

AFIT/DS/ENG/96-14

PERCEPTUAL FIDELITY FOR
DIGITAL COLOR IMAGERY

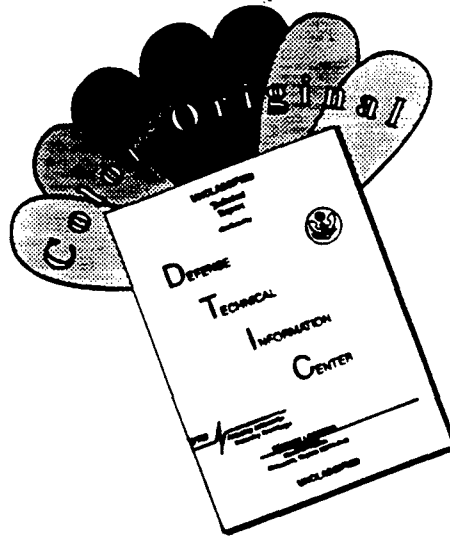
DISSERTATION
Curtis Eli Martin
Captain, USAF

AFIT/DS/ENG/96-14

Approved for public release; distribution unlimited

19970211 039

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

The views expressed in this dissertation are those of the author and do not reflect the official policy or position of the Department of Defense or the U. S. Government.

AFIT/DS/ENG/96-14

PERCEPTUAL FIDELITY FOR DIGITAL COLOR IMAGERY

DISSERTATION

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology

Air University

In Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

Curtis Eli Martin, B.S.E.E., M.S.E.E.

Captain, USAF

December, 1996

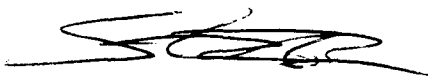
Approved for public release; distribution unlimited

PERCEPTUAL FIDELITY FOR DIGITAL COLOR IMAGERY

Curtis Eli Martin, B.S.E.E., M.S.E.E.

Captain, USAF

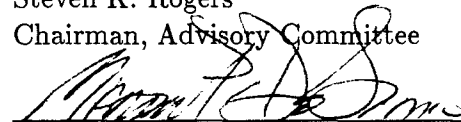
Approved:



Steven K. Rogers
Chairman, Advisory Committee

25 Nov 96

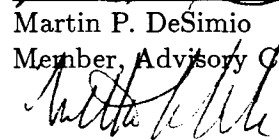
Date



Martin P. DeSimio
Member, Advisory Committee

25 Nov 96

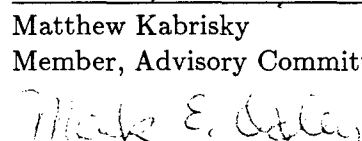
Date



Matthew Kabrisky
Member, Advisory Committee

25 Nov 96

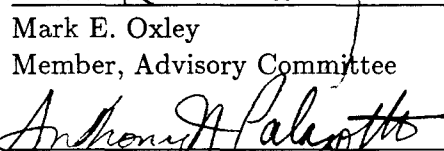
Date



Mark E. Oxley
Member, Advisory Committee

25 Nov 96

Date

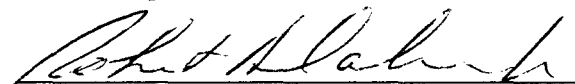


Anthony N. Palazotto
Dean's Representative

4 Dec 96

Date

Accepted:



Robert A. Calico, Jr.
Dean, Graduate School of Engineering

Acknowledgements

"A thousand times every day I remind myself that my inner and outer life depends on the labors of other men...." -*Albert Einstein*

I may not reach a thousand times a day, but it is certainly appropriate to recognize here those who have aided me in accomplishing this work.

First, I thank Dr. Steve Rogers for his superb guidance through my entire stay at AFIT. Particularly in the doctoral program, Dr. Rogers has given me the freedom to chart my own course, while ensuring that the final results would satisfy the degree requirements. For the self-confidence I have thus gained, I am grateful.

It has also been a unique privilege to develop a close association with Dr. Matthew Kabrisky throughout this research. His exhaustive knowledge and keen insight has provided me with direction now, and a model to follow in the future.

To Dan Zambon and David Doak, I am grateful for outstanding computer support. Their superb administration of the computer resources greatly simplified my work.

There are many other faculty, staff, and fellow students who have contributed tremendously to this work, including those who willingly participated in various experiments. Without them, I could not have obtained such intriguing results.

I am deeply grateful to my dear wife Melody. Her unfailing support and confident encouragement have carried me forward as I have progressed through this program. I have truly been blessed to find such a companion.

Finally, I give thankful praise to my God. At some future time I hope to know more fully the extent of His influence in bringing me to this place, at this time, and in these circumstances. For now, I am content to believe, and ascribe to Him the glory.

Curtis Eli Martin

Table of Contents

	Page
Acknowledgements	iii
List of Figures	viii
List of Tables	x
Abstract	xi
 I. Introduction	 1
1.1 Historical Background	1
1.2 Problem Statement and Scope	3
1.2.1 Problem Statement	3
1.2.2 Scope	3
1.2.3 Research Contributions	3
1.3 Dissertation Organization	4
 II. Color HVS Model Contributions	 6
2.1 Introduction	6
2.2 Physiologically Motivated Human Color Vision Models	9
2.2.1 Physiology	9
2.2.2 The Faugeras Color HVS Model	12
2.2.3 Other Color HVS Models	14
2.2.4 Summary	15
2.3 Color Mach Bands	16
2.3.1 Introduction	16
2.3.2 Background	19
2.3.3 Procedure	21

	Page
2.3.4 Discussion	24
2.3.5 Summary	25
2.4 Color Blindness and Color HVS Models	25
2.4.1 Introduction	25
2.4.2 Background	26
2.4.3 Procedure	27
2.4.4 Results	30
2.4.5 Discussion	32
2.4.6 Summary	35
2.5 Summary	36
III. Multiple-channel HVS Model Validation	37
3.1 Introduction	37
3.2 Achromatic HVS Models	37
3.2.1 Early Vision Models	37
3.2.2 Multiple-channel "Cortical" Models	38
3.3 Illusory Contour Formation by a Multiple Channel HVS Model .	39
3.3.1 Introduction	39
3.3.2 Methods	40
3.3.3 Results	41
3.3.4 Discussion	43
3.3.5 Summary	44
3.4 Summary	44
IV. Multiple-channel Color Image Fidelity	45
4.1 Introduction	45
4.2 Previous Image Fidelity Measures	45
4.2.1 Introduction	45

	Page
4.2.2 Early Vision-based Achromatic Image Fidelity Measures	45
4.2.3 Multiple-channel Achromatic Image Fidelity Measures .	46
4.2.4 Color Fidelity Metrics	55
4.2.5 Summary	57
4.3 The New Multi-channel Color Fidelity Measure	57
4.3.1 Introduction	57
4.3.2 The Color Visible Differences Predictor	58
4.3.3 Parameter Values	66
4.3.4 Summary	67
4.4 Demonstration	67
4.4.1 Introduction	67
4.4.2 Banding Distortions	69
4.4.3 Compression Distortions	77
4.4.4 Summary	79
4.5 Discussion	79
4.5.1 Introduction	79
4.5.2 General Observations	79
4.5.3 Parameters	80
4.5.4 Single Number vs. Difference Map Fidelity Measures . .	81
4.5.5 Summary	82
4.6 Summary	82
V. Conclusion	84
5.1 Research Contributions	84
5.2 Future Directions	85
Appendix A. Generation of Gabor Filters	86
Bibliography	92

	Page
Vita	99

List of Figures

Figure	Page
1. Simplified block diagram of the human color vision system.	9
2. Simplified model of the human color vision system.	12
3. Classic brightness Mach band illusion.	17
4. Color Mach band illusion.	23
5. Images used in the image preference experiments.	31
6. Simple achromatic HVS model.	37
7. "More anatomically correct" HVS model [37]	38
8. Illusory contour illusions	40
9. Gabor filters used in processing illusory contour stimuli	41
10. Illusions processed by the Gabor HVS model	42
11. Illusions processed by the CSF/Gabor HVS model	42
12. Plots of adjacent rows or columns in the CSF/Gabor HVS model outputs	43
13. Common structure of recent multiple-channel perceptual image fidelity measures.	47
14. Example of contrast masking	49
15. Sample threshold elevation function	51
16. Psychometric function used in the VDP	53
17. The Faugeras model of the human color vision system.	58
18. Basic elements of the VDP algorithm	59
19. The Color Visible Differences Predictor	60
20. Depiction of the filters used in the color perceptual fidelity measure for a 256×256 image.	63
21. Map of images shown in Figures 22 -27 to demonstrate the color VDP algorithm.	68
22. Color VDP output: mandrill with banding distortion added to A component.	70
23. Color VDP output: mandrill with banding distortion added to C_1 component.	71

Figure		Page
24.	Color VDP output: mandrill with banding distortion added to C_2 component.	72
25.	Color VDP output: mandrill with wide banding distortion added to C_2 component.	73
26.	Color VDP output: mandrill with banding distortion added to Y component of NTSC color space.	74
27.	Color VDP output: mandrill image compressed to 2 bits per pixel using the Hall perceptual coding algorithm.	78
28.	One-dimensional cosine- and sine-Gabor functions.	87
29.	Depiction of a Gabor function in frequency space.	87
30.	Transformation of variables diagram.	89

List of Tables

Table		Page
1.	Results of the color naming portion of the Dvorine color perception test for three color blind subjects (A, B, and C). Errors are shown in bold face. . .	28
2.	Results of the color contrast portion of the Dvorine color perception test .	29
3.	Summary of image preference experiments.	33

Abstract

The problem of measuring the fidelity of digital color images in a manner that corresponds to human perceptual assessments is addressed. Experiments are performed to validate human visual system (HVS) models, which provide access to a "perceptual space" in which visual distortions may be measured, and then a model is proposed for assessing the perceptual fidelity of digital color image. Color Mach bands are produced in the first experiment, demonstrating that, as in the brightness channel, low spatial frequency attenuation occurs in the chromatic channels of the HVS. In the second experiment, a correlation between the chromatic channels of the HVS model and color discrimination axes of color blind observers is demonstrated. Removing variation from one of the chromatic channels of a natural image produces a color-distorted image which the color blind subjects cannot distinguish from the original. Removing variation from the other chromatic channel produces an image that appears colorful to normally-sighted observers, but monochrome to the color blind observers. The third experiment shows that a Gabor filter-based HVS model produces illusory contours in several illusory contour stimuli. These results provide a unique validation of multiple-channel HVS models which process the image in multiple spatial frequency bands that are tuned to match measured sensitivities of neurons in the primary visual cortex of cats and monkeys. Finally, the multiple-channel processing used in the illusory contour experiment is combined with the color vision model from the first two experiments to produce a multiple-channel, color HVS model for measuring perceptual fidelity of color images. A demonstration of the model shows that the structure of the new model is correct. However, inaccurate parameter values for the multiple-channel processing of the chromatic channels cause over-prediction of visible differences in these channels. Suggestions for improving the estimates of the faulty parameters are provided.

PERCEPTUAL FIDELITY FOR DIGITAL COLOR IMAGERY

I. Introduction

1.1 Historical Background

Throughout history, man has sought to improve his ability to communicate with others by creating visual reproductions (images) of the world. From cave paintings to digital cameras, advances in imaging technology have been driven largely by a compelling desire for a high-quality reproduction of the world. An image that more accurately represents what is seen by an observer has the power to convey more information to its viewers. It can give them more of a sense (qualia) of actually *being* in the place of the observer and perceiving the scene with their own eyes. The quest to obtain more complete communication through higher quality images has motivated the development of increasingly sophisticated methods in imaging technology.

Through most of this development, the primary judge of image quality has been the human observer. This reliance on human assessments is natural, given the fact that human observers are nearly always the end users of images. However, in many modern systems, the large volumes of imagery that are generated make it impractical to use human observers to judge image quality. With the advent of digital image technology and increased understanding of the human visual system, the ability to measure the visual quality of an image without the aid of a human observer is beginning to become a reality. This dissertation describes the development of such a measure.

In image processing literature, there is some ambiguity in the use of the term *quality*. In order to understand this research in context, it is important to make a distinction in the usage of this term. Broadly, image quality measures may be classified as either absolute or relative measures. Absolute quality refers to the assessment of an image based on the image itself, without direct comparison to other images. This type of measure may be used to determine how well the image represents the real world for its intended purpose. In contrast, relative quality refers to the assessment of the quality of an image as

compared with a reference image. Most frequently, the two images in this type of quality measure consist of an original image and a perturbed version of that original, and the assessment is intended to describe how accurately the perturbed version represents the original. Depending upon how the measurement is taken, such relative quality measures may be called measures of fidelity (faithfulness to the original) or distortion (difference from the original). In this dissertation, the term *fidelity* will be used to refer to relative quality, while *quality* by itself will generally refer to an absolute quality assessment.

Because human observers are the end users of nearly all image producing or processing systems, the standard for evaluating the effectiveness of a quality or fidelity measure is how well it corresponds with actual human assessments. Useful machine-based measures of image quality or fidelity should correlate well with such assessments. The term *perceptual measure* is frequently used to describe these measures, referring to this goal of correlation with human performance. Early approaches used simple mathematical metrics such as mean square error to compute fidelity, without taking into account the operation of the human visual system (HVS), through which humans must perform their assessments. Perhaps not surprisingly, these early fidelity measures were found to be poor perceptual measures, failing to correlate well with the assessments of human observers [24]. More recent efforts have focused on incorporating models of the human visual system into image quality and fidelity measures. While these efforts have produced encouraging progress, a number of significant mileposts remain to be reached. In particular, most of the previous work has been limited in application to fidelity measures for monochrome imagery. The techniques developed in this domain have reached a rather high level of sophistication, in terms of both modeling the HVS and achieving results which correlate well with human assessments. However, relatively few techniques have been proposed for assessing fidelity of color images, and approaches to measuring the quality of either monochrome or color images remain few in number. The research reported here directly addresses the first of these problems—assessing fidelity of color images.

1.2 Problem Statement and Scope

1.2.1 Problem Statement. This dissertation focuses on the problem of developing a perceptual fidelity measure for color digital images.

1.2.2 Scope. The perceptual fidelity measure will be limited in the following ways:

1. The availability of a reference image for each image to be measured is assumed. This limits the application of the fidelity measure to evaluating systems which produce distortions in digital images, such as compression systems.
2. It is assumed that each color image to be measured, along with its reference image, are already in digital format prior to entering the perceptual fidelity measure.
3. A display model is required to establish a "viewing space" of the images. The display model used in this research is a very simple model of a color computer monitor display with a D_{6500} white point. No effort is made in this model to account for nonlinearities in an actual physical display.
4. The fidelity measure is developed for application to still images only. This assumption allows temporal characteristics of the human visual system to be ignored.

1.2.3 Research Contributions. In successfully developing a perceptual color image fidelity measure, this research has produced the following original contributions:

1. Experimental validation of color HVS models. This validation consists of demonstrations in two key areas:
 - (a) A demonstration of color Mach bands, created by forming a Mach band stimulus pattern in one of the color-mediating channels of a color HVS model. Prior to this research, such a clear demonstration of color Mach bands did not exist [71]. The appearance of color Mach bands in the stimulus created for this research supports the assertion that low spatial frequencies are attenuated in the color channels of the HVS [57].

- (b) A demonstration of correlation between the chromatic channels of the color HVS model and the color perception deficits of some color blind observers. This demonstration is accomplished by transforming the colors of a complex image into the perceptual space of the HVS model, removing the variation in the image from one of two chromatic channels of the model, and transforming the result back to the original color space. Removing variation from one of the two chromatic channels produces an image that is clearly different than the original to normal observers, but is indistinguishable from the original by the color blind observers. When the variation is removed from the other chromatic channel, the resulting image appears colorful to normal observers, but monochromatic to the color blind observers. These results suggest that the color HVS model accurately models the separation of color information that is performed by the HVS [56].
- 2. Experimental validation of a multiple-channel model of cortical processing in the HVS through the use of illusory contours. Formation of illusory contours in the output of the multiple-channel HVS model, in locations corresponding to those where illusory contours are perceived by human observers, supports the use of multiple-channel models of visual cortical processing [58].
- 3. The combination of a multiple-channel HVS model for assessing perceptual fidelity of monochrome images with a model of the color processing elements of the HVS to produce a model for assessing the perceptual fidelity of color images.

1.3 Dissertation Organization

The dissertation is organized into five chapters. The following three chapters elaborate on each of the three contributions enumerated above, each providing sufficient background material to put the work into proper context. Chapter II discusses color HVS models, while Chapter III focuses on multiple channel visual models for monochrome images. Chapter IV describes the combination of a multiple channel monochrome HVS model with the color HVS model to produce a multi-channel color HVS model for assessing fi-

delity of color images. Finally, Chapter V summarizes the research with a review of the contributions made.

II. Color HVS Model Contributions

2.1 Introduction

The human perception of color is a remarkable phenomenon that has received scholarly attention for centuries. Indeed, among other characteristics of the physical world, Newton ventured an explanation of colored light at the start of the eighteenth century [63]. A century later, Young proposed the revolutionary idea that the perception of color could be mediated by just three different types of retinal receptors [107]. Yet another century after that, Helmholtz provided the first suggestions of how these three different types of receptors would vary in response to light of different wavelengths [100]. These ideas of Young and Helmholtz, substantiated by subsequent experiments, have led to what is now commonly called the Young-Helmholtz trichromatic theory of color vision. This theory leads to the powerful result that any wavelength distribution of light may be expressed, in terms of human perception, as a set of three numbers, called tristimulus values, which are computed relative to three standard "primary" distributions. Further, the theory allows algebraic operations which have perceptual significance to be performed on these three numbers. For example, two different wavelength distributions of light which have the same tristimulus values will be perceived as the same color (in fact, indistinguishable from each other), and tristimulus values of two wavelength distributions can be added to produce the tristimulus values of the optical mixture (sum) of the two distributions. Thus, the collection of all colors may be conveniently represented as a three-dimensional vector space, where unit vectors represent unit amounts of the three primary distributions; changing the primaries amounts to a transformation (rotation) of the axes of the vector space. (For a complete discussion of colorimetry and its psychophysical foundations see [105].)

It should be noted that the Young-Helmholtz trichromatic theory is more a theory of the *receptors* of human color vision than a comprehensive theory of color vision. The trichromatic theory does not fully account for all the complexities of color vision; many aspects of how we see a colored world still remain to be uncovered [108]. Nevertheless, the Young-Helmholtz theory is a powerful tool which allows a stimulus to be characterized in perceptual terms. The ability to express color in terms of a three-dimensional space

with convenient mathematical properties has made it possible to express color images in a simple digital format (three color coordinates for each pixel), and it has tremendously simplified the process of manipulating these color images by allowing transformations to be expressed in terms of operations on the three color components.

An important goal in the progress of digital color image processing techniques has been the development of *perceptual* color spaces. Rather than expressing colors as unit amounts of arbitrary primary stimuli, these spaces seek to express color in unit perceptual amounts, so that a unit change in any direction from a given point (color) in the perceptual space corresponds to a color change (in brightness, hue, or saturation) that is just perceptible to the human observer [36]. Such a uniform space allows mathematical distances to be computed between two colors that roughly correspond to the qualitative differences expressed by the human observer.

Increasing understanding of the processes of color vision has helped to guide the development of these spaces by providing insight into the relationship between physical and perceptual descriptions of light. Early experiments in developing this relationship include color matching experiments, which led in 1931 to the adoption of standard human observer color matching functions by the Commission Internationale de l'Eclairage (CIE) [105]. By providing a set of standard primaries related to the human perception of color, this specification was the first attempt to provide a means for expressing a physical wavelength distribution of light in the perceptual terms of a tristimulus space. In the specification of this 1931 standard, the CIE proposed a color space referred to as the XYZ color space, in which a convenient planar representation of color known as a chromaticity diagram was realized. While (x, y) chromaticity coordinates (based on the XYZ tristimulus space) are still used today for specifying the color of a stimulus independent of its brightness, the chromaticity diagram and the CIE XYZ system upon which it is based lack the desirable property of perceptual uniformity—a perceptible change in color does not correspond to a unit change in either (x, y) chromaticity coordinates or XYZ tristimulus values [6].

One widely popular system for organizing color in terms of perceived differences is known as the Munsell Color System. In the Munsell Book of Color [105], color patches are organized in an arrangement such that loci of constant hue, saturation, and luminance form

a polar coordinate system. While this arrangement provides a psychophysical measurement of unit differences in hue, saturation, and brightness (perceptual values), the CIE standard tristimulus values (physically measurable values) of the color patches do not correspond directly to the polar arrangement. Several attempts have been made to develop simple transformations to relate CIE tristimulus values to the Munsell system. One of these that has received considerable attention is the Lab, or cube-root color coordinate system [32]. This system agrees closely with the psychophysical Munsell results, and its coordinates also correspond well with several physiological results [36].

Another commonly-used color system is the National Television Systems Committee (NTSC) receiver primary color coordinate system, frequently referred to as the YIQ system for the names commonly used to identify the three channels. This system is somewhat simpler than the Lab system, as it is accomplished by a single linear transformation from the CIE standard XYZ tristimulus space. The great triumph of the NTSC approach is a system which produces one high resolution luminance signal and two chrominance signals that can be dramatically downsampled spatially without serious visual effect, providing a means for broadcasting color television signals that are compatible with already existing black-and-white systems [3]. Despite this success, Frei points out that, like the CIE XYZ coordinate system, the color separations of the YIQ space do not correlate well with the human perception of color [28].

With the exception of the Lab system, all the color systems mentioned so far were developed solely on the basis of psychophysical measurements. Advancements in the understanding of the physiology of the visual system have encouraged the development of more sophisticated models which mimic the structure of the visual system more completely. Thus, current color HVS models rely upon both physiological and psychophysical results. Physiology suggests the structure of the model, while psychophysical results provide system performance data for setting parameters of the model.

This chapter considers both spatial and spectral aspects of a physiologically motivated color HVS model which has been used successfully in a wide range of engineering applications. First, Section 2.2 gives a top level overview of the main physiological elements of the human color vision system and describes the implementation details of the

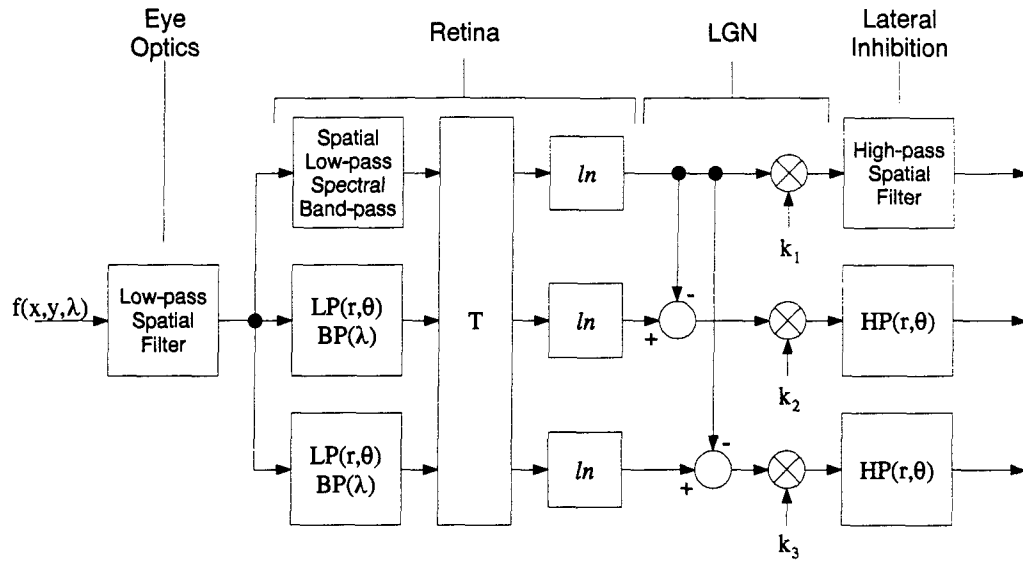


Figure 1. Simplified block diagram of the human color vision system. (Figure adapted from [36]).

color HVS model. In Section 2.3, the spatial aspects of the chromatic channels of the color HVS are explored. Specifically, color Mach bands are produced, lending support to the idea that low spatial frequencies should be attenuated in the chromatic channels of a perceptual model of the color HVS. Section 2.4 examines the separation of chromatic information into two chromatic channels by the model, showing that the chromatic channels correspond to color discrimination axes of observers with a certain type of color deficiency. Finally, Section 2.5 summarizes the significance of these results in the context of a perceptual color fidelity measure.

2.2 Physiologically Motivated Human Color Vision Models

2.2.1 Physiology. For the purposes of this discussion, the structure of the color HVS may be represented schematically as shown in Figure 1. The models examined in this chapter are inspired by this structure. The block diagram provides a good framework for a functional description of the operation of the various elements of the color visual system, generally following a summary by Hall [36]. The input signal $f(x, y, \lambda)$, represents a spatial and spectral distribution of light. For convenience, a rectangular coordinate system (x, y)

is assumed for the space-domain representation of the input signal, while polar coordinates (r, θ) are used to describe the spatial filters in the spatial frequency domain.

The spatial filter on the input represents the optics of the eye, including the pupil, the lens, and the ocular media. Like all optical systems, these optics attenuate high spatial frequencies, hence they may be modeled with a low-pass linear spatial filter. The eye is also known to introduce chromatic aberration into the visual signal; this phenomenon is ignored in the models discussed below.

At the back of the eye, the optical signal is sensed by a closely-packed array of photo-receptive neurons (rods and cones) in the retina. The rods are ignored here because it is assumed that the color images are viewed under photopic illumination conditions, in which the cone system dominates. Three types of cones are present in the normal human retina, differentiated from each other by their spectral sensitivity to light. The first bank of filters in Figure 1 represents these three different cone types. The spectral filters represent the spectral sensitivities, while the spatial filters model spatial effects due to the spatial distribution and interactions of the three types of cones.

The three cone types are conveniently designated according to the relationship between the peak wavelengths of their sensitivity curves: 570, 540, and 445 nm [105]. Thus, the cones with a peak response at 570 nm are referred to as L (long wavelength) cones, while those that peak at 540 nm are called M (medium wavelength) cones, and those that peak at 445 nm are called S (short wavelength) cones. Faugeras points out that this designation in terms of peak wavelengths is more appropriate than the more common designation as "red," "green," and "blue" cones for two reasons: first, the peak wavelengths do not correspond to these spectral colors, and second, excitement of a certain type of cone doesn't necessarily indicate the color of the stimulus [25].

In addition to differences in spectral sensitivity, there is also a difference in the number and distribution of the three cone types throughout the retinal mosaic. The resulting finite spatial sampling of each of the three systems is modeled by the low-pass spatial filters represented in the three cone channels in Figure 1.

Signals produced by the cones are processed by several other layers of neurons in the retina, ending with ganglion cells, which send retinal output signals to the lateral geniculate nucleus (LGN) via the optic nerve. The block following the retinal sensing blocks in Figure 1 represents this processing in the interconnected retinal layers above the cones. Jameson suggests that the processing of these layers produces a set of linear spectral summations of the cone outputs, represented by a linear transformation matrix T [45].

The final component of the model associated with the retina is a non-linear function, reflecting a non-linear relationship between input visual signals and measured retinal outputs. The exact functional form for this nonlinearity has been the topic of debate [12, 78, 81, 83–85, 94]. Three main alternatives are Fechner's logarithmic relation, Stevens' (cube root) power law, and an adaptive form expressed by

$$V_s = K_m \cdot \frac{I}{I + S}, \quad (1)$$

where V_s is the output response, I is the input intensity, and S is an adaptation parameter. This debate has not yet been fully resolved, although there have been some recent attempts to unify the three models (see, for example, [106]). In this work the logarithm is used exclusively, largely because of the analytical simplicity it provides.

The set of linear adders and multiplicative constants that follow the nonlinearity in Figure 1 represents two sets of opponent cells and one set of non-opponent cells found in the lateral geniculate nucleus (LGN) [21]. The top path represents the non-opponent cells, which are believed to transmit brightness information to the brain. The two lower channels, on the other hand, carry chromatic information in the form of two color difference signals: red-green and blue-yellow. In the models, multiplicative constants are included at this stage to ensure that incremental changes in chrominance signals produce an equivalent incremental change in hue. In practice, the differencing and scaling operations are accomplished with a single matrix multiplication. The results of the experiments performed with color blind observers reported in Section 2.4 provide interesting insight into the chromatic separation achieved by this stage of the HVS.

The final stage of the visual system depicted in Figure 1 consists of a set of high-pass spatial filters. These filters are included to account for lateral inhibition and other mechanisms which attenuate low spatial frequencies in the visual signals. Hall indicates that this low frequency attenuation may actually occur at various stages along the visual path—as early as the retina for the luminance channel and as late as the visual cortex, if at all, for the chrominance channels [36:14]. The issue of whether or not there is low-frequency attenuation in the chromatic channels of the color vision system is addressed by the color Mach band experiment described below in Section 2.3.

2.2.2 The Faugeras Color HVS Model. Because the color HVS model described here was designed for use with digital color images, its structure is simplified somewhat from that shown in Figure 1 to that illustrated in Figure 2. This is accomplished by means

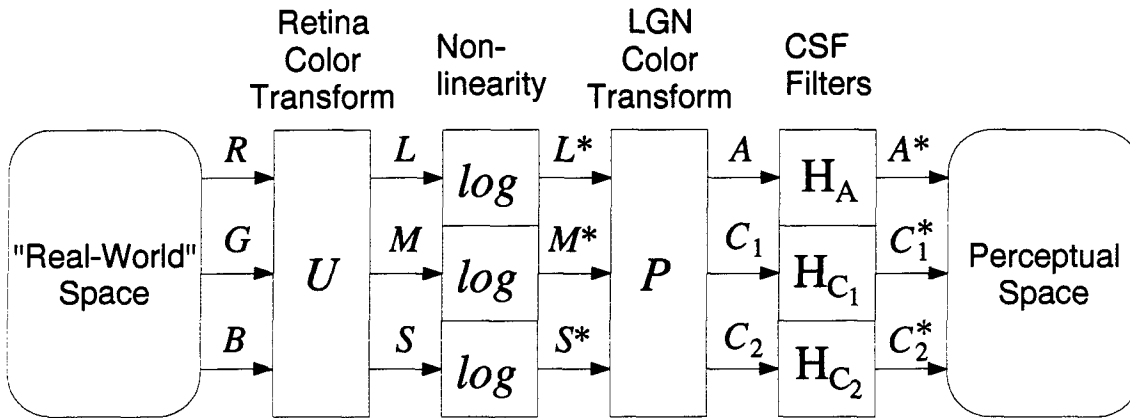


Figure 2. Simplified model of the human color vision system (adapted from [25]).

of two simplifying assumptions. First, it is assumed that the image is already expressed in terms of some tristimulus space, such as the CIE (R,G,B) space. This allows the cone absorption and neural interaction mechanisms of the retina, represented by the spectral filters and the transformation T in Figure 1, to be reduced to a single 3×3 transformation matrix U . The second, larger, assumption is that the low-pass filters shown in Figure 1 may be applied after the spectral transformations and the non-linear function, and combined with the high-pass spatial filters to form a single set of three linear, spatially invariant

bandpass filters applied to the outputs of the LGN stage. The validity of this assumption is addressed in Section 2.3.

As shown in Figure 2, the retinal stage of the Faugeras model consists of the linear transformation matrix U followed by a logarithmic non-linearity. The entries in the transformation matrix are determined based on a cone pigment absorption space developed by Stiles [86]. For a display monitor with a D_{6500} white point and images expressed in RGB coordinates, this transformation is given as [25]:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} .3634 & .6102 & .0264 \\ .1246 & .8138 & .0616 \\ .0009 & .0602 & .9389 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (2)$$

Of the different functional forms that have been suggested for the non-linear response of the retina, Faugeras chose the logarithm for its computational simplicity and to provide a homomorphic automatic gain control characteristic to the model [25]. After applying the logarithm, the outputs of the first stage of the HVS model are designated as L^* , M^* , and S^* , identifying the three types of cone photo-receptors in the retina.

The second stage of the Faugeras model, which accounts for the achromatic/chromatic separation performed by the LGN, is realized by a linear transformation matrix P that is applied to the outputs of the cone absorption stage. Faugeras gives this transformation as [25]:

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 13.8312 & 8.3394 & 0.4294 \\ 64 & -64 & 0 \\ 10 & 0 & -10 \end{bmatrix} \begin{bmatrix} L^* \\ M^* \\ S^* \end{bmatrix}. \quad (3)$$

As shown in Figure 2 and Equation 3, the outputs of this stage of the model are an achromatic channel, A and two chromatic channels, C_1 and C_2 . These channels are based on the physiological measurements of the LGN; the actual parameter values in the transformation matrix are derived from psychophysical brightness matching and color matching data [25]. The parameters for the achromatic channel are chosen in such a way as to produce an approximation to the photopic luminous efficiency function, $V(\lambda)$. Thus, this channel corresponds to the human perception of brightness. For the two chromatic chan-

nels, the parameters are fixed by solving an optimization problem to find a transformation which maps MacAdam ellipses onto circles of equal radius. MacAdam ellipses are loci of just-detectable color changes determined psychophysically through color change detection experiments [6]. Mapping these ellipses to circles produces an approximately uniform chromaticity space, in which a unit change in chromaticity coordinates corresponds to a just perceptible change in perceived color [25].

The final stage of the Faugeras model accounts for spatial effects of the HVS using linear, space invariant filters. As mentioned previously, Faugeras assumes that although these spatial effects occur all along the visual path, they may be lumped together into three filters applied to the output of the LGN stage [25]. Faugeras specifies the form of the transfer functions of these three filters by combining results of his own psychophysical experiments, which specify the low-frequency characteristics, with those of Campbell and Green [8], which specify the high-frequency characteristics [25].

Because the parameters of the Faugeras model are chosen based on measurements of human perception, the space described by A^* , C_1^* , and C_2^* may legitimately be called a *perceptual space*. With such a model of the visual system, it is possible to manipulate variables within the internal (perceptual) representation of the visual stimulus. This ability to probe perceptual HVS channels forms the basis for using the Faugeras model in the experiments detailed below.

2.2.3 Other Color HVS Models. Another color HVS model with strong perceptual correlation was proposed by Hall [36] at about the same time as Faugeras proposed his model. The Hall model is based on slightly different psychophysical and neurophysiological data than that used by Faugeras. However, it can be described using the same three-stage structure shown in Figure 2, and, despite arising from different experimental results, Hall's transformation matrices do not differ much numerically or perceptually from those of Faugeras—performing identical manipulations inside both color spaces produces almost indistinguishable results.

The filters used by Hall in the final stage of his model are also very similar to Faugeras' filters. This turns out to be advantageous, because Hall provides his filters in the form

of a mathematical expression, rather than the plots given by Faugeras. Hall's filters are based on a functional form obtained by Mannos and Sakrison through a unique approach involving human subjective quality ratings of complex images which agrees closely with several different measured human contrast sensitivity curves [10, 18, 48, 69, 98] (see [54] for a comparative plot of the results of these experiments). Although several other functional forms have been proposed for matching contrast sensitivity data [92], the Mannos and Sakrison form is perhaps the most commonly used functional form used to model the spatial processing in the HVS. Adjusting the parameters to yield unity gain at peak center frequencies of 8, 4, and 2 cycles/degree for the A , C_1 , and C_2 channels, respectively, the spatial filters are expressed as [36]

$$H_A(f_r) = 2.6[0.0192 + 0.113f_r] \exp[-(0.113f_r)^{1.1}], \quad (4)$$

$$H_{C_1}(f_r) = 2.6[0.0192 + 0.226f_r] \exp[-(0.226f_r)^{1.1}], \quad (5)$$

and

$$H_{C_2}(f_r) = 2.6[0.0192 + 0.452f_r] \exp[-(0.452f_r)^{1.1}], \quad (6)$$

where f_r is radial spatial frequency expressed in units of cycles per visual degree. Because of the similarity between these filters and the plots given by Faugeras, the functional forms of Equations 4 through 6 are adopted as the filters for the color HVS model used in this research.

2.2.4 Summary. This section has outlined a simplified, physiologically motivated HVS model consisting of a retinal cone transformation, an amplitude non-linearity, and a second colorimetric transformation representing the LGN. Spatial effects in each color component of this model are accounted for by a bandpass filter applied to the output of the LGN stage. The bandpass filters in the Faugeras model are closely represented by the functional form obtained by Mannos and Sakrison, used in the Hall model. In the following two sections, the colorimetric and spatial aspects of this model are explored further. Section 2.3 examines the question of whether or not low-frequency attenuation occurs in the chromatic components of the HVS by creating a stimulus that produces

a color Mach band illusion. Section 2.4 demonstrates the correspondence between the color components of the Faugeras model and the HVS by manipulating images in the Faugeras color space to produce colored images that are perceived as monochrome by anomalous trichromat observers. These two results support the use of the Faugeras model in developing a measure of perceptual color image fidelity.

2.3 Color Mach Bands

2.3.1 Introduction. The purpose of this section is to consider the question of whether or not low spatial frequencies are attenuated in the chromatic channels of the human visual system. The question is addressed through the creation of color Mach bands. Since first being described in 1865, Mach bands have been one of several illusions used to show that perceived brightness measured by the HVS is not a simple function of incident intensity [12]. Mach bands are regions of increased or decreased perceived brightness which appear at locations where a luminance gradient meets a plateau, as illustrated in Figure 3. Much effort has been expended to determine the causes of this illusion, in terms of both the stimulus conditions and the HVS mechanisms which bring it about [78]. While this work has been largely successful in outlining stimulus conditions which produce the Mach band illusion, as well as those which defeat it, there remains some disagreement as to what HVS mechanisms actually cause Mach bands to appear [40,79,80,99]. From the standpoint of linear filtering HVS models, the appearance of brightness Mach bands is attributed to attenuation of low spatial frequencies by the HVS [87].

Expanding consideration of the HVS to include color perception, it is reasonable to ask whether or not Mach-type phenomena can be produced using colored stimuli. Assuming that Mach bands are indeed caused by low spatial frequency attenuation, this is equivalent to asking whether or not the low spatial frequencies of the chromatic channels of the HVS are attenuated. Investigation of this question has caused a lively debate—a number of researchers claim to have produced color Mach bands, despite a general conclusion that they do not exist (for a summary, see [71]). Pease suggests that this disagreement is due at least in part to a poor specification of what is meant by the term “color Mach

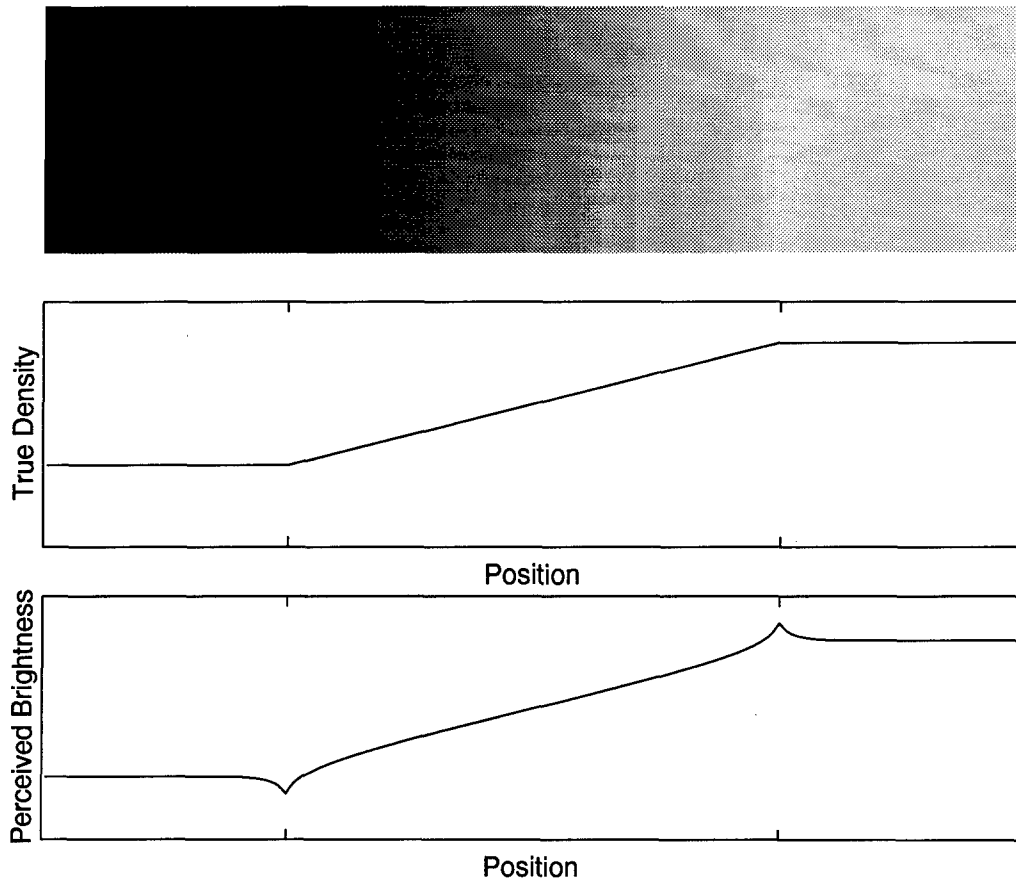


Figure 3. Classic brightness Mach band illusion. The density distribution in the top image is as shown in the middle plot, yet the apparent brightness of the stimulus is similar to that shown in the bottom plot, with bright and dark bands appearing at the knee points where the ramp meets the uniform areas (adapted from [87]).

band,” and proposes a distinction between brightness, hue, and saturation effects [71].¹ He argues that such a distinction in perceptual terms is both appropriate, because Mach bands arise as a result of human perception, and necessary, because it provides clarity in the description of the phenomenon. In these terms, Pease concludes from his own review of the literature that brightness and saturation Mach bands have been demonstrated, but that there is no evidence for the existence of hue Mach bands.

Keeping in mind the distinction between brightness, hue, and saturation effects suggested by Pease, this section considers the question of color Mach bands within the context of the perceptual HVS model of Faugeras. Assuming that the processing in the brightness channel produces the brightness Mach band illusion, it is reasonable to suppose that a Mach band illusion may be produced by constructing displays having uniform brightness which vary only in the chromatic channels. Mach bands arising from such stimuli would be accurately termed “chromatic Mach bands,” and may be manifest as a combination of hue and saturation effects [71]. Just as brightness Mach bands appear as a perceived shift along a brightness axis, chromatic Mach bands should appear as a perceived shift in color along a chromatic axis.

This section presents an approach to creating chromatic Mach bands on a standard computer workstation color display. Bands with the expected color shift are observed in the stimulus. For the purposes of this discussion, they will simply be called chromatic Mach bands to indicate a perceived color shift in the stimulus that has no change in brightness that could give rise to the bands. Although the appearance of the bands suggests that they could in fact be hue Mach bands, more careful analysis of the chromatic Mach bands is required in order to describe them with this term suggested by Pease.

Section 2.3.2 provides a context for this experiment by describing what a chromatic Mach band should look like and reviewing pertinent results already obtained. The production of the stimulus on a standard workstation color monitor and the appearance of

¹Brightness, hue, and saturation are the terms used by vision scientists to specify the perception of a color stimulus. Of these three, brightness is the most intuitive in its meaning. Hue corresponds most closely to what is commonly thought of as color (red, green, blue, etc.), while saturation describes the amount of white light that is present in a stimulus. For example, scarlet and pink are both red hues of different saturation: scarlet is highly saturated, while pink is desaturated.

the illusion is described in Section 2.3.3. Finally, Section 2.3.4 discusses the implications of these results.

2.3.2 *Background.*

2.3.2.1 *Description (Definition) of Chromatic Mach Bands.* Pease defines hue (or saturation) Mach bands as “regions with a change in the perceived hue (or saturation) localized near the flection points in a spatial distribution of illumination that varies in only one direction [71].” Subdividing the description of color into brightness and chroma (hue and saturation combined), chromatic Mach bands may be similarly defined as regions in an isoluminant distribution of light where the perceived chroma departs from the actual chroma variation. As in the case of brightness Mach bands, the perceived shift in a chromatic Mach band should represent an overshoot or undershoot where a ramp (in chroma) meets a plateau. A result of the over- and undershoot behavior in the brightness Mach band illusion is that the brightness of the plateau appears to match the brightness of a location in the ramp, and the band between this point on the ramp and the edge of the plateau appears to be the brightest (or darkest) area in the stimulus. Similarly, a chromatic Mach band could be identified by noting a shift occurring at the knee between a ramp and a plateau in chroma (keeping brightness constant everywhere), so that the chroma of the plateau seems to match that of a point on the ramp, and the chroma of the strip at the knee doesn’t appear anywhere else in the stimulus.

2.3.2.2 *Previous Research.* As mentioned earlier, Pease reviewed the work of several investigators of color Mach bands. Of these, only one [17] reported bands in isoluminant color ramp stimuli that could be hue Mach bands. However, because the bands were not described with sufficient detail, it was impossible for Pease to determine if they were indeed hue Mach bands. All other color Mach bands cited in Pease’s review were determined to be either brightness or saturation Mach bands [71].

Beginning with Mach himself, researchers have historically attributed the Mach band illusion to a lateral inhibition mechanism in the HVS [78]. This mechanism is frequently modeled by low spatial frequency attenuation in the modulation transfer function (MTF)

of the HVS. Through contrast sensitivity experiments, such attenuation has been measured for the brightness channel of the HVS. However, attempts to measure low frequency attenuation in the chromatic channels have not been conclusive. Pease cites four efforts to measure low frequency attenuation in the color channels of the HVS [39,82,95,96]. Of these, two ([95] and [82]) used equiluminous spatial sine-wave patterns that were modulated in chromaticity (the perceptual equivalent of both hue and saturation). Although these two experiments were similar in approach and technique, their results are contradictory: where one found evidence of low frequency attenuation [82], the other found none [95].

Another illusion caused by spatial brightness patterns is the Cornsweet illusion [12]. The stimulus for this illusion contains an abrupt luminance edge separating two uniform regions of equal luminance, with smooth, rapid transitions from the extremes of the edge to the luminance of the uniform regions. Although equal in luminance, the two uniform regions appear to differ in brightness; the region on the higher-luminance side of the edge appears brighter than the region on the other side. Using equiluminant stimuli with chromatic variations, Ware and Cowan measured a chromatic Cornsweet effect and found it to be somewhat smaller than the achromatic illusion [101]. As the Cornsweet effect is attributed to low frequency attenuation in the brightness channel of the HVS, the demonstration of the color Cornsweet effect suggests that such attenuation may also occur in the chromatic channels, although it may be somewhat less than that found in the achromatic channel.

Using an alternative approach to measuring contrast sensitivity, Faugeras provides further evidence that low frequency attenuation occurs in the chromatic channels of the HVS [25]. Based on his experiments, Faugeras also offers an explanation for why color Mach bands are difficult to perceive. Developing an HVS model that expresses color with an achromatic (brightness) component and two chromatic components, Faugeras found the spatial frequency response (MTF) of each of the chromatic channels to be roughly equal to frequency scaled (narrower) versions of the MTF of the achromatic system. These narrower MTFs give rise to point spread functions that are broader and lower in amplitude than the achromatic point spread function. Mediated by these broader and lower-amplitude point spread functions, color Mach bands would thus be expected to be both broader and

less visible than brightness Mach bands typically appear. As Faugeras never reported any attempts to produce color Mach bands himself, this experiment was performed during this research to see if chromatic Mach bands could be produced by creating a Mach band pattern in one of the chromatic components of the Faugeras color space.

2.3.3 Procedure. In an approach similar to that used by Stockham [87], the stimulus used in the chromatic Mach band experiment is created in the Faugeras perceptual space, then transformed into the tristimulus space of the display using the inverse of the HVS model. The stimulus is created in the Faugeras space by forming an image in which one of the two chromatic components (C_1 or C_2) is varied in the classic Mach pattern shown in Figure 3 (a linear ramp between two plateaus), while the achromatic component (A) and the other chromatic component are held constant. Then, the stimulus is transformed from the perceptual (A, C_1, C_2) space into the (R, G, B) monitor primary space by applying the inverse of the color HVS model. First, multiplying by the matrix P^{-1} (the inverse of the LGN achromatic/chromatic transformation—see Equation 3) transforms the stimulus into the (L^*, M^*, S^*) cone absorption space. Next, applying the exponent function (inverse of the log non-linearity) to the tristimulus values and multiplying by the inverse of the (R, G, B) to (L, M, S) transformation (see Equation 2) completes the transformation of the stimulus into the monitor's primary space. Note that the spatial filters in the final stage of the Faugeras model are omitted in this experiment. This is done intentionally, as the experiment is intended to probe these spatial processing aspects of the visual system.

After the colorimetric transformation, the color Mach band stimulus is displayed for observation on a standard Sun SPARCstation monitor screen. Several different monitors were used in different trials of the experiment, and no significant differences were observed between the displays. Two different stimulus sizes were used. The smaller stimulus subtends an overall visual angle of approximately 7.15° , while the larger one subtends approximately 19.5° . In both cases, the plateau regions occupy one half the total width of the image (one quarter of the width on either side). In the small stimulus, the Mach bands have a width of approximately 0.6° , while they appear to be 1.2° wide in the larger display.

An example of the color Mach band stimulus is shown in Figure 4, with the ramp and plateaus in the C_1 component, holding A and C_2 constant. On the monitor display, the lower C_1 plateau of this stimulus is a bright green like the color of ripe honeydew melon, while the upper C_1 plateau is the color of ripe cantaloupe. Between the plateaus, the ramp has the appearance of the thin layer found in a cantaloupe rind where the color undergoes a smooth transition from the ripe cantaloupe color to the green outer rind.

At this point it should be noted that all the color prints included in this dissertation are only intended to provide an approximate rendition of what is actually seen on the display screen. The subject of matching colors between a display and a color printer is a large research topic in itself, and no attempt was made in preparing this report to achieve an exact match. A clear example of the mismatch is seen in Figure 4, where the ramp is interrupted by banding artifacts that are due to quantization effects of the color printing process. Despite these defects, the reproduction is of sufficient quality to allow the color Mach bands to be perceived even in the printed copy.

In order to be sure that uniform brightness conditions were present, the luminance of the initial display was measured at several points in the stimulus using a Minolta spot photometer with a $1/3^\circ$ measuring spot. Remarkably, despite the fact that no attempt was made to calibrate the model to match the monitor screens used in the experiment (except for assuming a D_{6500} white point), the luminance was found to be uniform across the face of the stimulus, to within the calibration of the photometer. Because the luminance was measured directly from the displayed stimulus, it is clear that the desired uniform brightness condition was achieved.

The display was viewed by 10 normal trichromatic observers, all of whom reported a shift in the perceived color in the stimulus where the ramp meets the upper plateau (the cantaloupe side). This shift was most commonly described as a difference in the perceived colors of the plateau and the point where the ramp meets the plateau. Specifically, the color at the knee was described as appearing "cleaner" or "purer" than the plateau, which, in comparison, seemed to have a greenish tinge, matching the color on the ramp just to the left of the band. A similar color shift on the lower end of the ramp was more difficult

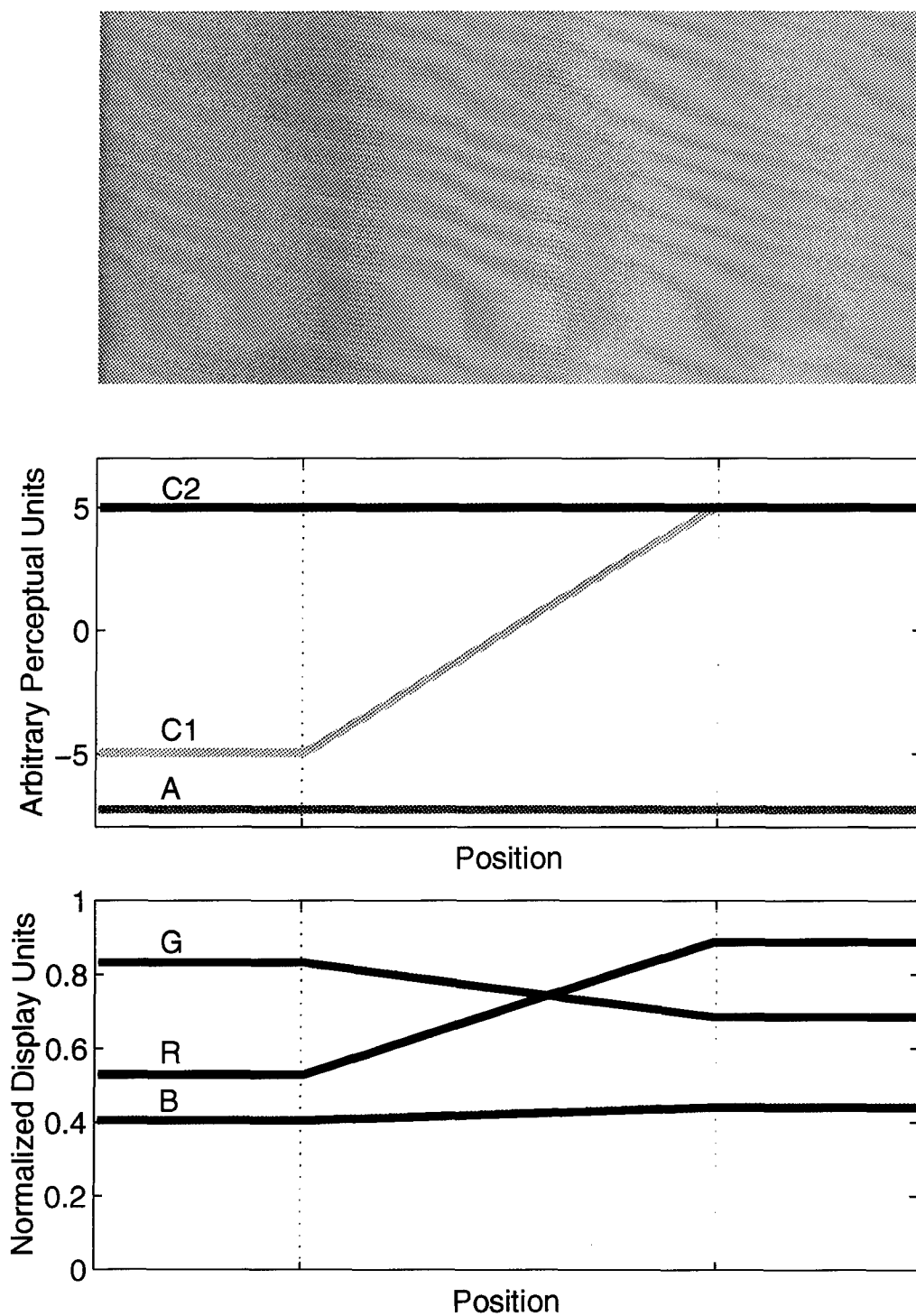


Figure 4. Color Mach band illusion (top), profile of color distribution in (A, C_1, C_2) perceptual space (center), and in (R, G, B) monitor primary space (bottom).

to perceive. This mirrors the experience of many observers who have difficulty perceiving the dark band in the brightness Mach band illusion.

Besides the combination shown in Figure 4, several other combinations of values were tested. The value of the brightness parameter (A) was varied, and ramps of varying steepness were tried in each of the two chromatic dimensions (C_1 and C_2). Although all the combinations tested were not recorded, the bands remained in each case, with the color of the plateau seeming to match a color on the ramp, and the color of the knee showing a notably "purer" character. It was noted, however, that for some sets of values, the strength of the illusion increased, indicating a dependence of the color Mach band illusion on the combination of color values used to create the stimulus.

In addition to the normal trichromats, the initial stimulus (with C_1 variation) was shown to two anomalous trichromat ("color blind") observers. These observers described the stimulus as simply a green square, with little or no variation in the stimulus at all. Because they could see no variation in the stimulus, these two observers were also unable to see the chromatic Mach bands seen by normal trichromats. When the variation was changed to the C_2 component, however, they could readily identify the changes in the stimulus, and they were able to perceive the chromatic Mach bands. These results prompted additional experiments to explore these observers' sensitivity to the C_1 and C_2 channels of the Faugeras model, which are reported in the following section.

2.3.4 Discussion. Beyond demonstrating the existence of Mach bands, the results of the above experiments contain significant implications about the Faugeras color HVS model, as well as the HVS itself. These implications are now discussed.

First, it is significant to note that the A component of the Faugeras model appears to be a reasonable correlate of brightness, as evidenced by the photometer measurements. Without any special calibration of the display, simply setting A to a constant produced a display with uniform brightness, which was surprising even to an experienced color display technician [77]. Further evidence that the model represents brightness well is provided by the color defective observers, who perceived only a uniform field when the C_1 stimulus was displayed, with very little (if any) change in brightness.

By themselves, the observations made by the color defective observers are intriguing. The fact that they could perceive changes along the C_2 axis, but not along the C_1 axis, suggests that the Faugeras model may provide some insight into the nature of color deficiencies. Further experiments to explore this possibility are discussed in the next section.

Finally, the importance of the fact that color Mach bands are evident in the displays should not be discounted. It supports the findings of Schade, Daw, and others who have reported color Mach bands, and bears out Faugeras' assertion that color Mach bands should be visible. As predicted by Faugeras, the chromatic Mach bands do appear broader and somewhat less visible than the brightness effect. Some recent researchers have disputed the lateral inhibition explanation for the existence of Mach bands [80], noting that the lateral inhibition model predicts "maximally induced" Mach bands in a step stimulus which are not observed [79]. Whatever the HVS mechanisms are that produce the Mach illusion, the results obtained here suggest that similar mechanisms exist in the chromatic channels of the HVS. Thus, models which use spatial filters with low frequency attenuation in the brightness channel should include similar attenuation in the chromatic channels as well.

2.3.5 Summary. This section has described the production of a color Mach band illusion. The Faugeras color HVS model allows the stimulus to be created in a space that may appropriately be termed a perceptual space, allowing color variations in the stimulus to be constrained to axes which correspond to chromatic channels in the HVS. Color Mach bands are observed in the stimulus, and they appear to include shifts in hue, which is contrary to most previous results. The existence of the color Mach bands suggests that each of the chromatic channels of the color HVS may have mechanisms similar to those in the brightness channel which cause the appearance of brightness Mach bands. The production of color Mach bands on a computer monitor display is a unique contribution of this research [57].

2.4 Color Blindness and Color HVS Models

2.4.1 Introduction. As noted in the previous section, when viewed by color defective observers, the color Mach band stimulus with variations in only the C_1 component

appeared to be a uniform green square. This result motivated an additional experiment to further examine the perceptions of these observers when presented with images which had been altered in the Faugeras perceptual color space. Using the Faugeras model, RGB images were transformed into the perceptual space of the model, where they were manipulated by setting one of the two chromatic components of the model equal to a constant (the brightness component was never changed). Then, after transforming the modified images back to the original RGB primary space, they were displayed on a color monitor. Setting the perceptual space C_1 component of an image equal to a constant produces an image that deuteranomalous trichromats cannot distinguish from the original full-color image, despite clear color differences apparent to normal trichromats. On the other hand, setting the C_2 component of the image equal to a constant produces an image that appears colorful to normal trichromats, but the deuteranomalous trichromats perceive as monochromatic. These results provide insight into the operation of the color HVS and support the use of the model in image processing applications. This section describes the experiments more fully and discusses the implications of their results.

Section 2.4.2 provides a brief review of normal and defective human color vision. Section 2.4.3 describes the experiments performed, Section 2.4.4 presents the results obtained, and Section 2.4.5 discusses the implications of these results.

2.4.2 Background. As discussed above, general agreement has been reached on a number of basic facts about the physiology and psychophysics of the color HVS. There has also been a significant amount of effort devoted to the study of color blindness. Something of a misnomer, the term "color blindness" refers to an inherited condition affecting certain individuals whose perception of color varies significantly from the majority. The purpose of this section is to summarize the generally accepted understanding of color blindness, as given in [67] (for more complete information, see [105]).

Hereditary color defects afflict around 8% of all males. About a quarter of these are dichromats, while the remainder (6% of all males) are referred to as anomalous trichromats. In dichromats, one of the cone types is almost completely unresponsive. These observers are able to match all spectral hues in color matching experiments using just two spectrally

fixed primaries. Based on which cone type is unresponsive, dichromats are subdivided into three categories. Protanopes, whose *L* type cones are unresponsive, typically confuse reds with greens. The *M* cones are unresponsive in deuteranopes, who confuse purples and greens. Finally, tritanopes, missing response of the *S* cone system, only have sensation of red and green.

The majority of color blind individuals actually lie between the extremes of dichromacy and normal trichromacy. These observers, called anomalous trichromats, require three spectrally fixed primaries to match all spectral hues in the color matching experiments, but their ability to distinguish colors varies significantly from that of normal trichromats, because the response of one set of cones is weak. The distinction between them follows the distinction made between dichromats above, based on the weak cone system. By far, the most commonly occurring color deficient individuals are deuteranomalous trichromats, comprising roughly 5% of all males. These observers, with a weak response from the *M* cones, characteristically confuse light shades of purple with light shades of green. Protanomalous trichromats (1% of males) have weak *L* cones and confuse pinks with light blue-greens. Tritanomalous trichromats, who are very rare, have weak *S* cones, and tend to confuse yellows with blues.

2.4.3 Procedure.

2.4.3.1 Subjects. Three male subjects, referred to as A, B, and C, participated in the experiment. All three were aware that they had color deficiencies, but these were not limiting to their engineering occupations. In conjunction with the experiments reported here, the three were given the 1944 Dvorine color perception test [22], which consists of a color naming portion and a color contrast portion. In the color naming portion, which uses a light and a dark tint of eight basic colors, all three subjects named more than half of the light tints incorrectly, and identified the dark violet tint as blue (see Table 1). In the color contrast portion of the test (the familiar "find the numbers" game), none of the three could distinguish between the color pairs of red/brown, green/orange, yellow/orange, green/yellow, or blue/violet; subject C had further difficulty distinguishing between red/gray, orange/brown, and gray/brown color pairs (see Table 2).

Table 1. Results of the color naming portion of the Dvorine color perception test for three color blind subjects (A, B, and C). Errors are shown in bold face.

Color	Subject's Response					
	Dark Tints			Light Tints		
	A	B	C	A	B	C
Red	Red	Red	Red	Gray	Pink	Beige
Green	Green	Green	Green	Gray	Pink	Beige/Green
Blue	Blue	Blue	Blue	Gray	Blue-green	Gray
Yellow	Yellow	Yellow	Yellow	Yellow	Yellow	Yellow
Brown	Brown	Brown	Brown	Pinkish	Tan/Greenish	Green
Orange	Orange	Orange	Orange	Tan	Tan	Green
Violet	Blue	Blue	Blue	Gray	Gray	Bluish Gray
Green	Green	Gray	Gray	Gray	Pink	Beige

Because the majority of color blinds are deuteranomalous trichromats [105], and all three subjects manifested basically the same difficulties, it is believed that the three subjects are all deuteranomalous trichromats.

2.4.3.2 Color Mach Bands. The current experiment was motivated by the recognition that color blind observers could not detect color differences in the color Mach band stimulus in which only the C_1 component was varied (see Section 2.3 above). On observing this stimulus, subjects A and B reported seeing a square of uniform color and brightness, while subject C indicated a difference in the perceived brightness (but not of the color) between the two sides of the stimulus. When the variation was switched to the C_2 component, all three subjects reported observing different colors in the stimulus. These results led to the conjecture that the perception of color for these subjects is mediated by only the C_2 channel. That is, the subjects seem to be unable to detect color variations in C_1 , but they do appear to be able to detect changes in C_2 . So, if the C_1 component of an image is altered (by, for example, setting it equal to a constant), the subjects should be unable to distinguish the altered image from the original. Conversely, if the C_2 variation is removed from an image, the resulting image should appear to have no color variation whatever for these subjects. The image preference experiment was designed to test this proposition.

Table 2. Results of the color contrast portion of the Dvorine color perception test for three color blind subjects (A, B, and C). Errors are shown in bold face, and blanks indicate charts in which the subject could not perceive any numerals.

Chart	Colors	Number	Response			Chart	Colors	Number	Response		
			A	B	C				A	B	C
A1	Rd/Bl	48	48	48	48	A31	Bl/Or	46	46	46	46
A2	Rd/Bn,Gn/Or	95				A32	Or/Bl	70	70	70	70
A3	Rd/Gy,Bl/Vt	26	2	2		A33	Bl/Gn	85	85	85	85
A4	Bl/Gn/Or	8	3	3	3	A34	Gn/Bl	38	38	38	38
A5	Rd/Yw	49	49	49	49	A35	Bl/Vt	56	56		
A6	Yw/Rd	77	77	77	77	A36	Vt/Bl	39			
A7	Rd/Bl	53	53	53	58	A37	Bl/Bn	22	22	22	22
A8	Bl/Rd	44	44	44	44	A38	Bn/Bl	80	80	80	80
A9	Rd/Or	87	87	87	87	A39	Bl/Gy	36	36	36	36
A10	Or/Rd	20	20	20	20	A40	Gy/Bl	52	52	52	52
A11	Rd/Gn	67		67	67	A41	Or/Gn	28			
A12	Gn/Rd	35	35	35	35	A42	Gn/Or	79			
A13	Rd/Vt	29	29	29	29	A43	Or/Vt	32	32	32	32
A14	Vt/Rd	83	83	83	83	A44	Vt/Or	63	63	68	63
A15	Rd/Bn	5		5		A45	Or/Bn	47	47	47	
A16	Bn/Rd	92				A46	Bn/Or	96	96	96	
A17	Rd/Gy	84	84	84		A47	Or/Gy	4	4	4	4
A18	Gy/Rd	59	59	59		A48	Gy/Or	50	50	50	50
A19	Yw/Bl	76	76	76	76	A49	Gn/Vt	25	25	25	25
A20	Bl/Yw	60	60	60	60	A50	Vt/Gn	93	93	93	93
A21	Yw/Or	3				A51	Gn/Bn	69	69	69	
A22	Or/Yw	74				A52	Bn/Gn	33		30	33
A23	Yw/Gn	62		68		A53	Gn/Gy	68	68	68	68
A24	Gn/Yw	99		69		A54	Gy/Gn	57	57	57	57
A25	Yw/Vt	24	24	24	24	A55	Vt/Bn	23	23	23	23
A26	Vt/Yw	7	7	7	7	A56	Bn/Vt	40	40	40	40
A27	Yw/Bn	98	98	98	98	A57	Vt/Gy	54	54	54	54
A28	Bn/Yw	37	37	37	37	A58	Gy/Vt	9	9	9	9
A29	Yw/Gy	58	58	58	58	A59	Bn/Gy	88	88	88	
A30	By/Yw	2	2	2	2	A60	Gy/Bn	65	65	65	

2.4.3.3 Image Preference. Three test images were chosen, and color-distorted versions of each of the three images were prepared for display to the anomalous trichromat subjects. To generate each distorted image, the original test image was transformed into (A, C_1, C_2) space using the HVS model, and either the C_1 or C_2 components were set to a constant. The resulting distorted image was then passed through the inverse of the HVS model to transform it back into the monitor's (R, G, B) primary space. Any values falling outside the closed interval $[0, 1]$ after this operation were assigned the nearest endpoint value (0 or 1).

Figure 5 shows the images used in the study. As in Section 2.3, it is recognized that a perfect match between colors displayed on the monitor and the printed page is virtually impossible to achieve. However, Figure 5 can still give the reader an idea of the distortions introduced by this process. The images were printed by converting them to color PostScript format and printing them on a Tektronix thermal wax transfer color PostScript printer. As a check on the color reproduction of this process, the resulting printed images were shown to subject A, who reported little difference between the displayed and the printed images—his color perceptions were largely the same for both media. Figure 5 has also been printed on a transparency slide with the same printer, producing similar results to those obtained with the color display.

In each experimental trial, both an original and a distorted version of an image were displayed side by side on a Sun SPARC-20 workstation 24 bit color monitor. The images occupied roughly 7° of the subject's field of view. The subjects were not informed of the color differences between the two images. In order to avoid any reference to color, each subject was simply asked which image he preferred. After he indicated his preference, the subject was asked to describe the differences in the two images qualitatively, and to justify his choice.

2.4.4 Results. Table 3 summarizes the results of the image preference experiments. In the C_1 experiments, the subjects had a difficult time distinguishing between the original and the constant C_1 images, choosing the color-distorted image as identical to or better than the original in half of the trials. The subjects consistently referred to the color

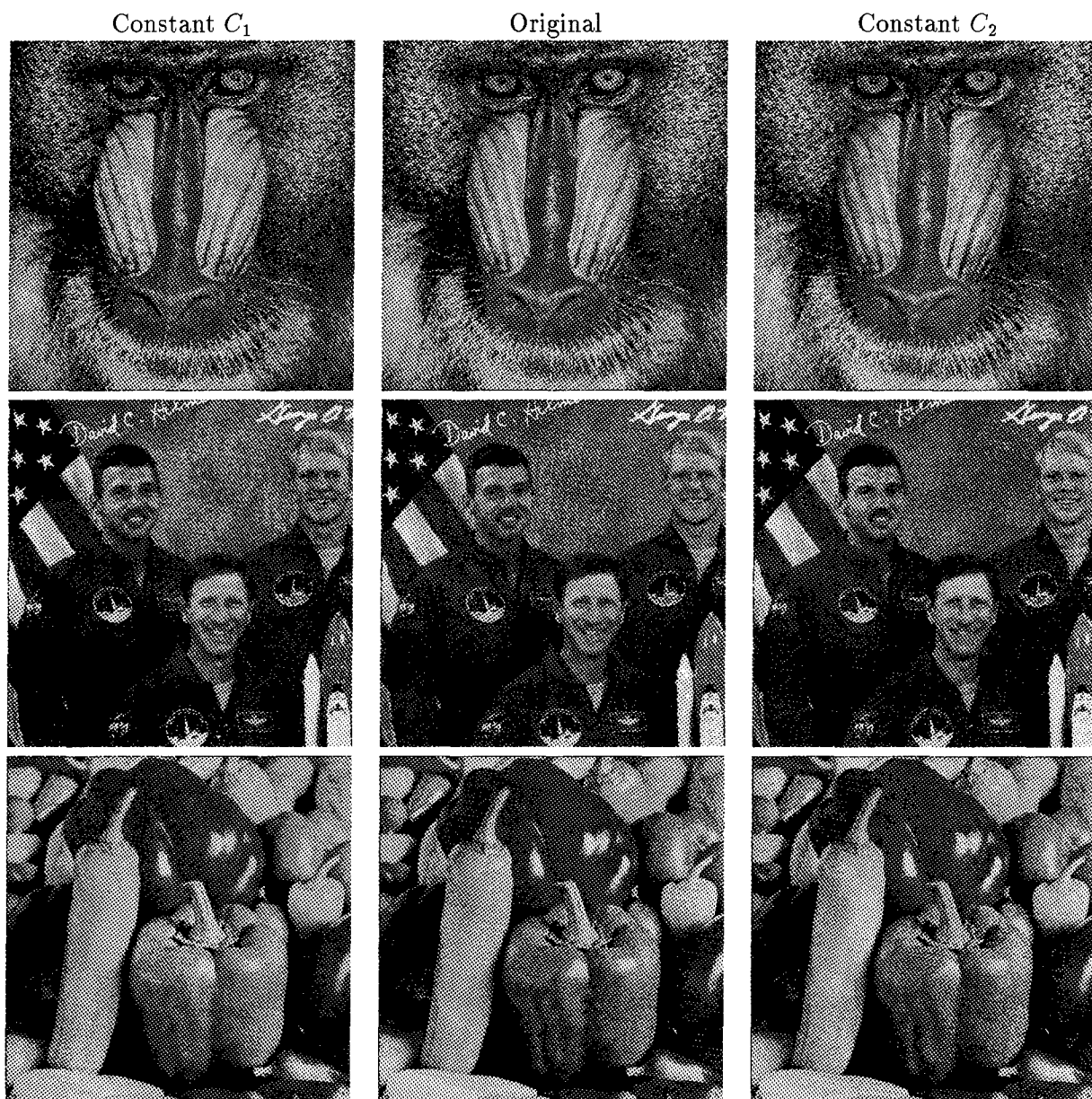


Figure 5. Images used in the image preference experiments. Top: Mandrill, Center: Astronauts, Bottom: Peppers. The Mandrill and Peppers images are from the USC-SIPI image database and are used by permission. The astronaut image is cropped from a NASA photo of the STS-26 crew.

of the nose in the constant C_1 MANDRILL picture as darker *red*, while normal trichromats describe it as dark *green*. All three subjects indicated that there was little perceptible difference between the original and the constant C_1 ASTRONAUT image. While they were clearly able to identify the distorted version of the PEPPERS image, the subjects still had difficulty naming the colors in it.

The constant C_2 comparisons provided equally interesting results. When asked to indicate a preference between the original and the constant C_2 ASTRONAUT images, both subjects A and B immediately stated that the constant C_2 image appeared to be monochrome, or black-and-white. Interestingly, subject C was able to accurately describe the colors in the constant C_2 ASTRONAUT image, despite his poorer performance in the Dvorine test. In the MANDRILL images, the impression of monochrome was still present for subjects A and C, although not as strong as in the ASTRONAUT case. As in the constant C_1 case, all three subjects were easily able to identify the constant C_2 version of the PEPPERS image.

2.4.5 Discussion. The image preference experiments validate the conjecture offered in Section 2.4.3. In the case of the constant C_1 images, all three subjects had a very difficult time distinguishing between the original and distorted MANDRILL and ASTRONAUT pictures, and although there was clearly a preference for the original PEPPERS image, the subjects had significant difficulty naming the colors in the constant C_1 version. This indicates that the subjects' ability to detect alterations in the C_1 component of the model is indeed impaired, although the fact that they could see a difference in the PEPPERS image suggests that the deficit cannot be characterized by a simple omission of the C_1 channel.

The results of the experiments with the constant C_2 images also largely agree with the predictions, particularly those involving the ASTRONAUT image. Without being solicited, the immediate reaction of subjects A and B to the constant C_2 image was that it looked monochrome, or black and white. In the constant C_2 MANDRILL image, neither subject A nor subject C was able to identify more than one colored area, while at least

Table 3. Summary of image preference experiments.

Constant	Image	Subject	Preference	Observations
C_1	Mandrill	A	Distorted	Images close, nose darker red in distorted image.
		B		Not shown this image.
		C	Distorted	Very small differences; nose darker red and cheeks brighter in distorted image.
	Astronauts	A	Original	Images close; normal brighter, richer.
		B	Distorted	Not much noticeable difference.
		C	None	Could not distinguish between the two.
	Peppers	A	Original	Difficult to name distorted colors w/o original.
		B	Original	Large red pepper changed color the most.
		C	Original	Distorted image intriguing—large green pepper looked pink.
C_2	Mandrill	A	Original	Only color evident is pink nose.
		B		Not shown this image.
		C	Original	Cheeks light green, nose and fur same color in distorted image.
	Astronauts	A	Original	Distorted image looked monochrome.
		B	Original	Distorted image looked black and white.
		C	Original	Green or brown uniforms in distorted image.
	Peppers	A	Original	Original appeared brighter/better.
		B	Original	Distorted looked like same image under yellow illumination.
		C	Original	Distorted looked old and tasteless.

three different colors (pink on the nose, light green on the sides of the nose, and red in the eyes) are evident to normal trichromats.

While the ASTRONAUT and MANDRILL images seemed to produce results most favorable to the conjecture that was made regarding the color perception of the three subjects, the PEPPERS images may provide the most insight. First, the fact that all three were able to easily distinguish the C_1 -altered image from the original suggests that the dynamic range of the C_1 component of the PEPPERS image is great enough that the distortions were of sufficient magnitude to be perceived. The distortions introduced by setting the C_2 component equal to the average value do not appear to be very large, as the test subjects were as able as normal observers to identify the alterations. Lacking variation in C_2 , there is apparently sufficient variation in the C_1 component that the subjects are still able to identify the colors in the constant C_2 image. And, since the subjects are clearly able to perceive changes in C_2 more easily, it is not really surprising that they could identify the color shift between the two.

One of the most interesting results of these experiments is the indication of what severe distortions can be tolerated by the deuteranomalous subjects. The most dramatic illustration of this is evident in the three ASTRONAUT pictures. To a normal trichromatic observer, the constant C_1 image seems closer to the original than does the constant C_2 version, but to the test subjects, exactly the opposite is true. For example, the red stripes of the flag appear redder to a normal observer in the constant C_2 image than in the constant C_1 image. But to subject A, the stripes in the constant C_2 image only appear dark—not red at all. Similarly, both subjects A and C identified the nose in the constant C_1 version of the MANDRILL as darker red than the original, while normal observers perceive it as dark green in color. The constant C_1 and C_2 images do not give normal observers color blind eyes, but they do yield hints about the dramatic distortions that color blind individuals are truly blind to.

Finally, note that both the color contrast and image preference experiments seem to indicate that the color deficit (the “lost” or severely reduced axis) in the color blind observers seems to be in (A, C_1, C_2) space, one step beyond the stage where the physiolog-

ical measurements indicate it would be (the (L, M, S) cone absorption space). Why this should happen is not immediately apparent.

2.4.6 Summary. This section has presented the results of experiments performed with deuteranomalous trichromat observers, using the Faugeras color HVS model. Two experiments were performed with three subjects: chromatic change detection and image preference. In both experiments, stimuli were produced by either creating or modifying an image in the perceptual space of the Faugeras color HVS model, then transforming the result into the appropriate space for display on a color monitor. Using the color Mach band stimulus, the chromatic variation experiment assessed whether or not the subjects could detect variations in either of the two chromatic components of the model's perceptual space, while the image preference experiment determined whether they could identify changes in either of the chromatic components in a complex image.

The chromatic variation experiment showed that the three observers could not distinguish between colors that only differed from each other in the C_1 component, but that they could correctly identify colors that only differed in their C_2 values. Based on this result, two hypotheses were proposed. First, it was hypothesized that the subjects would not be able to distinguish an original image from one in which all C_1 variation had been removed. Second, it was surmised that they would not be able to see any color in images from which all C_2 variation had been removed.

The results of the image preference experiment largely support these hypotheses. When asked to choose between the original and a constant C_1 version of three different images, the subjects showed a clear preference for the original of only one of the images. With both of the two other images, they either chose the modified version outright, or had a difficult time choosing a preference between the two versions. When given a similar choice between the original and a constant C_2 version of the same three images, the subjects showed a clear preference for the original, describing the altered versions as either completely monochrome or significantly reduced in color variation.

Significantly, the images used in the experiments were neither created nor manipulated in the portion of the model corresponding to cone outputs. Rather, the operations

were performed in the space corresponding to one stage later in the optic tract—the outputs of the LGN. Since deuteranomalous trichromacy is attributed to a weak M cone output, it is remarkable that manipulating the model analogs of the LGN outputs accounts so well for the deficiency in the color blind subjects.

2.5 Summary

The experiments reported in Sections 2.3 and 2.4 lead to several important conclusions. First, the experiments with the color blind observers provide validation of the Faugeras color HVS model from a unique perspective. These results represent the first time that an engineering HVS model has been intentionally tested with color blind observers. The close correspondence between the color channels and the deficits of the color blind subjects demonstrated by these experiments is most remarkable.

The second important conclusion stems from the color Mach band experiment. Despite a prevailing belief to the contrary, this experiment clearly demonstrates that color Mach bands do exist. Second, the appearance of color Mach bands in a Mach stimulus created in the chromatic channels of the HVS model provides evidence for including low spatial frequency attenuation in these channels of the model. In his measurements, Faugeras found low spatial frequency attenuation in the chromatic channels [25], but few other researchers have corroborated his findings. In fact, several researchers have concluded that there is no low spatial frequency attenuation in the chromatic channels (for example, see [95]). The color Mach bands produced in these experiments provide additional support for Faugeras' findings.

With the support of the results from these two experiments, the Faugeras color model, with the spatial filters proposed by Hall, is chosen for use in the perceptual image fidelity measure. In Chapter IV, this model is combined with a multiple-channel HVS model based on measurements of the response of simple cells in the primary visual cortex to produce a multiple-channel color HVS model for assessing perceptual fidelity. First, though, Chapter III provides additional background and verification for the multiple-channel approach.

III. Multiple-channel HVS Model Validation

3.1 Introduction

The previous chapter considered color aspects of the human visual system. In this chapter, recently advanced multiple channel achromatic HVS models are considered. Recent physiological evidence suggests that the primary visual cortex performs a space/spatial frequency analysis of the input scene, distributing the information in the scene among multiple channels which are tuned to respond selectively to different spatial frequencies and orientations [14, 46, 70]. This chapter provides a review of the physiological basis for these models, and examines one such model by looking at its response to illusory contour stimuli. The significant result is that processing of these stimuli by a multiple-channel model gives rise to output patterns which may be interpreted as the formation of illusory contours.

Section 3.2 reviews the development of achromatic HVS models, outlining the salient psychophysical and physiological results that motivate the models. Then, Section 3.3 presents the results of the illusory contour experiment.

3.2 Achromatic HVS Models

3.2.1 Early Vision Models. One of the early efforts using an HVS model to process images in a perceptual domain was described by Stockham [87]. In his work, Stockham used an HVS model composed of a logarithm followed by a linear filter, as depicted in Figure 6. Stockham based his model on an analysis of both the physics of image formation and the neurophysiology and psychophysics of vision [87]. Using results of contrast sensitivity experiments to specify the linear filter, several researchers have used this same structure in a number of different image processing applications [36, 64, 65, 87].

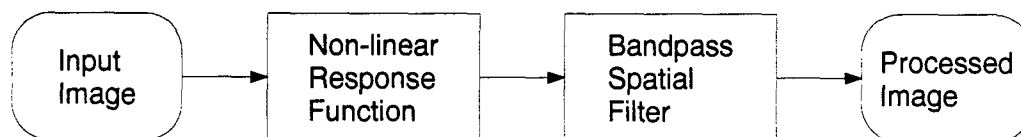


Figure 6. Simple achromatic HVS model.

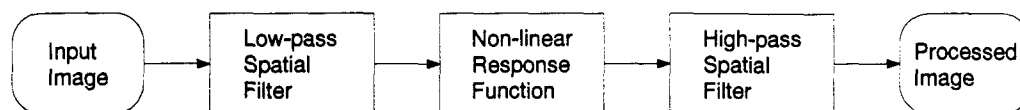


Figure 7. "More anatomically correct" HVS model [37].

Because of its simplicity, Hall and Hall found the model of Figure 6 to be inadequate [37]. Recognizing that high frequency attenuation due to eye optics and spatial sampling effects of the retina actually occurs in the visual system before the non-linear response of the retinal ganglion cells, they proposed a model in which the low pass spatial filtering is performed before the non-linearity is applied, leaving just a high pass filter after, as shown in Figure 7. In light of the structure of the color vision system shown in Figure 1, it may be noted that if an achromatic image is presented to the color system, the chromatic channels become zero, leaving active only the brightness channel (A), which has the structure shown in Figure 7. Hall and Hall performed in-depth analysis of this model [37], producing a fundamental result that the more complicated model provided a mechanism to control the high frequency rolloff of the overall system function based on input contrast [36]. However, in a subsequent study with typical complex images, the differences between these two models were not sufficient to warrant the use of the more complex model [35]. The majority of physiologically motivated models of the early vision system use the simpler structure reflected in Figure 6.

3.2.2 Multiple-channel "Cortical" Models. Even as early vision models were being employed in early perceptual image processing applications, psychophysical and neurophysiological evidence was building to suggest that the HVS contains channels tuned to spatial frequencies [7,9,41-44]. In fact, these initial results actually helped to motivate the use of Fourier analysis techniques in modeling the HVS [37].

Over the past two decades, the concept of multiple, spatial frequency tuned channels in the HVS has continued to mature. One model that has drawn considerable attention uses Gabor functions (Gaussian modulated sinusoids) to describe the measured receptive fields of simple cells in the visual cortex [46]. The development of this model may be traced through a series of papers, ending with results by Jones and Palmer, in which they found

that the receptive field profiles of a great majority of cat simple cells that they measured could be well fit by two-dimensional Gabor functions [14, 15, 46, 49, 55, 76].

Motivated by these results, Gabor models have been used successfully in a diverse range of applications, including image analysis and compression [16], face recognition [50], and forward-looking infrared image segmentation [2]. Several other HVS models have been developed which, while satisfying other constraints such as computational speed and interpolation capability, retain multiple channels with at least approximately Gabor shaped filters [73, 102, 103]. These have also found use in perceptual image fidelity measures for monochrome images [13, 91, 104].

Before proceeding, it should be acknowledged that despite the close fits obtained by Jones and Palmer and the success of the models based on Gabor filters, there may be other descriptions that perform equally well or better. Based on their analysis, Stork and Wilson conclude that the evidence is insufficient to favor Gabor functions over these alternative models, which include difference of Gaussians, difference of offset Gaussians, and derivatives of Gaussians [88]. The purpose of the present investigation is not to debate this point. Rather, recognizing their prior success, Gabor (and Gabor-related) models are accepted here as providing a reasonably close account of the spatial processing of the HVS, and examines the results of processing illusory contour stimuli with such a model.

3.3 Illusory Contour Formation by a Multiple Channel HVS Model

3.3.1 Introduction. Human vision researchers have been intrigued by visual illusions for many years because they provide valuable insights into the operations performed by the visual system. In addition, they provide a unique means for validating human visual system models—good HVS models are expected to account for the illusory effects. This section uses three visual illusions to examine a Gabor-based multiple channel HVS model. The processing of the multiple channel model gives rise to contours in the output that correspond to the illusory contours in the input illusions. These results support the notion that illusory contours are perceived as a consequence of processing applied by cortical neurons to the input visual stimulus.

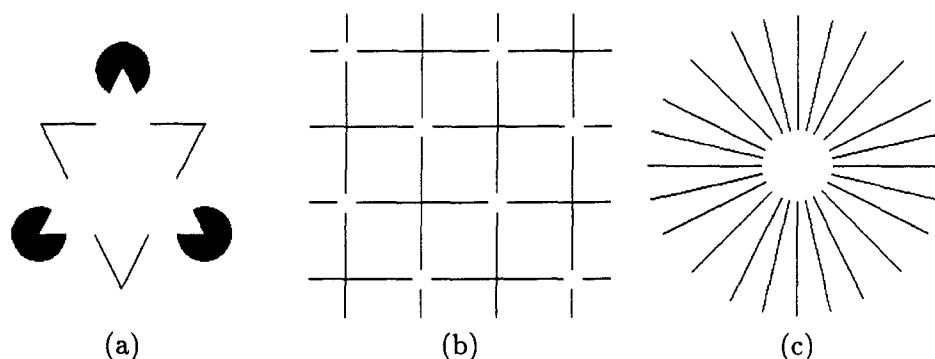


Figure 8. Illusory contour illusions: (a) Kanizsa triangle [47], (b) Ehrenstein "grid" illusion [23], and (c) "sunburst" illusion [23].

3.3.2 Methods. Three illusory contour illusions are examined: the Kanizsa triangle (Figure 8a) and the Ehrenstein "grid" and "sunburst" patterns (Figures 8b and 8c). Though created with simple black-and-white stimuli, each of these images contains white illusory objects that actually appear to be brighter than the white background. Because of the perceived difference in brightness between the illusory object and the white background, this white-on-white edge is referred to as an illusory contour.

The HVS model used to process the illusions is obtained by following the simple model shown in Figure 6 with a bank of Gabor filters. The logarithm is used for the non-linearity shown in Figure 6, while the bandpass spatial filter is taken from the contrast sensitivity function (CSF) given by Mannos and Sakrison [54]. Following Daugman [16], the Gabor filter bank is composed of a total of 60 Gabor filters: 30 pairs in conjugate phase, ("sine-Gabors" and "cosine-Gabors") uniformly distributed on a log-polar grid, with radial spatial frequency centers of 8, 16, 32, 64, and 128 cycles per image, and orientations of 0° , 30° , 60° , 90° , 120° , and 150° with respect to the horizontal. Each filter has a 1.5 octave radial spatial frequency bandwidth and a 30° orientation bandwidth, as depicted in Figure 9.

In order to observe the overall effect of the analysis of the visual stimulus by the bank of Gabor filters, the individual filter outputs are added together to produce a single, composite system output. This is not meant to imply that the visual system actually combines the outputs of the various spatial frequency channels in this way—a complete

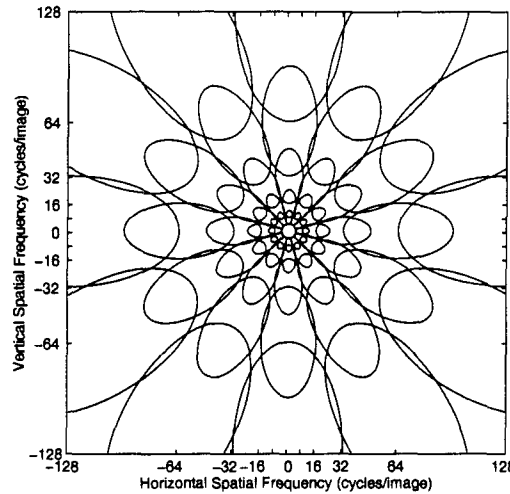


Figure 9. Gabor filters used in processing illusory contour stimuli. The ellipses represent $1/e$ levels of the 2-D Gaussian functions which compose the Gabor filters. So-called cosine Gabors are obtained by adding two Gaussians with the same radial spatial frequency center that are separated by 180° in orientation, while sine Gabors are obtained by taking the difference of the two.

understanding of how the different channel outputs are used in forming the perception of the object in view does not presently exist. Forming a linear combination is just a simple approach to assembling the outputs of the various filters to get an idea of the information that is preserved and the artifacts that are introduced by the filter bank processing. Additionally, the linear combination approach allows the analysis and reconstruction steps to be realized with a single filter, obtained by summing the impulse responses of all the individual filters in the filter bank.

In previous work, Ginsburg credited the low spatial frequency attenuation of the CSF with aiding in the formation of the illusory triangle [31]. By examining the outputs of the HVS model with and without the CSF component, the effects of the CSF can be more carefully noted.

3.3.3 Results. Figure 10 shows the results of applying just the Gabor filter bank to the three illusions of Figure 8. In all output images, the outputs are scaled to the full brightness range of the display device. Filtering the illusory images with both the CSF filter and the composite Gabor filter produces the results shown in Figure 11. Comparing

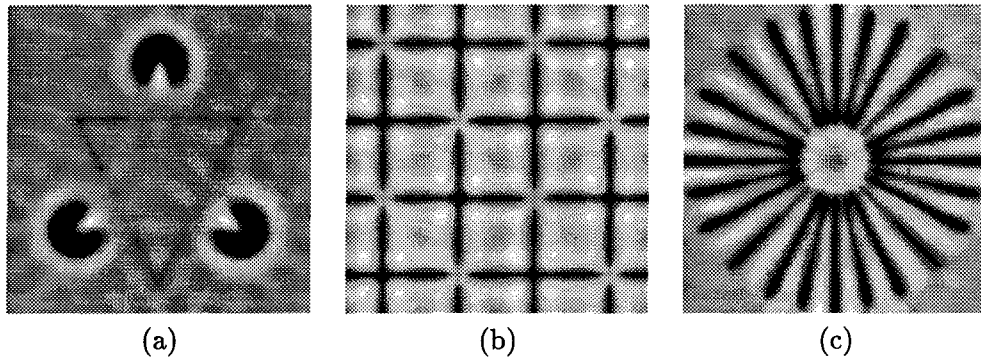


Figure 10. Illusions processed by the Gabor HVS model: (a) Kanizsa triangle, (b) Ehrenstein illusion, and (c) sunburst illusion.

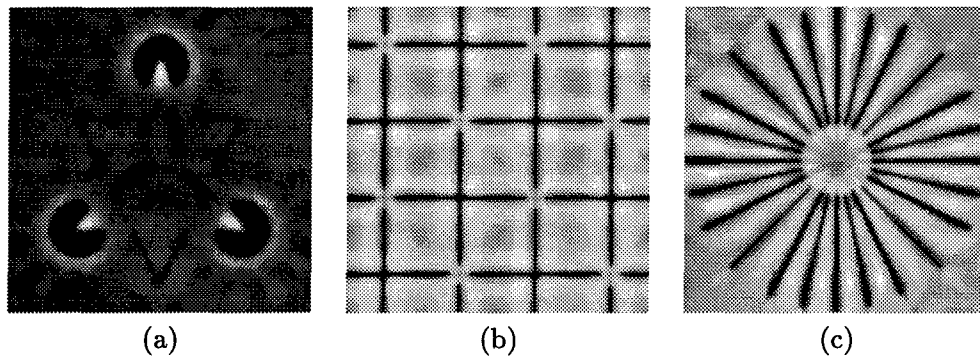


Figure 11. Illusions processed by the CSF/Gabor HVS model: (a) Kanizsa triangle, (b) Ehrenstein illusion, and (c) sunburst illusion.

Figures 10 and 11, it is apparent that the CSF filter principally serves to sharpen the edges in the image, including the illusory contour edges.

Note in Figures 10 and 11 that the HVS model output has clearly evident regions of increased relative response at the locations where the illusory contours appear in the original stimuli. Given the appearance of the illusions in the binary images, it is easy to think that these regions of increased brightness only appear to be brighter—that the illusion is occurring in the HVS outputs. However, by plotting horizontal or vertical cross-sections of the actual normalized outputs, it is easy to show that the brightness actually does increase relative to the surroundings at the locations of the illusory contours. In Figure 12, selected rows or columns from the CSF/Gabor processed illusions are plotted to show that the increase in relative brightness is actually present. The features in these

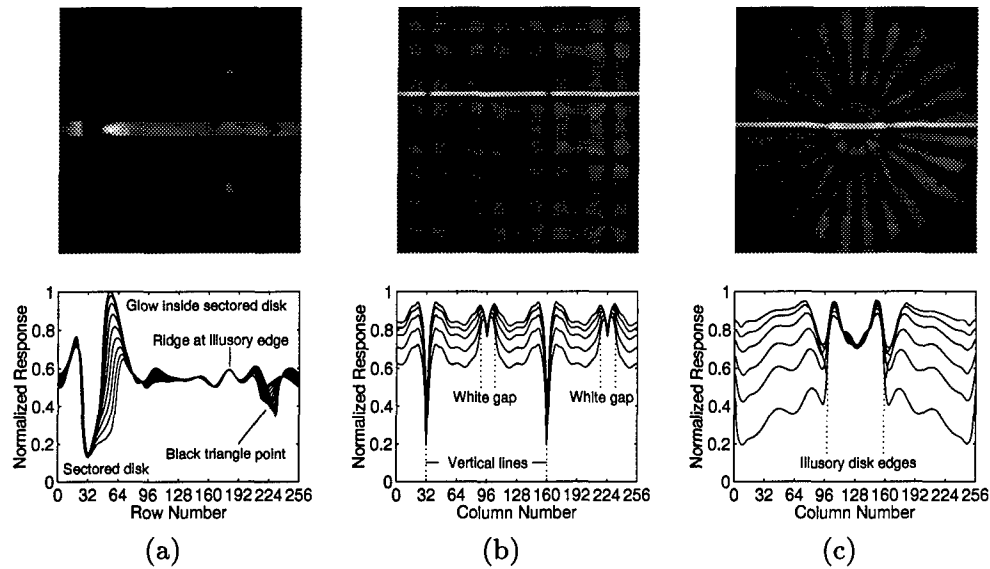


Figure 12. Plots of adjacent rows or columns in the CSF/Gabor HVS model outputs for the three input illusions: (a) columns 120-135 of the Kanizsa triangle, (b) rows 90-94 of the Ehrenstein illusion, and (c) rows 123-128 of the sunburst illusion. The plotted rows or columns are highlighted in the images above the plots.

plots clearly show that the output is increased at the locations where the illusory contours appear. In the Kanizsa triangle plot, a clearly defined ridge shows the increased response at the location of the illusory edge. The fact that the ridge occurs at the same row number and reaches the same amplitude in all 18 columns demonstrates the straight-line nature of the illusory contour. Similarly, the response in the Ehrenstein plot increases at the same place where the illusory edges appear in the stimulus, at the edges of the gaps. Finally, in the case of the sunburst illusion, while going from the outer edges to the center, the response first decreases, then increases at the locations where the central lines end. This serves to highlight the illusory edge, making the central disk appear brighter [12]. Because the illusory contour is a circle, the central peaks in Figure 12c would be expected to move toward the center. This is not observed in Figure 12c because only six rows are plotted for simplicity. As more lines are plotted, the expected shift of the central peaks is observed.

3.3.4 Discussion. As mentioned earlier, Ginsburg credited the low spatial frequency attenuation of the CSF with aiding in the formation of the illusory triangle [31].

The present study shows that the primary operation of the CSF is to highlight and sharpen edges, while Gabor processing actually produces illusory edges. With edges established, the HVS seems to have a mechanism for "filling in between the lines." This mechanism is perhaps best demonstrated by the Cornsweet effect [12]. Several engineering models of the HVS have been proposed which include such a mechanism (see, for example, [33, 34]). Thus, it may be argued that Gabor-like processing occurring in the simple cells of the visual cortex provides the edges, both illusory and real, to this "filling in" mechanism, which produces the uniform, bright shapes evident in the three illusions considered here.

3.3.5 Summary. In summary, this study has demonstrated a Gabor HVS model that produces illusory contours in its output. While some researchers have questioned the appropriateness of a Gabor scheme for modeling visual cortical simple cells [88], it is nonetheless true that Gabor functions do provide at least a close approximation to the receptive field profile of simple cells. The significance of this demonstration is that Gabor HVS model processing of illusory figures produces outputs in which the illusory contours emerge.

3.4 Summary

This chapter has considered an aspect of multiple-channel achromatic HVS models which has not been considered previously: how the model accounts for illusory contours. It has been shown that illusory contours arise as a consequence of processing when illusory contour stimuli are passed through a simple Gabor HVS model. When the Gabor model is combined with a bandpass function to account for the contrast sensitivity of the HVS, the edges in the output are sharpened, including the illusory edges. Together with successful application of Gabor-like multiple-channel models in perceptual image fidelity measures for achromatic digital images, this evidence serves to recommend the adoption of a multiple-channel analysis stage in a color image fidelity measure. A color HVS model which includes such a multiple channel structure is the subject of the following chapter.

IV. Multiple-channel Color Image Fidelity

4.1 Introduction

Thus far, attention has been focused on specific aspects of HVS models, providing some unique approaches to verifying the models which have been proposed. In this chapter, the problem of assessing perceptual image fidelity of digital color images is addressed. The model components explored in the previous two chapters will be combined into a single, multiple-channel color HVS model to produce a new measure for perceptual digital color image fidelity.

Section 4.2 reviews previous efforts to define a perceptually meaningful fidelity measure, and Section 4.3 describes the approach proposed in this dissertation research. Section 4.4 demonstrates the performance of this new fidelity measure using a few color images distorted in various ways. Section 4.5 discusses the significance of these results, and provides an analysis of the strengths and weaknesses of the measure.

4.2 Previous Image Fidelity Measures

4.2.1 Introduction. In order to put the new multi-channel color fidelity measure into context, it is important to appreciate the measures that have been explored to this point. First, perceptual fidelity measures for achromatic images based on early vision models are briefly discussed. Then, the use of multiple-channel HVS models in achromatic fidelity measures is considered. Finally, previous efforts to measure perceptual fidelity for color images are summarized.

4.2.2 Early Vision-based Achromatic Image Fidelity Measures. Since the question of measuring perceptual image quality was first considered, a multitude of approaches have been proposed. Among these are several variations on the mean square error theme. These approaches essentially transform the two input images into some kind of perceptual space and compute the mean square error between them in that space. Recently, Eskicioglu evaluated many proposed fidelity measures in the context of an image compression application [24]. Using four different compression algorithms, he compressed three

different images, then computed the perceptual fidelity of the twelve resulting distorted images using the various fidelity measures. Then, he compared the resulting measures to assessments made by human observers.

The HVS models used in the different fidelity measures tested in the Eskicioglu study were limited to early vision models, comprised of a non-linearity and a contrast sensitivity-based spatial filter. The best overall performance in the study was achieved by a normalized, weighted mean square error metric computed in the cosine transform domain, using the CSF to provide the perceptual weighting of the transform coefficients. However, the authors noted that the inclusion of a CSF-based model does not always improve correlation with human assessments [24].

4.2.3 Multiple-channel Achromatic Image Fidelity Measures.

4.2.3.1 Overview. Recent attempts to measure perceptual fidelity differ from earlier approaches in two significant ways. First, the models used in the newer approaches account more completely for the physiological and psychophysical data that has been collected. In addition to the non-linearity and CSF stages found in the early models, the newer models include a multiple channel cortical stage, in which the image is analyzed by a bank of filters tuned to narrow ranges of orientation and radial spatial frequency. As discussed in Chapter 3.3, this approach is motivated physiologically by the finding of cells in the primary visual cortex that are selective to stimuli of specific orientations and spatial frequencies. The recent models also include mechanisms to account for visual masking effects that have been measured through many recent psychophysical studies [26, 27, 30, 51, 72, 75, 89, 93, 97].

The second difference between recent and earlier approaches to measuring fidelity is a tendency toward producing a map of the perceivable differences rather than a single number representing overall fidelity. Because they identify locations in the image where differences can be perceived, these maps provide more insight to image researchers in a number of applications. With such a map, researchers can focus their attention on methods for reducing the visibility of errors in the regions identified in the map. The relationship

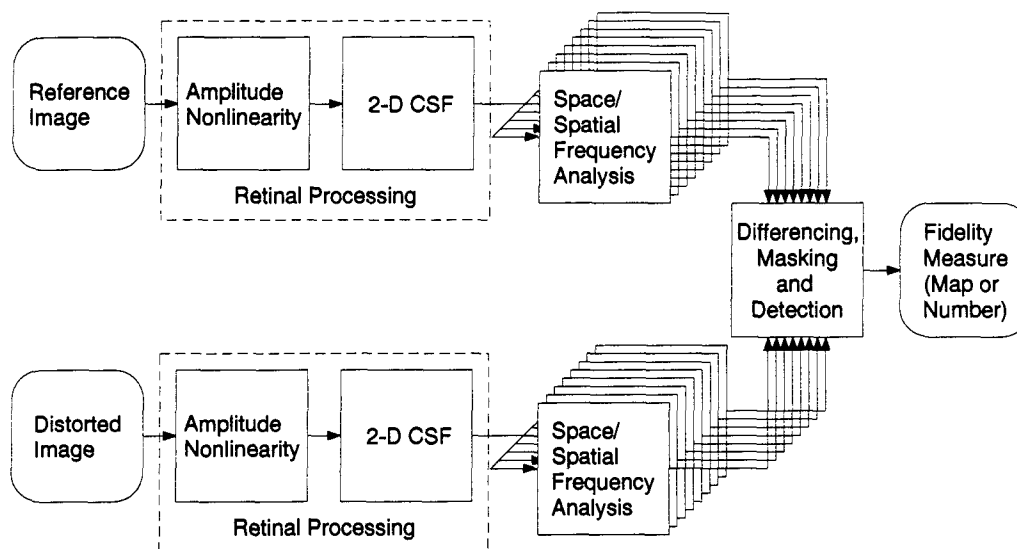


Figure 13. Common structure of recent multiple-channel perceptual image fidelity measures.

between visible difference maps and single-number fidelity measures remains an active area of research [13]; the present research follows the recent trend and produces a visible difference map.

This section considers three recent approaches to measuring perceptual image fidelity: the Visible Differences Predictor (VDP) proposed by Daly [13], a perceptual image fidelity measure by Westen *et al.* [104] which follows much of Daly's work, and a third approach by Heeger and Teo [38], which they call a perceptual image quality measure (because it is a relative quality measure, the Heeger and Teo approach is referred to here as a fidelity measure to maintain consistency with the convention established in Chapter I). While some implementation details differ between them, these three approaches share a common basic processing structure, depicted in Figure 13. The images are first processed by a retinal model that generally consists of non-linear amplitude processing and linear filtering. Following the retinal stage, the images pass through a cortical stage, which consists of a space/spatial frequency transform followed by differencing, masking, and detection operations. As they will be incorporated into the color fidelity measure, the processing elements of the cortical models used in these three models will now be reviewed.

4.2.3.2 *Space/Spatial Frequency Analysis.* The first step of the cortex stage is a space/spatial frequency transform, which analyzes the output of the retinal stage into local spatial-frequency, orientation, and phase components. This is accomplished using a hierarchical bank of spatially oriented band-pass filters, which are tuned to match measured spatial frequency selectivity of simple cells in the visual cortex, as discussed above in Section 3.3. Thus, passing through this filter bank, the retina model-preprocessed images are expanded into a set of filtered images, which Daly refers to as cortex bands [13]. For simplicity, the subbands in all three approaches will be referred to here as cortex bands.

The actual functions used to describe the filter bank in the three approaches vary: both Daly and Westen *et al.* use an approximation to the Gabor transform called the Cortex transform [103], while Heeger and Teo use a steerable image pyramid approach [73]. However, the bandwidth and center frequency specifications of the filter banks are similar, arising from experimental evidence [4, 20, 42, 66, 74, 89]. In radial spatial frequency, the subbands have one octave bandwidth, with center frequencies equally spaced on a log scale. In orientation, the subband centers are equally spaced at thirty degree intervals, with a uniform thirty degree bandwidth. Teo and Heeger also produce two subbands in quadrature phase for each orientation and radial spatial frequency pair [38].

In order to use psychophysical measurements of masking effects, the outputs of the cortex bands are expressed in terms of contrast by Daly [13] and Westen *et al.* [104]. This is accomplished by dividing the output of each subband either by a low-pass filtered version of the retina-preprocessed image (pixel by pixel) to obtain a measure of local contrast, or by an average (DC) value of the low-pass filtered image for a global contrast measure. Expressed in contrast units, the resulting subband outputs are passed to the contrast masking stage.

4.2.3.3 *Masking.* Masking refers to a reduction in the ability to detect a target pattern when it is superimposed on a background pattern, as illustrated in Figure 14. These effects are dependent upon both the contrast and the spatial frequency content of the masking signal, relative to that of the signal of interest. For a given spatial frequency, as the contrast of the masking signal increases, it becomes more difficult to detect the

is discriminant function is that usually the decision boundary lies in the tails of the

As a result, a large number of samples must be used in order to ensure that enough samples are found near the decision boundary to obtain the necessary resolution on the discriminant function. Several "rules of thumb" have been proposed for the number of samples required to obtain accurate performance from a neural network. For example, one rule is that the number of samples required to reach an accuracy level equal to 10 times the number of weights in the network [1:153].

One final comment is worth mentioning before leaving the subject of neural network approximation of a posteriori probability functions. When using a technique that relates the conditional pdfs to estimate the Bayes error rate, an assumption must be made about the a priori probability of each class in order to convert the pdf estimates to posteriori estimates. In a neural network approach, the a priori information is obtained directly from the proportion of data from each class presented to the net. Finally, the neural network learns the a posteriori probabilities without assuming a strict functional

Figure 14. Example of contrast masking. The print is most difficult to read when the spatial frequency of the grating matches the predominant frequencies in the print. (Figure adapted from [1].)

signal of interest. This dependence of the detection threshold as a function of masking contrast for a signal of a fixed spatial frequency is frequently described by a threshold elevation function.

A number of different experiments have been performed to measure masking effects using a wide variety of target and background patterns, including sinusoidal grating patterns and static and dynamic noise. Daly provides an excellent review of these experiments, unifying the results of the different experiments to produce a single expression of the threshold elevation function. As a function of normalized contrast of the masking signal m_n , Daly expresses the threshold elevation due to masking as [13]

$$T_e(m_n) = (1 + (k_1(k_2 \cdot m_n)^s)^b)^{1/b}, \quad (7)$$

where T_e is the threshold elevation, and k_1 , k_2 , s , and b are parameters which control the shape of the function. This function describes the increase of the detection threshold for a target spatial frequency as the contrast of the background masking pattern increases above that of the target. Both Daly and Westen *et al.* use this functional form to model masking effects, with very little change between the subbands of their models.

An example of the threshold elevation function described by Equation 7 is shown in Figure 15. The knee point of the curve is controlled by the parameters k_1 and k_2 in Equation 7, while the non-zero, high contrast slope is controlled by the parameter s . The knee point parameters are described in terms of two other parameters, W and Q , by

$$\begin{aligned} k_1 &= W^{(1-1/(1-Q))} \\ k_2 &= W^{(1/(1-Q))}. \end{aligned} \quad (8)$$

The parameter W represents the detection signal to noise ratio from noise masking experiments, while Q is the high-contrast slope of the T_e function when the two asymptotes of the function intersect at a value of 1.0 on the normalized mask contrast axis [13]. Finally, the parameter b in Equation 7 arises from a generalization by Daly of a common noise masking model [75].

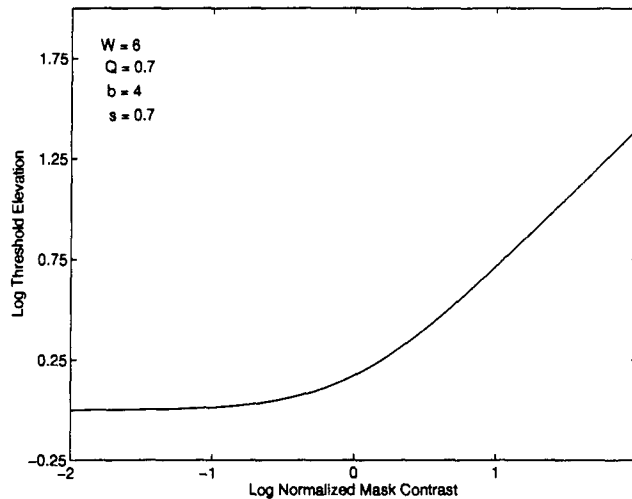


Figure 15. Sample threshold elevation function described by Equation 7. The parameter values are as shown.

It is important to note that masking effects are image dependent, and when assessing perceptual fidelity, distortions and image content can interchange roles as target and masker. Therefore, both Daly and Westen *et al.* use Equation 7 twice to compute two masking thresholds for each location in a subband, deriving the m_n signal from the two input images in the two computations. Then, the actual masking threshold is obtained by taking the minimum of these two thresholds. This symmetric approach proves to handle a wide range of different types of image distortion. The resulting threshold elevation mask is applied by dividing the difference in contrast at each location in a subband by the masking threshold value for that location. Thus, locations where the contrast difference is large and the threshold is small are likely to be identified as locations where distortions can be detected.

In contrast to the approach used by Daly and Westen *et al.*, there is no explicit masking mechanism included in the model proposed by Heeger and Teo [38]. However, contrast masking data forms the basis for calibrating the cortical stage of their model. For each input image, the cortical stage of the Heeger and Teo algorithm produces a local normalized energy measure for each location in each orientation subband. Heeger and Teo use divisive normalization to provide contrast gain control [38].

4.2.3.4 Computing the Detection Map or Error Measure. With masking

effects accounted for, the remaining task for the three measures discussed here is to recombine the information in the various subbands to produce the final output: an indication of the perceptual fidelity. As mentioned above, this final output may be in the form of an image map, where visible differences are identified, or it may be a single number which represents an overall fidelity assessment. The methods used by the three approaches to produce this final output will now be outlined.

The simplest of the three approaches is a single-number perceptual error measure (*PEM*) proposed by Westen *et al.* [104]. This measure is given as a generalized mean square error:

$$PEM(I_1, I_2) = \left[\sum_{i,j} \left| \sum_{k,l} |\Delta MLBC_{k,l}[i,j]|^\alpha \right|^\beta \right]^\gamma, \quad (9)$$

where I_1 and I_2 represent the two input images, and $\Delta MLBC_{k,l}[i,j]$ represents the masked difference in local band-limited contrast for the k, l th Cortex band, computed from the two input images using the masking threshold function described in the previous subsection. The parameters α , β , and γ are used to fit the performance of this measure to assessments made by human observers on a set of test images. With values of $\alpha = 1.5$, $\beta = 1.0$, and $\gamma = 0.33$, the Westen *et al.* perceptual error measure achieved a correlation coefficient of -0.84 with the human assessments, compared with a correlation coefficient of 0.78 achieved by a conventional peak-to-peak signal to noise ratio approach [104].

In contrast to the single-number *PEM*, the VDP algorithm produces a detection map showing locations in the image where distortions can be perceived. This map is created using a psychometric function and the technique of probability summation. The psychometric function, which describes the probability $P(c)$ of detecting a signal of contrast c , is given by Nachmias [62] as

$$P(c) = 1 - \exp(-(c/\alpha)^\beta), \quad (10)$$

where the parameters α and β control the location and slope of the transition region of the psychometric function (see Figure 16). The slope parameter β has been found to be

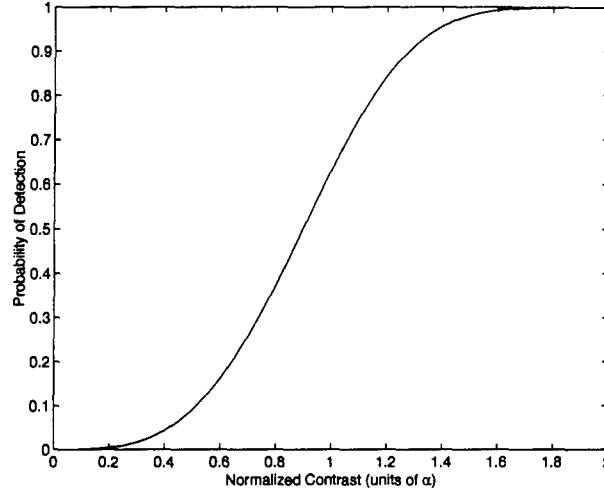


Figure 16. Psychometric function used in the VDP to compute the probability of detecting a signal of contrast c , relative to a masking threshold α .

roughly constant [59], and the location parameter α corresponds to the detection threshold described in the previous section. Thus, the probability of detecting a distortion at each location in a cortex band is computed by substituting the ratio of the contrast difference to the masking threshold at that location for c/α in Equation 10. This operation yields a probability of detection image $P_{k,l}[i, j]$ for each cortex band

$$P_{k,l}[i, j] = 1 - \exp \left[\left(\frac{\Delta C_{k,l}[i, j]}{T_{em}^{k,l}[i, j] \cdot T(0)} \right)^\beta \right], \quad (11)$$

where $\Delta C_{k,l}[i, j]$ is the difference in contrast and $T_{em}^{k,l}[i, j]$ is the threshold elevation mask for the (k, l) th cortex band, and $T(0)$ represents a threshold due to internal noise of the visual detection mechanisms. Because of the calibration techniques used, $T(0) = 1.0$ in the VDP [13].

The cortex band probability of detection images are combined into a single overall probability of detection image through the technique of probability summation. This operation is realized as a product series:

$$P_t[i, j] = 1 - \prod_{k,l} (1 - P_{k,l}[i, j]), \quad (12)$$

where $P_t[i, j]$ represents the total probability of detection image [13]. Pixel values close to one in this image indicate locations where the distortions are above the visibility threshold, while values close to zero identify distortions below threshold.

In order to retain information about whether an error is lighter or darker than the original image, Daly computes a sign for each image location. For each cortex band, the sign of the contrast difference at each location, $\Delta C_{k,l}[i, j]$, provides the sign for that location. For the overall probability of detection image, the sign for each location is determined as the sign of a weighted sum of the signs from the various cortex bands at that location, using the cortex band probabilities of detection as the weights. Thus, the signed probability of detection image for the k, l th cortex band, $P_{k,l}[i, j]$, may be expressed as [13]

$$SP_{k,l}[i, j] = \text{sign}(\Delta C_{k,l}[i, j]) \cdot P_{k,l}[i, j], \quad (13)$$

and the overall signed probability of detection image $SP_t[i, j]$, may be expressed in terms of these signed probability of detection images as [13]

$$SP_t[i, j] = \text{sign} \left(\sum_{k,l} SP_{k,l}[i, j] \right) \cdot P_t[i, j], \quad (14)$$

where P_t is obtained as given in Equation 12. The overall signed probability of detection image serves as basis of the perceptual fidelity measure in the VDP algorithm [13]. For each pixel location it provides the probability that a difference will be perceived, as well as an indication of whether that difference is lighter or darker than the original.

As with the VDP, the Heeger and Teo approach presents the description of perceptual image fidelity in the form of a probability image. This is accomplished by applying a maximum *a posteriori* detection rule to the normalized images produced by the cortical stage of the Heeger and Teo model. Unfortunately, Heeger and Teo do not provide sufficient detail in their report to give a more complete description of their detection mechanism.

4.2.3.5 Summary. This subsection has presented the details of three recent approaches to measuring perceptual fidelity of achromatic images using multiple channel HVS models. In particular, it highlights the important role of contrast masking in pro-

cessing the multiple spatial frequency and orientation specific channels. These ideas will be incorporated into the new color fidelity measure that is introduced in Section 4.3.

4.2.4 Color Fidelity Metrics. In contrast to the large number of measures that have been proposed for achromatic imagery, very few perceptual fidelity measures have been proposed for application to color imagery. The few measures that have been proposed are generally extensions of the early vision models described in Section 4.2.2, and are closely related to the mean-square error idea. For example, Faugeras looked at the metric

$$d(I_1, I_2) = \sum_{i,j} \delta[i, j] \quad (15)$$

where $\delta[i, j]$ is computed as the Euclidean norm of the difference vector of the $[i, j]$ th pixel in the Faugeras perceptual space:

$$\delta[i, j] = (\Delta A^*[i, j]^2 + \Delta C_1^*[i, j]^2 + \Delta C_2^*[i, j]^2)^{1/2}, \quad (16)$$

where $\Delta A^*[i, j] = A_1^*[i, j] - A_2^*[i, j]$, and $\Delta C_1^*[i, j]$ and $\Delta C_2^*[i, j]$ are similarly defined for the C_1 and C_2 channels (the asterisk is a reminder that the difference is taken after the CSF filtering is performed on each channel).

Faugeras also considered metrics with a form similar to that of Equation 15, where the $\delta[i, j]$ was replaced with either the sum of the absolute differences in the three channels,

$$\delta[i, j] = |\Delta A^*[i, j]| + |\Delta C_1^*[i, j]| + |\Delta C_2^*[i, j]|, \quad (17)$$

or an expression that selects the maximum channel difference for each location:

$$\delta[i, j] = \max\{|\Delta A^*[i, j]|, |\Delta C_1^*[i, j]|, |\Delta C_2^*[i, j]|\}. \quad (18)$$

Using a set of images distorted with different amounts of additive noise, Faugeras found that all three of these metrics produced the same quality rankings as human assessments [25].

Using a model very similar to that of Faugeras, Hall constructed a chromatic perceptual mean square error metric, $PMSE_c$, defined as the sum of normalized mean square

errors for three perceptual channels. Using Faugeras' notation to denote the three perceptual channels, this is expressed as [36]

$$PMSE_c = NMSE_A + NMSE_{C_1} + NMSE_{C_2}, \quad (19)$$

where

$$NMSE_A = \frac{\sum_{i=1}^M \sum_{j=1}^N [A_1^*[i, j] - A_2^*[i, j]]^2}{\sum_{i=1}^M \sum_{j=1}^N [A_1^*[i, j]]^2}, \quad (20)$$

and $NMSE_{C_1}$ and $NMSE_{C_2}$ are similarly defined for the C_1 and C_2 planes (again, the asterisk is a reminder that the difference is taken after the CSF-based filtering has been performed). Hall compared this metric with an NMSE metric computed as $PMSE_c$ in the (R,G,B) space (no color transformation nor spatial filtering) and a normalized MSE metric computed in (R,G,B) space after applying a Laplacian operator to the three color planes to highlight the importance of edges to human observers. In a subjective ranking test with a small set of compressed images, the $PMSE_c$ metric was found to have the highest correlation with human rankings [36].

Note that there is a subtle difference between the Hall metric and the one explored by Faugeras. The Hall metric treats the difference between the two HVS processed images as three separate color planes, computing an energy-normalized MSE in each of the three color channels and adding the three MSE values together. In contrast, the Faugeras metric treats the difference between the two images as a single array of vector quantities. The magnitude of the error vector in the perceptual color space is computed for each pixel, and the sum of the resulting array of magnitudes is taken as the fidelity measure. Which of these two approaches has stronger psychophysical or physiological support is unknown.

Aside from the work of Faugeras and Hall, the problem of perceptual fidelity for color images seems to remain a relatively untouched field. Some experiments have measured how humans distinguish between different colored fields, while others have shown that colored elements in a scene can affect the way colors are perceived in other parts of the scene, but this work has yet to be fully utilized in an application involving the measurement of perceptual differences between two complex colored images. Also noticeably absent are models of the human color vision system which perform local spatial frequency analyses

of the chromatic channels of the HVS. The following section proposes a model to fill this void.

4.2.5 Summary. This section has reviewed previous work in assessing perceptual image fidelity. A large number of approaches have been proposed for achromatic images, most of which use limited HVS models. Recent approaches in measuring achromatic image fidelity use more extensive multiple-channel HVS models which incorporate results of contrast masking experiments to produce a spatial map of perceptible differences between the two images. In the realm of color image fidelity assessments, the few approaches that have been proposed are limited to relatively simple early vision models. These results provide the backdrop for introducing a new color fidelity measure that incorporates features of the multiple channel achromatic measures with the color model examined above in Chapter II.

4.3 The New Multi-channel Color Fidelity Measure

4.3.1 Introduction. Drawing upon previous work and the results presented in Sections 2.3, 2.4, and 3.3, this section proposes a multiple-channel model for measuring perceptual image fidelity of color images. Conceptually, the model extends the ideas used in Daly's Visible Differences Predictor to the color domain by applying the VDP algorithm to each of the three color channels of the Faugeras perceptual color space, producing a color visible differences predictor.

Before proceeding with the description of the color VDP algorithm, note that this approach rests upon several assumptions. Chief among these is the assumption that local spatial frequency analysis similar to that documented for the brightness channel also occurs in the chromatic channels. This assumption is required because complete measurements of spatial frequency tuning in the chromatic channels of the HVS have not been reported. Given the the measurements of Faugeras and the color Mach band results documented in this dissertation, in which spatial processing similar to that found in the brightness channel is indicated in the chromatic channels, it seems reasonable to assume that the chromatic channels may also undergo local spatial frequency processing similar to that of the brightness channel, as modeled by Daly in the VDP [13].

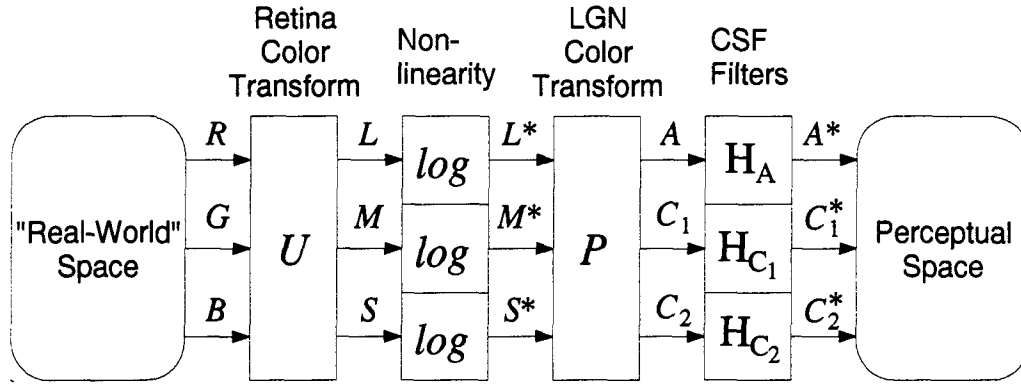


Figure 17. The Faugeras model of the human color vision system.

A second significant assumption implicit in the proposed approach is that the brightness and chromatic channels operate more or less independently, without any cross-channel interactions. In terms of masking effects, this approach is justified by the work of Mullen and Losada, who concluded that separate pathways detect color and luminance contrasts, with cross-masking occurring only in high-contrast situations [61].

With these assumptions in mind, the extended color visible differences algorithm will now be presented. Section 4.3.2 describes the algorithm from a system perspective, identifying the processing elements that are adapted from the two models to produce the new fidelity measure. The description of the system is completed in Section 4.3.3, where the parameter values for the new algorithm are chosen and justified.

4.3.2 The Color Visible Differences Predictor.

4.3.2.1 Overview. For reference, the Faugeras color model from Figure 2 is repeated here as Figure 17, and the basic elements of the VDP algorithm are shown in Figure 18. Recall that the Faugeras color model performs a linear color transformation from (R, G, B) color space to a retinal cone color space, which is followed by a logarithmic point non-linearity to model the non-linear response of the retina to input illumination. The logarithm operation is followed by another color transformation which represents a combination of cone outputs measured at the output of the LGN. Finally, bandpass spatial

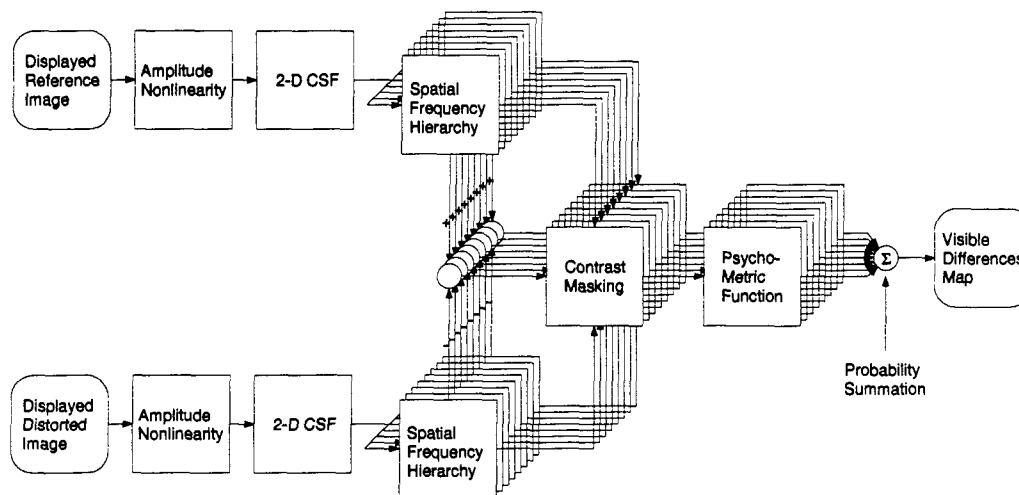


Figure 18. Basic elements of the VDP algorithm. Each arrow represents an image-sized array, and summations are performed on a pixel-by-pixel basis.

filters are applied to the three color channels output from the LGN stage to account for spatial effects of the visual system.

As described in the previous section, the first stage of the VDP algorithm consists of an amplitude nonlinearity and a linear filter representing the contrast sensitivity function. The following stage decomposes the input images into discrete spatial frequency and orientation bands using the cortex transform. Expressing the output of each cortex band in units of contrast, the difference between the two images at each location is computed within each cortex band. A psychometric function is then applied to these contrast differences, using a contrast masking function to control the detection threshold at each position in each cortex band. For each image location, the total probability of detection of an error is finally obtained by forming a probability sum (a product series) of the probabilities of detection from each band for that location.

In the new color fidelity measure proposed here, the multiple channel processing steps of the VDP algorithm are applied to the three output channels of the Faugeras color vision model, as shown in Figure 19. A complete description of the resulting model will now be given.

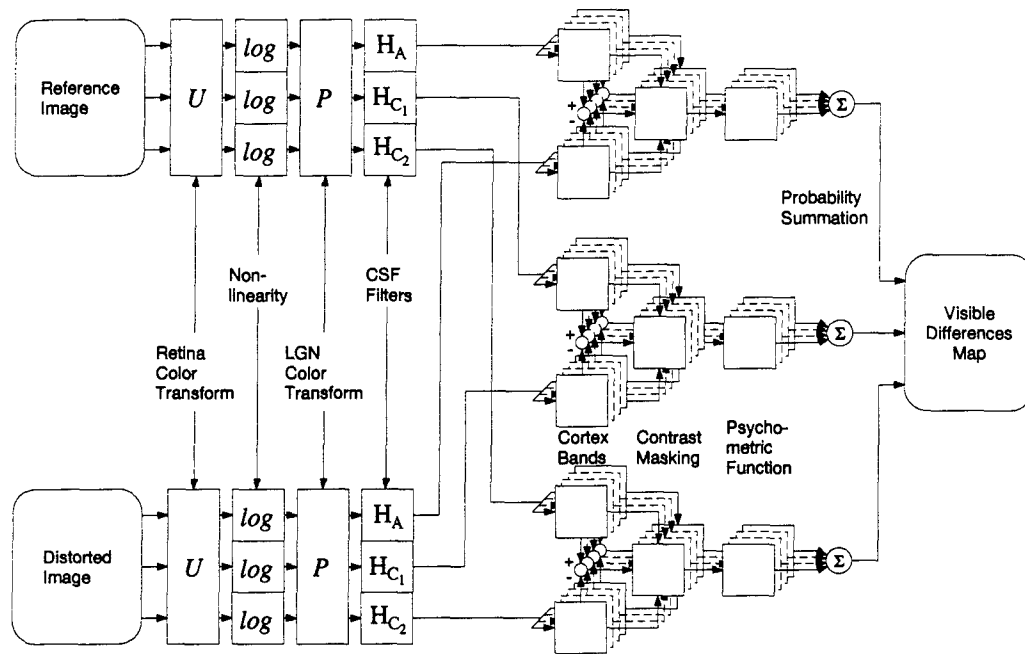


Figure 19. The Color Visible Differences Predictor

4.3.2.2 Early Vision Components. The early vision components of the model consist of the colorimetric transformations U and P , the amplitude non-linearity function, and the CSF filters. Because of the color blind tests reported in Section 2.4 the colorimetric transformations and the logarithmic non-linearity from the Faugeras model are used. Recalling from Equations 2 and 3, the transformations for a display with a D_{6500} white point are given as

$$U = \begin{bmatrix} .3634 & .6102 & .0264 \\ .1246 & .8138 & .0616 \\ .0009 & .0602 & .9389 \end{bmatrix}, \quad (21)$$

and

$$P = \begin{bmatrix} 13.8312 & 8.3394 & 0.4294 \\ 64 & -64 & 0 \\ 10 & 0 & -10 \end{bmatrix}. \quad (22)$$

In the VDP, Daly uses a shift invariant, invertible non-linearity which resembles the cube root function (asserted by some to be the “correct” functional form [36, 85]) and

is adaptive to local light levels [13]. However, in seeking a non-linear function to use in the structure shown in Figure 19, the analysis is complicated because the non-linearity is sandwiched between the retinal and LGN colorimetric transformations of the Faugeras model. To avoid this difficult analysis, the logarithm is used in the color VDP algorithm; a degree of local adaptive character is maintained in the color VDP algorithm by computing contrast locally, as described below.

The CSF filters used in the model have the functional form proposed by Mannos and Sakrison [54], tuned to peak at 8 cycles/degree for the brightness channel (A), 4 cycles/degree for the C_1 channel, and 2 cycles/degree for the C_2 channel:

$$H_A(f_r) = 2.6[0.0192 + 0.113f_r] \exp[-(0.113f_r)^{1.1}], \quad (23)$$

$$H_{C_1}(f_r) = 2.6[0.0192 + 0.226f_r] \exp[-(0.226f_r)^{1.1}], \quad (24)$$

and

$$H_{C_2}(f_r) = 2.6[0.0192 + 0.452f_r] \exp[-(0.452f_r)^{1.1}], \quad (25)$$

where the radial spatial frequency f_r is given in cycles/degree. A graphical comparison of these functions with those used by Faugeras shows the two to be in close agreement. The CSF filter used in the VDP algorithm is adapted to allow a range of viewing distances to be specified. In order to keep the overall system as simple as possible, this feature is not included here.

4.3.2.3 Multi-channel Analysis. After passing through the early vision stage of the model, the two input images enter the multiple-channel cortical stage. This stage applies the multiple-channel processing of the VDP separately to each component of the early vision stage (A , C_1 , and C_2), in a multiple-channel analog to the $PMSE_c$ approach used by Hall (processing components separately, then combining the results). This approach is justified by the findings of Mullen and Losada, who concluded that color and luminance contrasts are processed by separate pathways in the HVS which do not interact with each other, except for possibly in cases of high contrasts [61]. Thus, the three components of the Faugeras model (A , C_1 , and C_2) are individually analyzed with a bank of

spatial frequency and orientation tuned filters, and the outputs of these filters are expressed in terms of local contrast. The contrast difference in each subband is then computed and scaled by a masking function which sets the detection threshold in a psychometric function. The psychometric function produces a probability of detection image for each subbands, and these images are in turn combined via probability summation into a single probability of detection image for each component. Finally, an overall probability of detection image is obtained by applying the probability summation technique to the three component images. The specific elements of this multi-channel processing are now considered in detail, highlighting differences between the approach used here and that of the original VDP algorithm.

The first difference is in the choice of filters which produce the local spatial frequency analysis. Rather than the Cortex transform used by Daly, Gabor filters are used in the color VDP to form the filter bank. The most dramatic difference between these two approaches is that the filters in the Cortex transform are carefully formed to ensure that they sum to one over the entire spatial frequency plane, while the Gabor filters do not possess this property. This property is included in the Cortex transform to ensure that an image can be perfectly reconstructed from a Cortex transform analysis. However, as there is no intent in this application to reconstruct the analyzed component images, this constraint is not required. Notwithstanding, the Gabor filter bank is still constructed to possess the commonly cited attributes of measured cortical simple cell receptive field profiles: even spacing of the radial frequency centers on a log-frequency scale, 1.5 octave radial frequency bandwidth, and thirty degree orientation bandwidths spaced thirty degrees apart, as shown in Figure 20 (bandwidth figures reported here are the distance between $1/e$ points of the Gaussian functions composing the Gabor filters.) A total of 5 radial frequency centers with 6 different orientations is used in the present color VDP, along with a Gaussian, low-pass baseband filter. Only cosine Gabor filters are used, and the $1/e$ point of the baseband filter is chosen so that it matches the lower $1/e$ frequency of the lowest resolution Gabor filters. Appendix A provides implementation details of computing these filters.

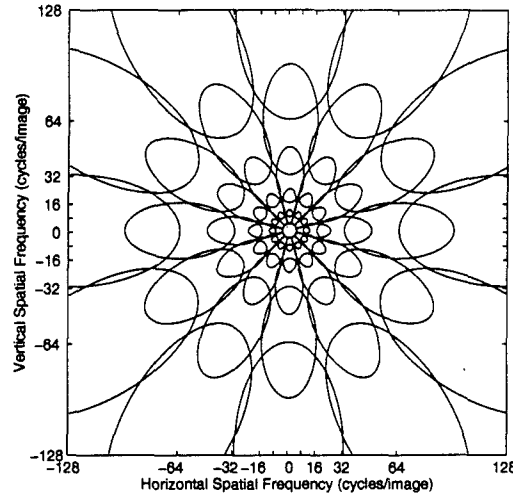


Figure 20. Depiction of the filters used in the color perceptual fidelity measure for a 256×256 image.

4.3.2.4 Contrast Units. In order to apply the contrast masking techniques developed by Daly, the subbands must be expressed in units of contrast. Following a Michelson definition of contrast, the contrast in the k, l th subband may be written as a function of pixel location as

$$C_{k,l}[i,j] = \frac{B_{k,l}[i,j] - \overline{B_{k,l}}}{\overline{B_{k,l}}}, \quad (26)$$

where $B_{k,l}[i,j]$ is the value of the i, j th pixel and $\overline{B_{k,l}}$ is the mean of all the pixels in the filtered image for the k, l th subband. However, because of the Gabor filters, the mean (DC) value for each of the subbands approaches zero, causing this expression to be indeterminate. Daly proposes two solutions to this difficulty, both of which leave the mean in the numerator as zero. The first solution, a "global" contrast, replaces the denominator mean with the input image mean, which is equivalent to the baseband mean $\overline{B_{00}}$:

$$C_{G_{k,l}}[i,j] = \frac{B_{k,l}[i,j]}{\overline{B_{00}}}. \quad (27)$$

The second solution, called "local" contrast, replaces the denominator with the baseband image and computes the ratio on a pixel by pixel basis:

$$C_{Lk,l}[i,j] = \frac{B_{k,l}[i,j]}{B_{00}[i,j]}. \quad (28)$$

As pointed out earlier, the logarithmic non-linearity does not adapt to local brightness as Daly's non-linear function does. In order to provide a degree of adaptability in the color VDP algorithm, the local contrast approach of Equation 28 is chosen for use in expressing subband output in terms of contrast.

4.3.2.5 Contrast Masking. The next element of the cortical processing stage is the contrast masking function. This function sets the detection threshold in the psychometric function that determines the probability of detection of a contrast difference between the two images. To obtain the threshold, Daly developed a model which unifies the results of a large number of different contrast masking experiments. This model uses a mutual masking approach that relies upon the content of both images to determine the masking threshold for each image location. In the color fidelity measure, this approach is adapted to account for the use of Gabor filters to define the subbands.

The mutual masking threshold elevation image is obtained as follows. First, for each subband, a threshold elevation image $T_e^{k,l}[i,j]$ is computed from each input image. Then, the masking image for that subband $T_{em}^{k,l}[i,j]$ is obtained by taking the minimum of these two threshold elevation images at each location:

$$T_{em}^{k,l}[i,j] = \min\{T_{e1}^{k,l}[i,j], T_{e2}^{k,l}[i,j]\}. \quad (29)$$

Thus, at each location in each subband, the reference and distorted images may interchange roles of masking and target signals.

The threshold elevation images for each subband are obtained using Daly's threshold elevation function, which was described earlier in Section 4.2.3. Using the functional form of Equation 7, a masking threshold image is computed for each subband. This is accomplished by expressing the normalized masking function m_n in Equation 7 as a function of location

for the k, l th subband, $m_n^{k,l}[i, j]$, producing

$$T_e^{k,l}[i, j] = (1 + (k_1(k_2 \cdot m_n^{k,l}[i, j])^s)^b)^{1/b}. \quad (30)$$

Daly obtains the normalized masking function for the k, l th subband from an input image by [13]

$$m_n^{k,l}[i, j] = |\mathcal{F}^{-1}\{\mathcal{L}[u, v] \cdot \text{csf}[u, v] \cdot \text{Gabor}^{k,l}[u, v]\}|, \quad (31)$$

where $\mathcal{L}[u, v]$ is the Fourier transform of the input image after being processed by the amplitude nonlinearity, and u, v are Cartesian frequency components. Note that the image itself is the masking signal; the CSF and Gabor terms serve to normalize each frequency component according to the HVS detection threshold for each spatial frequency in a uniform field, and the inverse Fourier transform provides a weighted sum of the frequency components in the k, l th subband.

As noted earlier, four parameters control the shape of the threshold elevation function: k_1 , k_2 , s , and b . The parameters k_1 and k_2 , which control the knee point of the threshold elevation curve, are completely determined by the value of two other parameters, W and Q , as given in Equation 9. The values chosen for these parameters in the color VDP algorithm are given below in Section 4.3.3.

With a normalized masking function from each input image for each subband, a threshold elevation image for the k, l th subband is computed for each image via Equation 30. Using the two threshold elevation images thus generated, the overall mutual masking threshold elevation image for the k, l th subband is obtained by taking the minimum of the two threshold elevations at each location, as indicated above in Equation 29.

4.3.2.6 Psychometric Function and Probability Summation. Once the masking threshold images have been computed, probability of detection images are computed for each subband of the three Faugeras color channels, and the subband images in each channel are combined through probability summation to produce an overall probability of detection image for each channel. These operations are accomplished as described above in Section 4.2.3.4. The probability of detection image for each subband is generated

through the psychometric function as given in Equation 11, and the probability of detection image for each color channel is then obtained by computing the product series of all the subbands in that component as described in Equation 12.

The in-channel probability summation operations produce three probability of detection images (or visible differences map) which reflect the visibility of differences in the A , C_1 , and C_2 color channels. Applying the probability summation technique to these three images, a single visible differences map for the two input images is formed, as in

$$P[i, j] = 1 - [(1 - P_t^A[i, j]) \cdot (1 - P_t^{C_1}[i, j]) \cdot (1 - P_t^{C_2}[i, j])] . \quad (32)$$

This overall map indicates the probability of detecting any kind of visible difference, whether it be a change in brightness or chroma.

In the context of color images, signs such as those used by Daly to indicate the nature of a visible change are not particularly useful. Because visible differences may occur in any or all three of the three Faugeras color channels, a single sign for the overall probability of detection image carries no meaning. Therefore, no signs are computed in the color VDP algorithm.

4.3.3 Parameter Values. The previous section outlined the structure of the color VDP algorithm. The purpose of this section is to identify and specify the values of the algorithm parameters.

The first set of parameters are the colorimetric transformation matrices of the Faugeras color model. The values of these parameters, derived for an RGB display with a D_{6500} white point, are given in Equations 21 and 22. The bandpass CSF spatial filters that follow the colorimetric transformations are given in Equations 23-25. The parameters for these filters are chosen to produce a peak frequency of 8 cycles/degree for the A channel, 4 cycles/degree for the C_1 channel, and 2 cycles/degree for the C_2 channel. These values were determined by Mannos and Sakrison [54] and Hall [36].

Following the CSF filters is the Gabor filter bank. As discussed above, these filters are chosen to produce logarithmically spaced radial frequency centers with 1.5 octave

radial frequency bandwidth, and thirty degree orientation bandwidths spaced thirty degrees apart. These values follow most closely the results cited by Daugman [15, 16, 46].

The final set of parameters required for the color VDP algorithm are those involved in computing the threshold elevation masking images. The shape of the threshold elevation function in Equation 30 is controlled by four parameters: k_1 , k_2 , b , and s . In turn, k_1 and k_2 are specified by two other parameters, W and Q , as described in Equation 9. Thus, the final set of parameters to be specified are W , Q , b , and s . The physical significance of these parameters are described above in Section 4.3.2. In the VDP, Daly chose the values of $W = 6$, $Q = 0.7$, and $b = 4$, and varied s depending on the subband, from 0.7 for the baseband to 1.0 for the middle frequencies [13]. In the color VDP algorithm, the same values are used for these parameters in each of the subbands of all three color components, with the exception that a fixed value of 0.8 is chosen for the slope parameter s . This simplification is used because the experiments required to better specify s are considered beyond the scope of this dissertation.

4.3.4 Summary. This section has outlined a new, multiple channel approach for measuring perceptual fidelity of color images. This new approach is called the color visible differences predictor, as it applies the multiple-channel operations of Daly's visible differences predictor to the color components of the Faugeras color model. In the following section, the performance of this new approach is explored.

4.4 Demonstration

4.4.1 Introduction. In the previous section, a color VDP algorithm was developed by adapting the VDP algorithm so that it could be applied to the three channels of the Faugeras color vision model. This section demonstrates the performance of this color VDP algorithm by applying it to a few images which have been distorted in various ways. In each test case, the original and distorted images are presented together with four outputs from the color VDP algorithm, as depicted in Figure 21. The visible differences for the individual channels of the Faugeras space (A , C_1 , and C_2) are included to show the influence of the three channels on the overall visible differences image. The overall visible differences image,

Original	Distorted
Visible Differences A Component	Visible Differences C ₁ Component
Visible Differences C ₂ Component	Overall Visible Differences

Figure 21. Map of images shown in Figures 22-27 to demonstrate the color VDP algorithm.

computed as described in Section 4.3.2.6, is identified as the perceptual fidelity measure. All four visible differences images are displayed as grayscale images; white identifies where the probability of detecting a visible difference is high, while black indicates areas where differences fall below the visibility threshold.

Before proceeding to the demonstration, the reader is reminded that the match between what is displayed on the monitor screen and what is shown in the color prints in this section is only approximate at best. All the experiments were performed assuming the display is a color (RGB) monitor with a D_{6500} white point. This assumption yielded valid results for the monitors used in the color vision experiments discussed in Chapter II, so it was included in the experiments detailed here. The same monitors were used in displaying the results documented here. While the color prints given here provide a reasonably faithful representation of what is observed on the screen, the reader should keep in mind that they are not exactly the same.

Two types of distortions were used in these experiments. First, an artificial banding distortion was produced by adding a banding pattern to the test image in various ways. These results are described in Section 4.4.2. The banding distortion is similar to one used by Daly to demonstrate the VDP [13]. After the banding distortions, the color VDP algorithm is applied to a compressed color image. Section 4.4.3 shows these results. This example is included to provide a test of the algorithm in a typical application.

4.4.2 Banding Distortions. The first set of distortions examined are generated by adding a banding pattern to a single color component of the test image. The banding pattern has both positive and negative values, so that both increases and decreases in the distorted component are represented. Figures 22-26 show the results of these tests. In Figures 22, 23, and 24, the banding is added to the A , C_1 , and C_2 components of the original, respectively. In Figure 25, a lower spatial frequency banding pattern is added to the C_2 component. Finally, Figure 26 shows the results of adding the banding distortion to the NTSC Y component of the original image.

The results shown in Figure 22 for the banding distortion added to the brightness (A) channel are the most encouraging of the set. In this example, the banding distortion

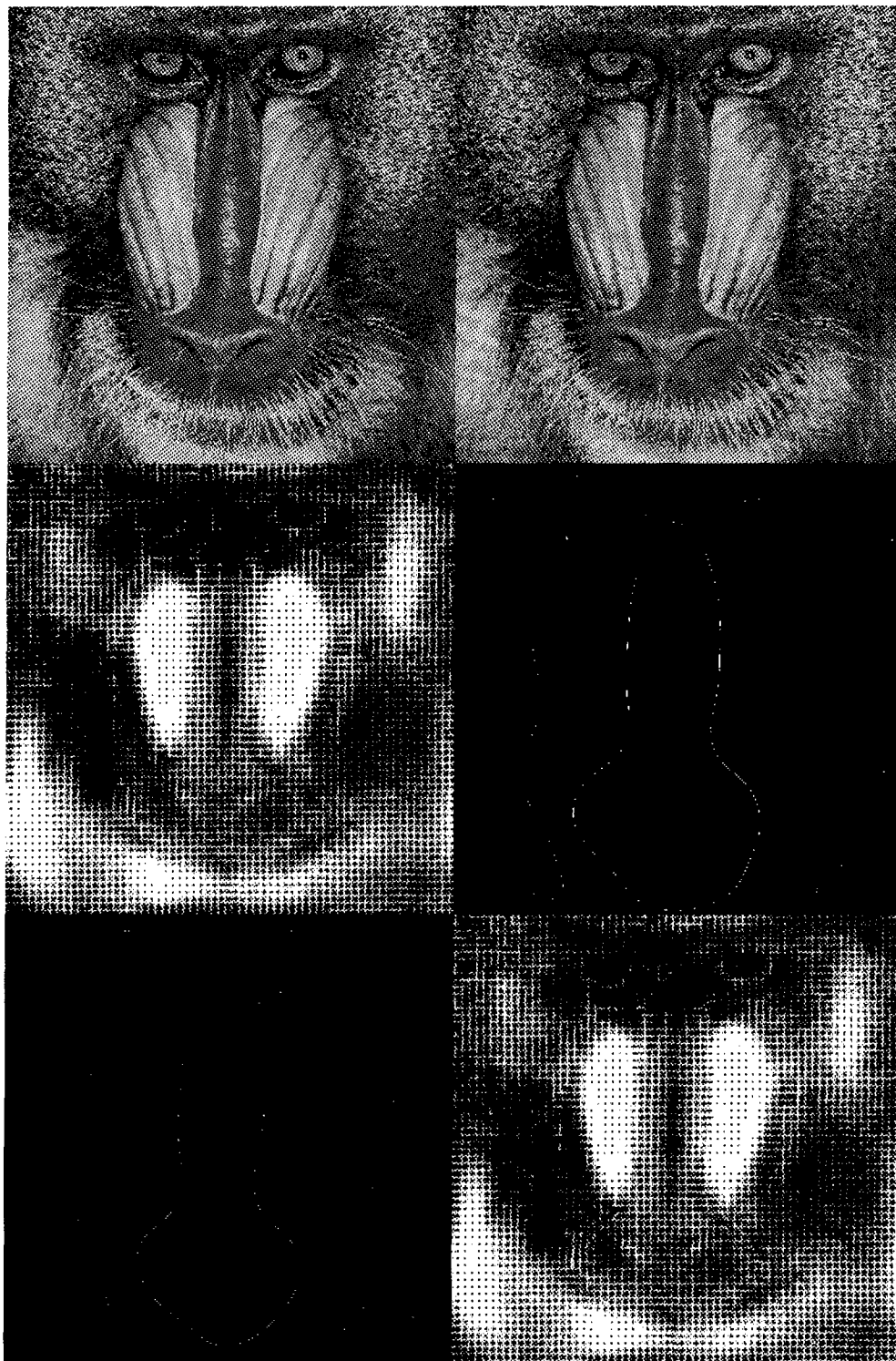


Figure 22. Color VDP output: mandrill with banding distortion added to *A* component.

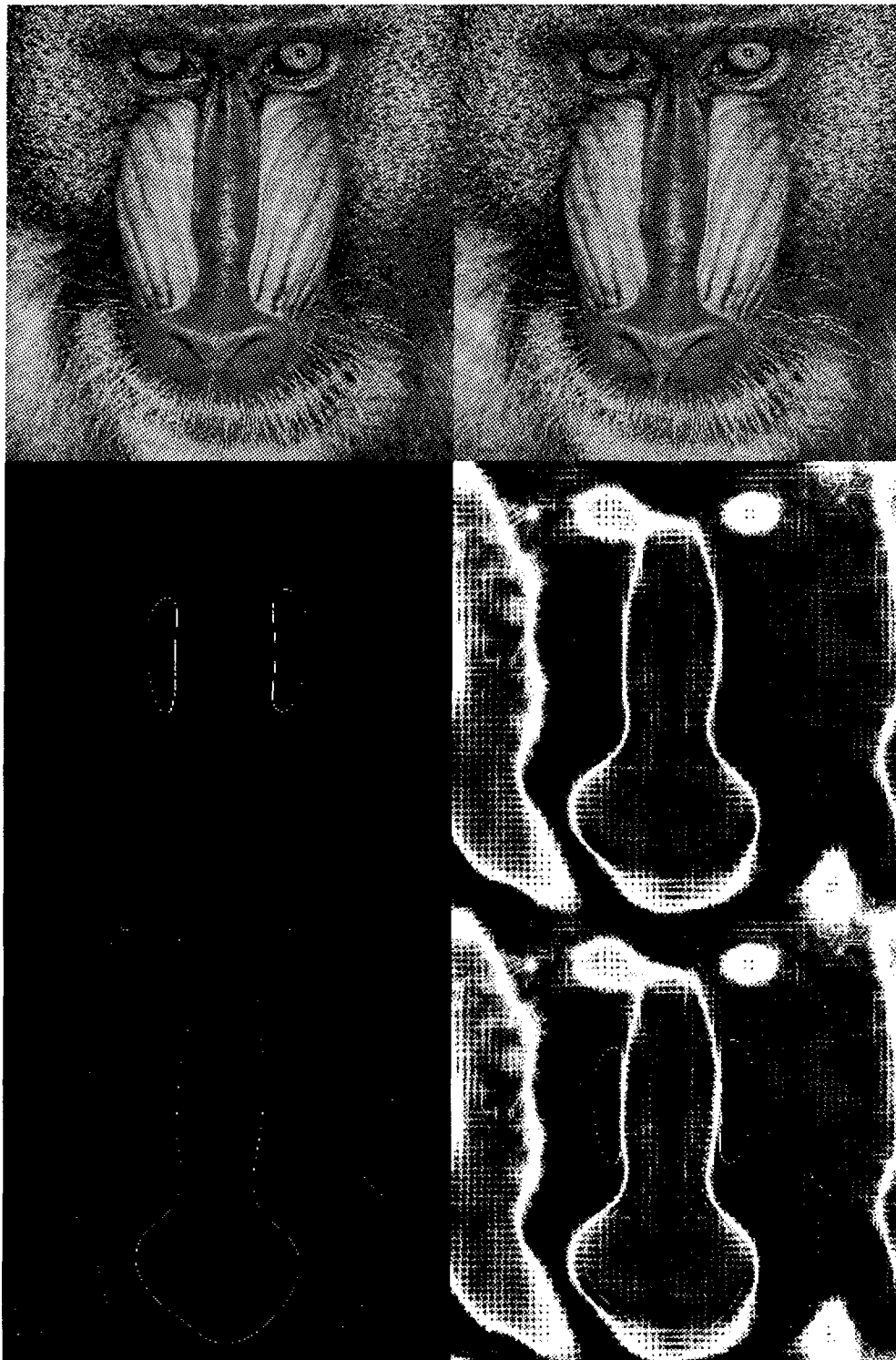


Figure 23. Color VDP output: mandrill with banding distortion added to C_1 component.

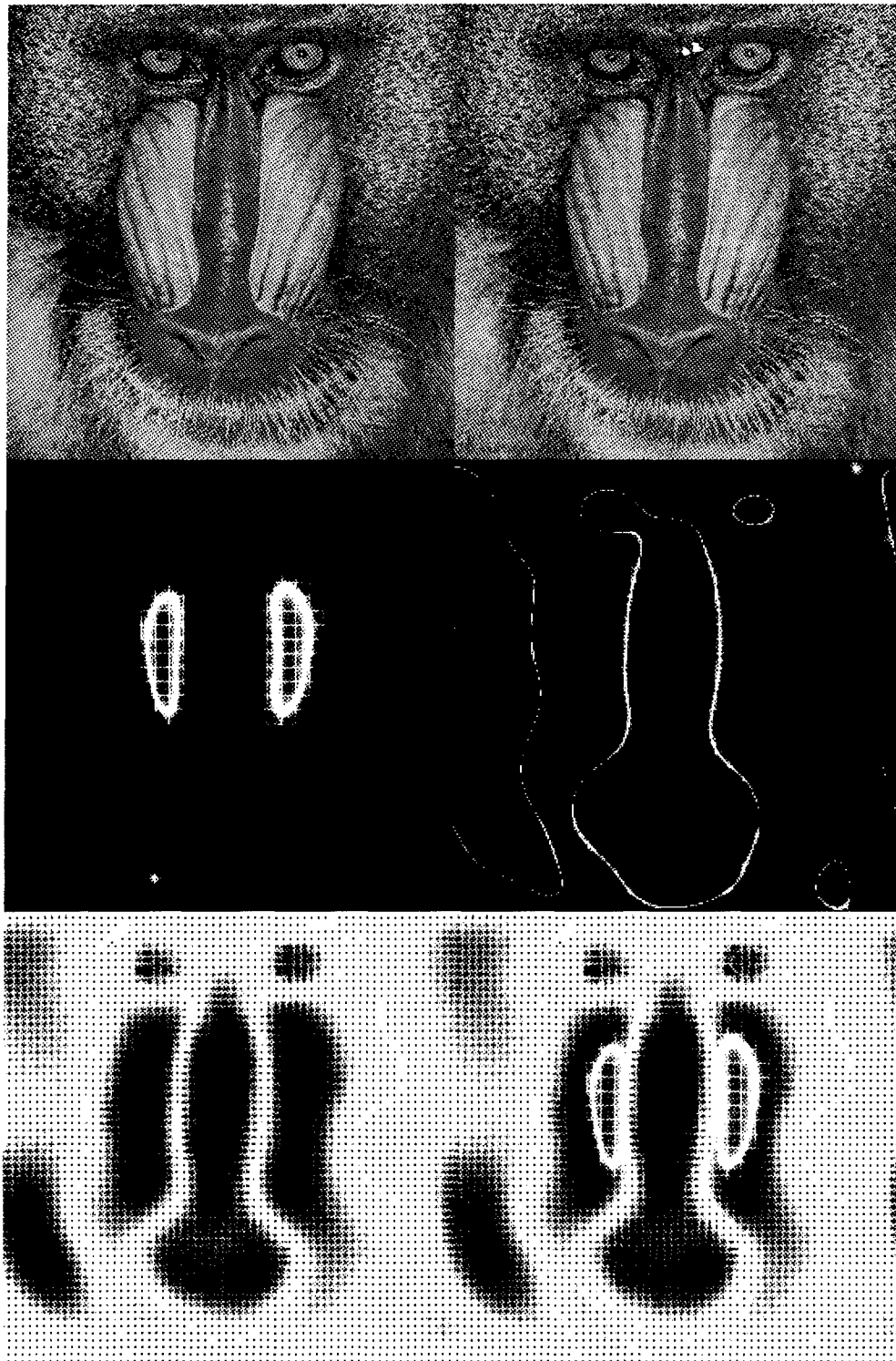


Figure 24. Color VDP output: mandrill with banding distortion added to C_2 component.

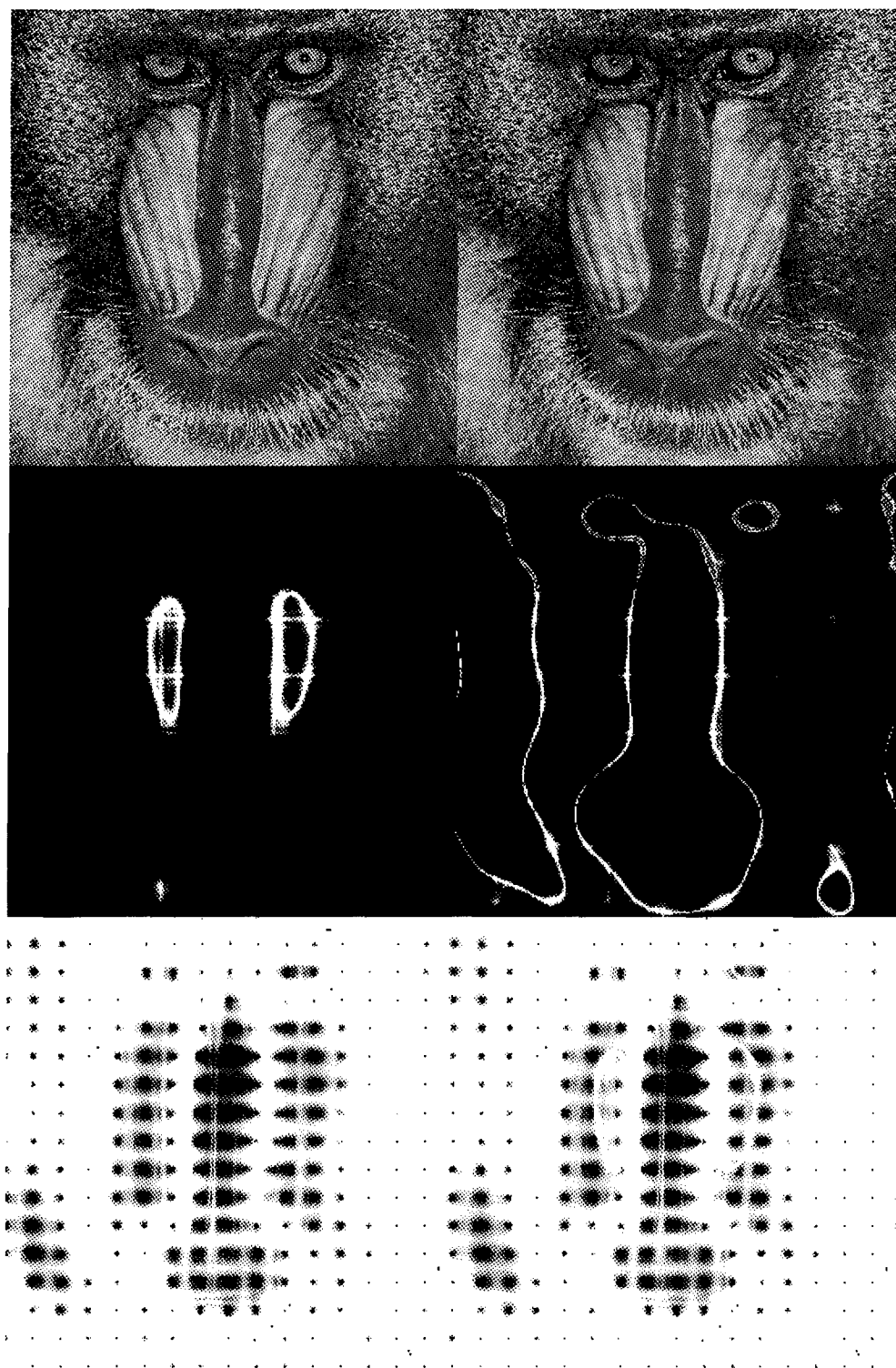


Figure 25. Color VDP output: mandrill with wide banding distortion added to C_2 component.

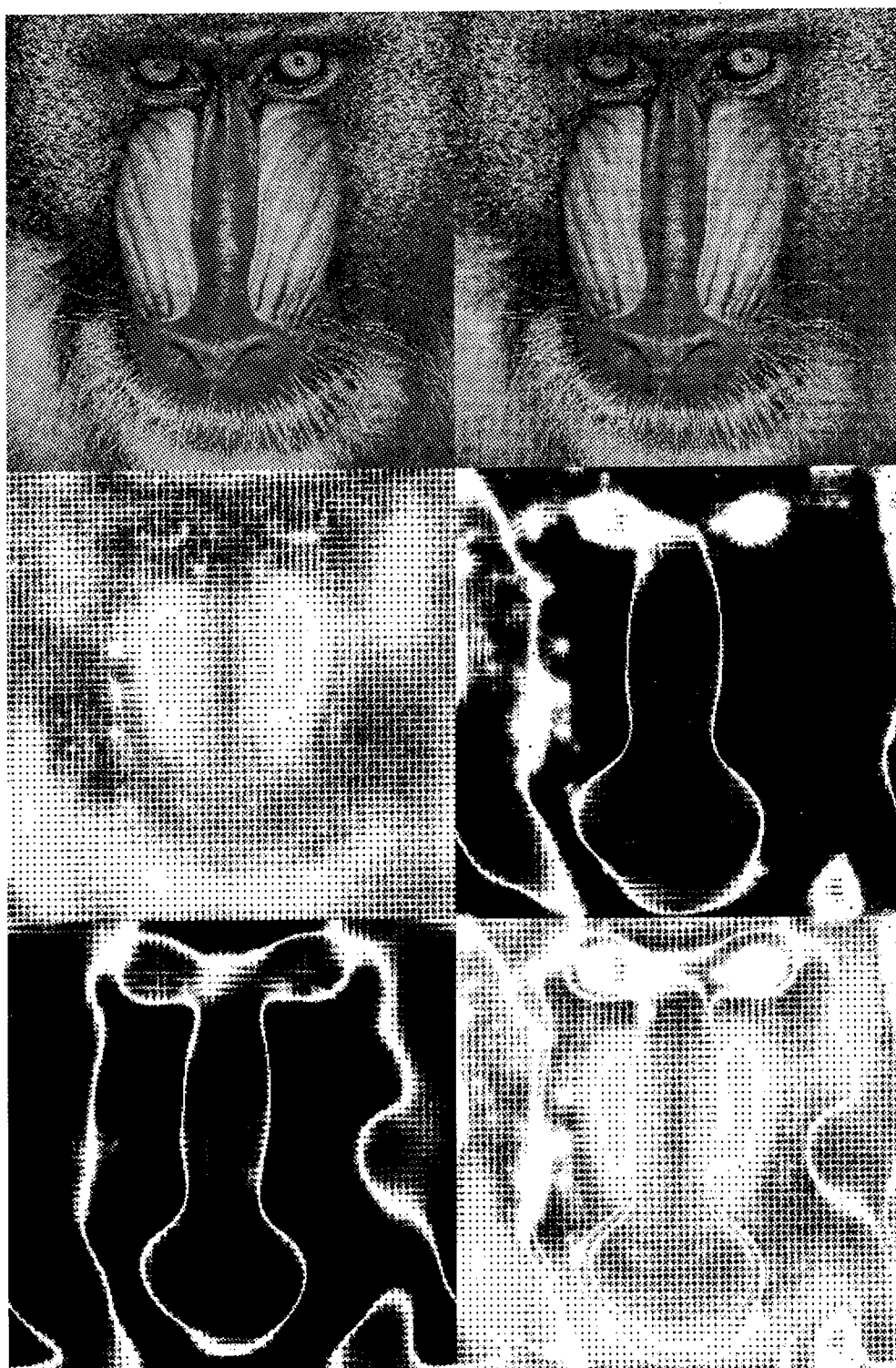


Figure 26. Color VDP output: mandrill with banding distortion added to Y component of NTSC color space.

is least visible in the dark areas around the eyes and at the corners of the mouth, and it is most visible on the nose. The predictions of the color VDP algorithm are consistent with these observations. In the output visible differences map, bright areas in the locations of the blue parts of the nose and the light orange area to the left of the mouth identify these areas as the locations where differences are most visible. The banded nature of the distortion is seen in much of the remainder of the map, while dark areas around the eyes and the corners of the mouth correspond to the locations where the distortion is least visible. With distortions only occurring in a single channel, this image provides a good test of the independence of the processing in the three channels of the algorithm. As expected, no visible differences are found in the two chromatic channels for this pair of images.

In the test where the banding is added to the C_1 channel (Figure 23), the results are still promising, although they are not quite as good as was hoped. The algorithm still only responds in the channel where the distortion was added, and it does a reasonably good job of identifying the banded distortion appearing mostly in the red part of the nose. However, the algorithm also suggests that the banding should be visible over a fairly substantial area on the left side of the image, as well as in the location of the eyes. This prediction does not correspond well with the findings of direct observation. The cause for this inaccuracy is attributed to the choice of parameters used in setting visibility thresholds in the chromatic channels of the model. Lacking sufficient data to provide adequate direction in setting these chromatic channel parameters, the same values were used in the chromatic channels as in the brightness channel. This gross assumption probably contributes significantly to the over-prediction of the model for this image. Nevertheless, the fact that the banding is identified in the location where it appears most visibly suggests that the approach of the color VDP algorithm does have some merit.

In Figure 24, similarly ill-chosen parameter values are likely the reason for over-prediction of the visible differences. In this case, the banding distortion was added to the C_2 component of the original image. On the video display screen, this distortion is very difficult to detect. The algorithm, however, suggests that the distortion should be evident almost everywhere in the image. Parameters relating to detection thresholds in the C_2 channel are thus highly suspect in this case. The separation of the three color channels is

still fairly good, although in this case a small area of error is made in the A channel on the sides of the nose. The cause for this anomaly is not immediately apparent.

The results shown in Figure 25 for the wide banding distortion added to the C_2 component seem to correspond more closely with the observed distortions. This image was used because of the low visibility of the narrow banding distortion added to the C_2 component in Figure 24. Recalling that the peak sensitivity of the C_2 channel is in the neighborhood of 2 cycles per degree, it was reasoned that the narrow distortion did not appear because it was beyond the cutoff spatial frequency for that channel. Thus, a distortion with lower spatial frequency content should become more visible. This reasoning is found to be correct in Figure 25, where the wide banding pattern is found to be much more visible, both in the displayed and the printed versions. The response of the color VDP algorithm also seems to be somewhat better in the wide banding case, although the visible differences are still somewhat over-predicted. However, at least some of the areas where the colors are left unchanged are correctly identified by the algorithm in this case.

In the final banding distortion experiment, the narrow banded pattern was added to the Y component of the NTSC YIQ color space. This was done to see if the color VDP algorithm would pick up distortions occurring in more than one color channel simultaneously. The results shown in Figure 26 are somewhat inconclusive, for two reasons. First, because it is difficult to get a feel for what the three channels of the Faugeras look like in various combinations, it is virtually impossible to identify the components of differences involving more than one channel. It is hard enough to identify changes when they are only in one of the chromatic channels! The second reason has to do with the choice of the Y component of the NTSC color space to receive the distortion. In retrospect, this was a somewhat poor choice, since, as a brightness correlate, the Y component is largely composed of the Faugeras A component, without much contribution from the C_1 and C_2 components. These considerations make it difficult to conclude with certainty that the color VDP algorithm accurately predicts what is perceived by the human observer. However, two aspects of the results shown in Figure 26 are hopeful: significant contributions to the overall visible differences map are made by the individual channel maps, and the

visible differences indicated by the overall map correspond quite well with what is observed in the input images.

4.4.3 Compression Distortions. While acknowledging some weaknesses, the banding distortion experiments in the previous section provide a fair degree of confidence in the color VDP algorithm. In this section, the algorithm is applied in a more realistic scenario—assessing the visible differences due to compression. Thus, Figure 27 shows the results of applying the color VDP algorithm to a compressed image. The image was compressed from twenty-four to two bits per pixel using a perceptual coding scheme developed by Hall [36]. This compression algorithm avoids blocky artifacts that arise in many compression schemes (including the common JPEG algorithm) by processing the entire image all at once, rather than in small blocks. Compression is achieved by allocating bits to Fourier coefficients based on the human contrast sensitivity function. Thus, quantization errors are spread across the entire image, with larger errors allowed for spatial frequencies to which the human visual system is less sensitive. While the resulting compressed image resembles the original quite well, it is not without artifacts—the perceptual bit allocation gives rise to a texture that is superimposed on the image.

The visible difference maps for the A , C_1 , and C_2 channels in Figure 27 show that visible differences are distributed over most of the image in all three channels. The red and blue parts of the mandrill's nose are the only areas that are identified by the color VDP algorithm where the differences are below threshold for the chromatic channels. However, when combined with the output of the A channel, only small areas of the red part of the nose remain below the visibility threshold. Close examination of the two input images in Figure 27 bears out the predictions of the color VDP. The compression-induced texture is evident over the entire image, although it is most difficult to detect on the red part of the nose and the furry cheeks. When the input images are displayed on the monitor, the texture due to the compression appears in the fur more strongly than in the printed version, while the red nose area remains the location where the texture is least noticeable.

As in the case where banding was added to the Y component of the NTSC color space, it is difficult to assess the accuracy of the predictions of the visible difference maps for the

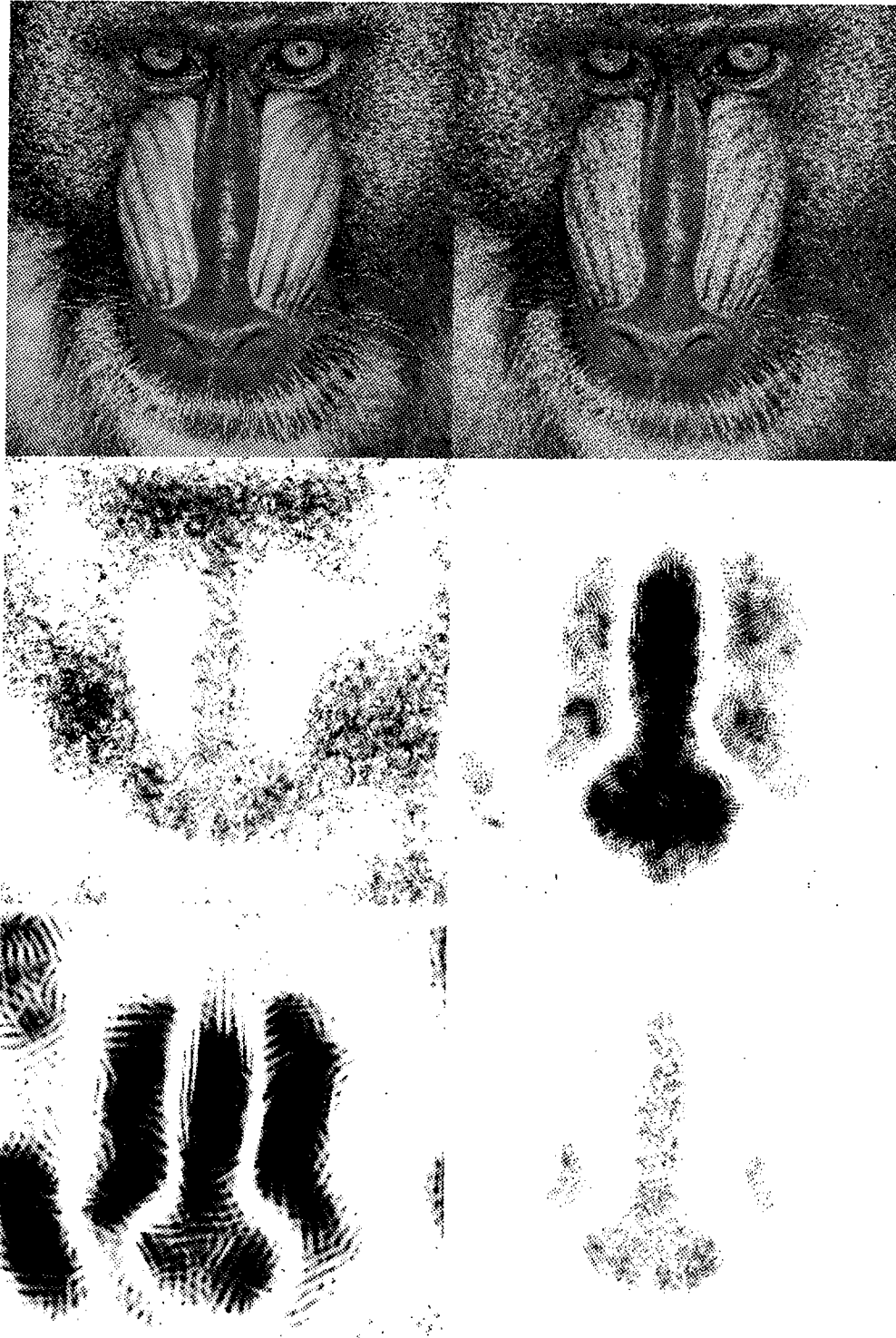


Figure 27. Color VDP output: mandrill image compressed to 2 bits per pixel using the Hall perceptual coding algorithm.

chromatic channels in this application. As noted earlier, the threshold parameters chosen for these channels tended to cause the algorithm to over-predict the chromatic visible differences. Because of the nature of the compression algorithm, it is anticipated that distortions at any given image location will not be restricted to occurring within a single color channel. In this situation, it is virtually impossible to determine which distortions are occurring in which of the three channels by direct observation. Therefore, it is very difficult to look at the visible differences map for a single channel and determine whether or not it provides a reliable prediction of visible differences that may be attributed to that channel. Nevertheless, despite these difficulties, the outputs of the color VDP shown in Figure 27 seem reasonable for this image.

4.4.4 Summary. This section has provided a demonstration of the color VDP algorithm proposed in Section 4.3. Several distorted test images were produced by distorting a single reference image in various ways. Included in the test set were images that were distorted in just one of the three Faugeras color channels, as well as an image that was distorted in more than one channel simultaneously, and an image that was distorted by an image compression algorithm. The results of applying the color VDP algorithm to these images provide valuable insight into the operation of the algorithm and the overall problem of assessing visible differences for color images. These results are discussed in the following section.

4.5 Discussion

4.5.1 Introduction. The thrust of this chapter has been the development of a new perceptual image fidelity measure for color digital image. The previous sections have provided background for the current effort, presented the new approach, and demonstrated how it works using a set of test images. In this section, the effectiveness of the approach is discussed, as well as ways in which it may be improved.

4.5.2 General Observations. Before focusing on specific aspects of the new approach, a few observations of a general nature are warranted. First, it is acknowledged that

the examples in Section 4.4 hardly constitute a complete test of the color VDP algorithm. In order to fully develop the color VDP algorithm, many more experiments are required.

The initial results are promising. The algorithm successfully identifies areas where distortions have been added, in most cases highlighting areas with the highest visibility of distortions. Considering the assumptions made to develop the color VDP approach, these results are positive. Clearly, there is more work to be done to refine the model, but the results suggest that the multi-channel processing of the Faugeras color channels is a viable approach. The parameters chosen may not be optimal in the sense of matching the HVS, but the structure of the approach is sound.

One of the greatest difficulties in evaluating this fidelity measure, or any perceptual measure for that matter, is in developing and performing the tests to determine how well the measure performs compared with human assessments. The difficulty is compounded when trying to compare this approach with previous approaches. In order to have a valid comparison, it must be established that the tests of the two measures follow the same procedures with substantially the same kinds of images and distortions. Without the ability to control or duplicate many of these conditions, it is difficult to identify one given approach as better than another with a high degree of confidence.

4.5.3 Parameters. As noted in the previous section, the color VDP algorithm seems to over-emphasize differences in the two chromatic channels of the Faugeras space. This may be attributed directly to the fact that, with the exception of the CSF filters, the parameters for these two channels were set to the same values as those used in the brightness channel. While it simplifies the specification of the model, this is likely a faulty assumption. In both previous color HVS models discussed in Chapter II, spatial bandpass filters are included for each of the three perceptual channels, suggesting that similar processing takes place in all three channels. However, the center frequencies and cutoff frequencies of the three bandpass filters are different, because the processing is not exactly the same. This alone should be enough to discourage one from using the same parameter values in the multi-channel processing applied to the three color channels. It is

reasonable to assume that similar operations occur, but to suggest that the processing is exactly the same requires a stretch of imagination.

Recent investigation has found several color masking experiments performed within the past two decades to quantify masking in chromatic stimuli [5, 11, 19, 52, 53, 60, 61, 68, 90]. Analysis of these experiments and corresponding results should provide some improvement in accounting for masking effects in the chromatic channels of the color VDP. The analysis is complicated, though, because the experimental results must be related to the Faugeras color model in order to be included in the present approach. Given the successful results with the Faugeras model discussed in Chapter II, it is possible that the results of some of these color masking experiments may be improved by performing the experiments in the context of the Faugeras model. A third approach follows that of Mannos and Sakrison [54], choosing parameter values by observing their effect on the performance of the color VDP algorithm itself. Any or all of these approaches may be followed to produce improvements in the color VDP algorithm.

4.5.4 Single Number vs. Difference Map Fidelity Measures. Another important problem that has not been addressed extensively is the utility of a visible difference map as a measure of perceptual fidelity. Difference maps have the advantage that they can identify image locations where distortions should be most apparent, as well as having the ability to indicate the nature of distortions occurring in the input images. This can provide important insight to a user who is working to minimize distortions caused by some kind of image processing system.

On the other hand, difference maps have a distinct disadvantage in that they cannot be readily compared to provide a ranking of images based on their perceptual fidelity. With visible differences maps, there is still the problem of determining some means of expressing the overall effect of all the distortions in an image in such a way that meaningful comparisons can be made between images. These means could be either subjective (e.g. some kind of human judgment applied to the maps) or objective (e.g. a computational formula).

In the original proposal of the VDP algorithm, Daly acknowledged this limitation of the difference map format, stating, "Once confidence is established in the accuracy of the VDP for threshold results, it can be used as a framework to study potential metrics that reduce the prediction to a single number [13]." A similar statement could be made for the color VDP algorithm. The framework of the model appears to be correct, but careful psychophysical tests are required to fine-tune the algorithm to achieve the desired accuracy.

One approach suggested by Daly is to find the maximum probability of detection in the overall difference map. If this peak probability of detection is below 1.0, then all the distortions are within the visibility threshold region. If the value is 0.5 or less, the two input images are identified as visually equivalent [13]. Such an approach at least provides a standard which can be used to evaluate parameter settings in an image processing system.

4.5.5 Summary. Reviewing the results presented in the previous section, this section has discussed the strengths and weaknesses of the color VDP algorithm as a perceptual fidelity measure. Generally, the structure of the approach seems to be sound, while more work to better specify the masking parameters for the chromatic channels is required. Several approaches for improving the specification of the parameters are suggested.

Also noted in this section are the pros and cons of using visible difference maps as measures of perceptual fidelity. While a map-based approach can be helpful in providing insight to the image system designer, the number of potential uses for such a measure increases significantly if there is way to develop a single-number metric from it. As the parameters become better defined, the color VDP algorithm may provide a valuable framework for the development of such a single-number measure.

4.6 Summary

This chapter has presented a new approach for computing the perceptual fidelity of color images. After reviewing previous efforts, it describes the new approach. Because Daly's visible differences predictor is applied to the three color channels of the Faugeras color model, the approach is called the color visible differences predictor. This

basic approach is based on the new experimental results described in Sections 2.3, 2.4, and 3.3 [56–58]. In the demonstration of the algorithm, the structure of the color VDP is found to be correct, but the parameters are poorly specified for the chromatic channels. In discussing these results, the chapter provides suggestions for improving the parameters, and concludes that the algorithm may provide a good foundation for the development of a single-number perceptual fidelity measure.

V. Conclusion

5.1 Research Contributions

The primary objective of this research was to produce a valid approach to assessing the perceptual fidelity of digital color images. In the quest to achieve this objective, several significant findings have been produced. Thus, the results of this research may be summarized as the following four original contributions:

1. A clear demonstration of color Mach bands, created by forming a Mach band stimulus pattern in one of the color-mediating channels of a color HVS model. The appearance of color Mach bands in this stimulus supports the assertion that low spatial frequencies are attenuated in the color channels of the HVS [57].
2. The modification of the colors in complex colored images to produce images that are perceived as colorful by normal observers, but monochrome by color blind observers. The modification is achieved by simply removing the variation of one of the color-mediating channels of a color HVS model from the image. The resulting perception of a monochrome image by color blind observers suggests that the color HVS model accurately models the separation of color information by the HVS into separate channels [56].
3. Experimental validation of a multiple-channel model of cortical processing in the HVS through the use of illusory contours. Formation of illusory contours in the output of the multiple-channel HVS model, in locations corresponding to those where illusory contours are perceived by human observers, supports the use of multiple-channel models of visual cortical processing [58].
4. The combination of a multiple-channel HVS model for assessing perceptual fidelity of monochrome images with a model of the color processing elements of the HVS to produce a model for assessing the perceptual fidelity of color images.

5.2 *Future Directions*

The contributions enumerated above provide foundation and direction for continued work in perceptual image fidelity, distance, and quality measures. All three of these are areas of great interest to the image processing community. Multiple-channel processing in the context of the Faugeras color model as performed in the color VDP algorithm provides a basis for developing each of these measures.

In the area of perceptual fidelity measures, two specific objectives remain open to exploration. First, the color VDP algorithm requires a refined model of masking effects in the chromatic channels. Section 4.5.3 provides some specific suggestions for developing this model. In addition, developing a meaningful technique of producing a single-number fidelity metric remains an important objective. The accomplishment of these objectives will lead to the development of a powerful tool for image processing system designers.

With an ever-increasing number of image databases for numerous applications, there is a growing need for a meaningful perceptual image distance metric. Such a metric would provide an automatic means of sorting through a set of images based on similarity to an image of interest, permitting image retrieval using a key image, rather than key words. Some work has already been accomplished in this area (for example, see the World Wide Web site, <http://www.virage.com>). In many ways, the measurement of perceptual distance between images is not much different than the measurement of perceptual fidelity. As work continues, the model developed in this research may therefore find direct application in improved approaches to the problem of measuring perceptual image distance.

Finally, there is the problem of perceptual image quality—the assessment of the subjective quality of an image based upon the image itself. This problem is of particular interest to designers of remote imaging and image communication systems, which typically do not allow the luxury of a reference image at the receiver for use in evaluating the received image. Therefore, techniques of assessing perceptual image quality in an absolute sense are very important in these applications, and few approaches have been proposed thus far. The contributions of this research are certain to provide valuable insight to those who are seeking to develop more accurate measures of perceptual image quality.

Appendix A. Generation of Gabor Filters

The purpose of this Appendix is to describe how to construct a Gabor filter for application in the FFT domain. It is intended to be of some use to any who may follow that have not used Gabor filters before. While the approach described is related specifically to the Gabor filters used in this research, the basic principles involved in developing these filters may be readily adapted to almost any application. A more general treatment may be found in Gaskill [29].

The Gabor function may be simply defined as a sinusoid modulated by a Gaussian envelope. Two types of Gabor functions are frequently considered: cosine-Gabors and sine-Gabors. As the name implies, the cosine Gabor is the product of a cosine and a Gaussian function:

$$G_{\cos}(x) = \exp \left[-\pi \left(\frac{x}{a} \right)^2 \right] \cos(2\pi f x), \quad (33)$$

where a is the spread parameter of the Gaussian and f is the frequency of the sinusoid. The sine-Gabor is similarly defined:

$$G_{\sin}(x) = \exp \left[-\pi \left(\frac{x}{a} \right)^2 \right] \sin(2\pi f x). \quad (34)$$

These two functions are illustrated for the one-dimensional case in Figure 28. Note that the cosine-Gabor achieves its maximum value at the peak location of the the Gaussian, while the sine-Gabor passes through zero at that location.

In the spatial frequency domain, sine- and cosine-Gabor functions are expressed in terms of a pair of Gaussians which are offset an equal distance in opposite directions from the spatial frequency origin, as depicted in Figure 29. The cosine-Gabor is obtained by dividing the sum of the two Gaussians by 2, while the sine-Gabor is obtained by dividing the difference by $2j$. The offset (ρ) and orientation (θ) of the Gaussians is determined by the spatial frequency of the sinusoid. In general, the major axes of the Gaussians forming a Gabor function are not constrained to lie along the radial direction, as Figure 29 indicates. However, following physiological data [15], this constraint is enforced in all of the Gabors used in this research. (This also turns out to be a rather convenient simplification!)

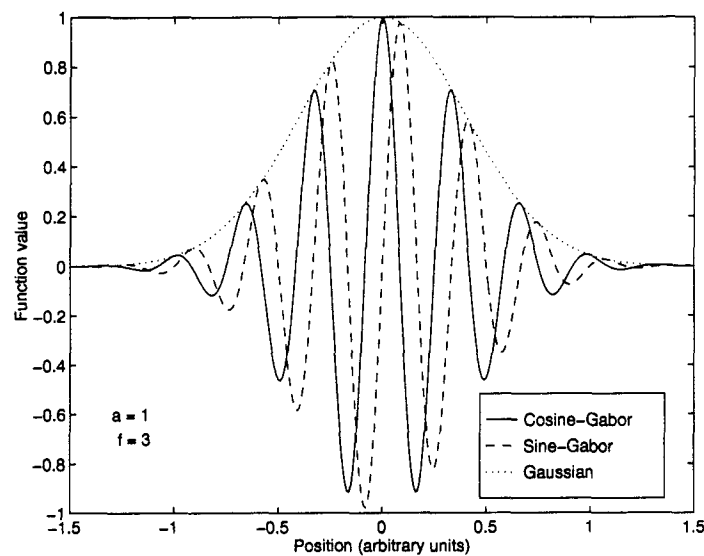


Figure 28. One-dimensional cosine- and sine-Gabor functions. a is the spread parameter of the Gaussian, f is the frequency of the sinusoid.

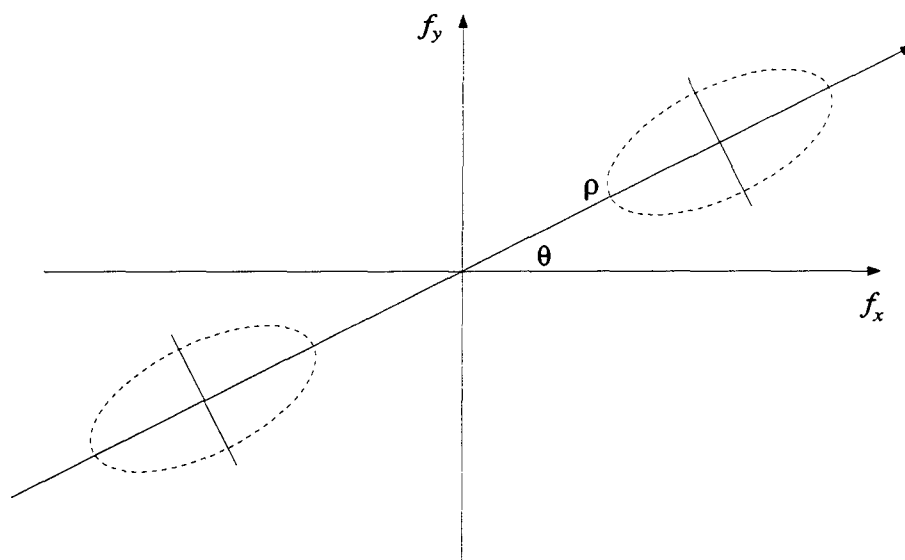


Figure 29. Depiction of a Gabor function in frequency space. The ellipses represent a constant contour, such as $1/e$, of the two Gaussian functions composing the Gabor.

With this background, it is now possible to consider the problem of computing a Gabor filter in the discrete Fourier transform domain. Essentially, this is accomplished by evaluating the two offset Gaussians at each spatial frequency represented in the discrete Fourier transform. This procedure will now be outlined.

First, the two-dimensional Gaussian function must be defined. In this research, Gaskill's definition is used [29]:

$$\text{Gaus}(u, v; a, b) = \exp \left\{ -\pi \left[\left(\frac{u}{a} \right)^2 + \left(\frac{v}{b} \right)^2 \right] \right\}. \quad (35)$$

This Gaussian function achieves a maximum value of unity at the origin, a value of $1/e$ at $u = \pm(a/\sqrt{\pi})$ and $v = \pm(b/\sqrt{\pi})$, and encloses a volume equal to the product of the spread parameters, $|ab|$. A sine- or cosine-Gabor function may be completely described in terms of this definition using four parameters: ρ , the radial offset from the origin to the peak of the Gaussian; θ , the orientation of the two-dimensional Gaussian; and a and b , the spread parameters describing the spread of the Gaussian along the radial direction (a) and perpendicular to it (b). Given these four parameters, the Gabor is obtained by computing one Gaussian with a positive offset along the radial axis, and another Gaussian with a negative offset. The Gabor function may then be obtained by forming the proper combination of these two Gaussians. Working from the definition of the Gaussian function in Equation 35, sine- and cosine-Gabor functions with arbitrary offset, orientation, and spread may thus be defined for any point in the spatial frequency plane. Because the application is ultimately in the discrete Fourier transform domain, an expression in terms of the rectangular spatial frequency coordinates (f_x, f_y) is desired.

Figure 30 provides a useful diagram for determining the appropriate modifications of Equation 35 to perform the desired rotation and shifts to produce the desired expression. First, note that by establishing the (u, v) coordinate system as shown in Figure 30, the Gaussian function represented by the ellipse in the figure is described by Equation 35. By expressing u and v in terms of f_x , f_y , ρ , and θ , the Gaussian function may be expressed in terms of the desired coordinate system. This transformation is accomplished by means of trigonometric identities.

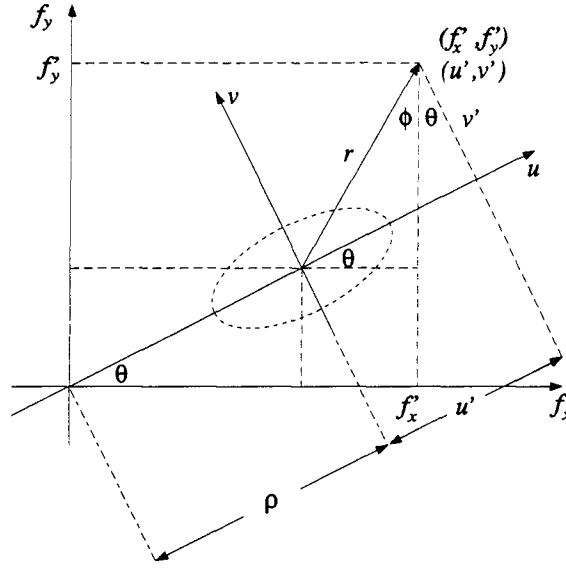


Figure 30. Transformation of variables diagram. The ellipse represents one of the two Gaussians composing a Gabor function. The other Gaussian is located in the opposite direction from the origin of the (f_x, f_y) plane.

The necessary transformation is obtained by choosing an arbitrary point in the spatial frequency plane, and then expressing the coordinates of that point in the (u, v) coordinate system, (u', v') , in terms of the coordinates of the point in the (f_x, f_y) coordinate system, (f'_x, f'_y) , together with the shift and orientation parameters, ρ and θ . Now, from the diagram in Figure 30, note that

$$\sin \phi = \frac{f'_x - \rho \cos \theta}{r}, \quad \text{and} \quad \cos \phi = \frac{f'_y - \rho \sin \theta}{r}. \quad (36)$$

The expression for u' in terms of f'_x, f'_y, ρ , and θ is obtained by first recognizing that

$$u' = r \sin(\theta + \phi). \quad (37)$$

Applying trigonometric identities and substituting the expressions from Equation 36 produces the following results:

$$u' = r(\sin \theta \cos \phi + \cos \theta \sin \phi) \quad (38)$$

$$\begin{aligned}
&= r \sin \theta \left(\frac{f'_y - \rho \sin \theta}{r} \right) + r \cos \theta \left(\frac{f'_x - \rho \cos \theta}{r} \right) \\
&= f'_y \sin \theta + f'_x \cos \theta - \rho.
\end{aligned}$$

Similarly, the expression for v' is obtained as follows:

$$\begin{aligned}
v' &= r \cos(\theta + \phi) \\
&= r(\cos \theta \cos \phi - \sin \theta \sin \phi) \\
&= r \cos \theta \left(\frac{f'_y - \rho \sin \theta}{r} \right) - r \sin \theta \left(\frac{f'_x - \rho \cos \theta}{r} \right) \\
&= f'_y \cos \theta - f'_x \sin \theta.
\end{aligned} \tag{39}$$

Substituting the results in Equations 39 and 40 into Equation 35, the Gaussian depicted in Figure 30 may be written as

$$\text{Gaus}(f_x, f_y; a, b) = \exp \left\{ -\pi \left[\left(\frac{f_y \sin \theta + f_x \cos \theta - \rho}{a} \right)^2 + \left(\frac{f_y \cos \theta - f_x \sin \theta}{b} \right)^2 \right] \right\}. \tag{40}$$

The other Gaussian composing the Gabor function depicted in Figure 29 is then obtained simply by replacing ρ by $-\rho$ in Equation 40. Thus, the cosine-Gabor function for a given set of ρ , θ , a , and b is written as

$$\begin{aligned}
G_{\cos}(f_x, f_y; a, b, \rho, \theta) &= \frac{1}{2} \exp \left(-\pi \left[\left(\frac{f_y \sin \theta + f_x \cos \theta - \rho}{a} \right)^2 + \left(\frac{f_y \cos \theta - f_x \sin \theta}{b} \right)^2 \right] \right) \\
&\quad + \frac{1}{2} \exp \left(-\pi \left[\left(\frac{f_y \sin \theta + f_x \cos \theta + \rho}{a} \right)^2 + \left(\frac{f_y \cos \theta - f_x \sin \theta}{b} \right)^2 \right] \right).
\end{aligned} \tag{41}$$

The sine-Gabor function is written similarly:

$$\begin{aligned}
G_{\sin}(f_x, f_y; a, b, \rho, \theta) = & \frac{1}{2j} \exp \left(-\pi \left[\left(\frac{f_y \sin \theta + f_x \cos \theta - \rho}{a} \right)^2 \right. \right. \\
& \left. \left. + \left(\frac{f_y \cos \theta - f_x \sin \theta}{b} \right)^2 \right] \right) \\
& - \frac{1}{2j} \exp \left(-\pi \left[\left(\frac{f_y \sin \theta + f_x \cos \theta + \rho}{a} \right)^2 \right. \right. \\
& \left. \left. + \left(\frac{f_y \cos \theta - f_x \sin \theta}{b} \right)^2 \right] \right).
\end{aligned} \tag{42}$$

Because the Gabor filters used in this research are constrained to have constant radial frequency bandwidths (measured in octaves) and orientation bandwidths, the spread parameters a and b are dependent upon the center radial spatial frequency ρ . For a 1.5 octave radial spatial frequency bandwidth and a 30° orientation bandwidth, $a = \sqrt{\pi}\rho/2$, and $b = a/2$. Thus, for each radial spatial frequency center, the necessary Gabor filter for each required orientation θ may be obtained by evaluating Equations 42 and 43 at each (f_x, f_y) pair (expressed in appropriate spatial frequency units) represented in the FFT domain, using the appropriate values of a and b for that radial spatial frequency center.

Bibliography

1. Anstis, Stuart M. "What does visual perception tell us about visual coding." *Handbook of Psychobiology* edited by Michael S. Gazzinga and Colin Blakemore, chapter 9, New York: Academic Press, 1975.
2. Ayer, Kevin W. *Gabor transforms for forward looking infrared image segmentation*. MS thesis, Air Force Institute of Technology, 1989.
3. Baldwin Jr., M. W. "Subjective sharpness of additive color pictures," *Proceedings of the IRE*, 39:1173-1176 (1951).
4. Blakemore, C. and F. W. Campbell. "On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images," *J. Physiology*, 203:237-260 (1969).
5. Bradley, A., et al. "Orientation and spatial frequency selectivity of adaptation to color and luminance gratings," *Vision Research*, 28:841-856 (1988).
6. Brown, W. R. J. and D. L. MacAdam. "Visual sensitivities to combined chromaticity and luminance differences," *J. Opt. Soc. Am.*, 39(10):808-834 (1949).
7. Campbell, F. W., et al. "The spatial selectivity of visual cells of cat and the squirrel monkey," *J. Physiology*, 204:120 (1969).
8. Campbell, F. W. and D. G. Green. "Optical and retinal factors affecting visual resolution," *J. Physiology*, 181:576-593 (1965).
9. Campbell, Fergus W. "The transmission of spatial information through the visual system." *The Neurosciences Third Study Program* edited by F. Worden and F. O. Schmitt, chapter 9, 95-103, Cambridge, MA: MIT Press, 1973.
10. Campbell, Fergus W. and J. G. Robson. "Application of Fourier analysis to the visibility of gratings," *J. Physiology*, 197:551-566 (1968).
11. Cole, G. R., et al. "Visual interactions with luminance and chromatic stimuli," *J. Opt. Soc. Am. A*, 7:128-140 (1990).
12. Cornsweet, Tom N. *Visual Perception*. New York: Academic Press, 1970.
13. Daly, Scott. "The Visible Differences Predictor: An algorithm for the assessment of image fidelity." *Digital Images and Human Vision* edited by Andrew B. Watson, chapter 14, 179-206, Cambridge, MA: MIT Press, 1993.
14. Daugman, John G. "Two dimensional spectral analysis of cortical receptive field profiles," *Vision Research*, 20:847-856 (1980).
15. Daugman, John G. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, 2(7):1160-1169 (1985).
16. Daugman, John G. "Complete discrete 2-d gabor transforms by neural networks for image analysis and compression," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 36(7):1169-1179 (1988).

17. Daw, Nigel W. "Visual response to gradients of varying colour and equal luminance," *Nature*, 203:215-216 (July 1964).
18. DePalma, J. J. and E. M. Lowry. "Sine wave response of the visual system, II. Sine-wave and square-wave contrast sensitivity," *J. Opt. Soc. Am.*, 52:328-335 (Mar 1962).
19. DeValois, K. K. and E. Switkes. "Simultaneous masking interactions between chromatic and luminance gratings," *J. Opt. Soc. Am.*, 73:11-18 (1983).
20. DeValois, R. L., et al. "Spatial frequency selectivity of cells in macaque visual cortex," *Vision Res.*, 22:545-559 (1982).
21. DeValois, R. L. and others. "Analysis of response patterns of LGN cells," *J. Opt. Soc. Am.*, 56(7):966-967 (Jul 1966).
22. Dvorine, Israel. *The Dvorine Color Perception Test*. Baltimore: The Waverly press, inc., 1944.
23. Ehrenstein, Walter. "Über Abwandlungen der L. Hermannschen Helliskeiterscheinung," *Zeitschrift für Psychologie*, 150:83-91 (1941). English translation by Anne Hogg (1987) "Modifications of the brightness phenomenon of L. Hermann," in *The Perception of Illusory Contours*, Edited by Susan Petry and Glenn Meyer, (New York:Springer-Verlag), pp. 35-39.
24. Eskicioglu, Ahmet M. and Paul S. Fisher. "Image quality measures and their performance," *IEEE Transactions on Communications*, 43(12):2959-2965 (1995).
25. Faugeras, Olivier D. *Digital Color Image Processing and Psychophysics Within the Framework of a Human Visual Model*. PhD dissertation, University of Utah, Salt Lake City, June 1976.
26. Foley, J. M. "Human luminance pattern-vision mechanisms: masking experiments require a new model," *J. Opt. Soc. Am. A*, 11:1710-1719 (1994).
27. Foley, J. M. and G. M. Boynton. "A new model of human luminance pattern-vision mechanisms: Analysis of the effects of pattern orientation, spatial phase, and temporal frequency." *SPIE Proceedings: Computational Vision Based on Neurobiology 2054*, edited by T. A. Lawton. 1994.
28. Frei, Werner. "Rate-distortion coding simulation for color images." *SPIE Proceedings: Advances in Image Transmission Techniques 87*. 197-203. August 1976.
29. Gaskill, Jack D. *Linear Systems, Fourier Transforms, and Optics*. New York: John Wiley & Sons, 1978.
30. Georgeson, M. A. and J. M. Georgeson. "Facilitation and masking of briefly presented gratings: Time course and contrast dependence," *Vision Research*, 27:369-379 (1987).
31. Ginsburg, Arthur P. "Is the illusory triangle physical or imaginary," *Nature*, 257(5523):219-220 (September 1975).
32. Glasser, L. G. and others. "Cube-root color coordinate system," *J. Opt. Soc. Am.*, 48(10):736-740 (Oct 1958).

33. Grossberg, S. and E. Mingolla. "Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading," *Psychological Review*, 92:173-211 (1985).
34. Grossberg, S. and E. Mingolla. "Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations," *Perception and Psychophysics*, 38:141-171 (1985).
35. Hall, Charles F. *Image Filtering Based on Psychovisual Characteristics of the Human Visual System*. Technical Report 740, Image Processing Institute, University of Southern California, March 1977.
36. Hall, Charles F. *Digital Color Image Compression in a Perceptual Space*. PhD dissertation, University of Southern California, 1978.
37. Hall, Charles F. and Ernest L. Hall. "A nonlinear model for the spatial characteristics of the human visual system," *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-7(3):161-170 (March 1977).
38. Heeger, David J. and Patrick C. Teo. "A model of perceptual image fidelity." *Proceedings of the 1995 IEEE International Conference on Image Processing*. 343-345. 1995.
39. Hilz, R. and C. R. Cavonius. "Wavelength discrimination measured with square-wave gratings," *J. Opt. Soc. Am.*, 60:273-277 (1970).
40. Hood, Donald C. and John A. Whiteside. "Brightness of ramp stimuli as a function of plateau and gradient widths," *J. Opt. Soc. Am.*, 58:1310-1311 (1968).
41. Hubel, D. H. and T. N. Wiesel. "Receptive fields of single neurones in the cat's striate cortex," *J. Physiology (London)*, 148:574-591 (1959).
42. Hubel, D. H. and T. N. Wiesel. "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *J. Physiology (London)*, 160:106-154 (1962).
43. Hubel, D. H. and T. N. Wiesel. "Receptive fields and functional architecture in two non-striate visual areas (18 and 19) of the cat," *J. Neurophysiology*, 28:229-289 (1965).
44. Hubel, D. H. and T. N. Wiesel. "Receptive fields and functional architecture of monkey striate cortex," *J. Physiology (London)*, 195:215-243 (1968).
45. Jameson, D. "Theoretical issues of color vision." *Handbook of Sensory Physiology VII/4*, edited by D. Jameson and L. M. Hurvich, Springer-Verlag, 1972.
46. Jones, J. P. and L. A. Palmer. "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophys.*, 58:1233-1258 (1987).
47. Kanizsa, Gaetano. "Margini Quasi-percettivi in campi con Stimolazione Omogenea," *Rivista di Psicologia*, 49:7-30 (1955). English translation by Walter Gerbino (1987) "Quasi-perceptual margins in homogeneously stimulated fields," in *The Perception of Illusory Contours*, Edited by Susan Petry and Glenn Meyer, (New York:Springer-Verlag), pp. 40-49.

48. Kelly, D. H. "Effects of sharp edges on the visibility of sinusoidal gratings," *J. Opt. Soc. Am.*, 60:98-103 (1970).
49. Kulikowski, J. J., et al. "Theory of spatial position and spatial frequency relations in the receptive fields of simple cells in the visual cortex," *Biol. Cybern.*, 43:187-198 (1982).
50. Lampinen, Jouko and Erkki Oja. "Distortion tolerant pattern recognition based on self-organizing feature extraction," *IEEE Transactions on Neural Networks*, 6(3):539-547 (May 1995).
51. Legge, G. E. and J. M. Foley. "Contrast masking in human vision," *J. Opt. Soc. Am.*, 70:1458-1471 (1980).
52. Losada, M. Angeles and Kathy T. Mullen. "The spatial tuning of chromatic mechanisms identified by simultaneous masking," *Vision Research*, 34:331-341 (1994).
53. Losada, M. Angeles and Kathy T. Mullen. "Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking," *J. Opt. Soc. Am. A*, 12:250-260 (1995).
54. Mannos, James L. and David J. Sakrison. "The effects of a visual fidelity criterion on the encoding of images," *IEEE Transactions on Information Theory*, IT-20(4):525-536 (July 1974).
55. Marčelja, S. "Mathematical description of the responses of simple cortical cells," *J. Opt. Soc. Am.*, 70:1297-1300 (1980).
56. Martin, Curtis E., et al. "Color blindness and a human color vision model." Submitted to *IEEE Transactions on Systems, Man, and Cybernetics*, Feb 1996.
57. Martin, Curtis E., et al. "Digital production of color Mach bands using a color human visual system model." Accepted for publication in *IEEE Transactions on Systems, Man, and Cybernetics*, Aug 1996.
58. Martin, Curtis E., et al. "Illusory contour formation by a multiple channel human visual system model." Submitted to *Vision Research*, Oct 1996.
59. Mayer, M. J. and C. W. Tyler. "Invariance of the slope of the psychometric function with spatial summation," *J. Opt. Soc. Am. A*, 3:1166-1172 (1986).
60. Mullen, Kathy T. "The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings," *J. Physiol.*, 359:381-409 (1985).
61. Mullen, Kathy T. and M. Angeles Losada. "Evidence for separate pathways for color and luminance detection mechanisms," *J. Opt. Soc. Am. A*, 11:3136-3151 (1994).
62. Nachmias, J. "On the psychometric function for contrast detection," *Vision Research*, 21:215-223 (1981).
63. Newton, Isaac. *Opticks: Or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*. London: S. Smith and B. Walford, 1704.
64. Nill, Norman B. "A visual model weighted cosine transform for image compression and quality assessment," *IEEE Transactions on Communications*, COM-33(6):551-557 (Jun 1985).

65. Nill, Norman B. and Brian H. Bouzas. "Objective image quality measure derived from digital image power spectra," *Optical Engineering*, 31(4):813-825 (Apr 1992).
66. Olzak, L. A. and J. P. Thomas. "Seeing spatial patterns." *Handbook of Perception and Human Performance* edited by K. Boff, et al., New York: Wiley, 1986.
67. Ophthalmology Branch, Clinical Sciences Division, USAF School of Aerospace Medicine,. *Aerospace Ophthalmology, II*, chapter 15. 1988.
68. Pandey, R. and R. L. P. Vimal. "Threshold elevation curves for the red-green channel estimated by oblique masking," *Invest. Ophthalmol. Vis. Sci. Suppl.*, 34:751 (1993).
69. Patel, A. S. "Spatial resolution by the human visual system," *J. Opt. Soc. Am.*, 56:689-694 (May 1966).
70. Pearlman, William A. "A visual system model and a new distortion measure in the context of image processing," *J. Opt. Soc. Am.*, 68(3):374-386 (March 1978).
71. Pease, Paul L. "On Color Mach Bands," *Vision Research*, 18:751-755 (1978).
72. Pelli, D. *Effects of Visual Noise*. PhD dissertation, Cambridge University, England, 1981.
73. Perona, Pietro. "Deformable kernels for early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5):488-499 (1995).
74. Phillips, G. and H. Wilson. "Orientation bandwidths of spatial mechanisms measured by masking," *J. Opt. Soc. Am. A*, 1:226-232 (1984).
75. Pollehn, J. and H. Roehrig. "Effect of noise on the modulation transfer function of the visual channel," *J. Opt. Soc. Am.*, 60:842-848 (1970).
76. Pollen, D. A. and S. F. Ronner. "Phase relationships between adjacent simple cells in the visual cortex," *Science*, 212:1409-1411 (1981).
77. Post, David L. "Personal communication." October 1995.
78. Ratliff, Floyd. *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. San Francisco: Holden-Day, 1965.
79. Ratliff, Floyd. "Why Mach bands are not seen at the edges of a step," *Vision Research*, 24:163-165 (1984).
80. Ross, John, et al. "The conditions under which Mach bands are visible," *Vision Research*, 29:699-715 (1989).
81. Rushton, W. A. H. "Peripheral Coding in the Nervous System." *Sensory Communication* edited by W. A. Rosenblith, Cambridge, MA: MIT Press, 1961.
82. Schade Sr., O. H. "On the quality of color television images and the perception of color detail," *J. Soc. Mot. Pict. Tel. Eng.*, 67:801-819 (1958).
83. Stevens, S. S. "The Psychophysics of Sensory Function." *Sensory Communication* edited by W. A. Rosenblith, Cambridge, MA: MIT Press, 1961.
84. Stevens, S. S. "Neural events and the psychophysical law," *Science*, 170(3962):1043-1050 (December 1970).

85. Stevens, S. S. *Psychophysics*. New York: Wiley, 1975.
86. Stiles, W. L. "A modified Helmholtz line element in brightness-colour space," *Proc. Phys. Soc. (Lond.)*, 58:41-65 (1946).
87. Stockham Jr., Thomas G. "Image processing in the context of a visual model," *Proceedings of the IEEE*, 60(7):828-842 (Jul 1972).
88. Stork, David G. and Hugh R. Wilson. "Do Gabor functions provide appropriate descriptions of visual cortical receptive fields?," *J. Opt. Soc. Am. A*, 7(8):1362-1373 (1990).
89. Stromeyer III, C. F. and B. Julesz. "Spatial-frequency masking in vision: critical bands and spread of masking," *J. Opt. Soc. Am.*, 62:1221-1232 (1972).
90. Switkes, E., et al. "Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings," *J. Opt. Soc. Am. A*, 5:1149-1162 (1988).
91. Teo, Patrick C. and David J. Heeger. "Perceptual image distortion." *SPIE Proceedings* 2179. 127-141. 1994.
92. Thomas, C. W., et al. "Models of contrast sensitivity in human vision," *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3):857-864 (May/Jun 1993).
93. Thomas, J. P. "Effect of static-noise and grating masks on detection and identification of grating targets," *J. Opt. Soc. Am. A*, 2:1586-1592 (1985).
94. Tomita, T. "Electrical Response of Single Photo Receptors," *Proc. IEEE*, 56:1015-1023 (1968).
95. van der Horst, G. J. C. and M. A. Bouman. "Spatio-temporal chromaticity discrimination," *J. Opt. Soc. Am.*, 59:1260-1266 (1969).
96. van der Horst, G. J. C., et al. "Transfer of spatial chromaticity-contrast at threshold in the human eye," *J. Opt. Soc. Am.*, 57:1260-1266 (1967).
97. van Meeteren, A. and J. M. Valetton. "Effects of pictorial noise interfering with visual detection," *J. Opt. Soc. Am. A*, 5:438-444 (1988).
98. Van Ness, F. L. and M. A. Bouman. "The effects of wavelength and luminance on visual modulation transfer." *Excerpta Medica Int. Congress series* 125, 183-192, 1965.
99. von Bekesy, Georg. "Mach- and Hering-type lateral inhibition in vision," *Vision Research*, 8:1483-1499 (1968).
100. von Helmholtz, Herman. *Handbuch der Physiologischen Optik*, 2. Hamburg: Voss, 1911.
101. Ware, Colin and William B. Cowan. "The chromatic Cornsweet effect," *Vision Research*, 23(10):1075-1077 (1983).
102. Watson, Andrew. B. "The Cortex transform: Rapid computation of simulated neural images," *Computer Vision, Graphics, and Image Processing*, 39:311-327 (1987).
103. Watson, Andrew. B. "Efficiency of an image code based on human vision," *J. Opt. Soc. Am. A*, 4(12):2401-2417 (1987).

104. Westen, S. J. P., et al. "Perceptual Image Quality Based on a Multiple Channel HVS Model." *Proceedings of the 1995 International Conference on Acoustics, Speech, and Signal Processing* 4. 2351-2354. 1995.
105. Wyszecki, G. and W. S. Stiles. *Color Science*. New York: John Wiley, 1975.
106. Xie, Zhenhua and Thomas G. Stockham Jr. "Toward the unification of three visual laws and two visual models in brightness perception," *IEEE Transactions on Systems, Man, and Cybernetics*, 19:379-387 (1989).
107. Young, Thomas. "The Bakerian lecture: on the theory of lights and colours," *Philos. Trans. R. Soc. Lond.*, 92:12-48 (1802).
108. Zeki, Semir. *A Vision of the Brain*. Boston: Blackwell Scientific Publications, 1993.

Vita

Captain Curtis E. Martin [REDACTED] 10 September 1968 in [REDACTED]. He graduated *magna cum laude* from Brigham Young University in 1991 with a Bachelor of Science degree in Electrical Engineering, and accepted a commission as a Second Lieutenant in the United States Air Force. While awaiting an active duty assignment, Lieutenant Martin worked as a software engineer at Space Dynamics Laboratory at Utah State University. In May of 1992, Lieutenant Martin was assigned to the Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio, to pursue a Master of Science degree in Electrical Engineering. Upon completion of the Master's program, Lieutenant Martin stayed at AFIT to pursue the PhD degree in Electrical Engineering. After graduation, Captain Martin will be assigned to the Optical Radiation branch of the Occupational and Environmental Health Directorate of Armstrong Laboratory at Brooks AFB, TX.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE December 1996		3. REPORT TYPE AND DATES COVERED PhD Dissertation
4. TITLE AND SUBTITLE PERCEPTUAL FIDELITY FOR DIGITAL COLOR IMAGERY			5. FUNDING NUMBERS	
6. AUTHOR(S) Curtis E. Martin				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Air Force Institute of Technology, WPAFB OH 45433-6583			8. PERFORMING ORGANIZATION REPORT NUMBER AFIT/DS/ENG/96-14	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Capt Rich Sumner NIMA/SUADE 2420 Vela Way Suite 1467-A5 Los Angeles AFB, CA 90245-4659			10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; Distribution unlimited			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The problem of measuring the fidelity of digital color images in a manner that corresponds to human perceptual assessments is addressed. Experiments are performed to validate human visual system (HVS) models, which provide access to a "perceptual space" in which visual distortions may be measured, and then a model is proposed for assessing the perceptual fidelity of digital color image. Color Mach bands are produced in the first experiment, demonstrating that, as in the brightness channel, low spatial frequency attenuation occurs in the chromatic channels of the HVS. In the second experiment, a correlation between the chromatic channels of the HVS model and color discrimination axes of color blind observers is demonstrated. Removing variation from one of the chromatic channels of a natural image produces a color-distorted image which the color blind subjects cannot distinguish from the original. Removing variation from the other chromatic channel produces an image that appears colorful to normally-sighted observers, but monochrome to the color blind observers. The third experiment shows that a Gabor filter-based HVS model produces illusory contours in several illusory contour stimuli. These results provide a unique validation of multiple-channel HVS models which process the image in multiple spatial frequency bands that are tuned to match measured sensitivities of neurons in the primary visual cortex of cats and monkeys. Finally, the multiple-channel processing used in the illusory contour experiment is combined with the color vision model from the first two experiments to produce a multiple-channel, color HVS model for measuring perceptual fidelity of color images. A demonstration of the model shows that the structure of the new model is correct.				
14. SUBJECT TERMS Color Vision, Color Mach Bands, Multiple-Channel Vision Models, Color Image Fidelity, Illusory Contours			15. NUMBER OF PAGES 113	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT SAR	