2004

WAVELET AND SINE BASED ANALYSIS OF PRINT QUALITY EVALUATIONS

Vijay Venkatesh Mahalingam
University of Kentucky, mvijay@engr.uky.edu

ABSTRACT OF THESIS

WAVELET AND SINE BASED ANALYSIS OF PRINT QUALITY EVALUATIONS

Recent advances in imaging technology have resulted in a proliferation of images across different media. Before it reaches the end user, these signals undergo several transformations, which may introduce defects/artifacts that affect the perceived image quality. In order to design and evaluate these imaging systems, perceived image quality must be measured. This work focuses on analysis of print image defects and characterization of printer artifacts such as banding and graininess by using a human visual system (HVS) based framework. Specifically the work addresses the prediction of visibility of print defects (banding and graininess) by representing the print defects in terms of the orthogonal wavelet and sinusoidal basis functions and combining the detection probabilities of each basis functions to predict the response of the human visual system (HVS). The detection probabilities for basis function components and the simulated print defects are obtained from separate subjective tests. The prediction performance from both the wavelet based and sine based approaches is compared with the subjective testing results .The wavelet based prediction performs better than the sinusoidal based approach and can be a useful technique in developing measures and methods for print quality evaluations based on HVS.

KEYWORDS: Wavelets, Printer, Image Quality, Human Vision, Psychometric Function, Perception Prediction.

Vijay Venkatesh Mahalingam

February 4, 2004.

WAVELET AND SINE BASED ANALYSIS OF PRINT QUALITY EVALUATIONS

By

Vijay Venkatesh Mahalingam

Dr. Kevin D. Donohue
(Director of Thesis)

Dr. William T. Smith
(Director of Graduate Studies)

February 4, 2004

RULES FOR THE USE OF THESES

Unpublished theses submitted for the Master's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author, and with the usual scholarly acknowledgments.

Extensive copping or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate school of the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

Name Date

THESIS

Vijay Venkatesh Mahalingam

The Graduate School

University of Kentucky

2004

WAVELET AND SINE BASED ANALYSIS OF PRINT QUALITY EVALUATIONS

THESIS

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering at the University of Kentucky

By

Vijay Venkatesh Mahalingam

Lexington, Kentucky

Director: Dr. Kevin D. Donohue, Associate Professor Electrical Engineering, Lexington, Kentucky

2004

DEDICATION

To Amma, Appa, Priya, Vasu

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Kevin D. Donohue for his unwavering support and guidance in this project. I cherish the many discussions that I have had with him throughout this research effort which has improved my understanding in the critical aspects of the subject and spurred me to think independently. Your presence was always encouraging and your counsel was always insightful. Thank you Sir, I have greatly enjoyed working with you.

I would also like to thank Dr.Laurence G.Hassebrook and Dr.Daniel D.Lau for agreeing to take part in my committee and provide their valuable insight.

This work would not have been possible without the financial assistance from Lexmark International, and the technical assistance that I have received through interaction with many people, especially Luke Cui.

I would also like to thank my lab mates Sriram, Pradeep, Prabhu, Ganesh, Srini, Geetha, Hima, IL-Won and Elios and the respondents of my study for their help and their patience in enduring me through these days.

TABLE OF CONTENTS

CHAPTER 3: Prediction and Analysis

CHAPTER 4: Results and Discussion

CHAPTER 5: Conclusions and Future Work

List of Tables

List of Figures

List of Files

VVThesis.pdf……1524 KB

CHAPTER 1

Introduction and Literature Review

Image quality assessment plays a very important role in image processing applications and is vital for developers of digital imaging systems. A great deal of effort has been made in recent years to develop image quality metrics [1 – 14]. These metrics are useful in developing quality control systems for imaging systems, benchmarking imaging systems and algorithms for evaluation between various imaging systems and developing imaging systems that achieve a given level of image quality with an optimal design and lowest possible cost. This introduction discusses definitions of image quality along with classification of image quality attributes and their description. A literature review describes various developments in image quality assessment and current trends in objective/subjective image quality assessment systems as it particularly related to print defects in images. This chapter concludes with a hypothesis statement and overall organization of the thesis.

1.1 Classification of image quality attributes:

Image and video signals generally require high storage space, transmission bandwidth and processing time. These signals undergo several stages of processing before they are delivered to the human observer which may introduce distortions that could reduce the final image quality. For example they are acquired by sensing devices such as cameras, which introduces error due to optics, color calibration, sensor noise and camera motion

etc. After acquisition they are usually compressed which introduces compression artifacts and further, transmission errors happen to distort the signal during the transmission. Finally the output devices like printers, monitors introduce distortions due to low reproduction efficiency, calibration errors etc. Mostly the artifacts introduced by these devices/processes are due to the limitations of the physical capabilities of these systems. These artifacts/attributes introduced at several stages play a very important role in perceived image quality. These artifacts can be broadly classified into different criteria [1]

1. The nature of the attribute (personal, aesthetic, artifactual or preferential), which affects its amenability to objective description.
2. The impact of the attribute in different types of assessment (first-,second-,third-party), which influences the difficulty of studying it; and
3. The extent to which the attribute is affected by imaging system properties, which largely determines the degree to which it concerns system designers.

Personal and aesthetic attributes such as appearances of the subject, memories associated with the image/photograph, are very subjective in nature and may bias an observer with a priori knowledge. On the other hand artifactual attributes, which include defects such as blurring, banding, graininess, streaking and a variety of digital artifacts degrades the image quality when it is apparent, making them easier to describe objectively when compared with other attributes.

The second type of classification is based on assessment made by individuals on the image quality. The evaluation of an image by a person associated with the formation of

the image is considered to be first-party assessment whereas the subjects in the image who render an opinion are considered as a second-party assessment. Finally, an evaluation by individuals not associated with the formation of the image is called third-party assessment. First party and to an extent second-party assessments are difficult to model when compared with the third-party assessments due to their subjective nature, as they involve a higher level of visual cognition and a priori information associated with them.

Finally the classification based on the imaging system properties is important as it allows identifying attributes of image quality that are of greatest interest to designers, since they fall under the control of system designers. For example attributes such as noisiness introduced by an imaging system affect the image quality whereas the memory associated with an image does not fit into this category.

1.2 Image quality description and approaches

The goal of image quality assessment research is to develop algorithms that can automatically predict the perceived image quality in agreement with the subjective scores from the human observers. Based on the previous discussions many definitions could be formulated for the notion of image quality depending on the applications. One such definition is “The quality of an image is defined to be an impression of its merit or excellence, as perceived by an observer neither associated with the act of image formation, nor closely involved with the subject matter depicted” [1]. In the compression scenario image quality is defined as “measure of the visible difference between original and distorted images or video sequences” [10].

Image quality assessment systems provide quality metrics that predict the perceived image quality. The success of the system depends on how well the system characterizes several attributes contributing to the perceived image quality. The common approach in image quality analysis is to characterize the attributes by objective metrics such as Mean Square Error (MSE) [2], Peak Signal to noise Ratio (PSNR), Modulation Transfer Function (MTF) of the system[1] and the Noise Power Spectra (NPS)[20] and the. These metrics, even while being comparatively easier to implement and providing useful information do not characterize some attributes well enough, as they do not take into consideration the processing performed by the Human Visual System (HVS). Hence they do not correlate well with the perceived Image Quality [2] by a human observer. Therefore models involving processing of HVS to predict perceived image quality have gained considerable attention in recent years. These methods have tried to incorporate several results from psychophysical and electrophysiological studies to model the HVS. They have made significant progress in providing more accurate image quality assessments when compared with the objective image quality descriptions.

1.3 HVS Feature Based Algorithms

A major emphasis in recent years has been given to deeper analysis and understanding of the processing of HVS. These research works are particularly influenced by the need for having a better description of image quality which is very helpful. For example, in compression/coding standards it directly leads to better performance and bandwidth gain. Electrophysiologists and psychophysicists have designed experiments to understand the operation of the HVS [4, 5, 6, 7]. Electrophysiologists have been concerned with the structural organization while psychophysicists have been concerned with the perceptual

process of the HVS. These studies have resulted in development of various models that are successful in predicting the response of the HVS to a set of simple stimuli. Various attempts have been made to combine a variety of psychophysical and electrophysiological results to obtain such a model [8, 9, 10, 11, 12, 13, 14]. The approach taken by most of these visual image quality models is to determine how the lower level physiology of the visual system limits the visual sensitivity. These limits are then converted to masking thresholds which specify the limits of acceptable distortion levels, the levels above which the perceptual image quality degrades appreciably. The standard model of visual detection based on these thresholds is mostly then modeled by a collection of spatial filters sensitive to narrow ranges of spatial frequencies and orientations that cover the input image space.

1.3.1 HVS features

The HVS based models mainly incorporate three types of processes that limit the visual sensitivity variations. They are luminance variations, spatial frequency and signal content masking and (in case of video) temporal variations.

The human eye is sensitive to relative luminance rather than the absolute luminance. For a wide range of background intensities, the ratio of the threshold value ΔI divided by I is a constant. This equation

$$\Delta I / I = K \qquad (1\text{-}1)$$

is called Weber's law[15]. The value for K is roughly 0.33. Most implementations of this aspect of visual sensitivity treat it as a point process.

The human visual system's Contrast Sensitivity Function (CSF) provides the characterization of its spatial frequency response [3, 14].The CSF exhibits a band-pass characteristic, much more sensitive to mid spatial frequency and less sensitive to low and high spatial frequency range. Experiments have been designed to estimate the CSF by using Sine-wave gratings [3, 14], Gabor functions [12] and recently wavelet basis functions [17, 23] over a wide range of frequencies depending on the applications. The CSF is probably more important in terms of the visual sensitivity and is exploited well by the researcher's based on the applications. The CSF shown below in Figure 1.1, adapted from [16] is a multivariate function of spatial frequency, temporal frequency, the orientation and viewing distance.

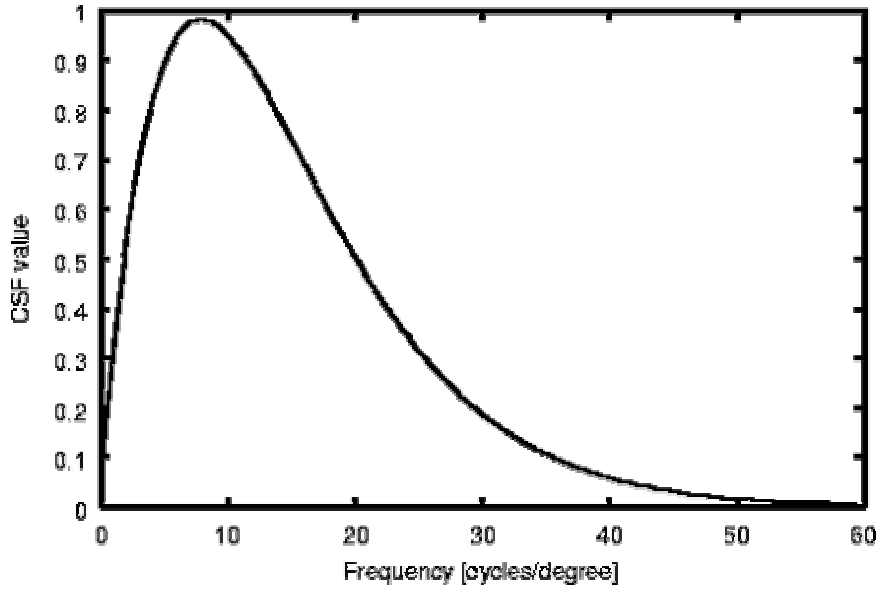

Figure 1.1 Contrast Sensitivity Function adapted from [14]

Signal content or contrast masking refers to the reduction in visibility of one image component caused by the presence of another image component with similar spatial location and frequency content. In other words the presence of a signal component in one

of the sub-bands, called masker contrast, will raise the threshold of detection for other signal components in the same sub-band [17]. The corresponding graph of this functional relation is shown below in Figure 1.2. As the masker contrast C_M increases the contrast masking takes effect and the contrast threshold of the signal C_T increases.

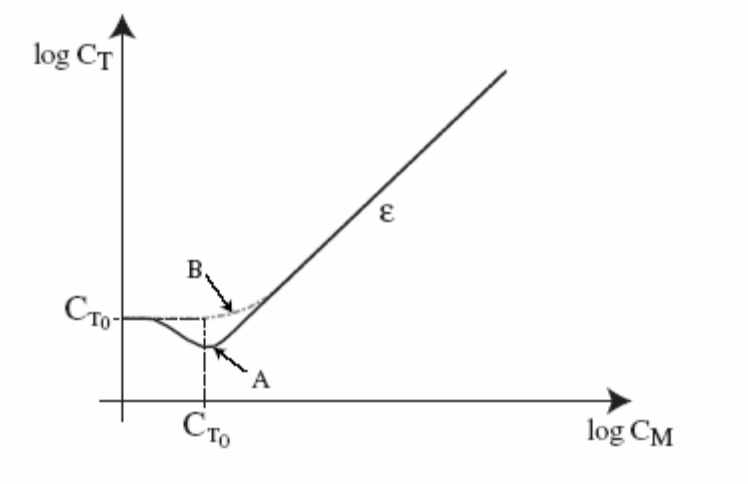

Figure 1.2 Contrast masking Function adapted from [17]

Graph A is obtained with a sinusoidal patterns while B is obtained with a noise pattern [17].

1.3.2 Review of models using HVS features

Understanding the operation of HVS has resulted in utilization of HVS features in vision models to predict the HVS response to a set of simple stimuli. These models have been focused on dealing with Image fidelity and Image quality aspects. Image fidelity is inferred as the ability to discriminate between two images whereas Image quality refers to the preference of one image over another image. The image quality models consist of several stages to model various types of sensitivity variations while the image fidelity models focus mainly on the discriminabiltiy criterion.

One of the early attempts was to use the CSF to weight different frequency components in an image for encoding of images [16]. Daly's Visual Difference Predictor (VDP) [8] an algorithm describing the human visual response was an important work in image quality framework. The VDP is a relative metric for prediction of visible differences between two digital images. The VDP included components for calibration of input images, a HVS model which addresses the three main sensitivity variations, namely, light level, spatial frequency and signal content, and a method for displaying the HVS predictions of the visible differences. The input to the VDP is two images, one of which acts as a reference image and the other is a distorted image by some processing. The output image is a map of the probability of detecting the differences between the two images as a function of their location of the images. Lubin's Visual Discrimination Model [10] is a computational method for predicting human performance in visual detection, discrimination and image-quality rating tasks. The model uses the concept of Just Noticeable Differences (JND), a discrimination threshold measure, and frequency channel vision modeling to yield robust estimates of their visual discriminabiltiy. Both VDP and VMD use a wide variety of results from the psychophysical and electrophysiological literature in order to provide a complete model of the HVS.

Watson [14] investigated the effect of luminance masking and contrast masking in the DCT domain with a motive to improve the coding efficiency of the DCT based JPEG standard. It follows an Image dependent perceptual approach and computes a visually optimal quantization matrix for any given image with minimal cost and produces better results than image independent quantization matrices. Villasenor et.al [18] have studied visibility of the wavelet basis functions and wavelet quantization noise in Y, Cb and Cr

color channels in order to achieve optimal compression using Discrete Wavelet Transform (DWT) based methods. A mathematical model for DWT noise detection thresholds as function of levels, orientation and visual resolution are constructed and used in computing a perceptually lossless quantization matrix for which the errors are below the human visual threshold.

Another important work in the field on human visual models was the Image Fidelity Assessor (IFA) developed by Taylor et.al [12]. The IFA evaluates the perceived image fidelity between two images based on multi-channel Gabor pyramid decomposition which acts as a human vision model. Psychometric functions were determined experimentally using Gabor patches, which closely match the receptive fields in the visual cortex, and described the discrimination ability of the HVS as a function of spatial frequency, orientation and adaptation level. This helped in having an integrated approach for vision modeling as the parameters of the model have a direct correspondence with the decomposition used in the model, a feature which was missing from earlier vision models. Wencheng Wu et. al [25]further extended the model to include three color channels and a new color descriptor and called it Color Image Fidelity Assessor (CIFA).

Several other multi-channel visual models based on gabor functions have been proposed [9, 26, 27]. Teo and Heeger proposed a model of perceptual distortion that is consistent with empirical findings in the physiology of visual cells and in spatial pattern psychophysics. Several other works which incorporate the perceptual process for measuring the image quality have been proposed with most of them focusing in the compression related scenario.

1.4 Print Defect Characterization

Characterization of artifactual printer defect patterns such as banding, graininess, mottle and streaking finds extensive applications in developing standards for printer quality analysis. More research work has been carried out in the field of print image quality analysis and especially towards characterization of print defects [19, 20, 21, 22, 23]. Predominantly these analyses are focused on characterizing print defects through objective quality metrics, estimation of visual threshold of those defects and also characterization based on Fourier based analysis.

Brigs et.al [19] show that the banding can be objectively quantified by their frequency analysis. Further they suggest that the reflectance range and the reflectance standard deviation of the reflectance profiles were useful in quantifying the overall banding severity. Kane et.al [20] illustrate the quantification of banding, streaking and graininess through a spectral separation technique followed by the estimation of individual noise power spectrum (NPS) in the flat field background by assuming that the components are additive. This procedure was tested with synthetic data sets, as well as data obtained from flat field prints and captures from digital devices and was found to be in agreement with visual observations. Mizes et.al [21] show that the streaking artifact can be characterized by a 1/f noise spectrum and provide the perceptibility limits of simulated 1/f noise samples. Cui et.al [22] report the visual thresholds measured for an inkjet banding profile which are lower when compared with reported results from the sinusoidal banding profiles. Michael et.al [23] introduced a perceptual print quality metrics evaluation by applying the Sarnoff JND vision model which is used in predicting digital video quality and concluded that the perceived distortion in black characters on white background is

less than for white characters on a black background. Recently more collaborative efforts have been carried out with ModelFest [28, 29] to evaluate computational models of early human vision. The phase one of the effort involves a collection of luminance contrast thresholds for 43 2-D monochromatic spatial patterns and evaluation of ModelFest data with five different spatial vision models. The work suggests the use of wavelet based models for early spatial vision modeling of the collected data. Most of these methods focusing on print defects, while providing useful information do not take into consideration the effect of HVS while analyzing the print defect patterns and hence do not correlate with the notion of perception. Hence the methods which incorporate the aspects of HVS gains importance. Recent works in the realm of print image quality are focused in comparing the effectiveness of the predicting the subjective image quality measurements carried out on printed hard copies with that of prediction based on objective image metrics based on digitized images of the hard copies [35]. This thesis work addresses those issues discussed in the international standards committee for a particular type of print defects and compares the effectiveness of two different approaches.

1.4.1 Motivation: Why HVS based and Why Wavelets?

Recent efforts have been done on characterizing the print defects such as banding graininess and streaking through wavelet based analysis [24] and Symlet patterns were shown to be robust in detecting and characterizing the print defect patterns in random backgrounds. From psycho-visual studies it is known that the HVS works with several perception channels, which are octave-wise spread over the spatial frequency range [5, 6, 7, 8]. This behavior closely resembles the dyadic structure of wavelet decomposition. The

Discrete Wavelet Transform (DWT) due to its orthogonal spatial frequency representation provides a better framework for characterizing signals with localized space-frequency properties which includes the print defects. These advantages help in simplifying the development of thresholds and combining the contrast detection probabilities to characterize the print defects by including the processing done by HVS.

1.5 Hypothesis

The objective of this thesis is to examine the feasibility of predicting the perception or visibility of print defect patterns, banding and graininess, under a flat field and noisy background conditions by a wavelet-based and sinusoidal-based approach and to compare their performance.

Subjective tests and metrics based on the wavelet and sinusoidal decompositions were designed to test the feasibility of this prediction and to compare the wavelet and the sinusoid based approaches. Separate tests were designed to estimate the detection thresholds for the individual wavelet, sinusoidal basis functions and simulated defect patterns and were used in the metrics for the prediction process.

1.6 Organization of the Thesis

Chapter 2 discusses the experiment set up and design to estimate the detection thresholds of the basis functions and simulated defect patterns. Chapter 3 presents the discussion of the prediction procedure and analysis of the experiments and Chapter 4 presents the results and discussions followed by Chapter 5 which presents the conclusion and the future research directions.

CHAPTER 2

Experiment Setup and Design

This chapter discusses the experiments used to estimate the HVS detection thresholds of the wavelet and sinusoidal basis functions along with the simulated defect patterns. The detection thresholds of the basis functions were used in the prediction process while the detection thresholds of the simulated defect patterns were used to compare the results of the prediction process with the results of the subjective testing. The chapter explains the general prediction procedure, experimental procedures involved with estimating the thresholds, the regulations involved with the procedure and the characteristics of the stimuli involved in the experiment. The discussion on the general prediction procedure employed and the relevance of the subjective testing procedure in the prediction process is presented in Section 2.1. Section 2.2 discusses about the subjective testing procedure and the measurement of the contrast thresholds. The next section, Section 2.3 describes the specifications of the experimental procedures followed in the subjective testing procedure. Section 2.4 describes the characteristics of the stimuli that were used in the experiment followed by the sensitivity study on the effect of viewing distance on the frequency content of the stimuli in Section 2.5. Finally the contrast computations of the stimuli are explained in Section 2.6.

2.1 General prediction procedure and subjective testing

The general prediction process of the print defects and the relevance of the subjective testing procedure in the prediction process can be explained with the help of a flow chart as shown in Figure. 2.1. The prediction of the HVS response for both, wavelet based and

the sinusoid based approach follow a general technique and differs only in the transform domain representation and calculation of contrast from the transform coefficients.

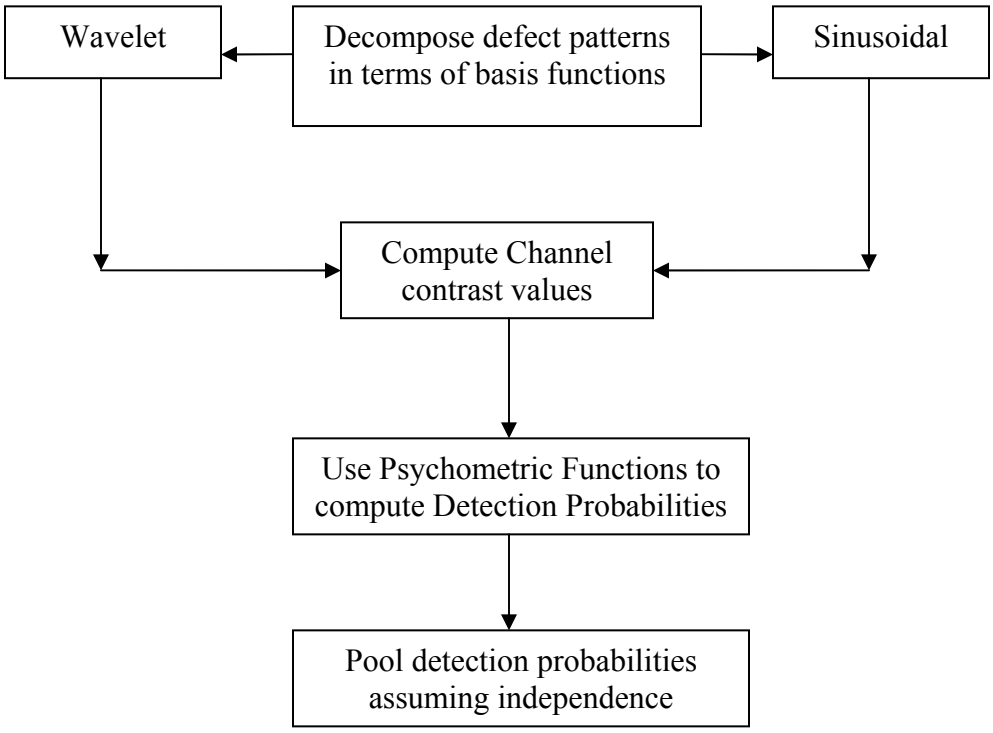

Figure 2.1 General Prediction Procedure

The general prediction procedure is explained below,

1. The first step in the prediction is the simulation of defect patterns, banding and graininess, at particular contrast level.
2. The second step is the decomposition of the image into corresponding basis function/transform domain coefficients according to either the wavelet or the sinusoidal based method.
3. The third step is computation of a contrast value from the resulting basis function/transform domain coefficients

4. The fourth step is to obtain a detection probability by substituting the computed contrast value in the respective psychometric function of the basis functions, which serves as a mapping function between the computed contrast values and detection probability.

5. The final stage comprises of combining the detection probabilities obtained together, referred to as pooling, to arrive at the predicted response.

The above steps were repeated for different contrast values until the contrast level yielding 92% detection probability was determined from the pooled probability values.

The pooling procedure mimics the process of integration of information from various channels in the brain. As there is no firm experimental evidence on how the brain integrates the data, this process is carried out by assuming total independence between the various visual channels. The subjective testing procedure on the basis functions helps us to estimate the contrast thresholds of the basis functions which are used to create the psychometric functions that serve as a mapping tool between the contrast values and the detection probabilities. The subjective detection thresholds of the simulated print defects are used to compare the results obtained from the prediction process.

2.2 Subjective testing procedure

The objective of the subjective testing procedure was to estimate the detection thresholds of various basis function stimuli and simulated print defect patterns. The estimated thresholds were then used to determine the feasibility of predicting the perception or visibility of print defect patterns. Predictions based on wavelet and sinusoidal basis approaches were compared. Therefore three separate subjective tests were performed.

One test presented the subjects with wavelet patterns at various scales and orientations, another test presented sinusoidal patterns at various frequencies and orientations and the final test presented simulated banding defect patterns at various scales and sizes and simulated grain pattern at various coarseness and sizes. A 2-alternative forced choice (2AFC) test was used with a Bayesian adaptive psychometric procedure referred to as QUEST [30], to determine the contrast thresholds and detection probabilities. Subjects used in the experiments were drafted for this study through announcements posted at the University of Kentucky and Lexmark International Inc. Results from 22 subjects taking 25 subjective tests, due to some overlap, were used in this study. Eight subjects were tested individually using wavelet patterns, eight with sinusoidal patterns and nine subjects with simulated grain defect patterns and simulated banding defect pattern. Subjective testing was performed in an office with no windows and a constant fluorescent light illumination.

2.2.1 Threshold Measurement

The subjective testing using the 2AFC procedure involves presenting a sequence of image pairs, one image containing only the background and the other containing a scaled stimulus added to the background. All the patterns were presented in both flat field and random noise backgrounds. The uniform background luminance was set to a mean gray level value of 128 and the random noise was generated as a white Gaussian noise with a -23dB contrast level. Following the presentation of the image pairs, the subjects used a mouse click to choose the image that appeared to contain the stimulus. The order of the image pair presentation was random for each trial and each image was presented for 1 second with a 0.5 second pause between images. These times were chosen to minimize

the time period of the subjective testing procedure, while being consistent with previous works [31]. The amplitude/contrast of the stimulus presented was varied adaptively from trial to trial based on the decision of the observer by a Bayesian psychometric procedure called QUEST [30] that sequentially estimated the 92% detection threshold. Prior information (in the form of prior probability density function) is used along with the data in the form of the likelihood function to guide the contrast value of the patterns used in each trial. The Quest function Q(T) [30] is given as

$$Q(T) = \ln f_T(T) + \ln f_{D/T}(D/T) \qquad (2\text{-}1)$$

Where $f_T(T)$ is the prior probability function of the threshold and $f_{D/T}(D/T)$ is the likelihood function of the threshold conditional upon the data D.

The initial guess to start the subjective testing procedure for each type of stimulus, varying in frequency, orientation, size and background, were estimated by empirical testing methods. The initial values chosen would be such that the testing procedure would reach the estimated thresholds within first six trials barring some finger errors by the subjects. Further trials allowed the adaptive testing process to estimate the exact threshold of the subject for various stimuli. The initial guess value for various stimuli used in the experiment play a very important role in the convergence of the experiments on an accurate threshold value within limited number of trials. Following each trial the Quest function is estimated depending upon the success or failure of detection of stimuli by the observer and is given as

$$Q_n[j] = Q_{n-1}[j] + \begin{Bmatrix} S[j-k] \text{ if sucess} \\ \text{F[j - k] if failure} \end{Bmatrix} \quad (2\text{-}2)$$

where *S[j-k]* is the success function and *F[j-k]* is the failure function both of which are obtained by shifting the psychometric function modeled by Weibull distribution[30].

Sixteen trials were used in each of the subjective tests and the final estimate of the threshold was taken to be the peak value of the likelihood function at the final trial. The procedure based on fixed number of trials for completing the experiment was chosen over a procedure based on confidence limits for the following reasons; simplicity of the experiment procedure, limited running time of the experiment which helps in maintaining the focus of the subject yielding better threshold estimates [31] and also it compares well with the conventional psychometric experiments based on staircase methods while intelligently making use of the knowledge obtained from previous trials inherent of the adaptive techniques.

The original stimuli were represented as gray scale images with gray-levels between 0 and 255 represented as

$$I(g) = B_g + S * (Stimuli) \quad (2\text{-}3)$$

Where *I(g)* is the displayed gray level image, *Bg* represents the background, either a mean gray level of 128 for uniform background or random noise field of -23dB contrast level for noisy background, *S* is the scaling factor of the stimuli which is varied by the adaptive procedure and the *Stimuli* refers to basis functions and the simulated defect patterns. For each trial, depending on the decision of the observer, the Scale factor *S* was

varied by using the adaptive Quest process. As mentioned above the gray-levels of the displayed final image was maintained between 0 and 255, the limits of saturation of the display system.

2.3 Experiment Procedures and Specifications

All the test patterns that were displayed as a grayscale images on a LaCIE 20" with physical dimensions of 371mm x 297mm. The resolution of the display was set to 1280 x 1024 pixels which yielded a resolution of 34.5 pixels/cm. All the stimuli were positioned at the center of the image of size 384 x 384 pixels with no additional fixation cues, therefore the edges of the images served as fixation cues. The background of the monitor outside the pattern display range was set to zero.

2.3.1 Display Visual Resolution

The visibility of the stimuli presented depends upon the display visual resolution in pixels/degree of the display system. For a given viewing distance and a display resolution, the effective display visual resolution visual angle is

$$d = x.r.\tan\left(\pi/180\right) \approx x.r.\pi/180 \approx x.r/57.3 \qquad (2\text{-}4)$$

where x is distance in cms, r is display resolution in pixels/cm and d is the display visual resolution in pixels/degree. In this study the viewing distance was set at 57 cms and therefore an image of size 384 x 384 pixels occupies 34.5 pixels/cm yielding a display visual resolution of 34.75 pixels/degree. At 57 cms viewing distance an image of the 384x384 pixels subtends 11.13 visual degrees. The spatial frequency content of the stimuli presented in the subjective testing procedure are expressed in cycles per degree (CPD) as this unit describes the frequency response, independent of the viewing distance.

With the above conditions the 384x384 image subtends 11.13 visual degrees at 57 cms viewing distance. Even though the viewing distance was kept constant throughout the subjective testing there may have been instances where the subjects did not maintain the specified distance. A sensitivity study analyzing the effects of change in viewing distance to the frequency components of the images displayed is presented in Section 2.5.

2.3.2 Display System Settings

The Gamma of the display, the relationship between the luminance generated by the device and applied voltage, was set to 2.7 to provide a smallest possible contrast level variations at the intended luminance level for integer increments of the display values in graphic buffer. The brightness and the contrast of the CRT were manually adjusted at the beginning of the testing so that the low-contrast stimuli were not visible for several nonzero values of the graphic buffer. At these settings the display system exhibited an approximate linear characteristic over the entire range of the display values. These settings were kept constant throughout all subjective tests. However in order to maintain consistency and a need to express those luminance values in an device independent representation the physical luminance of the display was measured at 17 uniformly spaced gray level increments from 0-255 using a Minolta Chromameter CS-100A before each subjective test. The luminance (Y-Value) was recorded with the results of each subjective test to convert the gray level graphic buffer values to luminance value, a device independent representation.

The background luminance for the 384x384 pixel field was set to a gray level of 128 or the mean gray level for the random field background throughout all subjective testing procedures. The display system parameters were adjusted properly so that the mean gray

level value corresponded to a luminance measure of 20cd/m^2. Although the luminance values were measured independently for each test, very little variations existed between measurements and this luminance measure represents the value for all the tests. The values of the display buffer gray levels and the measured Luminance (Y) value with the symmetric error bars of unit standard deviation is shown in Figure 2.2.

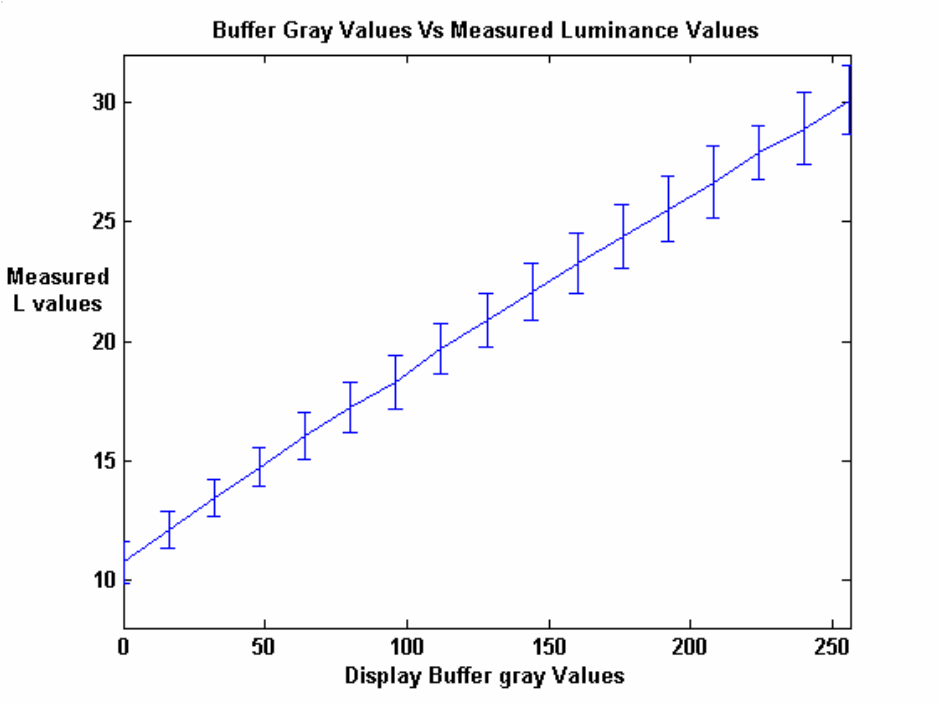

Figure 2.2 Display buffer values vs. Measured L values

The linear relationship between the display buffer gray levels and the measured luminance is evident from the above measurements. All the stimuli of intermediate gray levels ranging from 0-255 were converted into luminance values through linear interpolation with the help of the above relationship and were expressed in cd/m^2.

2.4 Stimuli Description and Generation

The subjective testing procedure used four stimuli sets in three separate tests, two basis functions, wavelets and sinusoid, and the simulated print defect patterns, banding and

grain under two different background conditions for the predicting the perceptibility of print defect patterns.

2.4.1 Wavelet Stimuli Description and Generation

The stimuli for the wavelet basis function tests consisted of 2-dimensional Symlets of 8th order[32], spanning from Levels-2 through 4 from a 4-level wavelet decomposition oriented along horizontal, vertical and diagonal direction, where Level-1 represents the smallest scale or highest frequency. Two examples of wavelet basis function from Level-4, the largest scale, and level-3 are shown below in figure 2.3(a) and 2.3(b).

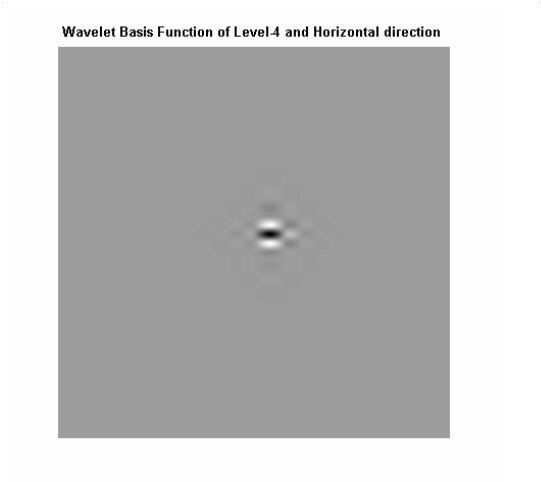

Figure 2.3: a) Wavelet basis function of level-4 and horizontal direction

Figure 2.3: b) Wavelet basis function of level-3 and horizontal direction

Figure 2.3 Wavelet basis function from Level-4 and Level-3

Totally 9 different stimuli were presented, each one varying across three levels and the three different orientations. Symlets are used in the subjective testing as they were shown to be more suitable in characterizing the print defects [24] due to their near symmetry in spatial domain and also due to their smoothness property (due to their high order). The higher order ensures a smoother spatial function with greater localization in the frequency

domain. The wavelet functions have a good localization in space as well as frequency domain. The low pass and high pass decomposition filter coefficients for an 8th order symlet are shown in Figure 2.4(a) and 2.4(b).

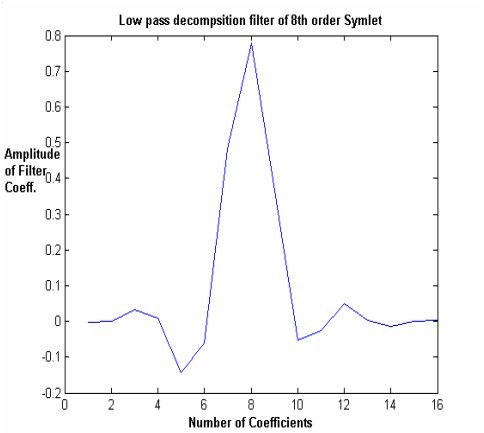

Figure 2.4: a) Low pass decomposition filter of 8th order Symlet

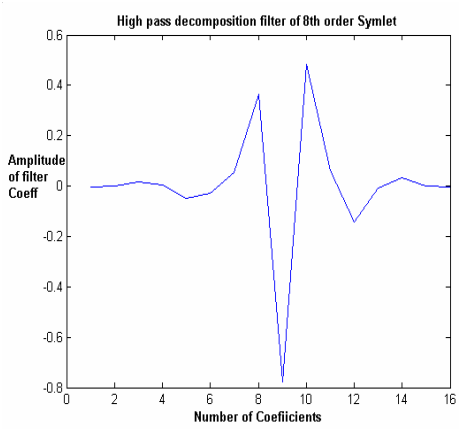

Figure 2.4: b) High pass decomposition filter of 8th order Symlet

Figure 2.4 Low-pass and high-pass decomposition filter coefficients of 8th order Symlet

The spectral characteristics of the low-pass decomposition filter and the high-pass above decomposition filters are shown below in Figure 2.5(a) and 2.5(b). Note the low-pass and the high-pass characteristics of the respective filters.

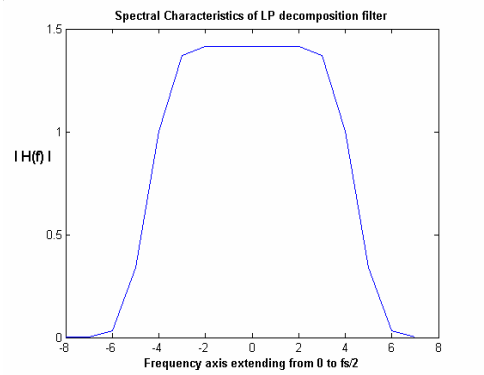

Figure 2.5: a) Spectrum of low-pass decomposition filter of 8th order Symlet

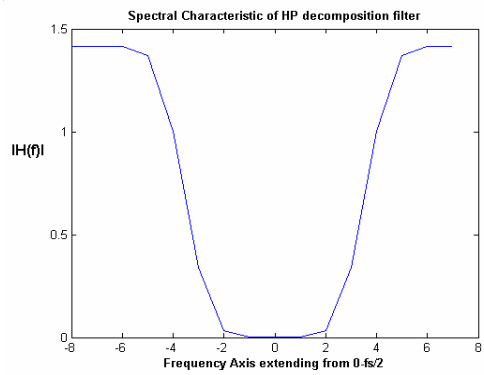

Figure 2.5: b)Spectrum of high pass decomposition filter of 8th order Symlet

Figure 2.5 Spectral Characteristics of low-pass and high-pass decomposition filters

The wavelet stimulus from the first level is not used because the stimuli from Level-4 to Level-2 cover approximately 1.07 – 8.58 cycles per degree (cpd) of spatial frequency range, so the contribution from the Level-1 was considered to be less significant due to very low sensitivity of the HVS for frequencies below 1 cpd as evident from the contrast sensitivity curve in Figure.1.1. In addition to this fact, it allowed for fewer wavelet patterns to be presented in the subjective tests. To ensure that the energy was not present at the first level of the image, the defect patterns were all filtered to zero out energy at this level.

The wavelet stimuli are generated by computing the impulse response of the wavelet filters by setting a single coefficient, in a specific level and orientation, to unit value in a 4-level DWT transform as shown in the Figure 2.6(a), setting all the other values to zero and computing the inverse DWT of the signal. The resulting symlet of a Level-4 wavelet pattern (largest scale corresponding to low frequency) with a horizontal orientation in flat and background is shown in Figure. 2.6(b). The DWT operates by an octave based division of the frequency spectrum, and the divisions on Figure 2.6(a) represent those octave bands.

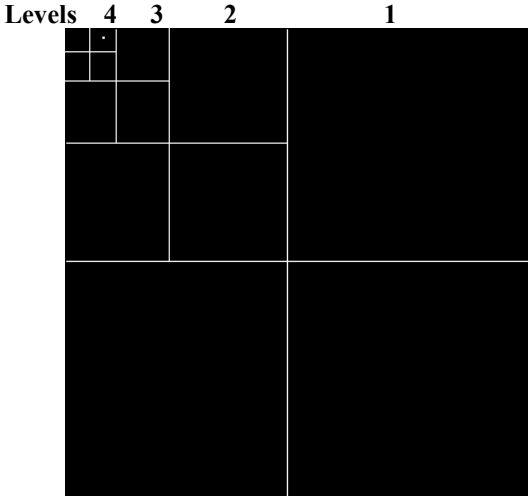

Figure 2.6: a) A 4-level DWT with a single coefficient in level-4 and horizontal direction, set to 1

Figure 2.6: b) Symlet basis function with horizontal orientation from level4 in flat field background

Figure 2-6 A 4-level DWT and corresponding basis function

At the first decomposition stage or at the 4th level of the transform, the band covers the spectral range extending from one half of the Nyquist rate (fs/4) to Nyquist rate (fs/2) which is equivalent to half of the display resolution or the sampling frequency. At the next level the band is lowered by a factor of two and spans fs/8 to fs/4 and so on. Thus for a display resolution of d pixels/degree, the spatial frequency f, of level L will be

$$f = d.2^{-L} \; cycles/degree \qquad (2\text{-}5)$$

The Symlet wavelet basis function generated was fixed at 384x384 size and were embedded in both flat field with a mean gray level of 128 and white Gaussian noise with -23dB contrast level as described in Eq. 2.3. The wavelet stimulus from the first level is not used because the stimuli from levels 2 through 4 cover approximately 1.07 – 8.58 CPD of the spatial frequency range. In our case, Level 4 of the wavelet stimulus consists of the octave band from [1.07 - 2.14]cpd , Level-3 consists of the band [2.14 - 4.29]cpd while Level-2 consists of the band ranging from [4.29 – 8.58] cpd. To ensure consistency

with this, all the spectral energy belonging to first level of the defect patterns were zeroed out through a filtering procedure.

2.4.2 Sinusoidal Stimuli Description and Generation

The stimuli for the Sinusoidal basis function were sine wave gratings added to a constant luminance background or white noise background oriented along horizontal and vertical directions given by

$$y = L + a * \sin(2 * \pi * f * t + \phi) \qquad (2\text{-}6)$$

Where L is the Luminance value and f is the frequency of the given sinusoid.

Figure 2.7:a) Sinusoidal basis function with horizontal orientation at 2.9 cpd in flat field

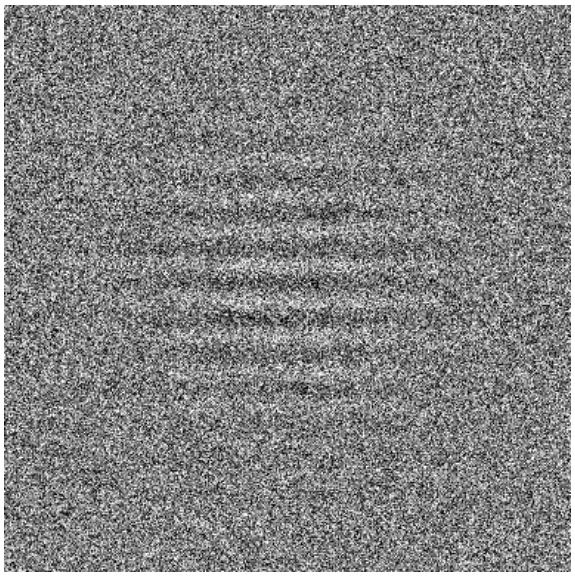

Figure 2.7:b) Sinusoidal basis function with horizontal orientation at 2.9 cpd in random field

Figure 2.7 Sinusoidal Basis functions under different backgrounds

The frequencies used for sinusoidal patterns for the subjective testing were [1.43, 2.87, 4.31, 5.75, 7.19, 8.62 and 9.34] CPD. To reduce the effect of abrupt changes along the edges of the 384x384 image field, a Gaussian window with standard deviation equal to one fourth of the image size weighted the sinusoidal patterns. Figures 2.7(a) and 2.7(b)

show examples of sinusoidal pattern in flat and random field backgrounds. The sinusoids have a highly localized form in the frequency domain as their Fourier decomposition would result in the most localized coefficients, and poor localization in space domain.

2.4.3 Banding defect Stimuli Generation and Description

Banding defects in print images are defined as one dimensional fluctuation in darkness [quote ISO/IEC DIS13660] characterized by Discrete Fourier Transform (DFT) of the luminance line perpendicular to the bands. The banding print defect was generated from a scanned banding profile presented in Cui et.al [22] as shown in Figure 2.8(a) Interpolation was used to stretch or compress this pattern to simulate banding at different frequencies. To simulate the effect of local banding artifacts, two different sizes of the banding pattern over regions of 2.76 and 1.84 visual degrees were generated. Also to

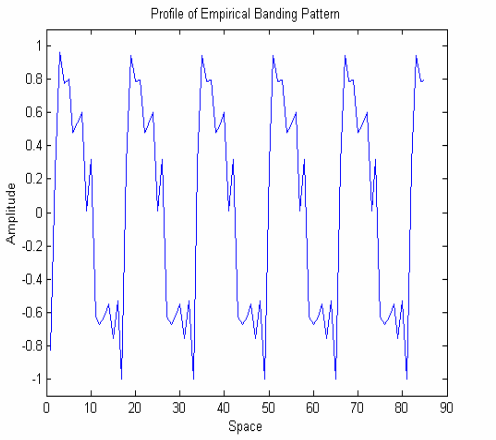

Figure 2.8:a) Measured banding profile

Figure 2.8:b) Simulated banding artifact in flat field occupying 2.76 degrees

Figure 2.8 Simulated local banding artifact

reduce the abrupt changes along the edges they were weighted by a Gaussian window whose standard deviations were correspondingly set to 0.69 and 0.46 degrees, one fourth

of the size of the local image size. These patterns were positioned at the center of an image of size 384x384 pixels (about 11 visual degrees) as shown in the figure 2.8(b). Three different frequencies of the banding pattern, 2.87, 5.75 and 8.62 cpd were generated for each size. The size and frequency of the banding patterns were chosen to be consistent with the wavelet basis functions used, such that almost all banding pattern spectral energy was distributed over the levels 2 through 4 of 4-level wavelet decomposition. This criterion is ensured by subjecting the simulated defect patterns to 4-level wavelet decomposition and zeroing out the coefficients in Level-1 and computing the inverse DWT to obtain a modified empirical banding pattern. This process created minor distortion of the high frequency defect patterns, by filtering out higher order harmonics and making the pattern more sinusoidal that the profile shown.

2.4.4 Grain defect Stimuli Generation and Description

Grain defects or graininess in print images are as perceived 2-D random fluctuations in blackness generally at frequencies greater than or equal to 1 cycle/mm [32] characterized by the standard deviation aperture product. A characterization for the graininess defect is the noise power spectrum (NPS)[20]. The NPS shows a linear spectral pattern on a log scale that is characterized by its spectral slope [21]. In other words the spectral roll-off characterizes the coarseness of the graininess pattern. The steeper the roll-off, the coarser the graininess pattern in the spatial domain. The graininess pattern was simulated by passing white Gaussian noise through a filter with a spectrum similar to what has been reported for measured data [20].A spectrum with a linear roll-off (on a log-linear scale) in both the horizontal and vertical direction was used with a negative slope of -0.17 dB/(cycles/mm). To simulate the effect of local grain artifacts, two different sizes of the

grain pattern over regions of 2.76 and 3.68 visual degrees were generated. To reduce the abrupt changes along the edges along the edges, the grain pattern is scaled by means of a modified Gaussian weighting function. The weighting function is defined by a Gaussian distribution with standard deviation equal to one fourth of the image size but with unit value for all distributed values when within a standard deviation from either side of the mean. The standard deviation of the Gaussian window was correspondingly set to 0.69 and 0.92 degrees. These patterns were positioned at the center of an image of size 384x384 pixels (about 11 visual degrees) as shown in the figure 2.9. Three different slopes were used to created 3 grain patterns of different coarseness for each image size.

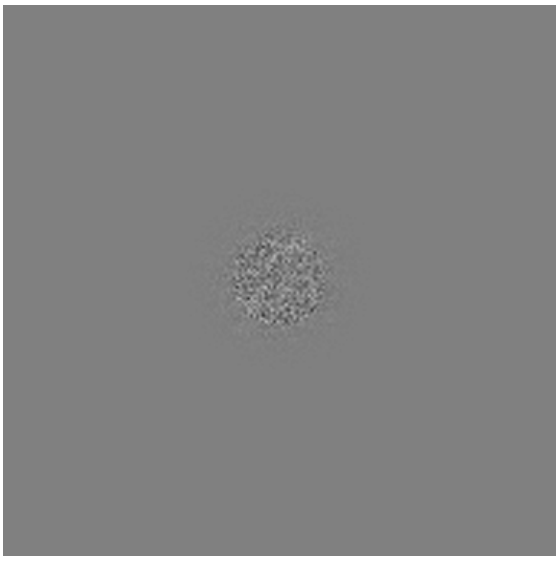

Figure 2.9:a) Simulated grain artifact in flat field occupying 3.68 degrees

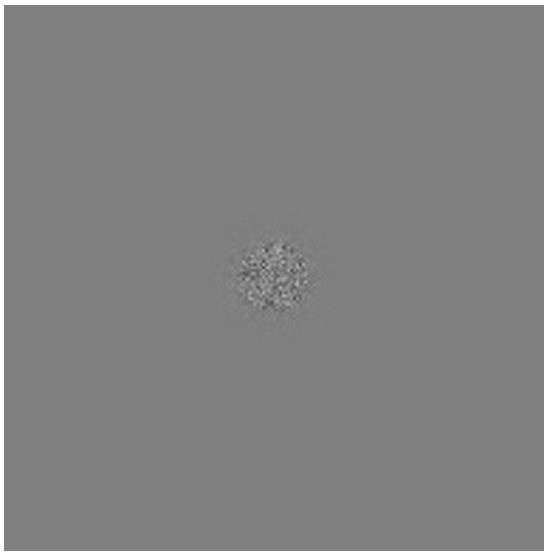

Figure 2.9:b) Simulated grain artifact in flat field occupying 2.76 degrees

Figure 2.9 Simulated Grain artifacts of 2 different sizes

To ensure that the spectral energy of the grain patterns were distributed within levels 2 through 4, they were subjected to 4-level wavelet decomposition and the energy at the first-level was filtered out.

2.5 Effect of change in viewing distance on the frequency content of the stimuli

Even though the viewing distance was kept constant throughout the subjective testing there may have been instances where the subjects did not maintain the specified distance by a small quantity. A sensitivity analysis was carried out analyzing the effects of change in viewing distance on the frequency components of the images displayed since the frequency content of the stimuli displayed is inversely proportional to the viewing distance.

The physical dimensions of the stimuli displayed of size 384x384 occupied an area of 11.13 cm x 11.13 cm in the display system. In the experiment the viewing distance was maintained at 57cms and frequencies of the stimuli were calculated based on this viewing distance. At the above specified condition a value of 1° visual angle is given by

57 *Tan(1°) = 0.9978 cm. An analysis was performed by varying the viewing distance to 57±10cms (even though the change in viewing distance was typically less than this magnitude) and the subsequent change in frequency of the stimuli is presented.

The unit of CPD is obtained from the formula

$$Cycles/Degree = \frac{Cycles}{Cms} * \frac{Cms}{Degree} \qquad (2\text{-}7)$$

The sinusoidal stimuli used for the experiments were [16, 32, 48, 64, 80, 96, 104] cycles occupying 11.13 cm x 11.13 cm viewing area. For example the frequency of the sinusoid of 16 cycles in the given field of 11.13 cms x 11.13 cms at 57 cms viewing distance in CPD is given by

$$CPD = \frac{1.4376 Cycles}{Cms} * \frac{0.9948 Cms}{Degree} = 1.43 CPD \qquad (2\text{-}8)$$

In terms of percentage change, a change in viewing distance by ±X% produces the same ±X change in the frequency content of the signal.

For example for a sinusoidal stimuli at 57 cms which has frequency of 2.86 cpd ,a change in viewing distance by -10 cms to 47 cms, approximately 17%, reduces the frequency of the stimulus to 2.35 cpd, the percentage of reduction being 17%.

The change in frequency for the sinusoidal stimuli is given below at the corresponding viewing distance is given below in Table 2.1.

Table 2.1: Effect of change in viewing distance on frequency of Sinusoidal stimuli

Freq. of Sine in 384x384 field.	At 47 cms, frequency in CPD	At 57 cms, frequency in CPD	At 67 cms, frequency in CPD
16	1.17	1.43	1.68
32	2.35	2.86	3.36
48	3.53	4.29	5.04
64	4.71	5.72	6.72
80	5.89	7.14	8.40
96	7.07	8.58	10.08
104	7.67	9.29	10.92

The same analysis is carried out for the wavelet stimuli by observing the band of frequencies at various levels of decomposition. At the first level of decomposition the wavelet band extends from 0-fs/4 in the approximations and fs/4 to fs/2 in the details and follows octave based decomposition for subsequent levels. The octave frequency bands associated with the corresponding decomposition levels for the wavelet stimuli at different viewing distances is provided below.

Table 2.2: Effect of change in viewing distance on the frequency of wavelet stimuli

	At 47 cm, freq in CPD		**At 57 cm, freq in CPD**		**At 67 cm, freq in CPD**	
Level 1	0-fs/4	fs/4-fs/2	0-fs/4	fs/4-fs/2	0-fs/4	fs/4-fs/2
	0 - 7.07	7.07-14.15	0 - 8.58	8.58-17.16	0 - 10.08	10.08-20.1
Level 2	0-fs/8	fs/8-fs/4	0-fs/8	fs/8-fs/4	0-fs/8	fs/8-fs/4
	0-3.53	3.53-7.07	0-4.29	4.29-8.58	0-5.04	5.04-10.08
Level 3	0-fs/16	fs/16-fs/8	0-fs/16	fs/16-fs/8	0-fs/16	fs/16-fs/8
	0-1.76	1.76-3.53	0-2.14	2.14-4.29	0-2.52	2.52-5.04
Level 4	0-fs/32	fs/32-fs/16	0-fs/32	fs/32-fs/16	0-fs/32	fs/32-fs/16
	0-0.88	0.88 -1.76	0-1.07	1.07-2.14	0-1.26	1.26-2.52

The inference made from this analysis is that, even if there were changes of the order of ±8 cms which corresponds to approximately 14% of the original viewing distance of the observer in the experiment, the frequency content of the stimuli were placed within

designated bands calculated based on the 57 cms viewing distance. In the case of the sinusoidal stimuli the impact of the viewing distance also has the same effect when grouping into octave bands as was done while doing the prediction procedure.

2.6 Contrast Computations of the Stimuli

The stimuli that are presented to the observers are represented in terms of their contrast values. The detection thresholds for the stimuli estimated are based on the contrast measures that are defined on the stimuli. The contrast sensitivity function describes the relationship between the contrast perception and the spatial frequency expressed in CPD and in this study these are represented in terms of detection thresholds. Generally speaking, contrast is a measure of the luminance variation relative to the average luminance in the surrounding region. The definition of a proper contrast measure is difficult over an entire range of image types, but for simple stimuli the Michelson's contrast is used.

The contrast values for all stimuli in these tests excepting the grain pattern were recorded and are represented in decibels as

$$c = 20 \log_{10}\left(\frac{\max(L(x,y)) - \min(L(x,y))}{L_{bg}}\right) \qquad (2\text{-}9)$$

where $L(x, y)$ is the luminance of the pixel values over the x-y plane and L_{bg} is the mean background luminance. The contrast was computed globally over the whole image and recorded with the subjects' response. This value represents the maximum contrast value of the image.

For grain patterns this measure cannot be used as they do not possess simple structure and are random in nature. Hence a contrast measure which characterizes the nature of the grain pattern based on the energy is used. The contrast measure of the grain pattern is defined as

$$C = 20 \log_{10} \left(\frac{\left[\sum_{x,y} (L(x,y))^2 \right]^{1/2}}{\pi r^2} \right) \qquad (2\text{-}10)$$

where *L(x, y)* is the luminance of the pixel values over the x-y plane and r encompasses the image field within the one standard deviation area, of the modified Gaussian function used to weight the grain pattern.

CHAPTER 3

Prediction and Analysis

This chapter discusses the prediction process involved in computing HVS detection thresholds, from the thresholds estimated through the subjective testing procedure. The two different prediction procedures, one based on a Wavelet based decomposition and the other based on Fourier/Sinusoidal based decomposition are explained. Section 3.1 discusses the prediction of the HVS response from psychometric functions and threshold estimation procedure from the results of the subjective testing procedure. Section 3.2 discusses the prediction procedure based on a wavelet based approach and Section 3.3 explains the prediction procedure based on a sinusoidal based approach for the banding defect. The final Section 3.4, presents the prediction process based on wavelet approach for the grain defect.

3.1 Prediction of HVS response

A psychometric function [30], describes the relation between a physical measure of a stimulus, usually the intensity, and the probability of a particular psychophysical response. Threshold characterization using psychometric functions is used so that the detection probabilities can be obtained for any contrast value and combined together in a prediction model to compute the probability of the pattern being detected. The psychometric function has the same shape under all conditions when expressed as a function of log intensity [30]. Under different conditions it differs only in position along the log intensity axis. This relation allows us to describe any particular psychometric

function $P_T(x)$, characterized by a threshold T, in terms of a canonical form $\Psi(x)$, by the relation

$$P_T(x) = \Psi(x - T) \qquad (3\text{-}1)$$

where x is the log intensity of the stimulus and the parameter T may be chosen as any convenient point in the function Ψ.

In our study we use the Weibull psychometric function and 92% point is chosen as the threshold T. Therefore the detection probability for i^{th} basis function estimated from the QUEST method can be substituted into the associated psychometric function of a particular stimulus as

$$P_i(c;T_i) = 1 - (1-\gamma) * \exp(-10^{\frac{\beta}{20}(c-T)}) \qquad (3\text{-}2)$$

where c is the contrast value of the basis function, and β is the parameter which specifies the slope of the psychometric function. The parameter γ specifies the probability of pattern being detected at zero intensity. For 2AFC tests the value is 0.5 since the subject is forced to make a decision between two choices. For 2AFC tests, having γ=0.5 and choosing β=3.5 yields threshold intensities with the probability of success being 0.92 [30]. All the parameter values except γ were expressed in decibels. The psychometric function has a 0.5 probability value when the pattern is not detectable and therefore varies from 0.5 to 1 with increasing contrast values. Various options were available to modify the psychometric function. In our method, in order to combine these probabilities from the contrast values of the decomposed image components, low contrast value corresponding to probabilities under 0.6 were scaled to zero. Therefore the shape of the psychometric function was maintained for probability values greater than 0.6 and were

linearly distorted down to 0 for probability values lesser than 0.6. The modified psychometric is then given as

$$P_i(c;T_i) = 1-(1-\gamma)*\exp(-10^{\frac{\beta}{20}(c-T)}) \text{ for } \mathrm{P_i}(c;T_i) > 0.6 \qquad (3\text{-}3)$$
$$= \left(\frac{P_i(c;T_i)-0.5}{0.1}\right) \text{for } 0.5 > \mathrm{P_i}(c;T_i) < 0.6$$

3.1.1 Final Threshold computation

The thresholds of the basis functions from the subjective testing are computed to create the psychometric functions which serve as a mapping tool between contrast values and detection probability as explained in Section 2.1.

The thresholds for the entire set of basis functions and the defect patterns were computed by two methods. In the first method the median likelihood values of the thresholds were found via a bootstrapping method to limit the effect of outliers and to obtain variability measure of the estimate. Bootstrapping is a resampling technique that assumes that the observations are independent and performs sampling with replacement from the acquired data where each observation has the same probability of being chosen each time [34]. Hence a bootstrap sample from an original sample of size *n* consists of randomly choosing one of the *n* values and repeating this *k* times, putting the chosen value back into the pool each time. The desired statistic is calculated from this bootstrap sample and the entire procedure is repeated *k* times. This ensures that each bootstrap sample is independent. In our analysis, for every stimulus, each bootstrap sample consisted of a set of threshold values, estimated through the QUEST procedure, drawn 64 times at random with replacement. For each drawing the median of the thresholds was taken as the

estimate. Then from all the 64 estimates, the median, upper quartile and lower quartile values were taken from the set. The median of those values was considered as the final estimate and was used in the psychometric function. The inter-quartile range was used as a measure of the variability. The bootstrapping procedure was done for each of the basis functions, wavelets and sinusoids, the grain defect pattern by drawing 8 samples at a time with replacement since 8 subjects were involved and for banding defect pattern, by drawing 9 samples at a time with replacement since 9 subjects were involved.

The median was taken as the estimate because of the presence of the outlier data in the subjective testing results, due to some users correctly speculating the presence of patterns in all the trials. To avoid this difficulty a second method was used in estimating the thresholds, in which the mean of the subjective thresholds were computed by excluding the outlier data from the subjective testing procedure. The advantage of computing the thresholds directly from the mean of subjective data is two fold. The first one is that the estimating mean directly from the data allows us to specify the 95% confidence limits on the threshold values which is a significant statistic measure. The second one is that it gives us an option of comparing it with the thresholds obtained from the bootstrapped median procedure which is a resampling procedure as opposed to the direct threshold computation from the subjective data. In this method the 95% confidence limits on the prediction threshold were used as a measure of variability.

3.2 Prediction of HVS response for banding defect based on a Wavelet approach

This section explains the several steps involved in prediction of HVS response for banding defect based on wavelet approach as discussed in Section 2.1.

The banding defect patterns were generated at different contrast levels consistent with the values used in subjective tests. These patterns were subjected to 4-level wavelet decomposition as described in [33], given by

$$y_{uv}^{(l+1)}(n,m) = \sum_i \sum_j y_{11}^{(l)}(2n-i, 2m-j) K_{uv}(i,j) \qquad (3\text{-}4)$$

where level *0* is the original signal and the wavelet kernel, K_{uv}, an 8th order Symlet [24, 33], is composed of 4 different combinations of high and low-pass filters applied in vertical and horizontal orientations. The subscripts on y indicate the orientation of its wavelet kernel and the superscripts represent the coefficients at level-l as described in [24]. Let the wavelet function (high-pass filter) be denoted by $h(i)$ and the scaling function (low-pass filter) be denoted by $g(i)$. Then a class of separable kernels in 2 dimensions for the wavelet decomposition can be defined as

$$\begin{aligned} K_{11}(i,j) &= g(i)g(j), & K_{12}(i,j) &= h(i)g(j), \\ K_{21}(i,j) &= g(i)h(j), & K_{22}(i,j) &= h(i)h(j). \end{aligned} \qquad (3\text{-}5)$$

Subscript 11 refers to the low-pass filter in both the vertical and horizontal directions. In going from one level to the next, y_{11} from the previous level is used, as indicated in Eq. (3.4). The scaling by 2 of the arguments of y_{11} denotes a dyadic sub-sampling that occurs in going from one level to the next. The levels represent octave sub-bands, since sub-sampling scales down the frequency axis of the wavelet kernel by a factor of 2, thereby

reducing its effective cutoff frequencies without changing the kernel coefficients. At the first level of the transform, the band covers the spectral range extending from one half of the Nyquist rate to Nyquist rate. This is equivalent to half of the display pixel rate. At the next level the band is lowered by a factor of two and so on.

3.2.1 Calculation of Contrast from transform coefficient

The next step is the conversion of the decomposed wavelet image coefficients to contrast values. Contrast at each spatial location within each channel is obtained by taking the ratio of the channel image to the base-band image. The mean background of the simulated image, which has the lowest frequency content of the image, is obtained from the y_{11} coefficients at the fourth level, in which each transform coefficient approximates a local DC component for a 16 by 16 pixel area because of the 4-level decomposition on an image of size 384x384 pixels. The pixel difference measures correspond to the y_{12}, y_{21}, and y_{22} coefficients at all lower levels and these are used as the numerator in a contrast ratio. These coefficients in the numerator represent the high-pass results of the wavelet decompositions, which approximate the local pixel differences representing a contrast measure. The wavelet the coefficients from y_{11}^4 are spatially aligned with the high-pass coefficients at the fourth level and lower. This spatial alignment of the coefficients is done by up-sampling and interpolation of the high-pass coefficients and the coefficients from level y_{11}^4 by the help of wavelet reconstruction filters from their respective levels to the Level 0, which equals the size of the original image. The contrast values for the wavelet basis coefficients can be computed from

$$c_{ij}^{(l)}(n,m) = \frac{2y_{ij}^{(l)}(n,m)}{\bar{y}_{11}^{(4)}} \qquad (3\text{-}6)$$

where $\bar{y}_{11}^{(4)}$ are the coefficients from $y_{11}^{(4)}$ expanded to level l through up-sampling and interpolation with the help of the wavelet reconstruction filters. Each level and orientation of the wavelet decomposition represents one of the basis function for which a psychometric function was computed. Therefore, the computed contrast ratio is substituted into the psychometric function to obtain probabilities for all 9 wavelet sub-bands and all spatial coefficients in each sub-band. These are pooled together to obtain the predicted detection probability for simulated defect pattern. For prediction performance, the defect contrast value that results in a predicted 92% detection probability was compared to the contrast ratio obtained through a subjective test on the defect pattern.

3.2.2 Pooling procedure to obtain probability of detection

The probability pooling was done assuming that each basis function contrast value from Eq. (3.6) represented independent visual channels. Therefore, the final probability of detection for a defect contrast level c was computed as the complement of the probability of not detecting a basis function component in any of the channels, which can be denoted by

$$P_d(c) = 1 - \prod_{n=1}^{N} \left(1 - \lambda_n\right) \qquad (3\text{-}7)$$

where N is number of channels and $\lambda(x_n)$ is the channel probability for basis function contrast c_n. The number of channels is 9, same as the number of basis functions used in

the subjective testing and the channel coefficients used in calculating the contrast of the stimuli. The channel contrast for the wavelet exploited the spatial orthogonality of the coefficients at each level and combined the independent probabilities over spatial contrast values:

$$\lambda_n = 1 - \prod_{k=1}^{K} \left(1 - P_n(c_k)\right) \qquad (3\text{-}8)$$

where c_k is the contrast values for the spatial wavelet coefficients corresponding to Eq. (3.6), n is the particular channel (a specific i, j, and l combination), and K is the total number of (spatial) wavelet coefficients at that level and orientation.

3.3 Prediction of HVS response for banding defect based on a Sinusoidal approach

The prediction based on sinusoidal decomposition was performed only on the banding defect as it is characterized by Discrete Fourier Transform (DFT) of the luminance line orthogonal to the banding defect. The sinusoidal based prediction follows the general prediction procedure described previously, but with a modification in the transform domain representation and method of calculating contrast.

In the first stage, banding defect patterns were generated at a given contrast level. A line was extracted perpendicular to the orientation of the banding defect from the center of the image and transformed in to sinusoidal components by a 1-dimensional DFT. This method of characterization of the banding defect is consistent with the definition of the banding defect [32].

3.3.1 Calculation of Contrast from transform coefficient

The contrast ratios were obtained by using the DC value to scale down each DFT component. The transform coefficients were grouped into four octave bands given by [1.07-2.14, 2.14-4.28, 4.28-9.56, 9.56-19.12 CPD] and the magnitude of the coefficients were summed within each octave band and the results were substituted into the psychometric functions obtained for the sinusoidal basis. The energy contrast value is given by

$$c_s(b) = \frac{2}{S(0)} \sum_{k=2^b f_o}^{2^{b+1} f_o} |S(k)| \qquad (3\text{-}9)$$

where *S(k)* are the DFT coefficients with index k ranging over octave bands starting with index of reference frequency f_o and b is an integer ranging from 0 to 3. The sinusoidal basis function corresponding to the center of the octave band was used to obtain probabilities from the psychometric function. The four octave bands were computed corresponding to a similar frequency range as the wavelet sub-bands, except that it included an additional low frequency band. A linear interpolation scheme was used to obtain psychometric functions at the frequencies of the sinusoidal component not tested. This modified way of calculation of detection probability was used over a more direct approach of computing the detection probability by applying each *S(k)* into its corresponding psychometric function and pooling those values as it yielded very poor prediction results with close to 100% under-prediction of HVS results. Thus, the modified method using octave bands described here was determined by assuming that the visual channels were octave based, and the detection of the pattern resulted from the contributions of all contrast energy in that channel. This modification resulted in good

performance for the sinusoidal based method and provided a better basis of comparing with the wavelet based method. The reason for the use of modified method based on octave bands for prediction in case of sinusoids as shown in Eq.3.9 was because of the characteristics of the sinusoidal basis function stimuli used in the subjective testing. The sinusoidal basis function stimuli were oriented along two directions and were not localized by windowing and hence could not provide similar spectral information as provided by the wavelets. The sinusoidal basis functions could not be made a local stimulus by windowing, because the print defects were being represented in terms of the global sinusoidal basis functions as necessitated by representation of the Fourier transform. On the other hand the stimuli representing the banding defect had to be local (implemented by windowing) to efficiently characterize the nature of the defect as it occurs practically.

3.3.2 Pooling procedure to obtain probability of detection

The probability pooling was done based on the assumption that the each basis function contrast value or in other words the sum of Fourier magnitude coefficients over the octave bands represented individual visual channels. As with the wavelet based process the final probability of detection for a defect at contrast level c was computed as the complement of the probability of not detecting a basis function component in any of the channels, denoted by

$$P_d(c) = 1 - \prod_{n=1}^{N} (1 - \lambda_n) \quad (3\text{-}10)$$

Where N is the number of channels and $\lambda(x_n)$ is the channel probability for basis function contrast c_n. The number of channels in this case is equal to 4 as there were four sub-bands used in the sinusoidal prediction procedure. The sinusoidal approach does not include spatial components and hence the channel probability $\lambda(x_n)$ is given as

$$\lambda_n = P_n(c_k) \quad (3\text{-}11)$$

where c_k is the contrast value from Eq. (3.8), and the psychometric function corresponds to the sinusoid at the center of the octave band.

3.4 Prediction of HVS response for Grain defect based on a Wavelet approach

The previous sections described the prediction of HVS response for banding defect based on a wavelet and sinusoidal method. This section explains the process of prediction of HVS response to grain defects, generated in two different sizes with three differing coarseness levels. The prediction of HVS response for grain pattern was carried out by a wavelet based procedure only. One main reason for that is the sinusoidal component analysis is more suitable for analyzing banding defect as it is characterized by FFT of the luminance line orthogonal to the banding defect and also due to the fact that the random structure exhibited by the grain defect does not provide any useful insight with the similar procedure. The other reason is that the subjective testing with the sinusoidal basis functions had stimuli oriented along two directions only and did not include any localized sinusoidal patterns. Hence there was not similar sufficient spectral information to cover

the entire spatial frequency range when compared with the information obtained from the subjective tests based on the wavelet basis functions. Therefore the grain patterns were analyzed only by a wavelet based prediction procedure and the results are compared for the two differing sizes of the grain pattern. The prediction of HVS response for grain pattern is similar to the procedure followed for the banding pattern. In the first stage the grain defects were generated with varying contrast levels by increasing the amplitude of the grain pattern and are subjected to 4-level wavelet decomposition.

The second step is converting the coefficients to a contrast measure, consistent with the definition of the contrast for the grain defect pattern. The contrast of the grain pattern from the wavelet coefficients are computed as

$$c_{ij}^{(l)}(n,m) = \frac{\sqrt[2]{\sum_{i,j} y_{ij}^{(l)}(n,m)}}{\bar{y}_{11}^{(4)}} \qquad (3\text{-}12)$$

The contrast calculation, detection probability and the pooling procedure were the same as followed in the banding defect pattern analysis.

CHAPTER 4

Results and Discussion

This chapter presents and discusses the results obtained from the subjective testing and the predicted thresholds obtained from the wavelet based and sinusoid based prediction methods. In the Section 4.1 the subjective thresholds obtained using a median operation to censor the effects of outliers is presented along with the predicted thresholds using the wavelet and sine based methods for the banding defect. The discussion of these results is presented in Section 4.2. The next section, Section 4.3, provides the results of the subjective thresholds obtained using a mean operation (after censoring the outliers) and the predicted thresholds using the two methods for the banding defect. The following Section 4.4 contains the discussion about these results. Section 4.5 deals with the grain defect, and presents results of the estimated subjective thresholds computed using mean and median operations and the prediction on the based on wavelet method. The final Section 4.6 is discusses the results of the grain pattern.

4.1 Comparison of Median subjective thresholds with the predicted thresholds for banding defect

The results of the median subjective thresholds, obtained as explained in Section 3.1, on the defect banding patterns are presented in Table 4.1, which include both defect sizes on

the flat field background at luminance level 20 cd/m^2 and random background with mean luminance 20 cd/m^2 and -23 dB of white noise. The final median subjective threshold computation is explained in Section 3.1.1

Table 4.1: 92% Median subjective detection thresholds in dB for subjective defect tests

Subjective median Values	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-35.7	-37.5	-31.5
Small banding in flat field	-34.0	-34.0	-30.5
Large banding in noise field	-34.0	-33.0	-31.5
Small banding in noise field	-32.7	-33.0	-28.0

The predicted values using the wavelet bases are shown in Table 4.2 along with the percent error relative to the median thresholds in Table 4.2.

Table 4.2: 92% detection Threshold in dB for predicted HVS response using wavelet bases with percent error relative to the median subjective thresholds in parenthesis.

Predicted values by Wavelet Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-40.1 (12%)	-35.5 (–5%)	-34.8 (10%)
Small banding in flat field	-40.0 (18%)	-33.5 (-1%)	-32.9 (8%)
Large banding in noise field	-35.5 (4%)	-32.5 (-2%)	-31.3 (-1%)
Small banding in noise field	-35.0 (7%)	-30.4 (-8%)	-29.9 (7%)

A systematic error of -1.33dB exists between the predicted and subjective thresholds. The predicted values using the wavelet bases after the removal of the systematic error is given below in Table 4.3

Table 4.3: 92% Detection Threshold in dB for predicted HVS response using wavelet bases after removing systematic error with percent error relative to the median subjective thresholds in parenthesis

Predicted values by Wavelet Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-38.8 (8%)	-34.2(-9%)	-33.5(6.5%)
Small banding in flat field	-38.7(13%)	-32.3(-5.5%)	-31.6(3.5%)
Large banding in noise field	-34.2(0.5%)	-31.2(-5.5%)	-29.9(-5%)
Small banding in noise field	-33.6(2%)	-29.7(-10%)	-28.6(2%)

The predicted values using the sinusoidal bases are shown in Table 4.4 along with the percent error relative to the thresholds in Table 4.1.

Figure 4.4: 92% Detection Threshold in dB for predicted HVS response using Sinusoidal bases with percent error relative to the median subjective thresholds in parenthesis

Predicted values by Sinusoidal Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-30.2 (-15%)	-30.1 (-20%)	-27.2 (-14%)
Small banding in flat field	-30.1 (-11%)	-30.0 (-12%)	-27.9 (-9%)
Large banding in noise field	-26.5 (-22%)	-25.6 (-22%)	-23.6 (-26%)
Small banding in noise field	-26.2 (-20%)	-25.5 (-23%)	-23.4 (-17%)

A systematic error of 5.75 dB existed between all the predicted and subjective thresholds. If this error is taken out, the then all errors range between -7 and 10 percent. The predicted values using the sinusoidal bases after removing the systematic errors are shown in Table 4.5 along with the percent error relative to the median thresholds in Table 4.1

Table 4.5: 92% Detection Threshold in dB for predicted HVS response using Sinusoidal bases after removing systematic error with percent error relative to the median subjective thresholds in parenthesis

Predicted values by Sinusoidal Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-35.9(1%)	-35.8(-4%)	-32.9(4%)
Small banding in flat field	-35.8(5.5%)	-35.7(5%)	-33.6(10%)
Large banding in noise field	-32.5(-3.5%)	-31.3(-6%)	-29.3(-7%)
Small banding in noise field	-31.9(-2.5%)	-31.5(-5.5%)	-29.5(4%)

4.2 Discussions on comparison between Wavelet and Sine based method for banding defect:

The error on the results for both wavelet and sinusoidal cases is comparable to the inter quartile ranges on the subjective defect tests (up to 12%) and the basis function tests (up to 11%). So it is realistic to conclude that the prediction performance was reasonable. The summing of energy in the octave bands for the sinusoidal based technique suggests that a Gabor patch should be used next time in a comparison with a wavelet based method as both of them share a great similarity in terms of their spectral characteristics. The contrast measures in the subjective tests were computed globally over a whole image

using the definition Eq.2.9, but the contrast computations from basis functions in the prediction process as explained in Section 3.3, characterize local pixels of the image. Hence the large systematic errors in sinusoidal based method are likely due to this inconsistency in scale of the contrast value computations. Once this error was removed the prediction performance was similar to that of the wavelet approach. The effect of the systematic errors on the performance of predictions has to be investigated in future research efforts.

The large errors for the low-frequency flat-field results are likely due the limited number for gray levels below the thresholds. The low frequency patterns for the wavelet bases were the easiest to detect and it was difficult to set up the experiment so that most people had several gray levels below which this pattern could not be detected. Some subjects were even able to correctly identify this pattern for all 16 trials. This resulted in significant quantization error for the low frequency wavelet patterns. This reason is supported by the observation that in most cases when white noise was added to the test patterns the prediction error reduced when compared with the results from the flat field background. The white noise had the effect of making the patterns harder to detect (reduced HVS sensitivity) and as a result more gray levels existed below the detection threshold for all subjects, especially for the lower frequency patterns. This effect is shown in Figure 4.1

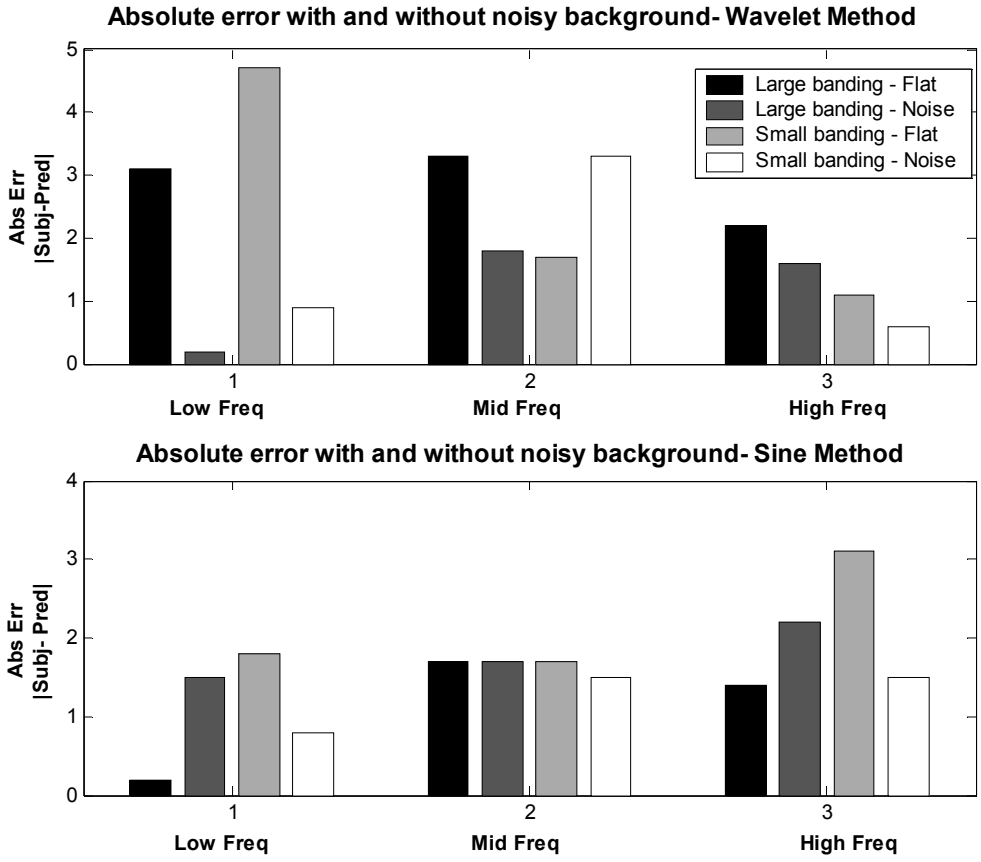

Figure 4.1 Absolute errors in flat field and noisy backgrounds for Median method

Overall both wavelet and sinusoidal based methods worked reasonably well suggesting the feasibility of developing measures and methods for print defect detection that reflect the response of the HVS. In addition, the performance prediction analysis done in these experiments used mostly different populations. There were 1 or 2 subjects common to all three tests. This suggests that the information obtained from one population on the basis functions was applicable to the larger population. It was not simply the case of learning the HVS capability of population and predicting the response of the same population

4.3 Comparison of Mean subjective thresholds with the predicted thresholds for banding defect

The results of the subjective test on the defect banding patterns are presented in Table 4.6, which include both defect sizes on the flat field background at luminance level 20 cd/m^2 and random background with mean luminance 20 cd/m^2 and -23 dB of white noise. The mean subjective thresholds are provided with their 95% confidence limits calculated with the help of the student's *t*-distribution. The level of confidence (precision) desired for the output variable $\bar{x}$ to within δ units at confidence level 1-α, using the following relation:

$$\Pr\left[\left|\bar{x}-\mu\right| < \delta\right] = 1-\alpha \qquad (4\text{-}1)$$

where

$$\delta = \left(t_{1-\alpha/2}, N-1\right)\frac{S}{\sqrt{N}} \qquad (4\text{-}2)$$

where S is the standard deviation and N is the number of trials, in this case the number of the valid subjective responses. The $\left(t_{1-\alpha/2}, N-1\right)$ can be obtained from the student's *t*-distribution table with α value set to 0.05 for 95% confidence intervals.

Table 4.6: 92% mean subjective detection thresholds in dB for defect tests with their confidence intervals

Subjective mean Values	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-34.90±1.20	-36.70±1.20	-31.70±2.30
Small banding in flat field	-32.40±0.90	-34.70±0.80	-29.40±2.30
Large banding in noise field	-33.90±1.70	-34.60±0.80	-29.80±1.50
Small banding in noise field	-32.30±1.50	-33.90±1.10	-28.70±2.0

The predicted values using the wavelet bases are shown in Table 4.7 along with the distance from the subjective mean thresholds in Table 4.6.

Table 4.7: 92% Detection Threshold in dB for predicted HVS response using wavelet bases with distance from subjective mean thresholds in parenthesis

Predicted values by Wavelet Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-40.2 (-5.3)	-37.00(-0.3)	-36.00(-4.3)
Small banding in flat field	-39.90 (-7.5)	-35.90(-1.2)	-35.6 (-6.2)
Large banding in noise field	-36.10 (-1.7)	-32.50 (+2.2)	-32.30 (-2.9)
Small banding in noise field	-35.50 (-3.2)	-31.0 (-2.9)	-30.70 (-2)

A systematic error of -2.7 dB exists between the predicted and subjective thresholds. The predicted values using the wavelet bases after removing the systematic error is shown in Table 4.8 along with the distance from subjective mean thresholds in Table 4.6.

Table 4.8: 92% Detection Threshold in dB for predicted HVS response using wavelet bases after removing systematic errors with distance from the mean subjective thresholds in parenthesis

Predicted values by Wavelet Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-37.5(-2.6)	-34.3(+2.4)	-33.3(-1.6)
Small banding in flat field	-37.2(-4.8)	-33.2(+1.5)	-32.9(-3.5)
Large banding in noise field	-33.4(+0.5)	-29.8(+4.8)	-29.6(+0.2)
Small banding in noise field	-32..8(-0.5)	-28.3(+2.9)	-28.0(+0.7)

The predicted values using the sinusoidal bases are shown in Table 4.9 along with the percent error relative to the thresholds in Table 4.6. A systematic error of 5.9 dB existed between all the predicted and subjective thresholds.

Table 4.9: 92% Detection Threshold in dB for predicted HVS response using Sine bases with distance from the mean subjective thresholds in parenthesis

Predicted values by Sinusoidal Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-30.3(+4.6)	-28.90(+7.8)	-26.10(+5.4)
Small banding in flat field	-29.80(+2.6)	-28.40(+6.3)	-26.10(+3.3)
Large banding in noise field	-26.50(+7.4)	-24.50 (+10.1)	-26.10 (+3.7)
Small banding in noise field	-26.10(+6.2)	-24.20 (+9.7)	-26.10(+2.6)

The predicted values using the sinusoidal bases after removing the systematic error is shown in Table 4.10 along with the percent error relative to the thresholds in Table 4.6.

Table 4.10: Detection Threshold in dB for predicted HVS response using Sine bases after removing the systematic error with distance from mean subjective thresholds in parenthesis

Predicted values by Sinusoidal Method	Low Frequency	Mid Frequency	High Frequency
Large banding in flat field	-36.2(-1.3)	-34.8(+1.9)	-32.0(-0.5)
Small banding in flat field	-35.7(-3.3)	-34.3(+0.4)	-32.0(-2.6)
Large banding in noise field	-32.4(+1.5)	-30.4(+4.2)	-32.0(-2.2)
Small banding in noise field	-32.0(+0.3)	-30.1(+3.8)	-32(-3.3)

4.4 Discussions on comparison of Wavelet and Sine based method for banding defect:

As mentioned previously the large errors for the low-frequency flat-field results are likely due to the limited number for gray levels below the thresholds. This limitation could not be overcome in this subjective testing procedure because of the limits of the display resolution of the display system. The lower frequency patterns of the wavelet and sinusoidal patterns were easier to detect and hence some subjects were even able to correctly identify this pattern for all 16 trials. Those measurements were classified as outliers in the testing and their thresholds were omitted while calculating the mean values upon which the psychometric functions were computed as mentioned in Section 3.1.1. The computed threshold's suffered from a small data set as there was one outlier in the subjective responses of each basis function variation reducing the data set from eight to seven. Hence the estimated thresholds may not be closer to the actual threshold. This

resulted in significant quantization error for the low frequency wavelet patterns. This reason is supported by the observation that in most cases when white noise was added to the test patterns the prediction error reduced .The white noise had the effect of making the patterns harder to detect (reduced HVS sensitivity) and as a result more gray levels existed below the detection threshold for all subjects, especially for the lower frequency patterns. Figure 4.2 shows the absolute errors between the subjective and the predictive procedures (with systematic error removed) in Flat field and noisy background conditions.

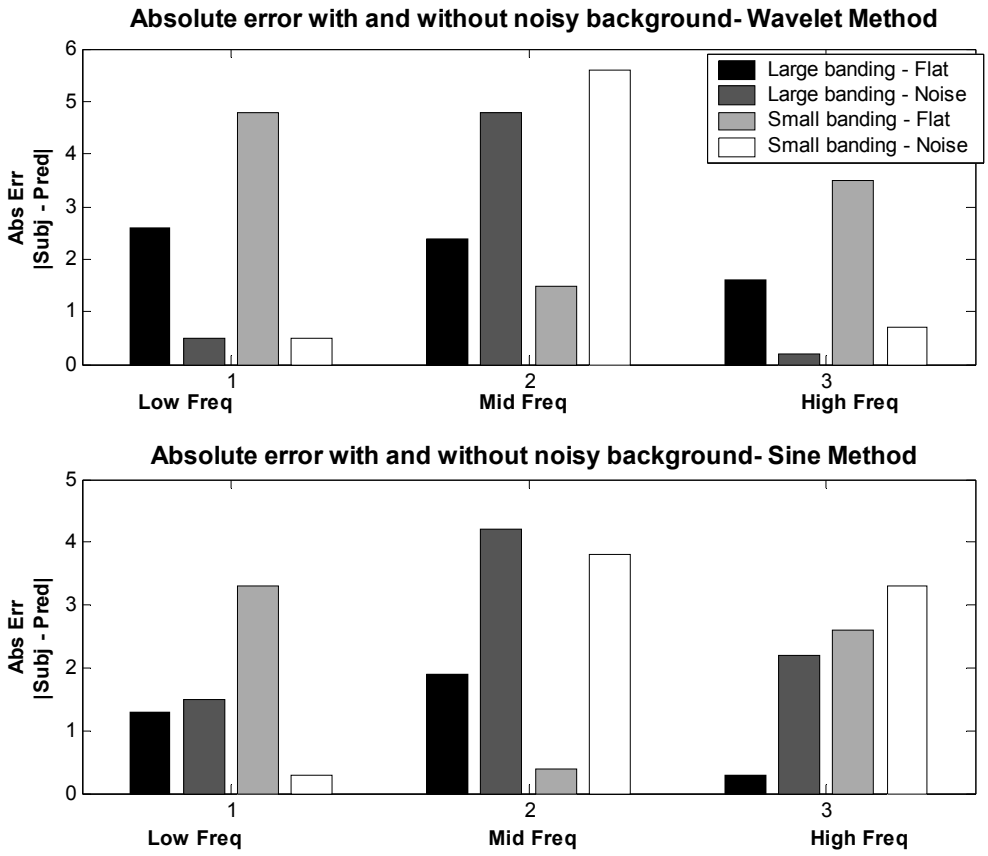

Figure 4.2 Absolute errors in flat field and noisy backgrounds for Mean method

The predicted thresholds based on the wavelet method in many cases falls just outside the 95% confidence interval estimates of the subjective thresholds suggesting that this method works reasonably well to suggest the feasibility of developing measures and methods for print defect evaluation that reflect the response of the HVS. The prediction results for the wavelet and the sinusoidal based methods perform better when the systematic errors are removed. Particularly in sinusoidal based prediction since the systematic errors are large there could be some inconsistencies while calculation of contrast measures in the prediction process which might have resulted in scaling problems. Hence more effort should be focused on understanding the effect of scale factor on the contrast values in future works.

Many sources of variability can be attributed as to why the wavelet based prediction do not fall within the 95% confidence estimates of the thresholds.

1. As mentioned above the lack of sufficient number of gray levels, due to the barrier of display resolution, below the threshold prevents us in estimating the exact thresholds. Because of this problem the estimated thresholds may slightly differ from the actual thresholds and hence the prediction might not fall within the confidence limits. One way to circumvent this problem is to increase the hardware capability to reproduce finer gray level variations of the stimuli in subjective tests.

2. The peak contrast measure may not be a good measure in characterizing the basis function stimuli. The peak contrast definition used here may be a good measure in representing a periodic image pattern such as a sinusoid but for a spatially limited pattern such as a wavelet basis function this contrast measure may not sufficiently characterize its response in the HVS. Different contrast measures such as energy

based measure may be used and compared between wavelet and sinusoidal based prediction models.

3. Assumption of the Weibull PDF for variation around the threshold values. A more empirical PDF, non-parametric in nature, may yield better prediction and more accurate results with pooling methods. The CDF was modified in a linear fashion to aid the prediction process; other methods of modifying the CDF have to be studied. The analysis of the prediction methods with a non-parametric and a parametric CDF should make an interesting study in the future research efforts.

4. Standardization of the subjective procedures could result in prediction of precise contrast thresholds. Standard testing methods as mentioned in ModelFest [28] should yield us better thresholds estimates and also allows us to compare the results of similar research efforts being carried out.

4.5 Results and discussion of Median subjective thresholds with the wavelet based prediction for Grain defect

The results of the 92% detection threshold in dB for the subjective testing using the median and the prediction thresholds by wavelet based method for the grain patterns of both sizes in the flat-field background at 20 cd/m^2 luminance level (Table 4.11).

Table 4.11: 92%Detection Threshold in dB for Subjective testing using the median and predicted HVS response using Wavelet bases for Grain Defect

Big Grain Pattern	**Coarse**	**Medium**	**Fine**
Subjective (median)	-36.5	-36.5	-36.5
Prediction Wavelet	-34.5(-4%)	-35.3(-3%)	-35.5(-3%)
Small grain pattern	**Coarse**	**Medium**	**Fine**
Subjective(median)	-36.5	-34.0	-36.5
Prediction wavelet	-33.6(8%)	-33.6(-1%)	-34.0(-7%)

Systematic error of 1.4 dB for the larger grain pattern and 1.93 dB for smaller grain pattern error existed between the subjective and the predictive thresholds. The results of the prediction of the grain patterns with systematic error removed is given in table 4.12

Table 4.12: 92%Detection Threshold in dB for Subjective testing using the median and predicted HVS response using after removing systematic errors using wavelet bases for Grain Defect

Big Grain Pattern	**Coarse**	**Medium**	**Fine**
Subjective (median)	-36.5	-36.5	-36.5
Prediction Wavelet	-35.9(2%)	-36.7(-1%)	-36.9(1%)
Small grain pattern	**Coarse**	**Medium**	**Fine**
Subjective(median)	-36.5	-34.0	-36.5
Prediction wavelet	-35.5(3%)	-35.9(-5%)	-35.9(2%)

Removing the systematic error limits the prediction error within 2% in case of larger grain pattern and -5% for the smaller grain pattern. The good prediction performance for the grain defect implies that the contrast measure for the grain defect based on the energy model characterizes the defect well. Though the wavelet based prediction performed well in case of the grain pattern, not much variation in threshold values could be obtained with the subjective experiments across the coarseness and size of the grain pattern.

4.6 Results and discussion of Mean subjective thresholds with the wavelet based prediction for Grain defect

The results of the 92% detection threshold in dB for the subjective testing using the mean with their 95% confidence limits and the prediction thresholds by wavelet based procedure along with the distance from the mean subjective thresholds for the grain patterns of both sizes in the flat field background at 20 cd/m^2 luminance level are presented in Table 4.13

Table 4.13: 92%Detection Threshold in dB for Subjective testing using the mean and predicted HVS response using Wavelet bases along with the distance from the mean subjective thresholds in parenthesis for grain defect.

Big Grain Pattern	**Coarse**	**Medium**	**Fine**
Subjective (mean)	-35.20±0.90	-35.70±0.70	-36.80±0.90
Prediction Wavelet	-36.20(-1)	-36.30(-0.6)	-36.30(0.5)
Small grain pattern	**Coarse**	**Medium**	**Fine**
Subjective(mean)	-36.50±1.0	-34.40±0.70	-36.80±1.00
Prediction wavelet	-34.30(-2.2)	-35.10(-0.7)	-35.00(1.8)

The performance of the wavelet based prediction was within the 95% confidence limits for all the case except the coarse grain pattern of smaller size falling short by about -1dB.

A systematic error of -0.3 dB and -0.4 dB existed for the large and small grain pattern. The results of the wavelet based prediction after removing the systematic error is presented along with the distance from the mean subjective thresholds in Table 4.14.

Table 4.14: 92% Mean detection Threshold in dB for Subjective testing and predicted HVS response using Wavelet bases after removing systematic errors along with the distance from the mean subjective thresholds in parenthesis.

Big Grain Pattern	**Coarse**	**Medium**	**Fine**
Subjective (mean)	-35.20±0.90	-35.70±0.70	-36.80±0.90
Prediction Wavelet	-35.9(-0.7)	-36.0(-0.3)	-36.0(0.8)
Small grain pattern	**Coarse**	**Medium**	**Fine**
Subjective(mean)	-36.50±1.0	-34.40±0.70	-36.80±1.00
Prediction wavelet	-33.9(-2.6)	-34.80(-0.4)	-35.4(1.8)

The good performance of the wavelet based prediction can again be attributed to the superior characterization of the grain pattern by an energy based contrast measure. The prediction performance suggests that the spatial pooling of orthogonal translation may follow the response of the HVS to the pattern size. However the prediction results for the smaller grain pattern are not within the 95% confidence limits suggesting that greater variations in coarseness and size must be examined along with the contrast scaling issues.

CHAPTER 5

Conclusions and Future Work

The chapter focuses on the contributions made and the inferences that can be derived based on the results of the subjective testing experiments on basis function and the simulated print defects. The outcomes of the experiments and a conclusion of the prediction performance of the wavelet based and the sinusoidal based method are discussed on section 5.1. The suggestions for performance improvements that can be incorporated to obtain more accurate results are presented in the following section. The next section, Section 5.3 describes the possible future research directions that could be followed in developing comprehensive standards for print quality evaluations.

5.1 Summary

The objective of this work was to explore the feasibility of predicting the perception or visibility of print defect patterns, banding and graininess, under a flat field and noisy background conditions by a wavelet-based and sinusoidal-based approach and to compare their performance. A novel method based on representing the print defects in terms of wavelet transform and a modified octave based sinusoidal method was introduced. Based on the performance one can conclude that the wavelet based method of prediction worked reasonably well when compared with the sinusoidal method, and hence can be an useful technique in developing measures and methods of print quality evaluations based on the HVS response.

Several different methods of pooling, using the single largest contrast value in the band limited contrast measure (equivalent to Minkowski's pooling with exponent of ∞), by using top 5 contrast values in the band limited contrast were tried in the prediction process. But the one reported, which uses all the contrast values in the contrast measure proved to be the better performer in the wavelet based prediction process.

The orthogonal wavelets have distinct advantages in developing models for probability pooling based on independence of the visual channels from the HVS model. In addition, orthogonality ensures the pattern energy was not over counted or under counted in each wavelet frequency or spatial band (this is not possible with the Gabor patch). The systematic error for the wavelet approach was within the inter-quartile distance of the error in case of the median based thresholds and just missed the 95% confidence intervals with the mean based thresholds under the noisy background conditions. The performance in the flat field background was hampered by the lack of gray levels below the actual thresholds and can be improved by using hardware with better display resolution. This suggests that the model for pooling probabilities and scaling of contrast values followed that of the HVS, at least better than the sinusoidal approach with no spatial component.

5.2 Suggestions for performance improvements

Adopting standardized subjective testing procedure by following the experiment conditions as suggested in ModelFest [30] should yield accurate estimates of thresholds and also allows the comparison of the results with a communal set of data.

Definite improvements may be achieved by more direct computation of detection probabilities over a range of contrast values between the 1% and 99% detection levels

from an empirical CDF collected through user inputs. The psychometric functions used to obtain intermediate probabilities from the QUEST procedure may not be very accurate for values far from the 92% detection threshold. Also 84% detection threshold measure instead of 92% detection threshold as suggested by recent efforts [29, 30] should provide us with a more stable measure of detection thresholds. The wavelet approach may have been relatively robust to this error because of its ability to capture the defect pattern energy in a few coefficients. In the future experiments better prediction and more accurate assessment of pooling methods may be obtained from non-parametric distribution functions for HVS detection probabilities. This may further improve the consistency of the performance and be more consistent with models for combining channel probabilities.

Use of different types of contrast measures for characterizing the stimuli and a prediction method consistent with the contrast definition used should provide more insight regarding the proper characterization of the stimuli.

5.3 Future Work

More interesting and insightful comparisons should be made between methods using Gabor patches and wavelets. Both methods are similar in that they can reflect spatial and frequency localization. As a matter of fact the Gabor patch is a wavelet; however it cannot operate as a tight set of orthogonal filters as can be done with the symlet. Future comparisons between these basis functions should allow for a more direct comparison of methods, and should indicate the advantage of orthogonality in predicting HVS response along with the impact on the shape of the basis function.

The wavelet based model shows more promise as they generate the orthogonal decomposition filters in a more systematic way. Some issues regarding the workings of human visual system itself have not yet been resolved, nevertheless these investigations based on subjective testing will go a long way in establishing a complete HVS model incorporating the higher and lower level vision processes that will have a tremendous impact in the larger context of perceptual image quality.

.

REFERENCES

1. Brian Keenan, Handbook of Image Quality, Marcel Dekker, Inc., 2002.
2. B.Girod, "What's wrong with Mean-squared Error", Digital Images and Human Vision, A. B. Watson (Ed.), Chapter 15, MIT Press, 1993, pg. 207-220.
3. B. A. Wandell, Foundations of Vision, Sinauer Associates, Inc., 1995.
4. A. B. Poirson, and B. A. Wandell, "Appearance of colored patterns: pattern-color separability", Journal of Optical society America A, vol. 10, no. 12, Dec. 1993, pg. 2458-2470.
5. A. B. Poirson, and B. A. Wandell, "Pattern-color separable pathways predict sensitivity to simple colored patterns", Vision research, vol.36, no. 4, 1996, pg. 515-526.
6. C. M. Bird, G. B. Hemming, F. A. Wichmann, "Contrast discrimination with sinusoidal gratings of different spatial frequency", Journal of Optical Society of America, vol.19, no.7, July 2002, pg. 1267-1273.
7. A. B. Watson, "The cortex transform: Rapid computation of simulated neural images", Computer Vision, Graphics and Image Processing, vol. 39, no. 3, September 1987, pg. 311-327.
8. S. Daly, "The visible differences predictor: An algorithm for the assessment of image fidelity", Digital Images and Human Vision, A. B. Watson (Ed.), Chapter 14, MIT Press, 1993, pg. 179-205.

9. D. J. Heeger and P. C. Teo, “A model of perceptual image fidelity”, Proceedings of IEEE International Conf. on Image Proc. Oct. 23 – 26 1995, Washington, D.C., USA, pg. 343-345.
10. J Lubin, “A visual discrimination model for imaging systems design and evaluation”, Vision models for Target Detection and Recognition, E. Peli (Ed.), Singapore: World Scientific, 1995, pg. 245 – 283.
11. S. J. P. Westen, R. L. Lagendijk, and J. Biemond, “Perceptual image quality based on a multiple channel HVS model”, Proceedings of IEEE International Conf. on Acoustics, Speech and Signal Processing, 1995, pg. 2351-2354.
12. C.C. Taylor, Z. Pizlo, J. P. Allebach, and C .A. Bouman, "Image Quality Assessment with a Gabor Pyramid Model of the Human Visual System" , Proceedings of the 1997 IS&T/SPIE International Symposium on Electronic Imaging Science and Technology, San Jose, CA, February 1997, Vol. 3016, pg. 58-69.
13. N. Jayant, J. Johnson, and R. Safranek, “Signal compression based on model of human perception”, Proceedings of the IEEE, vol. 81, Oct. 1993, pp. 1385-1422.
14. A. B. Watson, “DCT quantization matrices visually optimized for individual images”, Human Vision, Visual Processing, and Digital Display IX, Proc. SPIE, 3644, 1999, 168-174.
15. Weber E H, *De* Pulsu, Resorptione, Auditu et Tactu (Leipzig: Koehler), 1834.
16. J. L. Mannos, D. J. Sakrison, “The Effects of a Visual Fidelity Criterion on the Encoding of Images”', IEEE Transactions on Information Theory , Vol. 20, No 4, 1974, pg. 525-535.

17. G.E.Legge and J.M.Foley, “Contrast masking in human vision”, JOSA, Vol-70, No. 12, December, 1980, pg. 1458-1471.
18. A. B. Watson, G.Y.Tangand, J.A.Solomon and J.Villasenor, “Visibility of Wavelet quantization noise”, IEEE Transactions on Image Processing, Vol.6, 1997, pg. 1164-1175.
19. John C. Briggs, Mike Murphy, and Yichuan Pan, “Banding Characterization for Inkjet Printing”, IS&T, Portland, OR, 2000, pg. 84-88.
20. Paul J. Kane, Theodore F. Bouk, Peter D. Burns, and Andrew D. Thompson, “Quantification of Banding, Streaking and Grain in Flat Field Images”, IS&T, Portland, OR, 2000, pg. 79-83.
21. Howard Mizes, Nancy Goodman, and Paul Butterfield, “The Perceptibility of Random Streaking”, IS&T, Portland, OR, 2000, pg. 89-93.
22. Chengwu Cui, Dingcai Cao, and Shaun Love, “Measuring Visual Threshold of Inkjet Banding”, IS&T, Montreal, Canada, 2001, pg. 84-89.
23. Michael L.Brill, Jeffrey Lubin, “Perceptual scaling of quality metrics for hardcopy image evaluation”, International conference on Digital Printing Technologies, ICIP 1999, pg. 435 – 438.
24. K.D.Donohue, C. Cui, and M.Vijay Venkatesh, “Wavelet Analysis of Print Defects”, IS&T, Portland, OR, 2002, pg. 42-47.
25. Wenchung Wu, Z. Pizlo, and J.P. Allebach, “Color Image Fidelity Assessor”, Proceedings of the 2001 IS&T Image Processing, Image Quality, Image Capture, and Systems Conference, Montreal, April, 2001, pg. 148-151.

26. P. Teo and D.Heeger, “Perceptual image distortion”, IEEE conference on Image Processing, Nov 1994, pg. 982-986.
27. C.J.V. Branden Lambrecht**,** “Color moving pictures quality metric**",** ICIP-96, Lausanne, Switzerland, 1996. pg. 877-880.
28. Carney, T., Klein, S. A., Tyler, C. W., Silverstein, A. D., Beutter, B., Levi, D., Watson, A. B., Reeves, A. J. Norcia, A. M., Chen, C.-C, Makous, W. & Eckstein, M. P, “The development of an image/threshold database for designing and testing human vision models” Proceedings, Human Vision, Visual Processing, and Digital Display IX, SPIE, 1999.
29. A.B.Watson, “Visual detection of spatial contrast patterns: Evaluation of five simple models”. Optics Express 6(1), 2000, pg. 12-33.
30. A.B. Watson and D.G. Pelli, “Quest: A Bayesian adaptive psychophysical method”, Percept Psychophysics, 33(2), 1983, pg. 113-120.
31. B.Farell and D.G. Pelli, “ Psychophysical methods, or how to measure a threshold and why”, J.G.Robson and R.H.S Carpenter (Ed.), A Practical guide to vision research, Oxford University Press, 1998.
32. International Standard ISO/IEC 13660:2001.
33. Gilbert Strang and Troung Nguyen, Wavelets and Filter Banks, Wellesley-Cambridge Press, Wellesley, MA, 1996.
34. E.Noreen, Computer intensive methods for testing hypotheses, 2nd Edition, Chapman & Hall, London, 1989.
35. International Standard ISO/IEC JTC 1 SC28, Image quality for printer systems.

VITA

Vijay Venkatesh Mahalingam was born on 22nd December 1976 in Madurai, India. He received his Bachelor's Degree in Electronics and Communication Engineering from Madurai Kamaraj University, India in the year 1999. In pursuit of his higher education he attended The College of Engineering at University of Kentucky, Lexington. His research interests are in Signal and Image processing and he was a Research assistant in the Signal and Image Processing Lab. His research efforts, funded by Lexmark International., have led to two publications in PICS conferences and he is a student member of IEEE.