: Probing the subjective image quality*, Image Quality and System Performance III. L. C. Cui and Y. Miyake, 2006, pp. 1–7.

[87] University of Glamorgan, *Digital imaging and data compression*, 2005.

[88] Virginia E. Ogle and Michael Stonebraker, *Chabot: Retrieval from a relational database of images*, 1995.

[89] W. Osberger, A. J. Maeder, and D. McLean, *An objective quality assessment technique for Digital image sequences*, Proceedings ICIP-96 (IEEE International Conference on Image Processing) (Lausanne, Switzerland), vol. I, 1996, pp. 897–900.

[90] W. Osberger, A. J. Maeder, and D. McLean, *A computational model of the human visual system for image quality assessment*, Proceedings DICTA-97 (Auckland, New Zealand), 1997, pp. 337–342.

[91] Wilfried Osberger, Anthony J. Maeder, and Donald McLean, *An objective quality assessment technique for digital image sequences*, Proceedings ICIP-96, 1996, pp. 16–19.

[92] Greg Pass, Ramin Zabih, and Justin Miller, *Comparing images using color coherence vectors*, 1996.

[93] S. Pastoor, *3d-television: A survey of recent research results on subjective requirements*, **4** (1991), no. 1, 21–32.

[94] E. Peli, *Contrast in complex images*, Opt. Soc. Am. **7** (1990), no. 10, 2032–2040.

[95] Eli Peli, *In search of a contrast metric: Matching the perceived contrast of gabor patches at different phases and bandwidths*, Vision Research **37** (1997), no. 23, 3217 – 3224.

[96] M G Perkins, *Data compression of stereopairs*, IEEE Transactions on communications (1992), no. 40, 684–696.

[97] M. P. Peterson, *Maps and the Internet*, International Cartographic Association, 2005.

[98] Jonathan David Pfautz, *Depth perception in computer graphics*, Tech. report, 2000.

[99] Adrian Gh. Podoleanu and David A. Jackson, *Noise analysis of a combined optical coherence tomograph and a confocal scanning ophthalmoscope*, Appl. Opt. **38** (1999), no. 10, 2116–2127.

[100] Michel Pommeray, Jean-Claude Kastelik, and Marc Georges Gazalet, *Image crosstalk reduction in stereoscopic laser-based display systems*, Journal of Electronic Imaging **12** (2003), no. 4, 689–696.

[101] Charles Poynton, *Digital video and hdtv algorithms and interfaces*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.

[102] B. Reusch, *Fuzzy impulse noise reduction methods for color images*, Springer Berlin Heidelberg, 2006, ISBN 978-3-540-34780-4.

[103] William Reynolds and Robert V. Kenyon, *The wavelet transform and the suppression theory of binocular vision for stereo image compression*, 1996.

[104] Rhcastilhos, *Schematic diagram of the human eye in english*, December 2009, `http://en.wikipedia.org/wiki/Eye`.

[105] Greg Roelofs, *Linux Gazette: History of the Portable Network Graphics (PNG) Format*, Linux J., 19.

[106] Dietmar Saupe and Raouf Hamzaoui, *A review of the fractal image compression literature*, SIGGRAPH Comput. Graph. **28** (1994), no. 4, 268–276.

[107] H S Sawhney, Y Guo, K Hanna, R Kumar, S Adkins, and S Zhou, *Hybrid stereo camera: an ibr approach for synthesis of very high resolution stereoscopic image sequences*, In SIGGRAPH, Press, 2001, pp. 451–460.

[108] Kim Scarborough, *Anaglyph 3-d photo of edward kemeys's lion statue outside the art institute of chicago, in illinois.*, August 2009, `http://commons.wikimedia.org/wiki/File:Art_Institute_of_Chicago_Lion_Statue_%28parallel_stereo_pair%29.jpg`.

[109] G. Schaefer, *Cvpic colour/shape histograms for compressed domain image retrieval*, 2004, pp. 424–431.

[110] Willem van Schaik, *Portable Network Graphics*, 2000, `http://www.schaik.com/png/`.

[111] W. Schierber, *Fundementals of electronic imaging systems*, Springel, New York, 1993.

[112] O. Schreer, P. Kauff, and T. Sikora (eds.), *3d video communications*, Wiley, 2005.

[113] S. Sethuraman, Mel Siegel, and Angel Jordan, *A multiresolution framework for stereoscopic image sequence compression*, Proc. ICIP-94, vol. 2, November 1994, pp. 361 – 365.

[114] P. Seuntiens, *Visual experience of 3d tv*, Ph.D. thesis, Eindhoven University, Eindhoven, The Netherlands, 2006.

[115] Pieter Seuntiens, Lydia Meesters, and Wijnand Ijsselsteijn, *Perceived quality of compressed stereoscopic images: Effects of symmetric and asymmetric JPEG coding and camera separation*, ACM Trans. Appl. Percept. **3** (l 06), no. 2, 95–109.

[116] D. Shah and N. A. Dodgson, *Issues in multiview autostereoscopic image compression*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (A. J. Woods, M. T. Bolas, J. O. Merritt, & S. A. Benton, ed.), Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, vol. 4297, June 2001, pp. 307–316.

[117] H. Shao, X. Cao, and G.H. Er, *Objective quality assessment of depth image based rendering in 3dtv system*, 3DTV09, 2009, pp. 1–4.

[118] Einat Shneor and Shaul Hochstein, *Effects of eye dominance in visual perception*, International Congress Series **1282** (2005), 719 – 723, Vision 2005 - Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK.

[119] M. W. Siegel, Priyan Gunatilake, and Sriram Sethuraman, *Compression of stereo image pairs and streams*, In Stereoscopic Displays and Applications V, 1994, pp. 258–268.

[120] Mel Siegel, S. Sethuraman, J. S. McVeigh, and Angel Jordan, *Compression and interpolation of 3d-stereoscopic and multi-view video*, Stereoscopic Displays and Virtual Reality Systems IV, vol. 3012, February 1997, pp. 227 – 238.

[121] Donald E. Simanick, *The dangers of analogies*, August 2009, `http://www.lhup.edu/~dsimanek/scenario/analogy.htm`.

[122] Jon Siragusa, David C Swift, Bob Akka, Dave Milici, and Andy Spencer, *General Purpose Stereoscopic Data Descriptor*, 1997.

[123] M. Slanina and V. Ricny, *A comparison of full-reference image quality assessment methods*, Electronic Communication Systems and New Generation Technology (ELKOM), 2006.

[124] L. B. Stelmach and W. J. Tam, *Stereoscopic image coding: effect of disparate image-quality in left- and right-eye views.*, Signal Processing: Image Communications **14** (1998), 111–117.

[125] L. B. Stelmach, W. J. Tam, and D. V. Meegan, *Stereo image quality: effects of spatio-temporal resolution*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series (J. O. Merritt, M. T. Bolas, and S. S. Fisher, eds.), Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 3639, May 1999, pp. 4–11.

[126] Lew B. Stelmach, Wa James Tam, Daniel V. Meegan, André Vincent, and Philip Corriveau, *Human perception of mismatched stereoscopic 3d inputs*, ICIP, 2000.

[127] Markus Stricker and Michael Swain, *The capacity of color histogram indexing*, In Proceedings 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1994, pp. 704–708.

[128] Ivan E. Sutherland, *The Ultimate Display*, Proceedings of IFIP, 1965, pp. 506–508.

[129] Ivan E. Sutherland, *A Head-mounted Three Dimensional Display*, (1998), 295–302.

[130] Michael J. Swain and Dana H. Ballard, *Color indexing*, International Journal of Computer Vision **7** (1991), 11–32.

[131] D.J. Tolhurst and Y. Tadmor, *Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra*, Vision Research **37** (1997), no. 23, 3203 – 3215.

[132] C S Tong and M Pi, *Fast fractal image encoding based on adaptive search*, IEEE Trans. on Image Proc (2001), no. 10, 1269–1277.

[133] D. Tzovaras, N. Grammalidis, and M.G. Strintzis, *Disparity field and depth map coding for multiview image sequence compression*, 1996, pp. II: 887–890.

[134] E. Vrscay, *A hitchhiker's guide to fractal-based function approximation and image compression*, Math Ties: Faculty of Mathematics Alumni News, 1995.

[135] N. J. Wade, *On the late invention of the stereoscope*, Perception **16** (1987), 785–818.

[136] Gregory K. Wallace, *The JPEG still picture compression standard*, Commun. ACM **34** (1991), no. 4, 30–44.

[137] David Waller, *Factors affecting the perception of interobject distances in virtual environments*, Presence: Teleoper. Virtual Environ. **8** (1999), no. 6, 657–670.

[138] B. A. Wandell, *Foundations of vision*, Sinauer Associates, Inc., 1995.

[139] J. Wang and M. Fischler, *Visual similarity, judgmental certainty and stereo correspondence*, 1998.

[140] Z. Wang and E.P. Simoncelli, *An adaptive linear system framework for image distortion analysis*, 2005, pp. III: 1160–1163.

[141] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, Student Member, Eero P. Simoncelli, and Senior Member, *Image quality assessment: From error visibility to structural similarity*, IEEE Transactions on Image Processing **13** (2004), 600–612.

[142] Zhou Wang, Ligang Lu, and Alan C. Bovik, *Video quality assessment based on structural distortion measurement*, 2004.

[143] Zhou Wang, Hamid R. Sheikh, and Alan C. Bovik, *Bovik, objective video quality assessment*, in The Handbook of Video Databases: Design and Applications, CRC Press, 2003, pp. 1041–1078.

[144] Zhou Wang, Eero P. Simoncelli, and Alan C. Bovik, *Multi-scale structural similarity for image quality assessment*, in Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers, (Asilomar, 2003, pp. 1398–1402.

[145] C. Ware, *Information visualization, perception for design*, Morgan Kaufmann, 1999, ISBN 0-1-55860-511-8.

[146] A. B. Watson, *Proposal: Measurement of a JND scale for video quality*, 2000.

[147] M. A. Webster, O. H. MacLin, A. L. Rees, and V. E. Raker, *Contrast adaptation and the spatial structure of natural images*, Perception **25** (1996).

[148] Terry A. Welch, *High speed data compression and decompression apparatus and method*, 1983, `http://patft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=4558302.PN.&OS=PN/4558302&RS=PN/4558302`.

[149] S.J.P. Westen, R.L. Lagendijk, and J. Biemond, *Perceptual image quality based on a multiple channel HVS model*, Proceedings of ICASSP'95, 1995, pp. 2351–2354.

[150] Charles Wheatstone, *Constribdistions to the physiology of vision—part the second. on some remarkable, and hitherto unobserved, phenomena of binocular vision*, Philosophical Transactions of the Royal Society of London **142** (1852), 1–17.

[151] Wiley, *Human Visual System*, February 2009, `http://upload.wikimedia.org/wikipedia/commons/8/80/Wiley_Human_Visual_System.gif`.

[152] Hugh R. Wilson, Frances Wilkinson, and Wael Asaad, *Concentric orientation summation in human form vision*, Vision Research **37** (1997), no. 17, 2325 – 2330.

[153] Stefan Winkler, Animesh Sharma, and David Mcnally, *Perceptual video quality and blockiness metrics for multimedia streaming applications*, October 22 2001.

[154] Stefan Winkler and Pierre Vandergheynst, *Computing isotropic local contrast from oriented pyramid decompositions*, in Proc. ICIP, 1999, pp. 420–424.

[155] P. H. Winston and B. K. Horn, *Lisp*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1988.

[156] W. Woo and A. Ortega, *Overlapped block disparity compensation with adaptive windows for stereo image coding*, **10** (2000), no. 2, 194.

[157] Günther Wyszecki and W. S. Stiles, *Color science: Concepts and methods, quantitative data and formulae (wiley series in pure and applied optics)*, 2 ed., Wiley-Interscience, August 2000.

[158] S. Yano and I Yuyama, *Stereoscopic HDTV: Experimental system and psychological effects*, SMTPE J. **100** (1991), 14–18.

Appendix A

Symmetric and Asymmetric Compression - Objective Data

Table A.1: Symmetric vs. Asymmetric JPEG Compression Comparison Results, *Masha*, *Mannequin* and *Perseus*.

Test Image	File Size (Kb)	Compression (L,R)		PSNR (dB)	
		Symmetric	Asymmetric	Symmetric	Asymmetric
Masha	405	(6,6)	(10,4)	36.15	35.32
	190	(20,20)	(94,8)	30.55	19.55
	111	(31,31)	(99,16)	27.12	18.65
	70	(45,45)	(31,99)	24.09	18.37
	66	(47,47)	(33,97)	23.73	18.53
	56	(56,56)	(39,99)	22.40	18.22
	51	(62,62)	(44,99)	21.54	18.11
	45	(70,70)	(53,99)	20.45	17.91
	38	(84,84)	(76,99)	18.22	17.17
	33	(99,99)	(98,99)	15.71	15.74
Mannequin	238	(4,4)	(1,30)	41.01	37.17
	164	(8,8)	(2,99)	39.12	25.82
	108	(18,18)	(6,99)	36.55	18.82
	90	(24,24)	(9,99)	35.59	18.81
	76	(34,34)	(13,99)	34.23	18.80
	68	(42,42)	(15,99)	33.67	18.80
	60	(54,54)	(21,99)	31.84	18.79
	52	(66,66)	(28,99)	31.16	18.78
	50	(70,70)	(31,99)	29.97	18.76
	40	(80,80)	(54,99)	28.96	18.72
	26	(93,93)	(89,99)	22.84	18.47
Perseus	377	(5,5)	(2,20)	32.77	32.57
	233	(14,14)	(5,96)	27.59	18.46
	103	(25,25)	(11,99)	26.23	17.65
	87	(34,34)	(20,99)	24.92	17.3
	57	(40,40)	(25,98)	22.75	17.22
	50	(49,49)	(31,99)	21.32	16.92
	47	(58,58)	(44,96)	21.14	16.67
	43	(68,68)	(52,99)	20.33	16.49
	38	(82,82)	(74,98)	16.94	15.57
	31	(96,96)	(95,99)	15.53	14.59
Compression (6,6) is a JPEG compression (IJG) amount of 6 applied to the left image and a compression amount of 6 applied to the right. Compression (10,4) is a compression amount of 10 applied to the left image and a compression amount of 4 applied to the right. The higher the number the greater the compression.					

Table A.2: Symmetric vs. Asymmetric JPEG 2000 Compression Comparison Results, *Masha*, *Mannequin* and *Perseus*.

Test Image	File Size (Kb)	Compression (L,R)		PSNR (dB)	
		Symmetric	Asymmetric	Symmetric	Asymmetric
Masha	167	(42,42)	(1,59)	32.81	31.85
	138	(52,52)	(3,56)	29.26	18.16
	109	(62,62)	(1,70)	25.36	18.01
	91	(68,68)	(2,78)	23.73	17.68
	74	(74,74)	(5,81)	22.39	17.93
	65	(77,77)	(24,89)	21.16	17.85
	50	(82,82)	(62,98)	20.30	17.64
	39	(86,86)	(68,99)	18.70	17.54
	27	(90,90)	(81,99)	16.81	16.52
Mannequin	112	(63,63)	(2,62)	35.65	18.44
	101	(67,67)	(4,53)	35.46	18.18
	88	(72,72)	(3,71)	33.24	18.02
	79	(77,77)	(26,83)	30.45	17.94
	65	(80,80)	(35,82)	29.89	17.77
	50	(86,86)	(44,89)	28.61	17.76
	42	(89,89)	(64,92)	28.10	17.65
	31	(92,92)	(67,98)	26.33	17.23
	23	(96,96)	(85,99)	21.88	17.05
Perseus	132	(12,12)	(4,94)	27.18	18.38
	120	(21,21)	(8,97)	24.63	16.38
	105	(35,35)	(23,97)	22.45	16.33
	91	(42,42)	(28,99)	21.51	16.13
	78	(50,50)	(37,96)	21.03	15.88
	66	(57,57)	(46,99)	20.13	15.07
	50	(71,71)	(49,98)	19.21	14.93
	39	(84,84)	(68,96)	16.81	14.62
	26	(93,93)	(91,99)	15.08	13.78

Table A.3: Symmetric vs. Depth Map JPEG Compression SAD Results, *Masha*.

File	Compression (L,R)		Sum of Absolute Differences (SAD) (%)	
Size (Kb)	Symmetric	Asymmetric	Symmetric	Asymmetric
405	(6,6)	(10,4)	0.81	1.16
190	(20,20)	(94,8)	0.82	6.05
111	(31,31)	(99,16)	0.82	6.87
70	(45,45)	(31,99)	0.86	7.47
66	(47,47)	(33,97)	0.90	7.45
56	(56,56)	(39,99)	0.93	7.84
51	(62,62)	(44,99)	0.93	8.08
45	(70,70)	(53,99)	0.93	8.49
38	(84,84)	(76,99)	0.95	9.83
33	(99,99)	(98,99)	0.96	12.01

Appendix B

Symmetric and Asymmetric Compression - Subjective Data

Table B.1: *Masha*: Mean score (%) and Standard Deviation from Symmetric Subjective Trial.

Image	Symmetric Compression	File Size (kB)	Mean Score	St. Dev.
Masha	5	348	2.38	5.6282
	10	239	2.93	8.3923
	15	187	4.25	10.7650
	20	159	3.48	5.9008
	25	139	6.63	11.1486
	30	126	6.45	7.7458
	35	116	15.90	13.7464
	40	106	14.85	15.4779
	45	100	18.50	13.7281
	50	94	21.65	15.3565
	55	88	23.58	13.2159
	60	82	25.88	13.4626
	65	76	29.28	14.4257
	70	70	32.78	12.3278
	75	62	37.00	13.6495
	80	54	36.98	12.7490
	85	46	45.43	16.9341
	90	38	50.15	14.9058
	95	28	65.20	18.3767
	100	24	68.88	16.6967

Table B.2: *Mannequin*: Mean score (%) and Standard Deviation from Symmetric Subjective Trial.

Image	Symmetric Compression	File Size (kB)	Mean Score	St. Dev.
Mannequin	5	534	4.28	7.4764
	10	407	4.35	8.9573
	15	345	5.08	8.7366
	20	317	5.70	7.0971
	25	292	13.93	16.2266
	30	249	12.48	14.3259
	35	200	10.78	11.8808
	40	191	8.50	9.1399
	45	185	7.65	10.0551
	50	179	12.75	14.8475
	55	172	15.40	14.9423
	60	147	16.00	10.3006
	65	138	16.15	10.9744
	70	132	18.23	11.7941
	75	111	23.43	13.0106
	80	98	27.65	13.1199
	85	80	39.83	13.5588
	90	58	43.68	15.0731
	95	32	55.55	18.9046
	100	25	70.85	18.2385

Table B.3: *Perseus*: Mean score (%) and Standard Deviation from Symmetric Subjective Trial.

Image	Symmetric Compression	File Size (kB)	Mean Score	St. Dev.
Perseus	5	141	3.00	6.1644
	10	105	1.83	2.4378
	15	89	1.95	4.0696
	20	80	3.05	4.8302
	25	73	4.50	8.0192
	30	69	6.40	10.7150
	35	66	4.78	7.0146
	40	62	8.13	10.1582
	45	60	5.65	7.0875
	50	58	8.65	10.0296
	55	56	10.43	11.6770
	60	54	10.88	9.4846
	65	52	12.98	13.9127
	70	49	12.23	12.2275
	75	46	16.80	11.5807
	80	44	20.25	14.4980
	85	40	26.45	15.0451
	90	36	33.08	12.0456
	95	32	46.38	17.6398
	100	28	59.45	20.1086

Table B.4: *Masha*: Asymmetric Results from Subjective Trial.

Symmetric Compression	File Size (kB)	Asymmetric Compression	Asymmetric Difference	Mean Score	St. Dev.
(75,75)	62	(80,67)	13	25.1	7.5
		(67,80)	13	26.4	3.08
		(64,61)	23	27.55	5.4
		(60,86)	26	41	4.98
		(41,98)	57	66.9	8.4
		(97,40)	57	63.7	6.08
		(41,100)	59	67.35	6.98
		(100,40)	60	53.9	16.15
(65,65)	76	(64,61)	3	17.95	3.55
		(60,66)	6	18.9	3.43
		(44,80)	36	26.55	4.64
		(80,44)	36	28.1	4.01
		(41,82)	41	34.75	5.44
		(83,40)	43	27.55	3.61
		(28,100)	72	72.3	8.69
		(100,26)	74	66.7	11.63
(50,50)	94	(41,55)	14	14.6	3.49
		(57,40)	17	12.05	1.54
		(60,38)	22	14.9	3.7
		(37,61)	24	12.65	3.8
		(27,80)	53	28	3.95
		(84,26)	54	25.55	4.47
		(19,100)	81	56.6	15.77
		(100,18)	82	69.9	8.46
(30,30)	126	(23,40)	17	9.25	3.45
		(41,23)	18	12.95	3.65
		(60,18)	42	7.4	1.27
		(18,61)	43	15.35	2.92
		(80,14)	66	32.25	5.35
		(14,80)	66	29.95	5.84
		(100,11)	89	71.45	10.24
		(10,100)	90	70.6	9.44
(25,25)	139	(18,40)	22	5.4	1.73
		(41,28)	23	8.55	4.14
		(60,14)	46	11.8	4.1
		(14,61)	47	12.7	3.6
		(80,11)	69	23.95	3.68
		(11,80)	69	25.6	3.86
		(9,100)	91	71.7	6.59
		(100,8)	92	73.95	8.09

Table B.5: *Masha*: t-test Analysis - Comparison of Symmetric and Asymmetric Compression Approaches.

File Size (kB)	Compression	Mean	St. Dev	95% Confidence	Max	Min	t-test
62	Symmetric	38.23	12.84	5.63	70	16.5	0.001026
	Asymmetric	25.1	7.5	3.29	41	0	
76	Symmetric	29.28	12.63	5.53	69.5	11	0.001067
	Asymmetric	17.95	3.55	1.55	26	9	
94	Symmetric	21.65	13.72	6.01	52.5	2	0.004934
	Asymmetric	12.05	1.54	0.67	15	9	
126	Symmetric	6.45	6.41	2.81	20	0	0.526540
	Asymmetric	7.4	1.27	0.56	9	5	
139	Symmetric	6.63	9.5	4.16	29.5	0	0.560787
	Asymmetric	5.4	1.73	0.76	9	1	

Table B.6: *Mannequin*: Asymmetric Results from Subjective Trial.

Symmetric Compression	File Size (kB)	Asymmetric Compression	Asymmetric Difference	Mean Score	St. Dev.
(85,85)	80	(91,80)	11	23.5	3.58
		(79,92)	13	26.9	4.2
		(92,79)	13	26.85	2.83
		(71,96)	25	43.65	6.73
		(97,71)	26	54.95	6.74
		(70,97)	27	64.7	10.06
		(69,100)	31	75.8	8.78
		(100,69)	31	76.9	9.34
(80,80)	98	(79,84)	5	22.9	6.02
		(84,79)	5	20.35	3.5
		(90,71)	19	26.7	3.25
		(70,90)	20	30.75	4.2
		(91,68)	23	29.95	3.91
		(57,96)	39	39.75	4.68
		(100,57)	43	74.95	6.72
		(53,100)	47	78.05	9.89
(75,75)	111	(76,79)	3	16.05	2.68
		(79.75)	4	19.55	4.98
		(84,71)	13	23.65	4.15
		(70,84)	14	17.65	4.43
		(91,58)	33	18.45	3.02
		(45,96)	51	43.15	7.58
		(42,100)	58	70.55	10.64
		(100,39)	61	75.8	8.91
(50,50)	179	(34,71)	37	8.55	2.5
		(70,32)	38	6.95	2.39
		(79,31)	48	12.1	2.9
		(28,79)	51	7.95	2.56
		(91,25)	66	25	7.52
		(17,100)	73	67.8	8.04
		(18,96)	78	37.55	4.47
		(100,19)	81	78.7	10.01
(45,45)	185	(70,31)	39	9.2	1.64
		(31,71)	40	7.45	1.7
		(79,29)	50	9.15	1.5
		(26,79)	53	17.2	3.47
		(91,23)	68	15.25	5.29
		(17,96)	79	34.9	5.59
		(100,17)	83	79.4	8.73
		(16,100)	84	79.9	9.76

Table B.7: *Mannequin*: t-test Analysis - Comparison of Symmetric and Asymmetric Compression Approaches.

File Size (kB)	Compression	Mean	St. Dev	95% Confidence	Max	Min	t-test
80	Symmetric	39.83	11.41	5	66	24	0.000004
	Asymmetric	23.5	3.58	1.57	31	15	
98	Symmetric	27.65	9.88	4.33	46	10	0.004059
	Asymmetric	20.35	3.5	1.53	31	14	
111	Symmetric	23.43	11.3	4.95	42.5	0	0.010627
	Asymmetric	16.05	2.68	1.18	22	8	
179	Symmetric	12.75	12.21	5.35	51	0	0.026666
	Asymmetric	6.95	2.39	1.05	14	0	
185	Symmetric	7.65	7.7	3.37	27.5	0	0.906604
	Asymmetric	7.45	1.7	0.75	10	3	

Table B.8: *Perseus*: Asymmetric Results from Subjective Trial.

Symmetric Compression	File Size (kB)	Asymmetric Compression	Asymmetric Difference	Mean Score	St. Dev.
(80,80)	44	(85,73)	12	10.3	3.51
		(72,86)	14	12.8	4.55
		(87,71)	16	12.7	4.33
		(70,88)	18	14.6	5.67
		(75,100)	25	64	5.87
		(100,57)	43	46.7	6.5
		(56,100)	44	54.5	7.96
		(100,55)	45	45.85	10.94
(70,70)	49	(75,71)	4	10.95	2.98
		(70,76)	6	11.9	5.38
		(85,60)	25	14.15	3.75
		(59,86)	27	12.6	4.62
		(88,57)	31	15	5.03
		(56,88)	32	21.6	6.13
		(100,43)	57	56.55	6.94
		(43,100)	57	50.25	10.84
(60,60)	54	(70,64)	6	7.85	3.2
		(63,71)	8	6.75	3.84
		(77,57)	20	9.1	3.6
		(56,77)	21	9	4.51
		(85,46)	39	13.3	3.76
		(45,86)	41	10.5	5.42
		(100,33)	67	55.7	8.58
		(33,100)	67	50.45	8.29
(50,50)	58	(66,57)	9	6.65	3.83
		(56,68)	12	8.1	3.91
		(70,53)	17	5.9	2.27
		(51,71)	20	5.3	2.18
		(85,37)	48	9	4.86
		(36,86)	50	10.4	4.42
		(100,28)	72	41.95	11.44
		(28,100)	72	54.55	8.05
(40,40)	62	(56,57)	1	8.9	4.36
		(54,57)	3	6.95	4.06
		(70,42)	28	7.4	3.83
		(41,71)	30	5.6	3.08
		(85,31)	54	7.8	3.55
		(30,86)	56	13.1	3.84
		(100,24)	76	54	7.45
		(24,100)	76	57.75	7.7

Table B.9: *Perseus*: t-test Analysis - Comparison of Symmetric and Asymmetric Compression Approaches.

File Size (kB)	Compression	Mean	St. Dev	95% Confidence	Max	Min	t-test
44	Symmetric	20.25	13.65	5.98	52.5	0	0.007086
	Asymmetric	10.3	3.51	1.54	21	3	
49	Symmetric	12.23	9.97	4.37	38	0	0.602014
	Asymmetric	10.95	2.98	1.31	20	4	
54	Symmetric	10.88	8.43	3.69	24.5	0	0.042539
	Asymmetric	6.75	3.84	1.68	20	0	
58	Symmetric	8.68	8.12	3.56	25	0	0.058049
	Asymmetric	5.3	2.18	0.95	12	0	
62	Symmetric	8.13	7.25	3.18	23.5	0	0.114773
	Asymmetric	5.6	3.08	1.35	16	0	

Appendix C

Evaluation of Stereo Band Limited Contrast - Data

Table C.1: *Masha*: PSNR, SBLC and Mean score for Symmetric Human Trial (%).

Image	Compression	File Size (kB)	PSNR (dB)	1/PSNR	SBLC	Mean Score
Masha	5	348	36.1514	0.0277	0.7020	2.38
	10	239	33.4999	0.0299	0.7121	2.93
	15	187	30.5498	0.0327	0.7154	4.25
	20	159	28.5449	0.0350	0.7176	3.48
	25	139	27.1191	0.0369	0.7176	6.63
	30	126	25.5350	0.0392	0.7188	6.45
	35	116	24.0941	0.0415	0.7233	15.90
	40	106	23.8515	0.0419	0.7255	14.85
	45	100	23.7328	0.0421	0.7302	18.50
	50	94	22.9867	0.0435	0.7371	21.65
	55	88	22.4023	0.0446	0.7430	23.58
	60	82	21.9569	0.0455	0.7546	25.88
	65	76	21.5372	0.0464	0.7541	29.28
	70	70	21.2386	0.0471	0.7571	32.78
	75	62	20.4507	0.0489	0.7581	37.00
	80	54	19.4688	0.0514	0.7642	36.98
	85	46	18.2206	0.0549	0.7661	45.43
	90	38	16.8679	0.0593	0.7805	50.15
	95	28	15.7100	0.0637	0.7878	65.20
	100	24	12.0158	0.0832	0.8203	68.88

Table C.2: *Mannequin*: PSNR, SBLC and Mean score for Symmetric Human Trial (%).

Image	Compression	File Size (kB)	PSNR (dB)	1/PSNR	SBLC	Mean Score
Mannequin	5	534	41.0067	0.0244	0.6892	4.28
	10	407	39.1169	0.0256	0.7020	4.35
	15	345	37.2682	0.0268	0.7075	5.08
	20	317	36.5506	0.0274	0.7121	5.70
	25	292	36.0279	0.0278	0.7131	13.93
	30	249	35.5902	0.0281	0.7154	12.48
	35	200	34.4868	0.0290	0.7165	10.78
	40	191	34.2350	0.0292	0.7176	8.50
	45	185	34.2445	0.0292	0.7188	7.65
	50	179	33.6653	0.0297	0.7222	12.75
	55	172	33.1562	0.0302	0.7244	15.40
	60	147	31.8382	0.0314	0.7262	16.00
	65	138	31.5986	0.0316	0.7323	16.15
	70	132	31.1555	0.0321	0.7371	18.23
	75	111	30.2580	0.0330	0.7430	23.43
	80	98	29.9678	0.0334	0.7538	27.65
	85	80	29.5324	0.0339	0.7571	39.83
	90	58	28.9607	0.0345	0.7642	43.68
	95	32	25.8603	0.0387	0.7806	55.55
	100	25	22.8391	0.0438	0.8033	70.85

Table C.3: *Perseus*: PSNR, SBLC and Mean score for Symmetric Human Trial (%).

Image	Compression	File Size (kB)	PSNR (dB)	1/PSNR	SBLC	Mean Score
Perseus	5	141	33.1788	0.0301	0.5362	3.00
	10	105	30.8679	0.0324	0.5610	1.83
	15	89	26.5209	0.0377	0.5625	1.95
	20	80	24.0258	0.0416	0.5714	3.05
	25	73	22.5513	0.0443	0.5714	4.50
	30	69	22.5315	0.0444	0.5714	6.40
	35	66	22.4935	0.0445	0.5833	4.78
	40	62	21.9348	0.0456	0.5882	8.13
	45	60	21.6061	0.0463	0.5882	5.65
	50	58	21.9868	0.0455	0.5882	8.65
	55	56	22.1794	0.0451	0.5942	10.43
	60	54	20.1458	0.0496	0.5942	10.88
	65	52	17.0335	0.0587	0.5946	12.98
	70	49	16.6895	0.0599	0.5949	12.23
	75	46	16.1708	0.0618	0.6119	16.80
	80	44	15.9868	0.0626	0.7241	20.25
	85	40	15.4451	0.0647	0.9231	26.45
	90	36	15.3499	0.0651	0.9630	33.08
	95	32	15.5204	0.0644	0.9643	46.38
	100	28	13.9868	0.0715	0.9672	59.45