A HUMAN SPATIAL-CHROMATIC VISION MODEL FOR EVALUATING ELECTRONIC DISPLAYS

by

Charles J. C. Lloyd

Dissertation submitted to the Graduate Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Industrial Engineering and Operations Research

APPROVED:

R. J. Beaton, Chairman

H. L. Snyder

R. W. Conners

M. Deisenroth

A. M. Prestrude

April, 1990

Blacksburg, Virginia

A HUMAN SPATIAL-CHROMATIC VISION MODEL FOR EVALUATING ELECTRONIC DISPLAYS

by

Charles J. C. Lloyd

Committee Chairman: Robert J. Beaton

Industrial Engineering and Operations Research

(ABSTRACT)

This dissertation examines those attributes of full-color display systems (particularly color matrix displays) which degrade image quality. Based on this analysis, it is suggested that a comprehensive metric should measure image quality in terms of transmitted signal and noise modulation, both achromatic and chromatic. Moreover, it is suggested that these signal and noise measurements be weighted in terms of human spatial-chromatic visual characteristics.

A review of extant image quality metrics reveals several limitations of these metrics which make them unsuitable for the evaluation of color matrix displays. These limitations include the inability to account for chromatic modulation transfer and chromatic noise as well as the general inability to account for spatial and grey-scale sampling.

This work describes a new methodology for assessing image quality that can be applied to full-color as well as monochromatic, and sampled as well as continuous, display systems. Unlike most display quality metrics, the proposed methodology is not based on

the tools of linear systems analysis. Rather, it is based on more veridical models of the human visual system (HVS), including multi-channel models of spatial vision, the zone theory of color vision, physiological models of retinal processes, and models of the optics of the eye.

A display evaluation *system* consisting of the HVS model used in conjunction with a display simulator is described. The HVS model employs nine image processing stages to account for nonlinear retinal processes, opponent color encoding, and multiple spatial frequency channels. A detailed procedure for using the HVS model to evaluate display systems is provided.

The validity of the HVS model was tested by conducting contrast detection, discrimination, and magnitude estimation experiments on the model. The results of these experiments correspond closely with published human performance data. The utility of the display evaluation system was assessed by making image quality predictions for the display systems used in three image quality studies. Image quality predictions using the proposed system correlate strongly with ratings of image quality provided by human subjects. Results of these validation studies indicate that the proposed method of display evaluation is viable and warrants further development.

ACKNOWLEDGEMENTS

Soon after coming to Virginia Tech I became keenly interested in models of human visual processing and the design of visual display systems. This interest resulted largely from reading several excellent technical reports related to image quality which had come out of the Displays and Controls Laboratory over the previous decade. During my stay at Virginia Tech I had the good fortune of being able to work on several interesting funded projects in the Displays and Controls Laboratory. When the time came to start dissertation research, however, none of the ongoing contracts sufficiently matched the specific project I had in mind. But due to the support of several individuals I was given the unique opportunity to conduct my own line of research.

First of all I thank my wife Ching-Mang for her moral and financial support without which I could not have gone unfunded in order to pursue my highly specific interests. Second, I could not have conducted this research without the full support of Dr. Robert J. Beaton and Dr. Harry L. Snyder, who have devoted their time and resources to this project. Thanks also to Willard W. Farley and all of the members of the Displays and Controls Laboratory who have shared time, expertise, equipment, and facilities with me. Finally, thanks to Dr. Richard W. Conners and Sue-Ellen Cline who have supported me in the use of the Spatial Data Analysis Laboratory.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Trends in Image Quality

A great deal of work currently underway in the United States, Japan, and Europe is focused on the development and production of affordable full-color flat-panel displays. Currently available flat-panel devices provide some significant advantages over cathode ray tubes (CRTs), including reduced volume, weight, and power consumption, improved sunlight readability, and improved ruggedness. Unfortunately, color flat-panel displays generally exhibit poor image quality when compared to color CRTs, especially when used for the presentation of images where high resolution, full-color, and good control of grey scale are required. While ongoing research and development likely will improve image quality for these devices, it may be some time before the improved devices become available commercially.

Two major reasons for the relatively poor performance of currently available flat panel devices are reduced resolution and "sharp" pixel luminance profiles. A casual look at the display design literature suggests that image quality of color flat-panel devices rivals or surpasses that of color CRTs. Papers have been published describing color flat panels with "212 lines per inch" (Shibusawa, Kondo, Kigoshi, Inoue, Yoshifuku, Shimizu, and Hori, 1989), "1440 x 1100 dots" (Ichikawa, Suzuki, Matino, Aoki, Higuchi, and Oana, 1989), and "17-inch diagonals" (Friedman, Rahman, Peters, Wedding, and Repetti, 1989). However, a more careful examination of this literature reveals that color flat-panel displays which have *both* a high element density and a large active area do not yet exist.

It is difficult to compare color CRTs with color flat-panel devices by comparing simply

resolution, addressability, and active area size. This difficulty is exacerbated by the fact that designers of flat-panel devices use the term "pixel" to refer to a single color element; whereas the term pixel describes a group of red, green, and blue color elements on a CRT display. This inconsistent use of terms favors flat-panel devices by making it seem that the area of pixels is a factor of three or four smaller.

One way to simplify comparisons across displays that differ in multiple ways (e.g., resolution, addressability, size, technology) is to assume that each display will be viewed from a distance that results in an overall display angle of some constant value. The use of a constant display angle controls for variable display sizes and allows measurements of display features to be scaled in terms of visual angle. Measurements made in terms of visual angle allow for more direct comparisons with human performance data than do linear measurements. This approach for comparing devices is used below.

A 19-in. color CRT commonly used in graphics systems might have the following specifications: 1024 x 1024 pixels, 0.30 x 0.30 m active area, 0.26 mm mask pitch, and a 0.50 mm spot size (W. W. Farley, personal communication, June 5, 1989; Oakley, 1984). The resolution (pixel size), addressability (center-to-center spacing of pixels), and size of the smallest color element are shown in Table 1 for this device. The units of these measures are expressed in terms of visual angle in minutes of arc for an overall display angle of 24 deg (0.7 m viewing distance).

Also appearing in Table 1 are specifications for three of the latest high information content full-color, flat-panel displays for which a working prototype has been built (Moriyama et al., 1989; Shibusawa et al., 1989; and Yamauchi et al., 1989). As with the

Table 1. Comparison of Basic Parameters of Full-Color CRTs and Flat-Panels

Display Type	Resolution (pixel size)	Addressability (center-to-center)	Color Element Size
CRT	2.4 min dia	1.4 min	0.64 min dia
Flat-Panel (Moriyama et al., 1989)	5.2 min x 5.2 min	5.2 min x 5.2 min	1.7 min x 5.2 min
Flat-Panel (Shibusawa et al., 1989)	6.0 min x 6.0 min	6.0 min x 6.0 min	3.0 min x 3.0 min
Flat-Panel (Yamauchi et al., 1989)	5.1 min x 5.1 min	5.1 min x 5.1 min	5.1 min x 1.7 min

CRT, visual angles were calculated using an overall display angle of 24 deg. For these devices a "pixel" refers to a group of three or four color elements depending on the pixel geometry of the device.

Examination of Table 1 reveals that state-of-the-art flat-panel displays lag far behind commercially available color CRTs in terms of overall image transmission capability. For displays subtending the same visual angle, the best flat-panels have approximately one-fourth the addressability, pixels 2.3 times larger, and color elements 4.7 times larger than those of color CRTs. Additionally, it may take years of manufacturing process development before flat-panel devices of this capability can be produced economically.

The relatively low resolution of full-color flat-panel devices leads to the creation of image artifacts due to under-sampling and aliasing. Aliasing refers to the inadvertent production of sum and difference frequency components in an image which occurs when spatial frequencies in the original image are greater than half the pixel frequency (the Nyquist sampling limit) of the display. Aliasing noise common to flat-panel devices includes the generation of visible Moire´ patterns for periodic image components and "jaggies," "stair-stepping," or "barber polling" for lines which are not parallel to the tessellation structure of pixels on the display. Moreover, the visibility of pixel noise on low resolution displays is increased since it occurs at lower spatial frequencies where the visual system is most sensitive.

The relatively poor resolution of full-color flat-panel displays is compounded by the fact that these devices typically use rectangular pixel and color element luminance profiles. For displays with a sampling frequency less than approximately 20 pixels/deg (480 pixels for a 24 deg display) the luminance profile of pixels becomes important. Assuming a

square pixel profile and 20 pixels/deg, the first harmonic of 60 c/deg should be near the threshold of visibility. These high-frequency components of the image are a factor of six higher than the Nyquist sampling limit (10 c/deg) of the display. They should be considered as noise since they do not have any counterpart in the original image. With CRTs this type of pixel structure noise is less important since the Gaussian-shaped pixel profiles produce less spurious high-frequency energy.

With the use of rectangular luminance profiles, flat-panel devices generally have high-contrast lines between rows and columns of picture elements. These lines can produce strong luminance modulation at twice the Nyquist sampling limit of the display. Like the pixel edge noise discussed above, this pixel grid modulation should be considered noise since it has no counterpart in the original image. In contrast to a flat-panel device, a properly designed raster-scanned CRT which is adjusted for a resolution to addressability ratio (RAR) of close to 1.0 (Murch and Beaton, 1988; Murch, Virgin, and Beaton, 1985) will produce little visible noise.

The large-scale introduction of full-color flat-panel devices over the next few years will be a mixed blessing. On one hand, the reduced size, weight, and power consumption of these devices will allow them to be used in products such as portable personal computers and televisions for which CRTs would be impractical. On the other hand, the use of these devices will take the industry a step backwards with respect to the general trend in the improvement of image quality for displays in general.

If manufacturers could produce affordable display systems which exceed the capabilities of human vision in spatial, chromatic, and temporal bandwidth, there would be little need for a comprehensive model of the effects of color display system design

parameters on image quality. Given the capability to produce perceptually perfect displays there would be little need to examine the trade-offs among design factors since design trade-offs would be unnecessary. Unfortunately, even state-of-the-art color CRTs do not match human spatial, chromatic, and temporal capability for large area displays. With the expected general decrease in image quality that will accompany the increased use of flat-panel devices, the need for comprehensive methods of evaluation for color displays will become more acute.

Examination of Conventional Image Quality Metrics

Many metrics have been developed over the past three decades for assessments of image quality for various types of monochromatic display systems. Since these conventional metrics largely are inappropriate for assessment of flat-panel display systems, only a few of the most important of these metrics will be discussed here. Listed below are a number of critical reviews of many of the metrics which have been developed over the past two decades.

A good introduction to the general area of image quality assessment and to the research published prior to 1973 can be found in Biberman (1973). Side-by-side comparisons among the most important metrics have been conducted by Beaton (1984) and Task (1979). A recent review of a wide range of monochromatic image quality metrics has been conducted by Decker, Pigion, and Snyder (1987). Recent general discussions of the capabilities and limitations of conventional metrics as well as attempts to address color flat-panel display design issues were provided by Snyder (1985, 1988).

Three of the most important extant image quality assessment methods include the

modulation transfer function area (MTFA), signal-to-noise ratio (SN), and just noticeable difference (JND). The MTFA has been selected since it is characteristic of a large number of perceptually-weighted MTF-based metrics. Of the metrics in its class, the MTFA has demonstrated the best performance for metrics used in an image-dependent manner (Beaton, 1984). Additionally, the MTFA is the most thoroughly tested of all image quality metrics. The method has been so successful in predicting performance that it has been selected for use as an acceptance standard for video display terminals by the American National Standards Institute and the Human Factors Society (HFS, 1988).

The SN metric was selected since it differs qualitatively from most other metrics in that it attempts to account for noise directly rather than through the measurement of a separate contrast threshold function (CTF) or discriminable difference diagram (DDD) for each noise condition. Additionally, the SN metric demonstrated the best overall performance for image-independent metrics (Beaton, 1984).

Perhaps the most unique of all of the metrics developed to date is the JND-analysis approach developed by Cohen and Gorog (1974) and promoted by Carlson and Cohen (1978). This method was selected since suprathreshold as well as threshold visual responses are modeled. The recent papers of Barten (1989a, 1989b) provided additional support for the utility of incorporating suprathreshold response characteristics in quality metrics. Additionally, the method has strong intuitive appeal and seems to have good potential for extension to color display design issues.

The empirical modeling approach (Snyder and Maddox, 1978) has been selected since it has provided the most useful models for predicting human performance using dot-matrix displays.

MTFA. The MTFA metric was first introduced by Charman and Olin (1965) for use in evaluating photographic images. Since its introduction the metric has been evaluated by a number of researchers (e.g., Beaton, 1984; Blumenthal and Campana, 1981; Borough, Fallis, Warnock, and Britt, 1967; Gutmann, 1982; Snyder, 1973, 1974, 1976; Snyder, Keesee, Beamon, and Aschenbach, 1974; and Task, 1979). The MTFA provides a concise way to account for the overall capability of an imaging system as well as basic human visual performance.

The MTFA concept is illustrated in Figure 1 and is defined as the area bounded by the MTF of the imaging system, the CTF of the observer, zero spatial frequency, and the crossover frequency. The crossover frequency occurs where the MTF and CTF curves cross. This frequency defines the limiting resolution of the display device. Mathematically the MTFA can be expressed as:

$$\text{MTFA} = \int_{\omega=0}^{F_c} [F_s(\omega) - T(\omega)]\, d\omega \qquad (1)$$

where

ω = spatial frequency,

F_C = crossover frequency,

$F_S(\omega)$ = value of MTF at spatial frequency ω,

$T(\omega)$ = value of CTF at spatial frequency ω.

A concern with the MTFA metric expressed by numerous researchers is that the area between the CTF and the MTF is assumed to be perceptually homogeneous (Beamon and Snyder, 1975; Beaton, 1984; Gutmann, 1982; Gutmann, Snyder, Farley, and Evans, 1979; Task, 1979). With the MTFA metric, the area between the MTF and CTF is

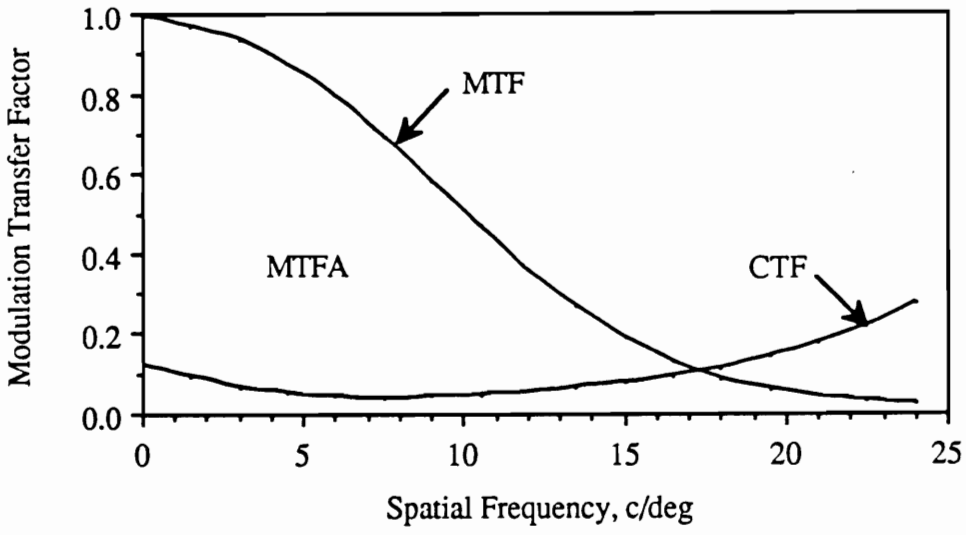


Figure 1. Illustration of the MTFA concept.

uniformly weighted. Experimental results obtained by Beamon and Snyder (1975) suggest that the area just above the CTF should be weighted more heavily than the area well above the CTF. Similarly, Gutmann (1982) has shown that displays with the same value of MTFA may differ in image quality. Though there is evidence that the MTFA is heterogeneous, nearly all studies using the metric have resulted in significant correlations with human performance.

SN. Building on the work of Hufnagle (1965), Beaton (1984) developed a signal-to-noise metric (SN) much like Hufnagle's Q3 metric. Beaton's SN metric represents an improvement over the Q3 metric since the SN metric is more easily calculated and shows a stronger correlation with performance. Like the MTFA and JND-analysis approaches, the SN metric uses the MTF to account for the imaging capability of the display. Unlike most other metrics, however, the SN metric relies on direct measurement of display noise rather than accounting for noise through the measurement of a CTF for each noise condition. Noise is accounted for within the metric using the Wiener noise power spectrum. Mathematically the SN metric is defined as:

$$SN = \frac{\int_0^{\omega_c}\int_0^{\upsilon_c} T_s(\omega, \upsilon)^2\, T_e(\omega, \upsilon)^2 \; \partial\omega\, \partial\upsilon}{\sqrt{\int_0^{\omega_c}\int_0^{\upsilon_c} W_i(\omega, \upsilon)^2\, T_e(\omega, \upsilon)^2 \; \partial\omega\, \partial\upsilon}} \qquad (2)$$

where

ω and υ are spatial frequency,

$T_e(\omega, \upsilon)$ = the two-dimensional CTF of the observer,

$T_S(\omega, \upsilon)$ = the two-dimensional MTF of the display,

$W_i(\omega, \upsilon)$ = the two-dimensional Wiener noise power spectrum.

In a side-by-side comparison among 16 image quality metrics, the SN metric yielded the highest overall correlation with performance (Beaton, 1984).

JND-analysis. Building on the previous work of Cohen and Gorog (1974), Carlson and Cohen (1978) developed a model to predict just noticeable differences (JNDs) in contrast for sine-wave gratings. Incorporated into their model is the idea that the human visual system separates the visual stimulus into independent bands of spatial frequency. At the heart of the JND-analysis approach is the use of Discriminable Difference Diagrams (DDDs). A DDD is a diagram which indicates how the size of just detectable modulation difference varies as a function of modulation and as a function of spatial frequency. Since contrast discrimination functions vary as a function of noise, a different DDD was developed for each of a number of noise conditions. Similarly, it was found that separate DDDs were required for evaluating CRTs as opposed to continuous photographic images. Thus, Carlson and Cohen developed and published 117 DDDs in their 1978 work.

The JND-analysis approach is perhaps the most sophisticated of the MTF-based metrics. The metric incorporates a more explicit model of human visual processing than any other display image quality metric. With the use of separate frequency bands and a different nonlinear JND function in each band the representation of supra-threshold human visual performance seems more reasonable than the perceptually homogeneous representation used in the MTFA metric.

While the JND-analysis approach is more sophisticated than other approaches, it is also more complex. Research reports describing the use of the metric indicate that the model

performs well. Unfortunately, few comparisons of this metric with other metrics have been published (Cohen and Gorog, 1978). Thus, it is difficult to assess whether the greater complexity of the JND-analysis approach is outweighed by better performance.

Taking off on the idea that suprathreshold visual sensitivity should be scaled nonlinearly, Barten (1987) has proposed the use of the Square-Root Integral (SQRI) metric. Like the MTFA and SN metrics the SQRI metric uses the MTF of a display and the CTF for human vision. Calculation of the SQRI metric is defined by Equation 3.

$$J = \frac{1}{\ln(2)} \int_0^{u_{max}} \sqrt{\frac{M}{M_t}} \, \frac{\partial u}{u} \tag{3}$$

where

M is the MTF of the display,

M_t is the CTF for human vision,

and u is spatial frequency.

Examination of Equation 3 indicates that the ratio of the MTF and CTF is compressed using a square-root function. Barten has demonstrated that the SQRI metric has good predictive capability under a variety of viewing conditions (Barten, 1987, 1989a, 1989b).

Empirical models. An important class of models which deserves attention is an approach used by Snyder and his students (Snyder and Maddox, 1978; Reger, Snyder, and Farley, 1989) for the prediction of task performance with dot-matrix displays. Snyder explains that since there is little theoretical basis for extending the MTF-based metrics to flat-panel devices, a "strictly empirical approach was taken to predict visual task performance from a pool of photometric and geometric predictor variables" (Snyder, 1988,

p. 48). In a series of studies reported by Snyder and Maddox (1978) a set of 20 predictors as well as performance on a speed-of-reading and a menu search task were measured for a number of simulated dot-matrix displays. Stepwise multiple-regression techniques were then used to find linear multiple-regression equations which would allow performance to be predicted from the photometric and geometric measurements. The equations derived for predicting reading and search performance in an early experiment are listed below (Snyder and Maddox, 1978, p. 149):

$$\text{Reading time (s)} = 1.43 + 0.023\ \text{VSQR} + 0.364\ \text{HMTFA} + 0.221\ \text{VMTFA} - 4.825\ \text{HMLOG, and} \quad (4)$$

$$\text{Search time (s)} = 0.78 + 0.024\ \text{VSQR} + 2.72\ \text{HLOG} + 0.193\ \text{VMTFA,} \quad (5)$$

where VSQR is the square of the (vertical fundamental spatial frequency minus 14.0), HMTFA and VMTFA are the horizontal and vertical pseudo-modulation transfer function areas, HMLOG is the base 10 log of the HMTFA, and HLOG is the base 10 log of (horizontal fundamental spatial frequency divided by modulation).

Snyder and Maddox obtained correlations of 0.76 for the reading task and 0.69 for the search task. Thus, approximately 50% of the variance in performance could be explained by the predictors. In a subsequent validation experiment the authors felt that the model provided satisfactory predictions of performance for two out of the three displays examined.

The empirical modeling approach deserves attention primarily because it provided the first metric for evaluating dot-matrix displays. The work of Snyder and his students (Abramson and Snyder, 1984; Decker, Dye, Kurokawa, and Lloyd, 1988; Decker,

Pigion, and Snyder, 1987; Lloyd, Decker, Kurokawa, and Snyder, 1988; Reger, Snyder, and Farley, 1989; Snyder and Maddox, 1978) constitute the most comprehensive examination of image quality issues for matrix-addressable devices published to date.

Though the empirical approach has shown merit, a number of important limitations must be pointed out. The most serious difficulty with the approach is that it produces models which are highly task, technology, and situation specific. Support for this concern is provided by the Snyder and Maddox experiments. In a follow-on experiment, Snyder and Maddox developed an extended prediction model using the same set of 20 predictor variables and a larger set of performance data. This larger data set included the data used to derive equations 4 and 5 as well as data collected using displays with higher "fundamental" spatial frequencies. The resulting prediction equations are listed below (Snyder and Maddox, 1978, p. 246):

$$\text{Adjusted reading time (s)} = 5.74 + 0.311\ \text{HFREQ} + 2.479\ \text{HMOD} + 4.365\ \text{HLOG} - 14.973\ \text{HFLOG} + 1.112\ \text{VMLOG, and} \quad (6)$$

$$\text{Search time (s)} = 7.27 + 0.027\ \text{HDIV} + 2.159\ \text{HLOG} + 5.916\ \text{VFLOG} - 0.339\ \text{VMTFA} - 0.054\ \text{VRANG} + 5.487\ \text{VMLOG,} \quad (7)$$

where HFREQ is the horizontal fundamental spatial frequency, HMOD is the modulation of the horizontal fundamental spatial frequency, HLOG is the base 10 log of (horizontal fundamental spatial frequency divided by modulation), HFLOG and VFLOG are the base 10 logs of the horizontal and vertical fundamental frequencies, VMLOG is the log base 10 of the VMTFA, HDIV is the horizontal fundamental frequency divided by modulation, VMTFA is the vertical pseudo-modulation transfer function area, and VRANG is vertical crossover frequency minus the fundamental spatial frequency. Correlations of 0.72 for the

reading task and 0.71 for the search task were obtained. Unfortunately, no external validation of this extended model was conducted.

Comparison of equations 4 and 6 reveals that they have no terms in common. Equations 5 and 7 share only two common terms. Thus there are obvious qualitative differences between the models generated from experiment to experiment. There are similar qualitative differences between equations 4 and 5 and between equations 6 and 7. In addition, variables which have positive coefficients in one equation can have negative coefficients in another, indicating the presence of "suppressor" variables. Increasing values of the pseudo-MTFA terms correspond with decreases in performance in three out of five cases, a result which is counter to expectations based previous work with the MTFA metric.

It is obvious that the form of the prediction equations is sensitive to a number of factors. If these equations were generalizable to a wide variety of situations, the form of these models should be stable. Snyder and Maddox address the issue of limited generality of the models and clearly state many of the constraints for their use. The models are applicable only to dot-matrix alpha-numeric characters and only for speed-of-reading or menu search tasks. The models would not apply to displays with visible raster structure or those for which the "fundamental" frequency was above threshold. Since the models do not deal with the issue of dynamic range, they are not applicable for the evaluation of image processing techniques nor the presentation of literal imagery.

Metrics for Color Display Systems

In sharp contrast to the great number of metrics available for the evaluation of

continuous monochromatic displays and imagery, there appears to exist no valid metric for full-color display systems. Listed below are a number of color display design issues which have received attention from researchers and display developers over the past few years:

1. Color element design for flat-panel devices (Kobayashi, 1983; Silverstein, Monty, Gomer, and Yeh, 1989);
2. Convergence for shadow-mask color CRTs (Barten, 1988; Pica and Klopfenstein, 1987; Silverstein and Lepkowski, 1986);
3. Color spot profiling methods for CRTs (Baur, 1985; Bortfeld and Beltz, 1987);
4. Impact of shadow-mask on effective resolution (Infante, 86; O'Callaghan and Veron, 1987);
5. Spatial color synthesis (Silverstein and Lepkowski, 1986);
6. Noise in color signals (Faugeras, 1976);
7. Bandwidth compression of color signals (Faugeras, 1976); and
8. Color contrast metrics (Post, Costanza, and Lippert, 1982; Lippert, 1987).

None of these researchers have attempted to build and validate a model which simultaneously accounts for many of the pertinent color display system variables which impact image quality, including luminance and chromatic modulation transfer, luminance and chromatic noise, effects of sampling, and the capabilities of the human observer.

Researchers have attempted to extend monochrome metrics to certain color display design issues. For example, Granger (1974a, 1974b, 1983) has shown a good correlation between the predictions of a modified SQF metric and quality word names. While the author suggested that the SQF metric applies to color images, close examination reveals that although the method uses a modified spectral luminosity function, the metric does not

include a term responsive to chromatic modulation transfer or chromatic noise. The author makes the critical assumption that "for a color image which is in at least fair color balance, only the luminance component contributes to the perception of image quality" (Granger, 1983, p. 2). While this metric may be useful for evaluating the quality of color photographic systems, it cannot address the chromatic resolution and noise issues described below. The basic assumption made by Granger must be questioned. Other researchers have suggested that the addition of color modulation to images "makes a dramatic difference in the quality of the perceived image" (Glenn and Glenn, 1987, p. 24). Likewise, Schreiber (1984) suggests that a color television picture is "far superior in perceived quality" than is a monochrome picture requiring the same overall bandwidth (Schreiber, 1984, p. 46).

Pica and Klopfenstein (1987) have attempted to use the JND-analysis approach for the evaluation of color displays. But as with the work of Granger, the evaluation only considers luminance and sampling factors and does not address the important issues of chromatic modulation transfer capability and chromatic noise.

All of the experiments discussed in this section deal with color display design issues although none develop a comprehensive image quality metric. Notice that most of these design issues involve the interaction of spatial and chromatic aspects of human vision. Before these issues can be fully addressed from an image quality assessment point of view, a model of human visual capability which simultaneously addresses spatial and chromatic factors must be developed.

Summary of Deficiencies With Conventional Metrics

While the above mentioned metrics have been shown to perform well for monochromatic displays, they have several fundamental deficiencies when it comes to the assessment of color displays. The following paragraphs discusses these difficulties.

Chromatic resolution and noise. None of the conventional metrics deals directly with the issue of color resolution: the ability of the display system to transmit chromatic modulation. Likewise, none of these metrics can account for the effects of chromatic noise. Both of these factors are expected to impact color image quality.

Sampling. All of the above metrics rely on the MTF of a display as the primary quantitative representation of imaging capacity. Little discussion of the procedure for determining the MTF is provided by the metric developers. The most common way of determining the MTF of a display involves the measurement of a line spread function (LSF) for an impulse signal (usually a single line) presented on the display and subsequent Fourier analysis. This method has been promoted by photometric measurement system manufacturers and through standards organizations such as ANSI (Human Factors Society, 1988) and EIA (Keller and Beaton, 1989).

A serious difficulty with any metric based solely on a single MTF is that it does not utilize spatial sampling or grey-scale quantization information. The difficulties associated with this limitation are clearly illustrated with an example. Consider two display systems each using a monochromatic CRT with a 0.3 mm spot. Display A is set up for 1 bit/pixel and 2 addressable pixels/in, while display B is set up for 12 bits/pixel and 200 addressable pixels/in. The difficulty with the current MTF-based metrics (e.g., the MTFA, SN, JND-

analysis, and SQRI metrics) is that they would predict that these two displays produced images of the same quality.

A method for incorporating spatial sampling information into an MTF-based metric has recently been used by Beaton (1988). In this method the coefficients of the discrete Fourier transform are constrained by the display sampling rate and the total number of pixels along each display dimension. Beaton showed a good correlation between the response of this metric and subjective evaluations of image quality using a simulated flat-panel display. Barten (1987) took a similar approach by constraining the upper limit of integration of the SQRI metric using the sampling limit of the display or the highest frequency to be displayed. These methods of accounting for the sampling rate of the display are straightforward uses of linear systems analysis techniques. Hopefully the techniques will be incorporated into all MTF-based metrics.

Structured noise. The most common form of display noise that has been addressed with monochromatic image quality metrics has been wide-band random noise. One metric that explicitly accounts for noise of this type is the SN metric. Other metrics are capable of accounting for this class of noise indirectly through adjustment of perceptual weighting factors such as the CTF. With the JND-analysis metric the effects of random noise can be accounted for through the selection of the appropriate DDD (if one exists).

As pointed out in the introduction, one of the most important determinants of image quality for full-color flat-panel displays is structured noise. "Structured noise" is defined here as any form of tessellations, lines, dots, or patterning (chromatic or achromatic) that appear on a display and which have no counterpart in the original image at the input to the system. Structured noise arises from unwanted patterning introduced by the tessellation of

pixels on the display as well as from the generation of aliases caused by the inappropriate sampling of images.

Structured noise has a significant negative impact on human visual performance, as demonstrated in numerous studies reviewed by Snyder (1980, 1988). Snyder et al. (1974) and Beaton (1988) demonstrated how the use of a modulation spectrum can be used to predict the visibility of structured noise. But apparently a metric which accounts for both the effects of structured noise as well as the modulation transfer characteristics of a sampled display device has not been developed.

The Need for a New Approach

The essential attributes of a color image quality metric seem apparent. The metric should account for the following factors:

1. Transfer of modulation: luminance and chrominance;
2. Random and structured noise: luminance and chrominance;
3. Human spatial-chromatic visual capability.

While each of these factors has been related to image quality in previous studies, they have not been combined into a single model nor cast in terms of an image quality metric. While fairly complete models of spatial-chromatic processing within the human visual system have been described, a method for using these models for the evaluation of display systems has yet to be developed.

Approach Suggested by Research in Visual Science and Image Processing

Channel models. An important theoretical development that has guided much of the research in the field of visual science for the past 25 years has been the idea that the human visual system decomposes the visual stimulus into a set of responses of a number of spatial channels, each sensitive to a narrow band of spatial frequencies and selective for a certain orientation.

Campbell and Robson (1968) used the linear systems analysis approach to show that the contrast threshold function (CTF) like that produced by Schade (1958) could be used to predict contrast thresholds for a variety of grating patterns. The authors measured CTFs for sine, square, rectangular, and saw-tooth gratings. They found that gratings of a complex waveform could not be distinguished from sine waves until the contrast had been raised enough so that the higher frequency components of the waveform reached their independent thresholds. For the higher frequency gratings, threshold detection depended only on the fundamental frequency present in the waveform. For frequencies below about 6 c/deg the contribution of the higher harmonics to detection performance was lower than would be predicted by a spatial filter followed by a simple peak detector. The authors hypothesized that the system must contain "a number of independent detector channels each preceded by a relatively narrow-band filter tuned to a different frequency. Each filter and detector would constitute a separate 'channel' and each channel would have its own contrast sensitivity function" (Campbell and Robson, 1968, p. 564).

A year later, Blakemore and Campbell (1969) found that with the use of adaptation gratings these channels could be desensitized temporarily. Using this experimental technique they estimated that the half-power bandwidth for spatial frequency channels

within the human visual system was just over one octave. By varying the orientation of the gratings, they determined that the channels were selectively sensitive to orientation as well as spatial frequency.

Sachs, Nachmias, and Robson (1971) conducted a study in which they measured threshold contrast for simple and complex grating patterns. The complex patterns were made up of two sine-wave components for which frequency and amplitude could be independently controlled. The hypothesis that the visual system behaves as a single broadly-tuned filter was tested against the hypothesis that multiple, narrowly-tuned channels are employed. The results of several experiments are compatible with the assumption that the human visual system contains a number of channels which are sensitive to narrow ranges of spatial frequencies. The output of each channel seems to be detected independently by higher visual centers.

Using a number of experimental conditions, essentially the same conclusions were drawn by numerous other researchers during the same time period (e.g., Graham and Nachmias, 1971; Pantle and Sekuler, 1968).

Multi-channel models. In recent years a number of researchers have tried to account for overall human visual capability using the channel approach. Models have been constructed in which multiple independent narrow-band channels work in parallel on the visual stimulus (Quick, 1974; Watson and Robson, 1981; Wilson and Bergen, 1979; Wilson and Gelb, 1984). Multi-channel theorists generally suggest that detection will occur when the response of one or more of the channels exceeds its individual threshold. The behavior of the ensemble of these channels has been shown to accurately predict the CTF of human vision.

Wilson and Giese (1977) demonstrated that human threshold visibility for frequency gradient patterns could be modeled by line spread functions of the difference-of-Gaussians (DOG) type. Based on these results, Wilson and Bergen (1979) developed a model for human threshold vision based on a set of four DOG mechanisms. The model assumes that four concentric DOG mechanisms service each position on the retina. Each of the four mechanisms is tuned to a different spatial frequency and each has a bandwidth of slightly over one octave. The gain and center frequencies of the mechanisms vary linearly as a function of retinal eccentricity. The responses of individual mechanisms were combined using Quick's (1974) approximation of spatial probability summation.

Wilson and Bergen (1979) measured the CTF for the composite model and compared this function with functions measured using a limited number of human subjects. Parameters of the model were then adjusted such that good agreement between model responses and the responses of human subjects was obtained. Responses to the Wilson and Bergen model were obtained using a computer simulation of the channels in one spatial dimension. Apparently the complexity of the multi-channel model was such that an analytical solution was impractical.

The most important conclusion of the Wilson and Bergen effort is that for the range of spatial frequencies considered, as few as four channels were sufficient to account for the human threshold data. The authors point out that all of the aspects of the model had been incorporated into models in previous studies. The unique aspect of their model is the simultaneous incorporation of multiple concepts into a single model. The discovery that four channels are sufficient followed from exercising a composite model and comparing the results with human performance data.

Anisotropy. Numerous studies have demonstrated that the sensitivity of the human visual system to grating patterns is decidedly orientation dependent (Appelle, 1972; Campbell, Kulikowski, and Levinson, 1966). It has been suggested (Wendland, 1983) that due to the anisotropic nature of the human visual system the performance of CRTs might be improved if the raster were oriented on the diagonal (as is done in half-tone printing) (Glenn, Glenn, and Bastian, 1985, p. 72; Snyder, 1985). A comprehensive image quality metric should reflect this orientation-dependent sensitivity.

Channel independence. Although researchers still argue over the exact orientation and frequency bandwidths of channels within the human visual system, the functional existence of independent channels is broadly accepted. The existence of independent, limited-bandwidth mechanisms should be reflected in an image quality metric since a single-channel model cannot properly account for selective effects of adaptation and masking.

In the evaluation of full-color flat-panel displays, selective masking becomes important due to the generation of large amounts of structured noise. For visual noise to have a masking effect on some particular channel, that channel must be sensitive to the noise. That is, the noise must have some power within the range of spatial frequency, orientation, and spatial location of that channel. Selective masking becomes important for the prediction of performance on certain tasks. For example, visible horizontal raster lines at 10 c/deg should have little effect on threshold detection of vertical sine-wave gratings at 5 c/deg, whereas the raster lines would increase the threshold for horizontal gratings at 8 c/deg. Since conventional metrics model the human visual system (and the display for that matter) as an isotropic, single-channel system, they cannot reflect these separable effects of masking. For this reason the present model uses separate independent channels.

Spatial-chromatic models. The tools of linear systems analysis have been quite useful for the evaluation of factors affecting the luminance resolution of displays. It is suspected that this same class of tools should also be useful for examining factors affecting chromatic resolution. While human responses to luminance modulation have been studied much more extensively, good sensitivity data for chromatically modulated stimuli are available.

Schade (1958) was the first researcher to measure human spatial-chromatic sensitivity using the tools of linear systems analysis. Schade found that the chromatic modulation response had a distinct band-pass nature much like that of the luminance response. Since 1958, a number of other researchers have measured this same function using both square- and sine-wave gratings (Granger and Heurtley, 1973; van der Horst, 1969; van der Horst and Bouman, 1969; van der Horst, de Weert, and Bouman, 1967). The chromatic modulation transfer function was found to be low-pass in all of these studies, prompting Granger and Heurtley (1973) to criticize the original work of Schade (1958) who had found a distinct resonance.

Unfortunately, the controversy over the shape of the chromatic functions did not end with the work of Granger and Heurtley. Several more recent studies have shown distinct band-pass characteristics. Using stimuli and apparatus more representative of display viewing conditions (i.e., larger field of view, higher luminance, chromaticities centered at white, etc.), Glenn et al. (1985) obtained band-pass isoluminance chromaticity CTFs which had peak sensitivities of 2.5 c/deg for the red-green channel and 1.3 c/deg for the blue-yellow channel. Likewise, Faugeras (1976) found very similar band-pass functions with the same ratios among the peak frequencies in each channel. Finally, Kelly (1989) showed a red-green chromatic CTF for normal unstabilized viewing that has a definite band-pass characteristic with a peak sensitivity at 0.5 c/deg

The recent study of human chromatic contrast sensitivity conducted by Mullen (1985) stands out from other studies of this phenomena primarily because of the several important precautions the author reports taking to insure that the chromatic channels were stimulated without simultaneous stimulation of the achromatic channel. In addition to using near-monochromatic stimuli with wavelength selection based on the human opponent-color spectral sensitivity functions, Mullen optically corrected for the chromatic aberration of the eyes of each subject. Most importantly, Mullen used an experimental technique to "null" the effects of the achromatic channel for each subject and *for each spatial frequency used.* For high spatial frequencies, where the achromatic channels are more sensitive, the achromatic response was nulled by repeatedly measuring the detection threshold using a range of red/green (or yellow/blue) luminance ratios for the chromatic gratings. The chromatic detection threshold occurred at that luminance ratio where contrast sensitivity was minimized. At low spatial frequencies where the sensitivity of the chromatic channels is greater, the same experimental procedure was used except that the maximum rather than the minimum threshold was taken as the chromatic threshold. This method of separately nulling the achromatic response for each spatial frequency distinguishes the Mullen work. Thus it seems likely that the results obtained by Mullen are more reliable than the results obtained in other studies, particularly those studies for which the authors publish few details of the experimental apparatus, method of calibration, or procedures employed (e.g., Faugeras (1976), Glenn et al. (1985), Kelly (1989), and Schade (1958)).

Results of the Mullen (1985) experiment indicate no drop in chromatic sensitivity as spatial frequency decreases down to 0.1 c/deg. These results are consistent with the results of Granger and Heurtley (1973), van der Horst (1969), van der Horst and Bouman (1969), and van der Horst, de Weert, and Bouman (1967). The Mullen data show a high

frequency cutoff (1% of peak sensitivity) of 11 to 12 c/deg for both the red-green and the yellow-blue channels. This value is somewhat lower than previous estimates and is likely due to the improvement in experimental technique.

The importance of separately nulling the achromatic response for each spatial frequency has been discussed also by Flitcroft (1989), who performed a thorough analysis of the impact of chromatic aberration on chromatic responses. The results of a simulation of the effects of chromatic aberration, defocus, and diffraction predicted that the relative responses of the long, medium, and short (L, M, and S) wavelength cone systems will be noticeably altered for spatial frequencies above 1 to 2 c/deg and severely disrupted at spatial frequencies above 4 to 5 c/deg. These findings support Mullen's suggestion that nulling the achromatic response at a single spatial frequency is insufficient. Researchers attempting to measure chromatic contrast sensitivity should null the achromatic response at each spatial frequency used.

Obviously there exists an unresolved controversy as to the nature of the chromatic transfer function at low spatial frequencies. However, this controversy should have little impact on this dissertation work since most of the display design issues addressed here involve only the middle and upper spatial frequency range. The above mentioned research efforts largely agree on the nature of the middle and upper-frequency portions of these functions. All of the studies show that the chromatic channels cut off at lower spatial frequencies than the luminance channel. The studies of Faugeras and Glenn et al. show that the red-green channel cuts off (- 3dB) approximately one octave below the luminance channel while the yellow-blue channel cuts off approximately one octave below the red-green channel. Similarly, the Mullen work indicates that the red-green and yellow-blue

channel cut offs (-3 dB) are 1.5 and 1.8 octaves below the cutoff of the achromatic channel.

This discussion of spatial vision and spatial-chromatic interaction applies only to static images viewed in a normal unstabilized manner. It must be pointed out that the nature of the CTF and spatial-chromatic interaction changes as a function of temporal factors. The findings of Kelly (1966, 1979, 1989), Koenderink and van Doorn (1979), Robson (1966), and Williams and Erickson (1974) have shown that temporal frequency has a strong effect on the shape of the CTF for both achromatic and chromatically-modulated gratings.

Image processing models. With the rapid development of image processing hardware and techniques, engineering models of the human visual system have shifted from simple analytical expressions toward treating the visual system as an image processor. A major advantage of this approach is that it provides the capability of exercising models which are difficult or impossible to deal with analytically. Additionally the approach largely frees the vision modeler from the constraints of linear systems analysis. For example, models which are multi-staged, non-linear, heterogeneous, and anisotropic can be used with no particular difficulty.

While image quality researchers in the fields of photography and electronic displays have used linear models almost exclusively, researchers in the field of image processing have been using nonlinear models of the human visual system for the assessment of image quality for almost two decades (e.g., Faugeras, 1976; Hall and Hall, 1977; Mannos and Sakrison, 1974; Saghri, Cheatham, and Habibi, 1989; Stockham, 1972; Xie and Stockham, 1989). The goals of image quality researchers in the field of image processing

have been largely the same as those of display researchers, that is, to derive objective measures of image quality. Unfortunately there has been little cross-talk between the disciplines.

The models generated by researchers in the field of image processing typically are more sophisticated in that they attempt to account for more of the physiological and psychophysical data than do the linear models used for the evaluation of display systems. For example, an early model developed by Stockham (1972) incorporated a logarithmic sensitivity prior to a linear spatial filtering stage. Justification for the nonlinear first stage was based on the supposed logarithmic nature of photoreceptor responses as well as supra-threshold discrimination data. Additionally, Stockham points out that the light distribution reaching the eye is the result of multiplicative superposition of object reflectance and scene illumination. Stockham argues that the use of a density rather than an intensity representation allows surface reflectance to be separated from scene illumination using linear spatial filtering. Thus, the use of a density representation followed by high-pass filtering allows the model to reflect brightness constancy effects in human vision.

Mannos and Sakrison (1974) discused the need for a distortion measure that correlates well with subjective image quality. In order to use Shannon's rate-distortion theory to determine an optimal image encoding scheme, a distortion measure which accounted for the difference between an original and a coded image had to be derived. Following the work of Stockham (1972) (and apparently unaware of the work with display image quality metrics), Mannos and Sakrison (1974) developed a human vision model using a nonlinear first stage followed by a linear spatial frequency filtering stage. This model was used as a pre- and post-processor for encoding monochromatic images at low information rates. Using this method, groups of images were generated using different settings of the

parameters of the human vision model for each image. The authors experimented with linear, logarithmic, and two power functions for the first stage. They also varied the shape of the linear spatial filter by manipulating three parameters that effectively controlled the peak frequency and the high and low frequency slopes.

For each combination of model parameters Mannos and Sakrison used the model to pre- and post-process images which were then rated by a group of subjects. The results of the quality ratings were used to "tune" the parameters of the human vision model so as to improve the model performance. Results of this model tuning effort produced a model with a power function of the form $F(u) = u^{0.33}$ first stage followed by a band-pass spatial frequency filtering stage. The form of the spatial pass band resembles the CTF for human vision with a peak frequency of 8.0 c/deg. The results of this study support the use of a nonlinear first stage of processing within an image quality metric.

In an attempt to explain the data of Davidson (1968) and Henning, Hertz, and Broadbent (1975), Hall and Hall (1977) embellished the nonlinear-linear models of Stockham (1972) and Mannos and Sakrison (1974). Citing optical factors as the cause, Hall and Hall placed a low-pass filter in front of the basic nonlinear-linear model. Based on the Fourier transform of the optical line spread function derived by Westheimer and Campbell (1962), this low-pass filter had a -3 dB frequency of 6.6 c/deg. Using this model Hall and Hall were able to predict the dependence of the describing functions on contrast at high spatial frequencies found by Davidson (1968).

Faugeras (1976, 1979) carried the nonlinear modeling approach of Stockham (1972) into the realm of color. Faugeras developed a four-stage model in which the first stage contained three spectral absorption processes each followed by a logarithmic stage.

Following the nonlinear stages were three opponent-color processes which derived luminance, red-green, and blue-yellow images. These three images were then processed by linear spatial filters. The Faugeras model can be viewed as a zone model of color vision which has been extended into the spatial domain.

The spatial-chromatic model of Faugeras was used to assess the effects of bandwidth compression and noise within the color signals of an image transmission system. Using a limited amount of data, Faugeras showed a good correspondence between image quality predictions made by the model and subjective evaluations of degraded imagery.

From an engineering point of view the Faugeras model is appealing since the same class of linear spatial filter used in monochromatic models is used for chromatic channels. The use of opponent-color signals is well founded in human physiology and psychophysics and is similar to the method of color encoding used for the transmission of color television images. Finally, the chromatic response of the model is directly comparable to the results of the isoluminance chromatic modulation threshold experiments discussed above.

Bumbaca and Smith (1987) have examined the utility of the Faugeras (1976) model for machine vision systems. The authors changed the model slightly by using the spectral response functions of Smith and Pokorny rather than the functions proposed by Stiles. A second change restricted the contribution of blue cones from the computation of luminance. Bumbaca and Smith then re-tuned the parameters of the model using the same procedure as Faugeras (1976). Unfortunately, it is impossible to evaluate whether these changes to the model resulted in any improvement in the model for evaluating images since no performance or perceptual response data are reported.

Additional support for the use of non-linear zone models of the human visual system comes from outside the field of image processing. The combination of a nonlinear cone response stage followed by an opponent color stage has been used by Valberg, Lee, and Tryti (1987), who simulated the responses of spectrally opponent cells in the lateral geniculate nucleus (LGN) of the monkey. Comparisons of the simulation results with actual data showed that the responses of 85% of the opponent cells that were measured could be satisfactorily simulated using the difference between hyperbolic functions of the receptor responses. The best fitting cone response functions were of the form:

$$Q = \frac{Q_{max} \times S_n}{S_n + s_n} \tag{8}$$

where Q_{max} is the maximum response of the cone, S_n is the receptor response due to spectral absorption, and s_n is the half-maximum response. Though the spectrally opponent cells found in the LGN of the monkey do not exactly correspond with the opponent color signals derived in the proposed model, Valberg et al., as well as other researchers (e.g., Marr, 1982, p. 263) suggest that the simple combination of the outputs of spectrally opponent cells with opposite cone inputs would result in the desired signal. Numerous researchers have identified (double-opponent) cells at higher levels of the primate visual system and have shown that these cells respond to spatial-chromatic differences (Livingstone and Hubel, 1984; Michael , 1983). These double-opponent cells would be responsive to chromatic edges and gratings and would be nonresponsive to luminance modulation.

Two primary lessons are to be learned from the results of image quality modeling efforts in the field of image processing. First, the value of incorporating nonlinear

processing seems clear. Second, the utility of incorporating an opponent-color stage of processing has been demonstrated for the examination of color image quality issues.

Though certainly more sophisticated, the nonlinear models used in the image processing field have not been as thoroughly evaluated as some of the linear metrics used for the evaluation of display systems. As discussed above, the MTFA metric has produced high correlations with both subjective image quality and task performance in numerous studies. Additionally, it has fared well in several side-by-side comparisons among competing metrics. In contrast, while the nonlinear models are sometimes evaluated using subjective quality data, studies that correlate model performance with human task performance are rare. Likewise, studies which make side-by-side comparisons among a number of competing models seem nonexistent. Much research work remains to be done if this class of model is to attain the status of an acceptance standard as has the MTFA metric.

DESIGN OF THE HUMAN VISION MODEL

Overview

This section of the dissertation describes a model of the human visual system (HVS) designed specifically to evaluate color electronic display systems. The following section describes how this HVS model can be used to evaluate a display. Together the model and procedure provide an objective method for measuring the ability of a display system to transmit chromatic and achromatic modulation as well as the amount of chromatic and achromatic noise generated by the system. Since the HVS model reflects threshold and suprathreshold human visual sensitivities, the model serves to weight these display measurements in terms of human visual performance.

Unlike the image quality metrics discussed in the introduction, the proposed methodology is not based on the tools of linear systems analysis. Rather, a more literal interpretation of the current thinking in visual science and retinal physiology is used. The HVS is modeled here as a sequence of nine image processing operations which successively abstract visual stimuli into the responses of a set of spatial-chromatic channels. The last sections of this dissertation demonstrate the correspondence between the responses of these channels and ratings of image quality.

The HVS model proposed here rests largely on the assumptions of zone theory commonly attributed to Muller and Judd (Wyszecki and Stiles, 1982). Like models adhering to the Young-Helmholtz trichromatic theory, the proposed model contains three types of cones exhibiting different spectral sensitivities. The cone absorption stage is followed by a nonlinear transduction stage which allows the model to reflect nonlinear

suprathreshold psychophysical behavior. The nonlinear responses from the transduction stage feed into an opponent-color stage in which achromatic, red-green, and yellow-blue difference images are derived. The basic zone model is extended into the spatial domain through the use of linear spatial filters. Since the model uses a number of discrete spatial mechanisms, each tuned to a separate "narrow" band of spatial frequencies, the model follows the multi-channel rather than single-channel theory of spatial vision. Spatial processing is done entirely in two dimensions to allow the model to use oriented spatial mechanisms. The scalar responses of the set of spatial mechanisms make up a response vector which can be used in the prediction of human performance.

Inputs to the HVS model consist of images of grating patterns as they would appear on the display system under evaluation. The model can provide two types of outputs. First, a unitary metric describing the overall capability of the display can be obtained. Such a metric is useful for making general comparisons among displays or correlating with measures of human performance. Second, the model can provide a response vector describing a display's noise and attenuation characteristics in each spatial frequency and chromatic band. This channel-specific information is useful for identifying specific attributes of a display design which impact performance.

Evaluation System: General Design Constraints

Due to the constraints of the computational facilities available for this project three primary constraints have influenced the design and validation of the model. First, the model applies only to static images. Thus, display quality issues involving motion and flicker cannot be examined with this model. Second, the maximum image size that the model can process is limited by memory and processing time constraints. Thus, the model

is not capable of evaluating those attributes of a display system which impact spatial frequencies below approximately 2.0 c/deg. Finally, this first version of the model is intended for examining display systems with a luminance level of approximately 50 cd/m^2. Future versions of the model should allow luminance level to be specified.

The primary intent of this work was to provide a tool that can be used for the evaluation of color display systems. In order that this tool be designed appropriately, it is necessary to examine the use of the tool from the display evaluator's point of view. The following paragraphs describe a number of more general operational requirements which have shaped the design of the proposed evaluation system.

Continuity. For the case of achromatic imagery the perceptually-weighted MTF-based metrics as well as the JND-analysis approach have performed well. For the sake of continuity with the existing knowledge base on image quality metrics, the present method should capitalize upon the best attributes of these previous methods. The evaluation system should apply to monochromatic as well as full-color displays. When used for achromatic imagery, the form of the proposed model should resemble that of the previous metrics.

Predictive capability. For monochromatic display systems the model should have an all-around predictive capability as good as or better than existing image quality metrics. Correlations obtained using existing perceptually-weighted, MTF-based metrics have generally been above 0.8 and are often above 0.9. Since this is one of the first attempts to predict image quality on color display systems, there is little basis for comparison. If correlations greater than 0.7 could be obtained for color imagery, the model would likely be useful in an engineering sense since it could explain more than 50% of the variance in subject responses.

Common language. An important attribute of the MTF-based metrics is that the MTF can be used to characterize the behavior of each of the system elements, including the human observer. An important advantage provided by the use of the MTF is that communications between systems designers and human factors engineers are significantly enhanced (Snyder, 1980). Using the MTF, the behavior of each component of the system can be characterized separately. Overall system performance can be determined by combining each of these separate descriptions. With the proposed evaluation system, the behavior of each component of the system (including the human observer) can be described in terms of the image processing operations which are performed. Each component of the system can be characterized independently of the other components of the system. As in the case of the MTF-based approach a common language can be used by display designers as well as visual system modelers.

Allowance for nonlinearity. There is ample evidence that the assumptions of linear systems analysis are violated at several stages of processing within the human visual system. Likewise, most display systems have components which depart significantly from these assumptions. Rather than limiting model development to the realm of linear systems, the more general approach will be taken here. The structure of the evaluation system should allow non-linear stages of processing (such as the intensity transfer function of a display, response of cones, or supra-threshold contrast discrimination functions) to be used for both the model of the proposed display device (the display simulator) and the model of the human visual system.

Display independence. The structure of the HVS model should not depend on the type of display under evaluation. This constraint would simplify use of the evaluation system since the same vision model could be used in evaluating a variety of display types.

Similarly, the evaluation system should be applicable to sampled (discrete element) systems as well as continuous systems.

Input specification. The requirements for use of the model by a display systems designer should be clearly defined. Display systems designers should know exactly which attributes of the display system should be measured (or specified) in order to use the model. In making this specification the designer should be allowed to work with design-relevant variables.

Speed and portability. To allow for portability to other locations and use by others, the model should run on a commonly available machine and not require the use of uncommon hardware or software. The model should be written in a common, high-level, structured language. The machine and software environment should allow for the use of a convenient facility for the display of color imagery.

Human Vision Model: Detailed Description

The proposed HVS model employs nine processing stages to account for nonlinear retinal processes, opponent color encoding, and multiple spatial frequency channels (see Figure 2). Each stage of processing is described below.

Input image. The input image constitutes the interface between the display simulator and the HVS model. This image precisely describes what appears at the display screen. Use of the input image, as opposed to an MTF, allows a display designer to accurately represent characteristics of display systems that do not meet the assumptions of linear systems analysis (e.g., linearity, shift invariance, and isotropy). For example, a display designer can simulate the effects of spatial and gray scale quantization, attenuation of

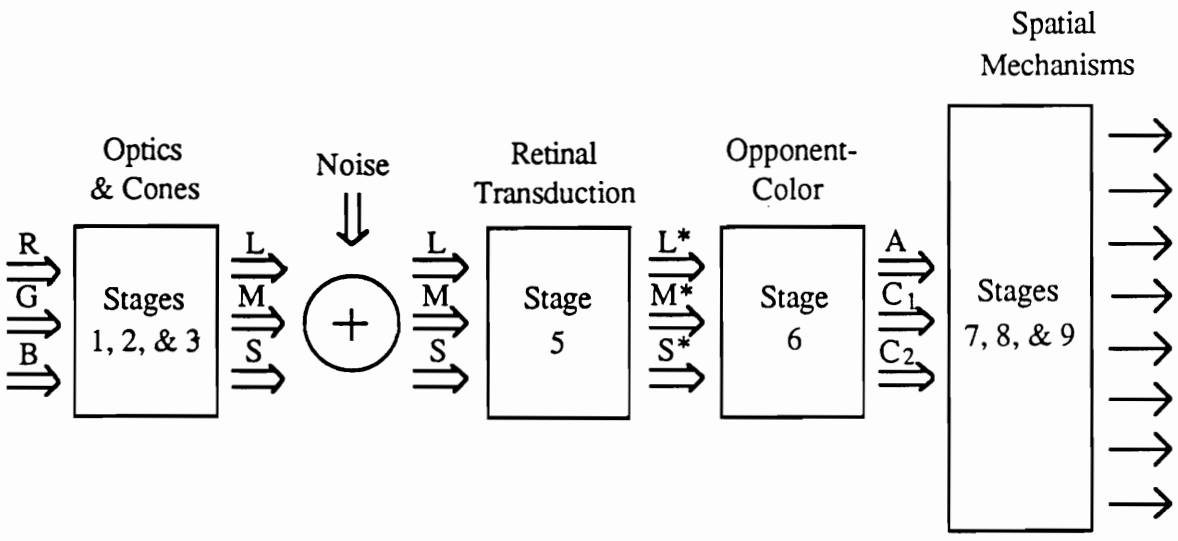


Figure 2. Block diagram of complete human vision model.

chromatic modulation, astigmatic pixel structure, and a nonlinear gamma function. This feature allows the proposed evaluation method to be applied to color as well as matrix-addressable displays.

The HVS model is designed to evaluate display systems that use three primary colors. The use of only three primaries allows input images to be conveniently described using a three-band image and a spectral power distribution for each band. For the current modeling work, each spectral image consisted of 128 x 128 pixels x 8 bits, and corresponded to the image produced by one of the display primaries. The sampling rate of the input image is 150 pixels/deg. This rate was set high enough that the smallest visible features of a display system could be represented without the risk of under sampling. Thus, the sampling frequency was chosen to be three times the high frequency cutoff (approximately 50 c/deg) for human vision.

Stage 1: Optical blur. The first stage of processing performed by the HVS model simulates biological blurring by the optics of the eye. Optical blur was computed by convolving each band of the input image with the peripheral (i.e., cornea, lens, and ocular media) point spread function (PSF) of the eye. The PSF was used since subsequent computational operations were performed in two dimensions. The PSF was computed using the inverse Able transform (Bracewell, 1986) of the LSF reported by Campbell and Gubisch (1966) for a 3.0 mm diameter pupil. Although the PSF was computed numerically, the resulting function is well fitted by Equation 9 ($R^2 > 0.9995$)

$$\text{Weight} = \text{Exp}(0.69593 - 178.88\, r + 2054.6\, r^2 - 10366\, r^3), \qquad (9)$$

in which r is the radius of the function in degrees of visual arc. Figure 3 shows the LSF

and derived PSF used in these computations.

In Figure 4, the PSF is plotted as a two-dimensional response surface at a sampling rate of 150 pixels/deg. The two-dimensional form of the PSF was generated by rotating the PSF described by Equation 9 about a vertical axis located at 0 deg. The input to the first stage of processing is a three-band (RGB) image of the surface of the display. The output is a blurred version of the same RGB image.

Stage 2: Linear spectral absorption. The second stage of processing in the HVS model simulates the absorption of light by each class of cone photoreceptors. Numerous researchers have proposed spectral response functions based on psychophysical experiments and direct measurements of photoreceptor pigments. The cone absorption functions selected were those derived by Smith and Pokorny (1972) (reported by Wyszecki and Stiles, 1982, Table 2(8.2.5)). The Smith and Pokorny functions are linear transformations of the 1951 Judd-corrected color matching functions. Research has shown that the Judd color matching functions more accurately account for the visual effects of short wavelength (blue) light than do the 1931 CIE functions. Thus, the Smith and Pokorny functions have been promoted for use in a system of photometry and colorimetry based on cone excitations (Boynton, 1983). Additionally, the Smith and Pokorny functions have been used previously in computer vision applications (Benzschawel and Guth, 1984; Bumbaca and Smith, 1987). Figure 5 plots the Smith and Pokorny functions. Calculation of cone responses is made using Equations 10 to 12.

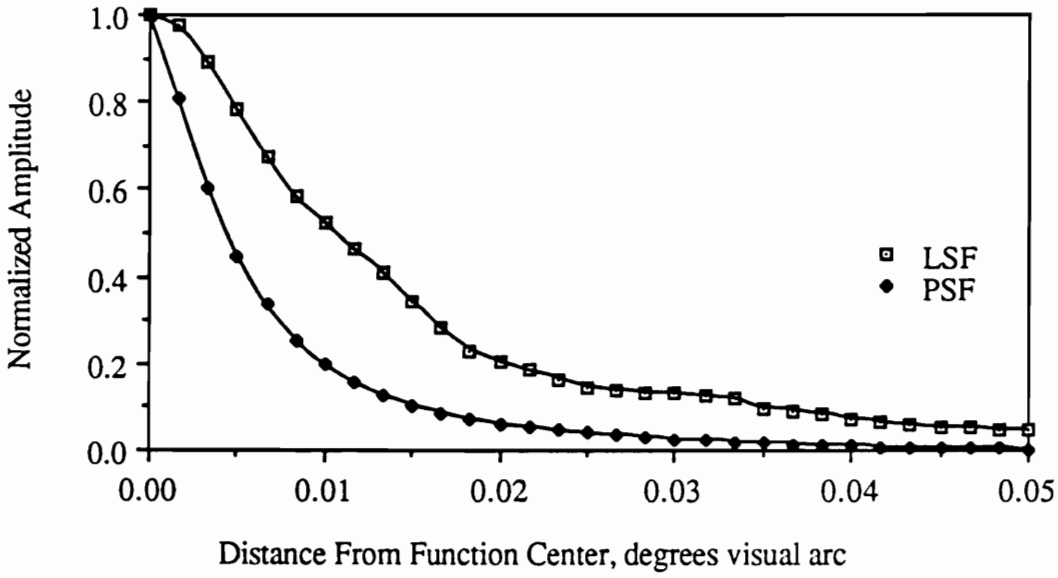


Figure 3. Point spread function used in simulating optical blur.

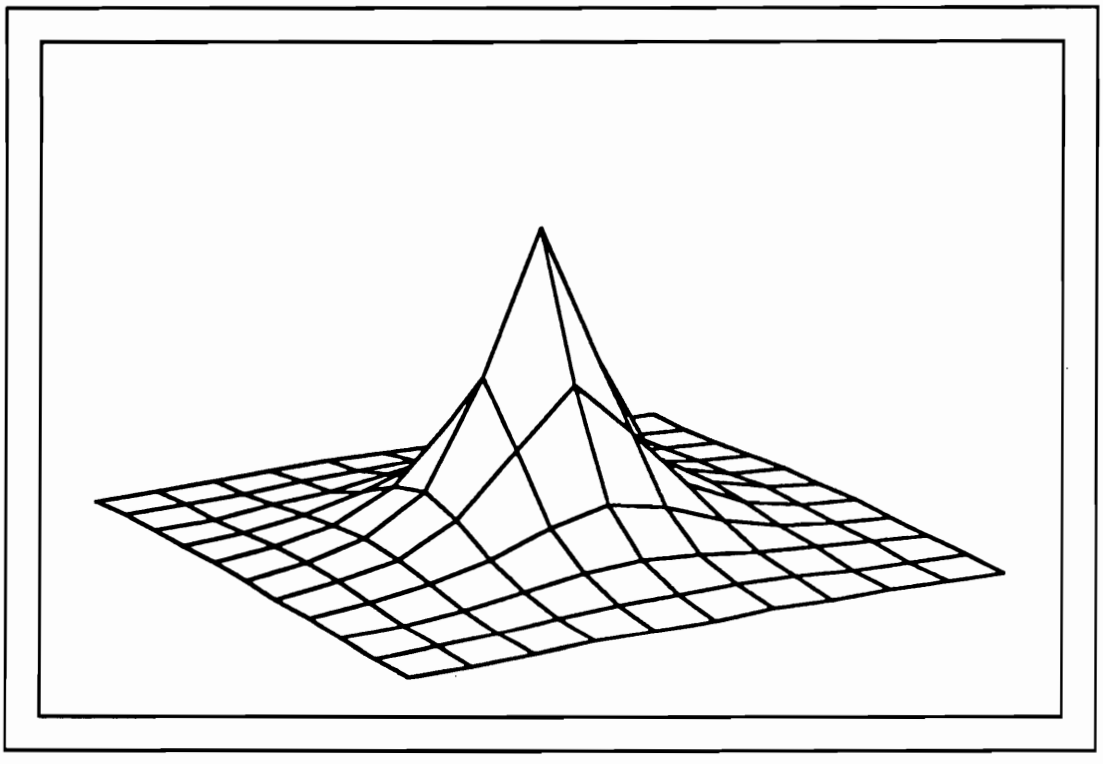


Figure 4. Oblique view of optical PSF at a sampling rate of 150 pixels/deg.

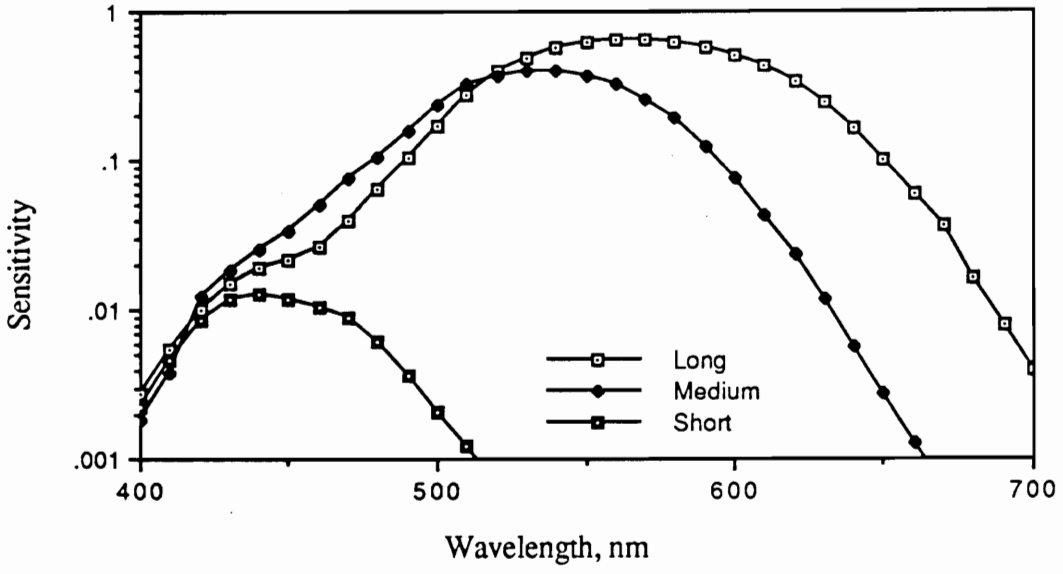


Figure 5. Spectral response functions for long, medium, and short wavelength cones. (Adapted from Smith and Pokorny, 1972)

$$L(x, y) = \int_{380}^{760} R(x, y, \lambda)\, l(\lambda)\, \partial\lambda + \int_{380}^{760} G(x, y, \lambda)\, l(\lambda)\, \partial\lambda + \int_{380}^{760} B(x, y, \lambda)\, l(\lambda)\, \partial\lambda \tag{10}$$

$$M(x, y) = \int_{380}^{760} R(x, y, \lambda)\, m(\lambda)\, \partial\lambda + \int_{380}^{760} G(x, y, \lambda)\, m(\lambda)\, \partial\lambda + \int_{380}^{760} B(x, y, \lambda)\, m(\lambda)\, \partial\lambda \tag{11}$$

$$S(x, y) = \int_{380}^{760} R(x, y, \lambda)\, s(\lambda)\, \partial\lambda + \int_{380}^{760} G(x, y, \lambda)\, s(\lambda)\, \partial\lambda + \int_{380}^{760} B(x, y, \lambda)\, s(\lambda)\, \partial\lambda \tag{12}$$

where $l(\lambda)$, $m(\lambda)$, and $s(\lambda)$ are the cone spectral response functions and $R(x, y, \lambda)$, $G(x, y, \lambda)$, and $B(x, y, \lambda)$ are the spectral power distributions for the red, green, and blue display primaries. The input to the spectral absorption stage of processing is the blurred RGB image. The output of this processing stage is a new three-band (LMS) image representing the responses of the long, medium, and short wavelength sensitive cones.

Stage 3: Response normalization. The third stage of processing simulates the local and global adaptation of cone photoreceptors to the overall level of light. Given that the first stage of neural processing within the HVS (the photoreceptor stage) has a smaller dynamic range than that of the visual stimulus, clipping necessarily occurs. If adaptation (gain control) used to control visual sensitivity were included after the photoreceptor stage it would be of limited utility due to the limited dynamic range of the receptors. On the other hand, including adaptation before response clipping would seem necessary in order to position the limited dynamic range of the receptors around the most useful portion of the dynamic range of the visual stimulus.

From a computational viewpoint the mechanics of visual light adaptation serve to center the nonlinear retinal transducer function (stage 5) at the mean intensity level of the input image. In the proposed model, cone responses are normalized so that their space-averaged value is centered on the retinal transducer function.

Stage 4: Internal noise. Internal noise is modeled as additive Gaussian noise on the outputs of each receptor mechanism. This single noise source determines the level of both contrast detection and contrast discrimination thresholds. Unlike previous models of visual threshold and suprathreshold contrast responses (Cannon and Fullenkamp, 1988; Carlson and Cohen, 1978; Legge and Foley, 1980; Nachmias and Sansbury, 1974; Pelli, 1985; Swanson, Wilson, and Giese, 1984; Wilson, 1980), noise is independent of signal strength and is added before the nonlinearity and spatial mechanisms.

A straight forward (although cumbersome) method of computing the model response to internal noise is simply to add noise to the cone mechanism responses for each input image processed by the model. However, since noise is a random variable, characterization of the stochastic response of the model for each input image would require excessive computational effort. Rather than perform repeated computations for each new input image, a deterministic response was assumed and computed in advance. The effects of internal noise added at this stage of the model are accounted for in the probability summation stage (Stage 9). Of course, the assumption is made that the response of the model to internal noise is independent of the content of input images.

Stage 5: Nonlinear retinal transduction. A unique attribute of the proposed model is the use of a single nonlinearity occurring prior to the spatial mechanisms to account for suprathreshold psychophysical performance. In contrast, existing models of contrast

discrimination and perception employ post-spatial mechanism nonlinearities to account for these data (Cannon and Fullenkamp, 1988; Legge and Foley, 1980; Nachmias and Sansbury, 1974; Swanson, Wilson, and Giese, 1984; Wilson, 1980). Since the nonlinearity in the model occurs prior to the spatial mechanisms, and significant spatial interactions occur at low levels within the vertebrate visual system, it is likely that the locus of the proposed nonlinearity is the retina. The proposed nonlinearity is thus called a retinal transducer function (RTF).

The proposed RTF was derived numerically as a piece-wise linear function using an iterative algorithm to determine the slope of each small segment of the function. Derivation of the RTF was constrained by two factors: the structure of the complete HVS model and a function describing the averaged data from six recent studies of contrast discrimination (Bradley and Ohzawa, 1986; Burton, 1981; Legge and Foley, 1980; Lloyd, Beaton, Parsley, and Simonetti, 1990; Nachmias and Sansbury, 1974; Pelli, 1985). To understand how the RTF was derived the reader must first have a clear idea of the structure of the HVS model. Thus, the details of the derivation of the RTF are presented in Appendix B. Briefly, the input range of the RTF was divided into 1024 intervals. The slope of the RTF in each of these intervals was then iteratively adjusted so that a just noticeable difference (JND) in the modulation of an input grating (determined from the function describing average contrast discrimination performance) produced a unit change in the output of the model. The resulting piece-wise linear RTF was transformed to log-log coordinates and fitted with a polynomial ($R^2 > 0.9995$). The resulting function is given as Equation 13. The RTF is plotted in linear coordinates in Figure 6.

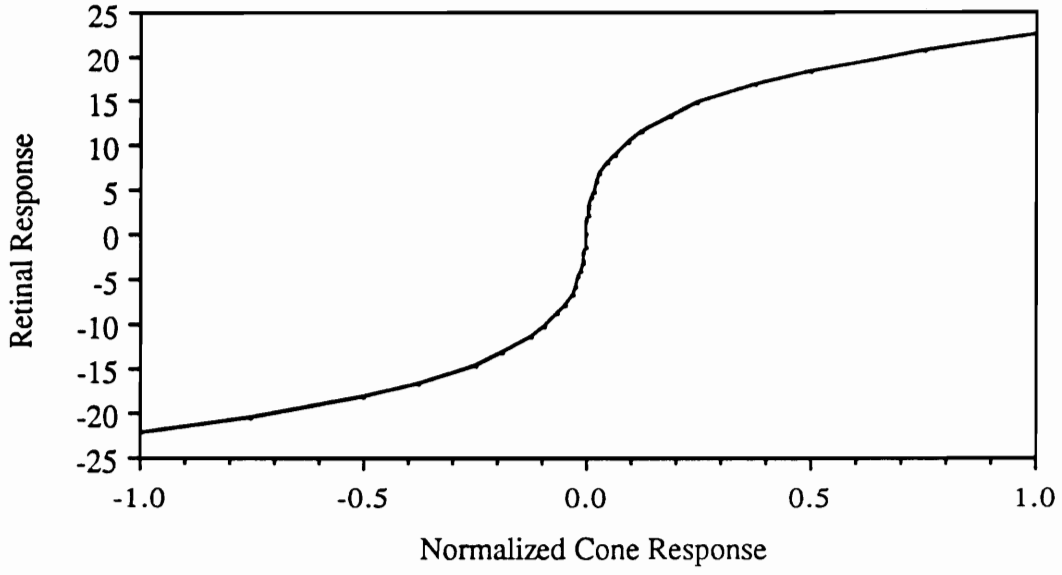


Figure 6. Nonlinear retinal transducer function.

for X >= 0,

$$\text{Response} = \text{Exp}(3.1272 + 0.34835\ \text{Ln}(X) + 0.020291\ \text{Ln}(X)^2 + 0.0061942\ \text{Ln}(X)^3)$$

and for X < 0,

$$\text{Response} = -\ \text{Exp}(3.1272 + 0.34835\ \text{Ln}(-X) + 0.020291\ \text{Ln}(-X)^2 + 0.0061942\ \text{Ln}(-X)^3) \quad (13)$$

where X is the normalized cone response.

Sage 6: Opponent-color signal extraction. A variation of the color-opponent model of Herring is used for the derivation of achromatic and chromatic signals. Equations 14 to 16 provide the formulae used in calculating achromatic (A), red-green difference (RG), and yellow-blue difference (YB) images:

$$A = 0.7186\ L^* + 0.2814\ M^* \quad (14)$$

$$RG = L^* - M^* \quad (15)$$

$$YB = L^* - S^* \quad (16)$$

where L*, M*, and S* are the nonlinear cone responses from stage 5. The coefficients used for the achromatic channel are empirical constants which maximize the correspondence between the spectral response of the achromatic channel and the 1931 CIE spectral luminosity function.

The input to this stage of processing is the three-band (L*, M*, S*) image resulting from the retinal transducer stage. The output is a new three-band image representing the

responses of the achromatic, red-green, and yellow-blue (A, RG, YB) opponent-color stages.

Stage 7: Linear spatial mechanisms. Sixteen oriented spatial mechanisms are used with the achromatic images produced by the previous stage of processing. These mechanisms were generated using four spatial frequencies centered at 3, 6, 12, and 24 c/deg and four orientations of 0, 45, 90, and 135 deg. The bandwidth of each mechanism is approximately one octave in spatial frequency. In contrast with achromatic spatial mechanisms, chromatic mechanisms are not orientation selective as indicated by physiological evidence showing that cells capable of signaling isoluminance chromatic modulation (double-opponent center-surround cells) within the striate cortex are not orientation selective (Livingstone and Hubel, 1984). Thus, red-green and yellow-blue images are operated on by two DOG mechanisms with center frequencies of 3 and 6 c/deg. The bandwidth of the chromatic mechanisms is approximately one octave in spatial frequency.

The achromatic spatial mechanisms were modeled as a 2-dimensional Gabor function. The Gabor function allows the parameters of spatial center frequency, spatial frequency bandwidth, preferred orientation, and orientation bandwidth to be specified separately. Daugman (1983) has promoted the Gabor function as a representation of visual cortical filters because these functions capture all of the fundamental properties of linear neural receptive fields: spatial localization, spatial frequency selectivity, and orientation selectivity. Also, Gabor functions have been used in various models of cortical processing (Daugman, 1980; Watson, 1983, 1987a, 1987b).

Mathematically, the Gabor function is defined as:

$$S(x, y) = G(x, y)\, H(x, y) \tag{17}$$

where

S(x, y) is sensitivity as a function of x and y position,

G(x, y) is a Gaussian windowing function, and

H(x, y) is a cosine grating function.

The Gaussian windowing function is described by the equation:

$$G(x, y) = K * \mathrm{Exp}\{-(x^2 + y^2) / \sigma^2\} \tag{18}$$

where K is constant used to set the gain of the function, and

$$\sigma^2 = s^2 / 2\{\ln(2)\}^{.5} \tag{19}$$

where s is the half-sensitivity width of the function.

The cosine function is described by the equation:

$$H(x, y) = \mathrm{Cos}(\alpha\omega y + \phi) \tag{20}$$

where

x and y are spatial position coordinates,

ω is the spatial frequency in cycles per degree,

α is a constant used to allow ω to be expressed in c/deg rather than the coordinate system of the images used for representing the profile, and

ϕ is the phase angle of the cosine grating selected so that a light bar coincides with the center of the Gaussian window.

The Gabor receptive field profile used for the 24 c/deg horizontal achromatic channel is plotted in Figure 7. The DOG receptive field profile used for the 12 c/deg red-green channel is plotted in Figure 8. Both of these profiles are sampled at 150 pixels/deg.

Computation of the response of a spatial mechanism was performed by centering the response profile of the spatial mechanism over the corresponding test grating as it appeared at the output of the display simulator. The sensitivity at each point of the spatial mechanism was then multiplied by the intensity of the signal corresponding to that point. The overall response of the mechanism was calculated as the sum of the products of sensitivity and intensity. Mathematically this operation is defined as:

$$R = \sum_{i=1}^{n} \sum_{j=1}^{n} I(i + k, j + l)\, S(i, j) \tag{21}$$

where

R is the response of the spatial mechanism,

I(x, y) is the intensity at each point within the neural image,

S(i, j) is the sensitivity at each point of the spatial mechanism,

n is the size (in pixels) of the spatial mechanism,

i, j are spatial position coordinates for the spatial mechanism, and

k and l are the spatial offsets required to center the spatial mechanism in the image.

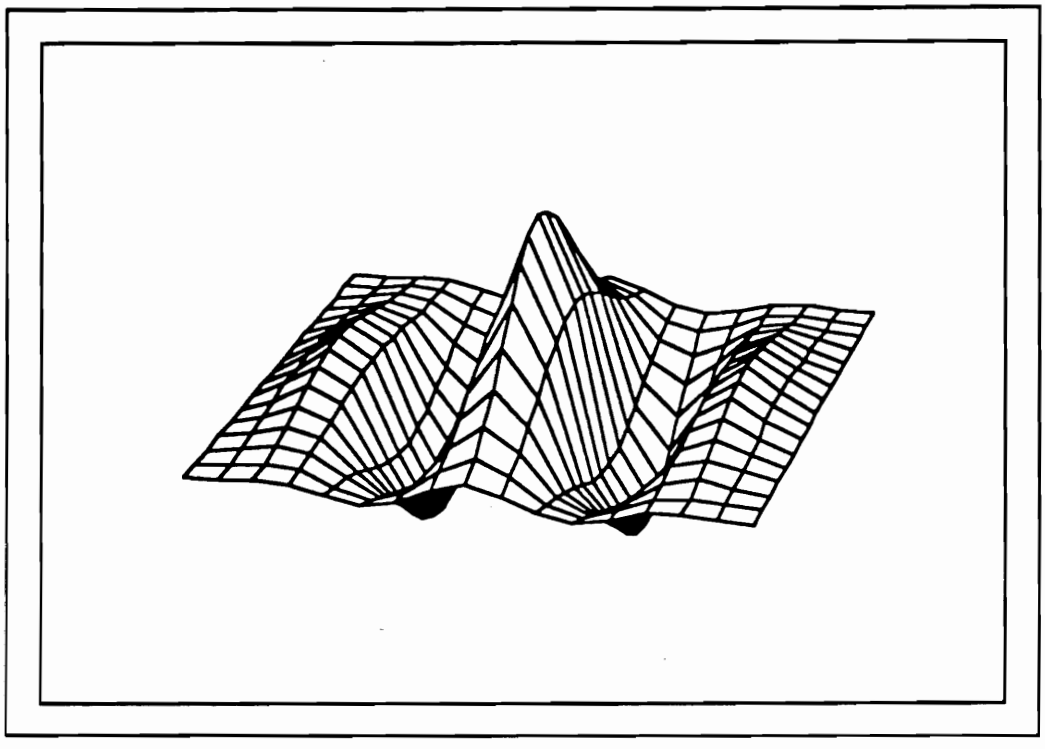


Figure 7. Spatial response profile for 24 c/deg Gabor spatial mechanism.

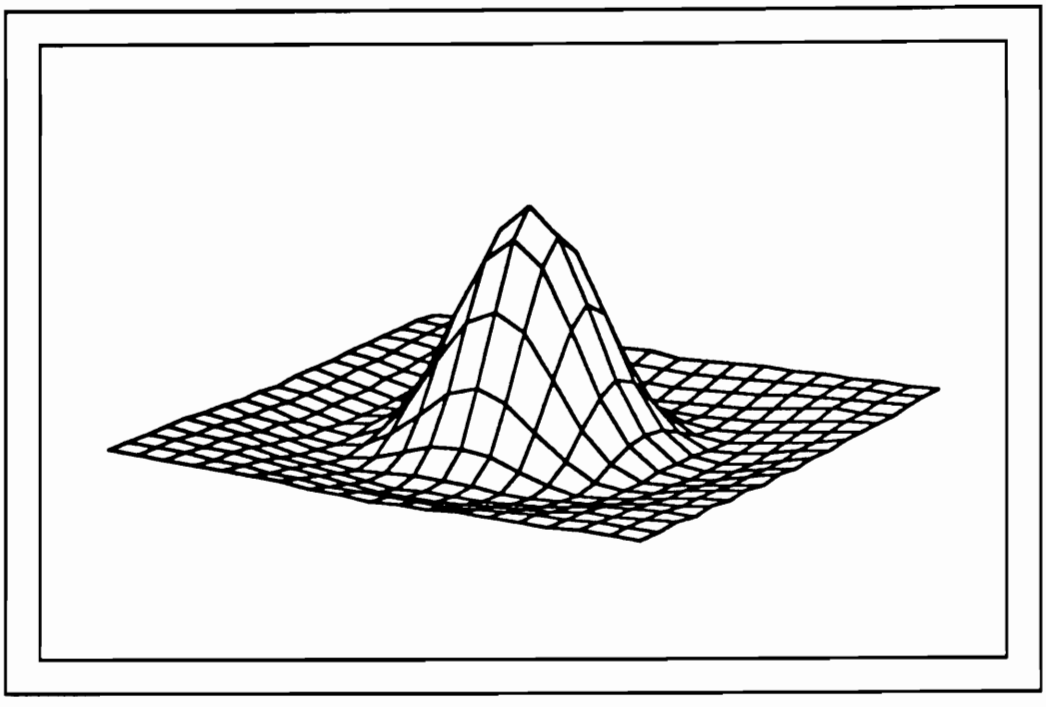


Figure 8. Spatial response profile for 12 c/deg DOG spatial mechanism.

The input to this stage of processing is a three-band image representing the responses of the achromatic, red-green difference, and yellow-blue difference mechanisms. The output consists of the scalar responses of the set of N spatial mechanisms. This set of responses can be thought of as an N-dimensional response vector.

Stage 8: spatial mechanism gains. It seems reasonable to assume that the neural pathways which transmit the outputs of spatial mechanisms have dynamic range characteristics that are independent of spatial frequency. If the dynamic range of these pathways is to be used effectively, then the maximum possible response of a spatial mechanism should be near the maximum signal capability of the pathway. Recall that modulation at high spatial frequencies is significantly attenuated by optical blur. It follows then, that the relative gains of spatial mechanisms must increase as spatial frequency increases. This increase in gain offsets the frequency-dependent decrease in modulation transfer due to the optics of the eye. Note that the relationship between gain and modulation transfer is not linear due to the intervening nonlinear RTF.

An analytic function relating channel gain to mechanism spatial frequency was determined using a series of grating patterns and corresponding spatial mechanisms varying in frequency from 3 to 48 c/deg. Gratings were presented with a modulation of 1.0 and the response of mechanisms of corresponding spatial frequency were calculated. As expected, these responses decreased as spatial frequency increased. Since the 3 c/deg mechanisms was used in the derivation of the RTF, the gain function was constrained such that the maximum response of each of the mechanisms equaled the maximum response of the 3 c/deg mechanism. Figure 9 shows the resulting gain function. A polynomial fit to these data is also presented in Figure 9.

The input to this stage of processing is the multidimensional response vector from the output of Stage 7. The output is the same response vector scaled in response amplitude as a function of spatial frequency.

Stage 9: Probability summation. The effects of random noise are taken into account using a variation of Quick's (1974) approximation of probability summation. Quick's approximation usually is employed to sum the outputs of multiple discrete spatial mechanisms that possess an additive random component (Cannon and Fullenkamp, 1988; Quick, 1974; Wilson and Bergen, 1979). In the proposed model, Quick's probability summation method is used to sum the responses of the same spatial mechanism over time rather than across various spatial mechanisms. Figure 10 illustrates this use of probability summation for predicting the results of a contrast detection or discrimination experiments using the two-alternative forced-choice (2AFC) psychophysical procedure.

The top half of the Figure 10 shows the processing performed on the pedestal level grating. The bottom half of the figure shows the processing performed on the signal + pedestal level grating. Although the processing performed in these two condition is identical, separate paths are drawn to represent the necessary separation of processing in time. The OR function at the output of the spatial mechanisms represents the supposed behavior of the subject in choosing the interval containing the largest of the pedestal + noise or the signal + pedestal + noise conditions. The noise in these two conditions is uncorrelated.

Model implementation. For the computations described above the model was implemented on an IBM PC-AT computer with a hardware numeric coprocessor and the "Turbo Pascal" computer language (version 4). Images used for computations within the

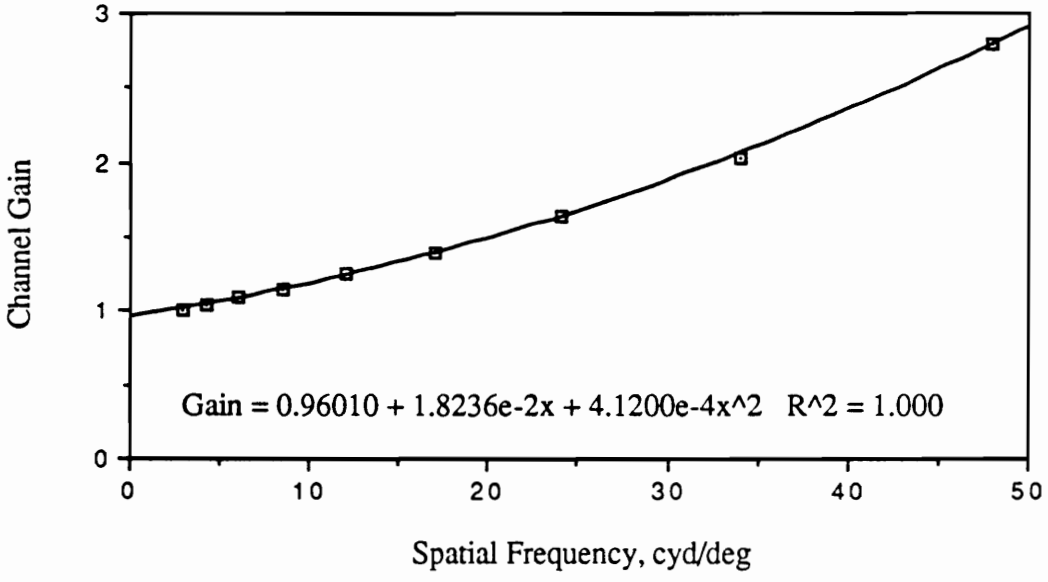


Figure 9. Spatial mechanism gain as a function of spatial frequency.

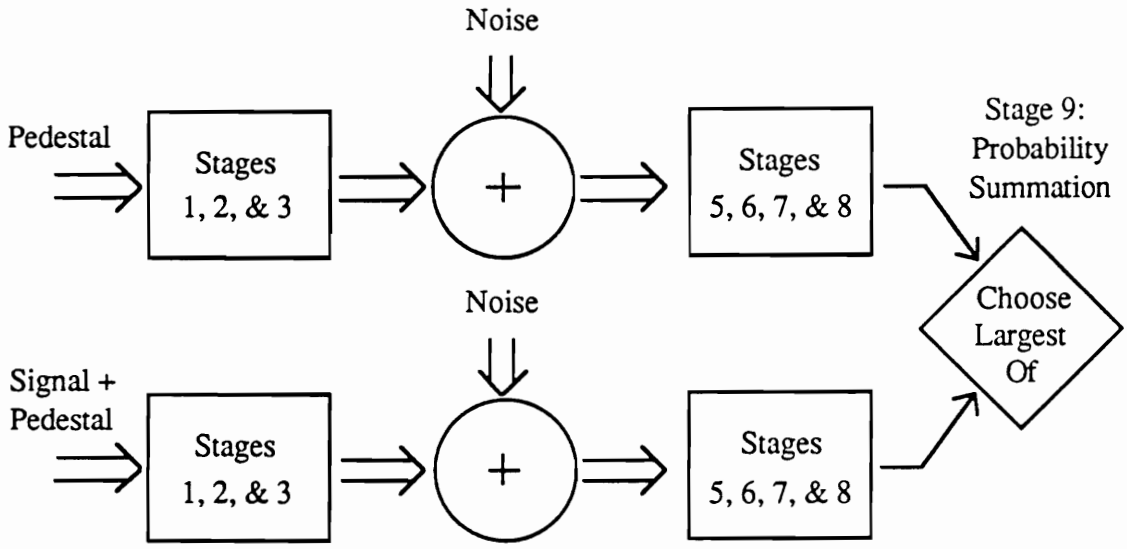


Figure 10. Illustration of the use of probability summation.

HVS model were implemented using 128 x 128 pixels x 16 bits. The precision of intermediate images was greater than that of input images in order to reduce quantization and round-off errors in intermediate stages of processing. Floating point calculations were performed using extended precision real numbers (19 decimal digit mantissa). In situations where many images had to be processed, four IBM PC-ATs were operated in parallel. The computational experiments discussed below required approximately 1000 hours of computing time to complete.

Spectral Effects of RTF

There is a wide-spread assumption by vision researchers that the first stages of human visual processing are linear. Support for this notion can be found in the fact that linear models are used in every one of the approximately 30 papers cited in this dissertation which describe visual channels or human contrast detection, discrimination, or magnitude estimation performance. Similarly, all of the extant display quality metrics cited above use the MTF of a display and thus assume that the first stages of human visual processing are linear. Given this pervasive assumption, an analysis of the effects of the proposed pre-spatial mechanism nonlinearity seems in order. The present analyses were conducted to determine the amount of spurious high frequency noise introduced by the nonlinear RTF. Additionally, the degree to which this noise mediates the performance of the model is discussed.

To gain a general appreciation for the impact of the nonlinear RTF on signals, sinusoidal signals were compressed using the RTF and plotted. Figure 11 shows the results of processing six signals ranging in modulation from 0.00316 to 1.0. The compressive nature of the RTF is revealed by the fact that although these sinusoidal signals

varied in modulation over a 316:1 range, the peak amplitudes of the signals processed by the RTF vary only over a 12.4:1 range. Figure 11 also reveals that as modulation is increased the shape of the processed waveforms become less sinusoidal. To aid in making comparisons among waveform shapes, three of the curves of Figure 11 are re-plotted in Figure 12 with amplitudes normalized to 1.0.

The curve for a modulation of 0.00316 corresponds very closely in shape with that of a pure sinusoid (not shown). This low distortion for gratings of low modulation reflects the effect of the linear center portion of the RTF. For signals with a modulation higher than 0.01 distortion of the waveform becomes apparent. For strongly modulated signals the peaks and valleys of the waveforms are flattened resulting in a waveform somewhat resembling a square wave.

The degree to which this signal distortion might mediate the performance of the model was assessed using a spectral analysis of this distortion. Spectral analyses were conducted using the six processed signals plotted in Figure 11. For each level of modulation 16 cycles of the compressed signal were generated (at 32 samples/cycle) and placed in a 512 element array. These compressed (real valued) arrays were Fourier transformed using a fast Fourier transform (FFT) procedure (Boreland's numerical methods toolbox) running on a Macintosh personal computer. The real and imaginary components of each resulting complex array were combined in quadrature to derive the amplitudes of the spatial frequency components of the compressed waveforms.

Results of these spectral analysis indicate that for the highest level of distortion, approximately 70% of the spectral energy in the waveforms occurred at the fundamental frequency. The remaining energy was distributed at odd integer multiples of the

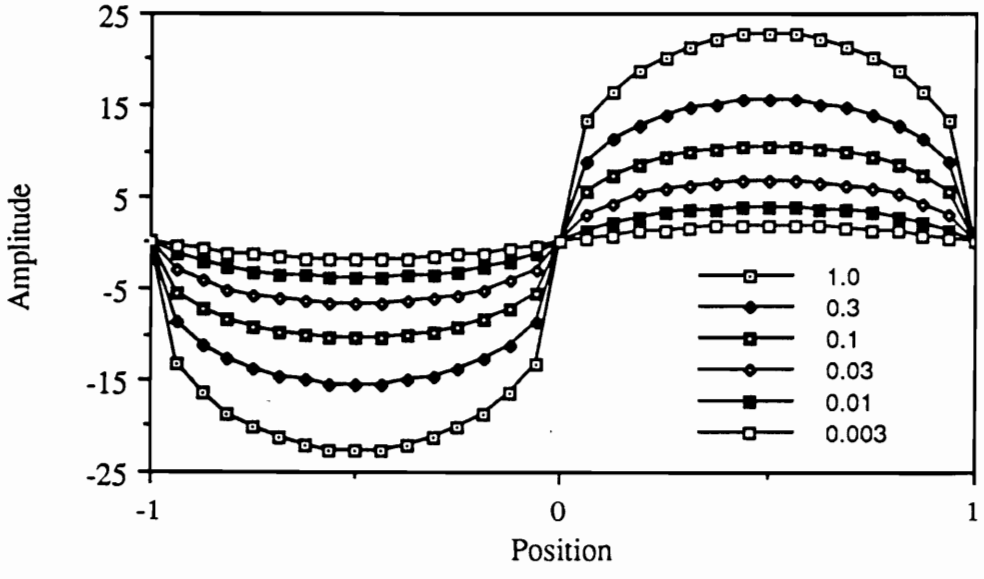


Figure 11. Results of the compression of sinusoidal signals using the nonlinear RTF for six levels of modulation. Parameter of graph indicates modulation of the input sinusoidal signal.

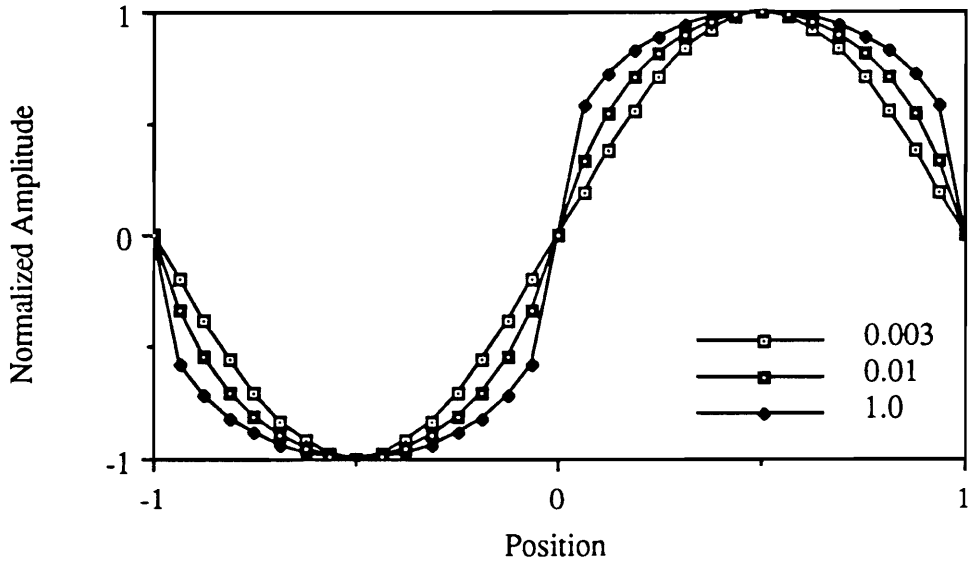


Figure 12. Results of the compression of sinusoidal signals using the nonlinear RTF for three levels of modulation. Peak amplitudes of curves have been normalized to 1.0. Parameter of graph indicates modulation of the input sinusoidal signal.

fundamental (3 x F, 5 x F, 7 x F, ...). The amplitudes of these harmonics decreased as spatial frequency increased. No spectral energy occurred at any other frequency. Note that this pattern of harmonics is similar to the harmonic structure of a square wave. The amplitudes of the fundamental and the first seven harmonics of the compressed waveforms are given for each level of modulation in Table 2. The ratio of the power of the harmonic and the power of the fundamental is plotted in Figure 13 for the first five harmonics. Examination of Figure 13 indicates that the amount of distortion is greatest for fully-modulated gratings and that this distortion decreases as modulation decreases. In the case of a fully-modulated grating, the amplitude of the first harmonic is 20% of the amplitude of the fundamental. For the case of a signal modulation of 0.00316 the amplitude of the first harmonic is 5% of the amplitude of the fundamental.

Results of these analyses indicate that a fully-modulated signal at 13 c/deg will produce a first harmonic at 40 c/deg with an amplitude of 0.197. This harmonic would be at detection threshold indicating that for frequencies above 13 c/deg the distortion products introduced by the RTF would not be visible. Thus, for high spatial frequencies distortion introduced by the RTF should have no impact on threshold or suprathreshold visual performance.

For spatial frequencies above 6 c/deg the sensitivity of the HVS rapidly decreases with increasing frequency. This fact combined with the decrease in the amplitude of harmonics with increasing frequency indicates that for threshold signals with a fundamental above 6 c/deg distortion products introduced by the RTF will be no more than 1/40 th of their threshold values. Thus, for spatial frequencies above 6 c/deg signal distortion by the RTF should have no impact on threshold performance.

Table 2. Results of spectral analysis of compressed sinusoidal signals at six levels of modulation. Entries for the harmonics indicate the ratio of power at the harmonic and power at the fundamental for each level of modulation

		Modulation Level					
Harmonic	Relative Frequency	0.00316	0.00100	0.0316	0.100	0.316	1.0
Fundamental	1 x F	21.79 (0.073)	47.35 (0.159)	85.78 (0.287)	137.0 (0.459)	203.7 (0.682)	298.5 (1.0)
1	3 x F	0.0484	0.1030	0.1480	0.1800	0.1970	0.1970
2	5 x F	0.0067	0.0324	0.0579	0.0793	0.0926	0.0961
3	7 x F	0.0001	0.0139	0.0296	0.0441	0.0538	0.0572
4	9 x F	0.0012	0.0068	0.0168	0.0263	0.0335	0.0363
5	11 x F	0.0012	0.0035	0.0098	0.0157	0.0209	0.0228
6	13 x F	0.0007	0.0017	0.0051	0.0086	0.0115	0.0125
7	15 x F	0.0003	0.0005	0.0016	0.0028	0.0038	0.0040

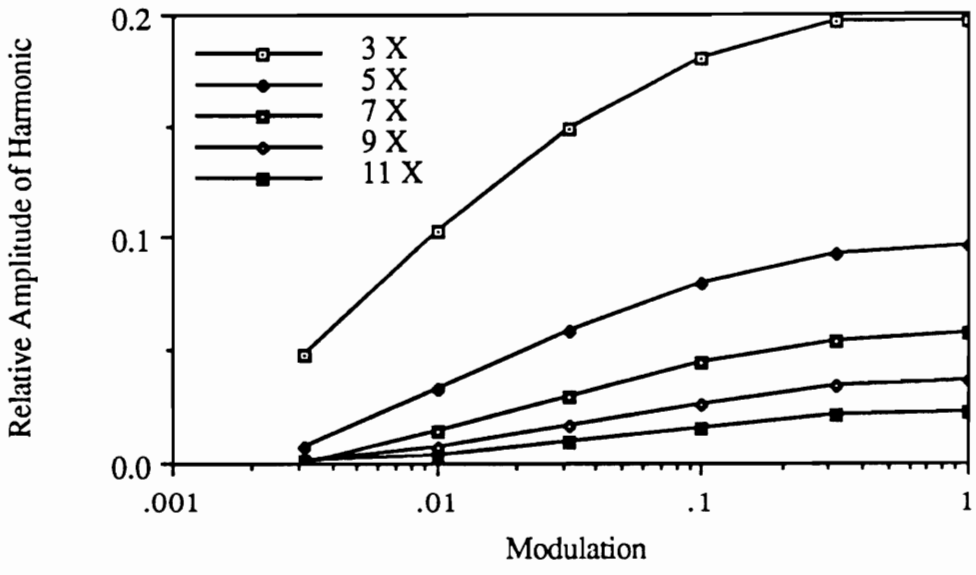


Figure 13. Relative amplitudes of the first 5 harmonics plotted as a function of modulation of the input grating. Parameter of graph indicates the relative frequency of the harmonic.

For threshold signals in the range of 1 to 6 c/deg the modulation of these signal is less than 0.004. Examination of Figure 13 indicates that the amplitude of the first harmonic would be approximately a factor of 20 below threshold. Thus, it seems that distortion products could not influence threshold performance at spatial frequencies above 1 c/deg.

For spatial frequencies below 1 c/deg human contrast sensitivity increases with increasing spatial frequency. It seems then, that the distortion introduced by the RTF might be visible for threshold level signals. To determine whether the distortion components would be visible for low spatial frequency signals, the relative amplitudes of these components have been plotted along with the contrast threshold data of Campbell and Robson (1968) in Figure 14. The fundamental of the compressed signal is at 0.1 c/deg while the frequencies of the first three harmonics are at 0.3, 0.5, and 0.7 c/deg. The amplitude of the fundamental has been set to the contrast detection threshold for a 0.1 c/deg grating. The relative amplitudes of the harmonics are indicated by the diamond symbols. This figure indicates that the harmonics of low spatial frequency signals should be at least a factor of 2 below threshold. It seems doubtful then, that the distortion products introduced by the RTF would affect threshold performance for low spatial frequencies.

The analyses performed here indicate that the distortion products introduced by the nonlinear RTF should have no impact on threshold or suprathreshold performance of the model for spatial frequencies above 13 c/deg. Moreover, these analyses indicate that this distortion should have little or no impact on threshold performance at any spatial frequency. The impact of this distortion on the suprathreshold performance of the model for frequencies below 13 c/deg is more difficult to assess and was not considered in this analysis.

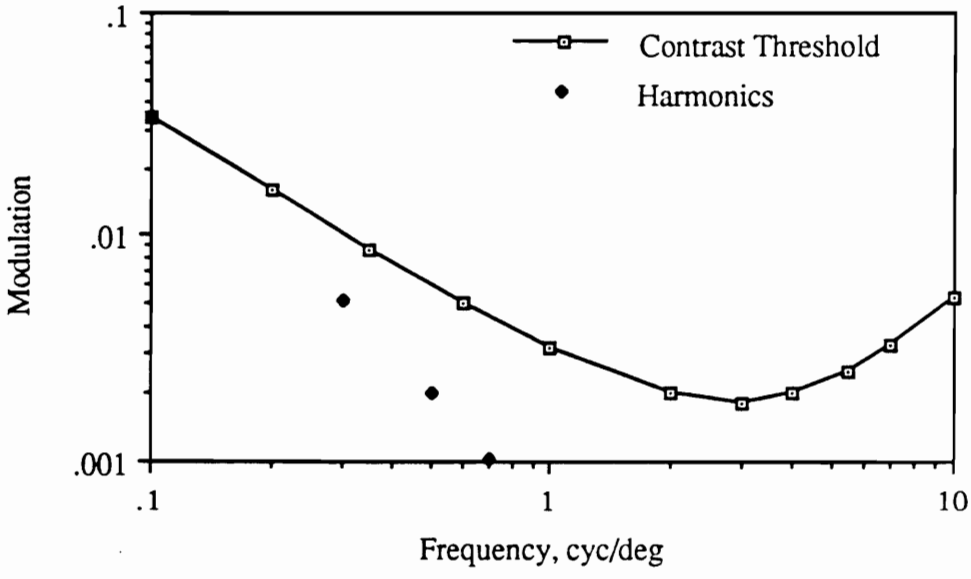


Figure 14. Amplitudes of the fundamental and first three harmonics (diamond symbols) for a compressed threshold sinusoidal signal (at 0.1 c/deg) plotted along with the contrast threshold data of Campbell and Robson (1968).

The results of the above analyses indicate that if the HVS employed a pre-spatial mechanism nonlinearity similar to the proposed RTF, this nonlinearity would have little or no measurable effect on threshold performance. Despite this, linear models have been used almost exclusively for the description of threshold human visual performance (e.g., Blakemore and Campbell, (1969); Campbell and Robson, (1968); Enroth-Cugell and Robson, (1966); Graham, and Nachmias (1971); Pantel and Sekuler, (1968); Sachs, Nachmias, and Robson (1971)). The fact that linear models have been used successfully does not, however, necessarily imply that nonlinear models cannot also predict this performance. The HVS model proposed here is one formulation of a nonlinear model which should produce the same threshold results as the linear models suggested by previous researchers. A significant advantage of the present model over existing threshold models is that the present model also explains suprathreshold visual sensitivity without requiring the use of a post-spatial mechanism nonlinearity.

DISPLAY EVALUATION PROCEDURE

Display system evaluation procedures using the proposed HVS model are described below. In general, the procedures involve the use of signal gratings consisting of cosine spatial frequency grating patterns at various frequencies, orientations, and chromatic content. A display simulator is used to present each of these gratings to the HVS model for the calculation of signal and noise responses. The display simulator is also used to present literal imagery to human subjects for experiments designed to test the validity of the HVS model.

For the evaluation of a visual display, the main objective is to compute HVS model responses for signal and noise resulting from a displayed image. Signal and noise calculations are based on processing of the grating signals. One grating signal corresponds in spatial frequency, orientation, and chromatic type with each spatial mechanism in the HVS model. For each grating signal the modeled visual mechanism (e.g., channel) most sensitive to that frequency, orientation, and chromatic type provides the signal or "target" response, while the remaining N-1 mechanisms of the HVS model provide noise or "non-target" responses.

It is important to note that noise is defined as any visual information contained in the simulated grating image that does not have a counterpart in the original grating signal. Since the spatial frequency and orientation sensitivity profiles of individual spatial mechanisms in the HVS model do not overlap significantly, only the target mechanism should respond to a noise-less grating signal. Thus, noise responses are assessed by examining the responses of non-target mechanisms.

For the N signal gratings used, one signal response and N - 1 noise (non-target) responses are calculated. Since it is unwieldy to work with N - 1 noise responses for each mechanism , the noise responses are combined as a root mean square (RMS) measure. Thus, the response to each signal grating is described by a signal response and an RMS noise response. The difference between signal and noise responses provides an index of perceivable visual information. The overall procedure for analyzing a single simulated display system is given in pseudo code in Figure 15.

The response of the HVS model to a "perfect" display system is determined by calculating signal and noise responses to grating signals directly, without the intervening processing stages of the display system simulator. The responses of four spatial channels to a perfect display are presented in Figure 16.

With no attenuation of signal gratings by the display system, the response of each channel in the HVS model is 51 JNDs. This value is indicated by the height of all bars in Figure 16. With no noise added to signals by the display, the basal response rate of the model is determined entirely by noise internal to the model. These basal responses are dependent on spatial frequency as indicated by the shorter bars in Figure 16. The numbers appearing above each bar indicate the JNDs of signal minus the JNDs of noise. These JND differences represent the JNDs available for transmitting signals over each channel. The JNDs available in each of the N channels should correspond to the overall capability of the display. Across 16 achromatic channels (4 spatial frequencies x 4 orientations) the total JNDs available is 667.

All physically realizable display systems will attenuate signals and add noise. Visible signal attenuation will reduce the corresponding mechanism signal responses from the 51-

```
For each grating signal  Do
Begin

    Simulate presentation of grating on display under evaluation

    For each spatial mechanism  Do
    Begin

        Calculate response of spatial mechanism to simulated grating

    End

    Separate target and non-target responses and store

End

For each spatial mechanism  Do
Begin

    Calculate RMS noise response using non-target responses

    Subtract noise response from signal response

    If noise response > signal response  Then
        Signal response - noise response = 0

End;

Sum the signal - noise responses for each mechanism
```

Figure 15. Display evaluation procedure expressed in pseudo code.

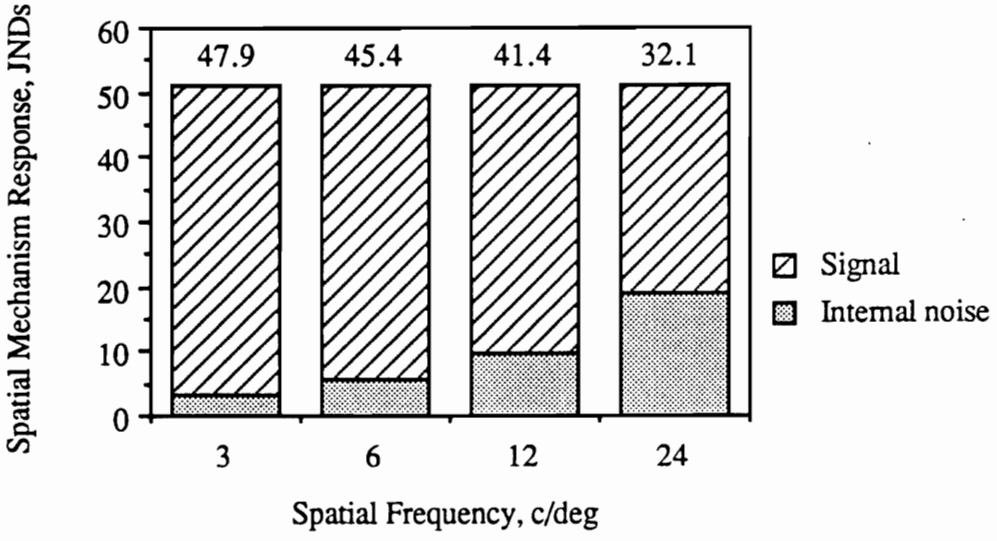


Figure 16. Spatial mechanism responses obtained without the use of a display simulator.

JND level calculated for a perfect display. Suprathreshold noise, on the other hand, will increase noise responses above the corresponding basal level. Both of these physical display characteristics decrease overall perceivable information levels.

Figure 17 diagrams the two distinct ways in which the display simulator is used within this dissertation. The lower path in the diagram indicates those components of the system used for evaluating displays. The upper path in the diagram indicates those components necessary for conducting human performance studies. Note that gratings are used for display system evaluations, whereas literal imagery is used for human visual performance evaluations.

The validity of the HVS model and display evaluation procedure was tested with the model (e.g., lower path in Figure 17) and with human observers (e.g., the upper path in Figure 17). The statistical correlation between these two evaluation approaches provides a quantitative index of validity.

In summary, the proposed procedure involves a signal-to-noise analysis of signal gratings which correspond with spatial frequency, orientation, and chromatic channels in the HVS model. This procedure measures and perceptually weights the ability of a display system to transmit signals. The ability of a display system to transmit signals strongly influences human visual performance.

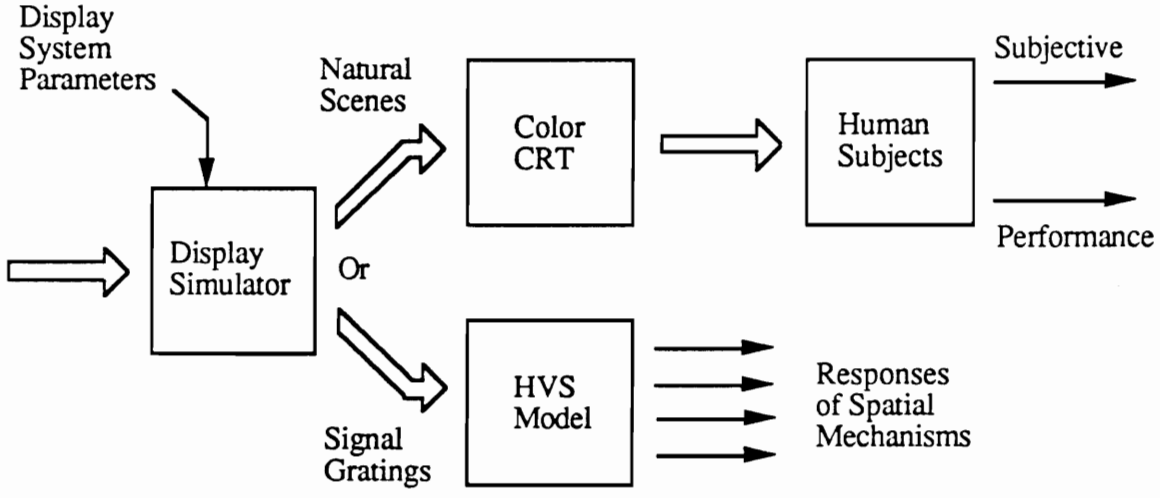


Figure 17. Block diagram of evaluation system showing data flow.

EXPERIMENT 1: PSYCHOPHYSICAL DATA

Two classes of validation experiments were performed on the proposed HVS model. First, the performance of the model was compared with extant data from classic psychophysical experiments published in the past decade. These comparisons are discussed in this section. Second, the utility of the model for evaluating display systems was assessed by comparing model responses with estimates of image quality made by human subjects. Experiments 2, 3, and 4 are of this second class.

Three aspects of human visual performance which have received much attention from visual scientists and which are important for the evaluation of display systems are contrast detection, contrast discrimination, and contrast magnitude estimation. Contrast detection responses have been used previously to perceptually weight engineering metrics of image quality such as the MTFA (Snyder, 1988). Although metrics such as the MTFA have demonstrated engineering utility it has been suggested that image quality judgements also involve suprathreshold human visual performance since most tasks performed using visual displays are performed at suprathreshold levels of modulation. Contrast detection, discrimination, and magnitude estimation data provide complimentary sources of information regarding real-world human visual performance.

Contrast Discrimination Performance

The contrast discrimination performance of the HVS model was characterized by calculating contrast discrimination functions of the form measured for human subjects. Starting with a grating of fixed pedestal modulation, a JND in modulation was determined using an iterative algorithm to find the increase in modulation over the pedestal level

required to produce a unit change in the output of the model. Given the derivation of the RTF (see Appendix B), a unit change in model output corresponds with a change in modulation of the input grating that is detectable 75% of the time using the 2AFC psychophysical procedure. Modulation JNDs were calculated in this manner for four spatial frequency channels using pedestal modulations ranging from 0.001 to 1.0. The resulting contrast discrimination functions are plotted in Figure 18. In this figure, the arrows along the ordinate indicate threshold modulations (using a pedestal modulation of 0) for the four curves.

The discrimination function for the 3 c/deg spatial mechanism has a slope of 0.78. Note, however, that the upper modulation portion of the 3 c/deg function was used in the derivation of the RTF. Thus, the fact that this curve matches human performance does not provide an independent validation of the model. Rather, it verifies that the computational implementations of the model work correctly.

Other attributes of the curves in Figure 18 provide independent validation of the model. First, the functions converge at high pedestal modulations. This result is consistent with recent studies showing that suprathreshold contrast discrimination performance is largely unaffected by a factors such as spatial frequency, display noise level, and luminance level which strongly affect threshold detection levels (Bradley and Ohzawa, 1986; Burton, 1981; Lloyd, Beaton, Parsley, and Simonetti, 1990; Maattanen, Koenderink, and Nienhuis, 1988; Pelli, 1985). Also, there is a characteristic dip in the functions for near-threshold pedestal modulations. This characteristic, which has been found in numerous studies, is attributed to the facilitative effect of the sub-threshold pedestal (Bradley and Ohzawa, 1986; Burton, 1981; Legge, 1981; Legge and Foley, 1980; Lloyd et al., 1990; Nachmias and Sansbury, 1974; Pelli, 1985; Wilson, 1980). Thus, the contrast

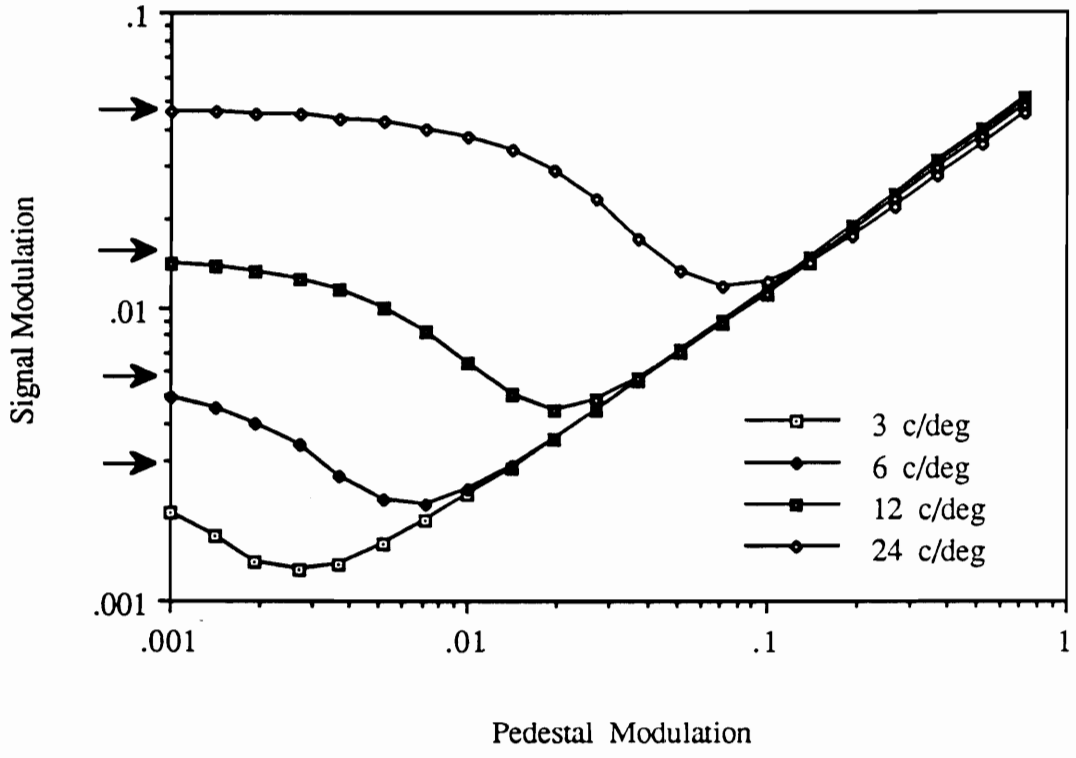


Figure 18. Contrast discrimination performance of the HVS model for four spatial frequencies.

discrimination functions calculated for the HVS model qualitatively and quantitatively agree with previous psychophysical research in this area of study.

Contrast Detection Performance

The contrast detection performance of the HVS model was calculated using a zero modulation pedestal and the same procedure used for calculating contrast discrimination performance. With zero pedestal modulation, the basal response of the model is determined solely by the level of internal noise. The contrast detection threshold occurs at the point where the response of the model to an input grating results in a model output one unit larger than the basal response rate. Contrast detection thresholds were calculated at nine spatial frequencies between 3 and 48 c/deg. The results of these calculations are presented in Figure 19. For comparison, Figure 19 also shows an estimate of the CTF for human vision determined by averaging the results of a number of studies appearing in the visual science literature (Farrell and Booth, 1984). The performance of the HVS model closely matches the human performance data over the 3 to 48 c/deg range covered by the model.

Contrast Magnitude Estimation

The contrast discrimination functions discussed above show the *change* in input modulation required to produce a unit *change* in model response. The data discussed here differ somewhat in that subjective magnitudes rather than the occurrence of changes in responses are characterized. In a typical magnitude estimation study, subjects are asked to scale their judgements of contrast magnitude. Ostensibly, these ratings correspond to some internal signal level. As a validation test of the proposed HVS model, it is desirable to

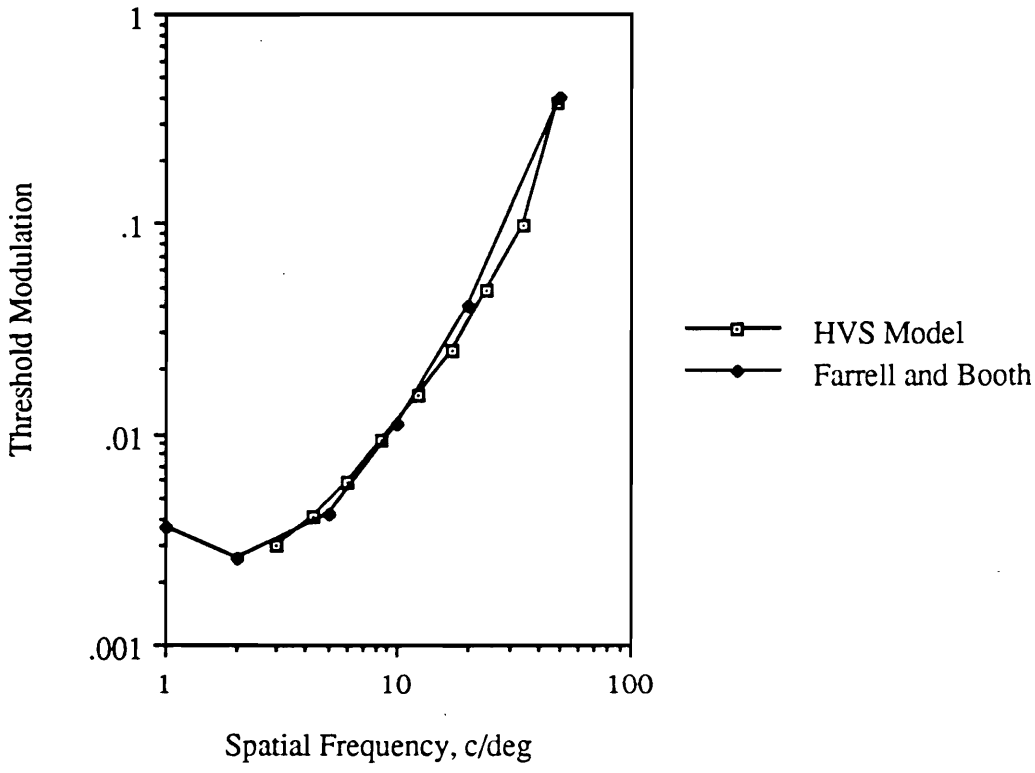


Figure 19. Contrast detection threshold of HVS model as a function of spatial frequency.

show that some level within the model correlates with contrast magnitude estimates. To test this assertion, model response magnitudes were calculated as a function of input modulation for four spatial frequency gratings. The resulting curves appear in Figure 20.

Similar to the contrast discrimination curves in Figure 18, the response curves in Figure 20 converge at high modulation. This data trend reflects a perceptual phenomenon known as "contrast constancy" (Cannon, 1985; Cannon and Fullenkamp, 1988). Contrast constancy refers to the tendency for gratings of equal physical contrast to appear equal even though factors such as spatial frequency and number of grating cycles induce large changes in threshold detection levels. In addition to contrast constancy effects, the slope of the upper modulation portion of the curves in Figure 20 is 0.35. This slope compares well with the slopes of 0.31 and 0.5 found in contrast perception studies (Cannon, 1985; Cannon and Fullenkamp, 1988).

The data shown in Figure 20 indicate that as modulation is decreased toward subthreshold levels the response of each spatial mechanism approaches a level determined by internal noise. In contrast perception studies, however, ratings of contrast magnitude go to zero for sub-threshold gratings. This obvious discrepancy between model responses and human performance can be explained simply.

For suprathreshold gratings the data support the hypothesis that magnitude estimates are based on mechanism responses. That magnitude estimates are zero for subthreshold gratings does not necessarily imply that mechanism responses are zero. Rather, a reasonable explanation is that subjects factor their ability to discriminate responses to gratings from responses to noise into their estimates of contrast magnitude. When a grating

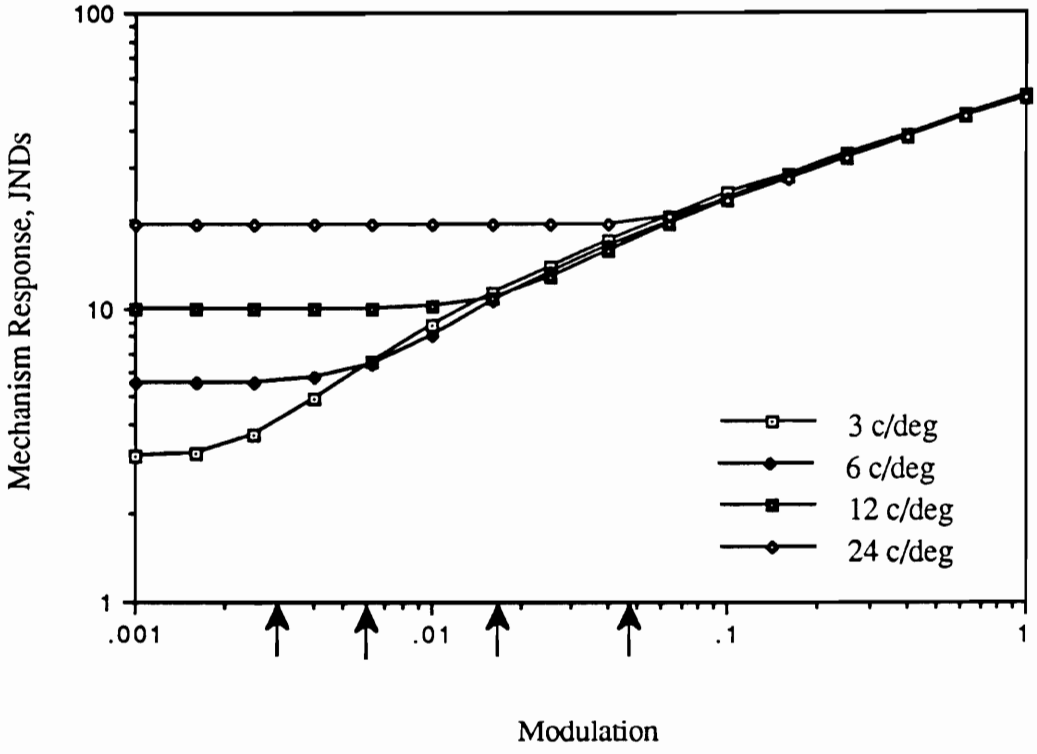


Figure 20. Spatial mechanism responses as a function of grating modulation for four spatial frequencies.

is below threshold, responses of a mechanism to the grating cannot be discriminated from responses to noise, thereby resulting in a magnitude estimate of zero.

Conclusions

Two primary data sets were used in the derivation of the proposed HVS model: the optical blur function of Campbell and Gubisch (1966), and the estimate of the 3 c/deg contrast discrimination function used in the derivation of the RTF. Additionally, an estimate of the contrast detection threshold for a 3 c/deg grating was used to set the level of internal noise. Given the limited amount of data used in developing the model, the finding that the model successfully predicts contrast detection, discrimination, and magnitude estimation data as a function of spatial frequency bolsters confidence in the structure of the model.

The experiments described above reveal that the responses of the proposed model are proportional to perceived contrast. It thus seems reasonable to view the HVS model as a means of mapping stimulus magnitudes into the perceptual space used by subjects. Since the contrast discrimination performance of the model reveals that JNDs in modulation map into unit differences in model response, it follows that the perceptual space defined by the model is scaled in terms of JNDs. Since the zero-level response of the model has been set arbitrarily, the absolute magnitudes of model responses are not intrinsically meaningful. The model user should thus attend to differences in model responses rather than their absolute levels.

EXPERIMENT 2: MONOCHROMATIC DISPLAY QUALITY

Purpose

As mentioned above, an important constraint on the proposed metric was that its predictions of imaging performance for monochromatic CRTs should be at least as good as those of extant MTF-based metrics. The present experiment assessed this predictive capability in comparison to those metrics examined by Beaton (1984).

Human Performance Data

In a monochromatic display evaluation, Beaton systematically varied the quality of photographic and CRT-based images. Twenty-five experimental conditions were used in which images were degraded using all combinations of 5 levels of blur and 5 levels of random noise. Both photographic and CRT-based images were evaluated by groups of photointerpreters who estimated image quality in each experimental condition. The subjective image quality ratings for photographic images are presented in Figure 21 while ratings for the CRT-based images are presented in Figure 22. Subjects used a NATO 0 to 9 scale for making ratings. Subjects were instructed to base quality ratings on "the usefulness of the image for military intelligence purposes" (Beaton, 1984, p. 79). Noise is expressed in terms of RMS transmissivity and blur is expressed in terms of the half-width (in pixels) of the effective blur filter applied to the images.

Beaton provided a thorough description of the method of image degradation used in the generation of his human performance data. These descriptions are complete enough that accurate simulations of the display systems can be constructed. Thus, comparisons of the

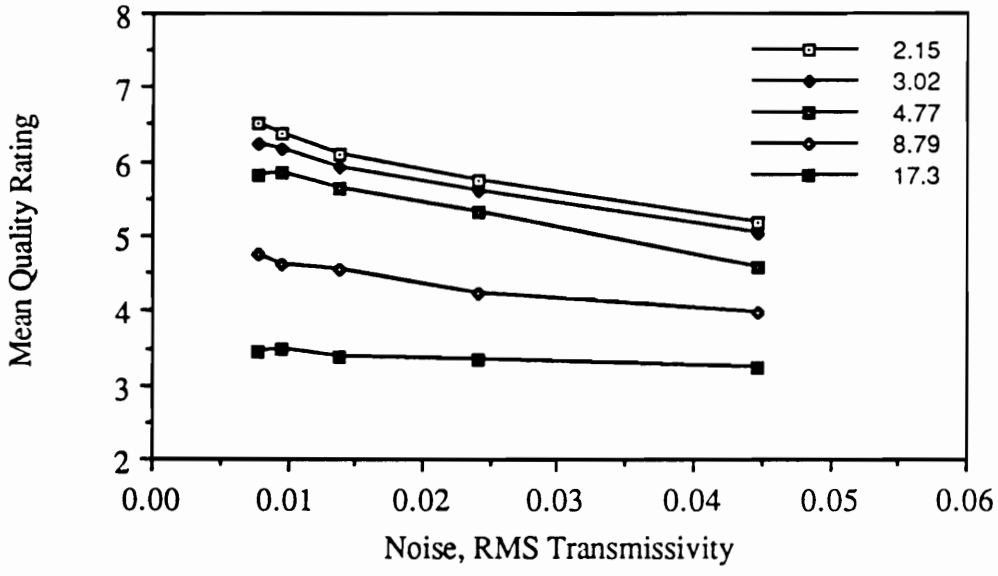


Figure 21. Subjective image quality ratings for the photographic images used by Beaton (1984). Parameter of graph is half-width (in pixels) of blur function.

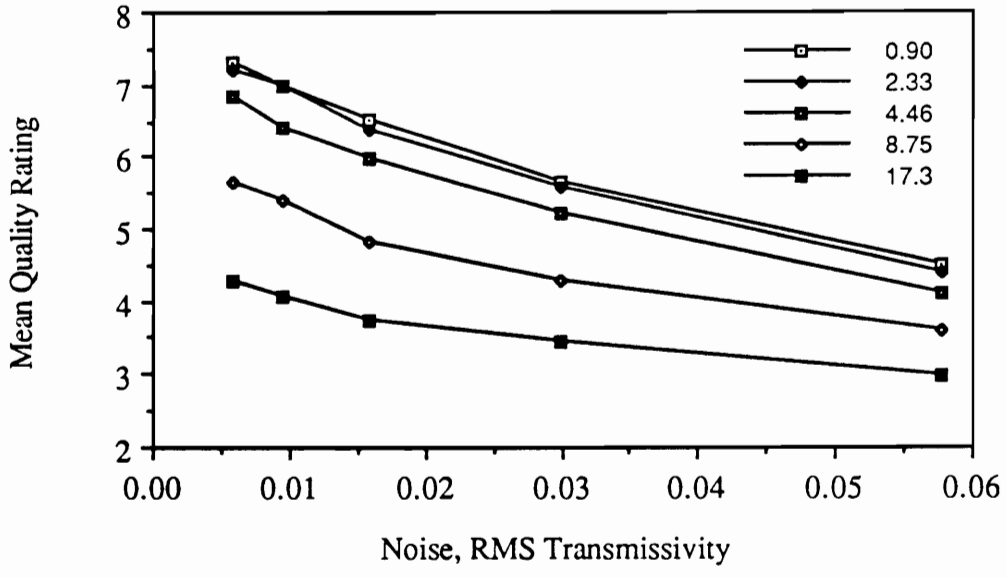


Figure 22. Subjective image quality ratings for the CRT images used by Beaton (1984). Parameter of graph is half-width (in pixels) of blur function.

performance of the proposed evaluation system can be made with the performance of the metrics evaluated by Beaton without having to collect additional human performance data. The most important advantage in using the Beaton data is that performance of the proposed metric can be compared easily with all 16 of the monochromatic metrics evaluated by Beaton.

One factor complicates use of the Beaton data. Since subjects in those experiments were allowed to select various image magnification levels, determination of image content in terms of spatial frequencies is somewhat arbitrary. This ambiguity was dealt with by assuming that subjects viewed displays from a fixed "optimal" distance determined by the Nyquist sampling limit of the original (4096 x 4096) images. Viewing distance was assumed to be that distance which resulted in an image sampling rate of twice the high-frequency cutoff for human vision (approximately 48 c/deg). Thus an image sampling rate of 96 pixels/deg was used. The MTFs of the display systems used by Beaton (provided in his Figure 8, p. 56) could then be scaled in units of spatial frequency.

Image Quality Predictions

Predictions of image quality for the 25 experimental conditions used by Beaton were made using the display evaluation procedure described above. Simulation of the displays involved blurring and adding noise to the set of grating signals as if these gratings were presented over each of the 25 displays. Since chromatic signals were not manipulated in this experiment the responses of the chromatic channels were not considered in the analysis. The predictions made by the HVS model for the 25 photographic display conditions are presented in Figure 23. Predictions for the CRT display conditions are presented in Figure 24. Inspection of Figures 21 to 24 indicates a strong correspondence

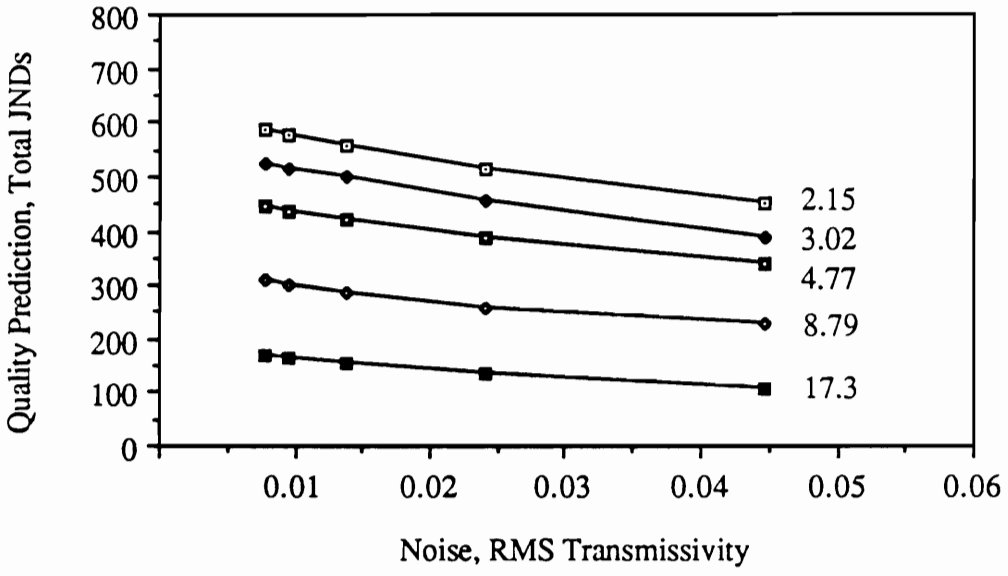


Figure 23. Image quality predictions made using the display evaluation system for the photographic images used by Beaton (1984). Parameter of graph is half-width (in pixels) of blur function.

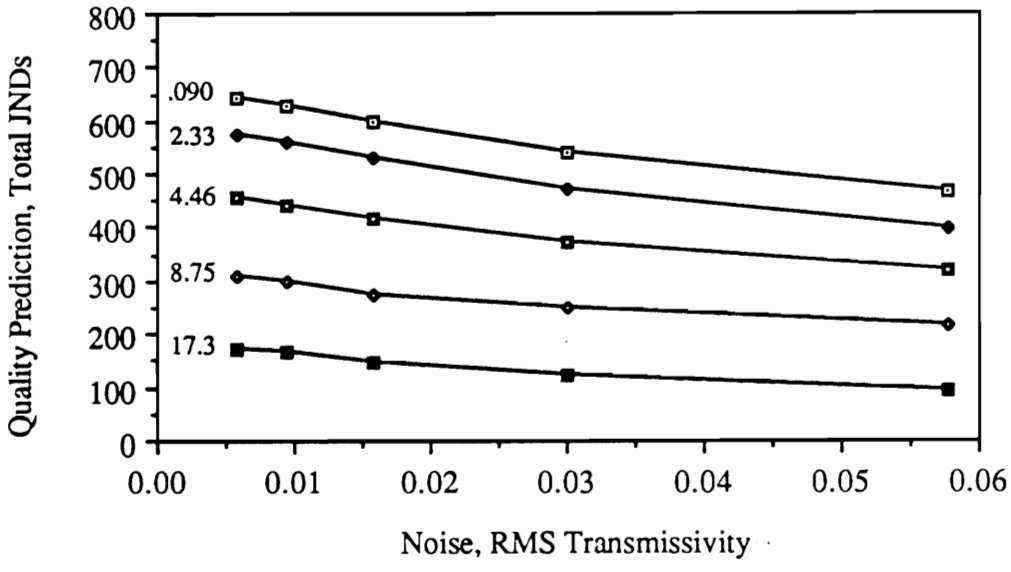


Figure 24. Image quality predictions made using the display evaluation system for the CRT images used by Beaton (1984). Parameter of graph is half-width (in pixels) of blur function.

between predictions of the proposed model and actual human performance data. The statistical correlation between predicted photographic image quality and ratings of the photographic images is 0.985 ($R^2 = 0.969$, N = 25). This correlation is higher than any found by Beaton using 16 monochromatic image quality metrics. Figure 25 graphically illustrates the correspondence between the predictions and ratings for photographic images.

The statistical correlation between predicted CRT image quality and ratings of CRT images is 0.892 ($R^2 = 0.792$, N = 25). This correlation is higher than the average correlation found by Beaton using 16 monochromatic image quality metrics. Figure 26 graphically illustrates the correspondence between the predictions and ratings for CRT images.

Conclusions

The high degree of correlation between predictions of the proposed metric and image quality ratings is encouraging. For photographic images the proposed metric correlates with performance more strongly than any other metric based solely on display MTFs. For CRT images, predictions of the proposed metric are not as strongly correlated with performance, however, the metric still accounts for 79% of the variance in these data. These results indicate that the display evaluation methodology performs at least as well as the best extant monochromatic image quality metrics.

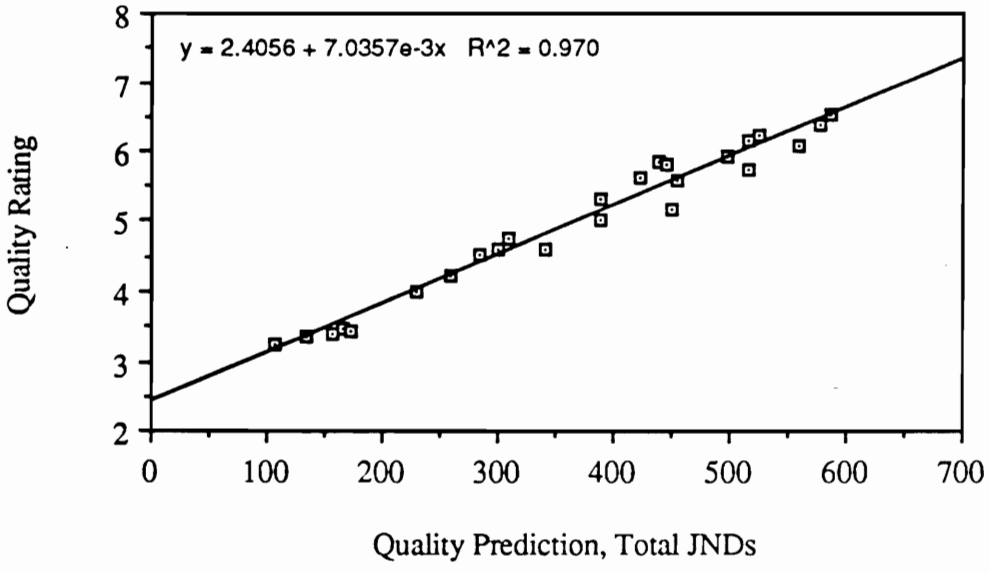


Figure 25. Correspondence between predictions and ratings of image quality for photographic images.

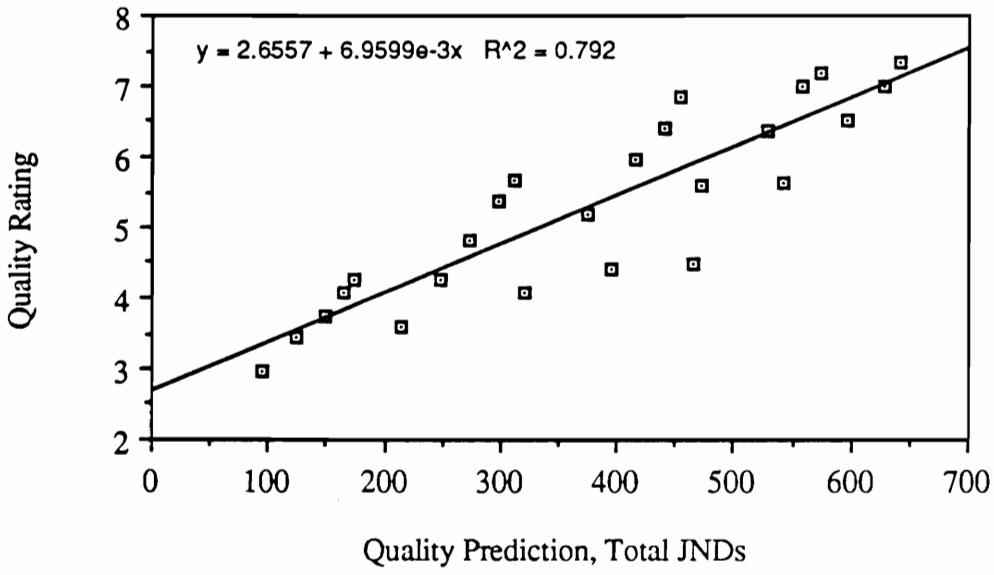


Figure 26. Correspondence between predictions and ratings of image quality for CRT images.

EXPERIMENT 3: COLOR MATRIX DISPLAY QUALITY

Purpose

The primary purpose of this research effort has been to develop a method for evaluating full-color display systems. Obviously the proposed model and method must be validated using color display systems. Numerous human performance data sets that can be used to assess the monochromatic performance of the proposed model are available in the literature. In contrast, few data are available for color displays. While research on the perceptual aspects of color displays has been conducted, few researchers have manipulated design variables which directly or indirectly impact chromatic modulation transfer or chromatic noise.

Human Performance Data

Two studies have been published recently which address the issue of image quality on color matrix display (CMD) systems (Rogowitz, 1988; Silverstein and Lepkowski, 1986). Both studies examine the impact of color pixel geometries on estimates of image quality. While chromatic modulation transfer was not directly manipulated, there is considerable variation in the amount of chromatic and achromatic noise among these four display types.

The pixel geometries used in these experiments were the delta triad, RGBG quad, diagonal line, and vertical line geometries. In these pixel grid geometries, square color elements are used resulting in "pixels" with a 3:1 length-to-width ratio for the diagonal and vertical line geometries. The pixel grid geometries are illustrated in Figure 27. Color element size is expressed in units of minutes of visual arc.

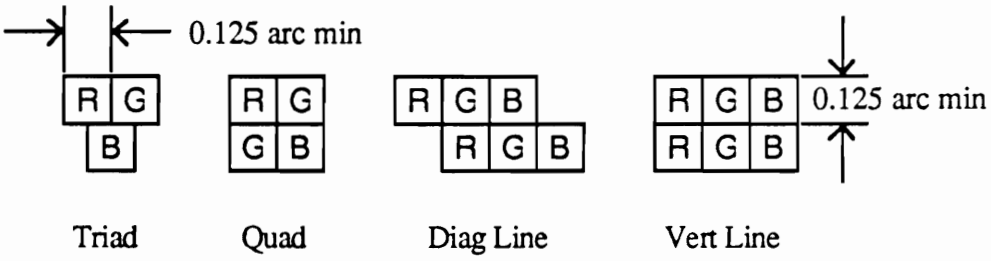


Figure 27. The four CMD pixel geometries evaluated by Rogowitz (1988) and Silverstein and Lepkowski (1989).

Results of the two studies for gray-scale CMDs indicate that subjects rated the delta triad highest in quality. The RGBG quad received the next highest rating, followed by the diagonal line geometry. The vertical line geometry received the lowest rating.

Image Quality Predictions

Predictions of image quality for these CMDs were made using the proposed evaluation method. The evaluation procedure was identical to that used in estimating monochromatic image quality. The dimensions and layout of color elements used in the CMD simulations were identical to those used by Rogowitz (1988). The image quality predictions made using the proposed model are presented in Figure 28.

Although Rogowitz and Silverstein and Lepkowski evaluated color display systems, the stimuli they used were monochromatic. Thus, the Rogowitz and Silverstein data sets do not contain variance attributable to chromatic modulation transfer characteristics. Therefore, the responses of the chromatic channels of the HVS model were not included in the present analysis. The quality predictions of Figure 28 are scaled in units of the total number of JNDs available on the 16 achromatic channels of the HVS model.

Conclusion

Predictions made using the HVS model correspond well with ratings made by subjects in the human performance studies. The correlation between predicted image quality and the Rogowitz data set is 0.99 ($R^2 = 0.98$), while the correlation with the Silverstein et al. data set is 0.89 ($R^2 = 80$). Though these correlations are based on few data points, the proposed model appears useful for predicting image quality on matrix addressable displays. Additionally, no difficulties with the use of astigmatic pixels were encountered.

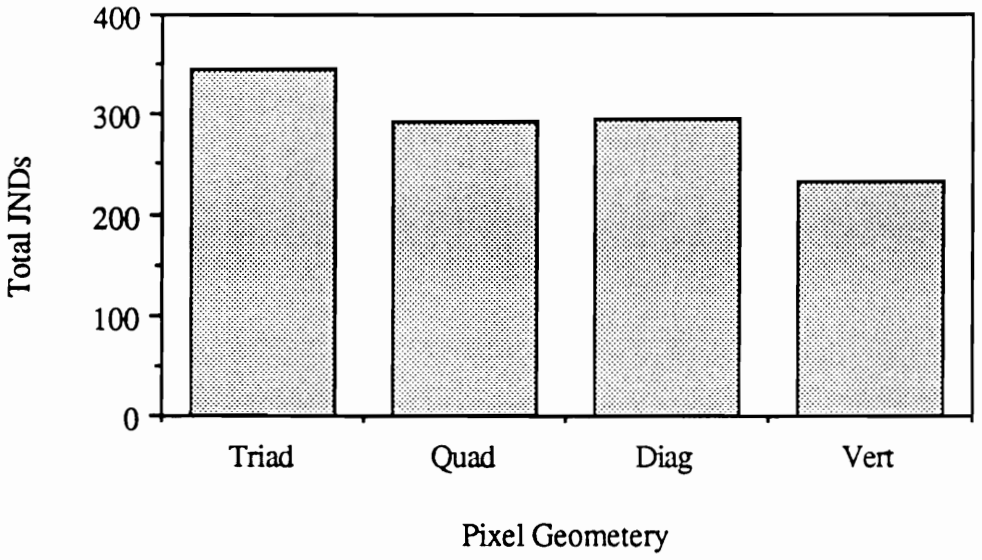


Figure 28. Image quality predictions made using the display evaluation procedure for the four CMD pixel geometries used by Rogowitz (1988) and Silverstein et al. (1989).

EXPERIMENT 4: CONTINUOUS COLOR DISPLAY QUALITY

Purpose

The overall goal of this experiment was to extend the validation of the proposed evaluation system into the realm of continuous full-color display systems. This experiment was designed to address three specific research questions:

Do manipulations of (isoluminance) chromatic attenuation and noise capabilities of a display significantly impact ratings of image quality?

If these manipulations significantly impact quality, what is the relative importance of chromatic content with respect to achromatic image content?

If these manipulations significantly impact quality, how well do predictions of the proposed method track ratings of image quality?

Method

Subjects. Eleven subjects (six female) between 18 and 38 (mean = 21.8) years of age participated in this experiment. Subjects were recruited from the Virginia Tech campus using an advertisement posted in the campus newspaper and they received pay for participation. Subjects were screened for normal near-field contrast sensitivity using a Vistech Consultants Inc. vision contrast test system. Subjects also were screened for normal color vision using a Farnsworth-Munsell 100-hue test for color discrimination.

Apparatus. All processing of images for this experiment was done using an IBM PC-AT equipped with a numeric coprocessor. A graphics controller (Vectrix Corporation, model "Pepe") was used to present images with up to 256 simultaneous colors on the

monitor. The graphics controller contained 1MB of memory. An addressability of 1024 x 1024 pixels was used for rendering images. Images were processed prior to the experimental sessions and stored on an IBM 3363 optical disk drive. Each image required 2 Mbyte of disk space. Images could be read from the optical drive to the graphics controller in 20 seconds using a Vectrix-supplied procedure that utilizes the direct memory access (DMA) capability of the IBM PC-AT.

A display viewing station allowed high-resolution color images to be viewed from a normal reading distance of 40 cm. The display viewing station consisted of a color CRT and an optical system designed to adjust accommodation and convergence to the desired reading distance.

The color CRT used here was a Barco Industries "Calibrator" color graphics monitor. The Barco monitor contains a delta-triad shadow-mask color CRT with a mask pitch of 0.31 mm . The monitor was configured for an active display area of 0.27 x 0.27 m. At a distance of 2.0 m the active area subtended 7.7 x 7.7 degrees of visual arc. The monitor was viewed through a stereo optical system containing a -505 mm focal length (-2.0 diopter) negative meniscus lens and a pair of first-surfaced mirrors for the optical path of each eye. At the 2.0 m actual distance of the monitor, the -2.0 diopter lens required subjects to focus at 0.40 m. One mirror in each optical path was fixed in a three-point mount to allow fine horizontal and vertical adjustment of each image. Horizontal adjustments were made so that the eyes converged at a distance of 0.40 m.

While this optical arrangement corrects accommodation and convergence to the desired distance, it does not change the angular subtense of images presented on the monitor. The display viewing station simulates the use of a 5.4 x 5.4 cm, 133 pixels/deg color display

viewed at a distance of 0.40 m. Figure 29 illustrates the configuration of the monitor and optical components used.

An adaptation of Robert's method for randomizing quantization noise (Schreiber, 1986, pp. 101-104) was employed for rendering images without producing brightness or color banding. With this scheme, the 8-bit color look-up table (CLUT) was partitioned such that the red and green primaries each used 3 bits (8 levels) while the blue primary used 2 bits (4 levels). Digital-to-analog converter (DAC) values were selected for the CLUT such that the luminance range of each primary was divided into luminance steps of equal size. Luminance was measured using a Minolta CS-100 colorimeter calibrated against a Hoffman Engineering LS-65-8B luminance standard. Chromaticity coordinates were also measured for each display primary using the Minolta colorimeter. The measured luminance and chromaticity coordinate values (for each gun full-on) are given in Table 3.

Robert's method involves the addition of uniformly distributed noise to each pixel in the image before the image value is rounded to the nearest step in the CLUT. The range of the uniform noise distribution was selected to be equal to the step size of the CLUT. While the addition of random "noise" to an image might be expected to degrade image quality, Schreiber shows that this technique does not increase the total amount of noise in the image. The advantage of the technique results from the spatial distribution of structured noise (banding) caused by sharp thresholding between DAC steps. Though the same total noise power is present in the resulting image, the random noise is less visible since its spectrum is flat as opposed to the spectrum of the structured noise which contains sharp peaks at spatial frequencies corresponding to the banding. The technique produced images with a low but perceptible level of "salt and pepper" noise.

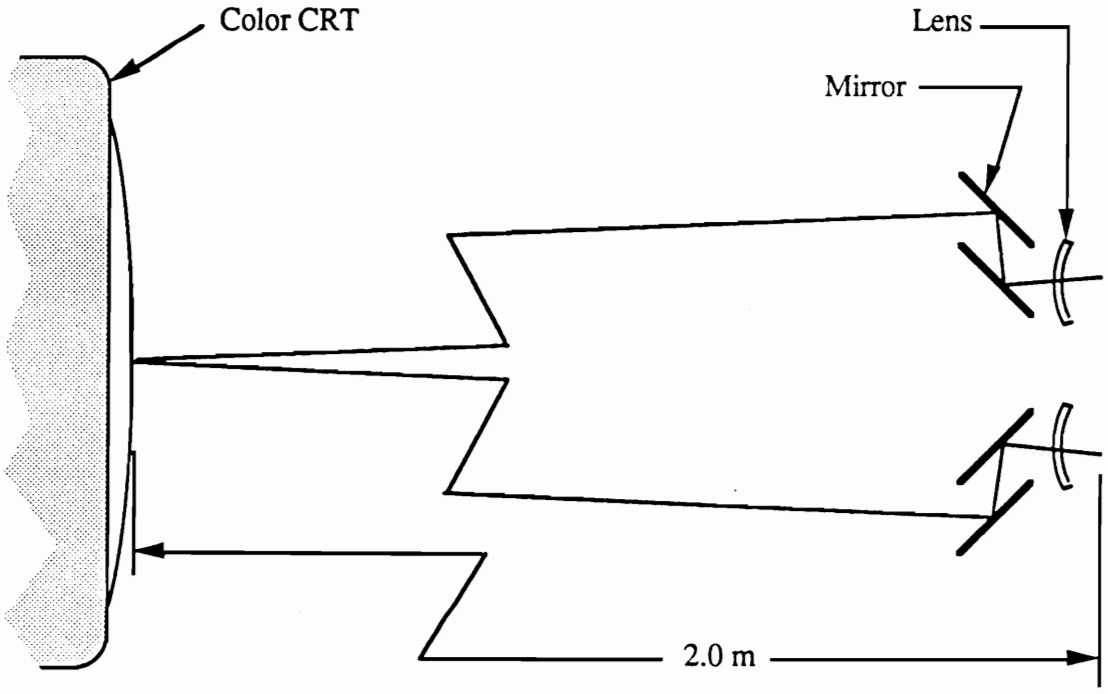


Figure 29. Arrangement of components for stereo optical system. For each eye a -2.0 diopter lens and a pair of first-surface mirrors were used to set accommodation and convergence to 0.40 m.

Table 3. Measured luminance and chromaticity coordinate values for phosphors on Barco monitor

Color	Luminance	x	y
Red	14.5 cd/m^2	.601	.355
Green	48.8 cd/m^2	.296	.589
Blue	5.76 cd/m^2	.149	.066

Images. The content of the image used in this experiment was selected using the following criteria:

The scene should contain common objects with colors highly familiar to subjects;
the scene must contain the colors red, green, yellow, and blue;
the colors should be of medium to high saturation; and
the scene must contain both low and high spatial frequencies.

A digital image or color photograph meeting these criterion could not be found. Thus a scene with the appropriate content was composed. The scene contained a fruit bowl set on a blue background. The fruit in the scene included an apple, an orange, a lemon, a lime, a peach, a banana, and green and purple grapes. An image of this scene was captured by arranging the selected objects on a digitizing table and digitizing the scene directly. No intermediate photographic step was used.

Images were captured using a Perceptics 9200 image capture system using a Sierra Scientific CCD camera. Images were represented using three 512 x 512 pixel x 8 bit sub-images or bands. Thus, three separate shots of the scene were taken using red, green, and blue filters in front of the camera lens. An infrared-reflecting dichroic mirror was used during image capture to reduce the sensitivity of the camera to the large amount of infrared radiation produced by the four quartz-halogen lamps used to illuminate the scene. This original 512 x 512 x 3-band image was mapped into the 1024 x 1024 images presented on the Barco monitor using linear interpolation and Robert's method described above.

Figure 30 shows a photograph of the scene as it appeared on the Barco monitor used for displaying images to subjects. This photograph is intended only to give the reader an

Figure 30. Photograph of the nondegraded image used in Experiment 4.

idea of the scene content. No attempt has been made to insure that the colors in the photograph accurately represent those that appeared on the monitor.

Experimental design. The experiment was conducted using a three-factor, full-factorial, within subjects design. The three factors were Attenuation level, Noise level, and Channel type. This experimental design is illustrated in Figure 31.

The "Attenuation level" was set to one of four levels determined by the ratio of modulation present in the degraded and original images. The attenuation levels used were 1.0, 0.5, 0.25, and 0.125.

One of four levels of static noise was used to degrade images. The noise was uniformly distributed with no correlation among pixels within the 512 x 512 x 3-band original image. Noise levels were set at 0, 25, 50, and 100% of the maximum possible range of intensity values in the image . Since the standard deviation of a uniform distribution is 29% of the range, these noise levels correspond to standard deviations of 0, 7.25, 14.5, and 29% of the intensity range.

In order to add chromatic noise and attenuation to the original image without also impacting the distribution of luminance, the original RGB image was first transformed into a three-band BW-RG-YB (achromatic, red-green, yellow-blue) image using the linear transformation:

$$\begin{bmatrix} BW \\ RG \\ YB \end{bmatrix} = \begin{bmatrix} 0.210, & 0.707, & 0.083 \\ 1.000, & -1.000, & 0.000 \\ 0.500, & 0.500, & -1.000 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (22)$$

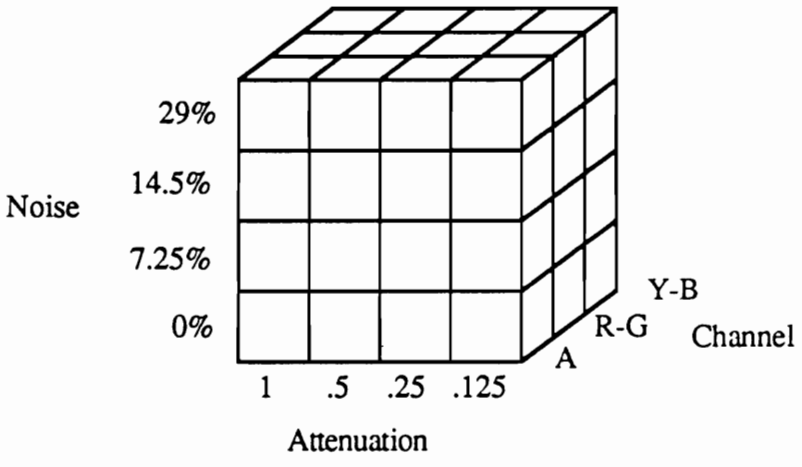


Figure 31. Illustration of the experimental design for Experiment 4.

The coefficients used to derive the BW channel are the relative luminances of the R, G, and B primaries for the Barco monitor.

Noise was added to one channel at a time by adding uniformly distributed noise to one of the transformed image bands. Similarly, the modulation in any one of these image bands could be attenuated by multiplying the transformed image values by an attenuation factor of less than 1.0. After attenuating or adding noise to one band of the transformed image, the image was transformed back into an RGB image using the inverse transformation:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.0, & 0.7485, & 0.0830 \\ 1.0, & -0.2515, & 0.0830 \\ 1.0, & 0.2485, & -0.9170 \end{bmatrix} \cdot \begin{bmatrix} BW \\ RG \\ YB \end{bmatrix} \quad (23)$$

Procedure: Subjects. Upon arrival at the laboratory subjects were asked to read and sign informed consent forms. Subjects then were given two vision screening exams described above. After completing the screening exams, subjects were seated at the viewing station which presented the original (nondegraded) image. As subjects examined the image, the experimenter read aloud instructions on the rating procedure to be used (see Appendix A). A series of 8 practice trials was then given to allow subjects to become comfortable with the equipment and the free-modulus rating procedure. The practice trials also allowed subjects to view the overall range of image degradation. Practice trials were randomly selected from among the 48 experimental trials.

After completing the practice trials a series of 48 experimental trials was given. Each subject completed the trials in a unique random order. For each trial, subjects were allowed to view the image for up to 30 seconds. The average time taken by subjects to view images

was 10 to 15 s. Ratings for each image were stated aloud by the subject and recorded by the experimenter. After making each rating, the subject's view of the display was blocked by a shutter while the next image was loaded to the graphics board. Due to time required to load images, an inter-stimulus interval of approximately 20 seconds was used.

After the presentation of all 48 images subjects were debriefed and paid for their participation. The average duration of an experimental session (from subject arrival to subject departure) was approximately 1.25 hr.

Results: Subjects. Two analyses were performed on the subjective rating data collected in this experiment. First, the data were analyzed using an analysis of variance (ANOVA) procedure to determine if the experimental manipulations reliably affected subjective ratings of image quality. The degree to which model predictions corresponded with subjective rating data was assessed by calculating the statistical correlation between these data.

Since subjects were free to choose their own response scale in the quality judgements, rating data from each subject were normalized to a standard range prior to analysis. The rating data from each subject were range-corrected by dividing each subject's scores by the standard deviation of those scores. Thus, the rating data submitted to the ANOVA had a range of approximately 0 to 3.5. Results of the ANOVA are shown in Table 4.

The ANOVA results indicate that each of the experimental variables -- channel, attenuation level, and noise level -- significantly affected ratings of image quality ($p < 0.0001$). Additionally, the results indicate that all of the interactions among these variables were significant at the same level of significance.

Figure 32 shows the means for the main effect of channel. A Newman-Keuls post hoc

Table 4. Analysis of Variance Summary Table for Image Quality Rating Data

Source	df	SS	MS	F	p
Subjects (S)	10	34.749	3.475	-	-
Channel (C)	2	255.022	127.511	258.46	< 0.0001
C x S	20	9.867	0.493	-	-
Attenuation (A)	3	81.890	27.297	82.38	< 0.0001
A x S	30	9.940	0.331	-	-
Noise (N)	3	51.333	17.111	102.36	< 0.0001
N x S	30	5.015	0.167	-	-
C x A	6	24.396	4.066	19.72	< 0.0001
C x A x S	60	12.371	0.206	-	-
C x N	6	7.476	1.246	6.69	< 0.0001
C x N x S	60	11.168	0.186	-	-
A x N	9	12.568	1.396	13.48	< 0.0001
A x N x S	90	9.321	0.104	-	-
C x A x N	18	8.569	0.476	4.77	< 0.0001
C x A x N x S	180	17.952	0.100	-	-
Total	527	551.636			

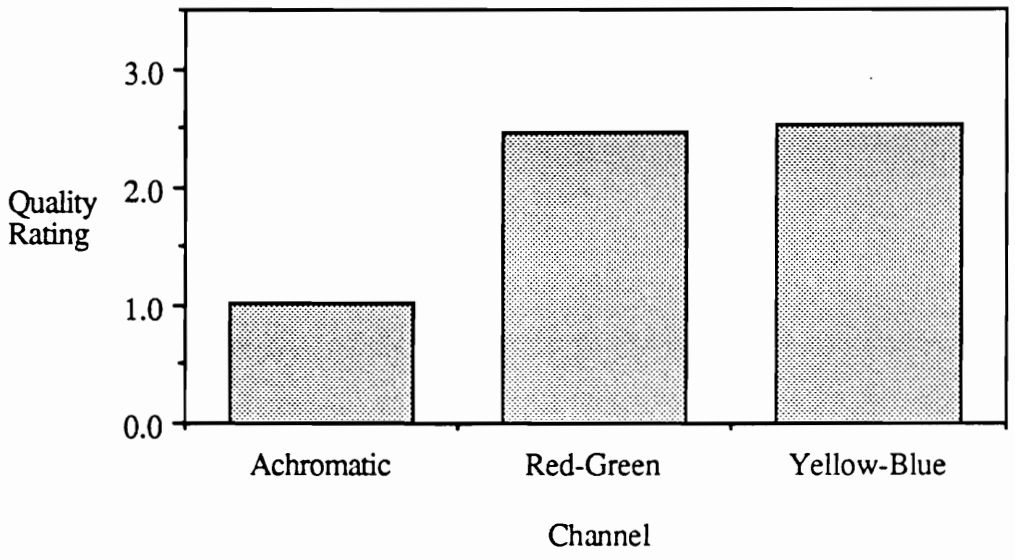


Figure 32. Image quality ratings as a function of channel.

test indicates that manipulations of the achromatic channel degraded image quality more strongly than manipulations of the chromatic channels ($p < 0.05$). The difference between the red-green and yellow-blue means was not significant ($p > 0.05$).

Image quality ratings are plotted as a function of attenuation level in Figure 33. As expected, quality ratings increase as the attenuation level approaches 1.0. A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the lowest two conditions (attenuation = 0.125 and 0.25). However, the remaining differences among these means were significant ($p < 0.05$).

Ratings of image quality are plotted as a function of noise level in Figure 34. Quality ratings decrease as the level of noise increases as expected. A Newman-Keuls post hoc test indicates that all of the differences among these means were significant ($p < 0.05$).

Quality ratings are plotted as a function of noise level for each channel in Figure 35. Quality ratings decrease as noise increases for all three channels; however, the rate of this decrease is not the same for each channel. A post hoc simple-effects F-test indicates that noise significantly degraded quality for all three channels ($p < 0.01$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the second and third conditions for the achromatic channel (noise = 7.25 and 14.5). However, the remaining differences among these means were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the first two conditions for the red-green channel (noise = 0 and 7.25). However, the remaining differences among these means were significant ($p < 0.05$).

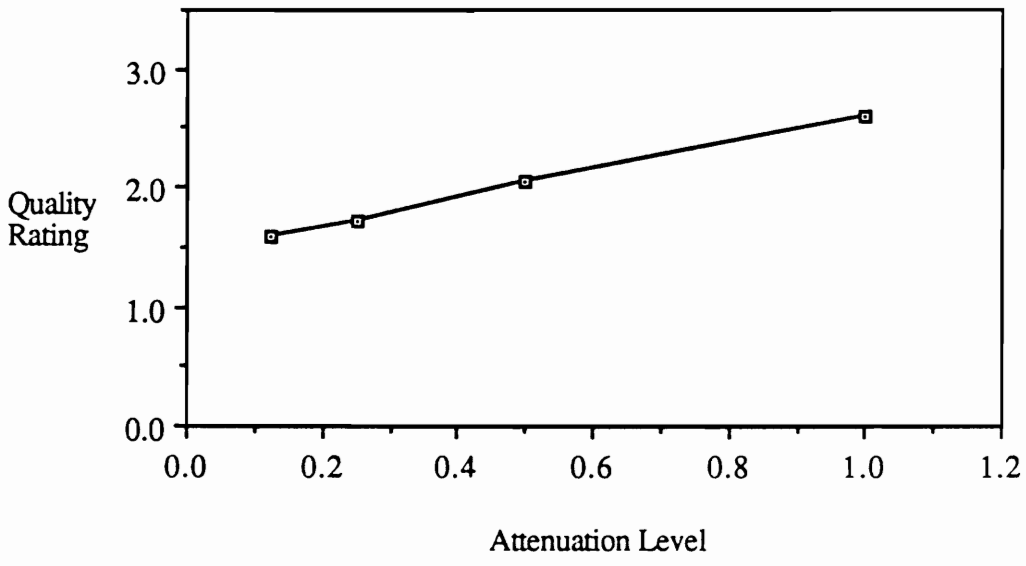


Figure 33. Image quality ratings as a function of level of attenuation.

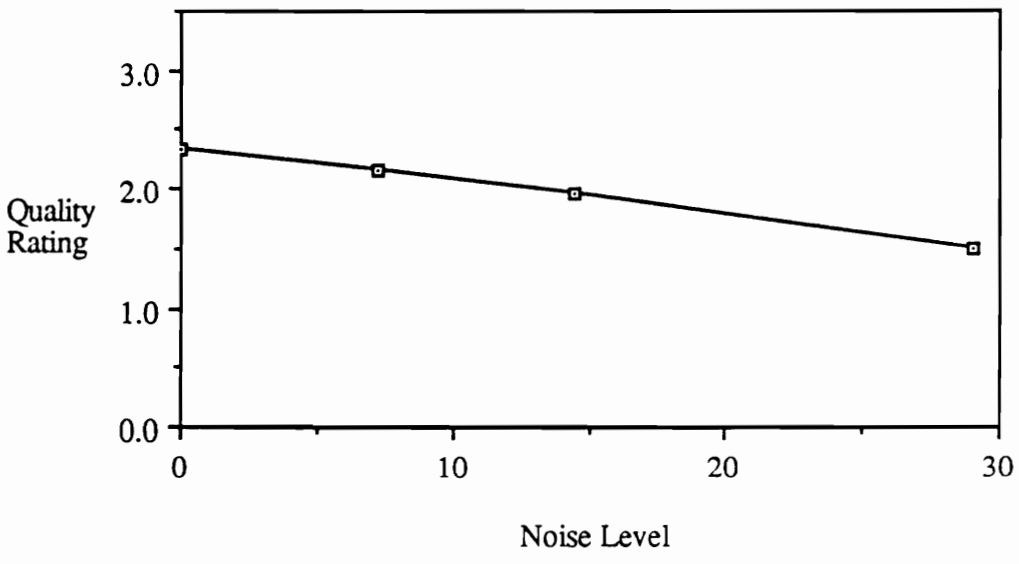


Figure 34. Image quality ratings as a function of level of noise.

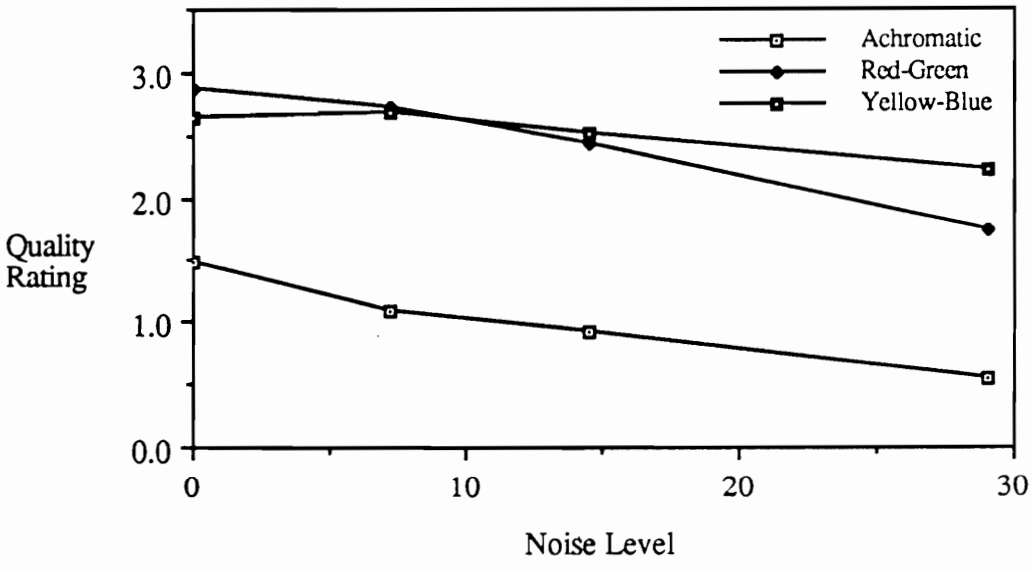


Figure 35. Image quality ratings as a function of channel and level of noise.

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the first three conditions for the yellow-blue channel (noise = 0, 7.25, and 14.5). The mean for the highest noise condition differs significantly from the other three means ($p < 0.05$).

Quality ratings are plotted as a function of channel and attenuation level in Figure 36. This figure reveals that the rate of decline of quality ratings with decreasing modulation transfer is greatest for the achromatic manipulations. A post hoc simple-effects F-test indicates that attenuation significantly degraded quality for all three channels ($p < 0.01$).

A Newman-Keuls post hoc test indicates that all of the differences among the means for the achromatic channel were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the lowest two conditions (attenuation = 0.125 and 0.25) for the red-green channel. However, the remaining differences among these means were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the lowest three conditions (attenuation = 0.125, 0.25, and 0.5) for the yellow-blue channel. The mean for the no-attenuation condition differs significantly from the other three means ($p < 0.05$).

Quality ratings are plotted as a function of noise level for each level of attenuation in Figures 37. This figure shows that fully-modulated images were more strongly impacted by display noise than were attenuated images. However, a post hoc simple-effects F-test indicates that noise significantly degraded quality at all levels of attenuation ($p < 0.01$).

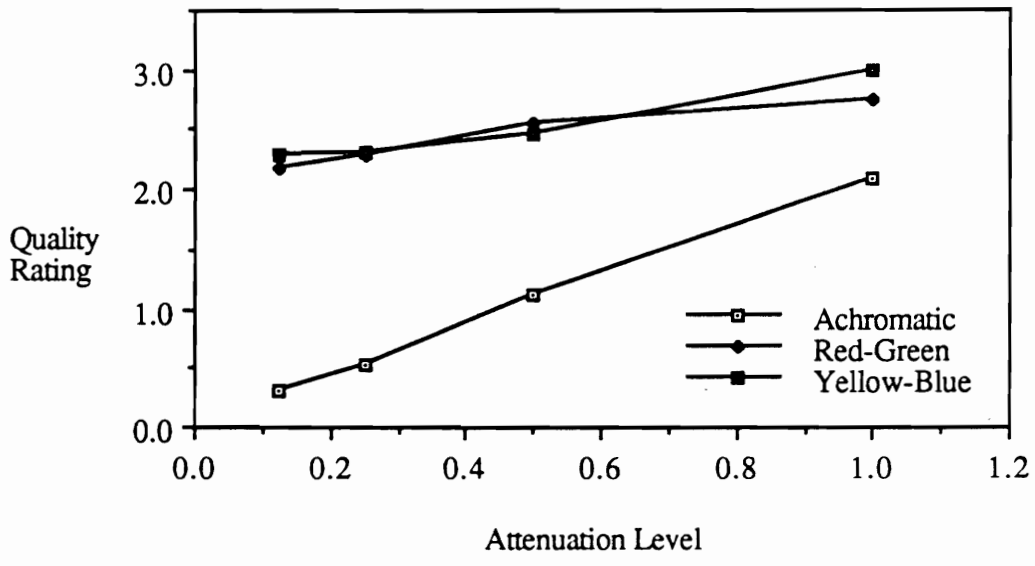


Figure 36. Image quality ratings as a function of channel and attenuation level.

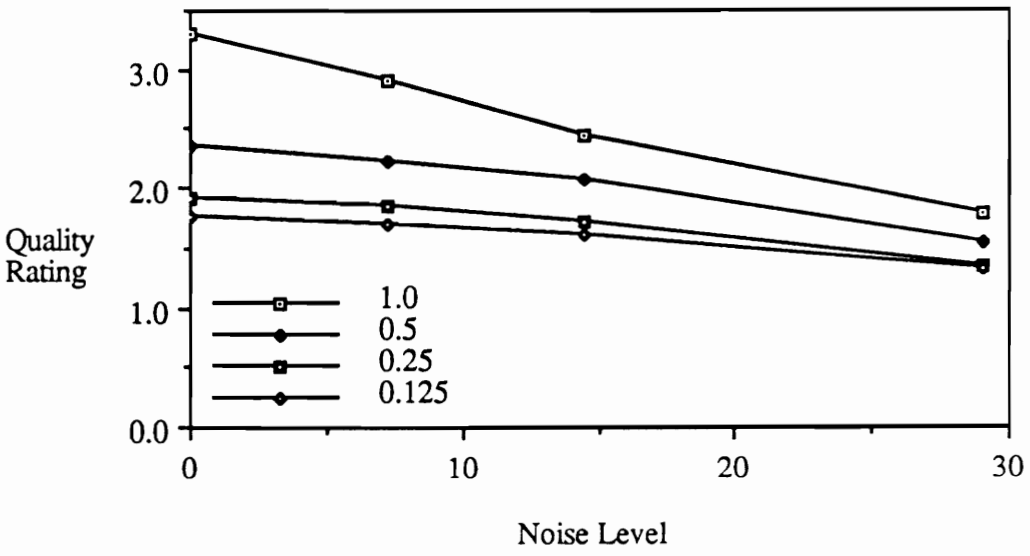


Figure 37. Image quality ratings as a function of level of attenuation and level of noise. Parameter of graph is attenuation level.

A Newman-Keuls post hoc test indicates that all of the differences among the means for the no-attenuation condition were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the lowest two conditions (noise = 0 and 7.25) when attenuation = 0.5. However, the remaining differences among these means were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the lowest two conditions (noise = 0 and 7.25) when attenuation = 0.25. Similarly, the test indicates no significant difference between the second and third conditions (noise = 7.25 and 14.5). However, the remaining differences among these means were significant ($p < 0.05$).

A Newman-Keuls post hoc test indicates no significant difference ($p > 0.05$) between the first three conditions (noise = 0, 7.25, and 14.5) when attenuation = 0.125. The mean for the highest noise condition differs significantly from the other three means ($p < 0.05$).

The three-way interaction involving the variables channel, noise level, and attenuation level is plotted in Figure 38. Comparison of the response of the achromatic channels with those of the chromatic channels reveals the stronger impact on quality of manipulations of the achromatic channels. The tendency for noise to impact unattenuated images more than attenuated images is most clearly illustrated in the responses to manipulations of the achromatic channel. For the achromatic channel a post hoc simple-effects F-test indicates that attenuation significantly degraded quality ($p < 0.01$) for the first three attenuation levels (attenuation = 1.0, 0.5, and 0.25). Noise had no effect on quality for the highest level of attenuation ($p > 0.05$). Additional F-tests indicate that noise significantly impacted quality

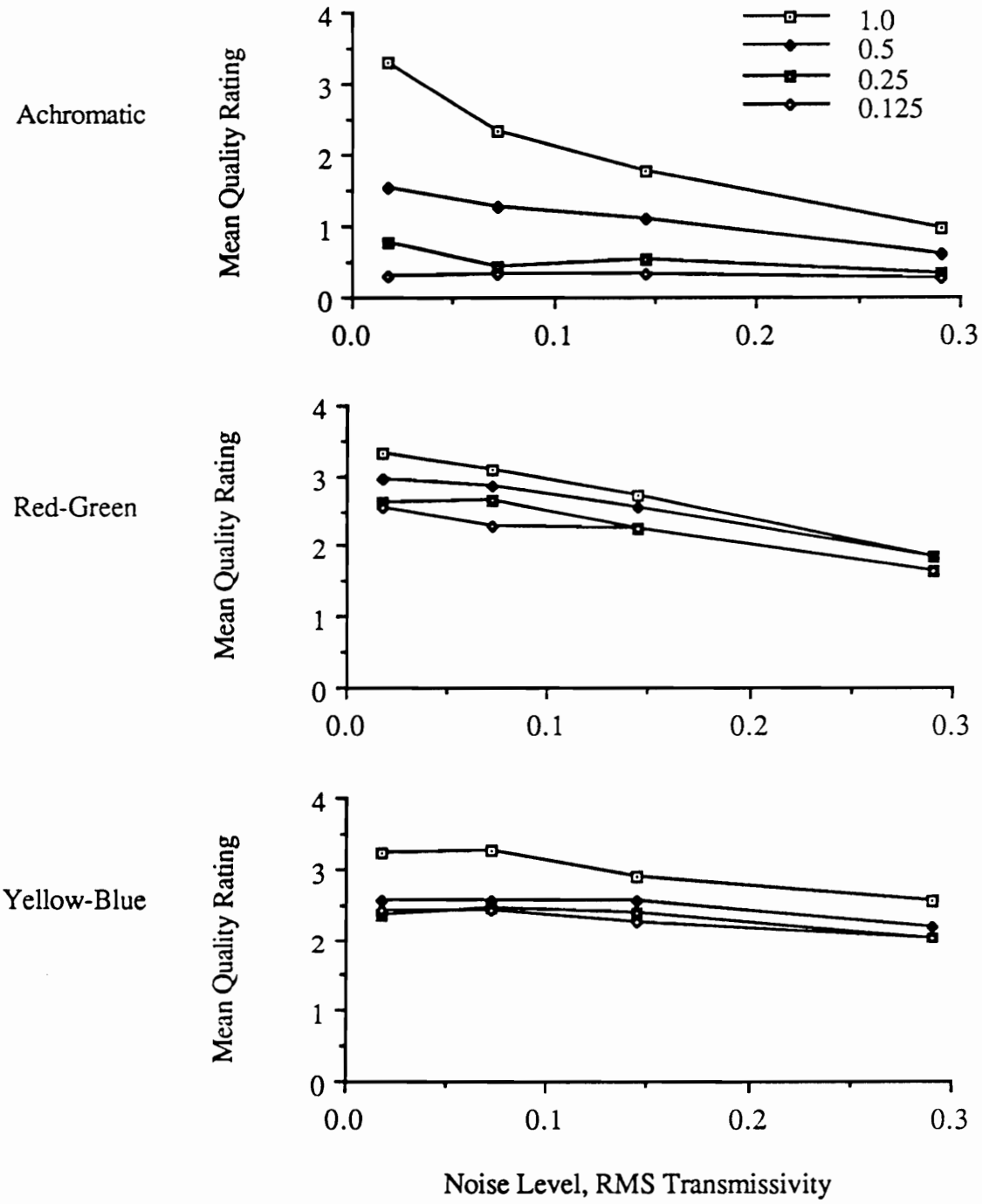


Figure 38. Image quality ratings for each channel as a function of level of attenuation and level of noise. Parameter of graph is attenuation level.

at all levels of attenuation for both the red-green and yellow-blue channels ($p < 0.05$).

A series of Newman-Keuls post hoc tests was made for each of the curves plotted in Figure 38. Results of these tests are indicated in Table 5. Groups of means which are not significantly different are placed within the same set of parentheses under the column heading "Groupings."

Discussion: Subjects. An important conclusion drawn from these quality rating data is that the achromatic channel contributes more to the perception of image quality than does either chromatic channel. This conclusion is not surprising and is likely the primary reason researchers have discounted the contribution of chromatic factors in the assessment of image quality. Though the contribution of the chromatic channels to the perception of quality is undoubtedly smaller than that of the achromatic channel, manipulations of both the red-green and the yellow-blue channels significantly influenced quality ratings in this experiment. Furthermore, quality ratings were influenced by both the attenuation of chromatic modulation and the addition of chromatic noise to either of the chromatic channels. These findings support the notion that a metric which is responsive to the transfer of chromatic modulation and the production of chromatic noise should be useful for the evaluation of color display systems. These data refute the suggestion of Granger (1974, 1983) that only the achromatic component of color images is important in the assessment of image quality.

The human performance data for that third of the trials corresponding to monochromatic manipulations (top frame of Figure 38) closely correspond with the image quality rating data collected by Beaton (1984) (see Figures 21 and 22). While the manipulations performed in these two studies are similar, there are two important differences. First, in the Beaton study monochromatic images were systematically degraded while in the present

Table 5. Results of Newman-Keuls post hoc tests for the three-way interaction among the variables channel, attenuation level, and noise level

Channel	Attenuation Level	Groupings
Achromatic	1.0	(1)* (2) (3) (4)
"	0.5	(1) (2, 3) (4)
"	0.25	(1, 3) (2, 3, 4)
"	0.125	(1, 2, 3, 4)
Red-Green	1.0	(1, 2) (3) (4)
"	0.5	(1, 2) (3) (4)
"	0.25	(1, 2) (3) (4)
"	0.125	(1, 2) (2, 3) (4)
Yellow-Blue	1.0	(1, 2) (3) (4)
"	0.5	(1, 2, 3) (4)
"	0.25	(1, 2, 3) (4)
"	0.125	(1, 2, 3) (3, 4)

* Means for conditions within each set of parentheses are not significantly different ($p > 0.05$).

study monochromatic experimental manipulations were made using *color* images. Second, Beaton changed the size of the effective blur function in order to change the modulation transfer capability of the simulated displays. Thus he used a spatial frequency dependent attenuation. In the present experiment attenuation was independent of spatial frequency. For each level of attenuation all spatial frequencies were reduced by the same amount. Thus, the monochromatic portion of the present experiment was not a simple replication of the Beaton study.

Procedure: Model. The display evaluation procedure described above was used to evaluate the same 48 display systems rated by human subjects. This evaluation was conducted in two steps. First, the ability of each display system to transmit modulation was assessed by using the display simulator to present test gratings and calculating the corresponding signal responses of the HVS model. Second, the perceptual impact of noise generated in each display condition was assessed by calculating the RMS response of the HVS model to a series of noise images representative of each condition.

Three types of sinusoidal test gratings were generated for use in evaluating displays. Each of these gratings was designed to maximally stimulate one channel of the HVS model while stimulating the other two channels as little as possible. With a linear HVS model the relative intensities of the R, G, and B primaries needed to produce isoluminance chromatically-modulated gratings can be determined with the use of a linear transformation. However, since the proposed HVS model employs a nonlinearity prior to the opponent-color stage, a linear transformation provides only an approximate solution. Starting with an approximate solution derived in this way the relative intensities of the grating components were iteratively adjusted until the gratings stimulated only the intended channels. The relative intensities resulting from this iterative process are given in Table 6.

Table 6. Relative Intensity of Red, Green, and Blue Components of Test Gratings

Channel	Red	Green	Blue
Achromatic	1.0	1.0	1.0
Red-Green	1.0	-0.3411*	0.0209
Yellow-Blue	-0.2515	0.2511	-1.0

* Negative values indicate 180 deg relative phase shift of image band.

Table 7 shows the responses of the achromatic, red-green, and yellow-blue 3 c/deg mechanisms to 3 c/deg gratings generated using the relative intensities listed in Table 6. These data indicate that responses of non-target channels were never more than 0.6% of the response of the target channels. Gratings generated using the data of Table 6 are thus suitable for independently stimulating the achromatic, red-green, and yellow-blue channels of this model.

Display evaluations were made using achromatic mechanisms with spatial center frequencies of 3, 6, 12, and 24 c/deg. Chromatic mechanisms had center frequencies of 3 and 6 c/deg. Higher spatial frequency chromatic mechanisms were not used for the chromatic channels since the contrast sensitivity data of Mullen (1985) indicate human chromatic sensitivity is reduced to half at less than 3 c/deg and is reduced to 1% of maximum sensitivity at 11-12 c/deg. The sensitivity of the 6 c/deg mechanism used here extends up to 10-12 c/deg.

Results: Model. In tuning the achromatic response of the model two parameters needed to be fixed: the level of the internal noise source and a display noise weighting parameter. In Experiment 2 these parameters were set using an estimate of contrast threshold provided by Farrell and Booth (1984), the CTF data collected by Beaton (1984), and the noise masking data of van Meeteren and Valeton (1988). With these data the model parameters were set independently of the display rating data collected by Beaton.

For the present study the levels of two additional parameters had to be determined. These parameters were a weighting factor for chromatic display noise and a parameter for setting the relative contributions of achromatic and chromatic channels. Unfortunately, these parameters could not be set independently of the display rating data due to a lack of existing data. The levels of these parameters were determined by iteratively adjusting them

Table 7. Responses of 3 c/deg achromatic, red-green, and yellow-blue mechanisms to achromatic, red-green, and yellow-blue 3 c/deg gratings

	Spatial Mechanism		
Grating	Achromatic	Red-Green	Yellow-Blue
Achromatic	51*	0.00	0.00
Red-Green	0.26	51	0.26
Yellow-Blue	0.05	0.10	51

* Responses units are JNDs

so as to maximize the statistical correlation between the predictions of the model and the subjective rating data. The results of this process indicate that model predictions correlate most highly with performance when achromatic channels are given 4.0 times the weight of each chromatic channel. This relative weighting of channels results in a best-fitting model with 4 achromatic mechanisms at each spatial frequency (3, 6, 12, and 24 c/deg), 2 red-green mechanisms (at 3, and 6 c/deg) and 2 yellow-blue mechanisms (at 3 and 6 c/deg) for a total of 20 spatial mechanisms.

The image quality predictions made using this model are plotted in Figure 39 as a function of channel, attenuation level, and noise level. Comparison of Figure 38 and 39 reveals a good correspondence between predictions and data. The statistical correlation between the predictions and data is 0.948 ($R^2 = 0.898$). The correspondence between these data is illustrated in Figure 40.

Discussion and Conclusions

Examination of Figure 39 indicates that the model somewhat underestimates the negative impact of high noise levels for the no-attenuation condition. A similar result is found in Figure 24 for the high-noise low-blur predictions made for CRT displays used by Beaton. The model also underestimates the negative impact of increasing levels of red-green noise as indicated by the shallow slope found in the center panel of Figure 39. Despite these mismatches between predictions and performance the model accounts for 90% of the variance in the data.

The data collected in this experiment provide answers to the three research questions the experiment was designed to address. Analysis of the subjective rating data clearly indicates

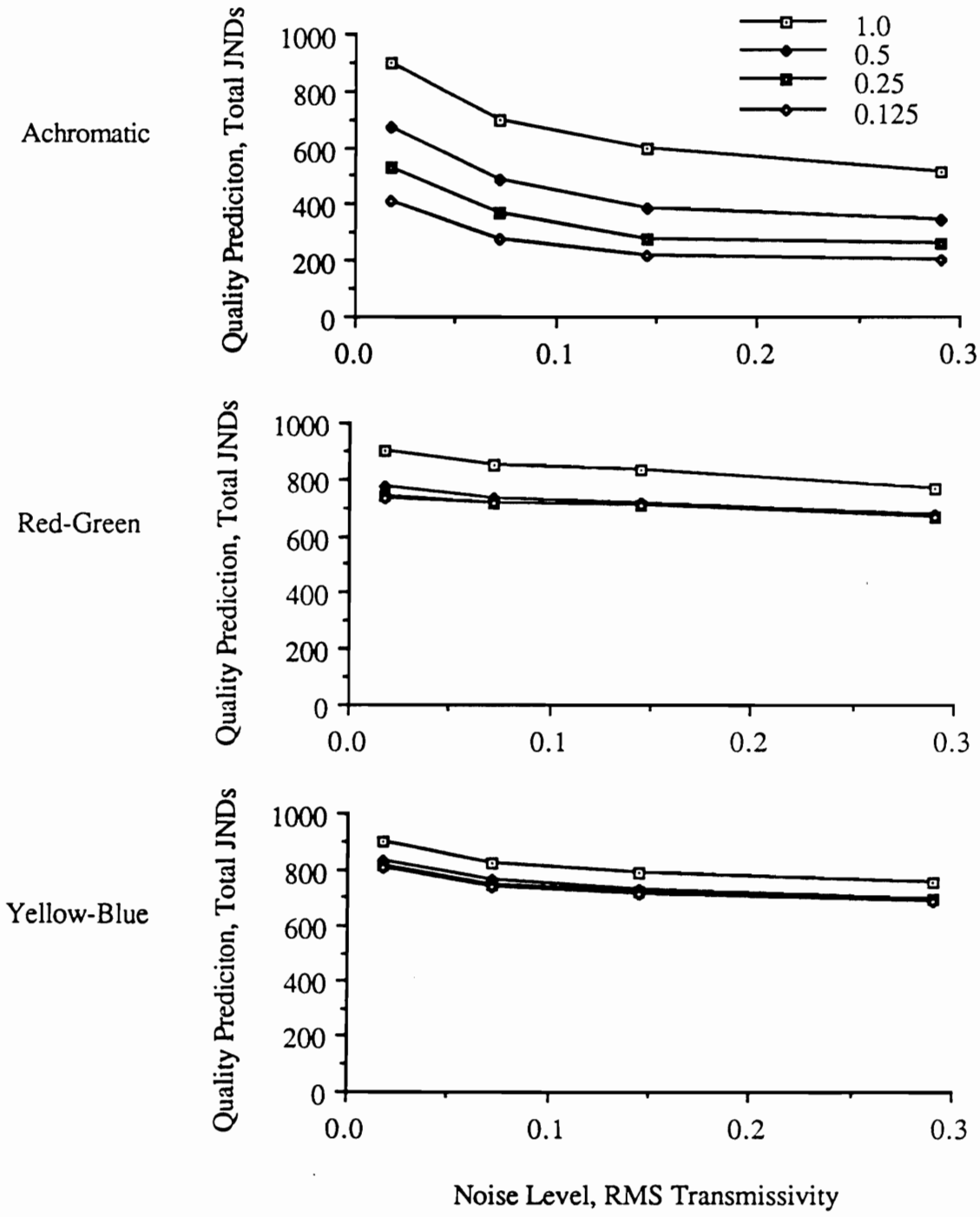


Figure 39. Image quality predictions as a function of channel, attenuation level, and noise level. Parameter of the graph is attenuation level.

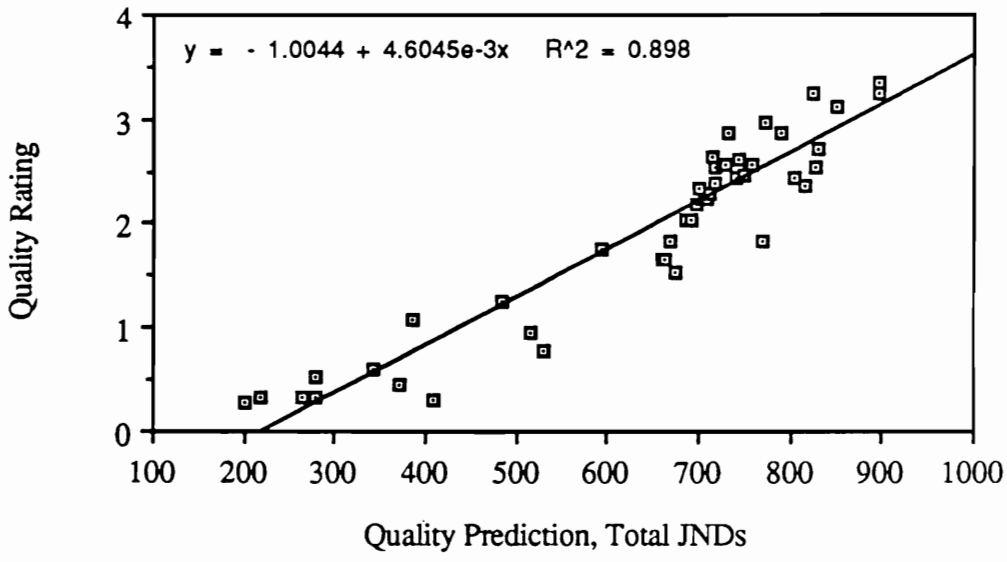


Figure 40. Correspondence between predictions and ratings of image quality for 48 color display systems.

that manipulations of both chromatic attenuation and chromatic noise significantly impact image quality. These data allowed the relative contributions of the achromatic and chromatic channels to be determined. The best-fitting model employs 16 achromatic, 2 red-green, and 2 yellow-blue channels and accounts for 90% of the variance in these data. This high correlation between predictions of the model and image quality ratings clearly supports the validity of the model.

GENERAL CONCLUSIONS

The primary goal of this dissertation was to develop and validate a method of evaluating color displays with a metric that accounts for chromatic as well as achromatic attenuation and noise. Moreover, it was considered necessary that the method apply to sampled as well as continuous displays. A valid metric with these capabilities is sorely needed to guide the many design trade-offs required in the development of new display technologies.

One display technology in particular, the CMD, is currently under intense development. This display technology employs sampled images and its development involves the trade-off of design factors which profoundly influence its spatial and chromatic characteristics. If the development of this and related display technologies is to be based on human visual capabilities, a number of engineering activities must occur. First, a valid model of human spatial-chromatic vision must be developed. The structure of this model must make it amenable to the analysis of practical display design problems. A procedure for using the model for display evaluations must be developed. Finally, both the model and the evaluation procedure must be validated. This dissertation has been an attempt to complete these activities.

At the core of the proposed display evaluation methodology is a veridical model of human spatial-chromatic visual processing. This model was adapted from current models of human spatial and chromatic vision, as well as from models of optical and retinal processes. Thus, the proposed model has high face validity. The validity of this HVS model was tested by comparing the results of contrast detection, discrimination, and magnitude estimation experiments performed on the model with the corresponding human

performance data. Results of these comparisons indicate a strong correlation between performance of the model and human performance.

An a priori constraint on the development of this system required that the method have predictive capability at least as good as extant metrics when used for evaluating monochromatic displays. Thus, comparisons of predictions of the proposed metric with subjective ratings of image quality were made for three types of monochromatic display system in Experiments 2 and 4. The correlations obtained were $R^2 = 0.970$ ($N = 25$) for predictions of photographic image quality, $R^2 = 0.792$ ($N = 25$) for predictions of CRT-based image quality, and $R^2 = 0.858$ ($N = 16$) for predictions of monochromatic perturbations of color image quality. The overall correlation for these three evaluations is $R^2 = 0.870$. In comparison, the *highest* overall correlation obtained by Beaton (1984) using 16 monochromatic metrics was $R^2 = 0.868$ for the SN metric. These results indicate the proposed metric performs at least as well as the best extant display image quality metrics when used for monochromatic display systems.

Image quality predictions were made also for two types of sampled display systems in Experiments 2 and 3. The correlation obtained between predictions and quality ratings for the monochromatic CRTs used by Beaton (1984) was $R^2 = 0.792$ ($N = 25$). Correlations between predictions and ratings of CMD quality were $R^2 = 0.98$ ($N = 3$) for the data of Silverstein et al. (1989) data and $R^2 = 0.80$ ($N = 4$) for the Rogowitz (1988) data. The overall correlation for these evaluations is $R^2 = 0.853$ ($N = 32$). Though this degree of correlation is encouraging, these results should be considered preliminary since relatively few data are available. Additionally, the CRT data do not provide a powerful test of the proposed method since the CRT was sampled only along the vertical dimension. Further experimental evaluation of the proposed method with sampled display systems is

recommended. Since CMD designers are the intended users of this system, follow-on testing should examine the effects of variables pertinent to CMD design (e.g., color element size, shape, luminance profile, and chromaticity; pixel geometry; and display addressability).

Predictions of image quality for full-color displays were made in Experiment 4. In this experiment the chromatic and achromatic attenuation and noise characteristics of 48 displays were varied. The correlation between predictions for these displays and subjective ratings of quality was $R^2 = 0.898$ ($N = 48$). Though this degree of correlation is high, caution should be exercised since these results are based on only a single study and some a posteriori setting of parameters in the model. Additional experimental evaluations of the chromatic responses of the model are recommend.

For follow-on evaluations it is recommended that the level of degradation of achromatic information be reduced so that the contributions of the chromatic variables are not overwhelmed. Additionally, the ranges of the independent variables should be adjusted so that their effects are more perceptually equivalent. Judging by the results of Experiment 4, the achromatic attenuation factor should probably be no less than 0.5, while achromatic noise should probably have a standard deviation of no greater than 15. Additionally, the levels of chromatic attenuation and noise might be more evenly distributed.

Sorely needed for progressing with this class of modeling are good estimates of human chromatic contrast discrimination curves of the type measured for achromatic stimuli. Given the relatively small contribution of chromatic manipulations to ratings of image quality, it seems that there must be relatively few JNDs on each of the chromatic channels.

This assertion cannot be tested, however, until the requisite human performance data have been collected.

Similarly, suprathreshold estimates of the magnitude of isoluminance chromatic modulation would be useful for assessing the appropriateness of the proposed HVS model. But, as with the chromatic contrast discrimination functions discussed above, these data do not currently exist.

While many advantages of the proposed methodology have been pointed out throughout this dissertation, one obvious difficulty with the approach must be discussed. The proposed HVS model and evaluation procedure are considerably more complex than any display quality metric introduced to date. This "feature" of the proposed methodology may be a disadvantage for several reasons. First, the relative complexity of the method makes it difficult to describe the method to display designers and other researchers, especially within the confines of a journal article or a conference presentation. Undoubtedly, the relative difficulty of describing (and thus understanding) the method reduces the likelihood of it being used by others.

Second, the computational complexity of the method would seem to make it difficult to use. While many of the extant metrics can be concisely described using a few equations, this concise description is quite deceptive. Use of any of the MTF-based metrics requires considerable hardware, software, and computational expertise. For example, the measurement of a line spread function (LSF) for the derivation of an MTF requires equipment such as a scanning microphotometer, luminance standard, optical table, and perhaps a positioning stage. Such equipment costs on the order of $10,000 to $100,000 and requires considerable training to operate. Additionally, calculation of the MTF, once

the LSF has been obtained, requires at least a personal computer and a person familiar with linear systems analysis techniques and the Fast Fourier Transform (FFT). Though extant image quality metrics certainly can be described more concisely than the proposed metric, their practical use is by no means a trivial proposition.

It seems likely that in the long run the complexity of the present method will be of minor importance. When an appropriate amount of testing of the method has been successfully completed and the procedures and algorithms have stabilized, the HVS model and associated procedures could be embodied in a single software program. Thus, much of the complexity of the proposed method could be made transparent to the user.

This dissertation demonstrates the feasibility and utility of employing a highly veridical model of the HVS for the evaluation of display systems. Moreover, several important advantages of the proposed method over extant metrics of image quality have been demonstrated. The proposed method for evaluating display systems looks promising. However, the strengths and weaknesses of the approach will be revealed only through additional testing. The outcome of these tests will likely determine the future of this line of research.

REFERENCES

Abramson, S. R., and Snyder, H. L. (1984). *Operator performance on flat-panel displays with line and cell failures* (Tech. Report HFL-83-3). Blacksburg, VA: Virginia Polytechnic Institute and State University, Human Factors Laboratory.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The oblique effect in man and animals. *Psychological Bulletin, 78*, 266-278.

Barten, P. G. J. (1987). The SQRI method: A new method for the evaluation of visible resolution on a display. In *Proceedings of the SID* (pp. 253-262). Playa del Rey, CA: Society for Information Display.

Barten, P. G. J. (1988). Effect of convergence errors on resolution. *In Proceedings of the SID* (pp. 3-5). Playa del Rey, CA: Society for Information Display.

Barten, P. G. J. (1989a). The effects of picture size and definition on perceived image quality. *IEEE Transactions on Electron Devices, 36* (9), 1865-1869.

Barten, P. G. J. (1989b). Evaluation of CRT displays with the SQRI method. In *Proceedings of the SID* (pp. 9-14). Playa del Rey, CA: Society for Information Display.

Baur, B. (1985). An advanced CRT-spot-contour measurment system. In *Proceedings of the SID* (pp. 55-57). Playa del Rey, CA: Society for Information Display.

Beamon, W. S., and Snyder, H. L. (1975). *An experimental evaluation of the spot wobble method of supressing raster structure visibility* (Tech. Report AMRL-TR-75-63). Dayton, Ohio: Air Force Aerospace Medical Research Laboratory.

Beaton, R. J. (1984). A human-performance based evaluation of quality metrics for hard-copy and soft-copy digital imaging systems. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Beaton, R. J. (1988). Linear systems metrics of image quality for flat-panel displays. In *Proceedings of the SPIE: Image Processing, Analysis, Measurement, and Quality: Vol 901*. (pp. 144-151). Los Angeles, CA. January, 1988.

Benzschawel, T., and Guth, S. L. (1984). ATDN: Toward a uniform color space. *Color Research and Application, 9* (3), 133-141.

Biberman, L. M. (1973). *Perception of displayed information.* New York: Plenum.

Blakemore, C., and Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203* (1), 237-261.

Blumenthal, A. H., and Campana, S. B. (1981). An improved electro-optical image quality summary measure. In *Proceedings of the Society of Photographic Instrumentation Engineers, Image Quality, 310.* (pp. 43-52). San Deigo, CA, August, 1981.

Borough, H. C.; Fallis, R. F.; Warnock, R. H., and Britt, J. H. (1967). *Quantitative determination of image quality* (Tech. Report D2-114058-1). Seattle, WA: Boeing.

Bortfeld, D. P., and Beltz, J. P. (1987). Semi-automatic measurement of the two-dimensional intensity distribution of a color CRT electron-beam spot. *SID International Symposium Digest of Technical Papers, XVIII*, 225-227.

Boynton, R. M. (1983). A system of photometry and colorimetry based on cone excitations: Relation to visual mechanisms. In *Proceedings of the CIE* (pp. B1/1 - B1/8). Paris: Bureau Central de la CIE.

Bracewell, R. N. (1986). *The Fourier transform and its applications* (2nd ed.). New York, NY: McGraw-Hill.

Bradley, A., and Ohzawa, I. (1986). A comparison of contrast detection and discrimination. *Vision Research, 26* (6), 991-997.

Bumbaca, F., and Smith, K. C. (1987). Design and implementation of a colour vision model for computer vision applications. *Computer Vision, Graphics, and Image Processing, 39,* 226-245.

Burton, G. J. (1981). Contrast discrimination by the human visual system. *Biological Cybernetics, 40,* 27-38.

Campbell, F. W., and Gubisch, R. W. (1966). Optical quality of the human eye. *Journal of Physiology, 186,* 558-578.

Campbell, F. W., Kulikowski, J. J., and Levinson, J. (1966). The effect of orientation on the visual resolution of gratings. *Journal of Physiology, 187,* 427-436.

Campbell, F. W., and Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, 197,* 551-566.

Cannon, M. W. (1985). Perceived contrast in the fovea and periphery. *Journal of the Optical Society of America, 2,* 1760-1768.

Cannon, M. W., and Fullenkamp, S. C. (1988). Perceived contrast and stimulus size: experiment and simulation. *Vision Research, 28* (6), 695-709.

Carlson, C. R., and Cohen, R. W. (1978). *Visibility of displayed information: Image descriptors for displays.* (Tech. Report ONR-CR213-120-4F). Princeton NJ: RCA Laboratories.

Carlson, C. R., and Cohen, R. W. (1980). A simple psychophysical model for predicting the visibility of displayed information. In *Proceedings of the SID* (pp. 229-246). Playa del Rey, CA: Society for Information Display.

Charman, W. N., and Olin, A. (1965). Tutorial: Image quality criteria for aerial camera systems. *Photographic Science and Engineering, 9,* 385-397.

Cohen, R. W., and Gorog, I. (1974). Visual capacity - an image quality descriptor for display evaluation. In *Proceedings of the SID* (pp. 52-62). Playa del Rey, CA: Society for Information Display.

Daugman, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research, 20,* 847-856.

Daugman, J. G. (1983). Six formal properties of two-dimensional anisotropic visual filters: Structural principles and frequency/orientation selectivity. *IEEE Transactions On Systems, Man, and Cybernetics, SMC-13* (5), 882-887.

Davidson, M. (1968). Perturbation approach to spatial brightness interaction in human vision. *Journal of the Optical Society of America, 58* (9), 1300-1308.

Decker, J. J., Dye, C. J., Kurokawa, K., and Lloyd, C. J. C. (1988). Effects of display failures and symbol rotation on visual search using dot-matrix symbols. In *Proceedings of the Human Factors Society 32nd Annual Meeting.* (pp. 1386-1390). Santa Monica, CA: Human Factors Society.

Decker, J. J.; Pigion, R. D., and Snyder H. L. (1987). *A literature review and experimental plan for research on the display of information on matrix-addressable displays* (Tech. Report 4-87). Aberdeen, MD: U.S. Army Human Engineering Laboratory.

Enroth-Cugell and Robson. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology, 187.* 517-552.

Farrell, R. J., and Booth, J. M. (1984). *Design handbook for imagery interpretation equipment* (Tech. Report D180-19063-1). Seattle, Washington: Boeing Aerospace Company.

Faugeras, O. D. (1976). Digital color image processing and psychophysics within the framework of a human visual model. Unpublished doctoral dissertation, University of Utah, Salt Lake City, Utah.

Faugeras, O. D. (1979). Digital color image processing within the framework of a human visual model. *IEEE Transactions On Acoustics, Speech, and Signal Processing, 27* (4), 380-393.

Flitcroft, D. I. (1989). The interactions between chromatic aberration, defocus, and stimulus chromaticity: Implications for visual physiology and colorimetry. *Vision Research, 29* (3), 349-360.

Friedman, P. S., Rahman, A., Peters, E. F., Wedding, C. A., and Repetti, A. A. (1989). 17-in. ac plasma RGB video monitor with demonstration of high-luminosity capability. *SID International Symposium Digest of Technical Papers, XX,* 344-346.

Glenn, K. G., and Glenn, W. E. (1987). From vision science to HDTV: Bridging the gap. *Information Display, 2,* 22-25.

Glenn, W. E., Glenn, K. G., and Bastian, C. J. (1985). Imaging system design based on psychophysical data. In *Proceedings of the SID* (pp. 71-78). Playa del Rey, CA: Society for Information Display.

Graham, N., and Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11,* 251-259.

Granger, E. M. (1974a). Specification of color image quality. Unpublished doctoral dissertation, The University of Rochester, Rochester, NY.

Granger, E. M. (1974b). Subjective assessment and specification of color image quality. In *Proceedings of the Society of Photo-Optical Instrumentation Engineers, 46,* 86-92.

Granger, E. M. (1983). A summary measure of image quality. In *Proceedings of the SPIE*. August, 1983.

Granger, E. M., and Heurtley, J. C. (1973). Visual chromaticity-modulation transfer function. *Journal of the Optical Society of America, 63* (9), 1173-1174.

Gutmann, J. C. (1982). *An experimental determination of the effects of image quality on eye movements and search for static and dynamic displays.* Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Gutmann, J. C.; Snyder, H. L.; Farley, W. W., and Evans, J. E. (1979). *An experimental determination of the effect of image quality on eye movements and search for static and dynamic targets* (Tech. Report AMRL-TR-79-51). Dayton, OH: Air Force Aerospace Medical Research Laboratory.

Hall, C. F., and Hall E. L. (1977). A nonlinear model for the spatial characteristics of the human visual system. *IEEE Transactions On Systems, Man, and Cybernetics, 7* (3), 161-170.

Henning, G. B., Hertz B. G., and Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research, 15*, 887-897.

HFS. (1988). *American National Standard for human factors engineering of visual display terminal workstations.* Santa Monica, CA: The Human Factors Society.

Hufnagle, R. E. (1965). A search for a summary measure of image quality. Part II. An invited paper presented at the Annual Meeting of the Optical Society of America, Philadelphia.

Hurvich, L. M., and Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review, 64* (6), 384-404.

Ichikawa, K., Suzuki, S., Matino, H., Aoki, T., Higuchi, T., and Oana, Y. (1989). 14.3-in.-diagonal 16-color TFT-LCD panel using a Si:H TFTs. *SID International Symposium Digest of Technical Papers, XX,* 226-229.

Infante, C. (1986). CRT systems. In *Proceedings of the SID: Seminar Lecture Notes* (pp. 8.1.2 - 8.1.126). Playa del Rey, CA: Society for Information Display.

Keller, P., and Beaton, R. J. (1989). The EIA standard for MTFs of monochromatic CRTs. *SID International Symposium Digest of Technical Papers, XX,* 204-207.

Kelly, D. H. (1966). Frequency doubling in visual responses. *Journal of the Optical Society of America, 56* (11), 1628-1633.

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69* (10), 1340-1349.

Kelly, D. H. (1989). Spatial and temporal interactions in color vision. *Journal of Imaging Technology, 15* (2), 82-89.

Kobayashi, S. (1983). Color LCDs: Colorimetric and ergonomic evaluation and optimization. In *Japan Display* (pp. 198-201).

Koenderink, J. J., and vanDoorn, A. J. (1979). Spatiotemporal contrast detection threshold surface is bimodal. *Optics Letters, 4* (1), 32-34.

Legge, G. E. (1981). A power law for contrast discrimination. *Vision Research, 21,* 457-467.

Legge, G. E., and Foley J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70* (12), 1458-1471.

Livingstone, M. S., and Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *The Journal of Neuroscience, 4* (1), 309-356.

Lloyd, C. J. C.; Beaton, R. J.; Parsley, M., and Simonetti, T. (1990). *Contrast detection and discrimination thresholds for chromatic and achromatic gratings* (Technical Report). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Lloyd, C. J. C., Decker, J. J., Kurokawa, K., and Snyder, H. L. (1988). Effects of line and cell failures on reading and search performance using matrix-addressable displays. In *Proceedings of the SID* (pp. 344-347). Playa del Rey, CA: Society for Information Display.

Maattanen, L. M., Koenderink, J. J., and Nienhuis, B. (1988). Contrast discrimination: Invariant to spatial parameters. *Vision Research, 28* (7), 811-818.

Mannos, J. L., and Sakrison, D. J. (1974). The effects of a visual fidelity criterion on the encoding of images. *IEEE Transactions On Information Theory, 20* (4), 525-536.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information* . New York: Freeman.

Michael, C. R. (1978). Color vision mechanisms in monkey striate cortex: Dual-opponent cells with concentric receptive fields. *Journal of Neurophysiology, 41* (3), 572-588.

Michael, C. R. (1983). Color processing in primate striate cortex. In J. D. Mollon, and L. T. Sharpe (Eds.), *Color vision: Physiology and psychophysics* (pp. 277-290). New York: Academic.

Moriyama, H., Uchida, H., Nishida, S., Nakano, H., Hirai, Y., Kaneko, S., and Tani, C. (1989). 12-in. full-color a-Si:H TFT-LCD with pixel electrode buried in gate insulator. *SID International Symposium Digest of Technical Papers, XX,* 144-147.

Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology, 359,* 381-400.

Murch, G. M., and Beaton, R. J. (1988). Matching display resolution and addressability to human visual capacity. *Displays,* Jan, 23-26.

Murch, G. M., Virgin, L., and Beaton, R. J. (1985). Resolution and addressability: How much is enough? In *Proceedings of the SID* (pp. 305-308). Playa del Rey, CA: Society for Information Display.

Nachmias, J., and Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039-1042.

Oakley, D. (1984). Raster graphics display technology: A review. *Displays,* Oct, 229-234.

O'Callaghan, J. P., and Veron, H. (1987). Parametric MTF analysis for shadow-mask CRT monitor configurations. *SID International Symposium Digest of Technical Papers, XVIII,* 214-217.

Pantel, A., and Sekuler, R. (1968). Size-detecting mechanisms in human vision. *Science, 162,* 1146-1148.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America, 2* (9), 1508-1532.

Pica, A. P., and Klopfenstein, R. W. (1987). Perceptual analysis and simulation of color monitor displays. *Journal of Imaging Technology, 13,* 163-165.

Post, D. L., Costanza, E. B., and Lippert, T. M. (1982). Expressions of color contrast as equivalent achromatic contrast. In *Proceedings of the Human Factors Society 26th Annual Meeting* (pp. 581-585). Santa Monica, CA: Human Factors Society.

Quick, R. F. (1974). A vector-magnitude model of contrast detection. *Kybernetik, 16,* 65-67.

Reger, J. J., Snyder, H. L., and Farley, W. W. (1989). Legibility of emissive and non-emissive flat-panel displays under fluorescent and daylight illumination. *SID International Symposium Digest of Technical Papers, XX,* 378-381.

Robson, J. G. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America, 56,* 1141-1142.

Rogowitz, B. E. (1988). The psychophysics of spatial sampling. In *Proceedings of the SPIE: Image processing, analysis, measurement, and quality: Vol 901.* (130-138). Los Angeles, CA. January, 1988.

Sachs, M. B., Nachmias, J., and Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the Optical Society of America, 61* (9), 1176-1186.

Saghri, J. A., Cheatham, P. S., and Habibi, A. (1989). Image quality measure based on a human visual system model. *Optical Engineering, 28* (7), 813-818.

Schade, O. H. (1958). On the quality of color-television images and the perception of color detail. *Journal of the Society of Motion Picture and Television Engineers, 67* (12), 801-819.

Schreiber, W. F. (1984). Psychophysics and the improvement of TV image quality. *Television Image Quality:SMPTE,* 36-53.

Schreiber, W. F. (1986). *Fundamentals of electronic imaging systems.* New York: Springer-Verlag.

Shibusawa, M., Kondo, J., Kigoshi, M., Inoue, K., Yoshifuku, T., Shimizu, K., and Hori, H. (1989). A 212-line/in. full-color TFT-LCD for cockpit use. *SID International Symposium Digest of Technical Papers, XX,* 230-233.

Silverstein, L. D., and Lepkowski, J. S. (1986). The perception of color primary spatial distribution in color information displays. *SID International Symposium Digest of Technical Papers, XVII,* 416-419.

Silverstein, L. D., Monty, R. W., Gomer, F. E., and Yeh, Y. (1989). A psychophysical evaluation of pixel mosaics and gray-scale requirements for color matrix displays. *SID International Symposium Digest of Technical Papers, XX,* 128-131.

Smith, V. C., and Pokorny, J. (1972). Spectral sensitivity of color-blind observers and the cone photopigments. *Vision Research, 12,* 2059-2071.

Snyder, H. L. (1973). Image quality and observer performance. In L. M. Biberman (Ed.), *Perception of displayed information* (pp. 87-118). New York: Plenum.

Snyder, H. L. (1974). Image quality and face recognition on a television display. *Human Factors, 16,* 300-307.

Snyder, H. L. (1976). *Visual search and image quality: Final report* (Tech. Report AMRL-TR-76-89). Dayton, OH: Air Force Aerospace Medical Research Laboratory.

Snyder, H. L. (1980). *Human visual performance and flat panel display image quality* (Tech. Report HFL-80-1). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Snyder, H. L. (1985). The visual system: Capabilities and limitations. In L. E. Tannas (Ed.), *Flat-panel displays and CRTs* (pp. 54-69). New York, NY: Van Nostrand Reinhold.

Snyder, H. L. (1988). Toward the determination of electronic display image quality. In W. Rouse (Ed.) *Advances in man-machine systems research, Vol. 4* (pp. 54-69). JAI Press.

Snyder, H. L., Keesee, R. L., Beamon, W. S., and Aschenbach, J. R. (1974). *Visual search and image quality* . (Tech. Report AMRL-TR-73-114). Dayton, OH: Air Force Aerospace Medical Research Laboratory.

Snyder, H. L., and Maddox, M. E. (1987). *Information transfer from computer-generated dot-matrix displays.* (Tech. Report HFL-78-3). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Stockham, T. G. (1972). Image processing in the context of a visual model. *Proceedings of the IEEE, 60* (7), 828-842.

Swanson, W. H., Wilson, H. R., and Giese, S. C. (1984). Contrast matching data predicted from contrast increment thresholds. *Vision Research, 24* (1), 63-75.

Task, H. L. (1979). *An evaluation and comparison of several measures of image quality for television displays.* (Tech. Report AMRL-TR-79-7). Dayton, OH: Air Force Aerospace Medical Research Laboratory.

Valberg, A., Lee, B. B., and Tryti, J. (1987). Simulation of responses of spectrally-opponent neurones in the macaque lateral geniculate nucleus to chromatic and achromatic light stimuli. *Vision Research, 27* (6), 867-882.

van der Horst, G. J. C. (1969). Fourier analysis and color discrimination. *Journal of the Optical Society of America, 59* (12), 1670-1676.

van der Horst, G. J. C., and Bouman, M. A. (1969). Spatiotemporal chromaticity discrimination. *Vision Research, 59* (11), 1482-1488.

van der Horst, G. J. C., deWeert, C. M. M., and Bouman, M. A. (1967). Transfer of spatial chromaticity-contrast at threshold in the human eye. *Journal of the Optical Society of America, 57* (10), 1260-1266.

van Meeteren, A., and Valeton, J. M. (1988). Effects of pictorial noise interfering with visual detection. *Journal of the Optical Society of America, 5* (3), 438-444.

Watson, A. B. (1983). *Detection and recognition of simple spatial forms.* (Tech. Report 84353). Moffett Field, CA: NASA Ames Research Center.

Watson, A. B. (1987a). Efficiency of a model human image code. *Journal of the Optical Society of America, 4* (12), 2401-2417.

Watson, A. B. (1987b). The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics, and Image Processing, 39,* 311-327.

Watson, A. B., and Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research, 21,* (7), 1115-1122.

Wendland, B. (1983). Extended definition television with high picture quality. *Journal of the Society of Motion Picture and Television Engineers, 92,* 1028-1035.

Westheimer, G., and Campbell, F. W. (1962). Light distribution in the image formed by the living human eye. *Journal of the Optical Society of America, 52* (9), 1040-1045.

Williams, L. G., and Erickson, J. M. (1974). *Contrast sensitivity as a function of spatial and temporal frequency, luminance, and stimulus position on the retina.* (Tech. Report F0259-IRI). Minneapolis: Honeywell.

Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics, 38,* 171-178.

Wilson, H. R., and Bergen, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research, 19,* 19-32.

Wilson, H. R., and Gelb, D. J. (1984). Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America, 1* (1), 124-131.

Wilson, H. R., and Giese, S. (1977). Threshold visibility of frequency gradient patterns. *Vision Research, 17,* 1177-1190.

Wyszecki, G., and Stiles, W. S. (1982). *Color Science: Concepts and methods, quantitative data and formulae.* New York: Wiley.

Xie, Z., and Stockham, T. G. (1989). Toward the unification of three visual laws and two visual models in brightness perception. *IEEE Transactions On Systems, Man, and Cybernetics, 19* (2), 379-387.

Yamauchi, S., Aizawa, M., Clerc, J. F., Uchida, T., and Duchene, D. (1989). Homeotropic-alignment full-color LCD. *SID International Symposium Digest of Technical Papers, XX,* 378-381.

Young, R. A. (1987). The Gaussian derivative model for spatial vision: I. Retinal mechanisms. *Spatial Vision, 2* (4), 273-293.

Young, R. A. (1986). *The Gaussian derivative model for machine vision: Visual cortex simulation* . (Tech. Report GMR-5323). Warren, MI: General Motors Research Laboratories.

APPENDIX A. Instructions for Participants

In this experiment you will be shown a series of approximately 50 images on this display system.

(Experimenter points out display system on which is displayed a sample image)

You will be allowed to view each of these images for up to 30 seconds. Your task is to assign a number to each image based on your impression of the quality of the image; that is, how well the image represents the original scene. High numbers should be assigned to good quality images while low numbers should correspond with poor quality images.

The range of the numbering system you use for making your ratings is entirely up to you. You should choose a numbering scheme with which you are comfortable. You may use whole numbers, fractions, or decimals. Additionally, you may use as many steps in your rating scale as you wish. Your primary goal is to try and make the number you assign match your perception of the quality of the image.

We will begin this experiment with a series of 8 practice trials. The purpose of these practice trials is to familiarize you with the equipment and task and to allow you to stabilize your rating scheme.

It takes the computer a certain amount of time to generate each image. Thus we will have a waiting period of approximately 20 seconds between each trial. The entire experiment should take approximately 50 minutes to complete.

Are there any questions?

(Present practice trials)

Are there any questions?

Once the regular trials have begun, please try to use your rating scheme as consistently as you can.

APPENDIX B: Derivation of the RTF

Estimation of Human Contrast Discrimination Performance

One of the primary purposes for using a nonlinear retinal transducer function (RTF) was to allow the HVS model to reflect nonlinear suprathreshold visual sensitivity without requiring the use of a post-spatial mechanism nonlinearity. To insure that the model reflects human contrast discrimination performance, a contrast discrimination function (CDF) was used to drive the derivation of the RTF. This CDF was estimated from the results of several recent studies of contrast discrimination performance. This section discusses the estimation of the CDF. The following section describes the procedure used to derive the RTF from the estimated CDF.

The results of numerous measurements of human contrast discrimination performance have been published in the visual science literature. The data from six of these studies are plotted in Figure 41 (Bradley and Ohzawa, 1986; Burton, 1981; Legge and Foley, 1980; Lloyd, Beaton, Parsley, and Simonetti, 1990; Nachmias and Sansbury, 1974; Pelli, 1985). Examination of Figure 41 indicates considerable variance among these data. An explanation for the differences among these data has not been given in the literature. Since there is no compelling reason to expect one data set to be more appropriate than the others, the current estimate of the CDF was based on all six data sets.

The dipper portion of the CDF was estimated by normalizing each of the six data sets to a common threshold value, plotting the data on a common graph, and hand-fitting a smooth curve through the data in the region of the dipper. The results of this procedure are

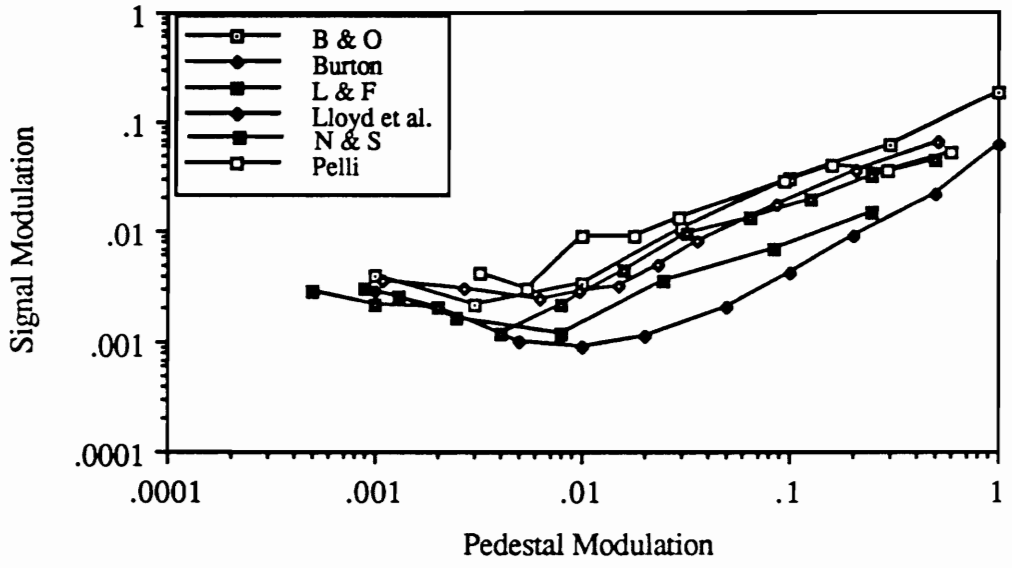


Figure Key

B & O	Bradley & Ohzawa, 1986, Figure 3-a.
Burton	Burton, 1981, Figure 3-c.
L & F	Legge and Foley, 1980, Figure 6.
Lloyd et. al.	Lloyd, Beaton, Parsley, and Simonetti, 1990
N & S	Nachmias and Sansbury, 1974, Figure 1
Pelli	Pelli, 1985, Figure 11-a

Figure 41. Plots of the six data sets used in determining an analytical function for contrast discrimination performance.

presented in Figure 42. The threshold value of 0.003 was selected since this value is near the peak of sensitivity for the HVS (occurring at approximately 3.0 c/deg).

The suprathreshold portion of the CDF was estimated by transforming the data to log-log coordinates, plotting the data (for pedestal modulations above 0.01) on a common graph, and fitting a least-squares regression line. The suprathreshold data and resulting regression equation are given in Figure 43. Since this regression equation was fitted in the log transform domain, the equation corresponds with a power function with an exponent of 0.77 in the linear domain. The subthreshold and suprathreshold portions of the CDF from Figures 42 and 43 are plotted together in Figure 44. Figure 44 thus provides an estimate of the complete CDF.

In the HVS model used here the threshold level of detection is determined by internal noise. Moreover, the characteristic dip in the CDF occurs as a result of probability summation. Thus, with the current model, the subthreshold and dipper portions of the CDF are not mediated by the form of the RTF. For this reason, the sub-threshold and dipper portions of the CDF were not used to drive the derivation of the RTF. It was thus necessary to estimate the CDF that would result if the effect of threshold were removed (e.g., as if internal noise were reduced to zero). An estimate of the CDF for the no-threshold condition was made by extrapolating the suprathreshold portion of the CDF to subthreshold levels using the regression equation of Figure 43.

Experimentation with the RTF algorithm described below indicated that it was unreasonable to expect that the CDF would continue to decline for very low pedestal modulations (e.g., modulation < 0.001). If the CDF were to decrease indefinitely, the slope of the center portion of the resulting RTF would approach infinity. To avoid this

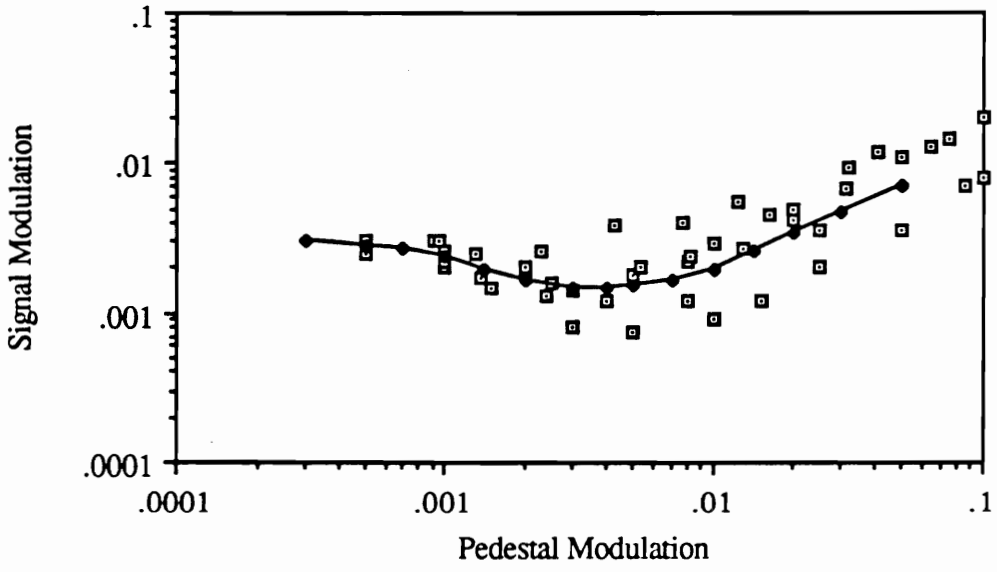


Figure 42. Contrast discrimination data of Figure 41 for pedestal modulations below 0.1. Smooth curve has been fit to the data by hand.

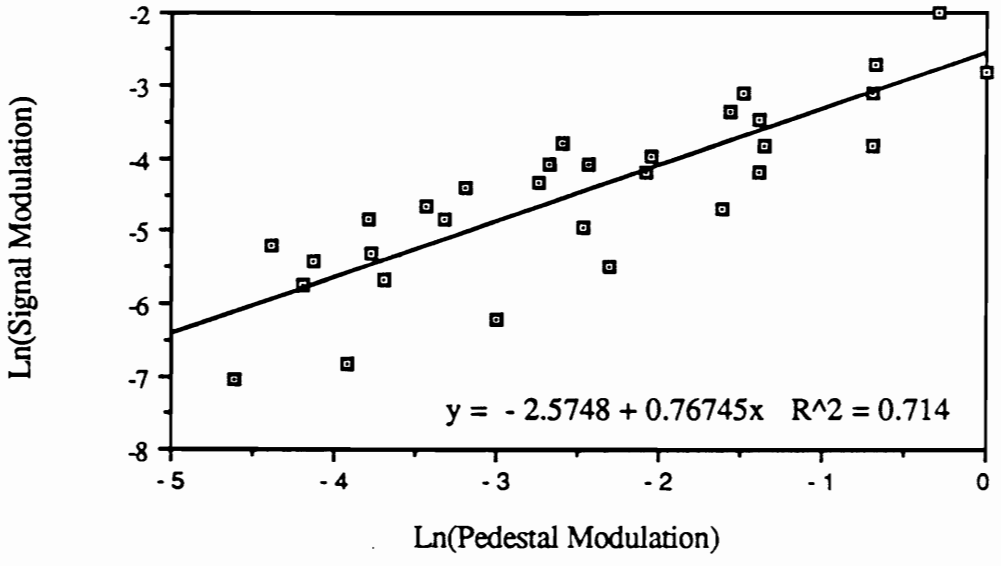


Figure 43. Contrast discrimination data of Figure 41 for suprathreshold modulations (modulation > 0.01).

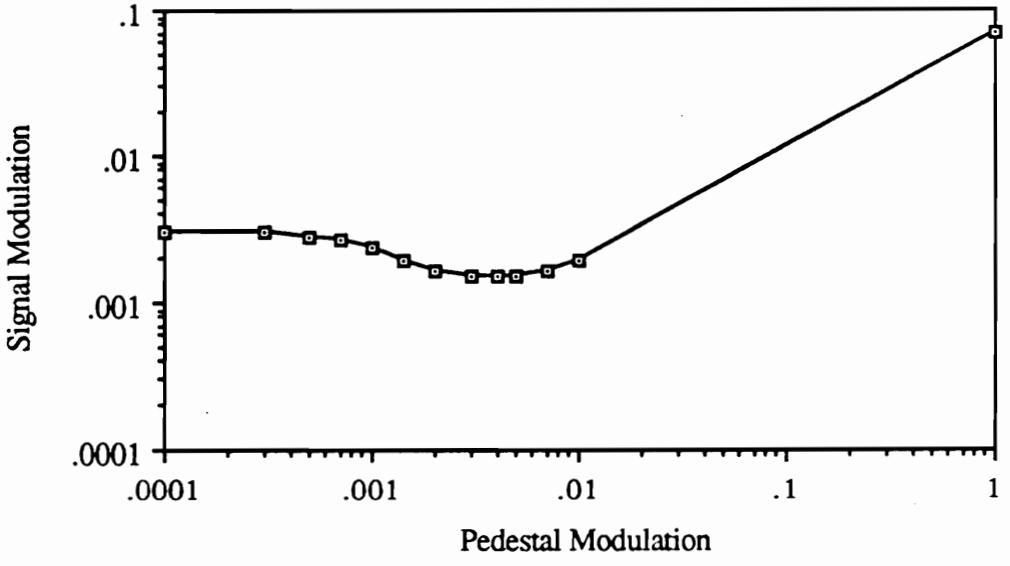


Figure 44. Estimated CDF for complete range of modulation.

physiologically implausible result, a lower limit of 0.0005 was used for the CDF. Since this limit is much lower than the lowest possible detection threshold, use of the limit is expected to have little effect on the performance of the complete model. This expectation is supported by the results of model validation studies (see pp. 75-78). The resulting estimate of the CDF used in the derivation of the RTF is given in Equation 24. This equation is plotted in Figure 45.

$$\text{Signal Modulation} = 0.07 * (\text{Pedestal Modulation})^{0.78} + 0.0005 \quad (24)$$

Generation of the RTF

The RTF was iteratively derived as a piece-wise linear function. Derivation of the RTF was constrained by two primary factors: the structure of the remainder of the HVS model (see pp. 38-59), and the estimate of the CDF derived above. Briefly, the input range of the RTF was divided into a large number of intervals. The slope of the RTF in each of these intervals was then iteratively adjusted so that a JND in modulation (determined from the CDF) at the input to the model produced a unit change in the output of the model. The overall procedure for generating the RTF is given in pseudo code in Figure 46.

The range of modulations covered by the RTF varies from 0.0 to 1.0. This range was divided into 512 intervals (modulation steps of 0.00195). To simplify this discussion, however, it is assumed that only 10 intervals were used (modulation steps of 0.1). Starting at the lowest level of modulation, the RTF was derived one interval at a time. For the first interval, two grating patterns were generated. Grating "A" was generated using a pedestal level of modulation of 0.1. Grating "B" was generated using a pedestal - signal level of modulation where the signal level was determined using the CDF. Each grating was then

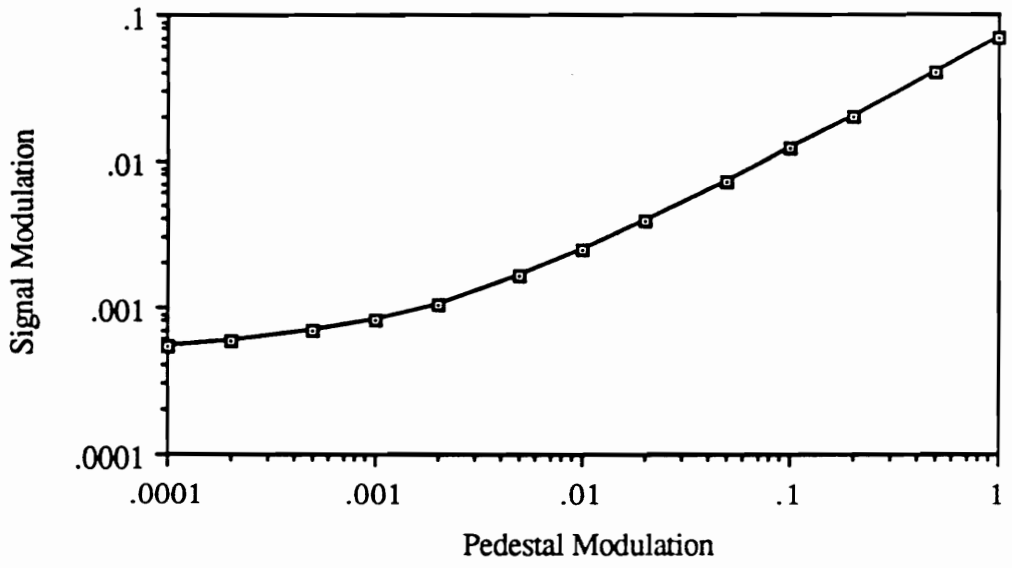


Figure 45. Estimated CDF for the case of zero internal noise.

```
For PedestalModulation from Low to High  Do
Begin

    SignalModulation := 0.07 * PedestalModulation^0.78 + 0.0005

    Generate grating A at PedestalModulation
    Generate grating B at PedestalModulation - SignalModulation

    Convolve grating A with optical LSF
    Convolve grating B with optical LSF

    While Abs(DeltaResponse - 1.0) > Precision  Do
    Begin

        Compress grating A with RTF
        Compress grating B with RTF

        ResponseA :=  spatial mechanism response to compressed grating A
        ResponseB :=  spatial mechanism response to compressed grating B

        DeltaResponse := ResponseA - ResponseB

        If DeltaResponse < 1.0  Then
            Increase slope of current segment of RTF
        Else
            Decrease slope of current segment of RTF

    End;
End;
```

Figure 46. RTF derivation algorithm expressed in pseudo code.

blurred by convolving with the optical line spread function of the eye.

Within the inner loop of Figure 46 the blurred versions of gratings A and B were "compressed" using the RTF. For the lowest level of modulation, however, only the first (positive and negative) intervals of the RTF were used. The responses of the spatial mechanisms to the RTF-processed gratings were then calculated and differenced. If the response to grating A minus the response to grating B was less than one, the slope of the first segment of the RTF was increased and the operations within the inner loop repeated. If the difference in responses was greater than one the slope of the RTF was decreased and the operations within the inner loop repeated. The slope of the first segment of the RTF was iteratively adjusted in this manner until a difference in response of 1.0 (plus or minus the value of the precision variable) at the output of the model was achieved. The RTF obtained for the first interval is illustrated in Figure 47. Note that due to symmetry, two segments of the RTF were generated for each modulation interval.

For the second interval two gratings were again used. Grating "A" was generated using a pedestal level of modulation of 0.2. Grating "B" was generated using a pedestal - signal level of modulation where the signal level was determined using the CDF. These gratings were then blurred by convolving with the optical line spread function of the eye. The same iterative algorithm was used to adjust the slope of the second two segments of the RTF such that a unit difference in the response of the model was obtained. The RTF obtained after completing the second step of this process is illustrated in Figure 48. Note that two segments were added to the RTF. Due to symmetry the slopes of these segments are the same.

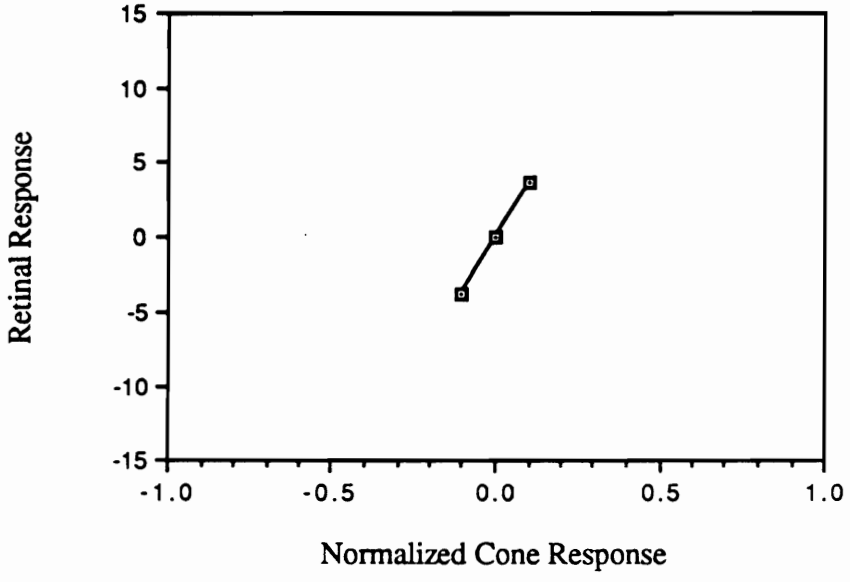


Figure 47. First two segments of the RTF.

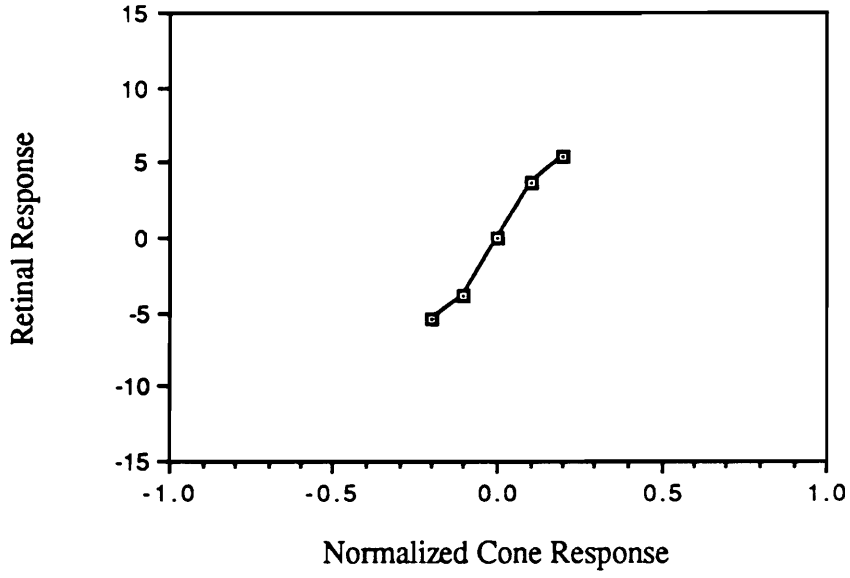


Figure 48. Second two segments of the RTF.

The remaining segments of the RTF were generated using the procedure described above. The complete RTF has twice as many segments as modulation intervals used in its derivation. Due to symmetry, the upper half of the RTF is a rotated version of the lower half. A complete RTF generated using 512 modulation intervals (1024 segments) is shown in Figure 6 (p. 48).

Vita for **Charles J. C. Lloyd**
328 Loudon Road, Blacksburg, Virginia 24060

Education

Ph.D. (May, 1988) from Virginia Polytechnic Institute and State University, Blacksburg, Virginia, in Industrial Engineering and Operations Research (Human Factors Option). Dissertation title: "A human spatial-chromatic vision model for evaluating electronic displays."

M. S. (August, 1984) from The University of Iowa, Iowa City, Iowa, in Industrial and Management Engineering (Human Factors Option). Thesis title: "The effects of irrelevant stimulus attributes in a same-different choice reaction time task."

B. S. (May, 1982) from the The University of Iowa, Iowa City, Iowa, In Psychology with minor in Physics.

September 1975 to February, 1978: Undergraduate in Electrical Engineering at Iowa State University, Ames, Iowa. No degree confered.

Work Experience

September 1985 to August 1986: Scientific programmer/analyst at Hercules Incorporated in Salt Lake City, Utah. Responsible for design and development of software system for performing data reduction, analysis, and report generation for small-scale rocket motor tests. Tests were conducted by ballistics range for characterization of propellants.

September 1984 to August 1985: Human Factors Engineer at Hercules Incorporated in Salt Lake City, Utah. Responsible for computer workstation furniture selection, training systems design, and human factors support during the design of a remotely-controlled solid rocket motor manufacturing facility. Facility is a 200 million dollar automated plant capable of producing motors containing up to hundreds of tons of nitroglycerin-based propellant.

Research Experience

March 1989 to Present: Conducted dissertation research which involved development of a spatial-chromatic model of human vision and a methodology for using this model for the evaluation of color electronic display systems. Conducted a series of computational experiments designed to characterize the performance of the model. Demonstrated the model's ability to predict human contrast detection, discrimination, and perception data. Also demonstrated the model's ability to predict image quality for monochromatic display systems. Currently working on characterization of the chromatic behavior of the model.

March 1989 to Present: Worked with Dr. Robert J. Beaton and Dr. Harry L. Snyder at Virginia Polytech on several projects including precision photometric measurement of 12-bit frame buffer display system, characterization of transmission of two liquid crystal

stereoscopic display shutters, and development of a large-area mechanical shutter for gating electronic and hardcopy displays.

September 1988 to February 1989: Research Assistant for Dr. Robert J. Beaton at Virginia Polytech. Designed and conducted study in which contrast detection and discrimination thresholds were measured for chromatic and achromatic Gabor patterns. Supervised the activities of three students who assisted in project. Designed and constructed a two field optical/mechanical apparatus for viewing stimuli. Apparatus allowed the modulation of each field to be independently controlled over a 1000:1 range and contained a binocular optical system for correcting accommodation and convergence to a specified viewing distance.

October 1988 to December 1988: Designed and conducted experiment involving human factors evaluation of real-time image processing operators for use in image enhancement and document restoration systems.

June 1988 to August 1988: Research Assistant for Dr. Harry L. Snyder at Virginia Polytech. Responsible for design and development of personal computer-based software system for conducting spectroradiometric measurements. System includes procedures for monochromator positional control, radiance calibration and measurement, CIE color coordinate calculation, and plotting of radiance scans. Gained experience with Data Translation data acquisition system and Gamma Scientific monochromator, photo-multiplier tube, and digital radiometer.

January 1988 to May 1988: Research Assistant for Dr. Harry L. Snyder at Virginia Polytech. Responsible for experimental design, design and construction of equipment, programming, running subjects, data analysis, and writing technical report for study evaluating the use of a touch screen system in adverse environmental conditions. Gained experience with Carroll-Touch IR touch sensitive device, EMS electroluminescent display, Rheem environmental control chamber, PDP-11/55, and DEC Laboratory Peripheral System.

June 1987 to December 1987: Research Assistant for Dr. Harry L. Snyder at Virginia Polytech. Responsible for acquiring and running subjects, data analysis, and technical report write-up for studies dealing with the effects of line and cell failures, failure mode, matrix size, display polarity, and font on reading and search performance. Gained experience with Tektronix high-resolution display and Gamma Scientific scanning microphotometer.

September 1986 to May 1987: Research Assistant for Dr. Harry L. Snyder at Virginia Polytech. Responsible for acquiring equipment, running subjects, and analyzing response data for study evaluating the usability of four electronic typewriters.

September 1983 to May 1984: Designed and implemented a microcomputer-based experiment management system. System controls the assignment of experimental conditions, presentation of stimuli, and the recording and compiling of subject response data for discrete-trial, human information processing experiments.

January 1983 to May 1984: Research Assistant for Dr. J. Richard Simon at The University of Iowa Human Factors Laboratory. Responsible for experimental design, hardware setup, programming, acquiring and running subjects, and data analysis for

experiments dealing with two-choice reaction time. Gained experience with experimental control and data collection using a PDP-11/34 minicomputer system.

Publications

Decker, J. J., Dye, C. J., Kurokawa, K., and Lloyd, C. J. C. (1988). Effects of display failures and symbol rotation on visual search using dot-matrix symbols. In *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA: Human Factors Society.

Downing J., Joma, G., Lloyd, C. J. C., Mohageg, M., and Snyder, H. L. (1987). *Typewriter evaluation study executive summary* (Tech. Report HFL-87-1E). Rochester: Xerox Corporation.

Lloyd, C. J. C. and Beaton, R. J. (1990) Design of a spatial-chromatic human vision model for evaluating full-color display systems. In *Proceedings of the SPIE, Vol 1249, Human vision, visual processing, and digital display*. Bellingham, WA: SPIE.

Lloyd, C. J. C., Beaton, R. J., Parsley, M., and Simonetti, T. (1990). *Contrast detection and discrimination thresholds for chromatic and achromatic gratings* (Tech. Report). Blacksburg, VA: Virginia Polytechnic Institute and State University.

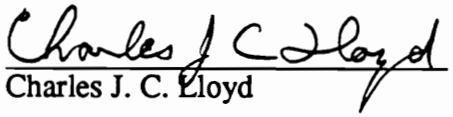
Lloyd, C. J. C., Decker, J. J., Kurokawa, K., and Snyder, H. L. (1988). Effects of line and cell failures on reading and search performance using matrix-addressable displays. In *SID International Digest of Technical Papers, XIX*, Playa del Rey, CA: SID.

Lloyd, C. J. C., and Raz, T. (1984). Experiment management system: A software tool for human factors research. In H. W. Hendrick and O. Brown, Jr. (Eds), *Human Factors in Organizational Design and Management* (pp. 229-233). New York: North-Holland.

Lloyd, C. J. C., Takahama, M., Farley, W. W. Jr., and Snyder, H. L. (1988). *Military evaluation of a commercial item touch panel evaluation study: Final report.* Contract number DAAL01-85-C-0411, US Army LABCOM, Fort Monmouth, NJ.

Professional Affiliations

Human Factors Society
Society for Information Display
Institute of Electrical and Electronic Engineers
Optical Society of America
The Society for Imagining Science and Technology
The International Society for Optical Engineering



Charles J. C. Lloyd