# Visual Quality with a Focus on 3D Blur Discrimination and Texture Granularity

by

Mahesh M. Subedar

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved November 2015 by the
Graduate Supervisory Committee:

Lina Karam, Chair
Glen Abousleman
Baoxin Li
Martin Reisslein

ARIZONA STATE UNIVERSITY

December 2015

ABSTRACT

Blur is an important attribute in the study and modeling of the human visual system. In this work, 3D blur discrimination experiments are conducted to measure the just noticeable additional blur required to differentiate a target blur from the reference blur level. The past studies on blur discrimination have measured the sensitivity of the human visual system to blur using 2D test patterns. In this dissertation, subjective tests are performed to measure blur discrimination thresholds using stereoscopic 3D test patterns. The results of this study indicate that, in the symmetric stereo viewing case, binocular disparity does not affect the blur discrimination thresholds for the selected 3D test patterns. In the asymmetric viewing case, the blur discrimination thresholds decreased and the decrease in threshold values is found to be dominated by the eye observing the higher blur.

The second part of the dissertation focuses on texture granularity in the context of 2D images. A texture granularity database referred to as GranTEX, consisting of textures with varying granularity levels is constructed. A subjective study is conducted to measure the perceived granularity level of textures present in the GranTEX database. An objective index that automatically measures the perceived granularity level of textures is also presented. It is shown that the proposed granularity metric correlates well with the subjective granularity scores and outperforms the other methods presented in the literature.

A subjective study is conducted to assess the effect of compression on textures with varying degrees of granularity. A logarithmic function model is proposed as a fit to the subjective test data. It is demonstrated that the proposed model can be used for rate-distortion control by allowing the automatic selection of the needed compression ratio for a target visual quality. The proposed model can also be used

for visual quality assessment by providing a measure of the visual quality for a target compression ratio.

The effect of texture granularity on the quality of synthesized textures is studied. A subjective study is presented to assess the quality of synthesized textures with varying levels of texture granularity using different types of texture synthesis methods. This work also proposes a reduced-reference visual quality index referred to as delta texture granularity index for assessing the visual quality of synthesized textures.

*In Memory of My Father*

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere and heartfelt thanks to Prof. Lina Karam for providing me the opportunity to pursue research work under her supervision. It is her constant guidance and support, at the same time giving the freedom to purse interesting research topics is what made this PhD journey gratifying and memorable. Her immense patience, caring nature and encouragement in conducting the research is what made this dissertation possible. Her flexibility to have long telephonic meetings even outside the office hours is what allowed me to purse PhD studies while working fulltime at Intel.

I would like to express my gratitude to members of my PhD committee, Prof. Glen Abousleman, Prof. Baoxin Li and Prof. Martin Reisslein for their guidance and constructive feedback. I am very grateful for having the opportunity to work closely with them on some of the past projects.

I would like to express my sincere thanks to Intel Corporation and to my past/present managers at Intel for supporting my PhD studies. Most importantly, I would like to acknowledge and thank Dr. Jorge Caviedes for being my manager and mentor at Intel, who encouraged me to purse the PhD studies. I would also like to thank my current manger Dr. Omesh Tickoo for providing me the flexibility to balance work and PhD studies. I am grateful to my talented colleagues Dr. Nilesh Ahuja, Dr. Teahyung Lee and Dr. Ibrahima Ndiour for being very supportive and sharing the workload during the stressful times.

I would like to extend my thanks to fellow IVU lab members for their support, and for participation in the subjective experiments which were grueling and time consuming. I would like to thank Milind Gide for being there for the discussions, and also for helping with any on-campus logistical issues. I would like to acknowledge

and thank Aliraza Golestaneh for working with me on the texture synthesis quality tests, and helping me with designing and conducting the experiments.

I would like to thank the Electrical Engineering department and ECEE Graduate Program Chair Dr. Joseph Palais for giving me the opportunity to pursue PhD studies at Arizona State University. I would like to express my thanks to ECEE graduate advisors Ms. Esther Korner and Ms. Sno Kleespies for their prompt responses to my queries and sending timely notifications on the deadlines.

Lastly, I would like to thank my family for their unconditional support in pursuing my dreams. My graduate education would not have possible without the numerous sacrifices from my parents. It is their belief and blessings which propelled me through these years. I would like to express special thanks to my wife Pornima, whose constant encouragement and enormous support is what allowed me to pursue PhD studies while working fulltime. I hope the compromises she had to make over these years is well-worth after culmination of this dissertation. My two wonderful kids whose smiling faces are my source of energy during the long work hours.

TABLE OF CONTENTS

LIST OF FIGURES

# Chapter 1

## Introduction

The capabilities and the limitations of the human visual system (HVS) are used in creating and consuming visual content. The perceived visual content evolves from images where the three dimensional world is projected onto two dimensional image frames with a finite sampling lattice. The HVS has a finite spatial resolving power close to 1 arcminute (arcmin) [2–4], which is approximately equal to 1.75 mm at a distance of 6 meters. This limitation is used to limit the spatial sampling frequency and to sample at a high enough resolution to minimize the spatial aliasing effects. Similarly, the perceived 2D video content results from displaying image frames in quick succession. The limitations observed in the temporal resolving power of the HVS [5] are considered in selecting video frame rates and display refresh rates. Perceieved 3D views can result from 2D images [6], where the images for the left and the right eye are rendered with horizontal parallax (disparity). The brain tries to fuse these images which gives the perception of depth, which is the third spatial dimension for the visual content. In this dissertation, the perceived visual quality and its modeling are studied in the context of 3D and 2D viewing to better understand the response of the HVS to selected physical stimuli.

The visual content is characterized by attributes such as color, texture, motion, blur, sharpness, contrast, depth, resolution and frame-rate. The usage of these attribute depends on the application. In this research work, blur and texture attributes are studied in the context of 3D and 2D visual content, respectively. The first part of the dissertation focuses on 3D stereo images, and a thorough study of blur discrimination for both the symmetric and asymmetric stereo viewing is presented. The second part of the dissertation focuses on texture attributes in the context of 2D images and a detailed study of texture granularity and its applications is presented.

## 1.1 Blur Discrimination Experiments

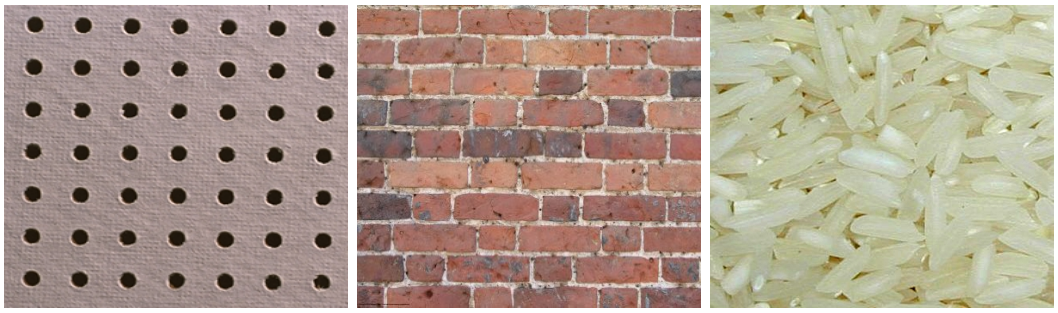
The quality of experience of consuming visual content depends on many factors from creation to distribution and to rendering of the content. Any artifacts or irregularities introduced in the content processing chain will degrade the overall quality of experience. Blur is an important attribute in characterizing the visual content. Blur is associated with loss of information or high frequencies in the visual content. The blur in the visual content is introduced by many factors including but not limited to motion blur, optical aberrations, compression, format conversion and display system limitations. In order to optimize the design of systems for applications such as acquisition, compression, distribution, and display of visual media, it is important to study and model the sensitivity of the human visual system (HVS) to blur. Blur discrimination studies help us to understand the level of blur which is visible to the HVS and helps in designing optimized systems along the video pipe.

In blur discrimination experiments, subjects are presented with two stimuli, one with a reference blur level and the other with a reference plus an additional blur referred to as target blur level. Using a Two-Interval Forced Choice (2IFC) or Two-Alternative Forced Choice (2AFC) subjective test methodology, subjects are asked to judge which stimulus has a larger blur. The threshold is measured as the just noticeable additional blur required in order to discern the reference blur level from the target blur level. In Figure 1.1, rectangular test patterns for a blur discrimination experiment are shown. The subjects are asked to observe the central vertical edge in order to judge which of the two test patterns has a larger blur. Depending on the application, the simulated blur can be modeled using different blur patterns such as Gaussian, rectangular and defocus blur. A Gaussian blur is typically used in these experiments as a representative of different blurring functions.

(a) Reference Pattern

(b) Target Pattern

Figure 1.1: Test stimuli with Gaussian blur for 2D blur discrimination experiments.



(a) Regular Texture (b) Hybrid Texture (c) Irregular Texture

Figure 1.2: Examples of texture regularity.

Blur discrimination in the context of 2D images has been extensively studied in the literature. In this dissertation, blur discrimination studies are extended to the 3D stereo viewing scenario.

## 1.2 Texture Attributes

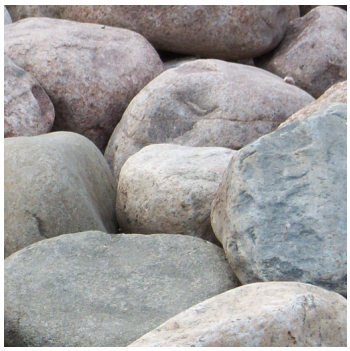
Two-dimensional textures represent spatial variations in the pixel values of the visual content. Textures are characterized by attributes such as regularity, directionality and granularity. Texture attributes are used in various applications such as image

(a) Low granularity (b) Medium granularity (c) High granularity

Figure 1.3: Examples of texture granularity.

retrieval, classification, synthesis and compression. Texture attributes can be characterized based on the observed primitives. The primitives represent smallest repetitive recognizable object in the texture. The texture regularity attribute provides information about the arrangement of the primitives. If the primitives are arranged in a very structured manner in terms of size, location, color, and shape then the texture is said to have high regularity. In Figure 1.2, examples of highly regular to irregular textures are shown. The observed texture primitives in these images are black circles, bricks and rice grains. It can be observed that the placement of these primitives define the regularity of textures. The directionality of the textures provides information about how the dominant structures or edges are aligned in the texture.

Texture granularity is another important attribute to characterize textures. The granularity level can be quantified based on the size of the primitive [7] in the texture. In addition, the spacing and density of the primitives can influence the perceived granularity of a texture. In Figure 1.3, examples of low, medium and high granularity level textures are shown. It can be observed that, for the low granularity level textures, one can easily recognize the objects in the image (e.g.,tennis ball). But, in the case of high granularity level textures, it is more difficult to discern the smaller size primitives. The granularity of the textures depends not only on the local



(a) Stack of boulders

(b) Surface of the boulder

Figure 1.4: Examples of global and local context of primitives for texture granularity.

but also on the global context of the primitive. In Figure 1.4, a stack of boulders (Figure 1.4(a)) is shown along with the surface of the boulder (Figure 1.4(b)). In the case of the stack of boulders (Figure 1.4(a)), the surface details of the boulder are not that important compared to the outline of the different boulders. On the other hand, if we are focusing only on the small patch of the boulders (Figure 1.4(b)), then the details on the surface of the boulder are more important. Hence, a granularity measure should consider the local as well as the global context of the primitives. In this dissertation, a detailed study of perceived texture granularity is presented along with an objective no-reference texture granularity index (TGI). The application of the proposed objective granularity index for texture compression and synthesis is discussed and demonstrated.

## 1.3 Contributions

In this dissertation one of the first psychometric experiments to measure the blur discrimination thresholds using stereoscopic 3D test patterns is presented. Both symmetric and asymmetric stereo viewing cases are considered. The binocular disparity is used as the only depth cue in the image. The 3D blur discrimination tests are carefully designed taking into consideration factors such as stereoscopic 3D perception, binocular rivalry, display cross-talk and vergence adaptation to the disparity.

The subjective test results indicate that the blur discrimination thresholds remain almost constant as we vary the disparity value, which corresponds to varying the object depth. This significant finding indicates that one can apply 2D blur discrimination models for stereoscopic 3D blur discrimination. It is shown that the Weber model provides a good fit to the subjective study measurements involving 3D stereo blur discrimination.

The study of asymmetric 3D stereo blur discrimination results showed that blur discrimination has a masking effect, in the sense that the eye observing the larger blur masks the blur observed by the other eye. This further indicates that asymmetric 3D stereo coding will introduce noticeable reduction in quality even if the compression quality of one of the streams (left or right eye) is kept higher.

Another significant contribution of this dissertation is to propose a no-reference perceptual texture granularity index and demonstrating its application for visual compression. This is the first ever work that quantifies the perceived granularity in textures. A texture database, referred to as GranTEX, is created consisting of texture images with different granularity levels. A subjective study is presented to quantify the perceived granularity level of the textures present in the GranTEX database. The presented database is useful in comparing the performance of objective granularity measures. It is demonstrated that the proposed Texture Granularity Index (TGI) correlates well with the perceived granularity level obtained from the subjective study. It is also shown that the proposed TGI outperforms other objective granularity measures published in the literature, including the one proposed for the MPEG7 standard.

In order to demonstrate the application of the proposed TGI for visual compression, a subjective study is conducted to assess the effect of compression on textures with varying degrees of granularity. It is found that, for a similar visual quality,

one can achieve almost double (∼1.8) the compression for the low granularity textures as compared to the high granularity textures, and around 1.5 times the compression for the medium granularity textures as compared to the high granularity textures. A logarithmic function model is proposed as a fit to the subjective test data. It is demonstrated that the proposed TGI metric and the proposed model can be used for rate-distortion control by allowing the automatic selection of the needed compression ratio for a target visual quality. The proposed model can also be used for visual quality assessment by providing a measure of the visual quality for a target compression ratio.

Finally, the effect of texture granularity on the quality of synthesized textures is presented. For this purpose, subjective studies are conducted to assess the quality of synthesized textures with different levels (low, medium, high) of perceived texture granularity using different types of texture synthesis methods. The results of the conducted subjective studies indicate that the non-parametric patch-based texture synthesis method outperforms other well-known parametric and pixel-based methods. This work also proposes a reduced-reference visual quality index referred to as the delta texture granularity index for assessing the visual quality of synthesized textures.

## 1.4 Dissertation Outline

The organization of the dissertation is as follows. The necessary background related to this work is presented in Chapter 2. This includes details on psychometric experiments, 3D video technologies and image texture models. A detailed study of 3D blur discrimination for stereo images is presented in Chapter 3. The proposed perceptual granularity metric is discussed in Chapter 4. An application of the proposed texture granularity index for visual compression is presented in Chapter 5. The relationship between texture granularity and synthesis quality is presented in Chapter 6. Conclu-

sions summarizing the contributions of this thesis and future directions are presented in Chapter 7.

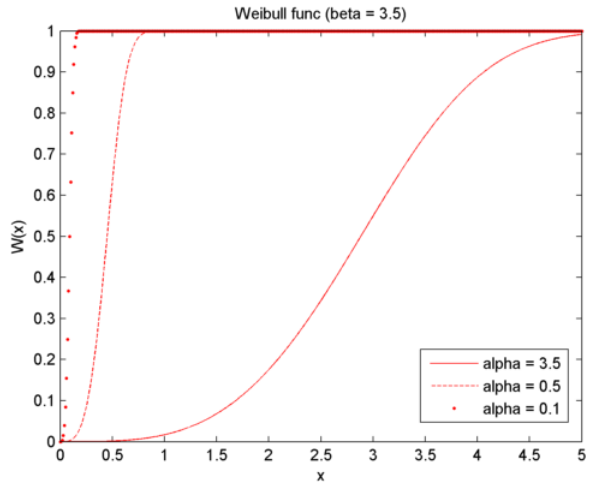
# Chapter 2

# Background

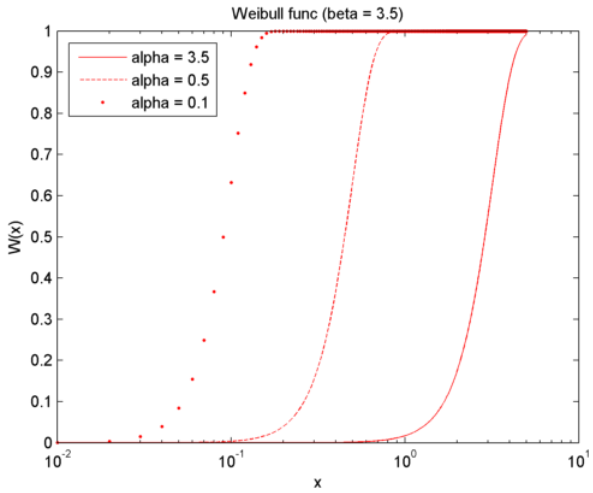
In this chapter, the background material related to this work is presented. A review of the QUEST [8] adaptive psychometric procedure is first presented. The QUEST procedure is used in the blur discrimination experiments. A broad overview of 3D technologies is presented next, which includes details on 3D visual perception, 3D video processing pipe and subjective quality assessment for 3D stereo content. The second part of this chapter focuses on image textures and provides a review on texture models, texture synthesis methods and the visual quality assessment of synthesized textures.

## 2.1 Adaptive Psychometric Experiments

In a psychometric experiment, the relationship between a physical stimulus and the subjective responses [9] is studied. The applied physical stimulus would stimulate the senses such as visual, hearing and touch. In a psycho-visual experiment, the visual stimulus is presented to the subject and the response of the Human Visual System (HVS) is studied. The response of the observer is obtained by varying the strength of the stimulus. In a yes/no experiment, the subjects are asked to judge the presence or absence of a signal. In a forced choice experiment, the subjects are asked to identify the stimulus in n-given alternatives. In the specific case of two-alternative-forced-choice (2AFC) method the subjects are asked to choose between the two alternatives. The selection of strength of the physical stimulus can be varied either in non-adaptive or adaptive steps. In a non-adaptive method, the stimulus is presented at predefined locations in the physical stimulus domain or in a predefined step sizes in an ascending or descending order. In an adaptive method, the stimulus strength in the physical domain is selected based on the previous responses from




(a) Linear x-axis




(b) Logarithmic x-axis

Figure 2.1: Weibull function plotted with linear and logarithmic x-axis.

the subject. The adaptive methods [10] tend to be more efficient in providing good convergence.

A plot of the correct response against the strength of physical stimuli results in a psychometric function, which describes the relationship between the physical measure of stimulus and the probability of a particular physical response. In the general form, a psychometric function [9] is given by the following equation:

$$\psi(x;\alpha,\beta,\gamma,\delta) = \gamma + (1-\gamma-\delta)\,F(x;\alpha,\beta) \tag{2.1}$$

where $F(x;\alpha,\beta)$ represents a two-parameter function which is typically a sigmoid function. The function $F(x;\alpha,\beta)$ is usually chosen to have a range [0, 1] and is represented by functions such as Weibull, logistic, or cumulative Gaussian. The parameter $\gamma$ gives the probability of correct detection in the absence of the stimulus signal. For an n-alternative forced choice experiment this is usually fixed at $\frac{1}{n}$ where $n$ is the number of choices (which is 0.5 for 2AFC method). The parameter $\delta$ represents the observer lapse, which is the probability of incorrect detection irrespective of the strength of the stimulus signal. Hence $(1-\delta)$ gives the upper bound for the curve representing the psychometric function.

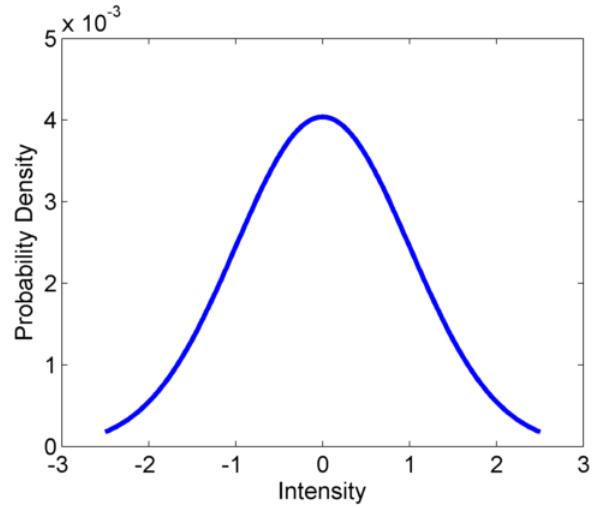
In the particular case of the Weibull distribution, $F(x;\alpha,\beta)$ is represented by

$$W(x;\alpha,\beta) = 1 - \exp\left\{-\left(\frac{x}{\alpha}\right)^{\beta}\right\} \tag{2.2}$$

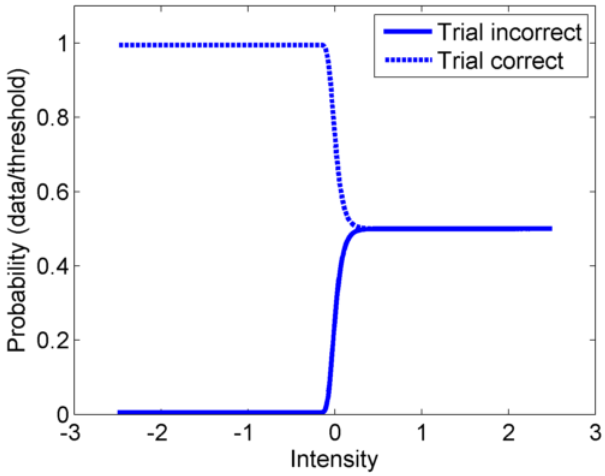
where $\alpha$ is the position and $\beta$ is the slope parameter. Plots of the Weibull distribution for different values of $\alpha$ are shown in Figure 2.1. The $\beta$ value is set at 3.5 in these plots. It is observed from the plots that in the case of a linear x-axis (Figure 2.1 (a)) the slopes of the curves are different, but in the case of a logarithmic x-axis (Figure 2.1 (b)) the plots show similar slope values. This indicates that on a logarithm x-axis the probabilities given by the psychometric functions are translations of one template curve along the axis. This property is used in the QUEST psychometric procedure described next.

### *2.1.1 QUEST Psychometric Procedure*

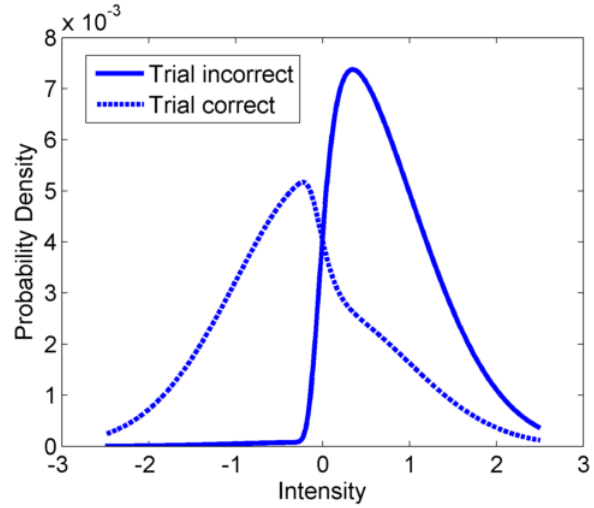
QUEST is an adaptive psychometric procedure [8]. This method calculates the stimulus strength (threshold) based on the most probable Bayesian estimate of the threshold from the available information. The information is derived from the thresholds observed in previous trials of the experiment and from the prior knowledge about the experiment. The prior knowledge can include the shape of the psychometric




(a) Prior probability density function




(b) Likelihood function




(c) Posterior probability density function

Figure 2.2: Application of Bayes rule to the QUEST probability density function.

function and the distribution of thresholds under study. This adaptive procedure helps to quickly converge to the correct threshold values within a few trials.

The QUEST method makes the following assumptions [11]:

- The psychometric function has the same shape under all conditions, when expressed as a function of log intensity.
- The subject's threshold does not vary from trial to trial.
- Individual trials are statistically independent.

The initial range for the threshold values is assumed to be known on a logarithmic threshold scale. A graphical representation of how Bayesian inference is applied for the QUEST method is shown in Figure 2.2. In Figure 2.2(a), a Gaussian initial probability density function (pdf) is shown and the first stimulus intensity is chosen corresponding to the mode of this pdf. The area under this curve sums up to one. Baye's rule is applied to find the maximum likelihood estimate of the threshold for the next trial. Assuming availability of prior knowledge about the experiment that the higher thresholds increase the probably of correct detection, two scenarios are possible: 1) if the subject responds correctly then the threshold is probably to the left of the current intensity (lower values) and the new pdf is skewed towards the lower intensity values; 2) on the other hand, if the subject responds incorrectly then the threshold is probably on the right side of the current intensity (higher values) and hence the pdf is skewed towards the higher intensity values (as shown in Figure 2.2 (c)). The QUEST method is used in the psychometric experiments presented in Chapter 3.

## 2.2 3D Video Technologies

The world observed around us is three dimensional, hence visualizing the 3D content gives us a sense of real presence. 3D visualization techniques mostly fall under stereo viewing techniques where two different views are rendered to each eye with a horizontal parallax between them. In this section an overview of 3D video technologies is presented.

### *2.2.1 3D Visual Perception Basics*

A brief description of 3D perception basics is provided here. The Human Visual System (HVS) perceives depth from monocular and binocular cues [6]. The monocular cues are derived from a single image viewed by one eye or by each eye independently. The monocular cues include head movement parallax, linear perspective, occlusion of more distant objects by near ones, shading or texture illumination gradient and lens accommodation (muscular tension to focus objects). For accurate depth estimation, however, our brain uses binocular vision, which is based on the fact that we sense the world through two 2D images that are projected on the retinas of the left- and the right-eye. Since our left and right eyes are separated by a lateral distance, known as the interpupillary distance (IPD), each eye sees a slightly different projection of the same scene. Binocular vision corresponds to the process in which our brain computes the disparity, which is the difference in distance between the center of fixation and the two corresponding points in the retinal image of the left and right eye.

In the natural 3D stereo (binocular) viewing, the eyes fixate on the object of interest through a process involving vergence and accommodation [12]. Since there is a lateral shift between the two eyes, which creates a disparity between the left and right eye images rendered onto the retina, the two eyes converge (inward movement)

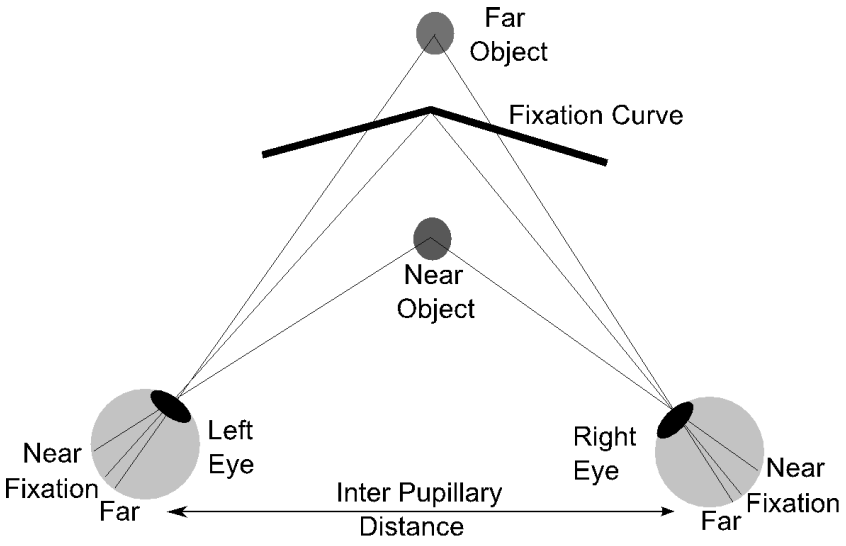



Figure 2.3: Stereoscopic vision.

or diverge (outward movement) in order to create a single retinal image called the cyclopean image. This process is called vergence. At the same time, the eyes change their focal length in order to create a sharper retinal image of the object of interest. This process is called accommodation. It is widely accepted that the vergence [13] is driven by the disparity and the accommodation is driven by the blur, and there is a cross coupling between the two processes. In the case of natural stereo viewing, the accommodation and vergence processes are coupled in order for the eyes to fixate and focus on the object of interest. The disparity between the left- and right-eye retinal images provides accurate depth information for the object of interest. The objects which are not in focus appear blurred in the retinal image and are modeled as defocus blur [14]. Defocus blur is an important depth cue for the brain providing relative depths of the objects in the scene. Hence, blur and disparity are complementary signals and they together provide accurate depth information to the brain.

The projection of images in stereo vision [6] is illustrated in Figure 2.3. The images projected on the fixation center have zero disparity and they correspond to

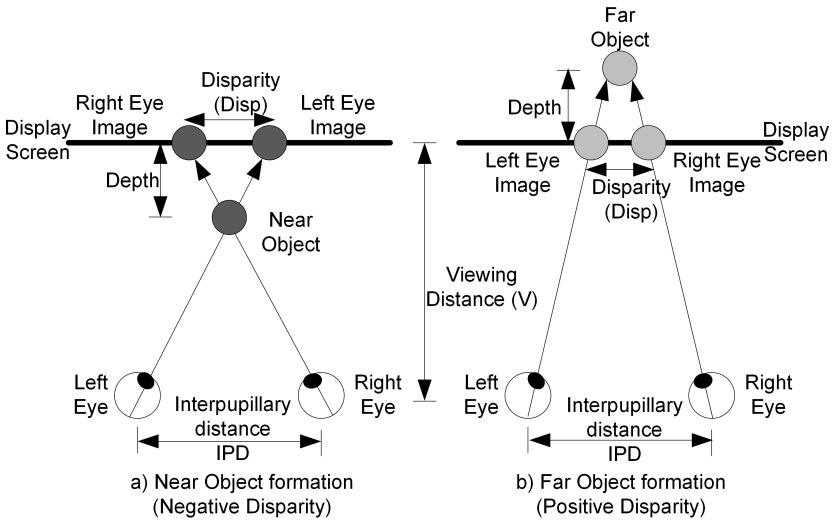



Figure 2.4: Depth perception in stereoscopic displays.

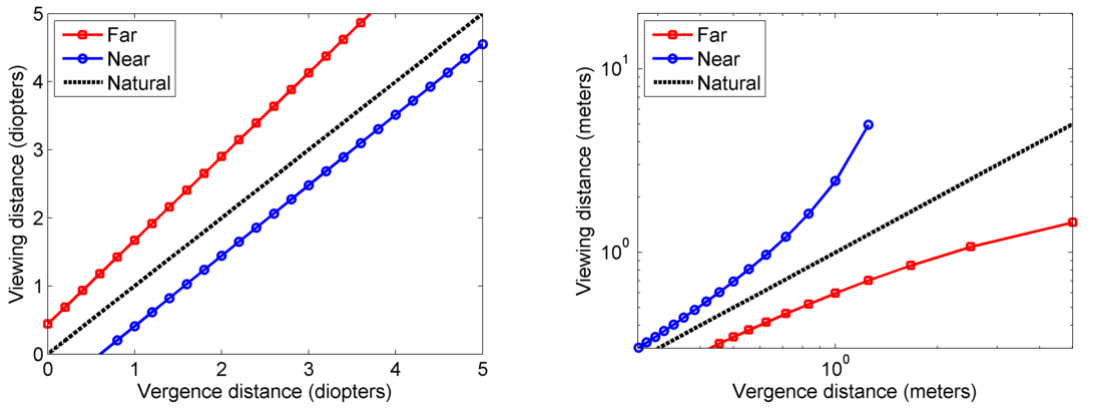
objects on the fixation curve. Objects with positive or negative disparity appear in front or behind the fixation curve, respectively. This binocular vision technique is used in stereoscopic 3D displays, where the left-eye image is horizontally shifted with respect to the right-eye image. The brain tries to fuse these horizontally shifted images to create a sense of depth. As shown in Figure 2.4, depending on the disparity, the object is perceived as near or far object. The rendered distance of the near object from the display screen is given by:

$$Depth_{Near} = \frac{V \cdot DISP}{(IPD + DISP)} \tag{2.3}$$

where $V$ is the viewing distance, $DISP$ is the disparity between left- and right-eye images, and $IPD$ is the interpupillary distance. Similarly, the rendered distance of the far object from the display screen is given by:

$$Depth_{Far} = \frac{V \cdot DISP}{(IPD - DISP)} \tag{2.4}$$

In the case of 3D stereo viewing, the left and right eye are always focusing on the display screen, but because of the lateral shift in the left and right eye images,




(a) Comfort zone measured in diopters

(b) Comfort zone measured in meters (Log-Log scale)

Figure 2.5: Comfort zone region where binocular fusion is possible.

the brain perceives near or far object. If the disparity values are above a certain threshold, the brain will not be able to create the cyclopean (single) image and hence, the binocular vision breaks down. In Figure 2.5, the plots of comfort zone [15] are shown in diopters (left panel) and in meters (right panel). On the horizontal axis the vergence distance is shown and on the vertical axis the viewing distance is shown. The black diagonal line shows the natural viewing condition where both the focal length and vergence distance are the same. The lines with squares and circles represent limits for near and far object regions, respectively, where a converged cyclopean image can be observed. The region between these two lines is referred as comfort zone. An object rendered outside this zone will lead to difficulty in fusing the left- and right-eye images. It can be observed from the right plot (Figure 2.5(b)) that the comfort zone gets wider as the distance from the screen gets larger. Rendering the objects outside the comfort zone causes eye-fatigue and eventual loss of depth sensation [16].

### *2.2.2 3D Video Chain*

The 3D video Quality-of-Experience (QoE) is affected by various factors along the delivery chain. A 3D video processing chain [17] consists of content acquisition/production, compression, transmission, and display (shown in Figure 2.6). On the acquisition and production side shooting conditions need to be managed to better assess the user experience [18]. Proper 3D shooting has to be aware of the final presence factor (ratio of screen width and viewing distance). The presence factor depends on the displays, and it is restricted as compared to the real 3D world. A set of 3D video acquisition factors directly affects the QoE, such as overall depth range, convergence, binocular rivalry. During the capture of multiple views of the scene, depending on the camera set-up, there will be misalignments and differences in luminance and chrominance values. These differences not only result in quality degradation, but also lead to incorrect or loss of depth information [19]. The 3D video acquisition for stereoscopic displays requires at least two views with a horizontal shift between them. A multi-view capture system captures two or more full resolution videos. The challenges in a capture system are temporal synchronization, camera calibration, matching white balance and simultaneous storage. It is difficult to perfectly match the camera systems and achieve zero vertical parallax, which requires a post processing rectification step. The video-plus-depth camera set-up consists of a high-resolution video camera with a time-of-flight depth sensor [20]. The obtained depth data in these systems [21] is very noisy and usually of very low resolution as compared to the acquired video resolution. A post-processing step [22] is needed to generate a high-resolution depth map. It is possible to convert between the video-plus-depth and multi-view video formats. A conversion algorithm between these two formats is required to resolve multiple correspondences and occlusion-uncovering regions [23]. Other possibilities

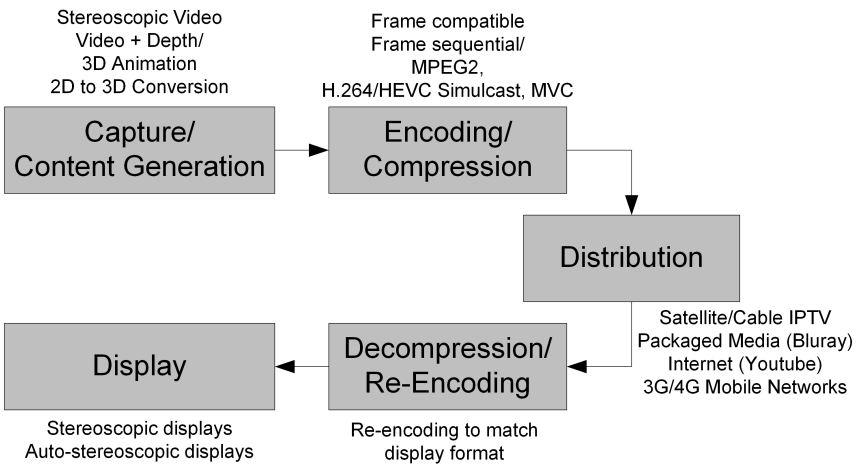



Figure 2.6: 3D video chain showing content generation, distribution and consumption.

for the generation of 3D content include 2D to 3D video conversion and 3D animation. A 2D to 3D video conversion algorithm makes use of information from motion parallax [24] and other available monocular cues in mostly a learning-based framework [25]. The 3D animation content is generated using 3D graphics and computer generated imagery (CGI) principles [26]. The 3D animation techniques have been very mature and are aiming for photo-realistic imagery.

The two or more views of a captured scene need to be stored or transmitted through a broadcast chain. In order to reduce the required doubling of the bandwidth for the left- and right-eye images, compression techniques are applied [27]. This can result in artifacts such as blur, blockiness, graininess, and ringing. In asymmetric video coding, one of the two video sequences (left or right) will be coded at a lower bit-rate or resolution. Transmission errors can also cause degradations in quality. As the impact of compression and transmission artifacts may be different on the left- and right-eye images, spatial and temporal inconsistencies between the left- and right-eye views can occur, which can result in eye-fatigue and loss of depth sensation.

Finally, the 3D video signal needs to be rendered on a display. Depending on the display technology, various artifacts can impair the perceived picture quality. A reduction in spatial or temporal resolution, loss in luminance or color gamut, and the occurrence of crosstalk are typical artifacts related to the display technologies. Visual discomfort is a physiological problem experienced by many subjects when viewing 3D video. There are many reasons leading to the visual complaints, but the most important cause for discomfort is the accommodation-vergence conflict which is discussed in Section 2.2.1.

### *2.2.3 3D Display Formats*

The stereo 3D displays consume the video in two main formats, namely Frame Compatible and Frame Sequential. In the Frame Compatible format, the spatial resolution of left- and right-eye views is reduced in order to keep the overall spatial and temporal resolution the same as the 2D video. This enables the use of existing infrastructures for compression and transmission. Depending on how the left- and right-eye views are spatially mapped, Frame Compatible [28] formats can be classified as side-by-side, top-bottom, row interleaved, column interleaved and checker-board patterns (shown in Figure 2.7). Figure 2.8 illustrates the Frame Sequential format, where alternating left- and right-eye views are displayed at full resolution in a time sequential manner. A display with a row-interleaved format is used in the psychometric experiments presented in Chapter 3.

Stereoscopic displays [29] need two different perspective views simultaneously, as described in Section 2.2.1. Several multiplexing methods can be used to carry the light to the appropriate eye. In color multiplexed displays, two views are shown with a different color each (e.g., red and cyan) and Anaglyph glasses with the corresponding color filters are used to separate the left- and right-eye (L-R) view. This technique

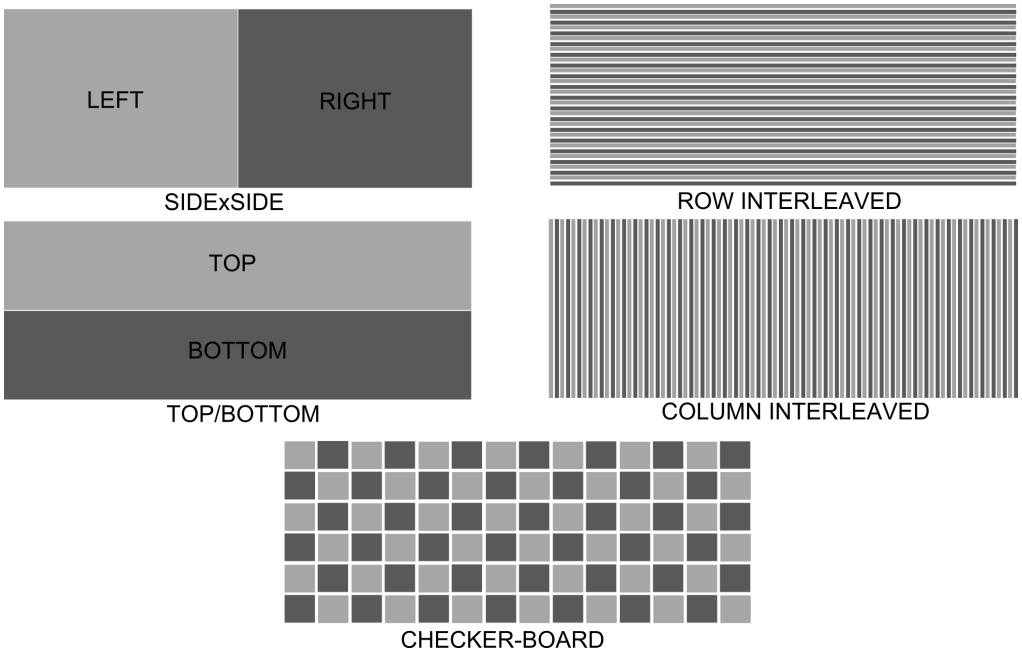



Figure 2.7: 3D frame compatible formats where left- and right-eye images are spatially subsampled.

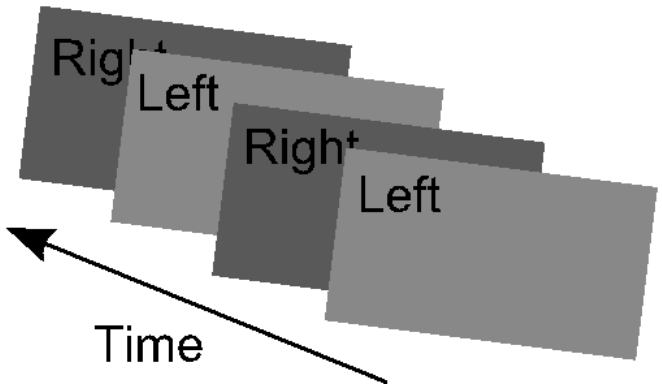



Figure 2.8: 3D frame sequential format where left- and right-eye images are temporally interleaved.

is simple but the observed visual quality is poor. The colors observed between the two views are different, which makes it tiring for the brain to match the two views. In the case of passive retarder stereoscopic displays [30], the L-R views are spatially multiplexed with different polarizations. This reduces the spatial resolution in either the vertical or horizontal direction depending on the interleave pattern. The polarized eye-wear does not need any synchronization with the display and it simply filters the two polarized views for the left and right eye. A passive stereoscopic display is used to render 3D stereo images in the psychometric experiments presented in Chapter 3. Active stereoscopic displays project the L-R views in a frame sequential format. The time-multiplexed L-R frames are filtered by the active glasses, which are synchronized with the display. Active stereoscopic displays require a higher frame rate for flicker free viewing. Auto-stereoscopic displays do not need specialized eye-wear and are classified into direction multiplexed, holographic and volumetric displays [31].

### *2.2.4 3D Subjective Quality Assessment*

3D stereo imposes new requirements and consequently opens new perspectives for quality assessment, as compared to 2D image and video quality assessment [17]. The Quality of Experience (QoE) of 3D videos triggers multidimensional factors from the human point of view. Numerous use cases require quality assessment for different needs all along the content delivery chain (shown in Figure 2.6). The artifacts introduced during each stage have different effects on the overall QoE. Separate methods of quality assessment need to be developed for different application scenarios so that current technologies can be improved.

Subjective quality assessment of 3D video is crucial in developing the 3D-TV infra-structure, that will provide optimum 3D QoE. Subjective evaluation of 3D QoE

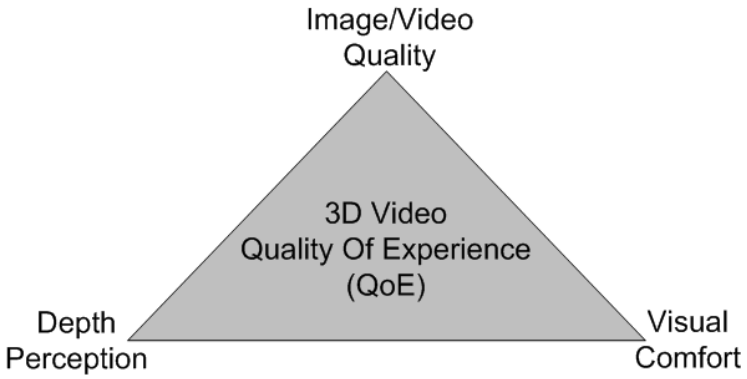



Figure 2.9: 3D stereo quality of experience.

is multifaceted, and should include measurement of depth perception, visual comfort along with 2D video quality (as shown in Figure 2.9). The 3D QoE tries to measure the sense of ultimate reality or true presence, along with physiological issues such as eye fatigue. 2D video quality is an important factor in the assessment of the quality of 3D video. Methodologies for subjective quality assessment for 2D video have been extensively studied, and were standardized as the ITU-T Rec.P.910 [32] and ITU-R Rec.BT.500 [33] standards. 3D video subjective quality assessment followed similar efforts and was standardized as the ITU-R Rec.BT.1438 [34] standard. But, these efforts are not sufficient to capture the overall 3D QoE. Depth perception tries to measure the added value of stereoscopic depth in 3D-TV systems. Subjective experiments were conducted in the literature to predict the depth perception. In [35], the relationship among the quality, depth perception, and naturalness is studied through subjective assessment of 3D video. It was concluded that the depth perception decreases as the quality decreases. When the quality is low, the perceived depth of the 3D video sequences gets closer to that of 2D video sequences. Visual discomfort is important to measure the overall 3D QoE. In [36], experiments are conducted to study visual fatigue in the context of multi-view acquisition. Standard subjective quality assessment protocols are used, but requesting the observer to rate visual fa-

tigue instead of overall quality. In these experiments, 8 to 10 seconds length video clips are considered, which is quite short to stress the visual fatigue.

Beside measurement methods, in order to conduct subjective quality assessments of 3D video, one has to carefully consider the following factors:

1. Content selection: The content selection is very important in order to correctly interpret the data obtained from the conducted 3D video subjective assessment experiments. When shooting or generating the content, the camera baseline width and disparity measurements will affect the comfortable viewing conditions. These viewing conditions need to be matched during the conducted subjective experiments. The viewing conditions here represent display size and the viewing distance, and if needed the content has to be reauthored for different viewing conditions.

2. Display technology: Currently, there are three main 3D display technologies available for 3D-TV, including active, passive and auto-stereoscopic displays. The underlying display technology affects factors such as spatial-temporal resolutions, crosstalk level, comfortable viewing position, that need to be considered in the subjective assessment.

3. 3D video pipe: The subjective quality assessment need to consider the underlying technologies in a 3D-TV video pipe. The selection of different formats (frame compatible, full-frame), compression methods, distribution channels will introduce different kinds of artifacts. These factors have to be defined and reflected in the subjective quality assessment.

4. Viewing conditions: Viewing conditions are strongly coupled with content selection and display technology. Display size and viewing distance need to be

defined based on the created content. Different luminance levels, lighting conditions and display calibration have to be considered based on the display technology.

There is a need for developing standardized subjective assessment methodologies that would encompass the above factors to study the overall 3D QoE. The existing subjective quality assessment methodologies for 3D video are largely the same as those for 2D video, which do not measure the multifaceted nature of 3D video quality.

## 2.3 Image Textures

Image textures are important in conveying spatial arrangement or structure of the scene. A low-level [37] representation of the image textures is described by the spatial variations in pixel intensity values. At a high-level, image textures are described by types of repetitive regions, referred to as primitives, and their spatial arrangements. In this section a review of texture analysis methods is presented. A summary of texture models is presented followed by a review of texture synthesis methods and their quality assessment. Texture models are useful in representing the textures in a compact form. Texture synthesis methods generate or synthesize the textures from the compact representation such as higher order statistics or image patches.

### *2.3.1 Texture Models*

The image textures represent intensity variations in the pixel values. The texture models are used to characterize and analyze texture images. The textures are modeled [37] using various approaches such as statistical, geometric, model-based and signal processing based methods.

*Statistical Methods*

The statistical methods try to characterize the textures based on the spatial distribu-

tion of pixel intensity values. In a co-occurrence matrix representation, the number of occurrences of pair of gray levels which are at a distance $d$ are calculated. The entries in the co-occurrence matrix are used to calculate the second order statistics [37] such as energy, entropy, contrast, homogeneity and correlation. These higher order statistics are useful in inferring the characteristics of the textures. Textures exhibit repetitive placement of patterns, which can be analyzed using autocorrelation features. The autocorrelation function can be used to represent regularity and the level of coarseness in the textures.

*Geometrical Methods*

The graphical methods model the textures as being composed of texture primitives. The geometric properties of the primitives are used to analyze the textures. One of the geometrical methods to represent the texture primitives is to use Voronoi tessellations [37] . The texture primitives are first extracted, for example using image gradients, and then tessellation is constructed to represent Voronoi polygons. Structural methods [37] are part of geometrical methods where primitives are extracted and then spatial placement rules are created. These method are effective for highly regular structures.

*Model-based Methods*

The model-based methods construct an image model to describe the textures. The model parameters capture the essential structure of the texture. Markov Random Field (MRF) based models [37] are popular way to model the textures. MRF models assume that the intensity at a pixel location depends on the intensities of only the neighboring pixels. Hence, MRF models capture the spatial contextual information in the images. Fractals are another popular tool to represent textures [37]. Fractals capture roughness and self similarity of the textures at different scales.

*Signal Processing Methods*

Signal processing methods compute the texture features from filtered images. Spatial domain filtering is a direct method to calculate the texture properties [37]. This analysis can also be carried out in frequency domain [37], as there is evidence that the human visual system analyzes frequency and orientation components. Fourier domain filtering is an elegant tool to analyze the global properties of the textures in the frequency domain. Gabor and wavelet filters are useful when there is a need for analyzing both local and global properties.

### *2.3.2 Texture Synthesis Methods*

Texture synthesis methods regenerate or synthesize the textures from small patches or compact representations. Texture synthesis has many applications, including compression [38], denoising [39] and hole filling [40]. Texture synthesis methods can be broadly classified into parametric and non-parametric approaches. These methods can be further classified into statistical and non-statistical based approaches.

In parametric based approaches, the texture is characterized and described through a set of perceptually motivated parameters. These global parameters are computed for the entire texture and are used for evaluating the similarity of the synthesized texture to the input. In non-parametric approaches, a region of the original texture called the "seed" or the "exemplar" is used. Non-parametric approaches can be further classified into pixel-based approaches and patch-based approaches. The pixel-based synthesis approaches [41–43] are slow since they synthesize only one pixel at a time. Also, they might lead to a loss of local structure and to a grainier texture compared to the original. The quality and speed of pixel-based approaches can be improved by synthesizing patches rather than pixels at a time. This is because contiguous pixels belonging to a particular patch in the input texture are

more likely to be contiguous in the output texture. However, in patch-based algorithms [44–46], when patches overlap with each other, handling the conflicting pixels in the overlapped regions becomes necessary. Blending the overlapped regions [46] can cause blurry artifacts. Instead of blending [44,45] one can find an optimal boundary between adjacent patches in the overlapped regions using methods such as graph cut [45].

In statistical based approaches, the texture is characterized and described through statistical features and statistical-domain transforms. Portilla and Simoncelli [47] used pairwise statistics to improve synthesis results for structured textures at the cost of a more complicated optimization procedure. Motivated by psychophysical and computational models of human texture discrimination [48, 49], Heeger and Bergen [50] proposed a method to analyze texture in terms of histograms of filter responses at multiple scales and orientations. Matching these histograms iteratively was sufficient to produce very good synthesis results for stochastic textures. However, since the histograms measure marginal, not joint, statistics they do not capture important relationships across scales and orientations; thus the algorithm fails for more structured textures.

### *2.3.3 Texture Synthesis Quality Assessment*

Subjective quality studies are important to quantity the fidelity of the synthesized textures. These studies which encompass textures with different attributes are needed to develop better objective quality metrics for texture synthesis quality. There are very few studies in the literature which attempt to measure the quality of the synthesized textures. In [51], a database is created which included 340 distorted images (including JPEG2000 compression, white Gaussian noise, Gaussian blur, sub-pixel shifts distortion types) extracted from 10 reference images, which were evaluated

by a group of 20 human subjects. They evaluated several existing mainstream full reference quality metrics with the help of a recently developed quality assessment framework (IVQUEST). As shown in their work [51], traditional objective metrics do not perform well on texture images and there is a need for specialized texture quality metrics. The subjective study presented by Swamy et al. [1] considered synthetic textures to measure the subjective quality of the synthesized textures. A parametric model is also presented to objectively measure the perceived quality of the synthesized textures.

In a practical setup it is important to conduct the subjective quality assessment using natural textures observed in the objects found around us. In [52], a subjective study using natural textures with different texture regularity attributes to assess the perceived quality of the synthesized textures is presented. In order to predict the quality of the synthesized textures, an objective metric based on texture regularity is proposed. It is also shown that popular objective quality metrics do not correlate well with the subjective scores. The main limitation of these methods is that they rely on pixel-based matches which is not appropriate for the synthesized textures. In this dissertation, the texture granularity attribute is used to investigate how the quality of synthesized textures is affected.

# Chapter 3

## 3D Blur Discrimination

In order to optimize the design of systems for applications such as acquisition, compression, distribution, and display of visual media, it is important to study and model the sensitivity of the human visual system to blur. In blur discrimination experiments, subjects are presented with two stimuli, one with a reference blur level and the other with a reference plus an additional blur referred to as target blur level. Using a Two-Interval Forced Choice (2IFC) or Two-Alternative Forced Choice (2AFC) subjective test methodology, subjects are asked to judge which stimulus has larger blur. The threshold is measured as the just noticeable additional blur required in order to discern the reference blur level from the target blur level.

The motivation for the 3D blur discrimination study is two fold – firstly, as a psycho-visual study we would like to understand how the artificial disparity introduced in the 3D displays affects the perceived blur; secondly, as an application of 3D video we would like to understand the relationship between disparity and blur in 3D displays to optimize the design of 3D video systems. Under real world viewing conditions, accommodation and vergence are coupled [53] in the sense that blur and disparity are complementary signals [13] which helps the eyes converge the focal and vergence lengths to almost the same value. In the case of 3D displays the eyes will always focus on the screen and an artificial disparity is introduced between the left and the right eye views, in which case the eyes try to compensate by changing the vergence length. Hence, when a disparity is introduced, the focal and vergence lengths will be different in the case of 3D displays. We would like to measure the blur discrimination thresholds that result when non-zero (positive and negative) disparity values are introduced in a 3D display viewing setup and compare the obtained results to the zero disparity case. Similar to existing 2D blur

discrimination studies, we adopt a Gaussian blur model. Gaussian blur was used in most of the previous studies on blur discrimination to study blur in vision such as detection and localization, and we are extending this to the 3D stereo viewing scenario. In designing applications involving a 3D video chain, one needs to know which attributes affect the visual quality. Disparity being one of the attributes of 3D video, we need to study how it affects the blur discrimination. One example would be how disparity affects sharpness enhancement algorithms, since different objects in the scene will be having different disparity or depth values. For example, in order to design an enhancement algorithm that results in perceptually similar levels of sharpness for all the objects in the scene, one needs to determine how blur is perceived at different disparity values. It is also interesting to see how the two eyes interact with each other when they are exposed to different levels of blur values. This scenario can arise when the two streams of the stereo video are coded at different quality levels. In this study we would like to answer some of these questions by studying the relationship between disparity and blur discrimination thresholds using symmetric and asymmetric stereo viewing cases.

One of the earlier studies [54] on blur discrimination reported that one can easily differentiate the blur at higher contrast ratios than at lower contrast ratios. Based on the hypothesis proposed in [55], the cue to measure blur in the edge is proportional to the distance between the maximum and minimum in the second derivative of the retinal stimulus. This hypothesis was proposed based on the blur discrimination thresholds measured using static test patterns for three types of blurring functions, namely Gaussian, rectangular and half-wave cosinusoidal. The study of the effect of motion blur on blur discrimination thresholds [56] indicated that motion produces equivalent spatial blur, and a motion deblurring model was proposed using temporal integration. The study of blur discrimination thresholds using

different blur levels and contrast values [57] indicated that the thresholds remain constant at higher contrast ratios and increase for very low contrast ratios. In [58] a review of past blur discrimination studies and a discussion of existing models for blur discrimination is presented. In the current study we throughly investigate blur discrimination in the presence of disparity cues for both 3D symmetric and asymmetric viewing cases.

In the literature, different models have been proposed to characterize the 2D blur discrimination thresholds. In [57] a divisible inhibition model is proposed, in which the blur discrimination threshold is characterized using a difference of Gaussian (DoG) operator between the two blur levels and a constant factor. In [58] the visible contrast energy (ViCE) model is proposed, in which the blur discrimination threshold is measured by calculating the contrast energy of edge differences after filtering by the contrast sensitivity function (CSF). In [59] the Weber's model is proposed to characterize the 2D blur discrimination threshold. In this model, the blur discrimination threshold is characterized by the total perceived blur and the Weber fraction.

A subjective study presented in [13] indicated that defocus blur and disparity are complementary signals to depth perception. The disparity is more precise at the fixation, and defocus blur is more precise as we move away from fixation. In our study, since the subjects are always focused on the display screen, defocus blur is not considered. Blur adaptation experiments were conducted by [60, 61]. In these experiments, the subjects were presented with an adaptation stimulus that is either sharper or more blurred than the test stimulus. The subjects were then asked to judge whether the test stimulus is sharper or more blurred. In these experiments [61], for the first trial initial adaptation was kept for a 1-minute duration and, for subsequent trials, the adaptations were kept for 3 seconds. It was found that the eye adapts to

the adaptation stimulus and that the thresholds get lower or higher depending on whether the adaptation stimulus is sharper or more blurred than the test stimulus. In our experiments, we tried to avoid any adaptation to the reference blur stimulus by keeping the stimulus times very small and by randomizing the reference and target blur level ordering for every trial.

Asymmetric stereo encoding is used to reduce the bandwidth. In asymmetric stereo image compression, two views are coded at different resolutions. The subjective studies conducted by [62] concluded that, if one of the views is coded at half the resolution as compared to the other view, the stereo compression quality is acceptable. Additional reduction in the resolution for one of the eyes, introduces a visible drop in the visual quality. In our study, in addition to the symmetric viewing case, we consider asymmetric stereo viewing by introducing different blur levels to each eye for the target object.

The objective of this chapter is to present the results of blur discrimination threshold measurements using stereoscopic 3D test patterns, and to compare these with the 2D blur discrimination threshold measurements. Asymmetric stereoscopic 3D is also prevalent; hence, in this study, a comparison of symmetric vs. asymmetric stereoscopic 3D blur discrimination thresholds is also presented. Finally, it is shown that the Weber's model can be used to fit the subjective study data obtained for the symmetric viewing case.

## 3.1 Setup

The current study investigates blur discrimination in the presence of disparity cues. For this purpose, stereo 3D test pattern objects are rendered at different depths to measure the blur discrimination thresholds. Blur in images provides a monocular cue for depth perception. But in this study, we consider disparity as the only depth

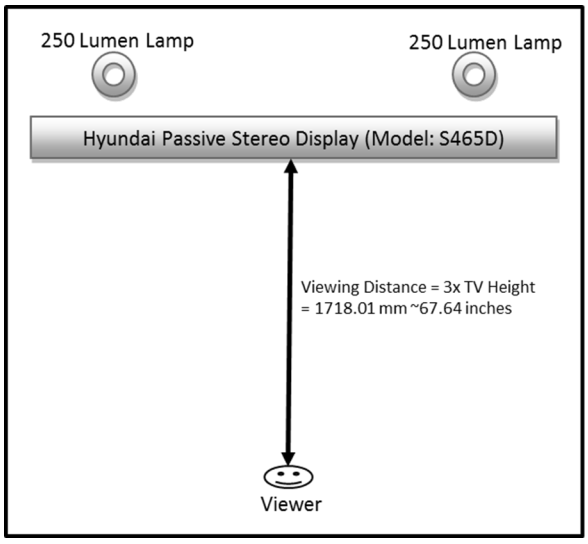



Figure 3.1: Setup of the experiment.

cue using synthetic stereoscopic test patterns. The experimental setup is shown in Figure 3.1. The stimulus is rendered onto a passive stereoscopic 3D display (Model: Hyundai S465D). The circular polarizing filter present in-front of the display screen renders the left and right eye images with opposite polarity. The viewers are required to wear passive polarized filter glasses, which will filter out appropriate images to the left and right eyes. Consequently, in our passive stereoscopic viewing setup, each eye will be exposed to view a different image. The image viewed by the left eye is a shifted version of the image viewed by the right eye, resulting in a non-zero disparity if the relative shift is non-zero. Positive and negative disparities are introduced by introducing positive and negative shifts, respectively, between the left-eye and right-eye images. It is important for the used polarized filter glasses to avoid crosstalk between the left- and right-eye images [63]. The higher amounts of crosstalk (greater than 2$\sim$3%) can cause ghosting at the edges and result in inaccurate blur discrimination threshold measurements.

The display screen has a resolution of 1920(w)x1080(h) pixels in the 2D mode and 1920(w)x540(h) pixels in the 3D mode. The display has a refresh rate of 60

frames per second, and is driven by a desktop computer. The viewing distance is fixed at 3 times the display height, which is equal to 1718 mm. The display is calibrated using PLUGE and Color-bar test patterns [64]. In the subjective testing room, two 250-lumen lamps are used. The lamps are placed behind the TV screen so that no reflection is observed on the display screen. The light meter measurement at the viewer's location is close to 50 lux. The luminance of the display is measured using the SectraScan PR650 [65] colorimeter instrument. The display gamma is measured to be 1.75. The display luminance measurements are taken without and with passive stereoscopic glasses, placed in-front of the instrument. The measurements without glasses are found to have a constant scale factor of 2.6 times measurements with glasses. This indicates that the Michelson contrast ratio remains the same with and without the glasses [66].

## 3.2 Stimulus

The stimulus used for the experiments is shown in Figure 3.2. A vertical edge is formed by placing two boxes together, forming a rectangular object. The background intensity is fixed at the average luminance value of 51 cd/m$^2$. The dimensions of the rectangular object rendered to each eye are 135.77(w)x33.95(h) arcmin. The left and right eye images are displayed in a row interleaved format on the Hyundai stereo display. The effective height of the perceived rectangular object after fusing the left and right eye images corresponds to 67.90(h) arcmin. A parallel camera model [6] is used to render the stereo images. In the parallel camera model the two lenses are parallel to the screen and the optical axes of the two cameras overlap at infinity. The binocular depth is introduced by shifting the left- and right-eye images in the opposite directions; the shift value is equal to the desired disparity value. A negative disparity corresponds to projecting the object in-front of the display screen, a positive disparity corresponds to projecting the object behind the display screen, and a disparity value

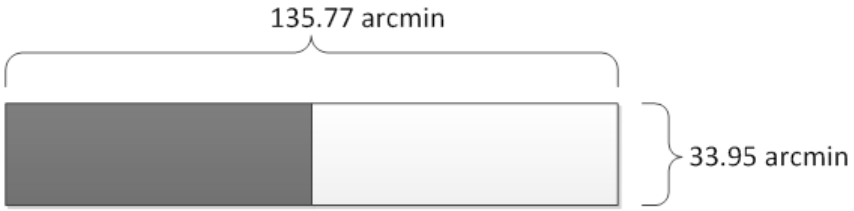



Figure 3.2: Stimulus presented to each eye.

of zero corresponds to projecting the object at the display plane. The disparity is forced not to be zero by the stereoscopic display viewing setup when the introduced shift between the left- and right-eye images is non-zero. Since positive and negative disparities are introduced by shits in the left-eye and right-eye images, it is clear that the subjects are not seeing just zero disparity. Furthermore, it was verified during the experiments that the objects are in fact projected at different depths and that the subjects are able to tell that the object appears to be in the front or in the back of the display screen.

In order to generate the blurred edge, a Gaussian low-pass filter is applied across the central edge along the horizontal direction. The strength of the filter is specified by the standard deviation of the Gaussian filter. A 2IFC method is used to present the stimulus. The stimulus is rendered onto the 3D display using the Psychometric Toolbox (PTB) [67]. PTB is a MATLAB toolbox with support for rendering stereo images onto the 3D displays. The left- and right-eye images for the reference object are filtered using a Gaussian low-pass filter with a standard deviation corresponding to the reference blur level. For every trial, the target object is generated by further filtering the reference object with a Gaussian filter of standard deviation equal to the additional blur level. The additional blur applied to the target object is calculated using the QUEST method [8]. QUEST is an adaptive psychometric method where the threshold for the current trial is calculated by the

most probable Bayesian estimate. The details of the QUEST method are presented in Section 2.1.1. The posterior probability density function is updated based on the previous responses. In this study, a psychometric function based on the Weibull distribution is used, as given below:

$$W(x) = 1 - \delta - (1 - \gamma - \delta) \exp\left[-10^{(x - x_{Th})\beta}\right] \tag{3.1}$$

where $x$ represents the log additional blur value under test and $x_{Th}$ represents the log just noticeable additional blur threshold which needs to be measured. The parameter $\beta$ represents the slope of the psychometric function, $\gamma$ represents the probability of correct detection at the zero threshold value and $\delta$ represents the false negative rate. The PTB toolbox has support for QUEST threshold estimation. The recommended QUEST parameter values for the 2AFC method are used in our subjective tests: $\beta$=3.5, $\gamma$=0.5, $\delta$=0.01. The probability of success when $x = x_{Th}$ is obtained by substituting the above parameters in Equation (3.1), and is equal to 0.8097. For every experiment at least 40 trials are performed. Additional trials are added if the standard deviation of the QUEST probability density function for the threshold is greater than 0.1.

The presentation order of the stimuli is shown in Figure 3.3. For every trial, the fixation frame is shown for an initial time of 1.5 seconds. It is followed by the first stimulus, the anchor frame and the second stimulus, each separated by a 0.5-second interval. Finally, a blank frame is shown, and the subjects are asked to choose the more blurred stimulus by pressing the left arrow (first stimulus) or right arrow (second stimulus) key on the keyboard. The fixation frame has L-shaped anchor marks and a central dot where the subjects need to focus before viewing the stimulus. The anchor marks and fixation dot are projected at the same depth as the stimulus. This provides the depth cue for the subjects before viewing the reference and target objects. For each trial, the location of the fixation center is

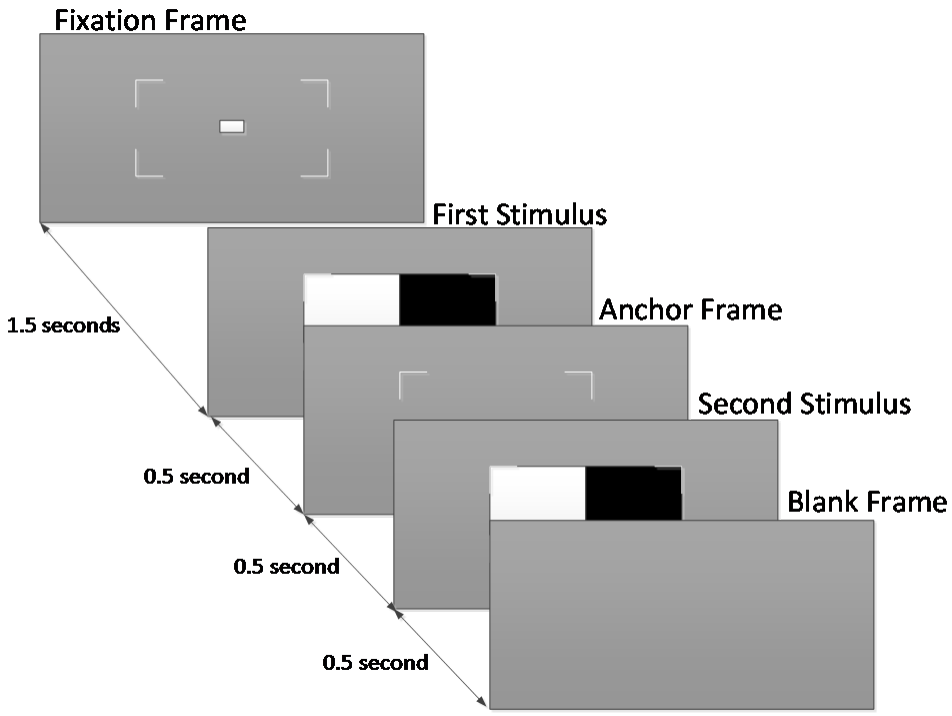


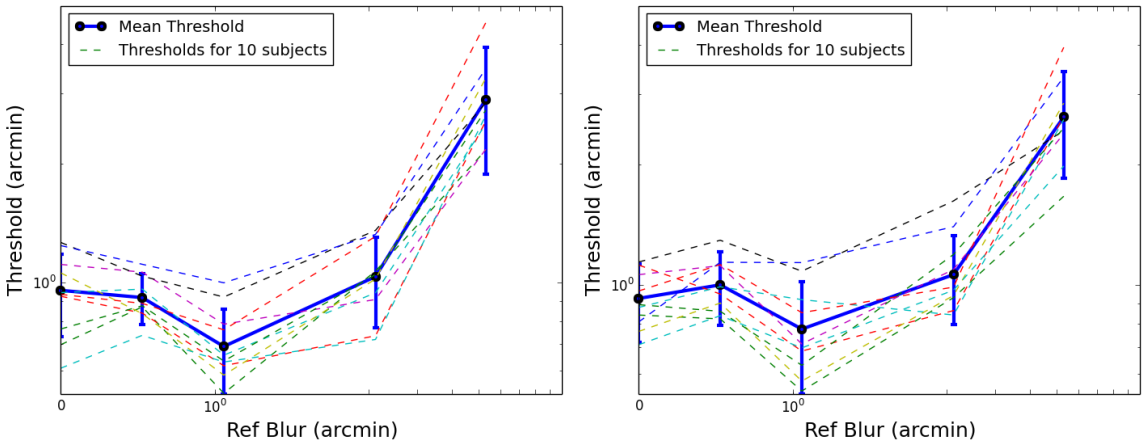

Figure 3.3: Display order for the stimuli.

slightly varied (within 16 arcmin) in the horizontal direction. In our pilot studies we found this to be helpful in reducing eye strain, by not concentrating at the same location. The shift in the central location needs to be small in order to avoid large changes in the viewing angles between trials. The anchor frame has only the L-shaped anchor marks, but no fixation dot is present at the center. The anchor frame provides a depth cue between the two stimuli. This helps in avoiding the need for accommodation time for the second stimulus. For every trial the display order for the reference and target objects are randomized. The contrast direction of the rectangular boxes (white—black or black—white) is also randomized from trial to trial. Before starting the experiments, the subjects are given a trial session to get familiar with the setup.

## 3.3 Symmetric Stereo Blur Discrimination

In the symmetric stereo viewing, both eyes will observe the same level of blur in the target object. In this subjective study, five reference blur levels and three disparity values are considered. This corresponds to a total of 15 combinations of reference blur and disparity values resulting in 15 experiments. The five reference blur levels considered are [0.0, 0.53, 1.06, 2.12, 5.30] arcmin. The reference blur level corresponds to the standard deviation of the Gaussian low-pass filter in arcmin unit. The three disparity values considered are [0.0, -21.22, 21.22] arcmin. The disparity value of 0.0 arcmin represents the 2D viewing case, where there is no shift between the left- and right-eye images. The disparity value of -21.22 arcmin represents projecting the object in-front of the display screen, and the disparity value of +21.22 arcmin represents projecting the object behind the display screen. Applying the parallel camera model and considering an average interpupillary distance of 64mm and a viewing distance of 1718 mm, the projected distance of the object in-front of the display screen is 244.11 mm (using Equation (2.3)) and the projected distance behind the display screen is 341.02 mm (using Equation (2.4)). The Michelson contrast ratio across the edge is fixed at 0.83, which is calculated as the ratio of the luminance difference over the sum of luminance values across the edge.

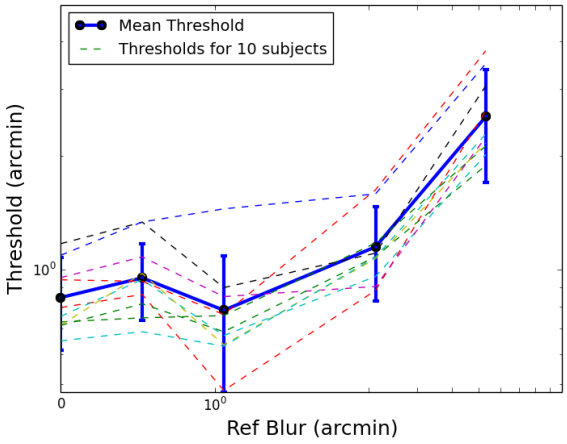
Ten subjects volunteered to participate in this subjective study. The age group of the subjects ranged from 22 to 34 years. Subjects were scanned for normal visual acuity, color vision and stereo vision conditions. As indicated above, a total of 15 experiments were conducted corresponding to five reference blur levels and three disparity values. Each experiment consisted of at least 40 trials, and the target blur level for each trial was calculated using the QUEST method. The threshold that is obtained from the experiments represents the just noticeable additional blur



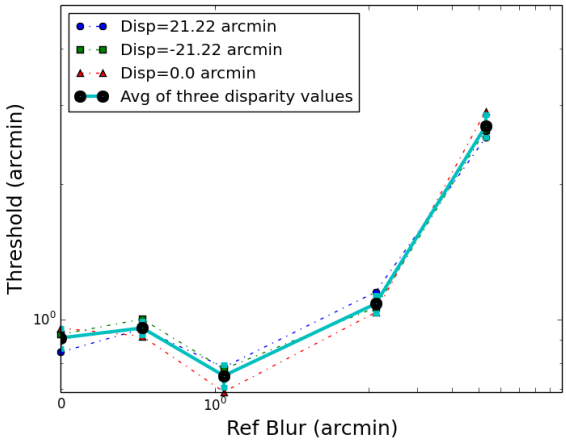

(a) Disparity = 0.0 arcmin




(b) Disparity = -21.22 arcmin




(c) Disparity = 21.22 arcmin




(d) Comparison of threshold values for different disparities

Figure 3.4: Threshold Vs. Reference blur (Log-Log scale, except near origin) plots for symmetric stereo blur discrimination.

required to differentiate between reference and target blur levels. All the subjects repeated each experiment four times in separate sessions; with the ten subjects every experiment is repeated 40 times. For each experiment, the average value of 40 measurements is considered after removing the outlier threshold values which are $2\sigma$ distance away from the mean threshold value, where $\sigma$ is the standard deviation of the subjective scores for the considered experiment.

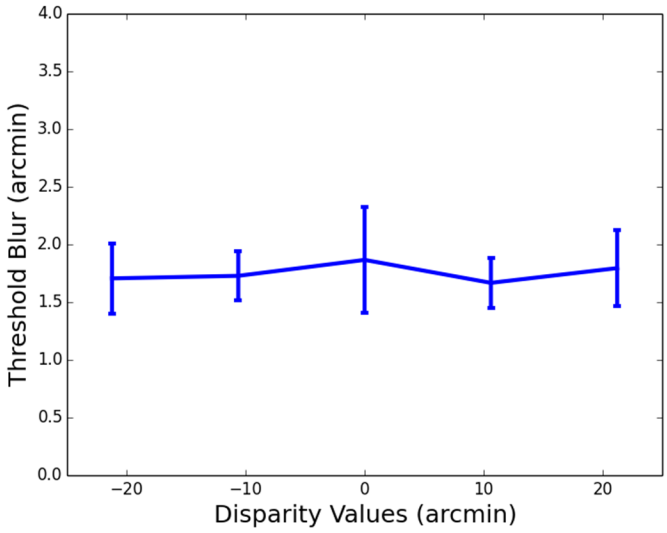



Figure 3.5: Blur discrimination thresholds for reference blur = 5.3 arcmin.

In Figure 3.4, the plots for blur discrimination threshold versus reference blur are shown. In Figure 3.4 (a)-(c), the plots of ten subjects for three disparity values are shown. The average value of the blur discrimination threshold for ten subjects is shown in solid lines. The vertical bars in the plots show the standard deviation of the threshold values. In Figure 3.4 (d), a comparison of the blur discrimination thresholds for three disparity values versus reference blur is shown.

It can be observed from the plots that the average threshold values for three disparity values have a similar trend, and that the threshold values remain constant or vary by a small amount across three disparity values. The vertical bars in Figure 3.4 (d) show the standard deviation of the threshold values for the three disparity values. The obtained standard deviation values between three disparity values are small and are equal to [0.04, 0.03, 0.04, 0.04, 0.15] for the considered five reference blur values. A dip in the threshold value is seen around the reference blur level of 1.0 arcmin. This is attributed to the intrinsic blur level or resolving power of the human visual system, which is close to 1 arcmin [2, 3]. This is consistent with the 2D blur

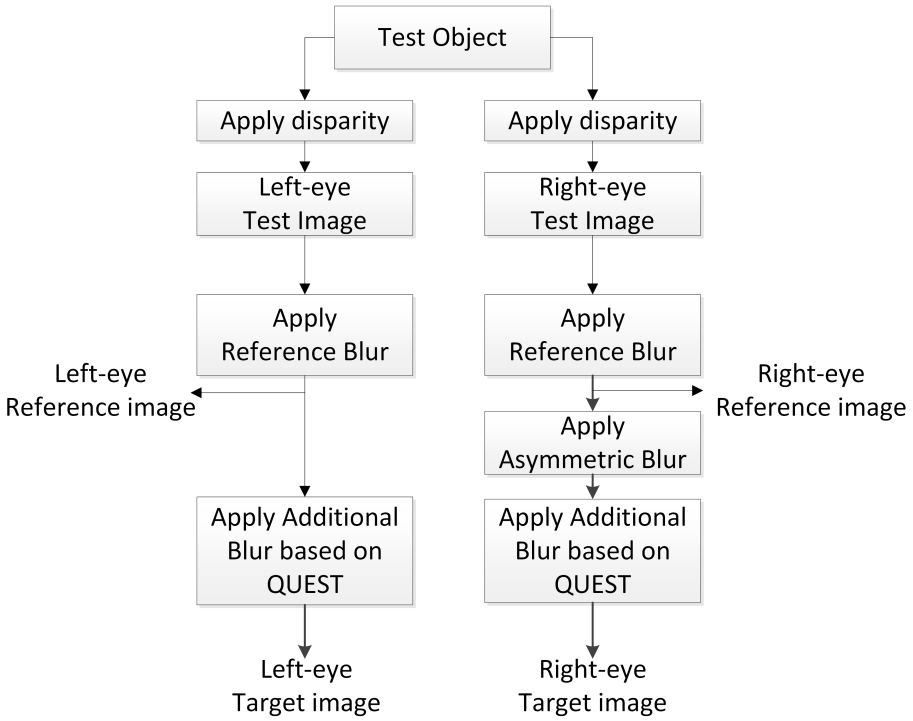



Figure 3.6: Steps to create target blur stimulus for the asymmetric blur stereo experiment assuming the right eye is the dominant eye.

discriminations studies. These plots signify that the blur discrimination thresholds remain constant as we vary the binocular disparity value.

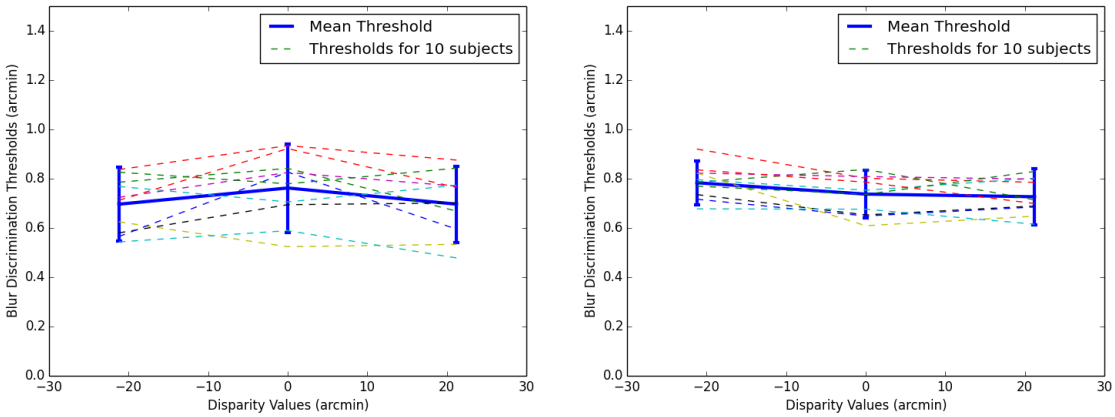
In order to confirm that the blur discrimination thresholds remain constant with change in disparity values, another test was conducted with additional disparity values. In this study we have used only one reference blur value (5.3 arcmin), but added additional disparity values [-21.22, -10.61, 0, 10.61, 21.22] arcmin, resulting in five experiments. Each experiment was repeated four times in separate sessions and the average threshold value of the four sessions is considered. Four subjects participated in this experiment. The plot of the blur discrimination threshold vs disparity value is shown in Figure 3.5. It can be observed that the blur discrimination thresholds are almost constant for different disparity values with a standard

deviation of 0.070. This further confirms that the blur discrimination threshold for the symmetric viewing case is not affected by the disparity value.

## 3.4 Asymmetric Blur Discrimination

In the asymmetric stereo blur discrimination experiments, the left and right eye observe different levels of blur in the target object. The steps to generate the reference and target images are shown in Figure 3.6. Based on the disparity value, the rectangular test object is shifted horizontally to obtain left- and right-eye test images. The Gaussian low-pass filter with a desired reference blur level is applied to the central vertical edge in each of the left- and right-eye test images. The objective of applying the same level of reference blur in the two images is twofold: 1) we wanted to model the reference blur as the natural blur observed in the scene, which is the same level of blur for both eyes; 2) we wanted to compare the asymmetric case with the symmetric case, hence we kept the same reference images for both the cases. An asymmetric Gaussian blur of 0.53 arcmin was used in these tests. A similar value was also used by [62] to generate images at half resolutions for asymmetric compression applications; however, no blur discrimination experiments were performed in that work. Finally, the target image is obtained by applying additional blur to both the left- and right-eye images. The additional blur value is estimated using the QUEST method. A dominant-eye test [68, 69] is carried out to determine the dominant eye of the subject. The asymmetric blur is applied to the dominant eye in order to be consistent across different subjects.

In the asymmetric blur discrimination experiments we have used reference blur levels of less than 1 arcmin. For higher reference blur levels in order to see a noticeable asymmetric blur difference, we have to apply a stronger Gaussian blur. From our pilot studies we have observed that applying an asymmetric blur difference




(a) Reference blur level = 0.0 arcmin (b) Reference blur level = 0.53 arcmin

Figure 3.7: Threshold Vs. Disparity blur plots for the asymmetric stereo blur discrimination.

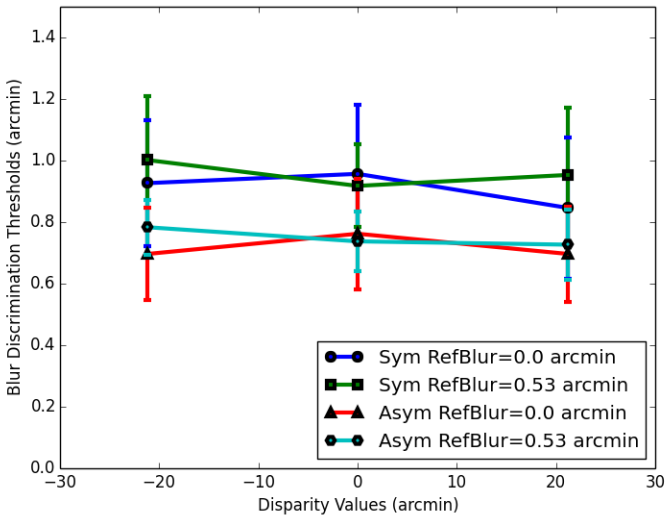



Figure 3.8: Comparison of symmetric and asymmetric stereo discrimination thresholds.

greater than 1.0 arcmin results in binocular rivalry and difficulty in fusing the left- and right-eye images. Hence, we have not considered higher reference blur values in this experiment.

In this subjective study six experiments were conducted which included two reference blur levels and three disparity values. The two reference blur levels considered are [0.0, 0.53] arcmin and the three disparity values considered are [0.0, -21.22, 21.22] arcmin. The Michelson contrast ratio across the edge is fixed at 0.83, similar to the symmetric blur discrimination experiment. Each experiment had at least 40 trials, additional trials were added if the standard deviation of the QUEST probability density function does not converge and is greater than 0.1.

Ten subjects volunteered to participate in this subjective study. The age group of the subjects ranged from 24 to 34 years. The subjects were scanned for normal visual acuity, color vision and stereo vision conditions. In Figure 3.7, the plots of the threshold value versus disparity for ten subjects is shown for different reference blur levels. Each subject repeated the 6 experiments four times, in separate sessions. The average threshold values among the four sessions is considered. It can be observed from the plots that the threshold blur values almost remain constant for the three disparity values. The standard deviation between three disparity values is equal to 0.02 and 0.04 for the considered reference blur levels 0.0 and 0.53 arcmin, respectively.

In Figure 3.8, a comparison of the obtained asymmetric and symmetric blur discrimination thresholds is shown for the same reference blur values. It can be observed that the asymmetric blur thresholds are smaller than the symmetric blur thresholds.The difference between the symmetric and asymmetric blur discrimination thresholds is almost constant for different disparity values. In the symmetric and

asymmetric viewing cases, the reference blur levels presented to each eye remain the same, but the target blur levels will be different in the two cases.

In the sysmmetric case, the target blur level is obtained by convolving two Gaussian kernels with standard deviations equal to reference and threshold blur values and is given by:

$$\sigma_{tar}^{sym} = \sqrt{(\sigma_{ref})^2 + (\sigma_{th}^{sym})^2} \tag{3.2}$$

where $\sigma_{th}^{sym}$ is the blur discrimination threshold for the symmetric case. In the asymmetric stereo viewing case, one of the eyes will observe a larger target blur compared to the other eye because of the additional 0.53 arcmin Gaussian blur applied to it. Hence, the target blur levels for the asymmetric stereo viewing case are given by:

$$\sigma_{tar}^{asym+} = \sqrt{(\sigma_{ref})^2 + (\sigma_{th}^{asym})^2 + 0.53^2}, \text{ for the eye with addtional blur} \tag{3.3}$$

$$\sigma_{tar}^{asym-} = \sqrt{(\sigma_{ref})^2 + (\sigma_{th}^{asym})^2}, \text{ for the eye with no addtional blur} \tag{3.4}$$

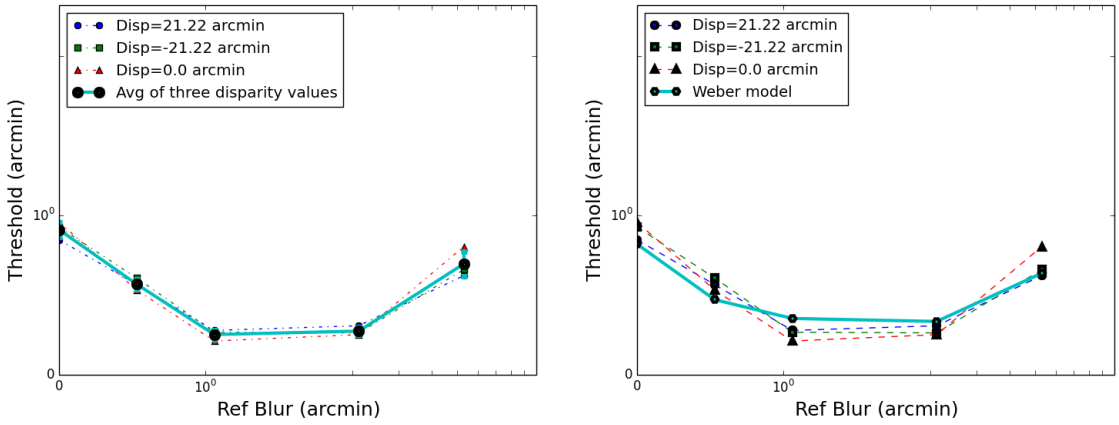
where $\sigma_{tar}^{asym}$ is the blur discrimination threshold for the asymmetric case, $\sigma_{ref}$ is the reference blur, and 0.53 arcmin is the additional blur applied to one the eyes in our experiment.

The obtained blur discrimination threshold values for both the symmetric and asymmetric stereo viewing cases are presented in Table 3.1 together with the corresponding target blur levels. For the considered cases shown in Table 3.1, it can be observed that the resulting asymmetric viewing target blur levels for the eye observing the larger blur are very close to the symmetric viewing target blur values. This indicates that the blur discrimination thresholds in the considered asymmetric case are dominated by the eye observing the larger blur level. Based on this observation, and using Equations (3.2) and (3.3), we obtain:

$$(\sigma_{th}^{asym})^2 = (\sigma_{th}^{sym})^2 - 0.53^2 \tag{3.5}$$

Table 3.1: Comparison of blur discrimination thresholds and target blur values for symmetric and asymmetric viewing case. All the values are in *arcmin* units.

| $\sigma_{ref}$ | Disparity | $\sigma_{th}^{sym}$ | $\sigma_{th}^{asym}$ | $\sigma_{tar}^{sym}$ | $\sigma_{tar}^{asym+}$ | $\sigma_{tar}^{asym-}$ |
|---|---|---|---|---|---|---|
| 0 | 21.22 | 0.8461 | 0.6966 | 0.8461 | 0.8753 | 0.6966 |
| 0 | -21.22 | 0.9266 | 0.6965 | 0.9266 | 0.8752 | 0.6965 |
| 0 | 0 | 0.9568 | 0.7623 | 0.9568 | 0.9284 | 0.7623 |
| 0.53 | 21.22 | 0.9534 | 0.7267 | 1.0908 | 1.0440 | 0.8995 |
| 0.53 | -21.22 | 1.0018 | 0.7835 | 1.1334 | 1.0843 | 0.9459 |
| 0.53 | 0 | 0.9177 | 0.7375 | 1.0598 | 1.0515 | 0.9082 |




(a) Blur discrmination treshold values modelled as an addtive term

(b) Fitting of Weber's model

Figure 3.9: Fitting of Weber's model to the subjective study data (Log-Log plot, except near origin).

## 3.5 Blur Discrimination Threshold Model

We are using Weber's model [58, 59] to characterize the subjective data from the symmetric stereo blur experiments. We show that this model provides a good fit to our subjective study data. Let $\sigma_{perRef}$ and $\sigma_{perTarget}$ denote the perceived reference and the perceived target blur levels. According to Weber's model:

$$\sigma_{perTarget} = \omega \sigma_{perRef}^{\rho} \tag{3.6}$$

where $\omega$ is the Weber fraction, and the exponent $\rho$ is a model parameter. In [59] the total perceived blur is modeled to consist of two values, namely intrinsic blur ($\sigma_{intr}$) and extrinsic blur ($\sigma_{extr}$) values. The extrinsic blur value corresponds to the actual Gaussian blur value (expressed in terms of the standard deviation of the applied blurring Gaussian function) that is applied to the considered visual pattern. In the literature [58, 59] the blur discrimination threshold is modeled as an additive term ($\sigma_{th\_add}$) to the reference blur ($\sigma_{ref}$). In this case, the extrinsic blur $\sigma_{extr}$ can be expressed as:

$$\sigma_{extr} = \sigma_{ref} + \sigma_{th\_add} \tag{3.7}$$

In our studies the blur discrimination threshold is measured as the standard deviation ($\sigma_{th}$) of a Gaussian kernel convolved with the reference blur ($\sigma_{ref}$). Thus, in our case, the extrinsic blur ($\sigma_{extr}$) is given by:

$$\sigma_{extr} = \sqrt{\sigma_{ref}^2 + \sigma_{th}^2} \tag{3.8}$$

Using Equations (3.7) and (3.8), we can remap our blur discrimination threshold ($\sigma_{th}$) to an additive one ($\sigma_{th\_add}$) as follows:

$$\sigma_{th\_add} = \sqrt{\sigma_{ref}^2 + \sigma_{th}^2} - \sigma_{ref} \tag{3.9}$$

In Figure 3.9 (a), the plots for the additive blur discrimination threshold ($\sigma_{th\_add}$) are presented for the three disparity values. Three important observations can be made: 1) the additive blur discrimination threshold values are in the same range as other studies in the literature [55]; 2) the threshold plots show a dip around reference blur level 1 arcmin; 3) the blur discrimination thresholds remain constant for the three disparity values. The blur discrimination threshold $\sigma_{th\_add}$ can be modeled as follows [58]:

$$\sigma_{th\_add} = -\sigma_{ref} + \sqrt{\omega^2(\sigma_{intr}^2 + \sigma_{ref}^2)^\rho - \sigma_{intr}^2} \tag{3.10}$$

In Equation (3.10), $\sigma_{intr}$, $\omega$ and $\rho$ are the free parameters, and $\sigma_{ref}$ corresponds to the standard deviation of the Gaussian blur applied to the reference object. After fitting the model to the subjective study data, the values of the free parameters $\sigma_{intr} = 2.246$, $\rho = 1.037$ and $\omega = 1.032$ are obtained. It can be observed from Figure 3.9 (b) that the Weber model provides a good fit to the subjective study data measured for the three disparity values, with a mean square error (MSE) of 0.007.

## 3.6 Discussion

We have compared blur discrimination thresholds for 2D and stereo-3D viewing cases. We also compared blur discrimination thresholds for symmetric and asymmetric stereo-3D viewing cases.

### *3.6.1 Disparity Effect*

As discussed in Chapter 2, under real-world viewing conditions, accommodation and vergence are coupled [53]. When using 3D displays, the eyes will always focus on the screen and an artificial disparity is introduced between the left- and right-eye views, for which the eyes try to compensate by changing the vergence length. We considered both negative and positive disparity values, along with zero disparity which corresponds to the natural viewing case in the sense that both accommodation and vergence are in sync when the disparity is zero. If the decoupling between accommodation and vergence had an effect on the blur discrimination threshold then the zero disparity case would have had a different blur discrimination threshold as compared to positive or negative disparity values. The subjective study results showed that the blur discrimination threshold values remain constant as we vary the disparity (or depth) of the synthetic test object. This indicates that the blur discrimination thresholds depend on the retinal image formed from the focused object. The 3D

blur discrimination threshold curves showed a dip around reference blur value of 1 arcmin, as also seen in the existing 2D blur discrimination experiments [58]. This dip for smaller reference blur levels [59] (near zero arcmin) is attributed to the large extrinsic blur values required to overcome the effect of intrinsic blur value. At higher reference blur levels, the effect of intrinsic blur diminishes and hence a continuous increase in the threshold blur is observed.

The past studies on blur discrimination have used an additive term to model the additional blur that is applied to the reference blur. In our study the additional blur is applied by convolving a Gaussian kernel with a standard deviation equal to the additional blur with a Gaussian kernel with a standard deviation equal to the reference blur. The mapping of our blur discrimination thresholds from a convolution model to the corresponding additive term model, indicated that the blur discrimination thresholds from our study are in the same range as existing studies [58].

### *3.6.2 Masking Effect*

In the asymmetric blur discrimination experiment, one of the eyes observes a greater (0.53 arcmin) blur than the other. Comparing the blur discrimination thresholds from this experiment with the symmetric case two possible extreme scenarios can be expected – 1) If the discrimination threshold was determined only by the eye with less blur, the blur discrimination thresholds would be identical to those obtained in the symmetric viewing experiment; 2) If the discrimination threshold was determined only by the eye with greater blur then the blur discrimination threshold for the asymmetric viewing experiment would be lower than the symmetric viewing experiment by the additional blur observed by one of the eyes. For the selected reference blur levels, the second scenario is being observed where the eye seeing the larger blur is masking the blur observed by the other eye. The decrease in the threshold

values for the asymmetric case as compared to the symmetric case depends on the value of the additional blur applied to one of the eyes (Equation 3.5). This is a strong indication that applying additional blur to one of the eyes is visible. In our subjective measurements we have applied the additional blur to the dominant eye. Our pilot studies showed that the threshold blur values remain the same even if we apply the additional asymmetric blur to the non-dominant eye. This also indicates that asymmetric coding applications using lower resolution for one of the eyes will have lower visual quality.

### *3.6.3 Eye Fatigue*

Stereoscopic 3D viewing can introduce eye fatigue, especially with asymmetric stereo-viewing. In our experiments we have considered disparity values which are within the stereo-fusion limits. During the pilot studies subjects would experience eye fatigue if they are exposed to asymmetric stereoscopic-3D content. This effect was severe for higher asymmetric blur differences. Hence, we chose a constant asymmetric blur difference of 0.53 arcmin between left and right eye, which did not introduce eye fatigue or strain.

### *3.6.4 Effect of Contrast*

The Michelson contrast ratio across the vertical edge is fixed at 0.83. The effect of contrast on blur discrimination thresholds has been studied in the literature for the 2D case. It is reported [55, 57] that for contrast ratios $> 0.1$, the blur discrimination thresholds do not vary significantly with the change in contrast values. Only for very low contrast ratio values ( $< 0.1$), the threshold will be affected and tend to have a higher value. Since we found that the change in disparity value does not affect 3D blur-discrimination thresholds, the effect of contrast should be similar to the 2D

case. It would be interesting to confirm the effect of contrast on the 3D stereo blur discrimination thresholds.

## 3.7 Conclusion

The blur discrimination thresholds using stereoscopic 3D test patterns are measured. The binocular disparity is used as the only depth cue in the image. Both symmetric and asymmetric stereo viewing cases are considered. The subjective test results indicate that the blur discrimination thresholds remain almost constant as we vary the disparity value, which corresponds to varying the object depth. These findings further indicate that one can apply 2D blur discrimination models for stereoscopic 3D blur discrimination. We have shown that the Weber model provides a good fit to the subjective study measurements. The comparison of symmetric and asymmetric blur discrimination thresholds shows that blur discrimination has a masking effect in the asymmetric case, in the sense that the eye observing the larger blur masks the blur observed by the other eye. Furthermore, the decrease in the threshold values for the asymmetric case as compared to the symmetric case depends on the value of the additional blur applied to one of the eyes.

# Chapter 4

## Texture Granularity

Textures represent spatial variations in the pixel values of the visual content. Textures can be characterized by different attributes such as granularity, regularity and directionality [70]. Texture granularity is an important attribute to characterize textures. The granularity level can be quantified based on the size of and spacing between the primitives in the texture image [7]. The texture primitive is defined as the smallest recognizable repetitive object in a texture.

In Figure 4.1, examples of low, medium and high granularity level textures are shown. It can be observed that the textures with the high granularity level (Figure 4.1(c)) have smaller size primitives, while the textures with the low granularity level (Figure 4.1(a)) have larger size primitives. In the low granularity level textures, one can easily recognize the objects in the image (e.g.,tennis ball). But, in the case of high granularity level textures, it is more difficult to discern the smaller size primitives.

The granularity of the textures depends not only on the local but also on the global context of the primitive. For example, in the case of the stack of boulders (shown in Figure 4.1(a)), the surface details of the boulder are not that important compared to the boundaries and shapes of different boulders. On the other hand, if we are focusing only on the image of a patch on the surface of the boulders, then the details on the surface of the boulder become more important. Hence, a granularity measure should consider the local as well as the global context of the primitives.

A granularity measure can be used in many applications such as texture classification, texture synthesis, texture compression and object recognition. Texture granularity is used in image segmentation applications [71] for characterizing the

(a) Low granularity level textures.

(b) Medium granularity level textures.

(c) High granularity level textures.

Figure 4.1: Sample textures in the GranTEX database.

texture properties. In segmentation applications, texture granularity is used along with other properties to define a descriptor which can be used to discriminate among different textures. In [71], the granularity is measured from gradient values in a small neighborhood. Texture granularity was also used as a feature descriptor in classification applications [72, 73] and in texture synthesis applications [74, 75] to define properties of the synthesized texture. In [75], neural network models are used to synthesize high granularity synthetic textures. Texture granularity is also used in the medical field [76–78] to diagnose and discriminate different medical conditions. The granularity along with color measures [78] are analyzed to discriminate granu-

larity in melanoma from similar areas in non-melanoma skin lesions. Texture feature vectors [76] were used to diagnose ultrasound liver images. In all these applications texture granularity is measured by quantifying image intensity variations within a small neighborhood of image pixels.

The MPEG7 [70, 73] standard provides the texture granularity level as one of the attributes in the meta data. The textures can be queried based on the granularity level of the images along with the regularity and the orientation information. One granularity metric is calculated from the Gabor filter responses [79]. The outputs of the Gabor filters from different bands are collapsed in the horizontal and vertical directions. The peaks and valleys of the autocorrelation function are used to detect the granularity of textures. A measure of texture granularity based on image particles was proposed based on the mean shift segmentation algorithm [80]. The granularity level obtained using this method is highly dependent on the correct segmentation of objects in the image. Another granularity measure [81] was proposed based on subtracting the mean gray level for the selected window location and then counting the number of sign changes. This method can capture the local granularity but will not be able to measure the global granularity property of the textures. Furthermore, none of the existing methods were proposed to measure the perceived texture granularity and no existing subjective study was conducted to quantify the perceived granularity of the textures.

This chapter is organized as follows. A database for texture granularity is presented in Section 4.1. Details about the subjective study that is conducted to measure the perceived granularity level of the textures are presented in Section 4.2. In Section 4.3, a texture granularity index (TGI) is proposed to automatically measure the level of perceived texture granularity. The correlation between the subjective granularity mean opinion scores (MOS) and the proposed objective TGI is also pre-

sented. It is shown that the proposed TGI outperforms the two other granularity metrics proposed in the literature. Conclusions are presented in Section 4.4.

## 4.1 GranTEX Database

The currently available image databases lack textures with different granularity levels which can be used to quantify and benchmark granularity measures. A database of image textures with different granularity levels is created, which we refer to as GranTEX. The aim of creating the GranTEX database is to facilitate quantifying the granularity levels of textures using subjective and objective metrics. A subjective study is conducted to measure the perceived granularity levels of the textures present in the GranTEX database.

The GranTEX database has 30 high resolution texture images which are representative of textures found in natural and man-made objects seen around us. All the texture images are cropped to a common size of 512x512 pixels. The images are stored in 24 bits-per-pixel RGB format. The granularity of the selected textures in the database can be classified into three levels namely low, medium and high. The low granularity textures have primitives with large sizes and occupy less than 50 objects in a 512x512 image. In the medium granularity level textures the primitive sizes occupy countable but many objects. In the case of high granularity level textures it is difficult to determine the texture primitive, and these have a large number of very small objects or particles. Examples for each category of textures is shown in Figure 4.1.

## 4.2 Subjective Study

A subjective study is conducted to measure the perceived granularity level of textures present in the GranTEX database. We used a single-stimulus subjective testing methodology where subjects are shown each texture image separately. The stimulus

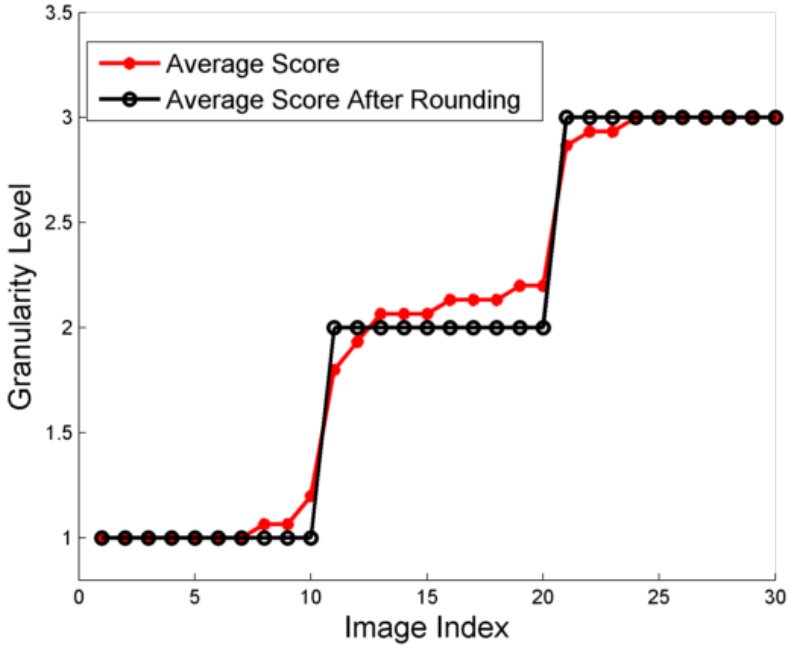



Figure 4.2: Subjective granularity mean opinion scores (MOS).

is presented on a desktop monitor with natural room lighting conditions. A MATLAB based GUI program is used to display the stimulus, and the subjects are asked to select appropriate radio button to quantify the granularity level of textures. A scale of three levels is used, where 1 refers to low, 2 refers to medium and 3 refers to high texture granularity. The display order of the texture images is randomized for every subject.

In this study 15 subjects volunteered to participate in the subjective experiment. Each subject was given a short training session to explain about the granularity level and the texture primitives before conducting the experiment. The age group of the subjects varied from 24 to 35 years. The subjects were checked for normal visual acuity and color blindness conditions. In Figure 4.2, the plot of the average granularity level subjective mean opinion scores (MOS) for 15 subjects is shown. It can be observed that the granularity levels of the selected textures fall into three categories – low, medium and high values. The images from index 1 to 10 have a

low granularity level and represent larger objects as shown in Figure 4.1(a). The images from index 11 to 20 have a medium granularity level and represent objects with medium sizes, as shown in Figure 4.1(b). The images from index 21 to 30 have a high granularity level and represent objects with finer details, as shown in Figure 4.1(c).

## 4.3 Objective Metric

This section presents a no-reference Texture Granularity Index (TGI) that can automatically quantify the perceived granularity level of textures. The wavelet transform of the texture image is analyzed at multiple resolutions. The peaks in the horizontal and vertical highpass subbands are identified. The average distance between the peaks in the horizontal and vertical directions represents the pseudo-periodicity of the primitives in the image. The pseudo-periodicity of peaks, at an adaptively determined dominant scale, is used as a measure of granularity in the image.

A block diagram of the proposed Texture Granularity Index (TGI) is shown in Figure 4.3. The input image is first divided into three undecimated wavelet subbands [82] at every level, namely, low-low (LL), horizontal-high (HH) and vertical-high (VH) subbands. The subband structure of the dyadic undecimated wavelet transform is shown in Figure 4.4. For the LL-subband, a low-pass filter is applied in both the horizontal and vertical directions. For the HH and VH subbands a high-pass filter is applied in the horizontal and vertical directions, respectively. The highest level at which the LL subband maintains important structural information is referred to as the dominant scale and is used to calculate the pseudo-periodicity of the primitives in the HH and VH subbands.

In order to determine the dominant scale, the structural similarity (SSIM) index [83] is calculated between the LL subband at the current wavelet decomposition level, and the original resolution image. If the SSIM index is above the empirically

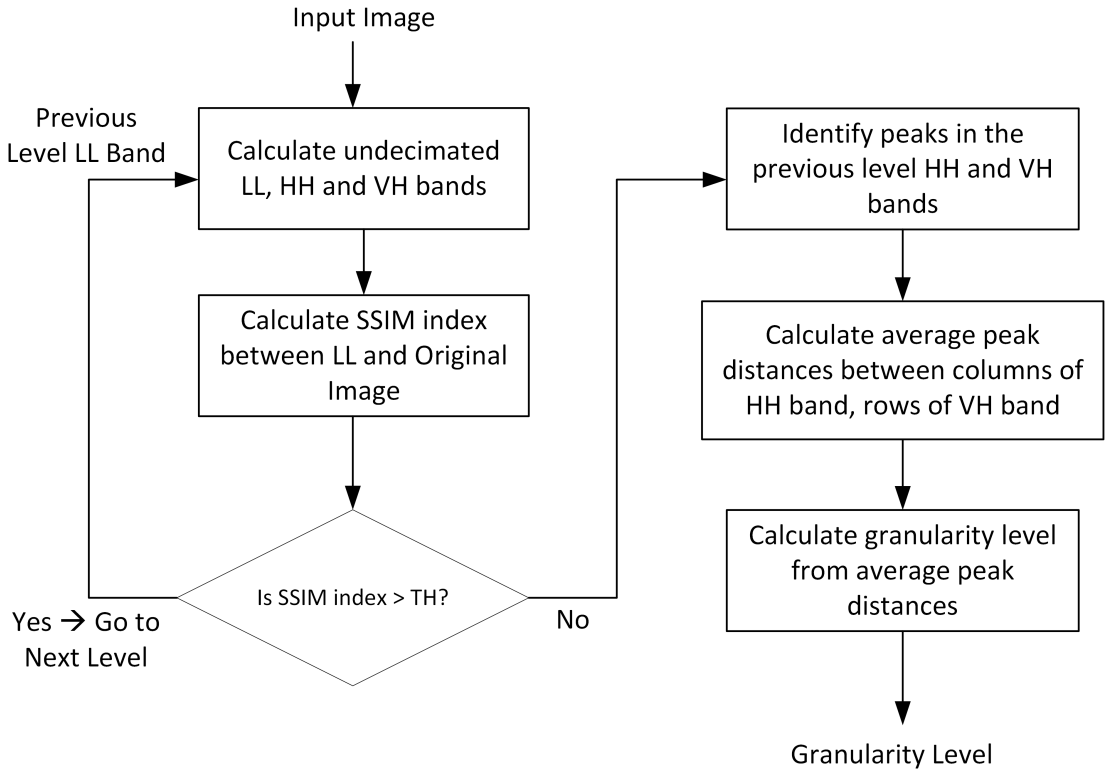



Figure 4.3: Block diagram of the proposed texture granularity index.

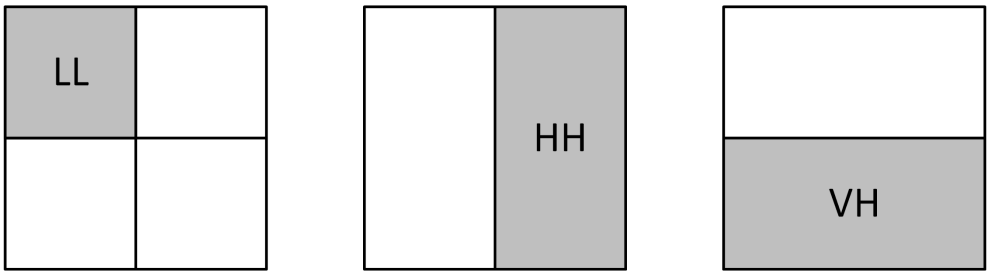



Figure 4.4: Wavelet subband structure.

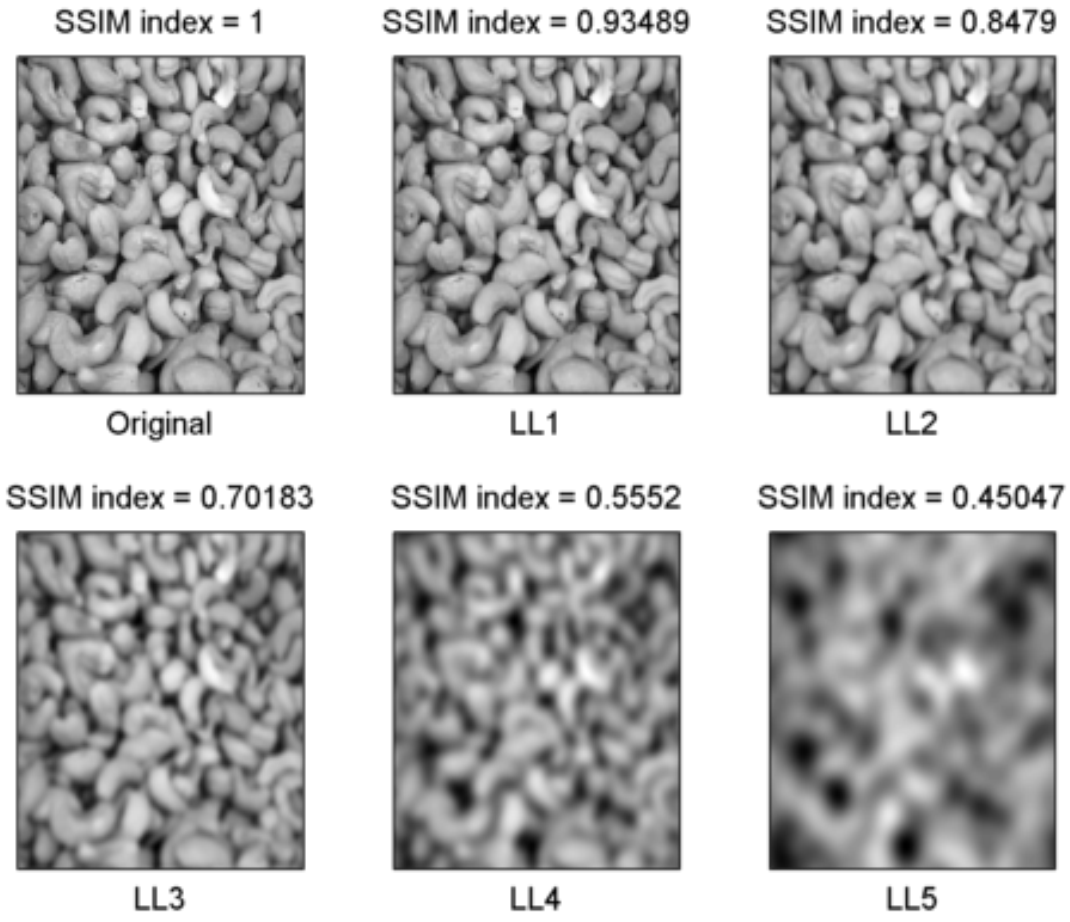



Figure 4.5: SSIM index values for LL subbands for the Cashews image.

calculated threshold, the next level is computed and analyzed. If the SSIM index is below the threshold, then the previous-level HH and VH subbands are used to calculate the pseudo-periodicity of the primitives. In Figure 4.5, an example image with LL subbands at multiple resolutions is shown along with the SSIM index. It can be observed that the image is loosing most of its structures or edges starting at the LL4 subband. An empirically calculated threshold value of 0.7 for the SSIM index is used in our experiments to decide whether the image is retaining important structures.

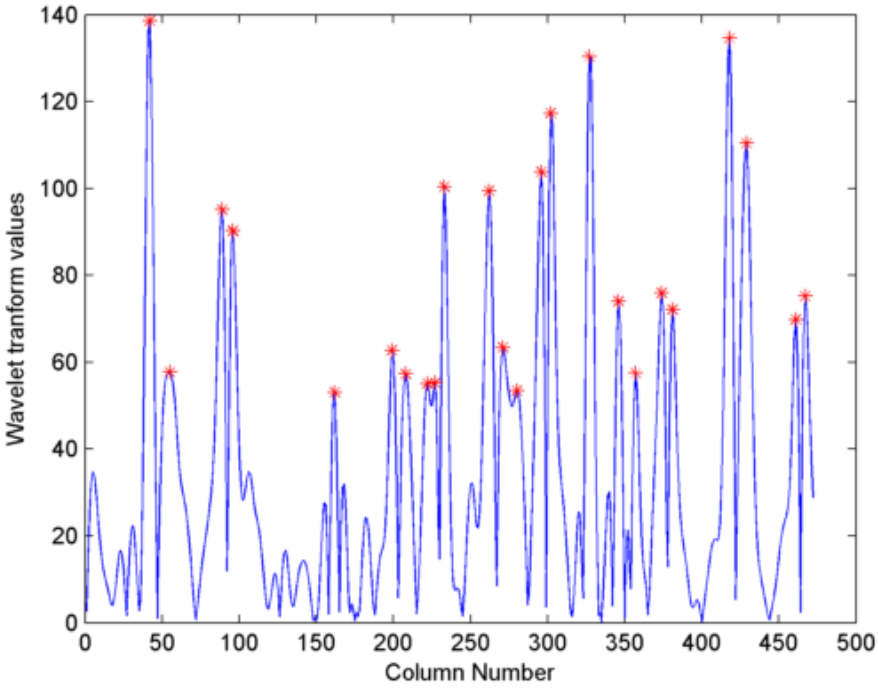



Figure 4.6: Peaks in the HH subband at Level-3, Row 28 for the Cachews image (border 20 columns are skipped on either side).

For computing the pseudo-periodicity, local peaks are detected by computing the local maxima of wavelet coefficients magnitude (absolute value) along the rows and columns of the HH and VH subbands, respectively. Outlier peaks due to noise are removed by filtering out local maxima with small values. The average distance between the peaks is calculated for every row (column) in the HH (VH) subband.The average across all the rows (columns) gives the pseudo-periodicity of the primitives in the horizontal (vertical) direction. An example showing detected peaks in the Cashews image at the Level-3 HH subband, row-28, is shown in Figure 4.6. The overall pseudo-periodicity for the considered texture image is calculated from the average of pseudo-periodicities in the horizontal and vertical directions.

In Table 4.1, the pseudo-periodicity values are presented in column 4. The perceived granularity level (MOS) and the pseudo-periodicity values are inversely

Table 4.1: MOS scores and objective metric values for textures in the GranTEX database.

| | **Texture Name** | **MOS Scores** | **Psuedo-peridicity Values** | **Proposed TGI** |
|---|---|---|---|---|
| 1 | rooftiles | 1 | 164.0593 | 0.00006 |
| 2 | bananas | 1 | 138.21715 | 0.00426 |
| 3 | pears | 1 | 128.58435 | 0.00961 |
| 4 | fruits | 1.0667 | 127.502 | 0.01042 |
| 5 | colorballs | 1.0667 | 120.5704 | 0.01678 |
| 6 | tennisballs | 1 | 107.98985 | 0.03474 |
| 7 | boulders | 1 | 99.77255 | 0.05208 |
| 8 | paving | 1 | 87.97995 | 0.08670 |
| 9 | mangoes | 1 | 69.2237 | 0.17168 |
| 10 | brickwalllarge | 1.2 | 54.22375 | 0.27309 |
| 11 | rooftilessmall | 1.8 | 46.29905 | 0.34112 |
| 12 | jackfruit | 2.0667 | 38.1117 | 0.42330 |
| 13 | leaves | 2.1333 | 36.74265 | 0.43830 |
| 14 | greenbeans | 1.9333 | 35.94465 | 0.44722 |
| 15 | corn | 2.2 | 34.0991 | 0.46834 |
| 16 | purpleflower | 2.1333 | 32.6047 | 0.48596 |
| 17 | peas | 2.0667 | 29.9109 | 0.51891 |
| 18 | pebbles | 2.1333 | 29.1614 | 0.52835 |
| 19 | cobra | 2.0667 | 28.60415 | 0.53545 |
| 20 | cachews | 2.2 | 28.31205 | 0.53920 |
| 21 | rug | 3 | 21.6498 | 0.62989 |
| 22 | fabric | 3 | 17.3998 | 0.69313 |
| 23 | concrete | 3 | 11.91145 | 0.78136 |
| 24 | lichen | 3 | 11.2303 | 0.79284 |
| 25 | curtain | 2.9333 | 11.1699 | 0.79386 |
| 26 | goldweave | 3 | 10.27745 | 0.80910 |
| 27 | palmhusk | 2.9333 | 10.2216 | 0.81006 |
| 28 | deadgrass | 3 | 9.9611 | 0.81455 |
| 29 | greengrass | 3 | 9.6486 | 0.81996 |
| 30 | snow | 2.8667 | 7.82215 | 0.85210 |

proportional, and there is an exponentially decreasing relationship between the pseudo-periodicity and the subjective granularity score values. The proposed texture gran-

ularity index (TGI) is given by the pseudo-periodicity value normalized to be in the 0 to 1 range as follows:

$$TGI = (1 - Periodicity/MaxPeriodicity)^{\beta} \tag{4.1}$$

where $MaxPeriodicity$ is a normalizing constant and $Periodicity$ is the pseudo-periodicity value calculated for the image as described previously. In our tests values for $MaxPeriodicity = 175$ and $\beta = 3.5$ are used. Table 4.1 (column 5) shows the TGI values for textures in the GranTEX database. The TGI metric scores are in the range 0∼1, the smaller TGI scores indicate low granularity level textures and the higher scores indicate high granularity level textures.

### *4.3.1 Performance Results*

The performance of the proposed TGI in terms of correlation with subjective scores is presented in this section in addition to comparison with the algorithm from the MPEG7 standard [70] and the mean-shift segmentation-based algorithm [80].

We have implemented the algorithm proposed in the MPEG7 standard [70] and used the recommended parameters in our tests. In this latter algorithm, Gabor filter responses at multiple scales and orientations are used to detect the texture granularity and the orientation. The parameters suggested in [70] are used, where the low (Ul) and high (Uh) center frequencies are set to 0.04 and 0.5, respectively, and the number of orientation parameter (K) is set to 6. The number of scales (S) parameter is set to 3 in order to match the three levels of granularity. For the mean-shift based algorithm [80], we used the implementation provided by the original authors with the parameters set as suggested in [80] (hs = 30, hr = 30, deltaS = -3, deltaR = -3). The number of layers for the analysis is set to 10.

In Table 4.2, the objective granularity metric values are presented. In column 3, the subjective granularity scores obtained from the subjective study pre-

Table 4.2: MOS scores and objective metric values for textures in the GranTEX database.

| | **Texture Names** | **Subjective MOS Range:[1∼3]** | **Proposed TGI Range:[0∼1]** | **MPEG7 Method [70] Range:[1∼3]** | **Mean-shift Method [80] Range:[0∼1]** |
|---|---|---|---|---|---|
| 1 | rooftiles | 1 | 0.00006 | 1 | 0.0533 |
| 2 | bananas | 1 | 0.00426 | 1 | 0.019 |
| 3 | pears | 1 | 0.00961 | 2 | 0.0929 |
| 4 | fruits | 1.0667 | 0.01042 | 2 | 0.0799 |
| 5 | colorballs | 1.0667 | 0.01678 | 1 | 0.1024 |
| 6 | tennisballs | 1 | 0.03474 | 2 | 0.0962 |
| 7 | boulders | 1 | 0.05208 | 1 | 0.0611 |
| 8 | paving | 1 | 0.0867 | 1.5 | 0.0162 |
| 9 | mangoes | 1 | 0.17168 | 1 | 0.0444 |
| 10 | brickwalllarge | 1.2 | 0.27309 | 1 | 0.5089 |
| 11 | rooftilessmall | 1.8 | 0.34112 | 1.5 | 0.0198 |
| 12 | jackfruit | 2.0667 | 0.4233 | 1.5 | 0.002 |
| 13 | leaves | 2.1333 | 0.4383 | 1.5 | 0.1115 |
| 14 | greenbeans | 1.9333 | 0.44722 | 1.5 | 0.1349 |
| 15 | corn | 2.2 | 0.46834 | 1 | 0.1609 |
| 16 | purpleflower | 2.1333 | 0.48596 | 1 | 0.252 |
| 17 | peas | 2.0667 | 0.51891 | 1 | 0.0623 |
| 18 | pebbles | 2.1333 | 0.52835 | 1 | 0.1685 |
| 19 | cobra | 2.0667 | 0.53545 | 2 | 0.0554 |
| 20 | cachews | 2.2 | 0.5392 | 1.5 | 0.0528 |
| 21 | rug | 3 | 0.62989 | 1 | 0.1954 |
| 22 | fabric | 3 | 0.69313 | 2 | 0.002 |
| 23 | concrete | 3 | 0.78136 | 1 | 0.036 |
| 24 | lichen | 3 | 0.79284 | 1 | 0.366 |
| 25 | curtain | 2.9333 | 0.79386 | 1 | 0.0933 |
| 26 | goldweave | 3 | 0.8091 | 1.5 | 0.0124 |
| 27 | palmhusk | 2.9333 | 0.81006 | 1.5 | 0.002 |
| 28 | deadgrass | 3 | 0.81455 | 1.5 | 0.0044 |
| 29 | greengrass | 3 | 0.81996 | 1 | 0.0813 |
| 30 | snow | 2.8667 | 0.8521 | 2 | 0.002 |

sented in Section 4.2 are also shown. The subjective granularity scores are in the range 1 to 3, where a score of 1 represents low granularity and a score of 3 represents

Table 4.3: Correlation between MOS scores and objective metric values.

| | **Pearson Linear Correlation Coefficient (PLCC %)** | **Spearman Rank Order Correlation Coefficient (SROCC %)** | **RMSE** | **MAE** |
|---|---|---|---|---|
| Porposed TGI | 97.8 | 91.34 | 0.0004 | 0.0003 |
| MPEG7 Method [70] | 2.6 | 14.16 | 0.3904 | 0.3505 |
| Mean-shift Method [80] | 16.05 | 15.05 | 0.1115 | 0.0763 |

high granularity. The proposed TGI metric values are shown in column 4, and its values are in the range 0 to 1. The smaller values represent low granularity textures, and higher values represent high granularity textures. In columns 5, the granularity metric values obtained from the MPEG7 method [70] are given, and the granularity values for this method are in the range 1 to 3. In column 6, the values of the mean-shift based granularity metric [80] are presented. The granularity metric values for this method are in the range 0 to 1.

In Table 4.3, the Pearson Linear Correlation Coefficient (PLCC) and the Spearman Rank Order Correlation Coefficient (SROCC) are presented in order to compare the correlation between the granularity MOS scores and the objective metric values. From Table 4.3, it can be observed that there is a high correlation between the proposed TGI and the subjective MOS scores with a PLCC value equal to 97.8% and an SROCC value equal to 91.34%. The granularity level scores obtained from the algorithm proposed in the MPEG7 standard [70] show a very poor correlation with the MOS scores. The PLCC value for this latter method is equal to 2.6% and the SROCC value is equal to 14.16%. The granularity levels obtained from the mean-shift segmentation based algorithm [80] also show a poor correlation with the MOS

scores. The PLCC value for the method of [80] is equal to 16.05% and the SROCC value is equal to 15.05%. The algorithm proposed in the MPEG7 standard [70] has also the limitation of not being able to handle irregular (in terms of shape and placement) textures. The mean-shift segmentation based algorithm [80] requires a reliable segmentation of the primitives, and hence can give incorrect results in the case of over segmentation due to variations within the primitive regions. These results indicate that the proposed TGI outperforms existing texture granularity metrics.

## 4.4 Conclusion

A novel no-reference perceptual texture granularity index (TGI) is proposed to predict the level of perceived granularity in a texture image. In the proposed Texture Granularity Index (TGI), the texture is analyzed at multiple resolutions to find the pseudo-periodicities of the primitives present in the texture. The estimated pseudo-periodicities are used to calculate the perceived granularity. A texture database, referred to as GranTEX, is created with images having varying levels of texture granularity. The presented GranTEX database also includes granularity level Mean Opinion Scores (MOS) for the textures present in the database. It is shown that the proposed TGI achieves a high correlation with the subjective scores and outperforms existing granularity metrics, including the algorithm from the MPEG7 standard.

# Chapter 5

## Texture Granularity and Visual Compression

In this chapter an application of the proposed perceptual texture granularity index (TGI) for visual compression is presented. For this purpose, a subjective study is conducted to assess the effect of texture granularity on compression in terms of bit-rate and perceived visual quality. We show that a logarithmic model whose parameters adapt to the level of texture granularity, provides a good fit to the obtained subjective data. The proposed logarithmic model is used for rate-distortion control by allowing the automatic selection of the needed compression rate (bits per pixel) for a target visual quality (MOS). The proposed model can also be used for automated reduced-reference visual quality assessment by providing a measure of the visual quality for a target compression ratio.

This chapter is organized as follows. In Section 5.1, a subjective study is presented to assess the effect of compression on textures with varying degrees of granularity. In Section 5.2, a logarithmic function model is proposed as a fit to the subjective test data. In Section 5.3, it is demonstrated that the proposed model can be used for rate-distortion control to select the compression ratio and for objective no-reference visual quality assessment of the compressed textures. Conclusions are presented in Section 5.4.

## 5.1 Subjective Study

Textures in images are important for the perceived visual quality. It is important for image and video compression algorithms to optimize the visual quality of textures for a given bit-rate. The texture granularity provides a measure of perceptually relevant details present in the image, which can be used to select the compression ratio for a given visual quality. In this work, subjective experiments are conducted to assess

the effect of compression on the visual quality of textures with different granularity levels. In this section the subjective study on the relationship between perceived texture granularity and compression quality is presented.

### *5.1.1 Motivation*

This work investigates the effect of texture granularity on the visual quality of JPEG2000-compressed textures. The subjective study results are used to derive a model that can be used for the automatic selection of the needed compression rate (bits per pixel) for a target visual quality (MOS), and for objective reduced-reference visual quality assessment by providing a measure of visual quality for a target compression bit-rate.

### *5.1.2 Experiments*

The subjective study is conducted using a five-point double stimulus impairment scale [33] (DSIS). The stimulus is presented on a desktop monitor with natural room lighting conditions. A MATLAB GUI program is used to render the side-by-side images onto the display. A total of 135 image pairs were used and the stimulus order was randomized for every subject. The subjects were asked to score the visual quality of images using a 5-point scale, where 1 represented poor quality and 5 represented excellent quality. A total of 15 subjects volunteered to participate in this subjective study. The age group of the subjects ranged from 21 to 38 years. The subjects were screened for normal visual acuity and no color-blindness conditions.

### *5.1.3 Results*

The plots of the compression rate in bits-per-pixel (bpp) versus the mean opinion score (MOS) for three granularity levels (low, medium and high) are shown in Figure 5.1. It is observed from the plots that as the granularity level of the texture

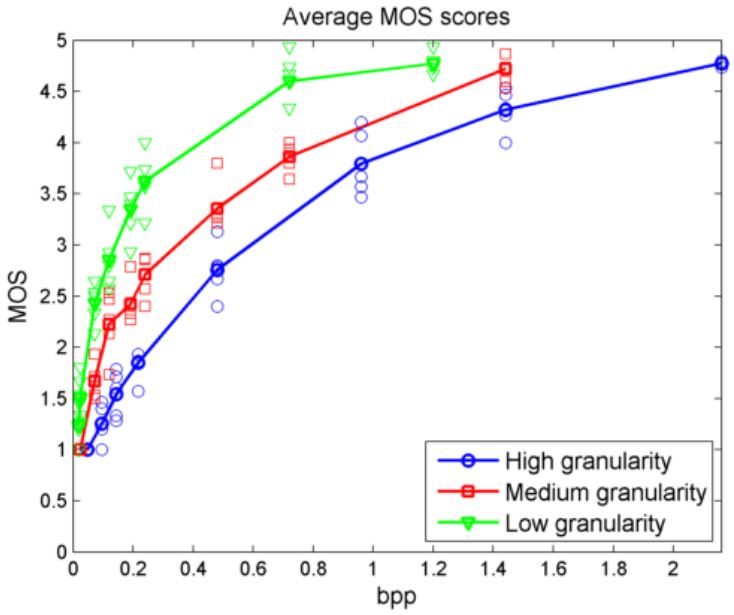



Figure 5.1: Mean opinion scores (MOS) for JPEG2000 compressed texture images.

increases the MOS decreases for the same compression ratio, and gets saturated at higher bit-rates. It is also observed from the plots that for a similar visual quality, we can achieve almost double (∼1.8) the compression for low granularity level textures as compared to the high granularity level textures, and around 1.5 times the compression for medium granularity textures as compared to the high granularity level textures. In the next section, we propose a logarithmic model as a fit to the observed subjective study data.

## 5.2 A Logarithmic Model for Texture Compression Quality

The subjective study results from the previous section indicate that, for the same compression rate, the visual quality of the textures varies with the texture granularity level. In this section we present a logarithmic fit to the subjective data to model the relationship between compression rate (bpp) and visual quality (MOS). It is shown that a constant visual quality can be achieved by adjusting the compression rate of the JPEG2000 encoder based on the granularity level of the textures.

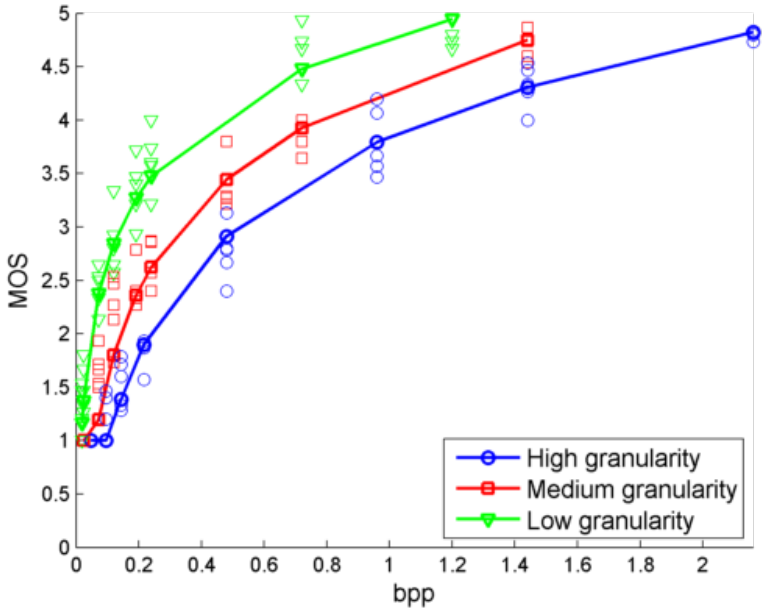



Figure 5.2: Logarithmic function model for subjective MOS vs rate (bpp) for three granularity levels.

In order to model the relationship between the compression rate (bpp) and the visual quality (MOS), the following logarithmic function is proposed:

$$MOS = \min(\max(b\ \log(a\, r), 1.0), 5.0) \tag{5.1}$$

where $r$ is the compression rate in bits per pixel (bpp), $MOS$ is the mean opinion score and ($b$, $a$) are the model parameters. The $MOS$ value is clipped within the valid range [1.0-5.0] using the min and max functions. The estimated model parameters for low, medium and high granularity level textures are given in Table 5.1. In Figure 5.2, the plots for the logarithmic fit to the subjective test results are shown.

We can rewrite Equation (5.1) to find the inverse relationship in order to determine the compression rate (bpp) that is needed to obtain a desired visual quality (MOS). This inverse relationship takes the form of an exponential function as given below:

$$r = (1/a)\exp(MOS/b) \tag{5.2}$$

Table 5.1: Logarithmic function model parameters.

| **Granularity Level** | **Parameter *b*** | **Parameter *a*** |
|---|---|---|
| Low | 0.914 | 186.637 |
| Medium | 1.186 | 38.095 |
| High | 1.27 | 20.66 |

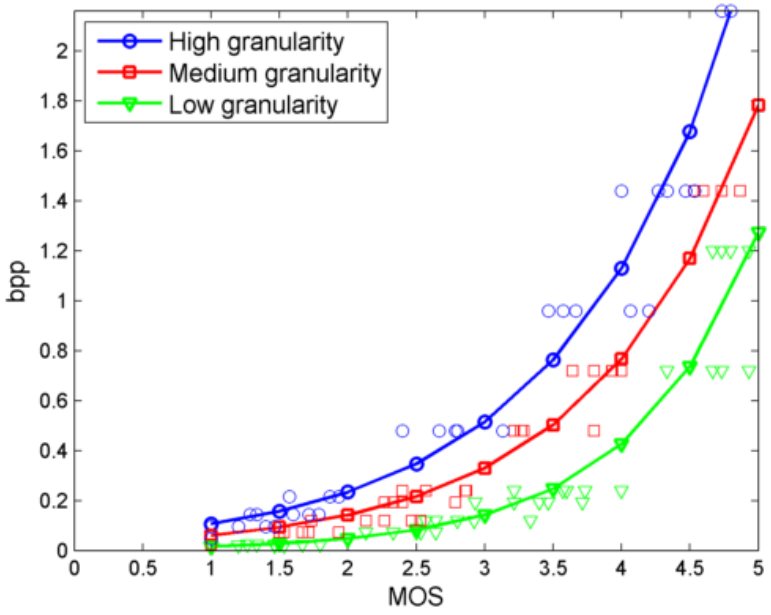



Figure 5.3: Exponential function model to fit rate (bpp) vs subjective MOS for three granularity levels.

In Figure 5.3, the plots for the exponential function model given by Equation (5.2) are shown to provide a good fit for the compression rate in terms of the subjective MOS for low, medium and high-granularity textures.

## 5.3 Rate Distortion Control for Texture Compression

It is observed from the subjective study results that the perceived visual quality of JPEG2000 compression varies with the granularity level of textures. In Section 5.2, a logarithmic function model and a corresponding exponential function model for the inverse relationship are proposed and can be used to predict the relationship between

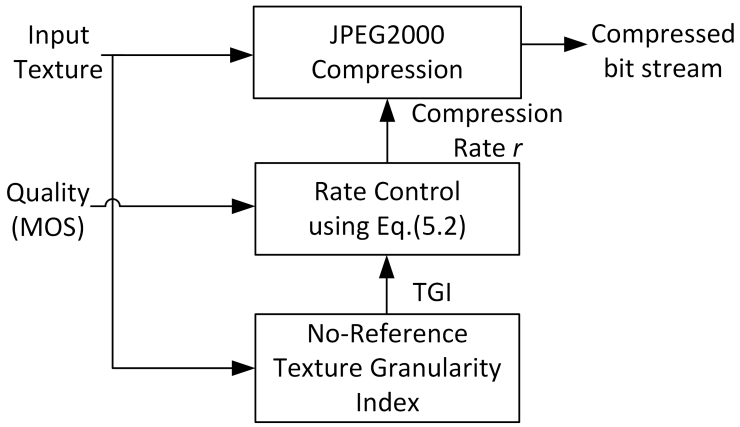



Figure 5.4: Constant-quality variable-rate compression scheme.

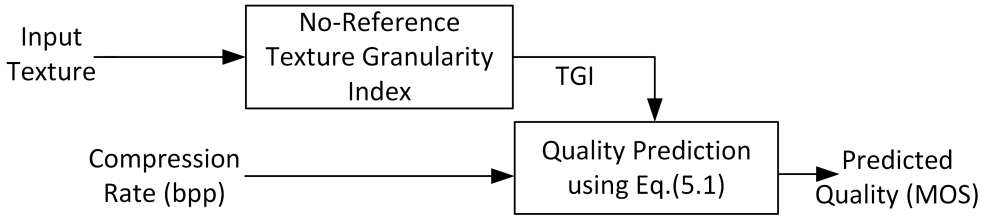



Figure 5.5: Reduced-reference visual quality assessment of compressed textures.

the compression visual quality and compression ratio for a given texture granularity level. In this section, the proposed texture granularity index (TGI) metric and the proposed models are used to demonstrate rate-distortion control for constant visual quality compression, while achieving bit-rate savings for low and medium granularity level textures as compared to high granularity level textures.

A block diagram of constant visual quality compression using rate-distortion control based on texture granularity is shown in Figure 5.4. The proposed no-reference TGI is calculated for the input texture image. The compression rate $r$ (bits per pixel) that is needed to achieve a desired visual quality (MOS) is calculated using Equation (5.2) with the parameters $a$ and $b$ determined as shown in Table 5.1 based on the texture granularity given by TGI. Alternatively, the logarithmic model

(a) Original textures

bpp=0.24, MOS=3.48 bpp=0.24, MOS=2.62 bpp=0.24, MOS=2.03

(b) JPEG2000 fixed compression rate

bpp=0.240, MOS=3.48 bpp=0.49, MOS=3.48 bpp=0.75, MOS=3.48

(c) JPEG2000 variable compression rate

Figure 5.6: Comparison of fixed vs. variable JPEG2000 compression rates for low, medium and high granularity level textures.

given by Equation (5.1) can be used to predict the visual quality (MOS) of the compressed texture for a given rate as shown in Figure 5.5.

A visual comparison of fixed versus variable compression rates using the proposed scheme is shown in Figure 5.6. In Figure 5.6(a), three original textures are shown with low (column 1), medium (column 2) and high (column 3) granularity levels. In Figure 5.6(b), the textures of Figure 5.6(a) are compressed using a fixed compression rate. It can be observed that the low granularity level texture (column 1) exhibits a better visual quality as compared to the medium and high granularity level textures. For the medium (column 2) and high (column 3) granularity level textures, some details are lost and edges show jagged artifacts. For the selected bit-rate (0.24 bpp), the estimated MOS values (using Equation (5.1)) for the low, medium and high granularity textures are 3.48, 2.62 and 2.03, respectively.

In Figure 5.6(c), variable compression rates are used for a target visual quality of MOS equal to 3.48. To achieve this target quality, the compression rates are calculated using Equation (5.2) resulting in $r = 0.240$, 0.4923 and 0.747 bpp for the low, medium and high granularity textures, respectively. From Figure 5.6(c), it can be seen that the variable compression rates for the considered textures achieve similar visual quality. In addition, it can be seen that the proposed scheme results in bit-rate savings of 0.51 bpp and 0.25 bpp for the low and medium granularity textures, respectively, as compared to the high-granularity textures.

## 5.4 Conclusion

An application of texture granularity to visual compression is presented. A subjective study was conducted to assess the visual quality of JPEG2000 image compression for textures with varying levels of granularity (low, medium and high). A logarithmic model whose parameters are adapted based on the level of texture granularity, is

proposed to fit the subjective study results allowing the prediction of the perceived visual quality of compressed textures for a target bit-rate. The inverse exponential model can be used to determine the bit-rate needed to achieve a target visual quality. The proposed Texture Granularity Index and models were used to demonstrate rate-distortion control to achieve constant visual quality and reduced-reference visual quality assessment of compressed textures.

# Chapter 6

## Texture Granularity and Synthesis Quality

Texture synthesis algorithms are used to generate or synthesize textures from small patches or from compact representations of textures. Texture synthesis algorithms are used in many applications including compression, noise reduction and view synthesis. Textures are very good candidates for lowering the bit-rate due to their high spatial frequencies. The texture-based image/video codecs attempt to save on bit-rate by synthesizing the texture regions at the decoder such that the synthesized regions perceptually resemble the original texture. This is achieved by sending a sample texture patch or synthesis parameters that represent the original textures, and that are typically very small in size as compared to the original image. In the view synthesis and hole filling applications texture synthesis techniques are used to extrapolate the texture regions using neighboring pixel values.

In this chapter the visual quality of synthesized textures for textures with different granularity levels, is studied. In order to compare texture synthesis approaches five different algorithms are selected. Finally, a reduced-reference objective delta texture granularity index is presented to measure the quality of the synthesized textures.

## 6.1 Subjective Study

In this section a subjective study is presented to measure the visual quality of texture synthesis algorithms for different granularity level textures. The texture synthesis methods produce different types of visual artifacts that lead to a loss in fidelity of the synthesized textures compared to the original. These artifacts include misalignment, blur, tiling and loss in periodicity of the primitives. The various parametric approaches [47, 50], and non-parametric approaches [41, 42, 44, 45], differ in their speed,

perceptual quality, and the amount of side information needed for synthesizing textures.

### *6.1.1 Motivation*

This study focuses on the effect of texture granularity on the quality of synthesized textures. Specifically, we would like to quantify how the granularity level of a texture has direct impact on the perceived loss in fidelity of the synthesized textures and how this loss in fidelity varies from one synthesis method to another. In other words, for synthesizing a texture of a given granularity, one method may have a perceptually better performance compared to another. For this purpose, subjective studies are conducted to assess the quality of synthesized textures with different levels (low, medium, high) of perceived texture granularity using different types of texture synthesis methods.

### *6.1.2 Selected Texture Synthesis Methods*

A brief description of five algorithms used in this study are presented here. In Algorithm-1 (Alg1), Efros et al. [44] developed a non-parametric patch-based texture synthesis algorithm called 'image quilting'. For every new patch to be synthesized and stitched, the algorithm first searches the source texture and chooses one candidate patch that satisfies the pre-defined error tolerance with respect to neighbors along the overlapped region. As mentioned already, patch-based approaches often introduce unwanted visual artifacts along overlapped boundaries. To disclose the minimum error path through the overlapped region, the minimum error boundary cut (MEBC) method is applied to smooth the transition between the overlapping boundaries of adjacent patches.

In Algorithm-2 (Alg2), Efros et al. [41] developed a non-parametric method that enforces statistics locally one pixel at a time to synthesize textures, which is more suitable for high granularity textures. The conditional distribution of each pixel, given all its neighbors synthesized so far, is estimated by searching the sample image and finding all similar neighborhoods. The texture is modeled as a Markov Random Field (MRF). This implies that the probability distribution of intensity values for a pixel, given the intensity values of its spatial neighborhood, is independent of the rest of the image.

In Algorithm-3 (Alg3), Wei et al. [42], extended Efros et al. [41]'s approach to multiple frequency bands and used vector quantization to speed up the processing. In their method instead of matching neighborhood pixels from a single image, they perform the matching based on two adjacent levels in Gaussian pyramids. They also apply tree structured vector quantization (TSVQ) to accelerate the algorithm by two orders of magnitude. They visit the pixels in a raster scan order as well as using a multi-scale framework. The method [42] works well for most of the textures. As shown in this paper, this method specifically works better for textures with high and medium granularity levels.

In Algorithm-4 (Alg4), Portilla et al. [47] proposed a parametric statistical model based on matching wavelet coefficients of multi-scale oriented filter responses. The parameters include the first and second order statistics of the filter coefficients of the neighboring orientation and scale. This method describes any texture through few parameters and thus achieves a very compact representation of textures. In this paper, this algorithm is shown to work well for textures with high granularity and good regularity.

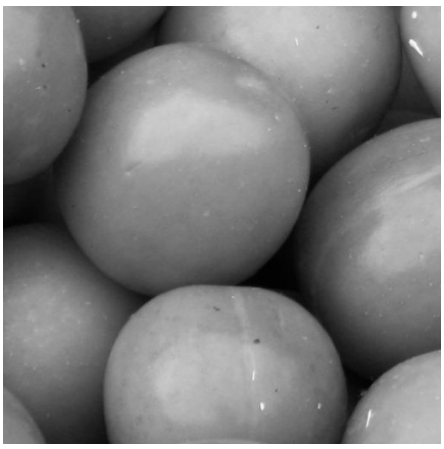
In Algorithm-5 (Alg5), Galerne et al. [84] proposed a parametric texture synthesis method. The method [84] is based on properties of two sample-based tex-

ture models namely random phase noise (RPN) and asymptotic discrete spot noise (ADSN). It is shown in this work that the method [84] works well for high-granularity textures.

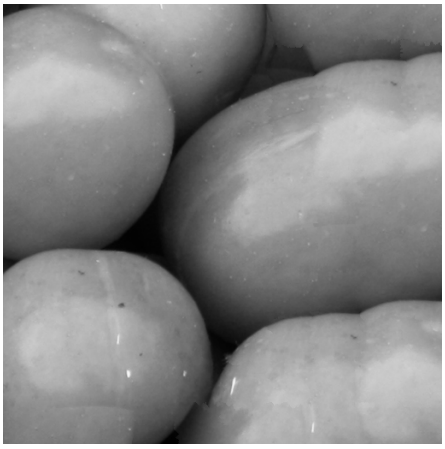
### *6.1.3 Experimental Setup*

We have selected 21 reference textures from the GranTEX database, seven from each granularity level (low, medium and high). In Figure 4.1, examples for low, medium and high granularity textures are shown. A low granularity texture will have large primitives (objects) as compared to medium and high granularity textures. For high granularity textures the primitive sizes are very small. The subjective study was conducted on 105 synthesized textures generated using 5 texture synthesis algorithms [41, 42, 44, 47, 84] which span the parametric, non-parametric, statistical, and non-statistical methods. The details of these algorithms are presented in Section 6.1.2. Since all the algorithms did not support synthesis of color images, the grayscale images were used in these experiments. All the algorithms used $64 \times 64$ size image patches as the input and generated $512 \times 512$ images as the output. The synthesized images for low, medium and high granularity textures are shown in Figures 6.1, 6.2 and 6.3, respectively.

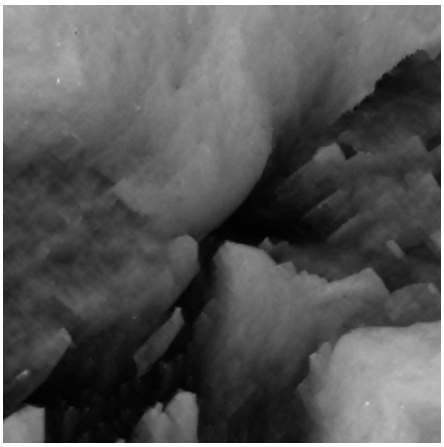
The subjects were presented both the original and synthesized textures side by side on an Alienware monitor (model: AW2310). A total of 105 image pairs were presented to every subject. The subjects were screened for normal visual acuity and color blindness conditions. Seventeen subjects volunteered to participate in this study. The display order was randomized from run to run. The experiments were administered following the subjective quality assessment procedure given in the ITU standard [32]. The subjects were asked to score the overall quality of the synthesized



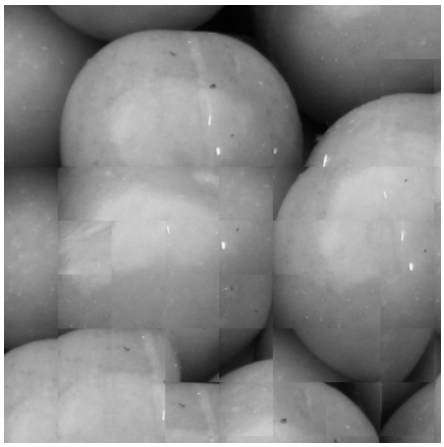
(a) Original Texture



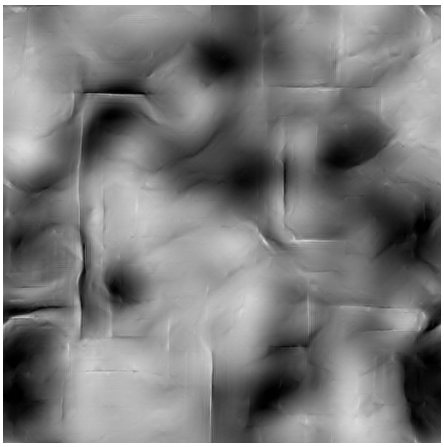
(b) Alg1 Synthesized Texture



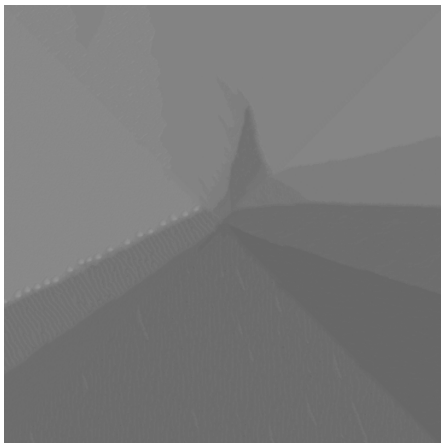
(c) Alg2 Synthesized Texture



(d) Alg3 Synthesized Texture

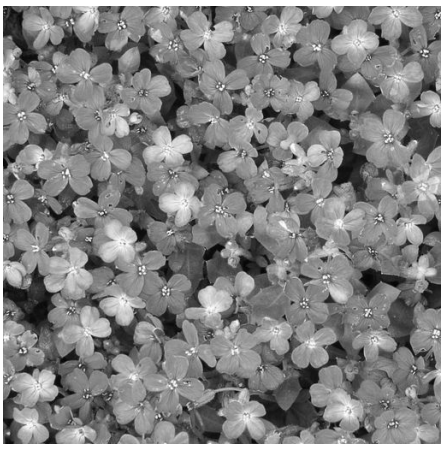


(e) Alg4 Synthesized Texture

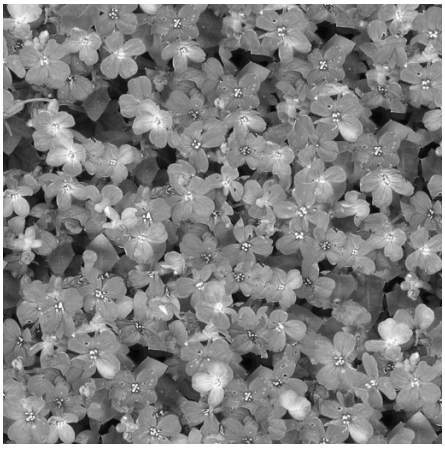


(f) Alg5 Synthesized Texture

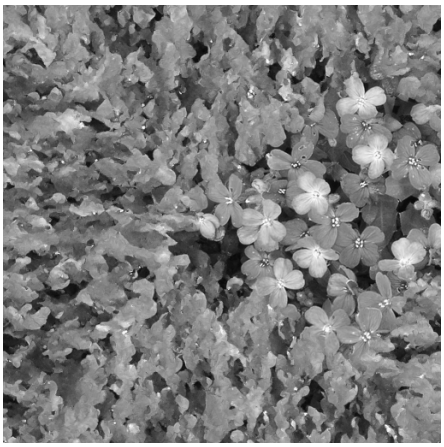
Figure 6.1: Low granularity level texture synthesis results.



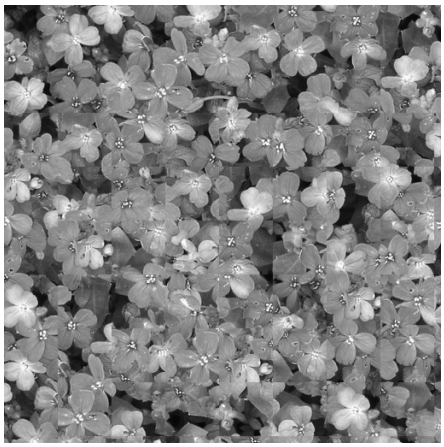
(a) Original Texture



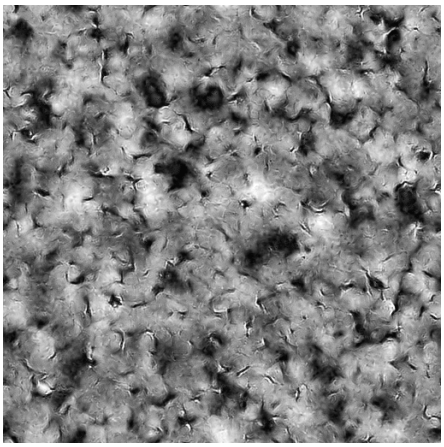
(b) Alg1 Synthesized Texture



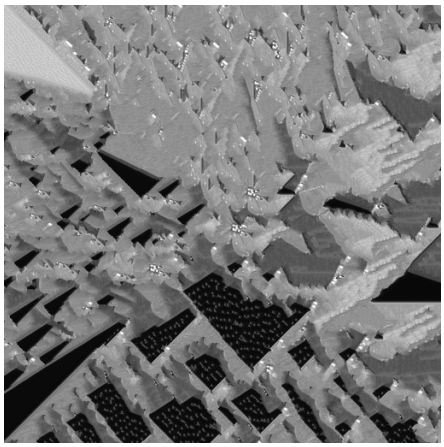
(c) Alg2 Synthesized Texture



(d) Alg3 Synthesized Texture

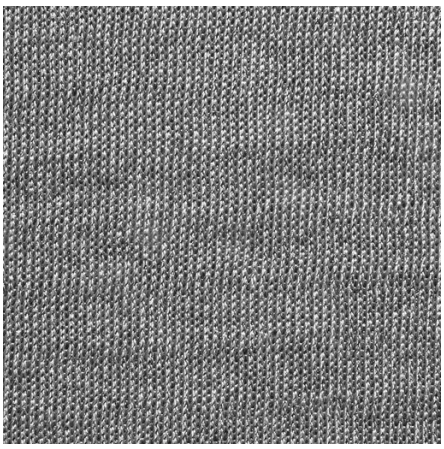


(e) Alg4 Synthesized Texture

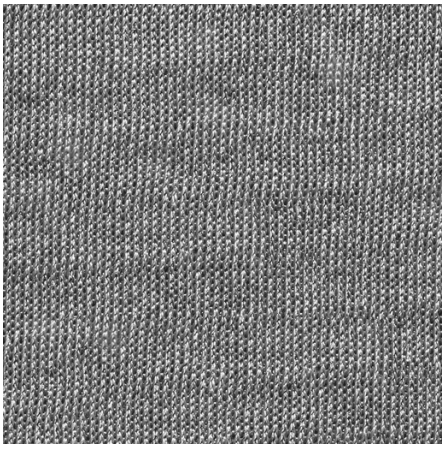


(f) Alg5 Synthesized Texture

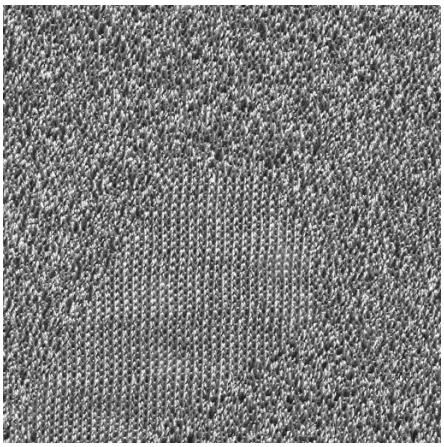
Figure 6.2: Medium granularity level texture synthesis results.



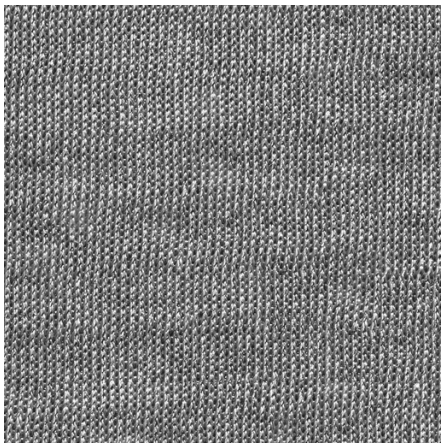
(a) Original Texture



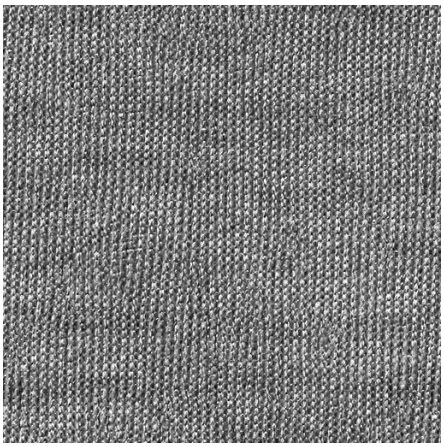
(b) Alg1 Synthesized Texture



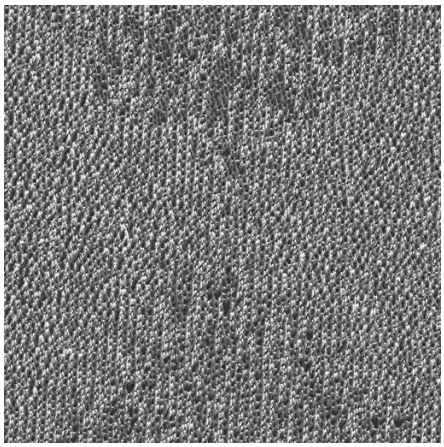
(c) Alg2 Synthesized Texture



(d) Alg3 Synthesized Texture

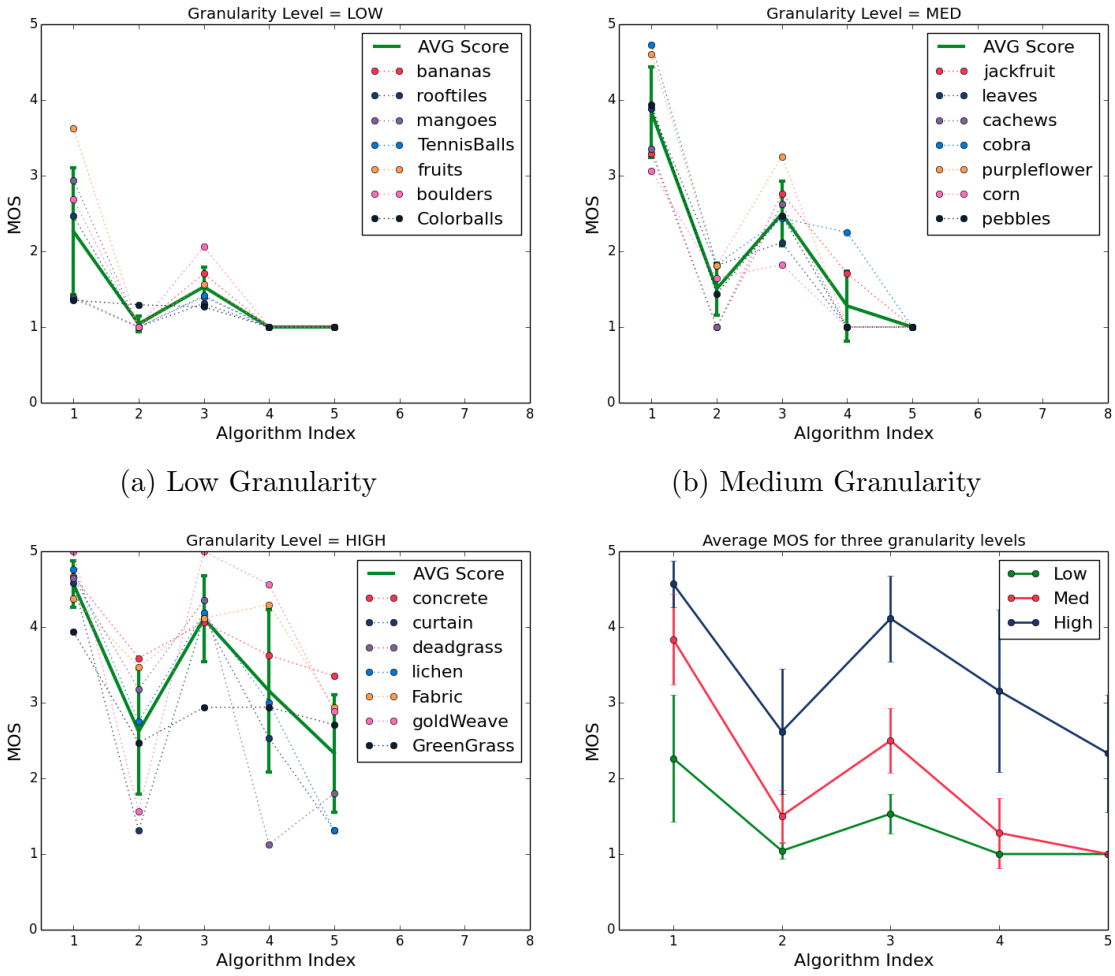


(e) Alg4 Synthesized Texture



(f) Alg5 Synthesized Texture

Figure 6.3: High granularity level texture synthesis results.




(a) Low Granularity

(b) Medium Granularity

(c) High Granularity

(d) Comparison of average MOS scores

Figure 6.4: The MOS scores for low, medium and high granularity textures.

texture for each observed texture using a five-point scale from bad to excellent, where a score of 1 represents the lowest quality and 5 represents the highest quality.

### *6.1.4 Results*

The Mean Opinion Scores (MOS) obtained from the subjective study represent the perceived quality of the synthesized textures when compared with the original textures. The plots for MOS versus algorithm index are shown in Figures 6.4 (a), (b) and (c) for low, medium and high granularity textures, respectively. It can be ob-

Table 6.1: MOS scores for low, medium and high granularity textures.

| Granularity | Alg1 | Alg2 | Alg3 | Alg4 | Alg5 |
|---|---|---|---|---|---|
| **LOW** | 2.4992 | 1.2294 | 1.8545 | 1.1150 | 1.0133 |
| **MED** | 3.7153 | 1.5004 | 2.5607 | 1.2143 | 1.0000 |
| **HIGH** | 4.1453 | 2.4706 | 3.7250 | 2.9129 | 2.1846 |
| **Average** | 3.4533 | 1.7335 | 2.7134 | 1.7474 | 1.3993 |

served that, for all the algorithms, the general trend is that the MOS is lowest for low granularity textures and highest for high granularity textures.

The overall trend in the synthesized quality of the textures is Alg1 > Alg3 > Alg4 > Alg2 > Alg5. Alg1 [44] is a patch-based non-parametric method which performs the best as compared to other methods. This method is able to preserve the local structure of the object when the object sizes are small. The considered textures in the database do not show regularity in the placement, hence it is important to consider the local structure of the objects. Alg3 [42] is a pixel-based method but uses a hierarchical approach and hence is able to capture some of the local structures in the texture. Alg2 [41] is a pixel-based method and performs poorly for most of the textures. Alg4 [47] and Alg5 [84] are parametric methods. Both of these methods perform poorly, especially for low- and medium-granularity textures.

The comparison of average MOS scores for low-, medium- and high-granularity textures is shown in Figure 6.4(d) and Table 6.1. Except for Alg1 (see Table 6.1), the synthesis quality for low- and medium-granularity textures is very poor for all the algorithms. These artifacts are caused because it is not possible to maintain the structure of the objects with smaller input patches. A granularity measure can be used to predict the smallest texture element or primitive size in the image and, thus, optimize the parameters for texture synthesis in order to achieve improved visual quality.

## 6.2 Delta Texture Granularity Index

Textures can be characterized in many different ways such as regular, irregular, stochastic, structured or grainy. The texture synthesis algorithms typically perform well on subset of these characteristic textures. The aim of the texture synthesis algorithm is to generate textures which are visually similar to the original textures, without introducing any structural, blur or tiling artifacts. It is shown in earlier studies [1] that many of the popular algorithms for objective visual quality assessment are not suitable for the texture synthesis quality evaluation. In most of these approaches a pixel-based or a local block-based statistical analysis is done which does not capture the global properties of the texture. It is shown in [85] that texture attributes can be used to quantify the quality of the synthesized textures. In this section, the difference in the Texture Granularity Index (TGI) (presented in Chapter 4) between the original and synthesized textures, which is referred to as delta-TGI, is used to assess the texture synthesis quality.

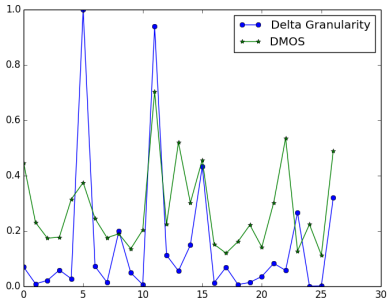
In this study the texture database provided in [1] is used. This database consists of texture images taken from the Brodatz database [86]. The provided database includes Differential Mean Opinion Scores (DMOS) between original and synthesized textures for 27 reference texture images. For this study, we selected two best performing algorithms [44, 87] from the database. It should be noted that the algorithm of [44] is the best performing Alg1 from the subjective study presented in Section 6.1. The algorithm of [87] (Alg6) is based on the algorithm [42], where a hierarchical approach is used to obtain per pixel texture synthesis. Each synthesized pixel stores the coordinates of the exemplar texture. The algorithm of [42] is the second best performing Alg3 from the subjective study presented in Section 6.1. In

Alg6, this approach is extended to achieve speed-up by introducing parallel texture synthesis.

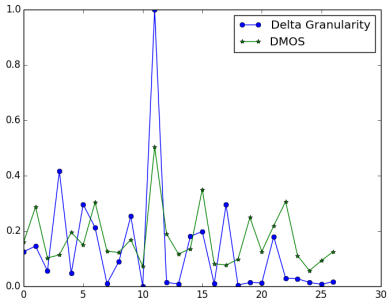
The plots of delta-TGI and DMOS values for the two selected algorithms is shown in Figure 6.5. A higher DMOS score or higher delta-TGI value indicates that there is a higher visible difference between the original and synthesized textures. It is observed from the plots in Figure 6.5 that for the majority of the textures the delta granularity values show a good correlation with the DMOS scores. The Pearson Linear Correlation Coefficient (PLCC) and Spearman Rank Order Correlation Coefficient (SROCC) between the proposed delta-TGI and DMOS scores are given in Table 6.2. For Alg1 [44], the proposed delta-TGI method gives a PLCC value equal to 64.73% and a SROCC value equal to 48.12%. For Alg6 [87], the delta-TGI method gives a PLCC value equal to 63.14% and SROCC value of 52.68. For Alg1, the existing popular CW-SSIM index [88] gives a PLCC value equal to 64.12% and a SROCC value equal to 13.98%. For Alg6, the CW-SSIM index [88] gives a PLCC value equal to -42.58 % and a SROCC value of -36.55%. It is observed from these results that even though the performance of the proposed delta-TGI and of the CW-SSIM seem to be comparable for Alg1 [44] in terms of PLCC, the proposed delta-TGI significantly outperforms CW-SSIM in terms of SROCC. Furthemore, for Alg6 [87], the proposed delta-TGI significantly outperforms the CW-SSIM index in terms of both PLCC and SROCC. In fact, for Alg6, CW-SSIM results in a negative correlation (in terms of PLCC and SROCC) with the subjective DMOS values.

## 6.3 Conclusion

Textures are important in maintaining details in image and video content. Texture synthesis methods can help to achieve high compression while maintaining the overall




(a) Alg1




(b) Alg6

Figure 6.5: Comparison of Delta-TGI and DMOS for two texture synthesis algorithms. The x-axis values represent the index of the texture images in the database of [1].

Table 6.2: Comparison of texture synthesis quality metric values.

| | **Pearson Correlation Coefficient (PLCC %)** | **Spearman Rank-order Correlation Coefficient (SROCC %)** | **RMSE** | **MAE** |
|---|---|---|---|---|
| Delta TGI (Alg1) | 64.73 | 48.12 | 0.07 | 0.06 |
| Delta TGI (Alg6) | 63.14 | 52.68 | 0.07 | 0.06 |
| CW-SSIM (Alg1) | 64.12 | 13.98 | 0.08 | 0.06 |
| CW-SSIM (Alg6) | -42.58 | -36.55 | 0.15 | 0.12 |

fidelity of the image and video content. Subjective and objective metrics for texture synthesis quality assessment are needed to design improved compression systems.

In this chapter we studied the visual quality of five texture synthesis algorithms for textures with different levels of granularity. It is found that the majority of the algorithms perform better for high granularity textures, as compared to low or medium granularity textures. The non-parametric patch-based method [44] outperforms other well-known parametric and pixel-based methods. This method leverages the local as well as global properties of the textures. The proposed texture granu-

larity index can be used to decide the patch size of the non-parametric algorithm in order to maximize the synthesis quality for a given granularity level.

A reduced-reference objective metric called delta texture granularity index (delta-TGI) is proposed to measure the perceived visual quality of synthesized textures. The proposed metric is shown to outperform existing popular full-reference visual quality metrics in terms of correlation with the subjective scores. However, the proposed delta-TGI metric does not capture some of the structural irregularities observed in the synthesized textures. It would be interesting to combine other attributes such as delta regularity values [52] in order to further improve the performance of objective visual quality assessment of synthesized textures.

# Chapter 7

## Conclusions and Future Research Directions

The research work presented in this dissertation focuses on the visual quality aspects of 3D stereo and 2D texture images. This work contributes to the areas of 3D blur discrimination, texture granularity analysis, visual compression, texture synthesis and texture quality assessment. This chapter summarizes the main contributions of this work, and provides few proposals for future research directions.

### 7.1 Summary of Contributions

- One of the important contributions of this work is the study of blur discrimination using 3D stereo test patterns. In this work, psychometric experiments are conducted to measure the blur discrimination thresholds for stereoscopic 3D viewing, as all the existing studies use 2D test patterns. The binocular disparity is used as the depth cue in these experiments. The subjective test results indicated that the blur discrimination thresholds remain almost constant as the disparity value is varied, which corresponds to varying the object depth. From these findings, it can be concluded that one can apply 2D blur discrimination models for stereoscopic 3D blur discrimination. It is shown that the Weber model provides a good fit to the subjective study measurements.

- The 3D blur discrimination experiments are presented for both symmetric and asymmetric stereo viewing cases. The comparison of symmetric and asymmetric blur discrimination thresholds showed that the blur discrimination has a masking effect in the asymmetric case, in the sense that the eye observing the larger blur masks the blur observed by the other eye. These findings further indicate that asymmetric 3D stereo coding will introduce noticeable reduction in quality even if the compression quality of one of the streams (left or right

eye) is higher. Furthermore, it is found that the blur discrimination threshold values for the asymmetric case decrease as compared to the symmetric case and that the amount of decrease depends on the value of the additional asymmetric blur that is applied to one of the eyes.

- A GranTEX texture database is presented consisting of texture images with different granularity levels. The GranTEX database has 30 images of textures typically found in natural and man-made objects seen around us and their corresponding subjective scores for the degree of perceived granularity. The textures in the database are grouped into three categories – low, medium and high granularity levels. Subjective raw (individual) scores and the mean opinion scores (MOS) are obtained through a subjective study that is conducted to quantify the perceived granularity level of the textures. The presented database is useful in comparing subjective and objective granularity measures.

- A novel no-reference perceptual granularity metric is proposed. It is shown that the proposed texture granularity index (TGI) has a much higher correlation with the subjective granularity scores, as compared to existing methods that attempt to quantify the granularity of textures, including the method proposed in the MPEG7 standard.

- A subjective study is presented to quantify the relationship between the perceived texture granularity level and the compression quality. The subjective visual quality (MOS) for the JPEG2000 compressed textures with different granularity levels is measured. The results of this study indicated that, for a similar visual quality, one can achieve almost double ($\sim$1.8) the compression for the low granularity textures as compared to the high granularity textures, and around 1.5 times the compression for the medium granularity textures as

compared to the high granularity textures. These findings are useful in texture compression and quality assessment.

- A logarithmic function model along with an exponential function for the inverse relationship is proposed to fit the subjective study results between visual quality (MOS) and compression rate (bits per pixel). The proposed model can be used for rate-distortion control by allowing the automatic selection of the needed compression ratio for a target visual quality. The proposed TGI metric and models are used to demonstrate variable bit-rate compression which achieves constant visual quality for textures with different granularity levels and at the same time achieves bit-rate savings for low and medium granularity level textures. For a given bit-rate, the proposed model also enables objective reduced-reference visual quality assessment for compressed textures based on the granularity level of the corresponding original uncompressed textures.

- The effect of texture granularity on the quality of synthesized textures is presented. For this purpose, subjective studies are conducted to assess the quality of synthesized textures with different levels (low, medium, high) of perceived texture granularity using different types of texture synthesis methods. The results of the conducted subjective studies indicate that the non-parametric patch-based texture synthesis method outperforms other well-known parametric and pixel-based methods.

- A reduced-reference visual quality index referred to as delta texture granularity index for assessing the visual quality of synthesized textures is proposed. The preliminary tests indicate that the reduced reference delta texture granularity index using the proposed TGI results in improved correlation with the differential mean opinion scores (DMOS) between original and synthesized textures, as compared to well known full-reference image quality metrics.

## 7.2 Future Research Directions

There are two potential tracks for the future research directions from this work - 1) focusing on the visual quality aspects of 3D video and 2) focusing on the visual quality aspects of texture compression and synthesis. In this section, some of the potential directions are proposed.

The 3D video technologies are still evolving, and display technologies which will minimize eye fatigue and other side effects of the 3D stereo viewing will help to propel this area of research. Based on the findings in this dissertation, the potential future directions related to blur discrimination are as follows:

- 3D depth perception is affected by both monocular and binocular depth cues. Understanding the blur discrimination performance in the presence of monocular depth cues along with the binocular depth cue (which is considered in this work) will be useful.

- The depth component in 3D stereo adds the sense of reality and there is a need to measure how this affects the quality of experience. These studies along with blur discrimination experiments will be helpful in designing a 3D stereo pipe with an improved Quality of Experience (QoE).

- The depth perception in 3D stereo varies for different color wavelengths [6]. The presented blur discrimination experiments using gray scale patterns need to be extended for color signals.

Texture analysis is an active area of research. Some of the potential research directions related to the work presented here are as follows:

- The future image and video compression standards will be using texture synthesis methods to reduce the bit-rate while maintaining a desrired visual quality. The existing well known image and video objective quality metrics will not work for texture synthesis quality assessment. The presented delta texture granularity index needs to be combined with other texture attributes to achieve better prediction of texture synthesis quality.

- Application of the proposed texture granularity index can be extended beyond visual compression and synthesis to other texture analysis tasks such as texture segmentation and retrieval.

- The different texture attributes can predict the compressibility of the textures. It would be interesting to combine multiple texture attributes in order to select an optimal compression ratio that will maximize the visual quality.

REFERENCES


[1] D. S. Swamy, K. J. Butler, D. M. Chandler, and S. S. Hemami, "Parametric quality assessment of synthesized textures," *IS&T/SPIE Electronic Imaging*, pp. 78 650–78 659, 2011.

[2] B. H. Walker, *Optical design for visual systems*. SPIE Press, 2000.

[3] R. J. Watt, *Visual processing: Computational, psychophysical and cognitive research*. Psychology Press, 1990.

[4] D. G. Curry, G. L. Martinsen, and D. G. Hopper, "Capability of the human visual system," *AeroSense 2003*, pp. 58–69, 2003.

[5] M. Robin and M. Poulin, *Digital television fundamentals*. McGraw-Hill Professional, 2000.

[6] D. F. McAllister, *Stereo computer graphics and other true 3D technologies*. Princeton University Press, 1993.

[7] R. M. Haralick, "Statistical and structural approaches to texture," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 786–804, 1979.

[8] A. B. Watson and D. G. Pelli, "QUEST: A Bayesian adaptive psychometric method," *Attention, Perception, & Psychophysics*, vol. 33, no. 2, pp. 113–120, 1983.

[9] F. A. Wichmann and N. J. Hill, "The psychometric function: I. fitting, sampling, and goodness of fit," *Perception & psychophysics*, vol. 63, no. 8, pp. 1293–1313, 2001.

[10] A. B. Watson and A. Fitzhugh, "The method of constant stimuli is inefficient," *Perception & Psychophysics*, vol. 47, no. 1, pp. 87–91, 1990.

[11] P. E. King-Smith, S. S. Grigsby, A. J. Vingrys, S. C. Benes, and A. Supowit, "Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation," *Vision research*, vol. 34, no. 7, pp. 885–912, 1994.

[12] G. K. Hung, K. J. Ciuffreda, M. Khosroyani, and B.-C. Jiang, "Models of accommodation," in *Models of the Visual System*. Springer, 2002, pp. 287–339.

[13] R. T. Held, E. A. Cooper, and M. S. Banks, "Blur and disparity are complementary cues to depth," *Current Biology*, vol. 22, no. 5, pp. 426–431, 2012.

[14] R. T. Held, E. A. Cooper, J. F. Obrien, and M. S. Banks, "Using blur to affect perceived distance and size," *ACM transactions on graphics*, vol. 29, no. 2, pp. 19–39, 2010.

[15] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *Journal of vision*, vol. 11, no. 8, pp. 11–19, 2011.

[16] M. Lambooij, M. Fortuin, I. Heynderickx, and W. IJsselsteijn, "Visual discomfort and visual fatigue of stereoscopic displays: a review," *Journal of Imaging Science and Technology*, vol. 53, no. 3, pp. 30 201–14, 2009.

[17] M. Barkowsky, K. Brunnstrm, T. Ebrahimi, L. Karam, P. Lebreton, P. Le Callet, A. Perkis, A. Raake, M. Subedar, K. Wang, L. Xing, and J. You, "Subjective and objective visual quality assessment in the context of stereoscopic 3D-TV," *3D-TV System with Depth-Image-Based Rendering*, pp. 413–437, 2013.

[18] B. Mendiburu, *3D movie making: stereoscopic digital cinema from script to screen*. CRC Press, 2012.

[19] S. Pastoor, "3D-television: A survey of recent research results on subjective requirements," *Signal processing: Image communication*, vol. 4, no. 1, pp. 21–32, 1991.

[20] P. Kerbiriou, G. Boisson, K. Sidibé, and Q. Huynh-Thu, "Depth-based representations: Which coding format for 3D video broadcast applications?" *Proceedings of SPIE*, vol. 7863, pp. 78 630–40, 2011.

[21] D. Kima, D. Mina, J. Ohab, S. Jeonb, and K. Sohna, "Depth map quality metric for three-dimensional video," *Proceedings of SPIE*, vol. 7237, no. 6, pp. 723 719–28, 2009.

[22] O. P. Gangwal and R.-P. Berretty, "Depth map post-processing for 3D-TV," *IEEE International Conference on Consumer Electronics*, pp. 1–2, 2009.

[23] H. Fradi and J. Dugelay, "Improved depth map estimation in stereo vision," *Proceedings of SPIE*, vol. 7863, pp. 78 631–78 637, 2011.

[24] J. Li and L. J. Karam, "Sparse depth estimation using multi-view 3D modeling," *IEEE International Conference on Emerging Signal Processing Applications (ESPA)*, pp. 151–154, 2012.

[25] A. Saxena, M. Sun, and A. Y. Ng, "Make3d: Learning 3D scene structure from a single still image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 5, pp. 824–840, 2009.

[26] A. Glassner, *Principles of digital image synthesis*. Morgan Kaufmann, 1995, vol. 2.

[27] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proceedings of the IEEE*, vol. 99, no. 4, pp. 626 –642, April 2011.

[28] A. Vetro, "Frame compatible formats for 3D video distribution," *IEEE International Conference on Image Processing (ICIP)*, pp. 2405–2408, 2010.

[29] N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett, "Three-dimensional displays: a review and applications analysis," *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 362–371, 2011.

[30] H. Urey, K. V. Chellappan, E. Erden, and P. Surman, "State of the art in stereoscopic and autostereoscopic displays," *Proceedings of the IEEE*, vol. 99, no. 4, pp. 540–555, 2011.

[31] W. Matusik and H. Pfister, "3D TV: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes," *ACM Transactions on Graphics (TOG)*, vol. 23, no. 3, pp. 814–824, 2004.

[32] ITU-T Recommendation, P.910, "Subjective video quality assessment methods for multimedia applications," 1999.

[33] ITU-R Recommendation, BT.500.11, "Methodology for the subjective assessment of the quality of television pictures," 2002.

[34] ITU-R Recommendation, BT.1438, "Subjective assessment of stereoscopic television pictures," 2000.

[35] K. Yamagishi, L. Karam, J. Okamoto, and T. Hayashi, "Subjective characteristics for stereoscopic high definition video," *Third International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 37–42, 2011.

[36] P. Campisi, P. Le Callet, and E. Marini, "Stereoscopic images quality assessment," *IEEE Signal Processing Conference*, pp. 2110–2114, Sept 2007.

[37] M. Tuceryan and A. K. Jain, "Texture analysis," *in The Handbook of pattern recognition and computer vision*, vol. 2, pp. 207–248, 1998.

[38] J. Byrne, S. Ierodiaconou, D. Bull, D. Redmill, and P. Hill, "Unsupervised image compression-by-synthesis within a JPEG framework," *IEEE International Conference on Image Processing*, pp. 2892–2895, 2008.

[39] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," *Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 60–65, 2005.

[40] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[41] A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," *IEEE International Conference on Computer Vision*, vol. 2, pp. 1033–1038, 1999.

[42] L.-Y. Wei and M. Levoy, "Fast texture synthesis using tree-structured vector quantization," *in Proceedings of the 27th ACM Annual Conference on Computer Graphics and Interactive Techniques*, pp. 479–488, 2000.

[43] X. Tong, J. Zhang, L. Liu, X. Wang, B. Guo, and H.-Y. Shum, "Synthesis of bidirectional texture functions on arbitrary surfaces," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 665–672, 2002.

[44] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," *in Proceedings of the 28th ACM Annual Conference on Computer Graphics and Interactive Techniques*, pp. 341–346, 2001.

[45] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick, "Graphcut textures: image and video synthesis using graph cuts," *ACM Transactions on Graphics (ToG)*, vol. 22, no. 3, pp. 277–286, 2003.

[46] E. Praun, A. Finkelstein, and H. Hoppe, "Lapped textures," *in Proceedings of the 27th ACM Annual Conference on Computer Graphics and Interactive Techniques*, pp. 465–470, 2000.

[47] J. Portilla and E. P. Simoncelli, "A parametric texture model based on joint statistics of complex wavelet coefficients," *International Journal of Computer Vision*, vol. 40, no. 1, pp. 49–70, 2000.

[48] J. R. Bergen and E. H. Adelson, "Early vision and texture perception," *Nature*, vol. 333, no. 6171, pp. 363–364, 1988.

[49] J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," *Journal of the Optical Society of America A*, vol. 7, no. 5, pp. 923–932, 1990.

[50] D. J. Heeger and J. R. Bergen, "Pyramid-based texture analysis/synthesis," *in Proceedings of the 22nd ACM Annual Conference on Computer Graphics and Interactive Techniques*, pp. 229–238, 1995.

[51] M. S. Gide and L. J. Karam, "On the assessment of the quality of textures in visual media," *in 44th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–5, 2010.

[52] S. Varadarajan and L. J. Karam, "A reduced-reference perceptual quality metric for texture synthesis," *in IEEE International Conference on Image Processing (ICIP)*, pp. 531–535, 2014.

[53] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of vision*, vol. 8, no. 3, pp. 33–63, 2008.

[54] J. Hamerly and C. Dvorak, "Detection and discrimination of blur in edges and lines," *Journal of the Optical Society of America*, vol. 71, no. 4, pp. 448–452, 1981.

[55] R. Watt and M. Morgan, "The recognition and representation of edge blur: evidence for spatial primitives in human vision," *Vision research*, vol. 23, no. 12, pp. 1465–1477, 1983.

[56] A. Pääkkönen and M. Morgan, "Effects of motion on blur discrimination," *JOSA A*, vol. 11, no. 3, pp. 992–1002, 1994.

[57] C. Chen, K. Chen, C. Tseng, S. Kuo, and K. Wu, "Constructing a metrics for blur perception with blur discrimination experiments," *Proceedings of SPIE*, vol. 7242, pp. 724 219–28, 2009.

[58] A. B. Watson and A. J. Ahumada, "Blur clarified: A review and synthesis of blur discrimination," *Journal of Vision*, vol. 11, no. 5, pp. 10–33, 2011.

[59] G. Mather and D. R. Smith, "Blur discrimination and its relation to blur-mediated depth perception," *Perception*, vol. 31, no. 10, pp. 1211–1220, 2002.

[60] M. A. Webster, M. A. Georgeson, and S. M. Webster, "Neural adjustments to image blur," *Nature Neuroscience*, vol. 5, no. 9, pp. 839–840, 2002.

[61] L. Sawides, P. de Gracia, C. Dorronsoro, M. Webster, and S. Marcos, "Adapting to blur produced by ocular high-order aberrations," *Journal of Vision*, vol. 11, no. 7, 2011.

[62] P. Aflaki, M. M. Hannuksela, J. Hakkinen, P. Lindroos, and M. Gabbouj, "Subjective study on compressed asymmetric stereoscopic video," *IEEE International Conference on Image Processing (ICIP)*, pp. 4021–4024, 2010.

[63] M. A. Weissman and A. J. Woods, "A simple method for measuring crosstalk in stereoscopic displays," *IS&T/SPIE Electronic Imaging*, vol. 7863, pp. 786 310–22, 2011.

[64] J. J. Kane, "Instrumentation for monitor calibration," *SMPTE Journal*, vol. 99, no. 9, pp. 744–752, 1990.

[65] PhotoResearch. (2014) PR-650 spectrascan colorimeter. [Online]. Available: http://www.photoresearch.com/current/pr650.asp

[66] G. Salvendy, *Handbook of Human Factors and Ergonomics*. John Wiley & Sons, 2012.

[67] D. G. Pelli, "The VideoToolbox software for visual psychophysics: Transforming numbers into movies," *Spatial Vision*, vol. 10, no. 4, pp. 437–442, 1997.

[68] G. Hagele. (2013) Dominant eye test. [Online]. Available: http://www.usaeyes.org/lasik/lasik-monovision-dominant-eye.htm

[69] W. R. Miles, "Ocular dominance in human adults," *The Journal of General Psychology*, vol. 3, no. 3, pp. 412–430, 1930.

[70] B. S. Manjunath, J.-R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703–715, 2001.

[71] A. Casals, J. Amat, and A. Grau, "Texture parameterization method for image segmentation," *European Conference on Computer Vision (ECCV)*, pp. 160 – 4, 1992.

[72] T. Sikora, "The MPEG-7 visual standard for content description-an overview," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 696–702, 2001.

[73] B. S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG-7: multimedia content description interface*. John Wiley & Sons, 2002, vol. 1.

[74] R.-I. Chang and P.-Y. Hsiao, "Artificial texture generation using force directed self-organizing maps," *IEEE International Conference on Neural Networks*, vol. 6, pp. 4123 – 4128, 1994.

[75] J. Ramfrez and M. Rivera, "Probabilistic rules for automatic texture segmentation," *International Conference on Artificial Intelligence (Lecture Notes in Artificial Intelligence)*, vol. 4293, pp. 778–88, 2006.

[76] J. W. Jeong and D. Kim, "The design of multi texture feature vector classifiers for the diagnosis of ultrasound liver images," *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 'Bridging Disciplines for Biomedicine' (Cat. No.96CH36036)*, vol. 3, pp. 1147 – 8, 1997.

[77] R. Mangoubi, M. Desai, N. Lowry, and P. Sammak, "Performance evaluation of multiresolution texture analysis of stem cell chromatin," *IEEE International Symposium on Biomedical Imaging*, pp. 380 – 3, 2008.

[78] W. Stoecker, M. Wronkiewiecz, R. Chowdhury, R. Stanley, J. Xu, A. Bangert, B. Shrestha, D. Calcara, H. Rabinovitz, M. Oliviero, F. Ahmed, L. Perry, and R. Drugge, "Detection of granularity in dermoscopy images of malignant melanoma using color and texture features," *Computerized Medical Imaging and Graphics*, vol. 35, no. 2, pp. 144 – 7, March 2011.

[79] P. Wu, B. Manjunath, S. Newsam, and H. Shin, "A texture descriptor for browsing and similarity retrieval," *Signal Processing: Image Communication*, vol. 16, no. 1, pp. 33–43, 2000.

[80] R. Dosselmann and X. D. Yang, "Mean shift particle-based texture granularity," *IEEE Instrumentation and Measurement Technology Conference (IMTC)*, pp. 101–6, 2008.

[81] P. Zinzindohoue, "Use of sign change method for granularity quantification and image segmentation-robustness," *Optik*, vol. 103, no. 2, pp. 55 – 8, 1996/08/.

[82] C. Q. Zhan and L. J. Karam, "Wavelet-based adaptive image denoising with edge preservation," *IEEE International Conference on Image Processing (ICIP)*, vol. 1, pp. 97–100, 2003.

[83] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[84] B. Galerne, Y. Gousseau, and J.-M. Morel, "Random phase textures: Theory and synthesis," *IEEE Transactions on Image Processing,*, vol. 20, no. 1, pp. 257–267, 2011.

[85] S. Varadarajan, "Texture structure analysis," Ph.D. dissertation, Arizona State University, 2014.

[86] P. Brodatz, *Textures: a photographic album for artists and designers.* Dover Publications, New York, 1966.

[87] S. Lefebvre and H. Hoppe, "Parallel controllable texture synthesis," *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 777–786, 2005.

[88] M. P. Sampat, Z. Wang, S. Gupta, A. C. Bovik, and M. K. Markey, "Complex wavelet structural similarity: A new image similarity index," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2385–2401, 2009.