UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: ONR-CR213-120-3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): IMAGE DESCRIPTORS FOR DISPLAYS		5. TYPE OF REPORT & PERIOD COVERED: Annual Report, (12-10-75 to 12-31-76)
		6. PERFORMING ORG. REPORT NUMBER: PRRL-77-CR-7
7. AUTHOR(s): Joseph J. Mezrich, Curtis R. Carlson and Roger W. Cohen		8. CONTRACT OR GRANT NUMBER(s): N00014-74-C-0184
9. PERFORMING ORGANIZATION NAME AND ADDRESS: RCA Laboratories, Princeton, New Jersey 08540		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: Task No. NR 213-137
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Arlington, VA 22217		12. REPORT DATE: February 1977
		13. NUMBER OF PAGES: 142
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): 139p.		15. SECURITY CLASS. (of this report): Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Modulation Transfer	Power Spectra
Displays	Images
Visual Perception	Sharpness
Pattern Recognition	Image-Quality Descriptors

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report summarizes our experimental and theoretical results on the perception of displayed information. This research is concerned with four issues: (1) the effect of color on pattern recognition, (2) the measurement of statistical estimates for luminance and chrominance information in pictorial scenes, (3) the relationship between perceived changes in image sharpness resulting from changes in display modulation transfer, and (4) the development of a unified display descriptor to model the perception of both luminance and chrominance information.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

299 000

20.

Our experiments to determine the utility of color provide results which establish an empirical basis for determining general principles about the role of color in pattern recognition. We have found that color may be irrelevant under conditions that represent global stimulus processing and when the shape and color dimensions are handled as separable. However, our results also indicate that color is involved in the shape recognition process when the separate dimensions of color and shape are integral or when the stimulus is locally processed. Thus, our experiments have shown that a single answer concerning the utility of color in pattern recognition is not to be expected.

The measured power spectra for both luminance and chrominance pictorial information were found to roll off at high spatial frequencies as the inverse square of the spatial frequency. We feel that these results support the hypothesis that edge transitions represent a significant feature of both the luminance and chrominance content of pictorial information. Unlike the luminance power spectra, however, which must have a maximum value at dc, the chrominance power spectra may have a value close to zero at dc. Also, our measurements on the rms modulation depths for pictorial scenes have shown that the average luminance fluctuations are large, whereas the average chrominance variations are relatively small.

In this report we present the results of a series of experiments that established the relationship between our sharpness descriptor, the visual capacity, and the subjective sharpness of displayed images. The just-noticeable difference (jnd) in image sharpness was measured as a function of display bandwidth. At high spatial frequencies these measurements were found to be in excellent agreement with the assertion that display sharpness is mediated by the perceived rms gradient content of an image: a quantity proportional to the square root of the visual capacity. The addition of chrominance information to a black-and-white image was not found to appreciably affect the perception of image sharpness. Finally, the measured results on the subjective sharpness for both pictorial images and single-transition luminance edges proved to be indistinguishable. We feel that this result supports the contention that edge transitions are important in determining image sharpness in pictorial scenes.

In this report a descriptor for the total channel capacity of the display-observer system is developed that includes the statistical properties of both luminance and chrominance information. This descriptor is based on a widely accepted model of the visual system that contains three independent channels: one channel that transmits the luminance information, and two channels that transmit the chrominance information. For the first time the nonlinearities associated with luminance perception are included in our model through a novel interpretation of recent psychophysical experiments. The model for the total information capacity is used to predict the optimum allocation for a commercial television system; the results are shown to be in good agreement with current U.S. television practice.

PREFACE

This Annual Report was prepared by RCA Laboratories, Princeton, New Jersey 08540, for the Office of Naval Research under Contract No. N00014-74-C-0184, task number NR 213-137. It describes work performed during the period 10 December 1975 to 31 December 1976 in the Materials Research Laboratory, J. J. Tietjen, Director. The Project Scientist is R. W. Cohen. Others who participated in the research are J. J. Mezrich and C. R. Carlson.

The authors wish to express their appreciation to the following: I. Gorog for his encouragement, advice, and critical evaluation of this work, Ralph Klopfenstein for his assistance in several numerical computations, Robert McNair for the development of the computer programs described in Section II, and Ronald Infanti for his help in the acquisition of the data presented in Section IV.

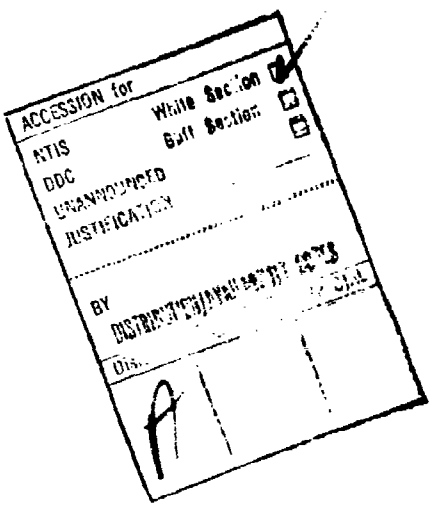


TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

*Commission Internationale de l'Eclairage (International Commission of Illumination).

**National Television System Committee. This committee of the Electronic Industries Association established the system, approved by the FCC (Federal Communications Commission), for commercial broadcasting in the U.S.

LIST OF ILLUSTRATION (Continued)

LIST OF TABLES

LIST OF NOTATIONS

A	Inspection area
$A(\nu)$	Contribution of internal visual noise to luminance contrast threshold function [Eq. (29)]
$B(\nu)$	Contribution of shot noise to luminance contrast threshold function [Eq. (29)]
$c_{eq,i}$	Equivalent chrominance sine-wave modulation [Eq. (35)]
$c_{Ti}(\nu)$	Threshold contrast-sensitivity function for chrominance sine-wave gratings [Eq. (34)]
$\overline{c^2_{m,i}}$	Mean square chrominance modulation for axis i = 1,2 in chromaticity diagram
C_v	Visual capacity [Eq. (5)]
$\overline{C}_I$	Average chrominance value for a scene, as percent saturation along the I_C axis [Fig. 14]
$\overline{C}_Q$	Average chrominance value for a scene, as percent saturation along the Q_C axis [Fig. 14]
$C_{m,I}$	Chrominance signal defined as $C_{m,I} = C_{T,I} - \overline{C}_I$ [Fig. 14]
$C_{m,Q}$	Chrominance signal defined as $C_{m,Q} = C_{T,Q} - \overline{C}_Q$ [Fig. 14]
$C_{T,I}$	Total chrominance signal along I_C axis [Fig. 14]
$C_{T,Q}$	Total chrominance signal along I_Q axis [Fig. 14]
d	Spacing between diffuser plates [Fig. 29]
e	Reduced perceived scene brightness [Eq. (2)]
E	Perceived intensity pattern on display
f	Spatial frequency on display [Fig. 31]
g	rms perceived gradient content [Eq. (5)]
Δg	Change in g necessary to produce one jnd in image sharpness [Eq. (7)]

H	Total channel capacity of display-observer system [Eq. (31)]
H_{C1}, H_{C2}	Channel capacity of two chrominance channels [Eq. (32)]
H_L	Channel capacity of luminance channel [Eq. (24)]
i	Index denoting perceived contrast levels
$\bar{I}$	Mean display luminance
$\overline{I^2}$	Mean square display luminance
I_C	Chromaticity axis [Fig. 14]
I_m	Luminance signal defined as $I_m = I_T - \bar{I}$ [Fig. 14]
I_{max}	Maximum value of luminance edge [Fig. 34]
I_{min}	Minimum value of luminance edge [Fig. 34]
I_Q	Chromaticity axis [Fig. 13]
I_T	Total luminance signal [Fig. 14]
jnd	Just-noticeable difference
k	Constant fraction parameter for realistic model of luminance level distribution [Eq. (16)]
k_w	Constant fraction parameter for Weber's law model of luminance level distribution [Eq. (9)]
L_o	Average luminance of displayed sine-wave gratings
m_o	Initial contrast in luminance contrast-discrimination experiments
$m_{eq}(\omega)$	Equivalent luminance sine-wave modulation [Eq. (19)]
m_i	Luminance sine-wave contrast required for perception of the i'th luminance level
$m_T(\omega)$	Threshold contrast-sensitivity function for luminance sine-wave gratings
MTF	Modulation transfer function
n	Dot number
$n(\omega)$	Number of perceivable luminance contrast levels at frequency ω [Eq. (25)]
$n_{Ci}(\omega)$	Number of perceivable chrominance contrast levels at frequency ω [Eq. (37)]

n_T	**Threshold dot number**
$N(\omega)$	Noise power spectrum
N_{TV}	Limiting display resolution expressed in TV-lines: $N_{TV} = \omega_M w/\pi$
N_{TV}^{eq}	Number of TV-lines for noiseless display required to give same H as noisy display
$N_v(\nu)$	Total visual noise power spectrum [Eq. (13)]
$N_{vI}(\nu)$	Contribution of internal visual noise to total visual noise power spectrum
$N_{vS}(\nu)$	Contribution of shot noise to total-visual-noise power spectrum
N_s	Luminance noise fluctuation [Eq. (26)]
$O(\nu)$	One-dimensional human-visual-system MTF
p	significance level
P	Probability of event
$P_c(t)$	Probability of color detection in time t [Eq. (1)]
$P_{c/s}(t)$	Probability of color-coded shape
$P_s(t)$	Probability of shape detection in time t [Eq. (1)]
$P(\delta\nu)$	Psychometric function defining the probability of correctly perceiving the sharper of two images that differ in bandwidth by $\delta\nu$
Q_C	Chromaticity axis [Fig. 13]
r	Viewing distance
$R(\omega)$	One-dimensional display MTF
R_{eff}	Effective overall display MTF for luminance signals
$R_{eff,Ci}$	Effective overall display MTF for two chrominance channels: i = 1,2
S	Perceived rms signal power [Eq. (5)]
S/N	Display signal-to-noise ratio

TR1	Reference 1
TR2	Reference 2
w	Display width
W	Total number of distinguishable states of display-observer system [Eq. (30)]
W_{C1}, W_{C2}	Contribution of two chrominance channels to the number of distinguishable states of display-observer system
W_L	Contribution of luminance channel to the number of distinguishable states of display-observer system [Eq. (22)]
α	Confidence factor
$\delta\nu$	Change in retinal display bandwidth, $\delta\nu = \nu_2 - \nu_1$, necessary to produce one jnd in image sharpness
ΔL^2	Change of mean square luminance in contrast-discrimination experiments [Eq. (10)]
Δm	Sine-wave contrast change required for observer discrimination [Eq. (12)]
$\Delta\nu$	Width of frequence-specific channel of visual system (luminance)
$\Delta\nu_{Ci}$	Width of frequency-specific channel of visual system (chrominance)
$\Delta\omega_o$	Change in display bandwidth [Fig. 40]
ν	Frequency coordinate on the retina: $\nu = \omega r/2\pi$
ν_o	Retinal frequency defined to be 3 cycles/degree-of-vision
ν_1	Retinal display frequency defined at $R(\omega r/2\pi) = 1/2$ where $\nu_1 < \nu_2$
ν_2	Retinal display frequency defined at $R(\omega r/2\pi) = 1/2$
$\Phi_L(\omega)$	Luminance power spectrum of natural scenes [Eq. (18)]
ω	Angular spatial frequency coordinate measured on display screen
ω_o	Minus 3-dB point of $R(\omega)$, i.e., where $R(\omega_o) = 1/2$
ω_o'	Display bandwidth at -3-dB point [Fig. 40]

ω_L	Lower cutoff frequency of luminance and chrominance power spectra
ω_M	Upper cutoff frequency for display MTF
θ	Viewing angle [Eq. (3)]

SECTION I

INTRODUCTION

All practical display systems are concerned with the efficient transfer of information to the human observer. This objective requires a detailed understanding of the perceptual capabilities of the observer, the performance characteristics of the display, the properties of the information to be transmitted, and the observer response intended from this information. Furthermore, these separate elements must be fused together in a manner that will result in the optimum display design for a given task.

Historically, displays have been developed empirically to maximize, under specific design constraints, the performance of the display. Although an understanding of both the relevant visual processes and the salient features of the information to be displayed were important in the overall conception of the display, there was no systematic methodology available to use in specifying detailed design requirements. Typically, design refinements were resolved by a process of iteration to achieve the best overall compromise in display performance. This approach, exemplified by commercial television, can be impressively successful. Thus today, almost 25 years after the adoption of U.S. commercial television standards, satisfaction with them is still nearly universal. However, although the empirical approach has resulted in a wealth of information about display design and certain aspects of human vision, it does not readily lend itself to generalization. In effect, each display must be optimized individually, a process that is both time consuming and expensive.

In this report, and in two previous reports [1,2] (hereafter referred to as TR1 and TR2, respectively), we have presented a family of mathematical descriptors that we have developed for the quantitative evaluation of the performance of displays viewed by human observers. This work follows the research approach of Rose [3], Schade [4], Biberman [5], and others [6-8] for advancing the understanding of displayed information by the development of models of the display-observer system.

This report is organized into a discussion of two parallel courses of research. In Section II, we present the results of a series of experiments designed to determine the general principles underlying the potential utility of color as an aid in the recognition of shapes in various contexts. In Sections III through V, we have continued to expand the development of the mathematical description of the display-observer system presented in TR1 and TR2.

Broadly speaking, the image descriptors we have developed are concerned with issues of information visibility. The various descriptor models address different spatial attributes whose visibility should be maximized for optimum display design. There remains the question, however, as to how the visible information is organized and handled by the observer to meet some performance goal. For example, the role of color in the recognition of visible patterns is beyond the scope of our descriptor formulation.

Rather than treat the recognition issue as a byproduct of the descriptors, we have, therefore, addressed it directly. Our approach is to provide an empirical basis for pointing out some general principles about the role of color in recognition. In Section II, experiments are presented that address three issues: (1) the potential for color as an aid to recognition when the primary attention is given to stimulus shape, and the questions of (2) whether demonstrated performance advantage due to color coding is a consequence of sensory or of cognitive issues, and (3) whether color involvement is demonstrable in steadily viewed displays where speed-stress is not a factor.

Our results indicate that color is involved in the shape recognition process when the separate dimensions of color and shape are integral or when the stimulus is locally processed. Integral dimension processing is suggested for alphanumeric recognition under brief exposure stimulation; local processing is suggested for the detection of large shape subtending 1° of visual angle steady viewing. On the other hand, color can be irrelevant under conditions that we feel represent global stimulus processing, and when the shape and color dimensions are handled as separable.

Whereas separable-dimensions processing is suggested for simple-shape recognition under brief exposure stimulation, global processing is suggested for the detection of small shapes (subtending 1/2° of visual angle) under steady viewing. Thus, our experiments show that a single answer concerning the utility of color in pattern recognition is not to be expected.

Our approach in the development of display descriptors has been to use the tools of statistical information theory. That is, we treat the displayed information, the effects of sampling, and the noise sources (both display and visual noise) as statistically averaged properties of the display-observer system. This approach has the important advantage of allowing one to determine the most likely behavior of the display for a given set of conditions. Further, as we show in Section III for pictorial scenes, the conclusions derived from statistical estimates may be expected to have wide applicability because the variations in these values are relatively small.

For the application of our display descriptors to actual displays, the statistical properties of the information to be presented on the display must be known. In TR1 we measured the ensemble-averaged power spectral density for luminance information, as represented by off-the-air commercial television. In that study we found that the power spectral density, for frequencies above a low-frequency cutoff, decreases as the square of the inverse spatial frequency. That is, the luminance power spectral density for actual scenes was identical in form to the power spectral density for randomly located luminance-edge transitions. As explained in TR1, this result supports the hypothesis that edge transitions represent a significant feature in natural scenes. In Section III of this report we have extended our luminance power spectral density measurements to show that the inverse square frequency roll-off described above is not affected by the finite angular extent of the scenes studied. In Section III we also present the first results of measurements of the chrominance power spectral densities for pictorial scenes. As in the case of the luminance power spectra, the chrominance power spectra were found to roll off at high spatial frequencies as the inverse square of the spatial frequency. However, unlike the luminance power spectra, which must have their maximum value at dc, it was found that the chrominance power spectra can have a value close to zero at dc. Finally, results are presented for the rms modulation depths for both the luminance and chrominance information in pictorial scenes. These results show that while the average luminance fluctuations in pictorial scenes are highly modulated, the average chrominance variations are relatively small.

An important aspect of our program has been the verification of our mathematical display descriptors. In Section IV we present the results of a series of experiments that established the relationship between our sharpness descriptor, the visual capacity, and the subjective sharpness of displayed images. In this study the relationship between the just-noticeable difference in image sharpness was measured as a function of display bandwidth. At high spatial frequencies these measurements were found to be in excellent agreement wih the assertion that display sharpness is mediated by the perceived rms gradient content of the image: a quantity proportional to the square root of the visual capacity. This result indicates that the visual capacity can be used as a normalizing function to predict the effective sharpness of displays with different MTFs. Another important result of these measurements was the observation that the addition of chrominance information to a black-and-white image does not appreciably affect the perception of image sharpness for that image. Finally, the measured results for a series of experiments on the subjective sharpness of both representative images and single-transition luminance edges proved to be indistinguishable. This result gives additional support to the contention that edges are a significant feature in determining the perceived attributes of actual scenes.

Our display descriptors for luminance perception, described in reports TR1 and TR2, were constructed on the assumption that the visual system was linear, an assumption that implicitly limited our analysis to signals of small amplitude. In Section V of this report we have removed this restriction. Instead, a model has been employed that takes into consideration the nonlinearity of brightness perception. This model, based on our interpretation of recent psychophysical experiments, also supports our previous assumption that the human visual system responds to the square of the signal amplitude.

Also, our previous efforts had been directed only at modeling displayed luminance information. In Section V of the report, however, we have extended our descriptor for the total channel capacity of a display to include both chrominance and luminance information. This descriptor is based on a widely accepted model of the visual system that contains three independent channels: one channel which transmits the luminance information, and two channels which

transmit the chrominance information [9]. The distribution of perceivable chrominance levels is included by using the measured results of appropriate psychophysical experiments. As a practical example, the total channel capacity is used to predict the optimum allocation for a commercial television system; the results are shown to be in good agreement with current U.S. television practice.

SECTION II

EFFECT OF COLOR ON PATTERN RECOGNITION

A. INTRODUCTION TO EXPERIMENTS

1. Overview

An observer engaged in visual pattern recognition is involved in decision making. That is, he must decide on the proper assignment of the inspected visual pattern to some category in his memory. It, therefore, appears reasonable that the more information this pattern presents to the observer, the more information he has to base his decision on; hence his decision making for recognition will be made more efficient. In particular, if patterns contained both shape and color information one would assume it likely that recognizability would be better than if color information were absent.

The difficulty with this assumption is that it does not take into account the inherent flexibility that permits a decision maker to reject or include color in the decision, depending on how the visual stimulus is processed. Although it is certain that color will benefit pattern recognition in many circumstances, it is also likely that there will be circumstances in which color will be disregarded in the recognition process. The thrust of our investigation of color effects in pattern recognition is to determine, in a general sense, when color is included in decision making for pattern recognition.

We describe, next, the experimental setup and a data collection algorithm used throughout our study. The three major sections which follow are concerned with alphanumeric recognition after brief exposure to stimuli, simple-shape recognition after brief exposure to stimuli, and a new paradigm for the investigation of shape recognition in steadily viewed displays in the absence of speed-stress. Our intent is to point out general issues of color involvement in pattern recognition; these issues are summarized in Section II.E.

2. Instrumentation and Procedures

Our experimental setup consisted of a NOVA* 2/10 minicomputer interfaced through a Lexidata** graphics display board to an RGB (red-green-blue) color

*Made by Data General Corp., Southboro, Mass.
**Produced by Lexidata Corp., Lexington, Mass.

monitor. For the display mode used in our experiments, a centered 128 x 128 point screen region could be addressed with red, green, or yellow (red + green) dots. Up to 16,384 [$(128)^2$] colored dots could be displayed at a time, and the entire display could be changed as often as every 1/60 s (TV field rate). Scan interlace was not used on the monitor.

The 1/60-s resolution of display change was important for our brief-exposure experiments; it also allowed us to use flicker photometry to set the red and green luminances to be equal. This was accomplished by a display of all the 128 x 128 points as red and green dots on alternating 1/60-s fields and, during steady viewing of this alternating pattern, adjustment of red-green ratio so that a nonflickering yellow field was seen. A silicon photocell was used to view this field alternation on an oscilloscope to check that the red and green phospor onset and offset times were equivalent and sufficiently rapid, i.e., that our setup did, indeed, provide a 1/60-s display resolution.

The observer sat behind a desk, his head about 10.5 ft from the monitor screen, and used a button box to communicate with the computer. For those experiments where brief-exposure stimuli were used (see Sections II.B and II.C), white fluorescent ceiling lights were kept on in the viewing area. The lights were situated behind the monitor screen, providing diffuse viewing-area illumination without screen glare to reduce the apparent contrast of the display. This was deemed the most suitable way to stretch the response-accuracy psychometric function without removing color legibility, so that 1/60-s changes in exposure duration were small enough to reveal performance trends. With overhead lights on, the blank monitor screen was at about 2 mL. A schematic diagram of the viewing area arrangement is shown in Fig. 1.

In all of our experiments, some type of threshold measurement was carried out, either to set a stimulus level for the experiment or as the main feature of the experiment. This amounts to finding the stimulus strength necessary to yield an appropriate level of detection accuracy. The procedure we use is the sequential estimation procedure introduced by Wetherill [10]. It is a particularly efficient means for estimating the desired stimulus parameter. One first decides on a step size to increase and decrease stimulus strength. The subject is run through a sequence of stimulus trials with the stimulus strength decreased by a step after a pair of correct responses, and with the stimulus

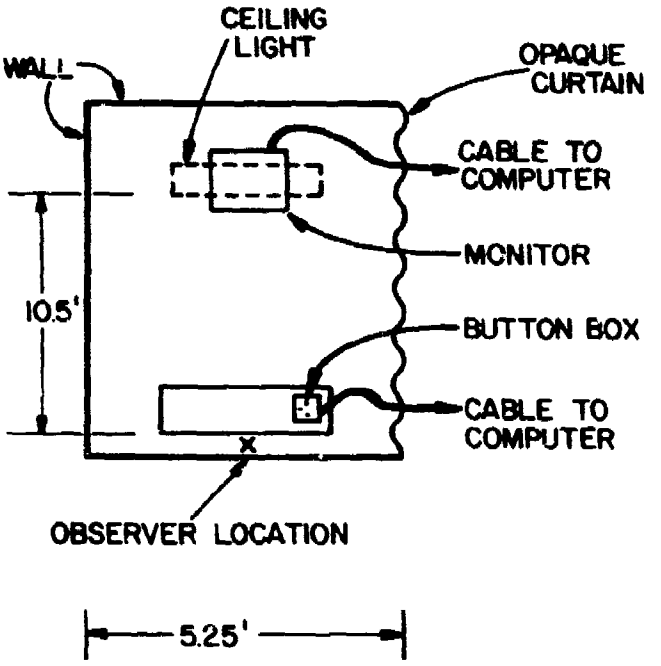


Figure 1. Viewing arrangement for experiments.

strength increased by a step after a pair of responses at least one of which is correct. One then has runs of increasing and decreasing steps in stimulus strength. The transition from runs of increasing to runs of decreasing strength is termed a peak, and the transition from decreasing to increasing runs is termed a valley. The average of the stimulus values at the peaks and valleys provides the estimate of the stimulus value at which accuracy satisfies the relation $P^2 = 1 - P^2$, or $P = \frac{1}{\sqrt{2}} = 0.707$. In our procedures we stop the threshold determination after eight peaks and valleys are produced. We throw out the first peak and valley to avoid starting point bias, and form our 70.7% estimate from the remaining peaks and valleys. In this procedure pairs of trials are inspected. Alternatively, quadruples of trials may be used to estimate the stimulus value where $P^4 = 1 - P^4$ is satisfied, i.e., where $P = 0.841$. The 84.1% estimate was used only in our experiment in Section II.C. In experiments where stimulus strength was varied by changing exposure duration,

the step size for the duration change was 1/60 s. In the experiment of Section II.D, dot number was the parameter of interest; the step size was set at 500 dots. These procedures were carried out by the computer.

B. INFLUENCE OF COLOR IN LETTER AND WORD RECOGNITION

1. Introduction

There are two ways in which color can be used for alphanumeric stimuli - as a redundant or a nonredundant code. An example of the nonredundant use of color is the commonly used red coloring of LED (light-emitting diode) displays. The use of color in this situation is primarily an issue of making display visibility possible in the sense of providing a match to the spectral sensitivity of the eye. Consequently, a poor color choice can usually be rendered acceptable simply by increasing the display brightness, or contrast. The color, per se, is irrelevant and does not aid the observer in deciding what the displayed message is. On the other hand, color can also be used as a redundant stimulus code. The question we will concern ourselves with here has to do with the potential for color as a recognition aid when the color is used in a redundant fashion.

This is not a new issue for experimental psychology. It was addressed, for example, by Eriksen and Hake [11], who concluded that completely redundant color coding of stimulus shape is an effective means for enhancing stimulus discrimination. The general topic has been discussed by Garner [12].

Nevertheless, there is reason to wonder whether such color benefit conclusions based on controlled laboratory experiments can be justifiably generalized to a nonlaboratory setting. The particular point that concerns us is one of attention priority. If one carefully instructs an observer in an experiment about the availability of correlated shape and color attributes of the stimulus, then insofar as the observer makes use of these instructions and attends to both stimulus attributes, one would expect a performance benefit due to the color coding of shape. Outside of the laboratory, however, one might often expect a natural tendency to attach priority to the processing of stimulus shape. Priority of attention to shape is particularly likely for alphanumeric stimuli, where a lifetime of experience orients one to read a

displayed message without concern for its color, or even for precise details of shape [13]. It is, therefore, worth asking whether, in a situation where priority of attention is given to shape processing for alphanumeric stimuli, a benefit of color coding is demonstrated. This is the question to which we address ourselves here.

2. Experimental Procedure

The sequence of photographs shown in Figs. 2 and 3 represents typical stimulus trials. The observer viewed a yellow fixation symbol (Figs. 2a or 3a) on the monitor and, when ready, pressed a button which caused the immediate presentation of a red or green stimulus centered directly above the screen location of the removed fixation-symbol arrow. The stimulus could be either a letter (Fig. 2b) or a four-letter word (Fig. 3b). When a predetermined brief duration had elapsed, the stimulus was immediately removed from the screen and replaced with a postexposure mask consisting of 1250 red and 1250 green dots, randomly located in a 2.6° x 2.0° region covering the stimulus location. Adjacent to the mask were two response-choice letters, one in red and one in green. The choices plus mask are shown in Figs. 2c and 3c. The subject's task was to inspect the choices and decide which of the two letters, disregarding color, had just appeared in the briefly displayed stimulus. Selection of the top or bottom choice on the screen was indicated by the pressing of one of two buttons. The correct response in Fig. 2c is clearly the top letter. The correct response in Fig. 3c is the top letter, since D appeared in the stimulus. Accuracy feedback was displayed (in yellow) immediately after a response was given (Figs. 2d and 3d).

In deciding between D or N in the response situation of Fig. 3c, there is no redundancy benefit to having processed "WORD" in the stimulus (Fig. 3b), since either choice letter could be used in the target letter location to make up a common word. All word stimulus trials had this property; the target letter and the incorrect choice could both be used in the critical letter position to construct a common word with the nontarget letters. The target letter position in the word was randomized for word stimuli in the experiment. This forced-choice procedure was introduced by Reicher [14] for investigating processes in word recognition, and is regarded as the safest procedure for minimizing guessing as a source of spurious results [15].

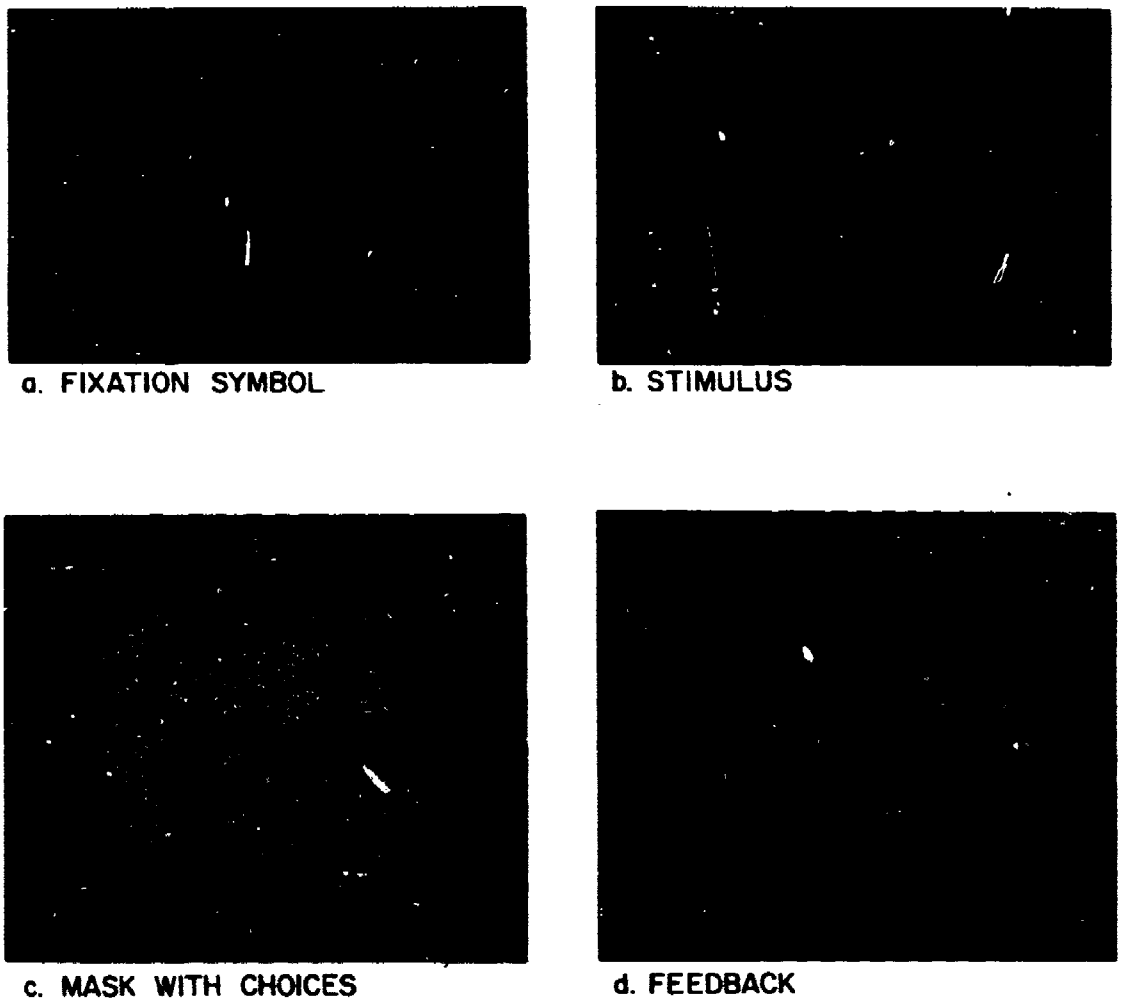


Figure 2. Letter stimulus trial. Fixation symbol (a) is followed by brief letter exposure (b); this is followed by dot mask and response choices (c). Accuracy feedback (d) follows response.

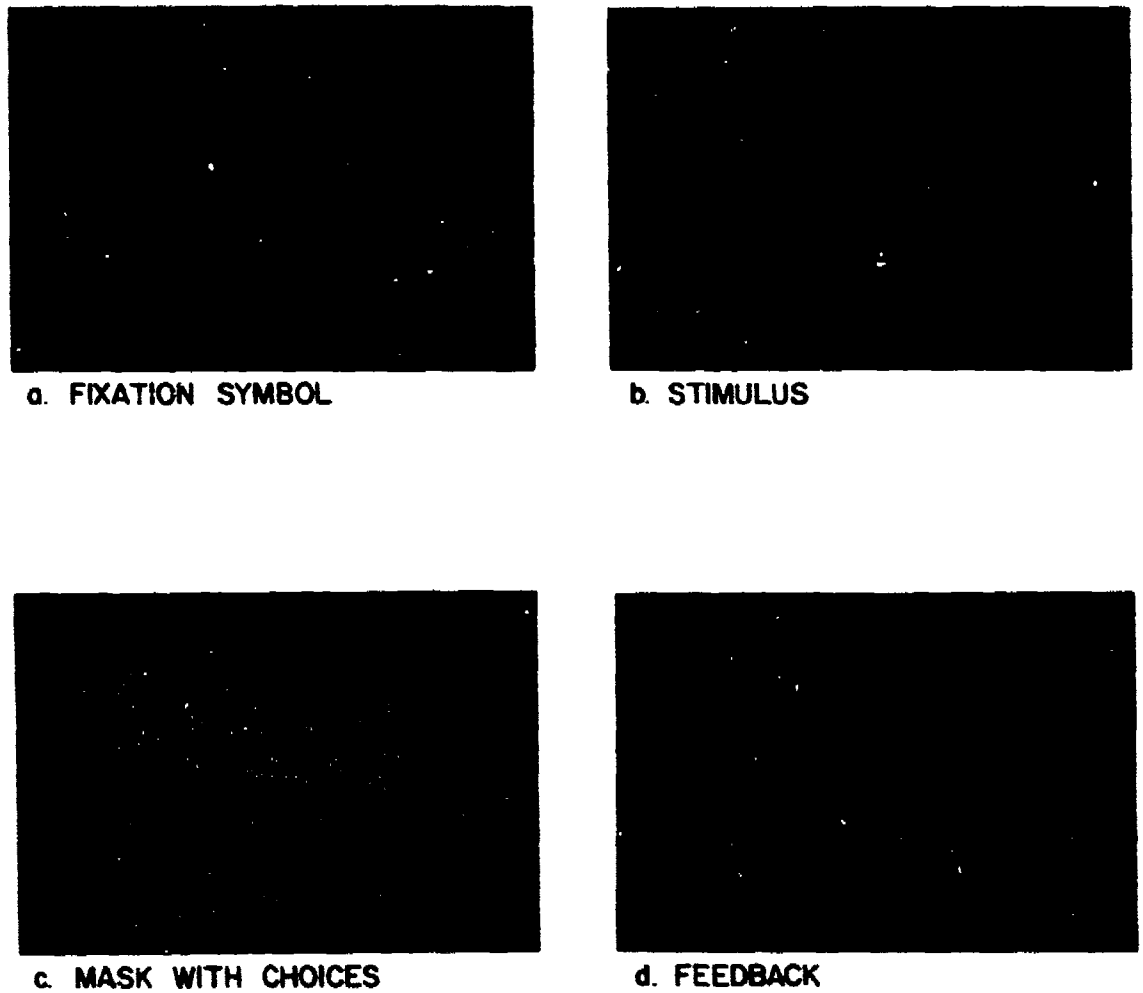


Figure 3. Word stimulus trial. Similar to Fig. 2, except that the stimulus (b) is a four-letter word.

Stimulus and choice letters (as well as the random dot mask) all had a luminance of about 10 mL, regardless of color. Letters subtended about one fourth of a degree, and were made up of dots subtending less than 3 min of arc. The viewing area was illuminated by overhead lights (see Section II.A.2.).

The stimulus color was equally likely to be red or green, and the correct response choice was as likely to be the same color as the stimulus as not. Placement of the correct choice in the top or bottom response position was random. The random dot pattern used for afterimage masking was different on each trial. The subject's task was to attend only to the letters (shape) of the stimulus and to select that choice letter which was thought to be presented in the stimulus. Subjects were instructed to disregard color, since choice selection based on color was as likely to be wrong as correct. The experimental situation, then, is one in which priority of attention is given to shape processing.

The color situations in this experiment are outlined in Table 1.

TABLE 1. STIMULUS-CHOICE COLOR COMBINATIONS

	Stimulus Color	Color of Correct Response Choice
1	Red	Red
2	Red	Green
3	Green	Green
4	Green	Red

There were four color situations and two stimulus types: letters and words. Consequently there were eight different conditions, randomly interspersed throughout an experimental session. There were 304 experimental trials per session; eleven observers were tested in different sessions. The exposure duration for stimuli was set individually for an observer so that recognition accuracy would be better than chance but less than perfect, so that the data could reveal systematic error trends. This exposure duration, typically 1/30 s, put accuracies in the neighborhood of 75% correct.

An experimental session lasted about an hour and consisted of the following: The observer was introduced to the experiment with a short instruction phase in which six trials, three with letter and three with word stimuli, were presented; stimulus exposure was set at 500 ms. Then a warm-up phase was run through. The warm-up phase had 32 trials - a random ordering of 16 letter-stimulus and 16 word-stimulus trials - with stimulus exposure set at 100 ms. Next, the proper stimulus exposure duration for the observer was determined. This consisted of applying the general sequential estimation procedure, outlined in Section II.A.2. (Instrumentation and Procedures) to temporal threshold determination. Only letter stimuli were used in this timing phase. After completion of the timing phase the main phase of the experiment was run. To minimize fatigue problems, a five-minute intermission occurred half way through the main phase.

3. Results

The first question of interest is this: Is the color used in the recognition task? In other words, is the observer more likely to be correct in his response if the correct choice is the same color as the stimulus (same condition) than if it is in the opposite color (opposite condition)? This question is answered in Table 2.

TABLE 2. COLOR EFFECT ON PERFORMANCE

Stimulus Type	Same Condition	Opposite Condition	Difference
Words:	73.1%	65.0%	8.1%
Letters:	74.4%	66.4%	8.0%

The table entries are percent correct, averaged across eleven observers. Of interest are the entries for the column marked "Difference." There was a systematic difference between the same and opposite conditions, amounting to an 8% difference in recognition accuracies. The effect was significant ($p < 0.01$, Wilcoxon signed-rank test).

The above results show that color is used by observers in the recognition task. We have involved observers in a task where priority of attention is to stimulus shape, a condition which presumably simulates the natural tendency in

dealing with alphanumeric material. One might wonder whether, in this somewhat realistic cognitive setting, information available in the stimulus color makes its presence felt, in the sense of affecting recognition performance. According to our results it does.

Of further interest is whether there are accuracy differences between red and green stimuli. That is, quite aside from the issue of color coding, is there any difference between red and green stimuli in terms of recognizability? From Table 2 we see that the average accuracy on word trials was 69% (the average of the 73.1% and 65% entries of Table 1), and the average recognition accuracy on letter trials was 70.4% (average of 74.4% and 66.4%). These average accuracies are, in turn, averages across the two stimulus colors, red and green. The breakdown by color is shown in Table 3.

TABLE 3. EFFECT OF STIMULUS COLOR ON PERFORMANCE

Stimulus Type	Red Stimuli	Green Stimuli	Average
Words:	75%	63%	69%
Letters:	74%	67%	70.5%

The red-green accuracy difference was significant for words ($p < 0.01$). However, this tendency was not systematic for letter stimuli ($p = 0.42$). Since the luminance values of the red and green stimuli were matched, one would not expect any differences between red and green stimuli to be due to the spectral sensitivity of the eye. The source of such an effect is probably cognitive, involving high-level processing. This is suggested by the different red-green accuracy differential obtained in word processing and in single-letter processing, respectively; word processing involves cognitive processes different from those that are operative in letter processing [16,17]. In any event, one can say that a benefit for red over green is strongly suggested. That is in agreement with results obtained in a different context, by Tyte et al. [18].

We next ask whether there is any response bias for color. This is, do observers have any systematic tendency to respond in favor of one color over the other? This is of interest in display design in terms of the likelihood

of false alarms with colored displays. Information concerning this question is recovered from the data by the following relations:

$$P(\text{respond red}) = \frac{1}{4}[A + (1-B) + C + (1-D)]$$

$$P(\text{respond green}) = \frac{1}{4}[(1-A) + B + (1-C) + D]$$

where

A = P (pick red choice given red stimulus)

B = P (pick green choice given red stimulus)

C = P (pick red choice given green stimulus)

D = P (pick green choice given green stimulus)

Applying these relations we found the following probability estimates:

For letters: P (respond red) = 0.54
P (respond green) = 0.46

For words: P (respond red) = 0.54
P (respond green) = 0.46

The tendency to respond red was significant for both letters and words ($p < 0.01$). The response bias in favor of red is interesting in light of the conventional use of red as a danger or emergency color code. Our results suggest a readiness to select this color even in a situation where the red coloring has no particular significance. This probably means that the culturally specified significance of red is maintained even when it is not intended in a particular setting. The choice of red color coding should, therefore, be reserved for messages where the cost of message rejection is high, but the penalty of a false alarm is not prohibitive.

The response bias for red is a different effect from the apparently better recognizability of red stimuli. This follows, as one would expect, a high degree of correlation between bias and accuracy if, say, the bias were the source of the apparent accuracy benefit of red coloring. For letter stimuli the correlation was 0.13, not significantly different from zero. For word stimuli the correlation was 0.57, a value only on the verge of significance ($p = 0.07$).

C. STATUS OF COLOR AS A SHAPE-INDEPENDENT STIMULUS FEATURE

1. Introduction

If color and shape are perceptually processed as independent dimensions of a stimulus, then the benefit of color coding of shape should be clear-cut. In particular, if in some inspection time t the probability of recognizing only the stimulus shape is $P_s(t)$ and the probability of recognizing only the stimulus color is $P_c(t)$ then, assuming processing independence, the probability of recognizing the stimulus given the availability of both color and shape is

$$P_{c/s}(t) = 1 - [1-P_c(t)][1-P_s(t)] > \max [P_c(t),P_s(t)] \qquad (1)$$

Color coding advantage can, of course, be demonstrated, as was shown in Section II.B.3. However, it is not obvious that such an advantage is necessarily a consequence of perceptual efficiency in the sense of Eq. (1), rather than, say, a consequence of enhanced memory representation of the stimulus so that the retrieval from memory necessary for response identification is more likely to be correct. The distinction is one of generality. If observed color advantage to stimulus identification is a consequence of Eq. (1), then perceptual issues govern the advantage. This advantage could, therefore, be assumed to be a consequence of a type of visual processing which is not likely to be unique to a particular experimental setting. On the other hand, if cognitive issues dominate, then a completely general experimental demonstration of color advantage is not likely to occur. It is, therefore, important to determine whether Eq. (1) can be shown to govern the identification of color-coded shape.

2. Experimental Procedure

We used an alphabet of four items in this experiment: a positively sloped diagonal line and a negatively sloped diagonal line, colored either red or green. The length of the lines correspond to 8 min of viewing angle. The stimuli are summarized in Table 4.

TABLE 4. STIMULUS ALPHABET

		Color	
		Red	Green
Slope	Positive:	1	2
	Negative:	3	4

An experimental trial consisted of the following sequence of events: The observer fixated a point on a monitor screen and pressed a button which caused our computer to immediately select one of the four members of our stimulus alphabet and randomly locate it in a centered 1.7° x 1.2° viewing region for a controlled duration. At the offset of the stimulus, a 2.6° x 2.0° region, encompassing the stimulus viewing region, was covered with a dense random mask of the diagonal colored lines. This mask was constructed by random placement of 240 stimuli of type 1, 2, 3, and 4 of the stimulus alphabet (960 lines in all). Different questions were asked for this type of presentation in different phases of the experiment, depicted in Figs. 4-7.

In one phase (Fig. 4), the observer was asked to identify the shape of the stimulus, disregarding color. This was a binary choice (positive or negative slope) indicated by a button press to the computer. In another phase (Fig. 5), the observer was asked to identify the stimulus color (red or green), disregarding shape, by means of a similar button press. Data from these two experiment phases were used to estimate the observer's psychometric curves for color or shape identification as a function of the stimulus viewing time, for the colored stimulus lines of this experiment.

The main point of this experiment was to determine how shape and color information are combined for color-coded shapes. The observed psychometric functions for pure shape identification and pure color identification were used to predict recognition performance for color-coded shape under the assumption that color and shape are independently processed. In the phase of the experiment where shape was color coded, a subset of the stimulus alphabet was used. For example, only stimuli 1 and 4 were presented in a phase (Fig. 6), and the observer had to decide which one was presented on a trial. In this case a positively sloped line was always red, and a negatively sloped line

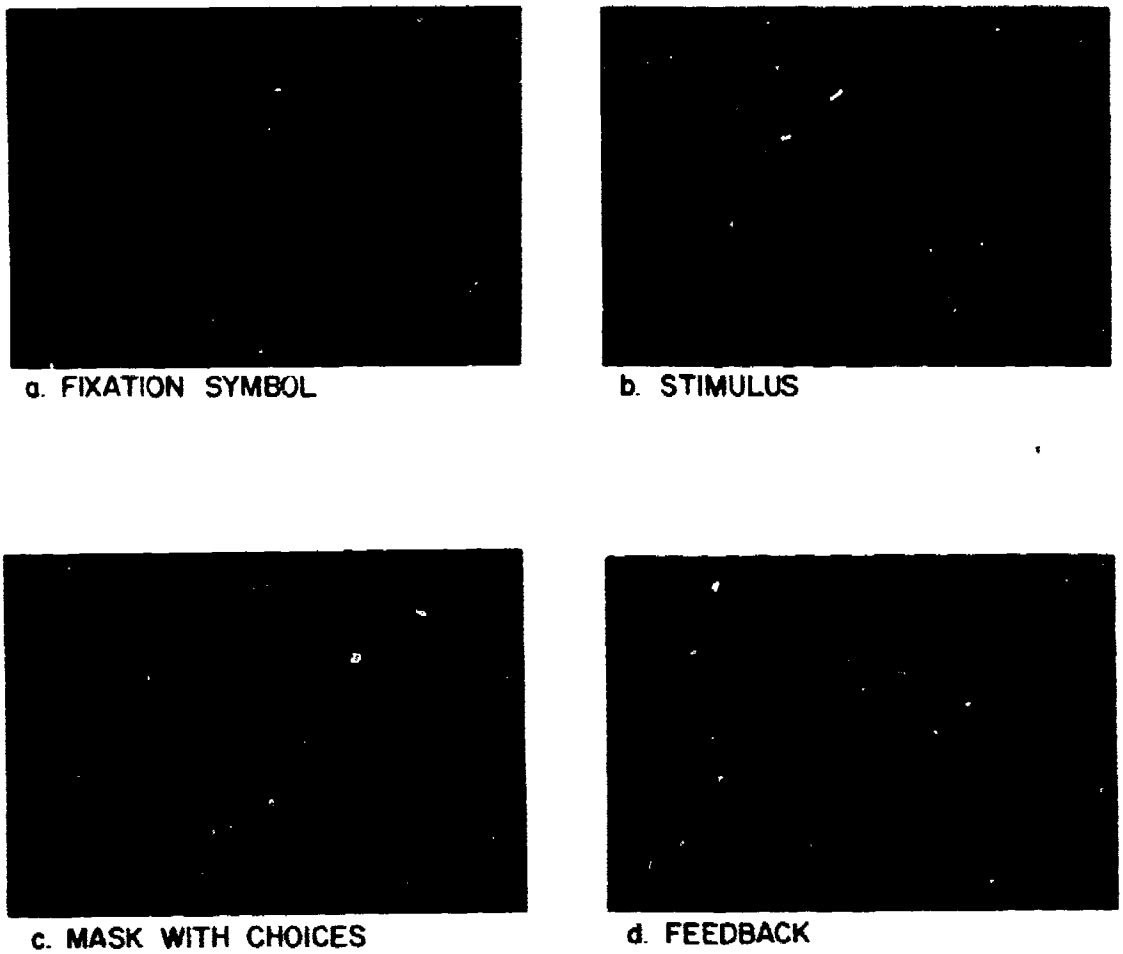


Figure 4. Shape-recognition trial. Fixation symbol (a) is followed by brief exposure of positively or negatively sloped diagonal line in a random location in the field of view (b); this is followed by a line mask and response choices (c). Accuracy feedback (d) follows response. The stimulus line is equally likely to be red or green and is equally likely to be positively or negatively sloped.

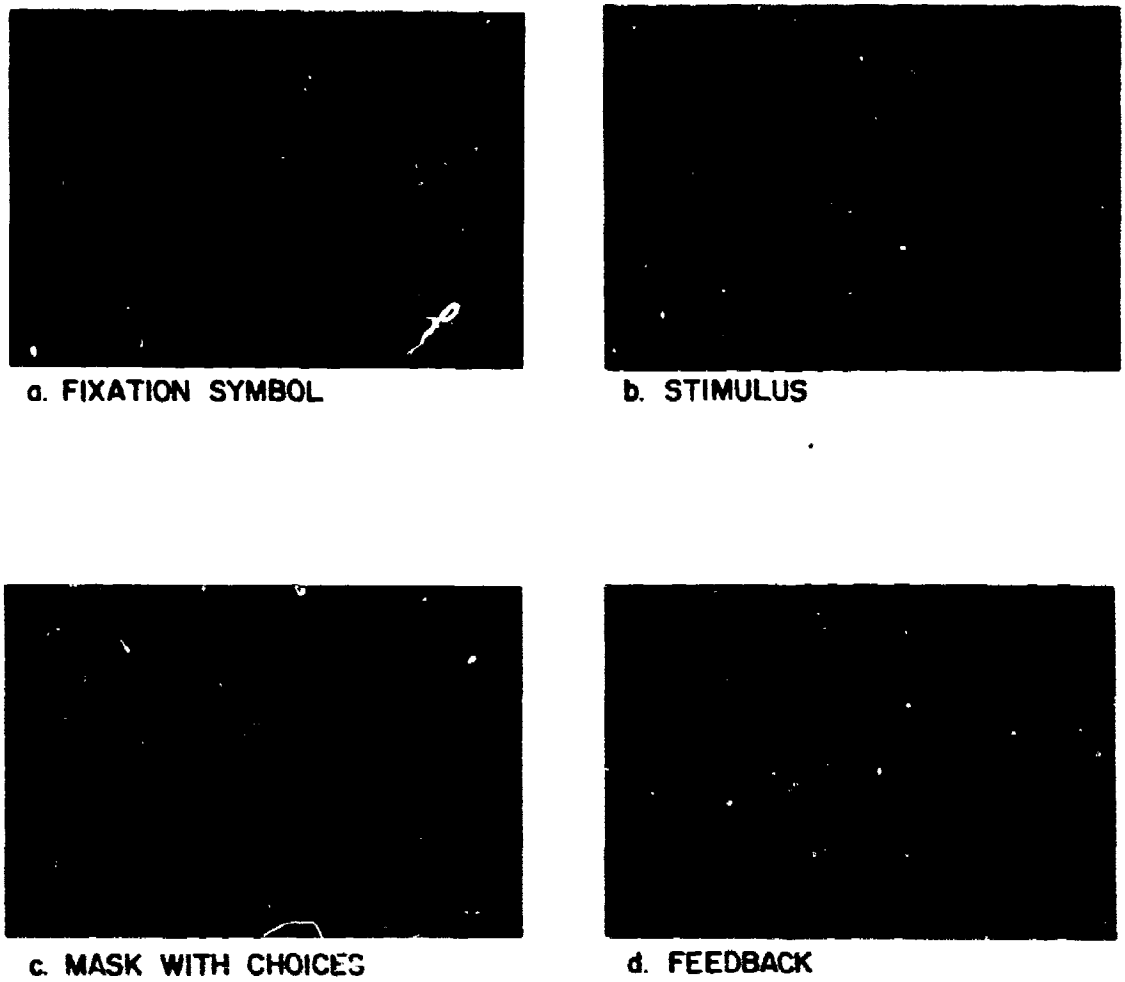


Figure 5. Color recognition trial. Similar to Fig. 4, except that the subject must respond to the stimulus color.

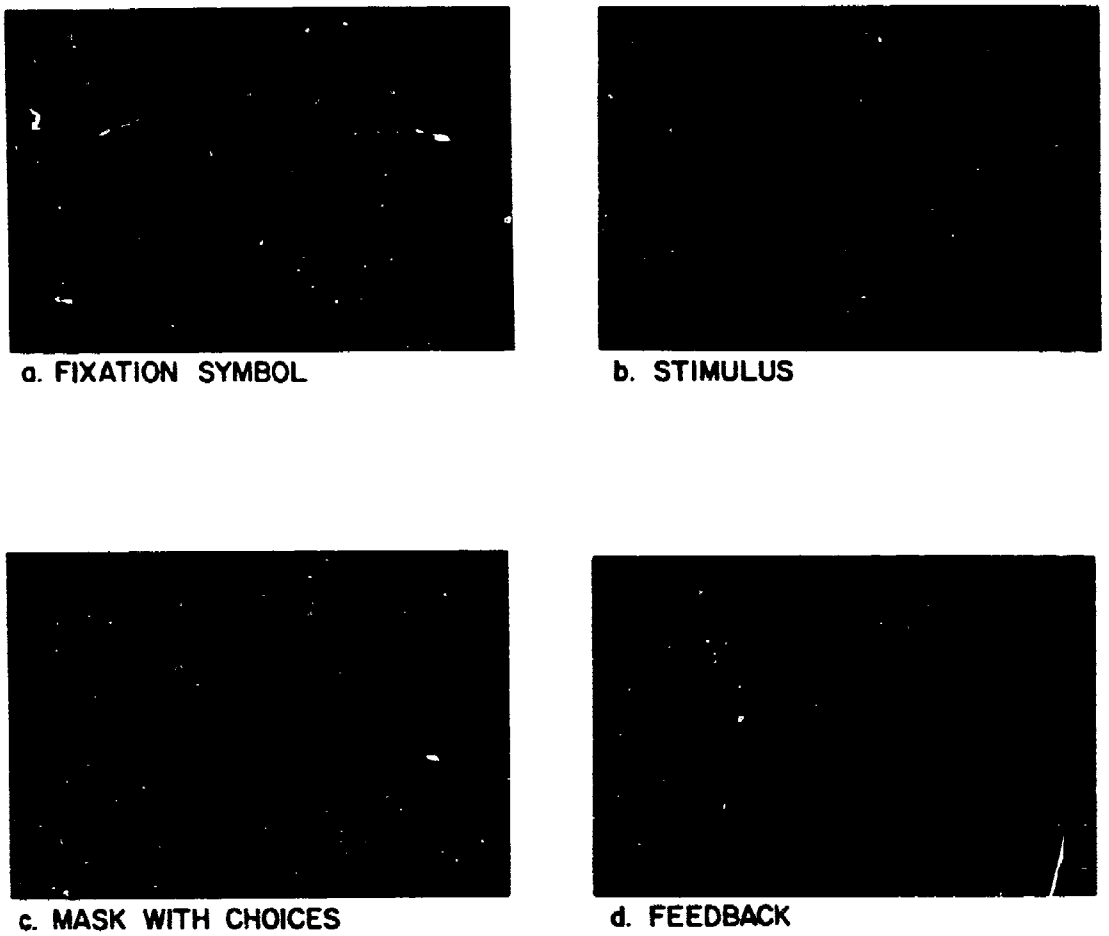


Figure 6. Color-coded shape trial: version A. Color and shape were completely redundant, so that a line of positive slope (b) was always red (PR), and a line of negative slop was always green (NG). The subject indicated which stimulus was presented, by giving the appropriate response (c).

was always green. The complementary experimental phase in which only stimuli 2 and 3 of the alphabet were presented was also run in each session (Fig. 7).

We estimate the desired psychometric functions by assuming that the function is cumulative normal [10]. In this case, the function plotted on probability paper is a straight line. Hence, the problem of estimating the psychometric function reduces to fitting a straight line on probability paper; two points suffice for this. The two points we chose to estimate were the stimulus exposure durations necessary to achieve 70.7% and 84.1% accuracy levels in the two-alternative-forced-choice situation of this experiment. The

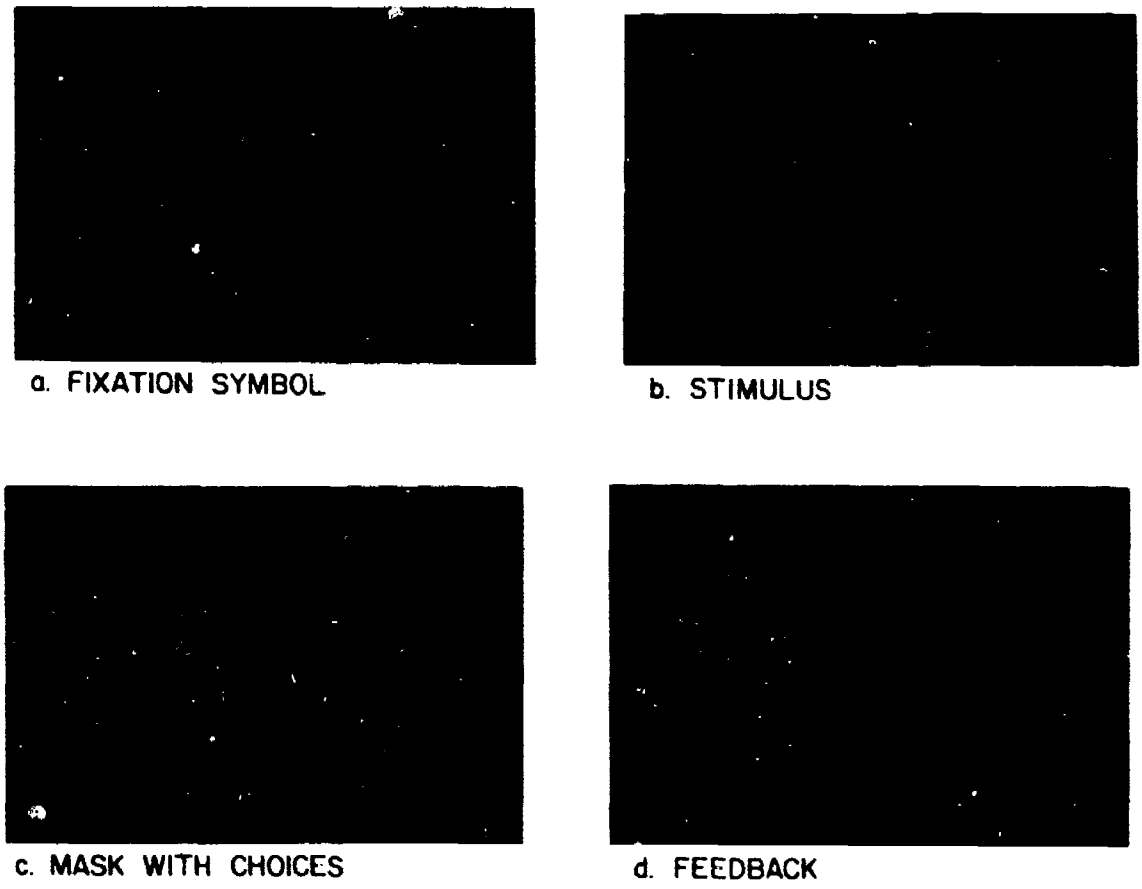


Figure 7. Color-coded shape trial: version B. Similar to Fig. 6, except that the line of negative slope was always red (NR), and the line of positive slope (b) always green (PG).

70.7% and 84.1% points are the most tractable (see Section II.A.2). Estimation trials for the 70.7% and 84.1% points were interleaved randomly.

After a number of practice runs through the experiment, the observer (J.J.M.) submitted to six replications of the experiment. Each data point used in the psychometric-function estimate was based on more than 700 stimulus trials.

3. Results

The experimental outcome, presented on probability paper, is shown in Fig. 8. The curves labeled SHAPE and COLOR, respectively, are the estimated psychometric functions (straight lines on probability paper) for these two attributes in the experiment. The SHAPE curve represents the accuracy level (% correct) as a function of exposure duration for the pure-shape-recognition task. The COLOR curve represents the corresponding information for the pure-color-recognition task.

Applying Eq. (1) to the SHAPE and COLOR curves, we get the predicted psychometric function for color-coded shape recognition. This predicted curve is shown, as are the data points for the color-coded shape task. Since the 84.1% exposure time is significantly longer than that predicted from Eq. (1) (Wilcoxon signed-rank test,* $p < 0.05$), the hypothesis that available color and shape information are independently processed and used in the recognition judgment is not supported. At the 70.7% accuracy exposure duration, only shape information is used (data point falls on SHAPE curve); this is expected, since color information is not available at this brief exposure. Around the 84.1% accuracy exposure duration one cannot decide whether the subject uses only the shape or only the color information of the stimulus, since the 84.1% points for color, shape, and color-coded shape are not significantly different. What is clear, however, is that the observer does not use both color and shape in an average stimulus trial.

Tangentially, it is worth noting that Fig. 8 shows different decision mechanisms to be involved in the recognition of shape and color, as indicated by the significantly different slopes of the color and shape probability plots of the psychometric functions ($p < 0.02$).

*Differences between estimated and observed exposure duration of less than 1 ms were considered not different in the analysis. A one-sided test was used here, since one does not expect durations shorter than predicted by Eq.(1).

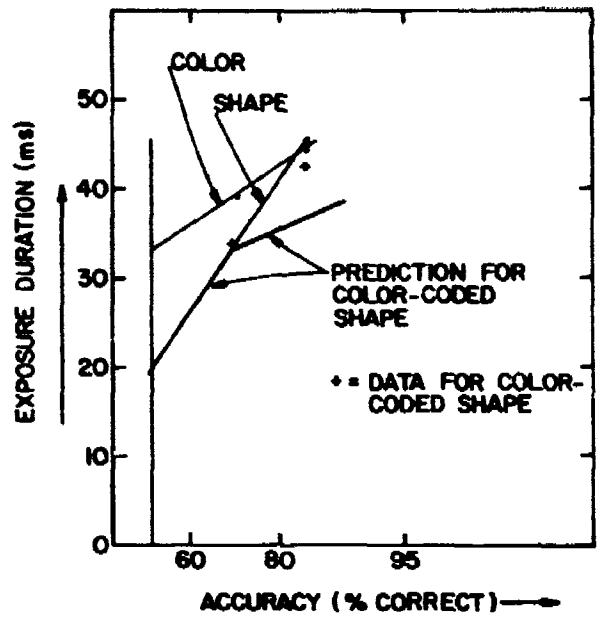


Figure 8. Results for color, shape, and color-coded shape recognition for observer J.J.M. Independent processing of color and shape in a color-coded shape-recognition task is not indicated.

Garner [19] has introduced a distinction between what he calls integral and separable stimulus dimensions. A multidimensional stimulus may have more than one dimension as far as the experimenter is concerned, but may or may not be dealt with as multidimensional by the observer. In Garner's terminology, if stimulus dimensions are dealt with as integral they are naturally brought together by the observer, so that the stimulus is not processed as an item with many attributes, but as a unit which may be decomposed into its attributes. In this case a performance benefit due to increased dimensionality is expected. In terms of multidimensional scaling, a Euclidean metric should hold for integral dimensions. If the stimulus dimensions are dealt with as separable, then the stimulus is, in fact, processed as an item with distinct attributes. In

this case dimensional preferences are expected, making a benefit of increased dimensionality less likely. A city-block metric should hold for separable dimensions.

Garner's concepts are well suited to the results described here and in the preceding section (II.B). Apparently what we have demonstrated in the current experiment, with color-coded shape recognition, is that the dimensions of color and shape are separable. As they are not united by the observer, dimensional preferences do exist. Consequently, a performance benefit due to color coding is not found. The preceding experiment with colored letter and word stimuli can, however, be interpreted as demonstrating that color and shape are integral dimensions, since the stimulus coloring intrudes into shape-recognition decisions. Hence, since the dimensions are integral, one should expect a benefit of increased stimulus dimensionality (provided the color is used in an informative sense).

When these two experiments are viewed together, the potential for color benefit clearly appears to be an issue of the cognitive handling of the available dimensions of color and shape. In the case of alphanumeric stimuli, color and shape appear to be integral dimensions; this would indicate a benefit of color for such display material. For simple shape recognition, color and shape appear to be separable dimensions with a consequent lack of benefit from color coding. This result strongly suggests that the likelihood of color benefit for a particular type of display corresponds to the likelihood of the color and shape information being dealt with as integral, rather than separable, dimensions.

D. INTRUSION OF COLOR INTO SHAPE-RECOGNITION DECISION PROCESSES IN STEADILY VIEWED DISPLAYS

1. Introduction

In the preceding sections we have explored the involvement of color in pattern recognition with brief-exposure stimuli. Since the beginnings of modern experimental psychology the use of brief exposures has been a popular device for degrading the visual stimulus so that systematic errors in performance

might be uncovered, and rules of visual processing thereby revealed [20]. It is in this spirit that we have used brief-exposure stimuli. However, such an approach introduces a temporal issue which quite possibly does not fully simulate the situation of an observer processing a steadily viewed display. For example, a potential complication is that all Fourier components do not reach visual decision centers at the same rate [21,22]. Hence the briefly viewed stimulus might tend to emphasize certain aspects of the stimulus energy spectrum which might be of far less importance under steady viewing.

Performance measures based on reaction time (RT) have become increasingly popular. However, it has recently become apparent that, for various reasons, generalizations based on RT measures are particularly prone to error. First, there is the point about different Fourier components of the stimulus being processed at different rates; this puts undue emphasis on the more rapidly processed components in a speed-stress situation. Differential frequency processing rates were recently demonstrated with an RT experiment [23]. A better-known issue is the complication of the speed/accuracy trade-off which involves subtleties of the interaction between response speed and response accuracy [24,25]. The use of RT measures, therefore, would not seem to offer us any advantage for understanding issues related to pattern recognition in steadily viewed displays.

Contrast-threshold techniques are appropriate for steadily viewed displays. However, their use in the investigation of color involvement in pattern recognition is problematic because the contrast threshold is more properly concerned with the stimulus strength required for visibility, rather than recognizability. In studying pattern recognition we are concerned with the high-level cognitive processing of visible information, as opposed to the sensory processing of information required to render the stimulus visible. The significance of contrast-threshold measures for display descriptors is discussed in Section V of this report.

The fact is that the study of pattern visual information processing for pattern recognition in steadily viewed displays has traditionally involved indirect physiological or physical measurements, such as eye movement and EEG recording, whose interpretation must often be dealt with at a qualitative

level. There does not appear to be a useful paradigm for the systematic psychophysical investigation of the problem. We have, therefore, devised one. Our approach, which has its origins in the work of A. Rose [3], involves the detection of structure in a random dot pattern. An example is given in Fig. 9.

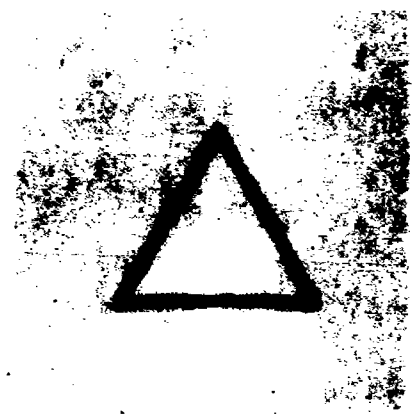
Figure 9a contains a random dot pattern; Fig. 9b, a line drawing of a triangle. Fig. 9c was produced by, in effect, placing the triangle of Fig. 9b over the dot pattern of Fig. 9a so that dots are obscured at the location of the triangle. The triangle portrayed in Fig. 9c can be seen by viewing this figure from a distance of a few feet (a poor reproduction of the figure may make the recognition of the triangle difficult). One can reproduce Fig. 9b from Fig. 9c by increasing the dot density in Fig. 9c so that individual dots are no longer resolved. Conversely, one can go from Fig. 9b to Fig. 9c by a decrease in dot density; one can also decrease the dot density further so as to render the triangle imperceptible. By varying the dot density, we effectively vary the signal-to-noise ratio of the bilevel display of the triangle. A measure of the observer's ability to recognize the triangle in the dot field can therefore be obtained in terms of the dot density required.

Clearly, the perception of the triangle in the dot pattern of Fig. 9c involves high-level cognitive judgments as to which dot arrangements are sufficiently haphazard to be considered noise and which ones appear sufficiently regular to be considered signal. These are active judgments concerned with pattern recognition, as opposed to pattern visibility. We are addressing issues of pattern recognition with this type of pattern, and we are addressing them by means of a steadily viewed display.

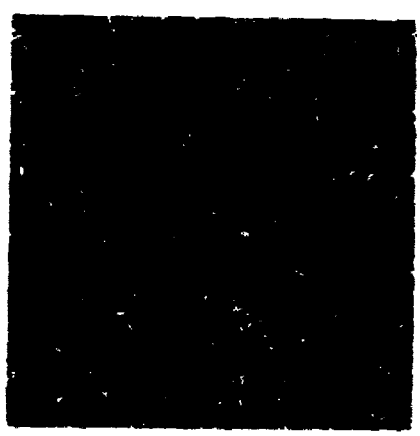
We have, therefore, a type of visual stimulus which, when steadily viewed, can either reveal a pattern or not, depending on the value of the dot density. The task for the observer is one of pure shape recognition. The question we will pose is, will color influence this ability?

All of the information appropriate for perception of the pattern portrayed in the dot field is contained in the presence or absence of dots, i.e., in the luminance profile of the dot field. As long as the dots are visible, their color is irrelevant. More to the point, if we color the dots without modifying the luminance profile of the display we do not alter the signal-to-noise ratio

(a)
RANDOM DOT FIELD



(b)
TRIANGLE TEMPLATE



(c)
TRIANGLE PORTRAYED
IN DOT FIELD

Figure 9. Portrayal of shape in a random dot pattern. The triangle (b) covers dots of the random dot pattern (a) so that the triangle shape is portrayed (c).

of the portrayed pattern. Hence, if we randomly color the dots of the type of display given in Fig. 9c we introduce random color noise without modifying the luminance signal-to-noise ratio of the pattern. By introducing this type of color noise we neither modify the observer's task, nor do we detract from the information available for shape recognition in the dot field. What we do is to open up a new dimension of the display, a dimension of color variation. The observer will operate efficiently if he disregards this dimension. It is not certain that he can. The degree to which noise in the color dimension intrudes into the shape-recognition processing that operates on the dot field's luminance profile gives us a measure of the impact of color information on shape recognition in steadily viewed displays.

2. Experimental Procedure

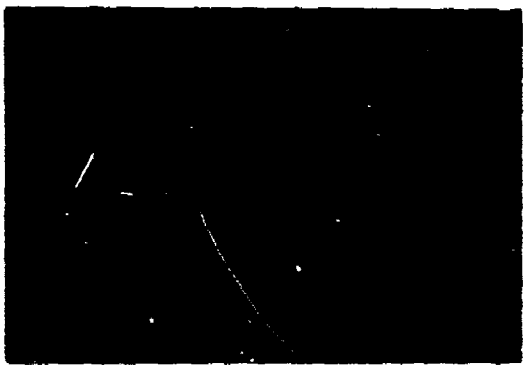
A typical stimulus trial is shown in Figs. 10 and 11. A "ready fixation" symbol, in yellow, was first displayed on the monitor (Figs. 10a and 11a). When the subject was ready to initiate a trial, a yellow, base-down triangle was presented in a random screen location for 500 ms (Figs. 10b and 11b). This triangle was presented to serve as a visual template for the immediately following steady display of a random dot field (Figs. 10c and 11c). The subject's task was to inspect the random dot field for as long as desired to decide whether or not the just-seen triangular shape was portrayed in the dot field. The triangle was equally likely to be present or absent; if it was portrayed in the dot field, its location was random. After the subject pressed a button to indicate presence or absence of the triangle, a yellow accuracy-feedback message was presented on the screen (Figs. 10d and 11d). The reason for randomizing the location of the template triangle and the subsequent dot-field portrayal of the triangle was to avoid the likelihood of enhanced appearance of the dot-field triangle because of location congruence with the template, which might allow the observer to bypass the dot-field analysis.

Two triangle sizes were used; a smaller triangle was about 1.2° on a side (Fig. 10b), and a larger triangle twice this size (Fig. 11b). Three versions of dot field coloring were used; the field could contain only red dots, only green dots, or a red-green mixture composed of an equal number of red and green

a. READY (FIXATION) SIGN

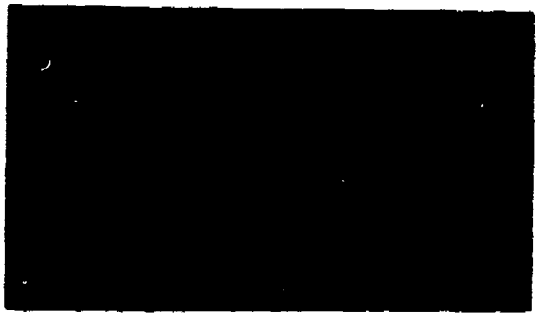
b. STIMULUS TEMPLATE

c. DOT FIELD INSPECTED FOR PRESENCE OF TRIANGLE



d. FEEDBACK

Figure 10. Small-triangle detection trial. The fixation symbol (a) is followed by a quick view of the triangle shape to be detected (b). The observer inspects the dot field (c) and then gives a button response to indicate whether or not he thinks the triangle is present. Accuracy feedback (d) follows the response.



a. READY (FIXATION) SIGN

b. STIMULUS TEMPLATE

c. DOT FIELD INSPECTED FOR PRESENCE OF TRIANGLE

d. FEEDBACK

Figure 11. Large-triangle detection trial. Similar to Fig. 10, except that the triangle size is doubled.

dots. The dot luminance (10 mL) did not depend on coloring. The two triangle sizes were used in separate back-to-back sessions for each observer. The ordering of which triangle size was used first was balanced across the observers. For each triangle-size session, the threshold number of dots required for triangle detection in the various colored dot fields was determined (71% accuracy level with 500 dot step size, determined by our standard sequential estimation procedure). The threshold determination trials for the red, green, and red-green-mixture dot fields were randomly interleaved in the sessions.

The experiment consisted of the two triangle-size sessions, each preceded by a short set of practice trials. This experiment took about an hour. Seven observers participated.

3. Results

The threshold number of dots required for triangle detection (average across observers) for the two triangle sizes as a function of dot colorings are given in Table 5.

TABLE 5. THRESHOLD DOT NUMBER FOR VARIOUS TRIANGLE SIZES AND DOT COLORINGS

		Triangle Size	
		SMALL	LARGE
Dot Color	Red:	2678	1857
	Green:	2571	1809
	Red-Green Mixture:	2643	2274

The difference for either size of target pattern, whether all-red or all-green dots were used, was not significant. The use of a monochrome dot field does not introduce a dimension of color variation, hence no difference is to be expected regardless of whether all-red or all-green dots are used.

For the red or green monochrome dot fields, the threshold dot density for pattern detection is expected to increase as triangle size decreases. However, the observed increase in density threshold is somewhat less than anticipated if the decision processes are the same for detecting both the large and small triangles. In inspecting the dot pattern for the larger triangle, for example,

one would assume the observer to be looking for some critical, triangle-defining dot-vacant region. Suppose such a potential signal region has area A. In the noise (signalfree) portions of the dot field a region of area A has, on the average, n dots with standard error (rms fluctuation) $\sqrt{n}$. Hence, a region of area A in the noise can contain $n \pm \alpha \sqrt{n}$ dots, with the probability of a particular dot number decreasing as α increases. The reader may recognize the expression $n \pm \alpha \sqrt{n}$ as the confidence interval about the mean value n at probability level determined by α. We may assume that, in trying to decide whether what looks like a dotfree region is not a fluctuation in the noise, the observer adopts a decision criterion, which amounts to specifying some value of α in the confidence interval. For a dot-vacant signal region to be distinguished from a noise fluctuation, the dot density must be sufficiently high so that $n - \alpha \sqrt{n} > 0$ for the observer-specified value of α. At threshold for signal detection, $n_t - \alpha \sqrt{n_t} = 0$, or $n_t = \alpha^2 = \text{constant}$.

Suppose, for detection of the larger triangle, the noise dot density is set at the threshold dot density of n_t dots per region of area A, and the smaller triangle is the signal pattern portrayed in the dot field. For the small pattern, the triangle-defining regions of interest will have area A/2, which contain, on the average, $n_t/2$ dots in the noise. Then the triangle will be below the detection threshold. For the smaller triangle to become apparent, the number of dots in the noise field should be double the number necessary to detect the larger triangle.*

However, the data indicate that, in going from the larger to the smaller triangle, the threshold dot number increases only by a factor of about $\sqrt{2}$, rather than a factor of 2. Hence, if the decision processes are the same for perceiving the triangles of either size, the smaller triangle is easier to detect than would be expected (or the larger triangle is more difficult to detect). One may conclude that the decision-making process is somewhat different when one is looking for the smaller triangle than when one is looking for the larger one.

The main object of this experiment was to determine if pure color noise intrudes into a shape-recognition task that depends solely on the processing

*This assumes that there is no inherent limitation of visual processing that prevents the inspection area, A, to be as large as the larger triangle.

of the luminance profile of the pattern. Such an intrusion effect is demonstrated for the larger triangle. The number of dots required for detection of the larger triangle in the monochrome dot field is 1833 (average of red and green thresholds), whereas for a red-green-mixture dot field 2274 dots are required. This increase in threshold dot number is significant (Wilcoxon signed-rank test, $p < 0.05$). From this it follows that the dimension of color variation may intrude into shape-recognition judgments. This outcome is consistent with results presented in Section II.B (above) that pertained to brief exposures for word and letter recognition. Consequently, if one makes available a dimension of color variation in a display, it potentially involves itself automatically in recognition judgments in the sense that no attention to its presence is required; indeed, it is not readily disregarded.

That it is the intrusion of the dimension of color variation that is involved, and not simply a tendency to attend to only one color at a time, is seen from the relationship of the dot number thresholds in the monochrome and multicolored situations. Given that 1833 dots are required for detection of the larger triangle in a monochrome dot field, one would expect an observer who simply attends to one color (red or green) in picking out the large triangle in the multicolored dot field, to require about 3666 dots for detection, i.e., 1833 dots in each color. Since only 2274 dots are required, a strict color-attention hypothesis is rejected.

While a color intrusion influence is found for detection of the larger triangle, it is absent for the smaller triangle. About 2625 dots were required for perception of the smaller triangle in the monochrome dot fields (average of red- and green-field thresholds) and about 2643 dots were required for small-triangle detection in the multicolored dot field. There is no significant difference between the dot number thresholds for the monochrome and multicolored dot fields insofar as detection of the smaller triangle is concerned. All subjects, when asked after completion of the experiment if they were aware of the multicolored nature of the red-green-mixture dot fields, reported that the multicoloring was quite apparent. Yet, the visible color variation dimension did not intrude into shape-recognition judgments for the smaller triangle.

As was pointed out in the discussion of triangle detection results for monochrome dot fields, there is reason to suspect that visual processing for

perception of the smaller triangle is somewhat different from that for the larger triangle. This is suggested by the relatively enhanced detectability of the smaller triangle as compared to what is expected from application of signal-detection-theory considerations. It appears likely that the relatively enhanced detectability of the smaller triangle in monochrome dot fields is related to the lack of color intrusion for the shape detection in the multicolored dot field. In particular, a likely source of these effects is that in processing for detection of the smaller triangle the dot field is globally processed as a unitary texture, in contrast to a local analysis for dot-field regularities which is presumably carried out for detection of the larger triangle.

That unitary texture processing can have dramatic influence is demonstrated by an effect we have produced using contrast reversal on a triangle-portraying dot field. In Fig. 12 we show the same dot field portraying the triangle, one field with white dots on black background and the other field with black dots on white background. It takes some adjustment of viewing distance, and some inspection time, to perceive the triangle in either dot pattern, but in any case the portrayed shape is decidedly more salient in the white-dots-on-black-background dot field. The explanation of this effect would take us too far afield from the current discussion [26]. The point we want to make is that this effect is a consequence of global texture processing and serves to illustrate that the consequence of such processing is not necessarily subtle.

If, in processing the multicolored dot field for detection of the smaller triangle, the dots are dealt with as a unitary texture, it is necessary for the observer to cognitively reject color differences that are visible so as to perceptually tie together all dots as members of the dot texture. If this is accomplished, then we would expect color variations in the dot field to be of no consequence. That is, detectability of the triangular shape should be no less difficult in the multicolored dot pattern than in the monochrome dot pattern. Furthermore, we might expect this global, interactive type of processing, which emphasizes cooperative influence among picture elements (dots), to yield better detectability than would be accomplished by a linear detector process of the type outlined in the statistical decision model described earlier. Why this global texture processing, with its probable effect

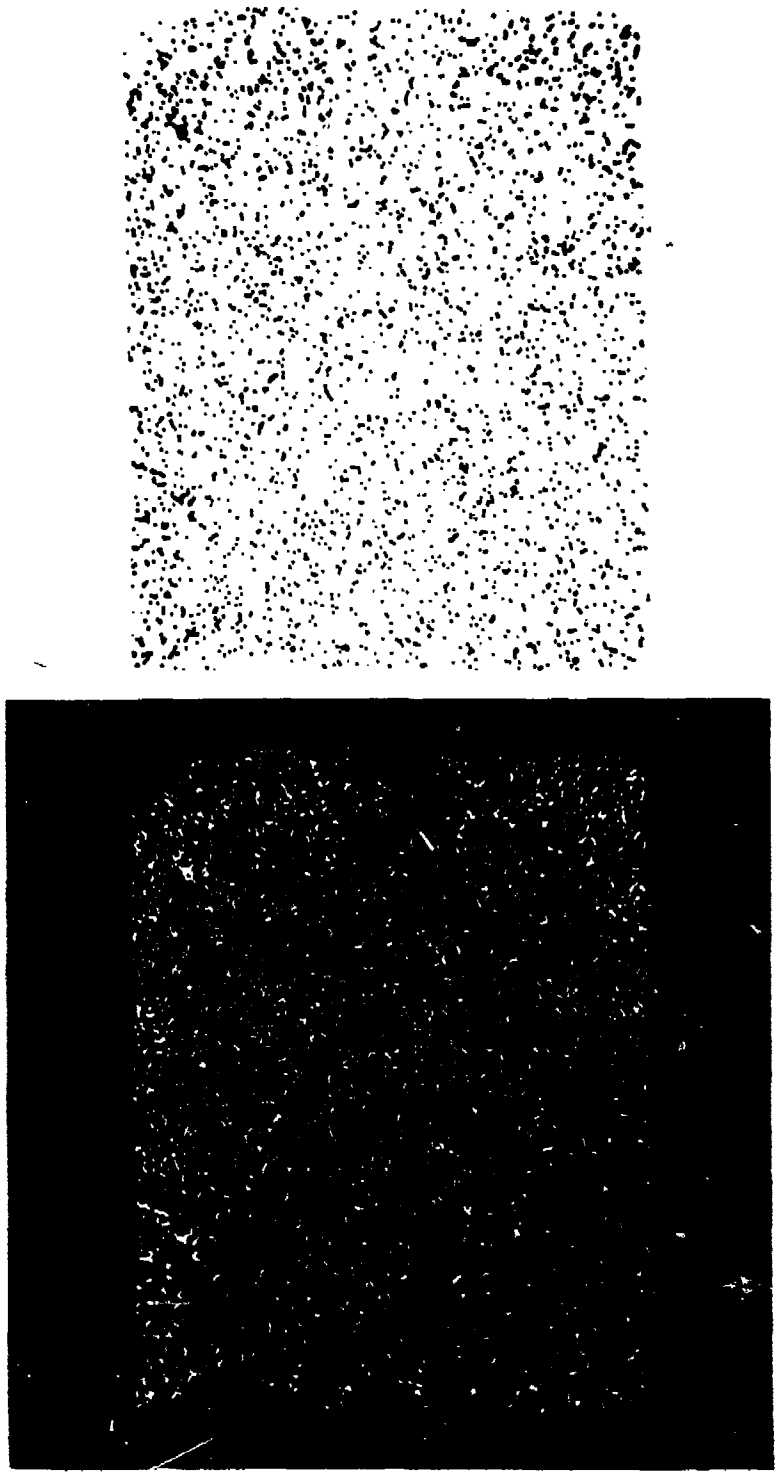


Figure 12. The contrast-sign effect (see text).

of preventing color intrusion and its probable enhancement of shape detection, should be evident for the small-triangle detection task and not for the larger triangle is a fundamental question that must be considered a topic for separate investigation.

However, we can say with some confidence that there are available to the observer local and global modes of visual processing for shape recognition. For local modes of processing, straightforward statistical decision-theory models of the shape-detection process should be appropriate. Of particular significance in local processing is the fact that intrusion of information available in a color dimension is expected for decision making about shape. Whether this yields a benefit or a deficit in performance presumably depends on whether the color information is relevant (as a redundant code) or irrelevant (as distracting noise). For global modes of processing, it appears that one can expect a somewhat enhanced recognition performance as compared to the local mode, but then the availability of a color dimension would seem to serve no purpose.

E. CONCLUSIONS

Our approach to determining whether a performance benefit, in terms of stimulus recognition, results from the color coding of shape, has been to look for some guiding principles, rather than find specific answers to specific problems. The principles that appear to follow from our work are these: When color and shape are treated as integral dimensions, or when the stimulus is locally processed, a benefit from stimulus color coding can be expected. When color and shape are treated as separable dimensions, or when the stimulus is globally processed, the availability of color coding may be irrelevant.

Hence, a universal answer to the question concerning the utility of color for pattern recognition is not expected. However, if one can determine, for a particular display-observer situation, the likely information processing mode, in terms of integrality vs separability of color and shape, or in terms of local vs global stimulus processing, a statement about the likelihood of color benefit can be made.

While our experiments do not allow us to characterize generally when integral dimension handling or when local processing will occur, we can make some reasonable inferences. We found, for simple shape discrimination (Section

II.C), that color did not enter into the discrimination judgments, whereas for word and letter recognition (Section II.B) color could not be disregarded. This suggests that the more complex the information to be processed, the more likely it is that color and shape will be handled as integral dimensions.

Using the Advanced Integrated Display System (AIDS) as an example, one might expect that color would be relatively unimportant for the right-hand engine-management screen, since it consists of relatively simple shapes arranged in a manner that places a minimal processing burden on the pilot. On the other hand, more complex displays, such as the lower moving-map or the upper, vertical-situation display, involve more than cursory simple-shape processing and might, therefore, benefit more from color coding.

Roughly speaking, one can distinguish local and global processing based on whether the observer inspects display details (local processing) or tends to take in the display as a whole (global processing). The engine-management display of the AIDS provides an example. The intent of this display is to allow the pilot to avoid inspecting display details, but rather glance at the display for malfunction indication - such as misaligned arrows. One may say that the display is meant to be globally, rather than locally, processed. In this case one would, therefore, anticipate a negligible benefit from color coding. On the other hand, the vertical-situation display is used for local processing, as when the signal carat indicates an item of interest. Here one would expect a benefit of color.

III. STATISTICAL PROPERTIES OF THE LUMINANCE AND CHROMINANCE VARIATIONS IN NATURAL SCENES

A. INTRODUCTION

Our approach in describing the performance of displays has been to treat the display-observer system as elements of a noisy communication channel. Generally, in order to maximize the amount of received information* and minimize the display cost, complexity, size, and other factors, an attempt is made to match the performance of the display-observer system to the properties of the information to be transmitted. However, in most practical systems a unique characterization of the information is not possible because a fraction of the information transmitted is constantly changing in an unpredictable manner, and because other random processes, such as noise, are inevitably present. Thus, it is possible only to describe the signal characteristics as random processes whose statistically averaged properties are known, or can be estimated. These average properties can then be used to describe the most likely system response for a given set of display-observer parameters.

In the development of the display descriptors presented in TR1, TR2, and Section V of this report we have assumed, based on preliminary experimental evidence presented in TR1, that statistical estimates for images can be properly incorporated into a mathematical framework to predict accurately the performance of the display-observer system. In Section IV further experimental results that support this assumption will be presented.

Our mathematical display descriptors require that the ensemble-averaged power spectral densities be known for both the luminance and chrominance variations in natural scenes.** It is also required that the absolute magnitudes of these power spectral densities be known. For luminance information,

*The form of the "information" must be defined according to the specific requirements of the channel. In both TR1 and TR2 this issue is discussed in detail with respect to displayed information.

**By "natural scenes" we mean those pictorial scenes that are of general interest to nonspecialized observers. This definition loosely defines one of many possible subsets of pictorial information. Other subsets, for example, are alphanumerics, line drawings, the art work of Jackson Pollock, and x-ray negatives. Each of these subsets of information has unique statistical properties that can be incorporated into our display formalisms.

the scale of modulation depth was determined by measuring the ensemble-averaged ratio of the rms luminance variations to the average luminance for a large number of individual scenes. For chrominance information the ensemble-averaged rms value for chrominance variations was measured in a calibrated system as the fraction of color saturation. These quantities are discussed in more detail in Section V.

In a strict sense it is possible to determine meaningful average properties for random processes only when the signals are stationary [27]. Although the signal sources discussed in this report are certainly not stationary, it will be shown that there is sufficient self-similarity among the scenes studied to warrant strong general conclusions to be drawn about their properties.

B. MEASUREMENT TECHNIQUES

Our general approach was to utilize as the signal sources the electronically generated luminance and chrominance signals both from off-the-air commercial television and from a three-vidicon television camera viewing 35-mm color slides. Television represents an appropriate means of producing these signals because, as part of the encoding process for transmission, it separates the pictorial information into three parts: a luminance signal, and two, roughly orthogonal, constant luminance chrominance signals. It will be shown in Section IV that signals of this form are required for a complete description of the information-handling capabilities of the human visual system.

The chromaticity signals we examined are shown in Fig. 13 as lines I_C and Q_C on a CIE chromaticity diagram. The solid straight lines show the chromaticity paths for systems with unity gamma. The curved dotted lines show the chromaticity paths for transmitted signals with a gamma of 1/2.2. In both cases the I_C and Q_C axes pass through the white point on the chromaticity diagram. Also shown on Fig. 13 are the NTSC red, green, and blue primaries. The saturation values given in this report are defined as percentages of these coordinates.

The chrominance signals were obtained by demodulating the appropriate video signals with a commercial vectorscope. This device has the advantage of allowing any chrominance axes through the white point to be easily selected.

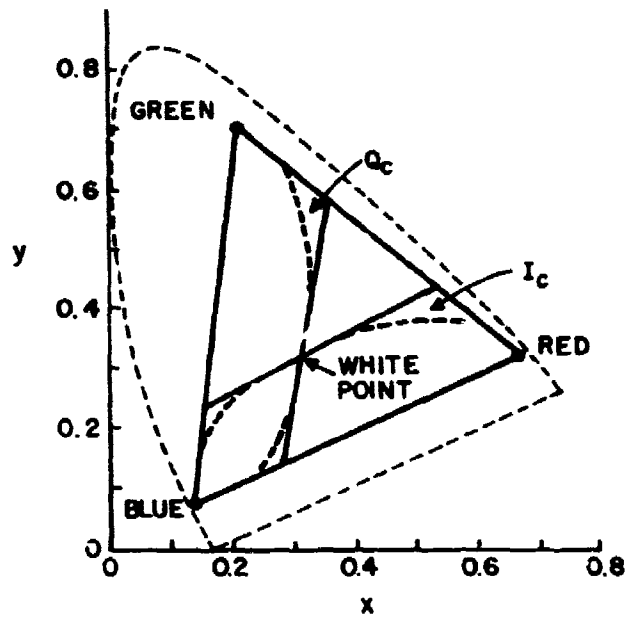


Figure 13. CIE chromaticity diagram showing the locations of the chromaticity axes I_C and Q_C. The lines through the white point represent chrominance paths for signals with unity gamma; the dotted lines through the white point indicate the chrominance paths for a gamma of 1/2.2. Also shown in this figure are the NTSC color primaries.

The I_C and Q_C axes were chosen for two reasons. First, these are the axes used in commercial television for transmitting the chrominance signals.* Second, during preliminary experiments it was found that the power distribution about the white point of Fig. 13 was highly anisotropic. The measured power distribution was roughly that of an ellipse with the major axis along I_C and the minor axis along Q_C. Thus, it was necessary to measure only the statistical properties along these two directions, since statistical estimates for other directions can be made from these values. The bandwidths for both the I_C and Q_C axes were approximately 0.5 MHz (-3 dB); the bandwidth for the luminance signals was 3 MHz (-3 dB).

For both the luminance and chrominance signals, all aspects of the waveforms that were not part of the actual scenes, such as color burst, sync, and blanking, were removed. Further, the total luminance, I_T, and chrominance, C_T,

*In the television literature they are known simply as the I and Q channels [28].

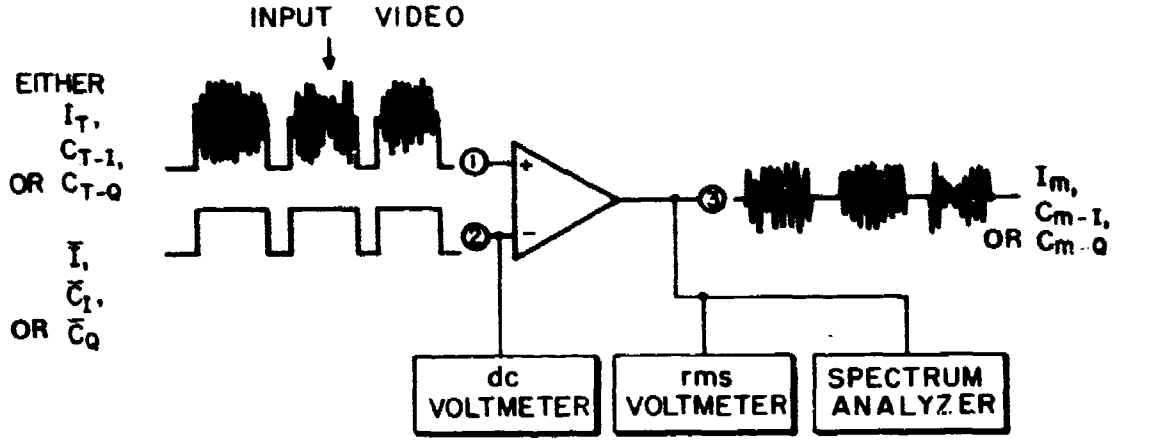


Figure 14. The luminance and chrominance signals were decomposed into an average term plus a modulation term. That is, $I_T = \bar{I} + \bar{I}_m$, $C_{T,I} = \bar{C}_I + C_{m}I$, and $C_{T,Q} = \bar{C}_Q + C_{m,Q}$, where $\bar{I}$ is the average luminance for the scene and C_I and C_Q are the average percent saturations for the scene along the I_C and Q_C axes, respectively. The properties of the average terms and the modulation terms were measured separately, as shown.

signals were decomposed into their respective average values, $\bar{I}$ and $\bar{C}$, and modulation terms, I_m and C_m. This operation is shown schematically in Fig. 14. $C_{T,Q} = \bar{C}_Q + C_{m,Q}$ or $C_{T,I} = \bar{C}_I + C_{m,I}$, depending on whether we are referring to the Q_C or I_C axes, respectively.

All of the power spectral densities in this report were obtained from the modulation terms I_m, $C_{m,Q}$, and $C_{m,I}$. This approach has the important advantage of minimizing the effects of the limited size of the scanned "window" of a scene on the measured power spectral densities. This is because the power spectral densities for the $\bar{I}$, $\bar{C}_Q$, and $\bar{C}_I$ terms are of the form $\text{sinc}^2(\omega w/2)$ (where w is the picture width), which rolls off at high frequencies as $1/\omega^2$ (an example of this spectrum is shown in Fig. 19). Since for scenes with relatively small modulation depths the power spectra from the average terms can mask details in the modulation power spectra, it is important that the average term be removed before the power spectra are obtained.

For both the luminance and chrominance signals the video bandwidths and signal-to-noise ratios were sufficiently high for their effects on the measured statistical signal properties to be negligible. The input signal amplitudes were normalized by maintaining a constant-peak-white-signal amplitude and a

constant-color-reference amplitude. The system gamma was approximately 1/2.2 for the commercially transmitted signals and approximately 1.0 for the 35-mm slides used in this study. The actual gamma of the television camera used to view the 35-mm slides was 0.6, but because the gammas for the color slides varied between roughly 1.4 and 2.2, the effective gamma of the overall system was close to 1.0. The effect of the different gammas on the chromaticity results can be seen in Fig. 13.* Because the average saturations of the images studied were relatively small, and because the perceptual space approximated by the CIE diagram is highly nonlinear, we conclude that the effects of the different gammas on the measured chromaticity results are insignificant.

The measurements were performed with the apparatus shown schematically in Fig. 14. For each image studied the amplitude of the average term (either $\bar{I}$, $\bar{C}_Q$, or $\bar{C}_I$) was determined by adjusting its value until the reading on the rms voltmeter was a minimum. The average rms value (either $\sqrt{\overline{I_m^2}}$, $\sqrt{\overline{C_{m,I}^2}}$ or $\sqrt{\overline{C_{m,Q}^2}}$) was then read from the appropriately calibrated dial.

The power spectra were obtained with an ac-coupled video spectrum analyzer. The spectra produced by a television video signal consist primarily of terms at discrete frequencies that are integer multiples of the horizontal line frequency ($\approx$ 15.8 kHz). Thus, the power spectra presented are limited at the low end to a frequency proportional to the inverse of the picture width, and at the high end by the bandwidth and signal-to-noise ratio of the system. Additional details about the properties of video spectra can be found in Ref. [29].

The 35-mm color slides used in this study were selected to represent a diverse range of subject material. We assume that the quality of the slides was sufficiently high for the imperfections in the slides not to affect our results significantly, although it is apparent that the saturation and contrast of actual images would exceed those of our slides.

For convenience, the 35-mm slides were organized into four categories: single-object scenes, two-object scenes, three-to-ten-object scenes, and crowd scenes. These classifications represent, roughly, the number of prominent

*All the luminance measurements were performed with 35-mm slides.

objects in the scenes. It should be emphasized that these classifications are only approximate and that, especially for the chrominance results, the power spectra should be inspected individually.

C. RESULTS

1. Average rms Modulation Depth for Luminance Information

The results of our measurements on the normalized rms modulation depth for luminance variations in natural images is shown in Table 6. Also shown, as entry No. 34, is the result for a page of newsprint. The measured results have all been normalized by dividing the rms luminance values for each scene by their respective average luminance values.

As indicated in the table, the normalized rms modulation depths for the individual slides studied ranged between 0.28 and 0.87, with an average value of 0.56 (excluding slide No. 34). The average value indicates that natural scenes are highly modulated. For comparison consider a scene composed of alternate black-and-white strips of equal width. The fractional rms modulation depth for this scene is 1.0.

In observing the video waveforms of these and many other scenes it was noted that the actual distribution of luminance levels is not uniform from black to maximum white. In general, the highlight portions of the scenes occupied a smaller fraction of the total area of the scenes than did the low-lights. By inspection it was estimated that the highlight amplitudes were roughly three-to-four times the average luminance amplitudes. This observation explains, in part, the large fractional modulation depths reported above.

Finally, the fractional rms modulation depth was measured for a page of newsprint and found to be 0.15. This value is considerably lower than that for any of the natural images studied. It is due to the relatively infrequent occurrence of printed alphanumeric figures.

2. Luminance Power Spectral Density Measurements

The measured luminance power spectra data are shown in Figs. 15-18. The numbers on the figures corresponding to specific data points refer to their respective scene description in Table 7. The average results for each of the four general classes of images is shown in Fig. 19.

TABLE 6. LIST OF SCENES AND THEIR RESPECTIVE NORMALIZED rms LUMINANCE MODULATION DEPTHS

No.	Description		$I_{\sigma}^2/\bar{I}$
1	Apartment building		0.87
2	Girl standing in room		0.86
3	Girl with RCA sign		0.86
4	Girl watching TV		0.80
5	Aeroplane and two people		0.76
6	Lady with checker board		0.76
7	Girl on striped blanket		0.72
8	Four people on beach		0.69
9	Man and women standing		0.66
10	Man and aeroplane		0.66
11	Red zinnia		0.66
12	Girl and dotted background		0.63
13	Crowd of Indians		0.60
14	Ten people		0.59
15	Girl and duck		0.57
16	Man and woman in room		0.56
17	Power lines		0.51
18	Girl in country		0.50
19	Stadium crowd		0.47
20	Fruit basket		0.47
21	Motel sign		0.47
22	Standing girl		0.47
23	Head of blond lady		0.45
24	Face of young girl		0.44
25	Manikin		0.43
26	Girl and tree		0.42
27	Soap box		0.42
28	Lady in kitchen		0.41
29	Fruit basket		0.39
30	Dog on grass		0.35
31	Four people		0.34
32	Aeroplane and mountains		0.31
33	Head of girl		0.28
		Average	0.56
34	Page of newsprint		0.15

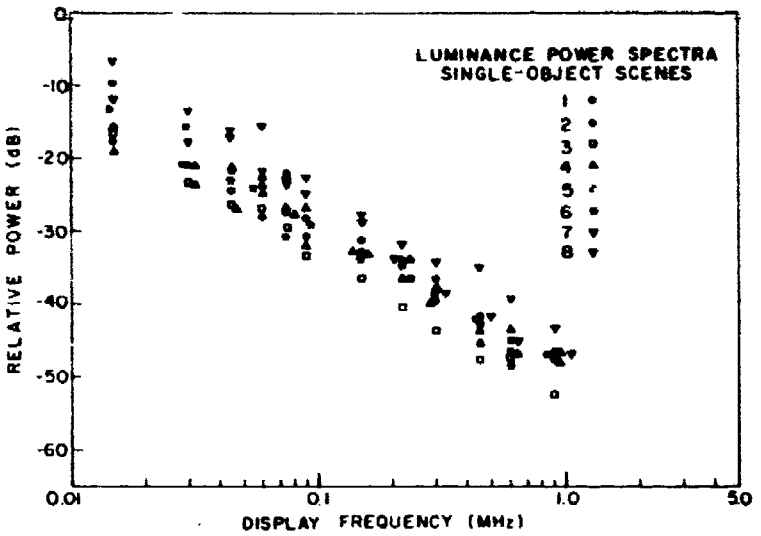


Figure 15. Average horizontal luminance-modulation power spectra for individual scenes that contain one prominent object. The numbers next to the symbols in the legend correspond to the scene numbers of Table 6.

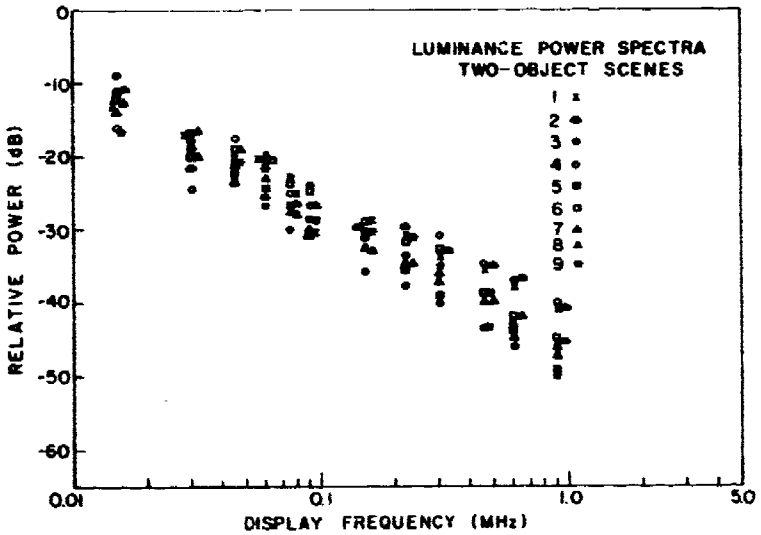


Figure 16. Average horizontal luminance-modulation power spectra for individual scenes that contain two prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 6.

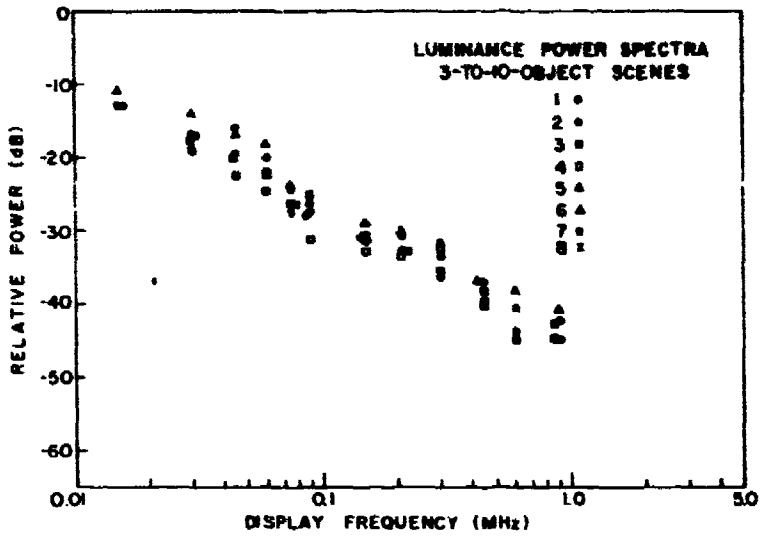


Figure 17. Average horizontal luminance-modulation power spectra for individual scenes that contain three-to-ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 6.

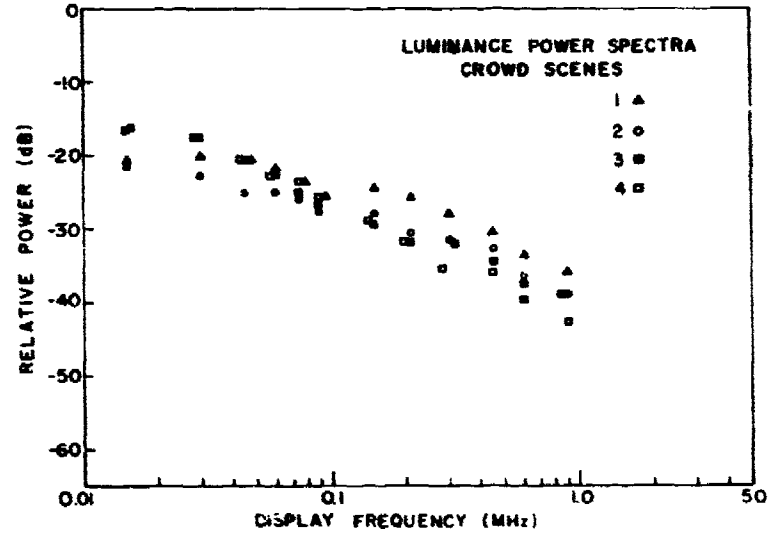


Figure 18. Average horizontal luminance-modulation power spectra for individual scenes that contain more than ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 6.

TABLE 7. LIST OF SCENES USED TO DETERMINE THE LUMINANCE POWER SPECTRA

Types of Scene	No.	Description
Single-Object Scenes	1	Girl in hat
	2	Red zinnia
	3	Apartment building
	4	Head of girl
	5	Manikin
	6	Soap box
	7	Head of girl
	8	Head of young girl
Two-Object Scenes	1	Child in sandbox
	2	Church in country
	3	Aeroplane and two people
	4	Girl on blanket
	5	Santa doll and toy
	6	Girl with RCA sign
	7	Girl with duck
	8	Girl with sunflower
	9	Man and woman
Three-to-Ten-Object Scenes	1	Nine people in a row
	2	Two people in a room
	3	Table with fruit
	4	Group of flowers
	5	Beach scene with four people
	6	Fruit bowl
	7	Girl watching TV
	8	Lady in a room
Crowd Scenes	1	Stadium crowd
	2	Field of tulips
	3	Crowd of Indians
	4	Line drawing with type

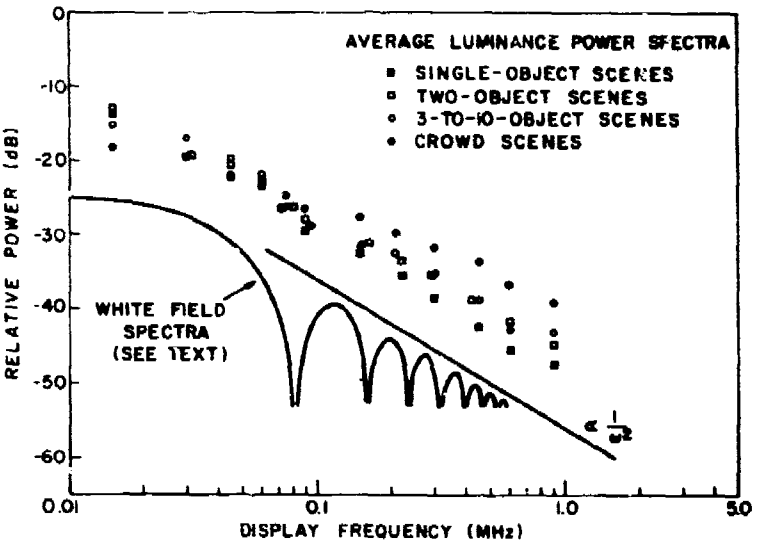


Figure 19. Luminance-modulation power spectra obtained from the average results of Figs. 15 through 18. At high display frequencies, all spectra roll off as $1/\omega^2$. Also plotted on this figure is the spectrum (vertical scale arbitrary) for the average luminance term $\bar{I}$.

The data points of Fig. 19 support the general conclusion of TR1 that at high frequencies the power spectra for natural scenes rolls off as $1/\omega^2$. As expected, however, Fig. 19 also indicates that the lowest frequency for which this relationship applies is a function of the scene content. As the scenes increase in complexity, from single-object scenes to scenes that contain many prominent objects, the frequency above which the spectra roll off as $1/\omega^2$ becomes progressively higher. This transition point, or low-frequency cutoff, ω_L, is roughly proportional to the inverse of the dominant length scale of the images under investigation. For off-the-air video we have found that the cutoff frequency is well approximated by $\omega_L = 2\pi/\text{picture width}$ (TR1).*

As described in TR1, we feel that the $1/\omega^2$ roll-off strongly suggests that edge transitions represent a significant feature of natural scenes. This conclusion follows from the fact that the power spectra for an ensemble of edges, randomly placed and with random amplitudes, will also have a high-frequency roll-off of the form $1/\omega^2$. The significance of edges can also be convincingly established by observation of one's environment or by viewing on an

*Note, however, that for all imaging systems the power spectra must have their maximum value at dc [25].

oscilloscope the waveforms of luminance signals from a television camera. In general, textural variations are of much lower modulation depth than are edge transitions.

3. Average rms Modulation Depth for Chrominance Information

Our findings on the average rms modulation depths for chrominance information are summarized in Table 8. The averages shown in this table for 36 slides were obtained, with few exceptions, from the list of slides given in Table 9. The off-the-air scenes were chosen to illustrate the range of possible chrominance values experienced on commercial television. They are not intended to represent an ensemble average of all scenes that are broadcast. Each entry in the table, for off-the-air scenes, is the average of approximately 15-min segments. It should be realized that the saturations of individual scenes within this averaging period varied from practically zero to almost 100%.

Several general conclusions may be drawn from the results presented in Table 8. First, for natural scenes the rms distribution of saturations about the white point is highly anisotropic. Typically the rms saturation along the I_C axis was three times that of the rms saturation along the Q_C axis. Second, we note from the table that individual scenes have an average saturation comparable to their chrominance rms modulation depths, but that for a large number of scenes the average chromaticity tends more toward white (zero saturation). This fact represents a major difference between the properties of luminance and chrominance information. For an ensemble of images the average luminance must be a large positive value, but for chrominance information the average saturation can be zero. From the entries in Table 8 we find the average saturation along the Q_C axis to be effectively zero, whereas the average saturation along the I_C axis varies between 3.5 and 15% (toward orange). Thus, for this ensemble of images the average hue is slightly orange. This result is a reflection of the fact that in manmade objects red, orange, and yellow are used freely to add "warmth" and "excitement." Third, we see from the results presented in Table 8 that, except for game shows and similarly garish settings, the average rms saturation for chrominance variations is quite small. We suspect that for the majority of non-manmade objects the average saturation would be even less.

TABLE 8. SUMMARY OF RESULTS ON THE rms MODULATION DEPTH OF CHROMINANCE INFORMATION

I. Average of 36 Slides

I_C Axis	Q_C Axis
$< \sqrt{\overline{c_m^2}} > = 9.5\%$	$< \sqrt{\overline{c_m^2}} > = 2.4\%$
$< \bar{c} > = 3.5\%$	$< \bar{c} > = 0.7\%$
$< \lvert\bar{c}\rvert > = 7.7\%$	$< \lvert\bar{c}\rvert > = 2.2\%$

II. Off-the-Air Video (All Entries Represent Approximately 15-min Averages)

DESCRIPTION	I_C Axis		Q_C Axis	
	$\sqrt{\overline{c_m^2}}$	$\bar{c}$	$\sqrt{\overline{c_m^2}}$	$\bar{c}$
Game Show (Video Tape)	23.0%	15.0% (orange)	7.5%	2.3% (magenta)
Game Show	29.0%	13.0%	16.6%	0.8% (green)
Talk Show	21.0%	12.0%	6.7%	2.5%
Talk Show	8.0%	11.0%	5.0%	≈ 0
Soap Opera	9.5%	3.4%	2.9%	0.5%
Soap Opera	8.3%	6.8%	3.6%	≈ 0
Archie Bunker (Film)	7.0%	9.2%	2.6%	1.1%

Note: All values are given as percent of saturation.

TABLE 9. LIST OF SCENES USED TO DETERMINE THE CHROMINANCE POWER SPECTRA

Types of Scene	No.	Description
Single-Object Scenes	1	Red zinnia
	2	Indian girl
	3	Soap box
	4	Girl in tall grass
	5	Head of blond lady
	6	Manikin
	7	Lady
	8	Girl and scarf
Two-Object Scenes	1	Santa doll with toy
	2	Aeroplane and two people
	3	Girl on blanket
	4	Girl watching TV
	5	Girl and sunflower
	6	Child in sandbox
	7	Flowers in vase
	8	Girl with RCA sign
	9	Man and woman standing
	10	Lady and soap box
	11	Girl in country
Three-to-Ten-Objects Scenes	1	Four people
	2	Lady in kitchen
	3	Man and woman in room
	4	Lady in room
	5	Two people in room
	6	Fruit and table setting
	7	Boy with kite
	8	Fruit basket
	9	Beach scene with four people
Crowd Scenes	1	Building
	2	Ten people
	3	Crowd of Indians
	4	Stadium crowd
	5	Tulip field
	6	Country scene in fall

4. Chrominance Power Spectral Density Measurements

The individual power spectral densities for the 34 slides listed in Table 9 are presented in Figs. 20, 21, 22, and 23 for chrominance signals along the I_C axis, and in Figs. 24, 25, 26, and 27 for chrominance signals along the Q_C axis. It should be emphasized that the classification of these images into four groups was performed by inspection of the images. Careful examination of the individual power spectral densities indicates considerable variation between slides in each group.

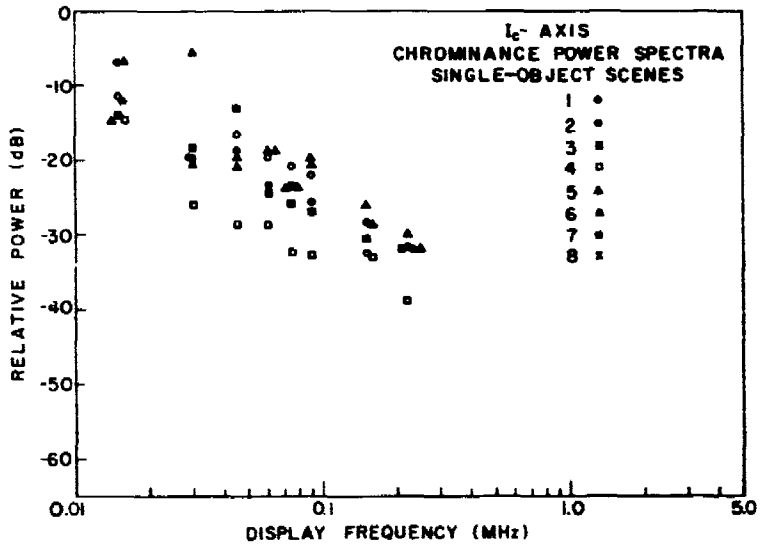


Figure 20. Individual chrominance power spectra for the $C_{m,I}$ terms of scenes that contain one prominent object. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

In Fig. 28 the average power spectral densities for all four classifications of images, and for both the I_C and Q_C axes, are presented. For both the I_C and Q_C axes, these results reflect the amount of structure in the images by the degree to which their spectra are flattened at the lower display frequencies. Also, as predicted from the results presented in Table 8, the power along the I_C axis is considerably greater than that along the Q_C axis. Finally, we note that the power spectra for frequencies above a lower cutoff frequency, as in the case of the luminance power spectra, roll off as $1/\omega^2$. This result is expected due to the distribution of colors in natural scenes. Transitions

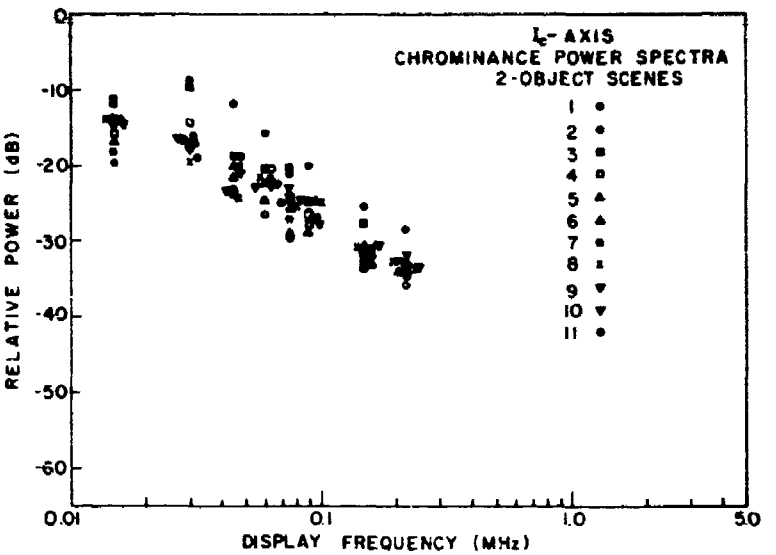


Figure 21. Individual chrominance power spectra for the $C_{m,I}$ terms of scenes that contain two prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

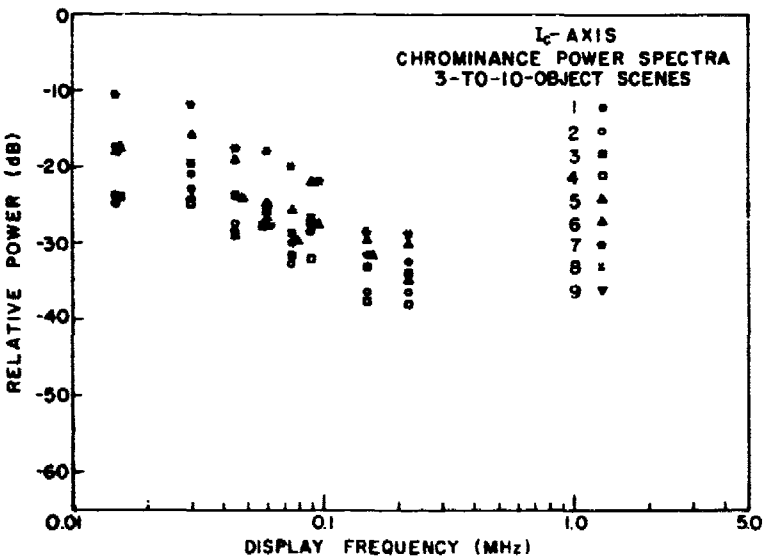


Figure 22. Individual chrominance power spectra for the $C_{m,I}$ terms of scenes that contain three-to-ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

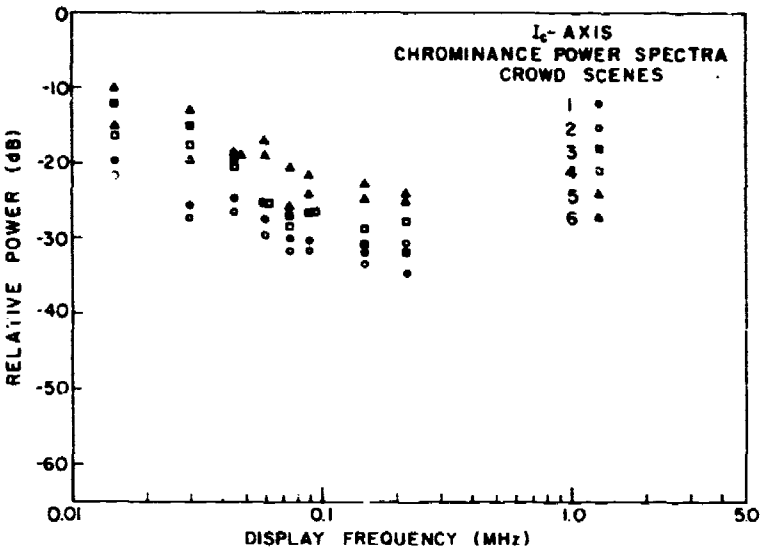


Figure 23. Individual chrominance power spectra for the $C_{m,I}$ terms of scenes that contain more than ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

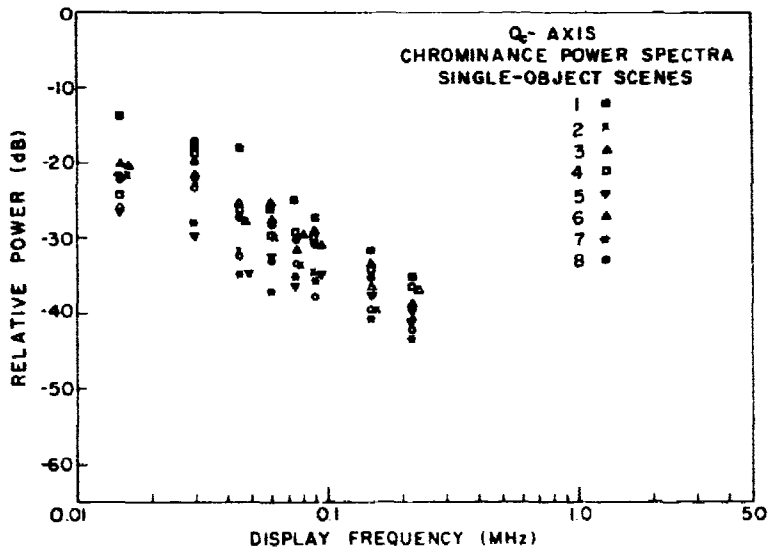


Figure 24. Individual chrominance power spectra for the $C_{m,Q}$ terms of scenes that contain one prominent object. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

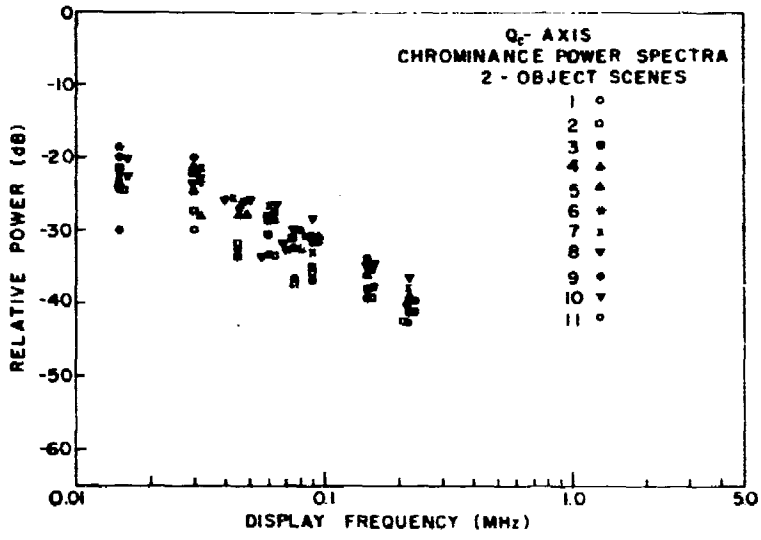


Figure 25. Individual chrominance power spectra for the $C_{m,Q}$ terms of scenes that contain two prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

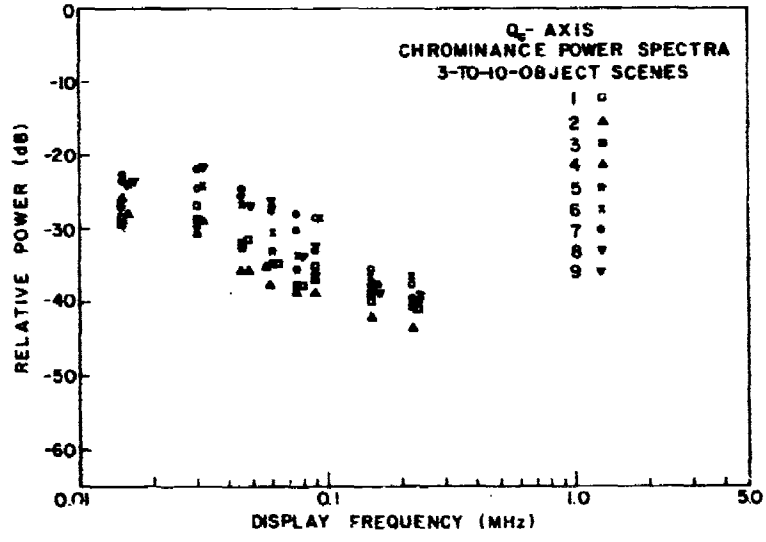


Figure 26. Individual chrominance power spectra for the $C_{m,Q}$ terms of scenes that contain three-to-ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

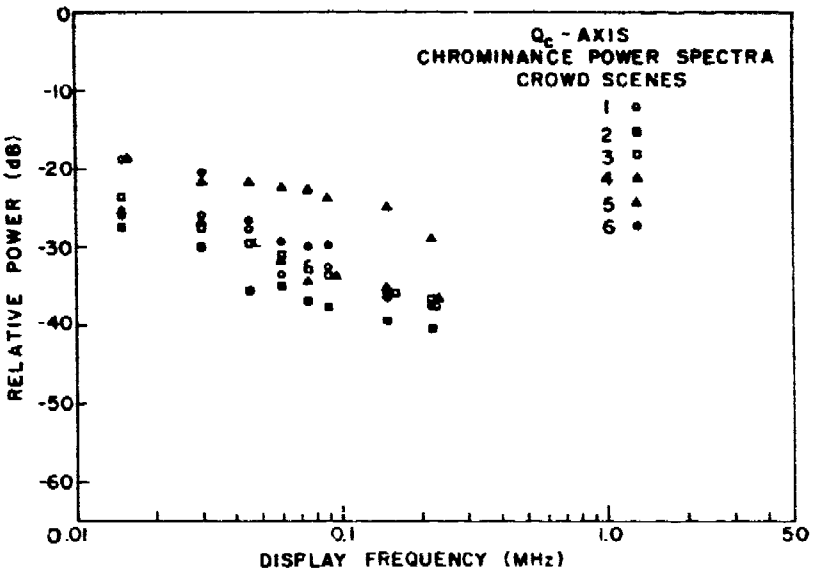


Figure 27. Individual chrominance power spectra for the $C_{m,Q}$ terms of scenes that contain more than ten prominent objects. The numbers next to the symbols in the legend correspond to the scene numbers of Table 9.

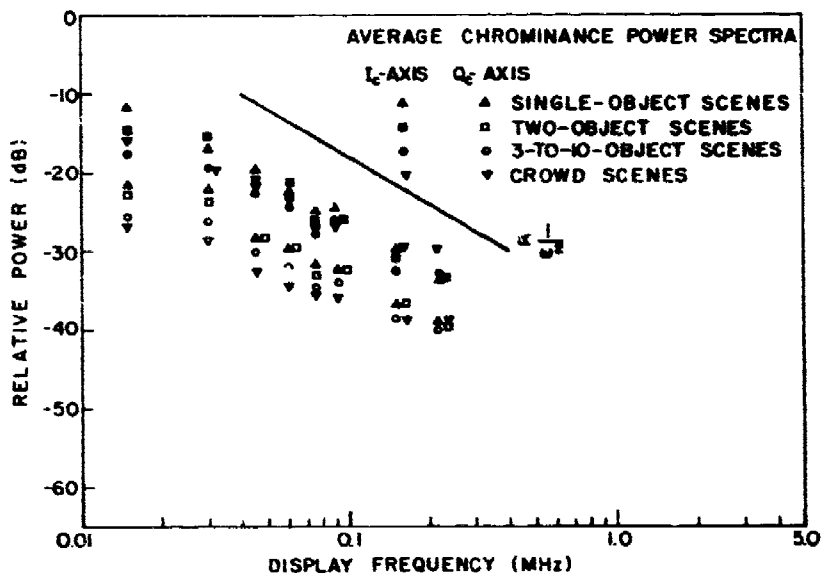


Figure 28. Chrominance power spectra obtained from the averaged results of Figs. 20 through 27. Note that at the higher display frequencies the spectra roll off as $1/\omega^2$. These results clearly show the large difference in power between the I_C and Q_C chrominance axes.

from one chrominance value to another almost always occur suddenly. Because the power spectral densities for each chrominance transition are of the form $1/\omega^2$, it follows that an ensemble average of scenes composed of randomly placed edges will also have a power spectral density of the form $1/\omega^2$. Thus, we find that the power spectra for luminance and chrominance information are similar in form.

D. SUMMARY OF RESULTS

The major conclusions of this study are:

(1) $0.3 \lesssim \langle \sqrt{I_m^2} \rangle / \bar{I} \lesssim 1.0$

Average of 34 slides = 0.56.

(2) The power spectral density for luminance information rolls off as $1/\omega^2$. For individual scenes with a finite sized aperture, there also exists a lower cutoff frequency that is determined by the picture width and scene content. The largest term in the power spectral density is always at dc.

(3) $3\% \lesssim \langle \bar{C}_I \rangle \lesssim 15\%$ (toward orange)

Average of 36 slides = 3.5%.

(4) $0 \lesssim \langle \bar{C}_Q \rangle \lesssim 2\%$ (toward green)

Average of 36 slides = 0.7%.

(5) $8\% \lesssim \langle \sqrt{C_{m,I}^2} \rangle \lesssim 30\%$

Average of 36 slides = 9.5%.

(6) $2\% \lesssim \langle \sqrt{C_{m,Q}^2} \rangle \lesssim 10\%$

Average of 36 slides = 2.4%.

(7) $2.0 \lesssim \langle \sqrt{C_{m,I}^2} \rangle \Big/ \langle \sqrt{C_{m,Q}^2} \rangle \lesssim 5.0$

The measured values along the I_C and Q_C axes represented, approximately, the directions of maximum and minimum rms saturations.

(8) The ensemble-averaged power spectral density for chrominance information rolls off as $1/\omega^2$ for high spatial frequencies. It tends toward zero saturation at dc.

(9) The $1/\omega^2$ roll-off in the power spectra for both luminance and chrominance information is in agreement with the assumption that natural scenes are primarily composed of randomly located edge transitions.

IV. SUBJECTIVE SHARPNESS OF DISPLAYED IMAGES

A. INTRODUCTION

For the proper design and evaluation of displays it is necessary to know the relationship between the objective physical variables of the display and the subjective perceptual variables of the observer. One of the most important subjective variables is image sharpness. In this section we will present the results of a series of experiments that relate the subjective impression of image sharpness to the modulation transfer function (MTF) of the display. Specifically, we have determined the relationship between the change in display MTF necessary to produce a just-noticeable difference (jnd) in image sharpness. The measurements were performed on a large, high-brightness display with still images, both monochrome and colored.

The only previous experiments to determine a scale of subjective image sharpness were performed by Baldwin [30], who used defocused black-and-white 35-mm motion picture film. For small-screen (18 cm x 19 cm), low-brightness ($\approx$ 3 mL) images he measured the jnd in image sharpness as a function of the display resolution. Baldwin found that the jnd in image sharpness was determined by a constant (0.005 degree) change in the linear size of the figure of confusion* of the projected images. That is, he found that as the images became sharper, a larger percentage change in resolution was required to produce a constant change in image sharpness. Because the bandwidths for each of the images used by Baldwin were greater than 6 cycles/degree-of-vision (well past the peak of the visual system's MTF), the general nature of his results is consistent with the current understanding of the suprathreshold spatial frequency response of the eye. For still lower spatial frequencies, however, his findings are not expected to apply. This is because, at low spatial frequencies, the response of the eye decreases. Thus, the change in resolution necessary to produce a jnd in image sharpness is expected to increase with decreasing resolution.

Since the early work of Baldwin (in 1940), it has become widely recognized that a more appropriate physical variable for quantifying image sharpness

*The figure of confusion is a rough measure of the width of the image point-spread function.

is the display MTF, and not simply the display resolution. For example, Schade [4] has shown that subjective sharpness is described more accurately by the noise equivalent bandwidth (N_e), which weights a given spatial frequency according to the square of the system's MTF at that frequency. Recently we have shown, in TR1 and TR2, that the noise equivalent bandwidth of Schade can be improved by generalizing its properties to the perceptual level. We call this descriptor the visual capacity, C_v, because it is the information theory equivalent of the channel capacity for a bilevel transmission system.

Unfortunately the measurements of Baldwin cannot be recast in terms of our current understanding of image sharpness because the display point-spread functions reported by him are ill defined. Furthermore, Baldwin did not record the MTFs of the film he used and, although he indicated that film jitter reduced the vertical resolution of his display, he did not quantify the resolution loss resulting from this cause. All of these reasons make it impossible to reconstruct the overall MTFs for his system.

In this section we report on a series of experiments which relate the subjective impression of image sharpness to the MTF of the display. We will show that our measured results are in good agreement with the assumption that a jnd in image sharpness is determined by a constant change in the rms gradient content of the images – a quantity proportional to the square root of our display descriptor, the visual capacity. We will show that the jnd for conventional images and for an image consisting only of a single luminance transition are practically identical. This measurement supports our contention that edges are a significant feature of natural scenes (see also Section III). Finally, we will show that our measurements calibrate the visual capacity. By this we mean that it is now possible, with the visual capacity as a normalizing stimulus, to specify in perceptual terms the subjective sharpness of displays with different MTFs.

B. EXPERIMENTAL APPARATUS

1. Introduction

The rudiments of the system used to obtain the results presented in this section are shown in Fig. 29. A variable-bandwidth, low-pass, spatial-frequency filter composed of two parallel diffusive plates was used as the display. The

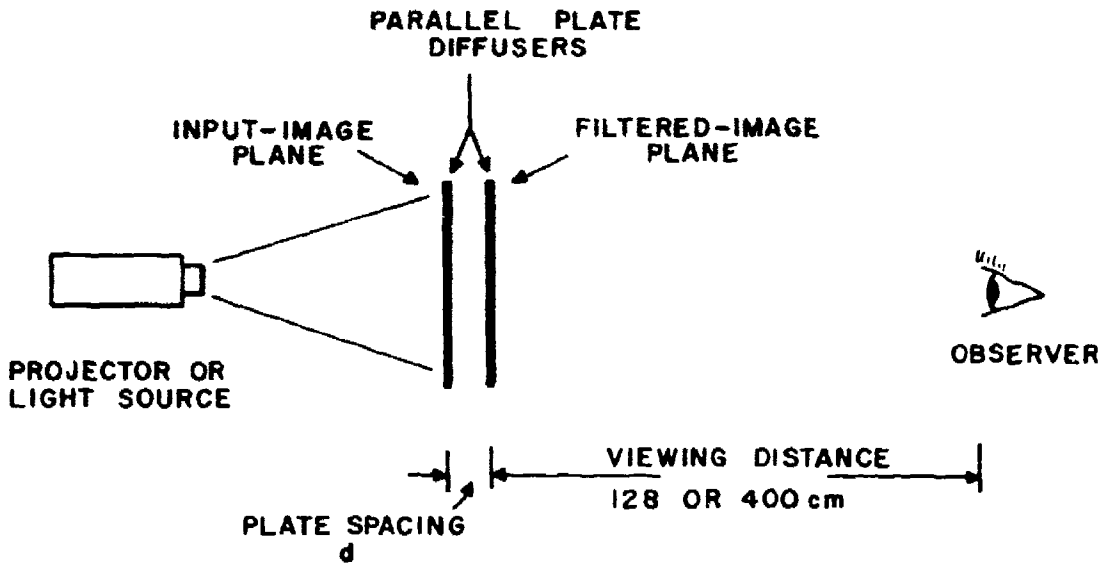


Figure 29. This figure shows the basic elements used in the performance of the experiments. Observers, sitting at viewing distances of either 128 or 400 cm, viewed images that were low-pass filtered by a display composed of two parallel diffusive plates. Images were produced either by the direct projection of 35-mm slides or by the back-illumination of images placed over the input plane of the display. The display was 59-cm wide by 43-cm high; the average screen brightness was always 35 mL.

bandwidths of output images, produced by either direct projection of 35-mm slides or by back-illumination of images placed over the input plane of the display, were controlled by the separation, d, between the two plates. Observers viewing the screen were asked to select the sharper of two images, presented successively, that differed in bandwidth by a small amount. At discrete display bandwidths, distributed throughout the bandpass of the human visual system, the change in bandwidth required to produce a jnd in image sharpness was measured. Each of these elements will be discussed in detail in the succeeding paragraphs.

2. Diffuser Display

In Fig. 30 photographs of the display constructed for these experiments are shown. The display w essentially free of spatial-frequency noise, and

Figure 30. These two photographs show the diffuser display constructed for these experiments. The top photograph shows the input side of the display with a superimposed luminance edge. The bottom photograph shows, from the observer's position, the same edge after filtering by the display.

introduced no spurious coloration into the displayed images. The average display brightness was always 35 mL.

The bandwidth of the display was varied by moving the output image diffuser plate with a motor-controller whose accuracy was better than $\pm$ 0.003 cm. During experiments a masking frame was placed in front of the display (the screen is not shown in Fig. 30) to prevent extraneous visual clues caused by the motion of the output image plate from being detected by the observers. The free aperture of the display was 59 cm x 43 cm.

The electronic circuitry that fed the motor-controller allowed access to two separate plate spacings by the flipping of a switch on a small hand-held box. Thus, an observer could sequence back and forth between two images that differed in bandwidth by a preset amount. Also on the hand-held box were two button switches that corresponded to the two bandwidth settings. During an experiment the observer would decide which of the two images he believed to be the sharper, and press the button corresponding to his choice. If his selection was correct, a green light flashed; if his selection was wrong, a red light flashed. After his selection the results were automatically recorded, and the order of presentation for the two images with different bandwidths was randomized. Therefore, an observer who pushed only one of the buttons on the hand-held box would guess, on the average, the sharper image 50% of the time.

The measured MTF for the display is shown in Fig. 31 as a function of the display frequency times the plate separation, fd. The solid line on the figure represents the analytical approximation for R, also given on the figure. It may be seen that the correspondence between the measured data points and the approximated curve is excellent. Figure 31 also shows that at all plate separations the measured MTFs are of similar form. The large changes in MTF shape that characterize lenses at different defocus positions [25] are not experienced with this system. This represents an important advantage in these experiments because it means that the MTFs can be specified accurately at each plate separation value. An example of the input-output properties for this display with a luminance edge input is shown in Fig. 30.

The limiting spatial frequency response for the diffuser display, referred to as the screen MTF, is established by the granularity of the diffusive material on the two parallel plates. We measured the screen MTF with the plate separation, d, set to zero and found it to be equal to the MTF of Fig. 31 with a plate

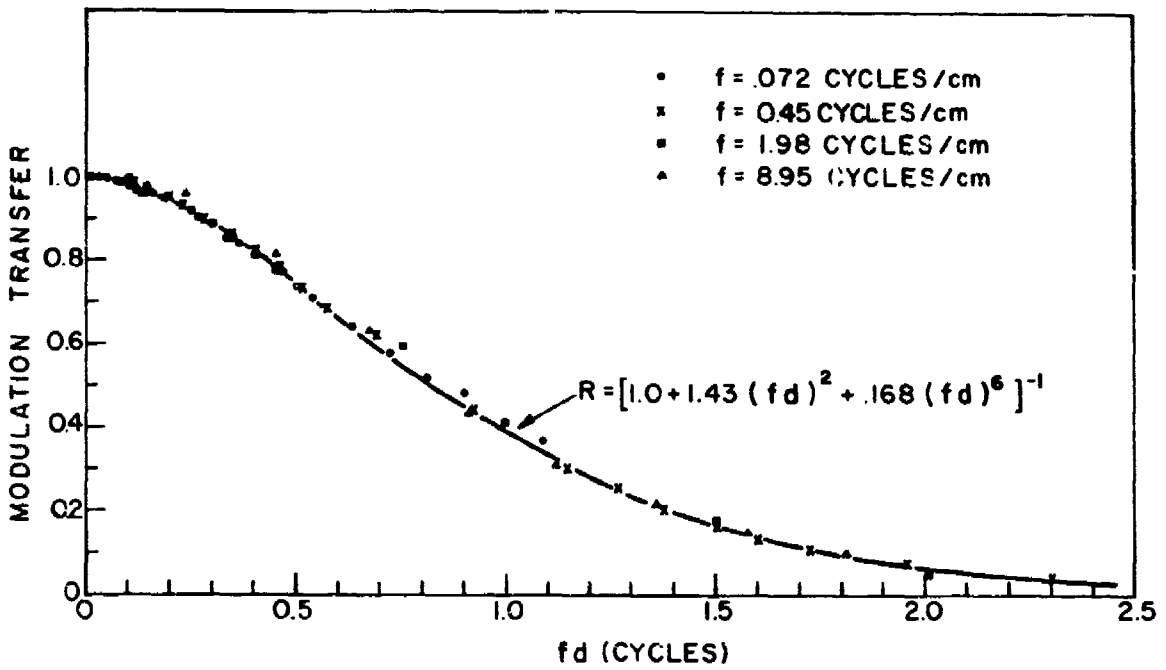


Figure 31. The modulation transfer function for the diffuser display (shown in Fig. 30) as a function of the spatial frequency on the display, f, times the plate spacing, d. These results show that for all plate spacings the form of the modulation transfer functions is identical.

spacing of 0.013 cm. Therefore, the screen MTF can be considered as a small off-set in the plate spacing, d. Where required, this off-set has been included in our results.

3. Scene Characteristics

Two types of still images were investigated in this study: representational scenes in black-and-white and color, and single-transition luminance edges. The luminance edges were studied because of the unique role they play in the derivation of our descriptors, and because of their importance in the composition of natural scenes (see Section III).

The representational scenes we used are shown in Fig. 32. The top picture, of a manikin, was studied in both black-and-white and color. The significant features of this image are its large characteristic feature size, the texture in the straw hat, and the large-area colors of extreme saturation. By accentuating

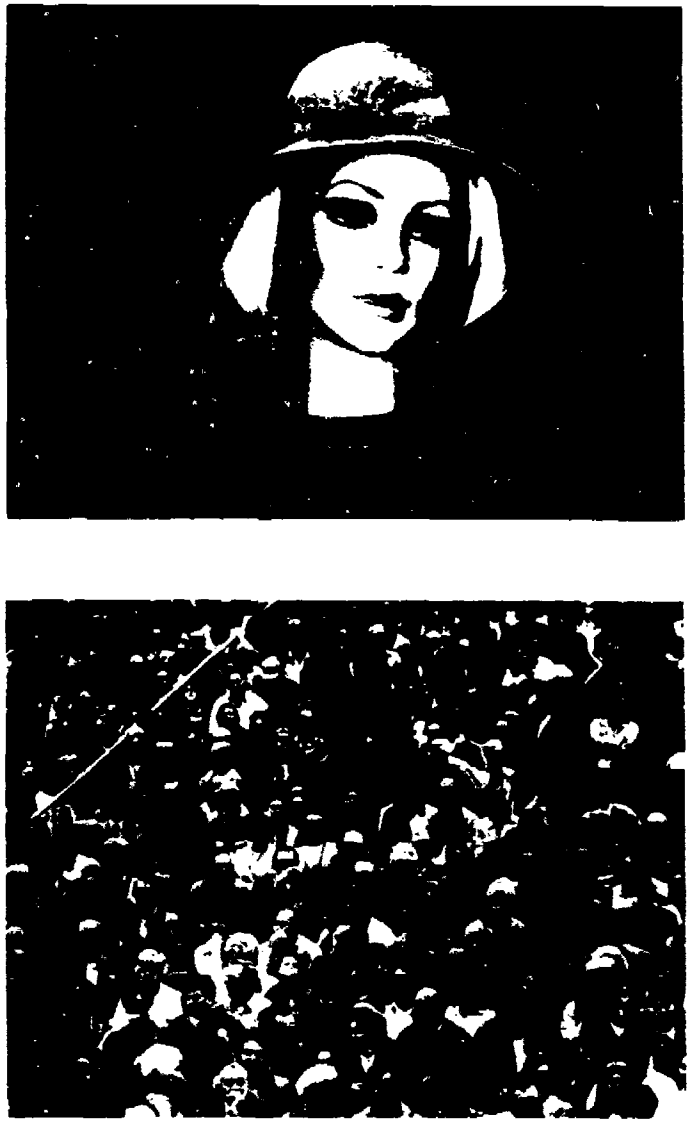


Figure 32. Photographs of the two representational scenes used in this study. The top photograph, of a manikin, was studied both in black-and-white and color.

the colors in this image we attempted to magnify possible differences in subjective sharpness between images in black-and-white and in color. The stadium crowd scene shown in Fig. 32 was represented by a characteristic feature size approximately one-tenth that of the manikin scene. This scene was studied only in color.

Two luminance edges were investigated, as shown in Fig. 33, with contrasts of 82 and 12%, respectively. The definition for edge contrast is given in Fig. 34 as $(I_{max} - I_{min})/(I_{max} + I_{min})$. These two edges were selected as representative of the range of contrasts that may reasonably be expected in actual images.

Figure 35 shows the average of the luminance power spectral densities in the horizontal direction for both the manikin and stadium crowd scenes as a function of the diffuser display frequency. These spectra were obtained by the use of the techniques outlined in Section III with a TV camera. That is, as shown schematically in Fig. 14, the power spectra were obtained from the modulation terms I_m after the average scene luminances $\bar{I}$ were subtracted off. As explained in Section III, this technique has the advantage of minimizing the influence of the finite picture size on the measurements of the power spectra.

The power spectra for the manikin and stadium crowd scenes clearly reflect the important features of each image. The spectrum for the crowd scene is relatively flat up to a frequency that is roughly equal to the inverse width of one spectator. Above this frequency the spectrum rolls off approximately as $1/\omega^2$. The power spectrum for the manikin scene rolls off approximately as $1/\omega^2$ at all but the lowest spatial frequencies, manifesting the larger scale of this scene. The average normalized modulation depth, $\sqrt{\overline{I_m^2}}/\bar{I}$, for the manikin scene was 0.43; for the stadium crowd scene it was 0.47.

At high spatial frequencies the power spectra for the two luminance edges shown in Fig. 33 also roll off as $1/\omega^2$. For spatial frequencies below roughly the inverse of one picture-width the spectra progressively flatten as the frequency is decreased. It is interesting to note that although the stadium crowd scene, the manikin scene, and the luminance edges appear very dissimilar, over a very wide range of spatial frequencies their average power spectra are of almost identical form.

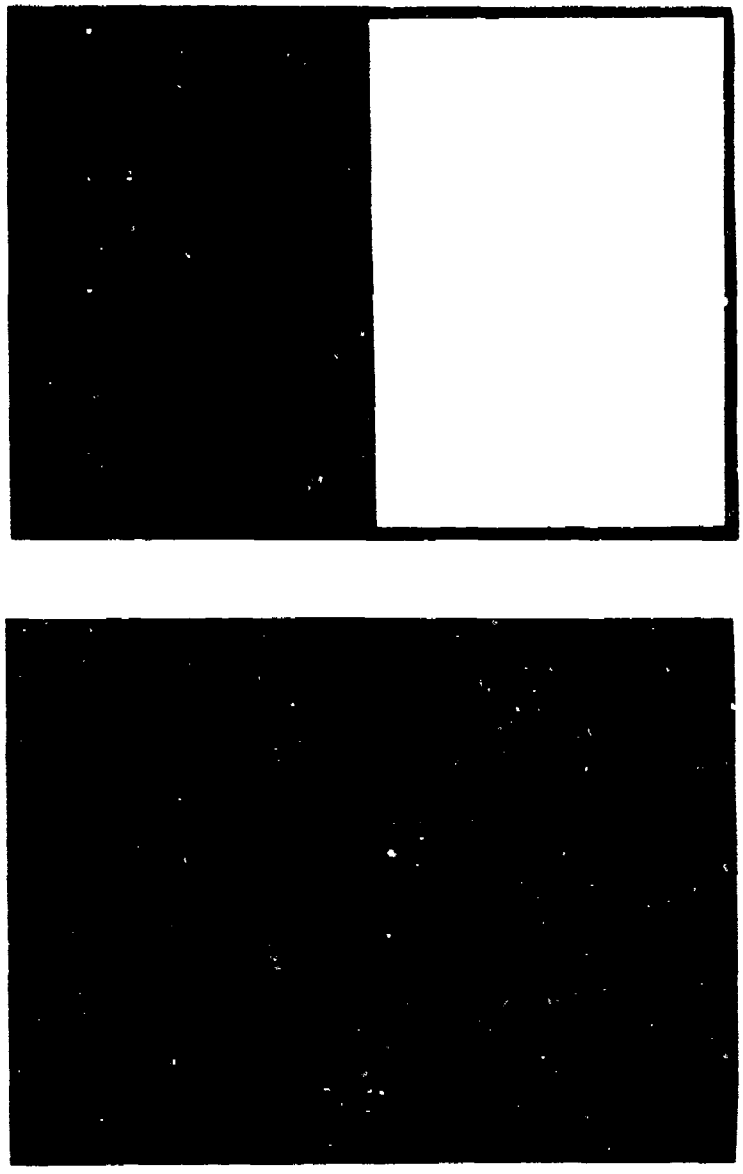
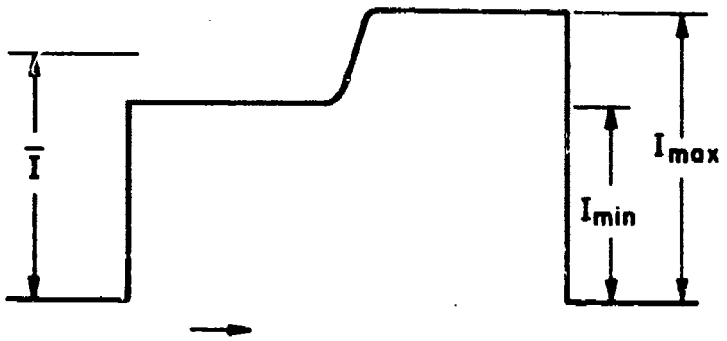


Figure 33. Representations of the two luminance edges used in this study.



$$\text{CONTRAST} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Figure 34. Definition of edge contrast.

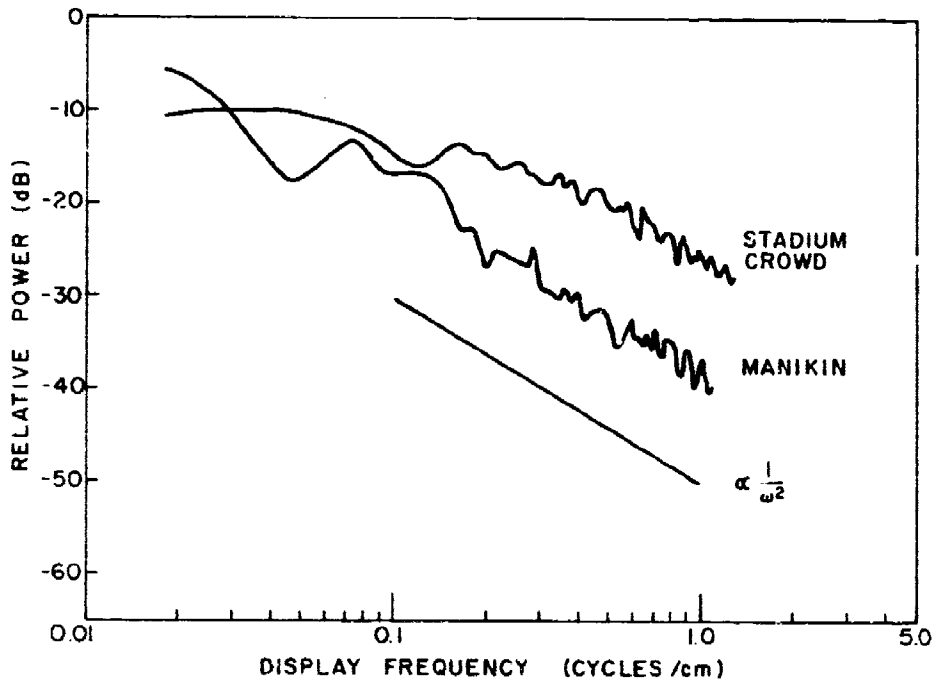


Figure 35. Average horizontal luminance power spectra for the crowd scene and the manikin scene as a function of display frequency. At high spatial frequencies both spectra roll off approximately as $1/\omega^2$. These spectra were obtained from the I_m luminance terms, as shown in Fig. 14.

4. Scene Modulation Transfer Functions

In Figs. 36 and 37 the various processes required to produce the images, described in the preceding section, are summarized. The luminance edges were formed by back-illuminating Wratten neutral-density filters that were placed over the input plane of the diffuser display (see Fig. 29). The overall MTF

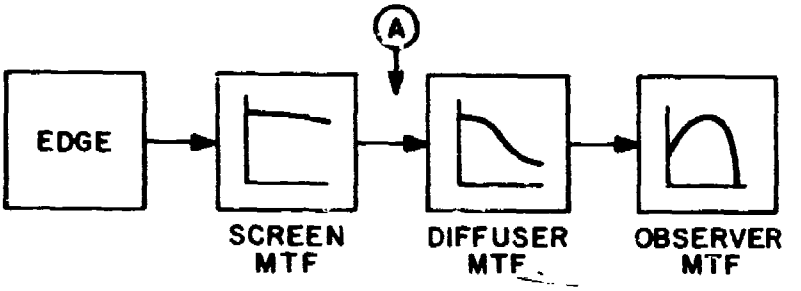


Figure 36. System elements necessary for the generation of luminance edges. Because the luminance edges were produced by the back-illumination of neutral-density filters placed directly over the input image plane, they were degraded only by the screen MTF before filtering by the diffuser MTF. Thus, the limiting quality of the edges at point A was excellent.

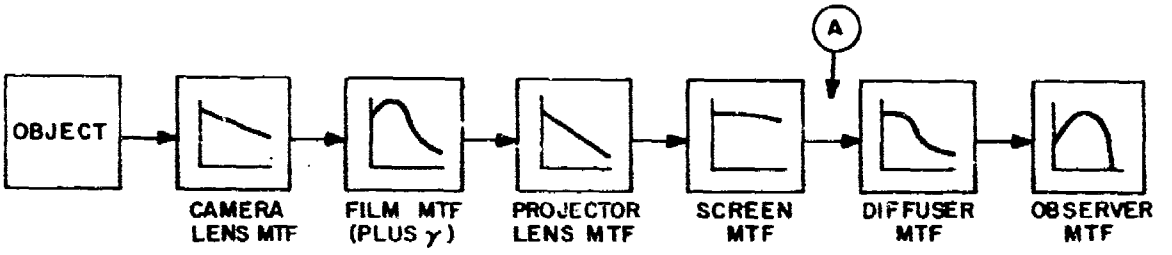


Figure 37. System elements necessary to produce the representational scenes of Fig. 32. The overall system MTFs were obtained for each image by measuring the MTFs at point A. Therefore, all degrading influences up to the diffuser display were accounted for. The measured MTFs are shown in Fig. 38.

for this arrangement is given simply by the product of the screen MTF and the diffuser display MTF, as shown in Fig. 36. Since the luminance edges were degraded only by the screen MTF, whose bandwidth was in excess of 50 cycles/cm (-3 dB), the limiting sharpness for these edges was excellent.

In Fig. 37 the basic steps necessary to produce the representational scenes of Fig. 32 are summarized. For the color and black-and-white images of the manikin, the MTFs were obtained by photographing, under identical conditions, a series of sine-wave gratings of different spatial frequency. These gratings were then used to measure the system MTF on the diffuser display with a plate separation of zero. Thus the resulting MTFs, shown in Fig. 38, include all of the degrading influences up to the diffuser display.

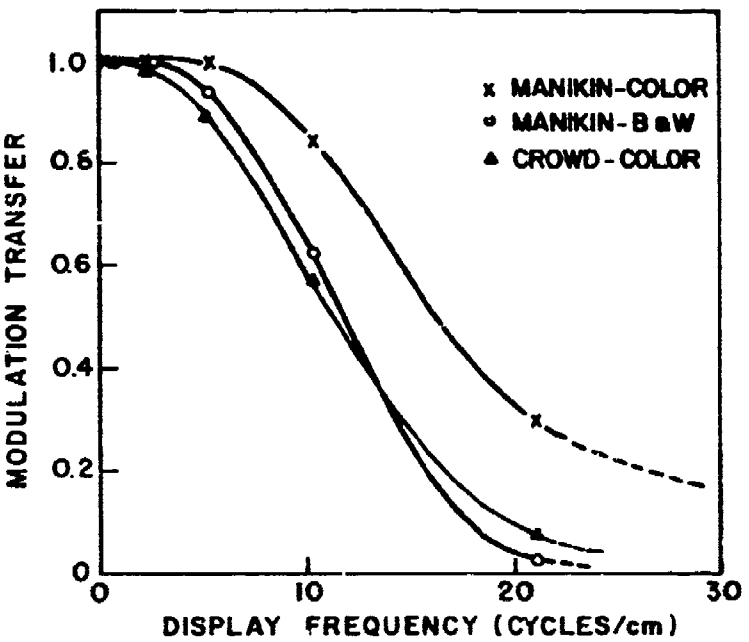


Figure 38. The measured MTFs for the two representational scenes used in these experiments (Figs. 32-37).

The 35-mm color slide of the manikin was produced with Kodachrome* 25, a direct-reversal film of excellent quality. The black-and-white 35-mm slide of the manikin was obtained with Panatomic*-X negative film, which was then printed on Kodak 5302 positive stock. The extra processing step explains why the MTF

*Registered trademark of Eastman Kodak Co., Rochester, N.Y.

for the black-and-white slide is less than the MTF for the color slide of the manikin. The crowd scene was obtained from a 3-in. x 5-in. master color print of excellent quality. The MTF reported in Fig. 38 for this scene represents the processing steps necessary to reduce it to a 35-mm color slide. It is assumed that the additional MTF loss inherent in the original was small compared to the MTF loss in subsequent processing steps. Most significantly, however, this slide was studied primarily at very low diffuser bandwidths where the effects of the processing MTFs are negligible.

It should be noted that, as discussed in Section III (see also Fig. 35), the expected form of the power spectra for these scenes, at high spatial frequencies, is $1/\omega^2$. This assumption makes it possible to reconstruct the spectra for these scenes on the diffuser display with the aid of Fig. 38.

5. Observer Modulation Transfer Functions

The visual MTF used in the computations of Section IV.D. was derived from the visual-contrast-sensitivity function shown as a solid line in Fig. 39. This curve was obtained in a previous study (TR2) with an average screen brightness of 35 mL and a display that subtended approximately 7.0° of visual angle.

The data points shown in Fig. 39 represent the measured contrast-sensitivity values for the two observers who participated in these experiments. These data were obtained by the method of adjustment. That is, at each spatial frequency the contrast of the sine-wave grating was reduced until the observer stated that the grating was just visible. These experiments were performed with the diffuser display at an average screen brightness of 35 mL; each point represents the average of ten experiments.

C. EXPERIMENTAL PROCEDURE

Observers viewing the diffuser display were asked to determine which one of two images, presented sequentially and with different bandwidths, appeared to be the sharper. The observers were instructed to look at the central third of the display and use whatever clues they felt would best enable them to distinguish the sharper image. The images were presented continuously on the screen, as their bandwidths were changed from one value to the other. The pace

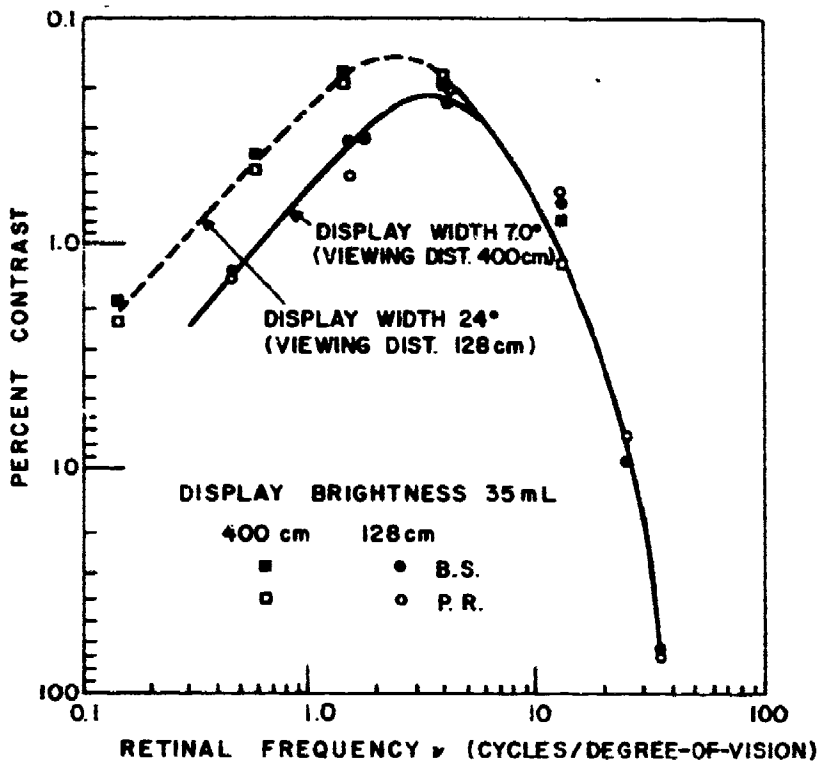


Figure 39. Measured contrast-sensitivity functions for the two observers who participated in this study. The solid line was used to define a visual MTF for use in analytical computations.

of each experiment was controlled by the observers, who flipped a switch on a small hand-held box to alternate between the two images presented. No-choice decisions were not allowed unless the observer was distracted.

Data points were obtained according to the following procedure: The bandwidth of the display was set at two values, ν_1 and ν_2, where $\nu_2 = \nu_1 + \Delta\nu$. The observer, sitting at a fixed viewing distance and viewing the screen normally, alternated between the two images *once*. Then he would decide which of the two images, the first or the second, was the sharper, and record his answer by pressing the appropriate button on the box that he held. After each choice he would be informed as to the correctness of his decision. Subsequently the order of the two images was again randomized. Thus, 50% of the time the first image would be the sharper and 50% of the time, the second image. After 50 transitions at these two display bandwidths the observer's percentage of

correct answers was recorded, and the value of $\delta\nu$ changed. This procedure was repeated for different values of $\delta\nu$ until the psychometric function, $P(\delta\nu)$, at selected values of ν_1 throughout the bandpass of the eye, was established. It was found from preliminary experiments that $P(\delta\nu)$, which is defined as the probability of correctly choosing the sharper of two images at a specific value of ν_1 as a function of $\delta\nu$, is well approximated by a normal error curve (a straight line on arithmetic probability paper). All the results presented in this section are for $P(\delta\nu) = 0.75$, assuming that the data points at each ν_1 are described by a normal distribution.

Two subjects, each with better than 20–20 vision (Snellen rating), were hired to perform these experiments. They were both well motivated and highly experienced in the performance of these tests. They used normal binocular vision to view the display.

The experiments were performed at viewing distances of either 128 or 400 cm. These viewing distances result in display widths of 24 and 7°, respectively. The effect of the different display sizes on the visual-contrast-sensitivity function is shown in Fig. 39. The illumination around the display was approximately one eighth of the average screen luminance. This ratio was chosen to assure both good screen contrast and good visual performance. In a previous study (TR2), intended to determine the effect of surround luminance on the visual-contrast-sensitivity function, we found the contrast-sensitivity function not to be affected severely by surround luminances of less than a factor of ten higher or lower than the mean display luminance.

D. THE PERCEIVED EDGE-GRADIENT CONTENT

In this section we will define a quantity proportional to the square root of the visual capacity; we will use this quantity to interpret the measured subjective sharpness results presented in the next section. In TR1 we showed that, as a direct consequence of the $1/\omega^2$ form of the power spectra of natural scenes, this quantity is proportional to the number of edge transitions that can be perceived across a display. In this context we will show that a constant change in the number of perceived edges can be related to the subjective sharpness of displayed images.

We start from the definition of the reduced perceived brightness (TR2, p. 6) expressed as

$$e = E/[\langle E^2 \rangle]^{1/2} \tag{2}$$

The mean-square perceived-gradient content is given by

$$\begin{aligned} g^2 &= \langle (\partial e/\partial \theta)^2 \rangle \\ &= \langle (\partial E/\partial \theta)^2 \rangle / \langle E^2 \rangle \end{aligned} \tag{3}$$

where θ is the viewing angle. This quantity can be interpreted physically as the average inverse square of the angular distance between perceived transitions.

If we now assume the scenes under consideration to have power spectra of the form $1/\omega^2$, we obtain

$$g^2 = \frac{r^2 \int_0^\infty \frac{d\omega}{\pi} |R(\omega)|^2 O^2(\omega r/2\pi)}{\int_0^\infty \frac{d\omega}{\pi\omega^2} |R(\omega)|^2 O^2(\omega r/2\pi)} \tag{4}$$

where r is the viewing distance. Note that the numerator is the visual capacity, C_v (TR1), while the denominator is the perceived mean-square signal power, S^2, for an input power spectrum of the form $\Phi(\omega) = 1/\omega^2$. Thus we have, for the rms perceived-gradient content,

$$g = rC_v^{1/2}/S \tag{5}$$

We assume that the perception of a change in sharpness of a display is governed by the quantity g in that a change Δg is required before an observer can perceive a sharpness difference. We further assume that the required change Δg is independent of the bandwidth.

The original assumption that the relevant psychophysical quantity involves the square of the input signal, in this case the perceived mean-square gradient, is supported by our quantitative interpretation of the suprathreshold sine-wave results of Nachmias and Sansbury [31], as described in Section V. The particular form of g is reasonable, since it is a measure of the average structure,

or luminance variation, perceived by the viewer in the displayed image. Thus our assumption about Δg is equivalent to postulating that a constant change in the average perceived structure of an image is required to produce equal changes in image sharpness.

In order to compute g the following analytic form for $R(\omega)$ was used:

$$R(\omega) = [1 + 0.951(\omega/\omega_o)^2 + 0.0495(\omega/\omega_o)^6]^{-1} \qquad (6)$$

where ω_o is the -3-dB point of $R(\omega)$, i.e., $R(\omega_o) = 1/2$. In Fig. 39 the closeness of this fit to the measured MTF can be seen. The form for $O(\nu)$ was taken from our threshold-contrast-sensitivity-function measurements (TR2, p. 95) with a display diameter of 6.5°. In Fig. 39 it is shown that this value for $O(\nu)$ is an accurate representation of the visual MTFs of the observers who participated in this study.

The psychophysical quantity Δg was computed by determining, at a specific ω_o, the change in display bandwidth $\Delta\omega_o$ required to produce a perceivable change in display sharpness. Thus, the condition to be satisfied is

$$\Delta g = g(\omega_o + \Delta\omega_o) - g(\omega_o) = \text{constant} \qquad (7)$$

The magnitude of Δg represents the only adjustable parameter. It determines the scale of the required bandwidth changes, but for small $\Delta\omega_o/\omega_o$ it does not alter the form of the variation of $\Delta\omega_o$ with ω_o. The computed values presented in the next section, were obtained by assuming that $\Delta g = 1/7.5$ deg^{-1}. It will be shown that this value represents an accurate approximation to our experimental results.

It is important to realize that the results obtained for large ω_o (i.e., $\omega_o r/2\pi\nu_o \gtrsim 4$, where ν_o = 3 cycles/degree-of-vision) are insensitive to the details of the form of the MTF of visual system $O(\nu)$, but depend crucially on the low-frequency behavior of the display. For example, at $\omega_o r/2\pi\nu_o = 6$ the -3-dB point of the display lies at about 18 cycles/degree-of-vision, where the sensitivity of the human visual system is small. However, the peak of the visual MTF lies at a display frequency at which $R(\omega)$ is greater than 0.97. Thus the *perceived* effect of changes in the display bandwidth parameter ω_o arises primarily from changes in the low-frequency part of the display's spectral range. Stated differently, the perceived sharpness improvement resulting from bandwidth increases is due mainly to the improvement of the low-frequency response.

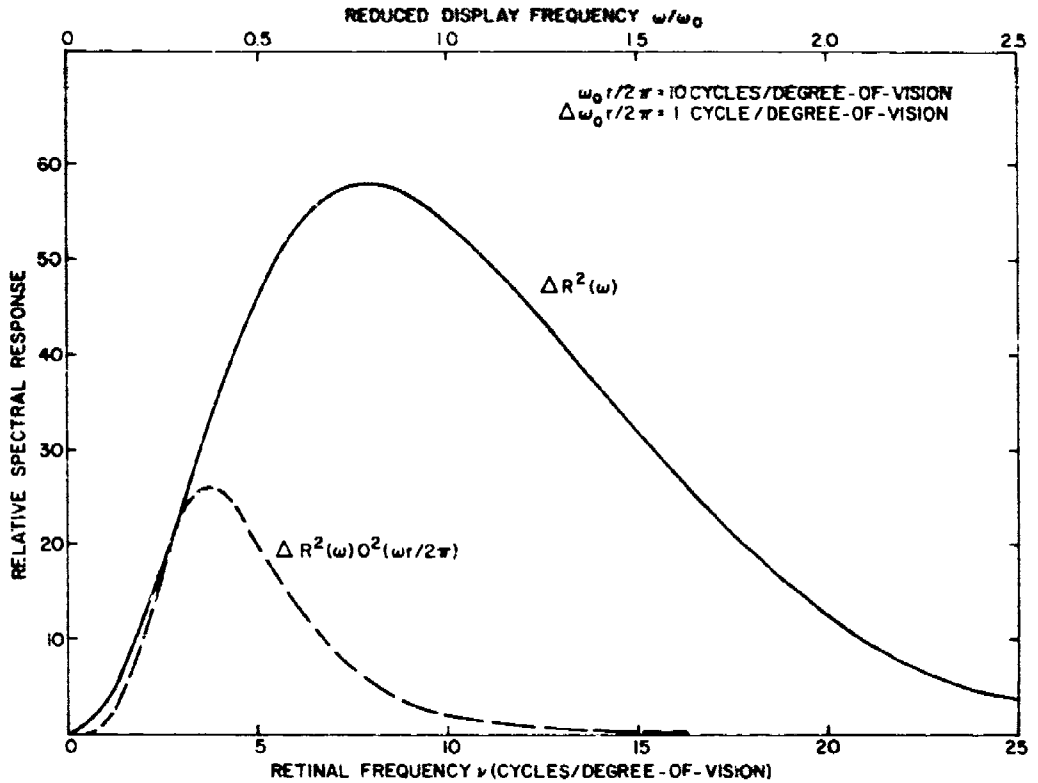


Figure 40. This figure shows the relative spectral increase when a display (for an MTF of the form shown in Fig. 31) with a bandwidth of 10 cycles/degree-of-vision is increased to one of 11 cycles/degree-of-vision. Although the increase in spectral response due to the increase in display bandwidth, $\Delta R^2(\omega)$, occurs throughout the bandpass of the eye, the perceived response, $[\Delta R^2(\omega)O^2(\omega r/2\pi)]$, is significant only at the lower spatial frequencies.

In Fig. 40 this point is illustrated for a display with bandwidth $\omega_o r/2\pi = 10$ cycles/degree-of-vision and a change in display bandwidth of $\Delta\omega_o r/2\pi = 1$ cycle/degree-of-vision. In this figure we have plotted the relative spectral response of the quantities

$$\Delta R^2(\omega) = |R(\omega)|^2 \quad \text{at } \omega_o = \omega_o' + \Delta\omega_o \quad - \quad |R(\omega)|^2 \quad \text{at } \omega_o = \omega_o'$$

and $\Delta R^2(\omega)O^2(\frac{\omega r}{2\pi})$ as a function of retinal frequency. It can be seen from Fig. 40 that although the change in display bandwidth $\Delta\omega_o$ results in an in-

crease in $\Delta R^2(\omega)$ throughout the bandpass of the eye, the perceived response $\Delta R^2(\omega)O^2(\frac{\omega r}{2\pi})$ is significant only near the peak of the visual MTF (which lies at low display frequencies). The contribution in spectral response for retinal frequencies between 10 and 35 cycles/degree-of-vision is negligible! This observation has strong implications for display design. For many displays of practical interest the perceived sharpness can be improved simply by increasing the low-frequency response, rather than by extending the display bandwidth.

On the other hand, for small display bandwidths $[(\omega_o r/2\pi\nu_o) \lesssim 1]$, the predicted subjective sharpness results are sensitive to the precise form of the visual MTF, $O(\nu)$. We will return to this point later.

E. RESULTS

The measured results are shown in Figs. 41-45, plotted with $(\nu_2-\nu_1)/\nu_2$ as a function of ν_1. The frequencies ν_1 and ν_2 are, respectively, the lower and higher bandwidths (defined at $R(\omega_o) = 1/2$) necessary for an observer to discern 1 jnd in image sharpness as described in Section IV.C. Each data point on the figures represents over 200 choices between two images with different bandwidths. The coordinators are chosen so that when $(\nu_2-\nu_1)/\nu_2 = 1$, the bandwidth must be extended from ν_1 to infinity to produce one additional jnd in image sharpness. Also shown in these figures are the calculated results based on the results of Section IV.D.

The quantity $(\nu_2-\nu_1)/\nu_2$ in these figures represents close to the minimum change in bandwidth that can be detected. The values were obtained under optimum conditions by highly trained observers viewing still images on a noiseless, large, bright display. When conventional displays, such as commercial television, are viewed, 1 jnd in image sharpness would represent an almost imperceptible change. However, a change in sharpness corresponding to 3 jnd's would, in most situations, represent a conspicuous improvement.

An obvious feature of the measured results is that, for the two observers studied, the measured values for $(\nu_2-\nu_1)/\nu_2$ at each ν_1 differ by as much as a factor of 2. This is because individual observers (a) will have different criteria for deciding a jnd in image sharpness, (b) will generally look at the display in different locations, and (c) will not be identical in their visual performance. In a preliminary study with ten observers we found that at ν_1 = 4 cycles/degree-of-vision for the manikin scene the measured jnd's varied by

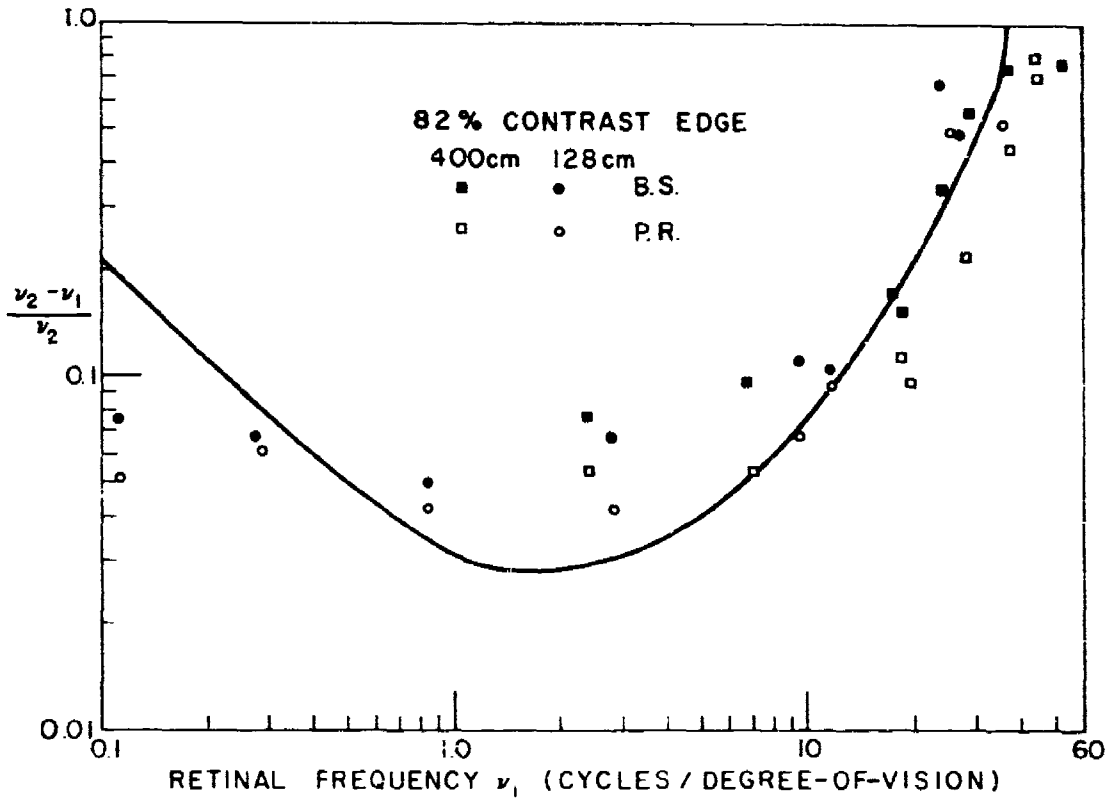


Figure 41. Measured results, $(\nu_2 - \nu_1)/\nu_2$ as a function of ν_1, for an 82%-contrast luminance edge. One jnd in image sharpness was defined as the difference in bandwidth ($\delta\nu = \nu_2 - \nu_1$) from ν_1 necessary for an observer to perceive a change in image sharpness 75% of the time. Each data point is the average of over 200 events, and all retinal frequencies are defined at the point at which $R(\omega) = R(2\pi\nu/r) = 1/2$.

4 to 1 from the most sensitive to the least sensitive observers. The two observers reported here (P.R. and B.S.) were both above average in sensitivity when compared with this group.

Our results differ from Baldwin's in that, independent of the size of the point-spread functions,* the linear change in their widths necessary to produce 1 jnd in image sharpness was not constant. We found that the linear change in the width of the point-spread functions increased at both high and

*For the diffuser display described in Section IV.B, the width of the point-spread function is directly proportional to the plate spacing, d. The relationship between plate spacing and display MTF is shown in Fig. 31.

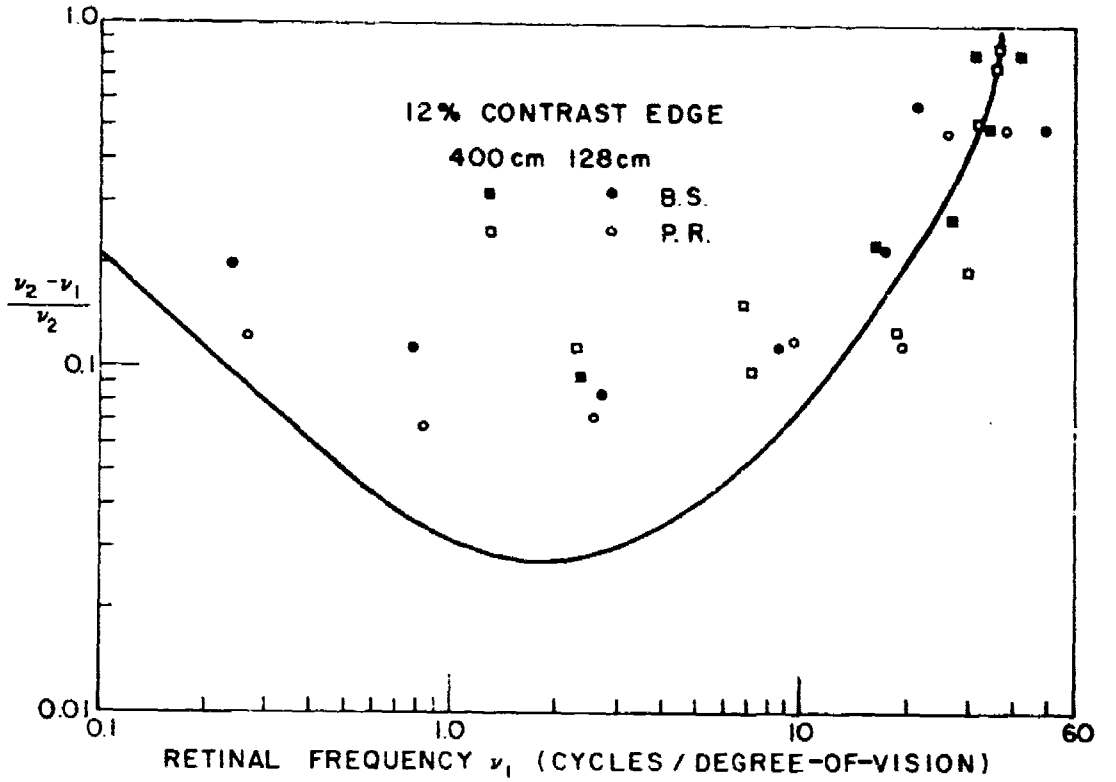


Figure 42. Measured results, $(\nu_2 - \nu_1)/\nu_2$ as a function of ν_1, for the 12%-contrast luminance edge. One jnd in image sharpness was defined as the difference in bandwidth ($\delta\nu = \nu_2 - \nu_1$) from ν_1 necessary for an observer to perceive a change in image sharpness 75% of the time. Each data point is the average of over 200 events, and all retinal frequencies are defined at the point at which $R(\omega) = R(2\pi\nu/r) = 1/2$.

low spatial frequencies with a minimum value in the range of 3-7 cycles/degree-of-vision. This is a result that was successfully predicted by the analysis presented in Section IV.D.

The measured results for each of the images studied are similar in form. Also, it can be seen that our analytical results are in good agreement with the measured results for display frequencies above 10 cycles/degree-of-vision. This is significant since most practical displays are situated in this frequency range. This is also the frequency range where the detailed shape of the visual MTF, $O(\nu)$, is least important, and where our conclusion as to the power spectra of natural scenes having the form $1/\omega^2$ is most valid. At the

lowest frequencies, however, our analytical prediction consistently overestimates the slope of $(\nu_2-\nu_1)/\nu_2$ vs ν_1. This difference may be attributed to several possible causes, including the assumptions of our model. Yet this flattening might in part be a consequence of a suprathreshold visual MTF that does not attenuate as quickly as we have assumed it would at low spatial frequencies.* This possibility is given direct support by the work of Georgeson and Sullivan [32]. They measured a suprathreshold MTF for the visual system, using a contrast matching technique, and found that the measured MTFs were flattened over the middle range of retinal frequencies (0.5-10 cycles/degree-of-vision) for sine-wave gratings with contrasts above threshold. If we were to incorporate into our formalism [Eq. (4)] a visual MTF with these properties, the resulting predictions for the scaling of $(\nu_2-\nu_1)/\nu_2$ vs ν_1 would come very close to the measured results. Or, conversely, if we were to determine the visual MTF that best fits our measured results, we would define the correct *operational* suprathreshold visual MTF to use in the analysis of natural images.

Now consider the results for the two luminance edges studied, shown in Figs. 41 and 42. These results are among our most important because they were obtained with essentially perfect input edges, all of whose unfiltered power spectra are of the form $1/\omega^2$. Thus, one of the assumptions of our model is met exactly, except at the lowest spatial frequencies where the power spectra are determined by the picture width. However, as our previous results (TR2, p. 95) with sine-wave gratings have shown, a display 24° wide is essentially of infinite width perceptually.

At the highest spatial frequencies the results of Figs. 41 and 42 are practially identical. Also, the correspondence between the measured values and the theoretical prediction is excellent. At lower spatial frequencies two observations can be made. First, the jnd for the 82% contrast edge is approximately 80% smaller than the corresponding jnd for the 12% contrast edge. This difference in sensitivity is quite small when compared to the almost 50-to-1 difference in signal power between these two images. Thus, our assumption of normalizing the reduced perceived brightness with the rms perceived brightness from the display is supported. Second, note that the

*An additional complication will be discussed below.

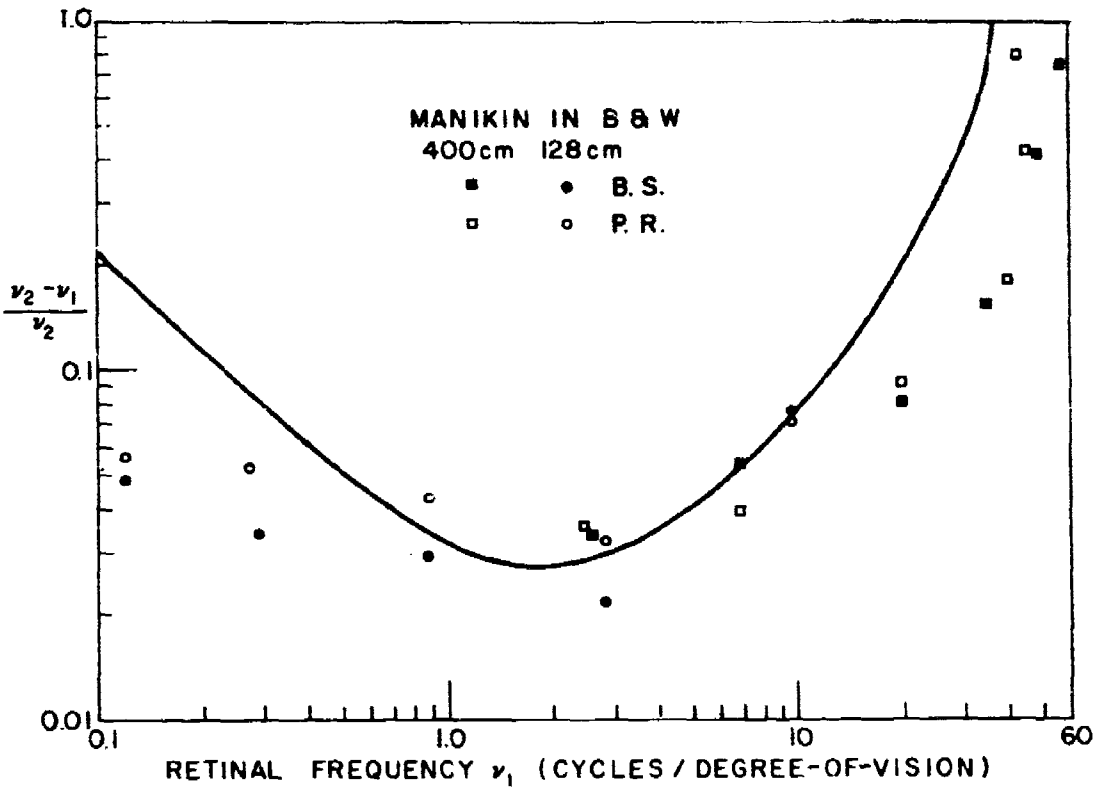


Figure 43. Measured results, $(\nu_2 - \nu_1)/\nu_2$ as a function of ν_1, for the manikin scene in black-and-white. One jnd in image sharpness was defined as the difference in bandwidth ($\delta\nu = \nu_2 - \nu_1$) from ν_1 necessary for an observer to perceive a change in image sharpness 75% of the time. Each data point is the average of over 200 events, and all retinal frequencies are defined at the point at which $R(\omega) = R(2\pi\nu/r) = 1/2$.

measured values of $(\nu_2-\nu_1)/\nu_2$ for the 82% contrast edge are practically independent of ν_1 from 0.1 to 2 cycles/degree-of-vision, whereas over this frequency range the measured values of $(\nu_2-\nu_1)/\nu_2$ for the 12% contrast edge decrease slightly with increasing frequency. That is, these results are in general agreement with the proposition that the visual MTF is a function of scene contrast; the shape of the MTF flattens with increasing contrast. Third, the measured results for the 82% contrast edge are almost identical to the measured results for the representational scenes. This fact is additional indirect evidence that edges are the most significant feature of natural scenes.

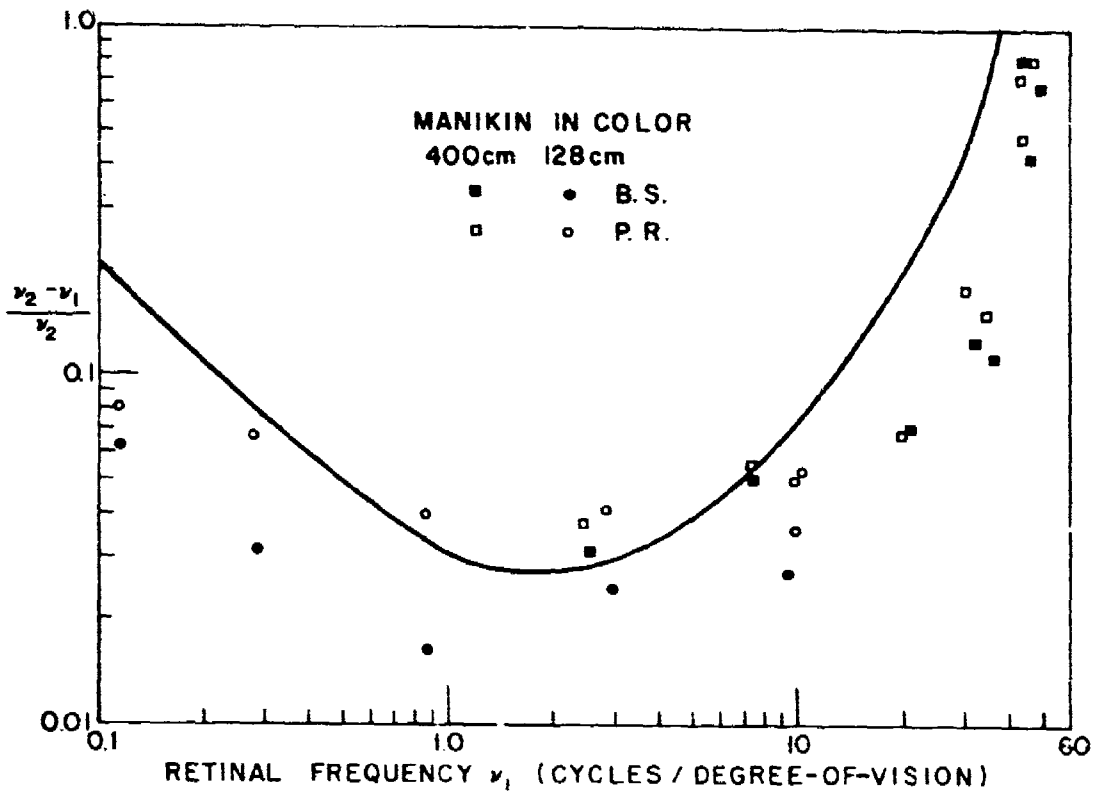


Figure 44. Measured results, $(\nu_2 - \nu_1)/\nu_2$ as a function of ν_1, for the manikin scene in color. One jnd in image sharpness was defined as the difference in bandwidth ($\delta\nu = \nu_2 - \nu_1$) from ν_1 necessary for an observer to perceive a change in image sharpness 75% of the time. Each data point is the average of over 200 events, and all retinal frequencies are defined at the point at which $R(\omega) = R(2\pi\nu/r) = 1/2$.

Next consider the results obtained by the use of the black-and-white and colored manikin scenes (see Figs. 43 and 44). For frequencies above approximately 5 cycles/degree-of-vision the measured results are consistently lower than the predicted results. This is an expected finding due to the limited bandwidth of these images, as shown in Fig. 38. It is possible to interpret this result $1/\omega^2$) as arising from a display whose actual frequency response is the product of the appropriate MTFs of Figs. 31 and 38. If the MTFs for the images are included in the analysis, the result is to produce the desired correspondence between theory and experiment.

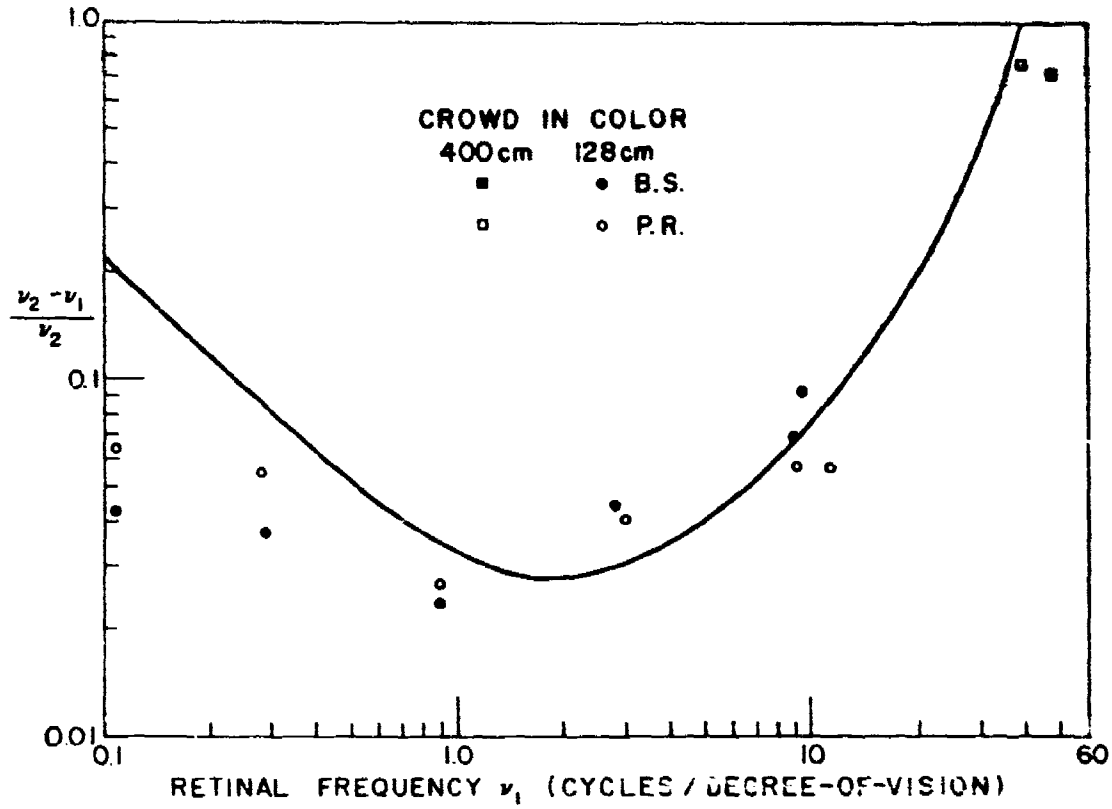


Figure 45. Measured results, $(\nu_2 - \nu_1)/\nu_2$ as a function of ν_1 for the crowd scene in color. One jnd in image sharpness was defined as the difference in bandwidth ($\delta\nu = \nu_2 - \nu_1$) from ν_1 necessary for an observer to perceive a change in image sharpness 75% of the time. Each data point is the average of over 200 events, and all retinal frequencies are defined at the point at which $R(\omega) = R(2\pi\nu/r) = 1/2$.

At the lowest spatial frequency the differences in the results of the colored and black-and-white manikin scenes are not large enough to allow specific conclusions to be drawn. Our general appraisal of these measurements is that the addition of color to an image does not appreciably improve the jnd in image sharpness. This conclusion is consistent with our current understanding of luminance and chrominance information processing in the human visual system (see Section V).

In Fig. 45 the results of our measurements of the stadium crowd scene are given. This slide was originally included because at low spatial frequencies it has a spectrum that is relatively flat compared to the manikin scene. However, the measured results for the crowd scene and the manikin scene are obviously indistinguishable, especially at low spatial frequencies. This result

is encouraging because it implies that the spectral variation among scenes will have to be considerably greater than the one that exists between these two scenes (see Fig. 35) in order to significantly affect the scales of subjective sharpness reported here. However, as we have shown in Section III, the power spectra for natural scenes are all of remarkably similar form. We therefore conclude that our measurements on subjective sharpness will be applicable to a wide variety of natural scenes.

Finally, throughout this section we have tacitly assumed that image sharpness is a meaningful concept, whatever the display bandwidth. In fact, however, the concept applies only as long as the images under study are recognizable as images. For display bandwidths that are extremely low, the images become amorphous shapes for which the commonly understood meaning of sharpness has little significance.* Indeed, at very low bandwidths the distinct impression of the experiment converts from a judgment of a semantic sharpness change to one of a semantic contrast change. This transition occurs at about 1 or 2 cycles/degree-of-vision. If we accept the proposition that the spatial-frequency-specific channels of the visual system are roughly 1 cycle/degree-of-vision wide [33], then this transition occurs when only one or two channels are excited. If the energy within each channel is summed before a detection apparatus is applied, then at the lowest frequency, where only one channel is excited, the measured change in bandwidth necessary to produce 1 jnd should be the same for either a scene with a broad spectrum (such as the crowd scene) or a scene with a narrow spectrum (such as a single sinusoid) at the same average contrast. This is true even though the latter experiment must involve only a change in the *contrast* of the sinusoid. We tested this hypothesis using a sine-wave grating at 0.14 cycles/degree-of-vision whose general impression of contrast (22%) was the same as that of the crowd scene with a display frequency of ν_1 = 0.14 cycles/degree-of-vision. For these conditions the measured value of $(\nu_2-\nu_1)/\nu_2$ necessary to produce 1 jnd in contrast change was 0.6. The comparable value for the crowd scene was 0.7. Thus, at the lowest spatial frequencies we find our measurements for the jnd in sharpness to be indistinguishable from the results of an experiment whose objective had been to measure the jnd in contrast of a single sinusoid.

*This also applies for the edges studied.

F. SUMMARY OF CONCLUSIONS

We have obtained the first results which relate the perception of image sharpness to the modulation transfer function of a display. Our specific findings are as follows:

(1) The jnd in image sharpness is not determined by a constant change in the linear size of the display point-spread function. The change in the point-spread function is a minimum at about 3-5 cycles/degree-of-vision and becomes progressively larger for display frequencies away from this point.

(2) The scales of subjective sharpness are not appreciably different for black-and-white and for colored images. From this we conclude that image sharpness is determined primarily by the luminance portion of images.

(3) The measured jnd's in sharpness for single-transition luminance edges and conventional images are indistinguishable. This finding supports our view that edges are the dominant feature in natural images.

(4) We have shown that at high frequencies our measured results are in good agreement with the assumption that a jnd in image sharpness is determined by a constant change in the perceived rms gradient content of images [Eq. (7)]. At low spatial frequencies our results support the contention that the roll-off in the visual MTF is more gradual for high scene contrasts than it is for scenes with low contrast. We believe that Eq. (7), for the change in the display MTF required for a perceivable change in sharpness, can be applied with confidence to displays with MTF forms different from that employed in these experiments.

(5) We have shown that the judgment of sharpness changes is not a strong function of scene content. This result is partially a consequence of the high degree of self-similarity in the spectral properties of natural scenes.

(6) At very low spatial frequencies ($\lesssim$ 1 cycle/degree-of-vision), we have established that the judgment of a sharpness change is indistinguishable from the judgment of a contrast change for a single sinusoid.

(7) Finally, we have obtained the scale value necessary to calibrate the visual capacity. It is now possible to specify, by means of visual capacity, the absolute perceived sharpness of different displays.

V. STATISTICAL THEORY OF DISPLAY DESCRIPTORS

A. INTRODUCTION

A major goal of our research during the current year was to determine, from an information theory point of view, the effect of color on perceived image quality. In our previous formulation (cf. TR1) of the information capacity of the display-observer system, only the contribution of the perceived luminance signal was considered. Accordingly, we have sought to extend the concept of the information capacity so as to encompass the psychophysical dimensions of luminance and chrominance. Such a total descriptor would not only allow us to assign a quantitative value to the relative performance characteristics of monochrome and color displays but would permit display designers to optimize the allocation of available resources to produce the best balance of luminance and chrominance signal capability. At the outset the treatment of such disparate visual dimensions as luminance and chrominance on an equivalent basis may seem akin to comparing apples and pears. However, it will be seen that the formalism of information theory, when combined with a simple, verifiable model of the human visual system, allows an unambiguous collective treatment of these two dimensions.

In order to proceed with the synthesis of a combined information theory descriptor for luminance and chrominance, it was necessary to develop a far less heuristic measure of the luminance channel capacity than the quantity introduced in TR1. In particular, the assumption that the visual system can be regarded as a linear system, characterized by a single modulation transfer function, has been removed. Instead, a nonlinear signal-detection model has been employed. This model makes use of existing sine-wave contrast-sensitivity data and, as will be seen, agrees well with the results of other psychophysical experiments. The essential aspects of the model are discussed under V.B. below. In V.C. the model is employed to develop the first measure of the information capacity of the display-observer system that takes into account the nonlinear characteristics of the human visual system. In V.D, the information theory approach is applied to the chrominance dimension in a manner consistent with that utilized for the luminance dimension. The result is an expression for the total channel capacity of the display-observer system,

including both luminance and chrominance. This expression is then applied to the solution of a practical design problem involving the optimum trade-off between luminance and chrominance signal capability.

II. THE DISTRIBUTION OF PERCEIVABLE SINE-WAVE LUMINANCE LEVELS

1. Signal-Detection Model

Our approach to the determination of the channel capacity of the display-observer system requires a knowledge of the number of perceivable sine-wave luminance levels as a function of retinal frequency ν. Therefore, before proceeding with the development of the formalism for the channel capacity, we shall describe a simple detection model that appears to account for the observed distribution of these levels.

It is well known that the human visual system can be characterized by a sine-wave threshold sensitivity function $m_T(\nu)$. This function depends somewhat on the mean luminance of displayed sine-wave gratings, background luminance, and display size, but, as was shown in TR2, the function is remarkably constant from person to person if the viewing conditions are held constant. Thus, $m_T(\nu)$ can indeed be regarded as a fundamental property of the visual system.

However, the threshold-sensitivity function is not sufficient to determine the number of perceivable luminance or contrast levels at a given retinal frequency. For example, if the visual system were linear, the contrast or modulation $m_i(\nu)$ required for the perception of i levels would be simply

$$m_i(\nu) = i m_T(\nu) \text{ where } i = 1, 2.... \text{ (linear model)} \tag{8}$$

On the other hand, a nonlinear Weber's law model for the distribution of levels would give an exponential relationship. For a Weber's law model, we have $\Delta m/m_i = (m_{i+1}-m_i)/m_i = k_w$(a constant). This difference equation is easily solved to give

$$m_i(\nu) = (1 + k_w)^{i-1} m_T(\nu) \text{ (Weber's law model)} \tag{9}$$

The actual distribution is more complicated than either of the above models. As may be seen in Fig. 46, the data of Nachmias and Sansbury [31] show that the required contrast change Δm for a jnd in contrast [79.4% correct response in a forced-choice experiment] is not independent of the value of the initial contrast, as predicted by the linear model Eq. (8), nor does it increase linearly with the initial contrast, as required by the Weber's law model. Indeed, the dependence of Δm on starting contrast is not even monotonic. The required Δm first decreases from its initial value m_T, goes through a minimum, and finally does approach a Weber's law characteristic at high values of the initial contrast.

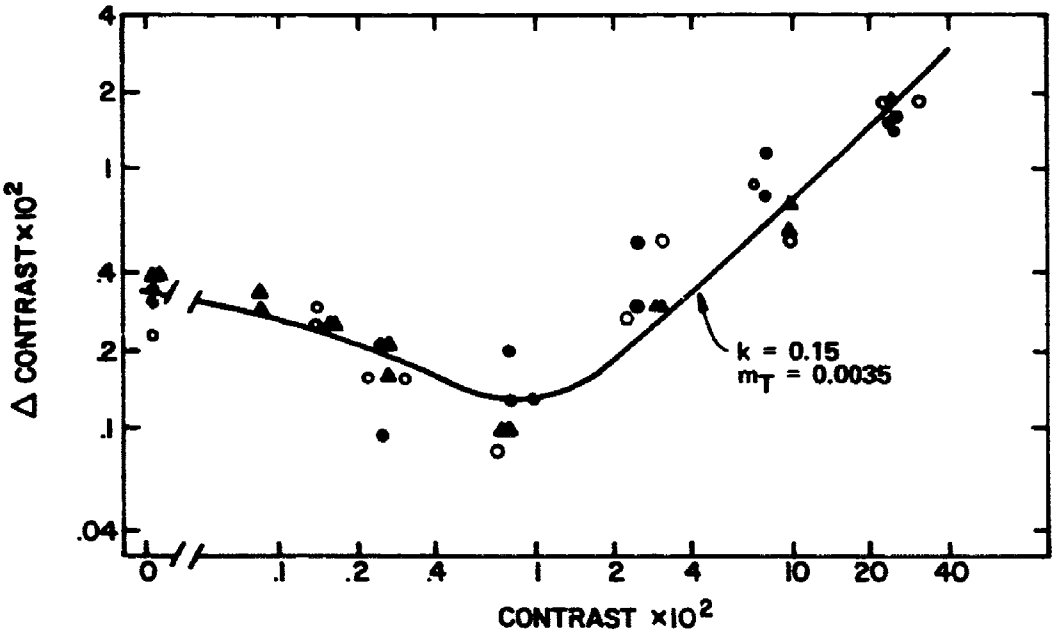


Figure 46. Required contrast increase for a jnd in contrast for 3-cycle/degree-of-vision gratings as a function of the starting contrast value. The data points are taken from the experiments of Nachmias and Sansbury [31]. Different symbols represent different observers. The solid curve is the theoretical fit to the experimental points, based on the indicated values of the threshold contrast m_T and the constant fraction k.

The observed behavior can be accounted for by a simple square-law detection model. Such a model has been discussed in the psychophysics literature [31,34] but, to the best of our knowledge, has not been applied to an analysis of the results of the measurements of the jnd in contrast. We assume that the relevant psychophysical stimulus in the discrimination experiments is the difference ΔL^2 in the mean-square luminance. If the initial contrast is m_o, the expression for ΔL^2 is

$$\Delta L^2 = \frac{1}{2}(m_o + \Delta m)^2 L_o^2 - \frac{1}{2} m_o^2 L_o^2 \tag{10}$$

where L_o is the average luminance of the displayed gratings. The crucial statement of the model is as follows: In order for a contrast difference to be perceived with a given probability, the change in mean-square luminance must be equal to a constant fraction of the interfering signal. In general, the interfering signal consists of the sum of contributions from visual noise, random display noise, and the initial value of the mean-square luminance. Mathematically, this statement may be written in the form

$$\Delta L^2 = k[N_v(\nu)\Delta\nu + N(2\pi\nu/r)2\pi\Delta\nu/r + \frac{1}{2} m_o^2 L_o^2] \tag{11}$$

where k is a constant fraction, $N_v(\nu)$ is the visual-noise spectral powe per unit retinal frequency, $N(\omega)$ is the display-noise power spectrum as a function of the angular frequency on the display screen $\omega = 2\pi\nu/r$ (r is the viewing distance), and $\Delta\nu$ is the width of the appropriate spatial-frequency-specific channel of the visual system [34].

For the contrast discrimination experiments of Nachmias and Sansbury [31], we may take $N(\omega) = 0$. Then, combining Eqs. (10) and (11) and solving for Δm, we easily obtain

$$\Delta m = -m_o + [(1 + k)m_o^2 + 2kN_v\Delta\nu/L_o^2]^{\frac{1}{2}} \tag{12}$$

Since, in the limit $m_o \to 0$, $\Delta m = m_T$, we can make the identification

$$m_T^2(\nu) = 2kN_v(\nu)\Delta\nu/L_o^2 \tag{13}$$

Thus, Eq. (12) becomes

$$\Delta m = - m_o + [(1 + k)m_o^2 + m_T^2]^{\frac{1}{2}} \tag{14}$$

For the case of small initial contrast, $m_o^2 << m_T^2$, Eq. (14) reduces to $\Delta m = m_T - m_o$, thereby predicting an initial increase in observer discriminability as m_o is increased from zero. On the other hand, for the case of large initial contrast, $m_o^2 >> m_T^2$, Eq. (14) becomes $\Delta m = [(1 + k)^{1/2} - 1]m_o$, indicating Weber's law behavior. The solid curve in Fig. 46 was generated from Eq. (14) by use of the value $m_T = 3.5 \times 10^{-3}$ for the threshold-con rast-sensitivity and k = 0.15. It is seen that the theoretical curve is in very good agreement with the experimental results. It should be pointed out that an analysis based on a linear detection model is incapable of reproducing the observed initial decrease of Δm and the consequent minimum at $m_o \tilde{=} 8 \times 10^{-3}$. On the other hand, models based on powers higher than the square-law assumed here may be able to account for the experimental observations, but in view of the success of the square-law hypothesis in explaining the observed $\Delta m(m_o)$, consideration of these more complicated models seems unwarranted.

Given the form of Eq. (7) for the required change Δm for a jnd in contrast, one can obtain the modulation level m_i necessary for the perception of i luminance levels at a given retinal frequency. First, Eq. (14) is written as a difference equation. Setting $m_o = m_i$ and $\Delta m = m_{i+1} - m_i$, we have

$$m_{i+1} = [(1 + k)m_i^2 + m_T^2]^{\frac{1}{2}} \tag{15}$$

Equation (15) is easily solved to give

$$m_i^2 = (m_T^2/k)[(1 + k)^i - 1] \text{ where } i = 1,2.... \tag{16}$$

Figure 47 compares the calculated distribution of levels obtained from Eq. (16), for k = 0.15, with the predictions of the linear model, Eq. (8), and the Weber's law model, Eq. (9), with $k_w = [(1 + k)^{1/2} - 1)] = 0.0724$. One can see that there are considerable differences in the predictions of the various models. Because of the initial increase in discriminability for small contrast values, the actual distribution places more perceivable levels at low

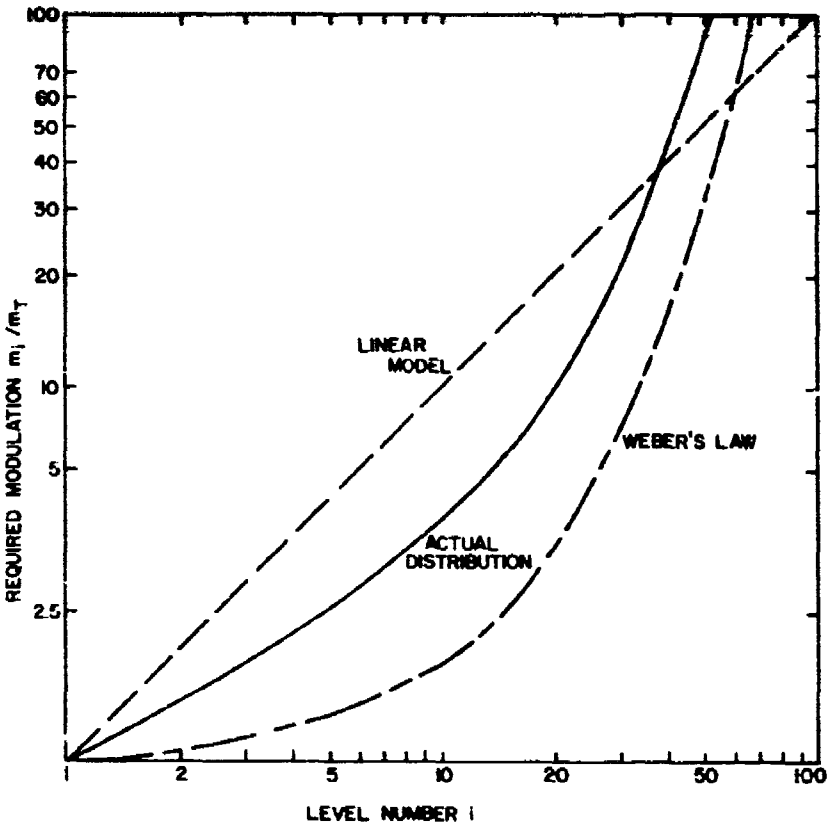


Figure 47. Computed distribution of perceived luminance levels for three model distributions. Eq. (16), with k = 0.15, was used to compute the actual distribution. Eq. (9), with $k_w = [(1 + k)^{1/2} - 1] = 0.0724$, was used to compute the Weber's law model.

contrast values (below about $40m_T$) than does the linear model. For contrast values above approximately $40m_T$, the increase of the required level spacing Δm at high contrast values dominates, so that the actual distribution requires a larger modulation depth to achieve a given number of perceivable levels than does the linear model. On the other hand, as can be seen from the figure, an extrapolation of the Weber's law regime ($\Delta m = k_w m_i$) of the actual distribution to small contrast values would consistently underestimate the modulation required to achieve a given number of perceivable levels. Thus, one cannot ignore the effect of the enhanced discriminability at low contrast levels in calculating the number of perceivable levels.

The result, presented in Eq. (16), may be employed to compute the total number of contrast levels that can be perceived, at a given retinal frequency,

without exceeding a specified contrast value. Only the threshold-contrast-sensitivity function $m_T(\nu)$ and the parameter k are required to perform the computation. In the following, we shall employ the result for $m_T(\nu)$, presented on p. 95 of TR2, for a 6.5° display* with an average luminance of 34 mL and surround brightness of 3.4 mL. These viewing conditions were selected as typical of a wide variety of display situations. The value k = 0.15 will be used and assumed to be independent of frequency. Although data for $\Delta m(m_o)$ was presented for only one retinal frequency (3 cycles/degree-of-vision), it is reasonable to assume that k is slowly varying with frequency. This assumption implies that each frequency-specific channel of the visual system is endowed with a detector of equal sensitivity.

As a result of assuming a frequency-independent k, the computed contrast sensitivity $1/\Delta m$ as a function of ν should become relatively independent of frequency as the initial contrast m_o is increased. This effect is shown in Fig. 48. As m_o is increased, the low-frequency roll-off of the threshold-sensitivity curve rapidly disappears, and the curves develop a wide-band region of almost constant sensitivity. (Naturally, at $\nu = 0$, the contrast sensitivity must always vanish.) The computed flattening of the sensitivity curves, and hence the assumption of a frequency-independent k, is consistent with the results of the suprathreshold contrast-matching experiments of Georgeson and Sullivan [32]. These authors observed that the relative perceived contrast of pairs of gratings of different retinal frequencies become relatively independent of frequency as the grating contrast increased.

2. Number of Perceivable Contrast Levels

To compute the maximum number of perceivable contrast levels as a function of ν, we set $m_i = 1$ and $i = n(\nu) - 1$ in Eq. (16). Solving for $n(\nu)$, we have

$$n(\nu) = 1 + \frac{\ln\left[1 + k/m_T^2(\nu)\right]}{\ln\,(1 + k)} \qquad (17)$$

where k = 0.15 (100% modulation)

*Rigorously, $m_T(\nu)$ is independent of display size. We have found (TR2, p. 95) that, for display sizes subtending more than about 6° of viewing angle, the measured $m_T(\nu)$ varies slowly with display size over the frequency range of interest.

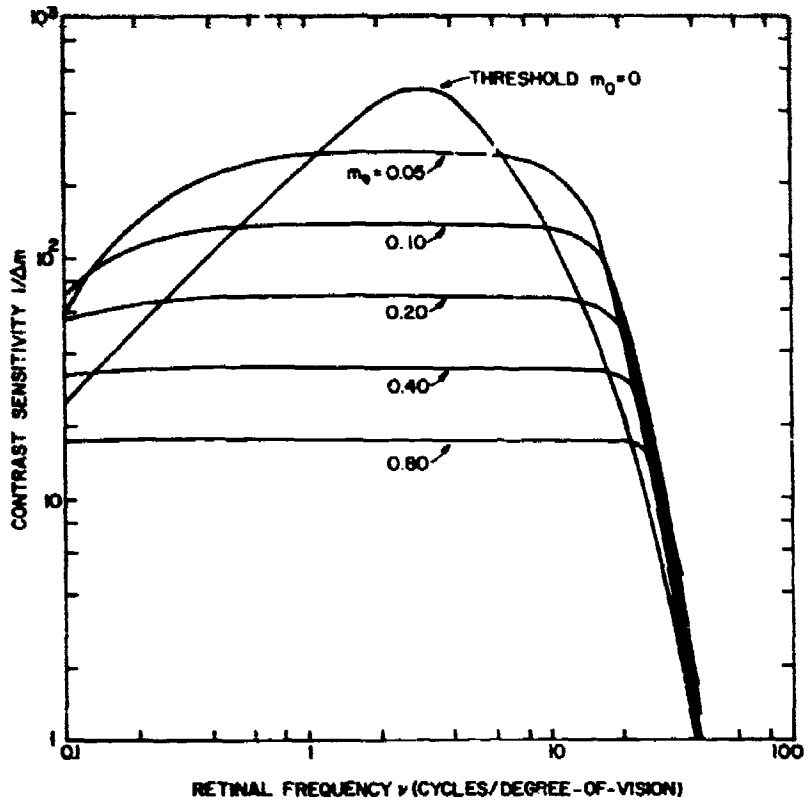


Figure 48. Computed contrast sensitivity $1/\Delta m$ as a function of retinal frequency ν for various starting contrast levels m_0. The quantity Δm is the required contrast increase for a jnd in contrast.

Equation (17) and the measured $m_T(\nu)$ were used to compute the $n(\nu)$ indicated by the upper curve in Fig. 49. The maximum value of $n(\nu)$ is approximately 76 and is achieved at the frequency $\nu \tilde{=} 3$ cycles/degree-of-vision, corresponding to the peak of the threshold-contrast-sensitivity curve. This result tells us that, in principle, we need not employ more than about six bits of picture information* for the assumed viewing conditions (34-mL display luminance, 3.4-mL surround luminance), provided the levels are spaced in accord with Eq. (16).

*It should be kept in mind that the analysis assumes a constant 79.4% correct response in the forced-choice contrast-discrimination experiments. If one were to impose a more rigorous constraint on the detectability of the contrast levels, computed values of $n(\nu)$ would naturally be larger.

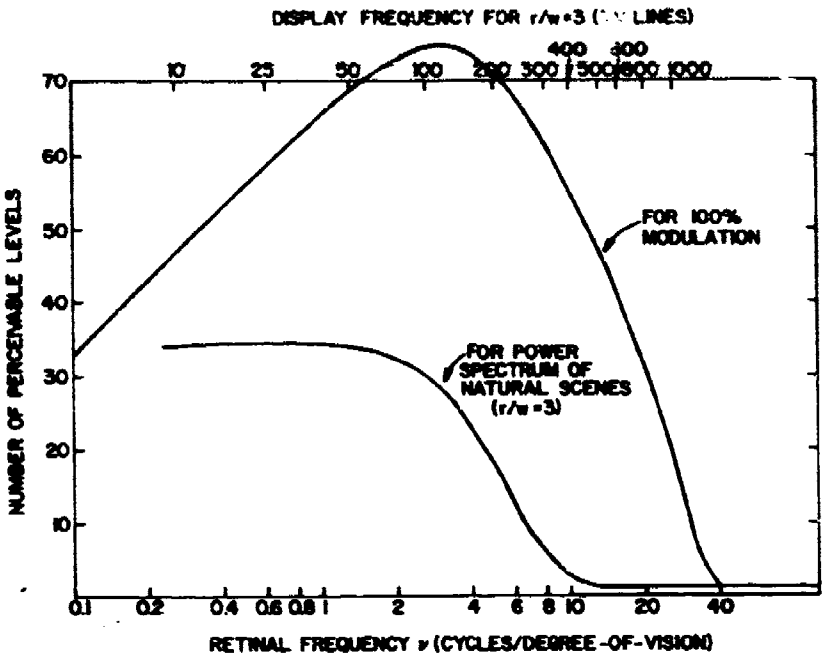


Figure 49. Computed number of perceivable luminance levels as a function of retinal frequency ν for gratings of 100% modulation and for the luminance power spectrum of natural scenes at the viewing distance $r/w = 3$. The corresponding display frequencies at $r/w = 3$ are indicated at the top of the figure.

Indeed, as will be shown below, the required number of contrast levels for the statistically averaged picture content is significantly less than even this modest figure.

Before computing the channel capacity of the display-observer system, we need to derive an expression for the number of perceivable contrast levels as a function of ν for the case where the value of the contrast at any frequency is established by the statistically averaged signal power spectrum. Real scenes do not consist of 100%-modulated sine-waves at all frequencies, so that Eq. (17) should be viewed only as providing an upper limit to the number of perceivable contrast levels. A consistent statistical communication theory approach to the calculation of the channel capacity requires the number of levels that can be transmitted through a communication channel to be computed by use of the power spectrum corresponding to the ensemble of possible inputs to the channel. The power spectrum has been the subject of considerable study during the course of this research, first in TR1 and, in more detail, in Section III of this report. We have found that, for display frequencies greater

than about 1 cycle/picture width, the modulated contribution to the luminance power spectrum for natural scenes, $\Phi_L(\omega)$, can be represented by an inverse-square-frequency characteristic*:

$$\Phi_L(\omega) = 2\omega_L[\bar{I}^2 - (\bar{I})^2]/\omega^2 \tag{18}$$

Here $\omega_L \tilde{=} 2\pi/w$ (w is the picture width) is the lower cutoff frequency, and $\bar{I}^2$ and $\bar{I}$ are the mean square and mean luminance, respectively. In order to change the power spectrum represented by Eq. (18) into an equivalent sine-wave contrast, we appeal to the concept of frequency-specific independent channels [33,34] in the human visual system. We imagine that each channel is centered about a retinal frequency $\nu = \omega r/2\pi$ and has a width $\Delta\nu$. In general, the power spectrum $\Phi_L(\omega)$ encompasses many of these channels, so that the total power in a single channel is approximately** $(1/\pi) \times (2\pi\Delta\nu/r) \times \Phi_L(\omega)$. This single-channel power can be regarded, for the purposes of detection, as equivalent to that of a single sine-wave of frequency ω whose spectral width is much less than the channel width. In other words, the independent-channel model implies that the detector for a particular channel cannot distinguish between the input for a single sine-wave whose mean-square luminance is $\frac{1}{2}m_{eq}(\omega)(\bar{I})^2$ and a continuum of frequencies with the same value of mean-square luminance contained within the channel. Thus, the power spectrum given in Eq. (18) is equivalent to a sine-wave contrast $m_{eq}(\omega)$ where***

$$\frac{1}{2}\, m_{eq}^2(\omega)(\bar{I})^2 = (2\pi\Delta\nu/r)2\omega_L[\overline{I^2} - (\bar{I})^2]/\pi\omega^2 \tag{19}$$

*The form of $\Phi_L(\omega)$, given in Eq. (18), satisfies the condition $\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi}\, \Phi_L(\omega) = \bar{I}^2 - (\bar{I})^2$ if the low-frequency behavior of $\Phi_L(\omega)$ is represented by a Lorentzian (cf. TR2, p. 46).

**We employ the Fourier representation in which the spectral power per unit of (positive) frequency is $(1/\pi)\Phi_L(\omega)$.

***To be precise, Eq. (19) may break down at frequencies sufficiently low for the channel width $\Delta\nu$ to become of the order of the channel frequency. In practice this occurs when $\nu \lesssim 1$ cycle/degree-of-vision. In that case, the specific form of the response of each channel may be important; the integral of the power spectrum over a single channel may therefore not be approximated by simply multiplying the power spectrum evaluated at the center frequency of the channel by the channel width, as assumed in Eq. (19). Nevertheless, for the sake of simplicity, we shall apply Eq. (19) at all retinal frequencies greater than that corresponding to the lower cutoff frequency ω_L.

Rearranging Eq. (19) and changing the display frequency ω to the retinal frequency coordinate ν, we can express m_{eq} as

$$m_{eq}^2(\nu) = (4/\pi)(r/w)\Delta\nu\,[\overline{I^2}/(\bar{I})^2 - 1]/\nu^2 \tag{20}$$

The number of perceivable contrast levels corresponding to the luminance power spectrum for natural scenes is now easily obtained from Eq. (16) by replacing m_i by m_{eq} and setting $i = n(\nu) - 1$. We also take into account the modulation transfer characteristics of the display by multiplying the power-spectrum Eq. (18) by $|R_{eff}(\omega)|^2$, where R_{eff} is the effective overall MTF of the display. We thus find, for $n(\nu)$,

$$n(\nu) = 1 + \frac{\ln\left\{1 + (4k/\pi)(r/w)\Delta\nu[\overline{I^2}/(\bar{I})^2 - 1]|R_{eff}(2\pi\nu/r)|^2/\nu^2 m_I^2(\nu)\right\}}{\ln\,(1+k)} \tag{21}$$

where $k = 0.15$ (power spectrum of natural scenes)

The lower curve in Fig. 49 represents the $n(\nu)$ computed from Eq. (21) for a viewing distance of three picture widths. In the calculation, we have assumed a perfect display ($R_{eff} = 1$); the channel width was taken to be 1 cycle/degree-of-vision, the value indicated by recent psychophysical measurements [33]. The value of the quantity $[\overline{I^2}/(\bar{I})^2 - 1]$ was set at 1/6, representative of the results of the statistical measurements described in Section III. A comparison of the two curves in Fig. 49 shows that the effect of the $1/\omega^2$ power spectrum is to reduce and flatten the curve of $n(\nu)$. The reduction is simply the result of the fact that the statistically averaged contrast level [Eq. (20)] is significantly less than 100%. The flattening is due to the exact compensation of the low-frequency linear roll-off of the threshold-contrast sensitivity [cf. Section V of TR2] by the inverse-square-frequency dependence of the power spectrum.

Figure 49 indicates that, for an average scene, little more than five bits of picture information are required, and that, at the viewing distance considered, negligible picture information is needed beyond a display frequency of 400 TV-lines.* The effect of viewing distance is shown in Fig. 50,

*Display frequency in TV-lines is defined here as the total number of half-cycles on the display screen: $N_{TV} = \omega w/\pi$.

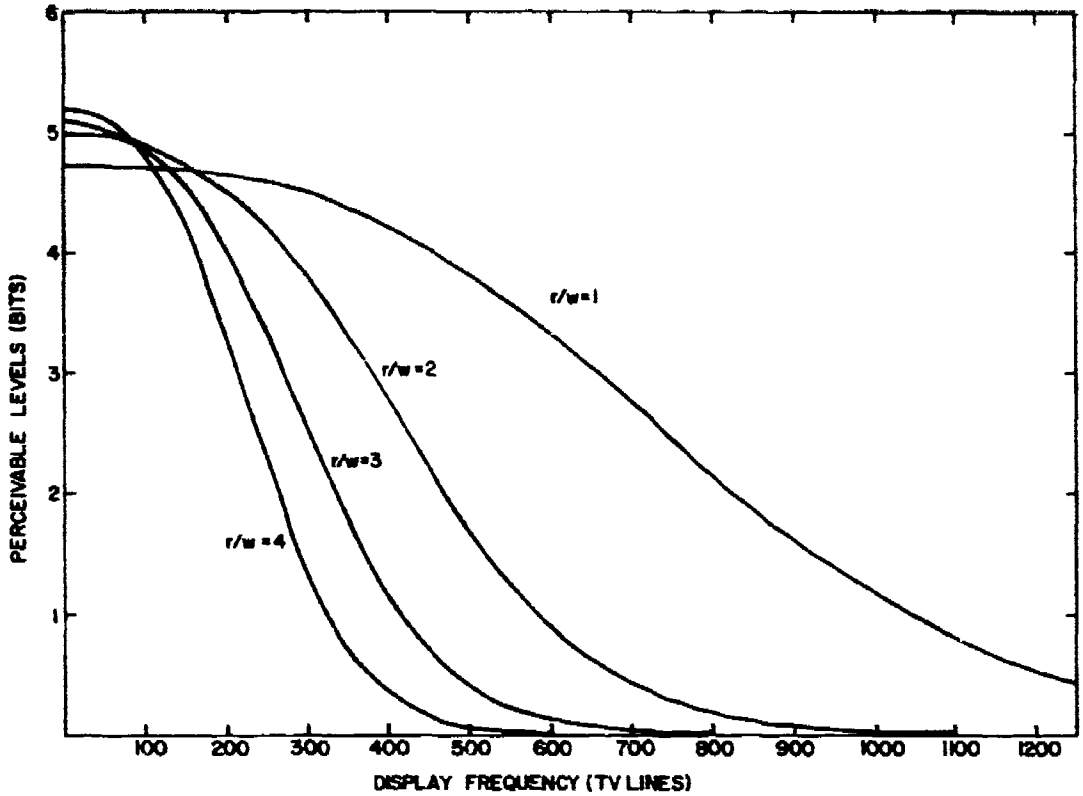


Figure 50. Computed base-2 logarithm of the number of perceivable luminance levels as a function of display frequency for various viewing distances.

where $\log_2 n(\nu)$ is plotted for r/w = 1, 2, 3, 4. One sees that the maximum number of bits remains approximately constant at about five bits. However, as one would expect, the display bandwidth requirements depend sensitively on the viewing distance, smaller bandwidths being required as the viewing distance increases.

C. THE LUMINANCE CHANNEL CAPACITY

1. Formalism

In information theory, the channel capacity is defined as the maximum number of units of binary information that can be transmitted through a communication channel. Numerous treatments of linear communications systems exist [35], but, as was seen in V.B.1 above, human perception of contrast

levels at a given retinal frequency is extremely nonlinear. In what follows, we shall take into account this nonlinear distribution by employing the result [Eq. (21)] for the number of perceivable contrast levels as a function of retinal frequency in deriving an expression for the channel capacity for luminance signals.

Before proceeding, we should like to reiterate a point made previously (in Section II.A of TR2) regarding the meaning of an information theory descriptor such as the channel capacity. Since the channel capacity is concerned only with the transmission of structural information, there is no formal explicit constraint that would require the transmitted picture to be perceived as a faithful rendition of the original. The channel capacity should be viewed as a measure of the ability of the channel - consisting of the display *and* the observer - to transmit structural information, measured in bits, from the input scene to the cognitive level of the brain. In practice, one usually expects to find a one-to-one relationship between the channel capacity and a measure of the faithfulness of reproduction such as the correlation fidelity (cf. Section II of TR2). However, the distinction between the concepts of faithfulness of reproduction and information transmission should be kept in mind.

In order to compute the channel capacity, we need an expression for the total number of distinguishable states W_L for the display-observer system. We consider only the one-dimensional case here: the extension to two dimensions by means of the method of Section IV of TR2 is trivial. For a given display frequency ω, there are $n(\omega)$ perceivable contrast levels. We choose a set of quantized display frequencies in a manner that will allow only integer multiples of the fundamental angular frequency $\omega_L = 2\pi/w$; i.e., $\omega = \omega_j = j\omega_L$ (j is an integer). Thus, in the frequency interval between ω_j and $\omega_j + \Delta\omega$, where $\Delta\omega = 2\pi\Delta j/w$, there are $[n(\omega_j)]^{\Delta j}$ distinguishable states of the display-observer system, provided $n(\omega)$ is slowly varying over the range $\omega_j \leq \omega \leq \omega_{j+\Delta j}$. For a system of independent frequencies, the total number of distinguishable states then is

$$W_L = \prod_{\omega_j} \left[n(\omega_j)\right]^{(w\Delta\omega/2\pi)} \tag{22}$$

The luminance channel capacity H_L is the base-two logarithm of W_L [36,37]:

$$H_L = \sum_j (w\Delta\omega/2\pi)\log_2 n(\omega_j) \tag{23}$$

When the important display frequencies are much larger than ω_L, the sum in Eq. (23) may be replaced by an integral. Proceeding to the limit of an infinitesimal $\Delta\omega$, Eq. (23) becomes

$$H_L = w\int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \log_2 n(\omega) \tag{24}$$

Equation (21) for the number of perceivable levels may be employed in Eq. (24) for the case of noiseless displays, $N(\omega) = 0$. However, we wish to include the effect of display noise on the luminance channel capacity. This task is easily accomplished by referring to Eq. (11), where it is seen that display noise is added to the visual-noise power spectrum in a manner that replaces the visual noise $N_v(\nu)\Delta\nu$ by the quantity $[N_v(\nu)\Delta\nu + N(2\pi\nu/r)2\pi\Delta\nu/r]$. Then, employing Eq. (13), which links $N_v(\nu)$ to the observable quantity $m_T^2(\nu)$, we find that Eq. (21) must be modified so that $m_T^2(\nu)$ is replaced by the quantity $[m_T^2(\nu) + 2k(2\pi\Delta\nu/r) \times N(2\pi\nu/r)|R_{eff}(2\pi\nu/r)|^2/(\bar{I})^2]$. Thus, expressed in terms of the display frequency coordinate ω, Eq. (21) for $n(\omega)$ becomes

$$n(\omega) = 1 + \frac{\ln\left\{1 + \frac{4k}{\pi}\frac{(\omega_L 2\pi\Delta\nu/r)\ [\overline{I^2}/(\bar{I})^2 - 1]\ |R_{eff}(\omega)|^2}{\omega^2[m_T^2(\omega r/2\pi) + 2k(2\pi\Delta\nu/r)N(\omega)|R_{eff}(\omega)|^2/(\bar{I})^2]}\right\}}{\ln(1+k)} \tag{25}$$

where $\omega_L = 2\pi/w$; $k = 0.15$

2. Properties of the Luminance Channel Capacity

Equations (24) and (25) constitute our result for the luminance channel capacity of the display-observer system. The quantity H_L is a function of the viewing distance and depends on the display modulation transfer function $R_{eff}(\omega)$, the display signal-to-noise ratio (through the noise power spectrum $N(\omega)$), and the statistics of the input scenes (through the quantities ω_L and $[\overline{I^2}/(\bar{I})^2 - 1]$). As a function of viewing distance, H_L exhibits the same general characteristics

as the visual capacity [38]; it rises from zero, achieves a broad maximum at a display-dependent viewing distance, and falls off at larger viewing distances. We have considered the simple case of a noiseless television system with an ideal, flat passband [$R_{eff}(\omega) = 1$ for $|\omega| \leq \omega_M$ and $R_{eff}(\omega) = 0$ for $|\omega| > \omega_M$]. We find* that, for a passband corresponding to $N_{TV} = \omega_M w/\pi = 300$ TV-lines, the computed maximum of H_L is approximately 1400 bits and occurs at a viewing distance of about two picture heights (3/2 picture widths). This value of H_L corresponds to an average of nearly five bits over the allowed passband, as one would expect in view of the results shown above (in V.B.2; cf. Fig. 50). The value of the viewing distance for maximum H_L is smaller by a factor of 2 than the value computed for maximum visual capacity [38]. The two descriptors represent two entirely different perceptual quantities, sharpness and total information, so that the viewing distance that optimizes one of these quantities need not optimize the other.

The dependence of H_L on display bandwidth at a fixed viewing distance $r/w = 3$ is shown in the uppermost curve of Fig. 51. Once again, for the sake of simplicity, we have assumed an ideal low-pass filter for $R_{eff}(\omega)$. As the bandwidth is increased, H_L first rises linearly with bandwidth but eventually bends over and approaches an asymptotic value. The computed asymptotic value of H_L at this viewing distance is approximately 1590 bits. As can be seen from Fig. 51, an increase in the display bandwidth beyond about 400 TV-lines has a negligible effect on H_L.

The effect of display noise on H_L can be computed by means of Eq. (25) for $n(\omega)$. We treat the case of a white-noise spectrum,

$$\begin{aligned} N(\omega)|R_{eff}(\omega)|^2 &= \pi N_s^2/\omega_M \qquad \text{for } |\omega| \leq \omega_M \\ &= 0 \qquad\qquad\quad \text{for } |\omega| > \omega_M \end{aligned} \tag{26}$$

where N_s^2 is the mean-square noise-luminance fluctuation, as measured on the display screen, and ω_M is the maximum display frequency. Substituting Eq. (26)

*As in the computations of Section V.B, we employed the measured $m_T(\nu)$ for 34-mL display brightness and 3.4-mL surround brightness. The channel width $\Delta\nu$ was taken to be 1 cycle/degree-of-vision, and the quantity $[\overline{I^2}/(\bar{I})^2 - 1]$ was set equal to 1/6.

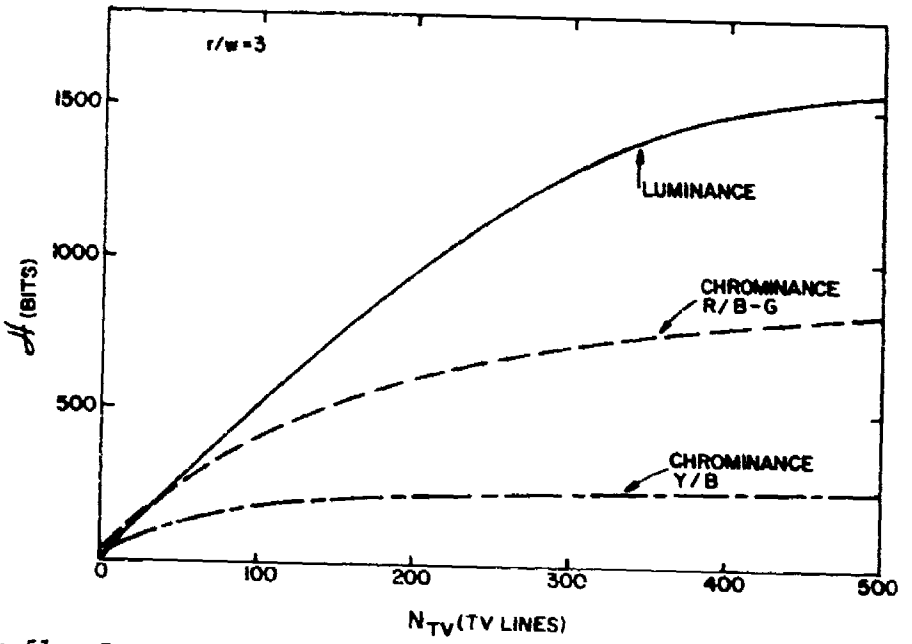


Figure 51. Computed channel capacities of the display-observer system as a function of display bandwidth at the viewing distance r/w = 3. Curves are shown for the one luminance and two chrominance channels, Red/Blue-Green and Yellow/Blue.

into Eq. (25), we see that display noise becomes an important factor when the following approximate equality holds:

$$m_T^2(\omega r/2\pi) \approx 2\pi k \frac{\Delta\nu}{(\omega_M r/2\pi)} \frac{N_s^2}{(\bar{I})^2} \tag{27}$$

Next, we define the display signal-to-noise ratio S/N as the ratio of the mean luminance $\bar{I}$ to the rms noise fluctuation N_s. Then, taking k = 0.15, $\Delta\nu$ = 1 cycle/degree-of-vision, Eq. (27) becomes

$$m_T^2(\omega r/2\pi) \approx 54\ [(\omega_M r/2\pi)(S/N)^2]^{-1} \tag{28}$$

This equation predicts that, for the example of a 300 TV-line display at a viewing distance of three picture widths, display noise becomes important when $S/N \approx 0.35\ m_T^{-1}(\nu)$. Now, the experimental values of $m_T^{-1}(\nu)$ average about 1/0.004 = 250 over the allowed frequency range. Thus, we expect the effect of display noise to become significant when $S/N \lesssim 90$, a value in agreement with practical display experience.

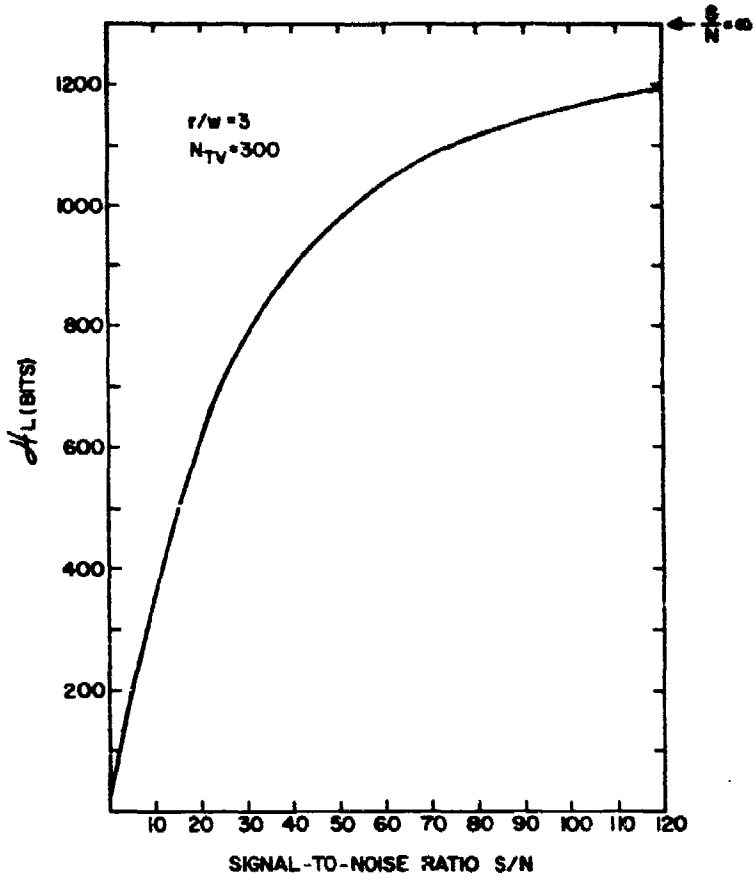


Figure 52. Computed luminance channel capacity H_L as a function of display signal-to-noise (S/N) ratio at the viewing distance r/w = 3. The display passband is represented by an ideal low-pass filter with a bandwidth of 300 TV-lines. The power spectrum of the noise was assumed to be white. The value of H_L for S/N = ∞ is indicated in the figure.

The result of a detailed calculation of H_L as a function of S/N for our hypothetical 300 TV-line display, viewed at r/w = 3, is shown in Fig. 52. It is seen that the computed H_L is reduced by only 10% from its noisefree value when S/N = 100. However, for values of S/N below about 50, H_L falls off rapidly. This result is illustrated in more practical terms in Fig. 53. There, we have plotted the equivalent noisefree bandwidth N_{TV}^{eq} against S/N. The quantity N_{TV}^{eq} represents the bandwidth necessary for a noisefree display to match the luminance channel capacity of a 300 TV-line display with the appropriate value of S/N. For example, a 300 TV-line display with S/N = 100 is equivalent,

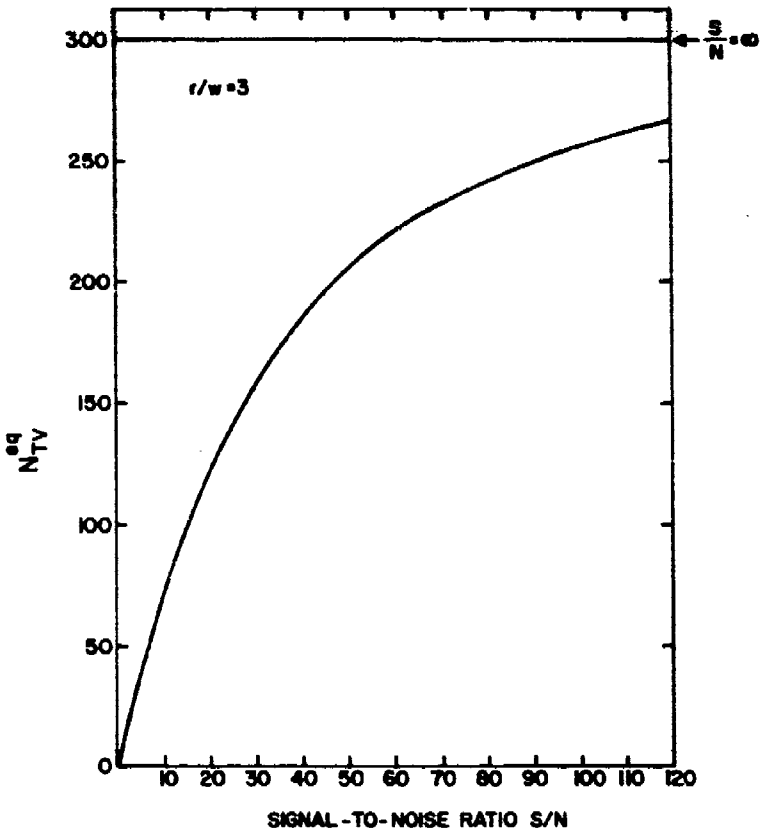


Figure 53. Computed equivalent noiseless display bandwidth N_{TV}^{eq} as a function of display signal-to-noise (S/N) ratio at the viewing distance r/w = 3. The display passband is represented by an ideal low-pass filter with bandwidth of N_{TV}^{eq}. The power spectrum of the noise was assumed to be white. The value N_{TV}^{eq} = 300 is indicated in the figure for S/N = ∞.

in luminance channel capacity, to a 256 TV-line noisefree display. Using the results of Section IV; we estimate that the 256 TV-line display is, perceptually, less than 2 jnd's from the 300 TV-line display. Thus, a display signal-to-noise ratio of 100 is very good indeed in this case. However, for values of S/N much below 100, N_{TV}^{eq} decreases rapidly. At S/N = 50, the equivalent bandwidth is below 200 TV-lines, as a result of which the display has been seriously degraded.

It should be borne in mind that the numerical values of H_L are derived from a specific input power spectrum: that for natural pictorial scenes.

Naturally, the values of H_L and conclusions derived from them may change markedly if other power spectra are employed. For example, the input from alphanumeric displays with random character positions may be characterized by a white-noise power spectrum for frequencies below roughly the inverse of the characteristic width of an alphanumeric character, a form entirely different from that for natural scenes [Eq. (18)]. For such an ensemble of input scenes, Eq. (25) for $n(\omega)$ must be modified by replacing the power spectrum term $2\omega_L[\bar{I^2} - (\bar{I})^2]/\omega^2$ by the appropriate white-noise power spectrum.

Finally, it should be noted that H_L is, in general, a luminance-dependent quantity. For displays operating above about 10 mL, the dependence of H_L on luminance is negligible. For smaller luminance values, it is possible to include the effect of luminance on H_L in a simple way by letting the threshold-contrast-sensitivity function $m_T(\nu)$ vary with luminance. We assume that the visual noise $N_v(\nu)$ consists of the sum of an internal-noise term $N_{vI}(\nu)$, a property of the visual system, and a shot-noise term $N_{vS}(\nu)$ that arises from the finite number of photons arriving at the retina. These contributions can be separated by examining the experimental $m_T(\nu)$ as a function of luminance. From fundamental statistical considerations, we expect $N_{vI}(\nu)$ to be proportional to the square of the average luminance, whereas $N_{vS}(\nu)$ should be proportional to the average luminance itself. Then, Eq. (13) for $m_T^2(\nu)$ can be written in the form

$$m_T^2(\nu) = A(\nu) + B(\nu)/\bar{I} \tag{29}$$

where $A(\nu)$ and $B(\nu)$ are independent of luminance. The functions $A(\nu)$ and $B(\nu)$ can be determined by an analysis of existing $m_T(\nu)$ data at various values of the mean luminance. We have verified the form of Eq. (29) using the data presented in Fig. 33 of TR1. As expected, the contribution of the last term in Eq. (29), due to photon shot noise, was found to be negligible for luminance values above about 10 mL. Thus, the function $A(\nu)$ represents the $m_T^2(\nu)$ employed in the numerical calculations presented here.

D. THE CHROMINANCE CHANNEL CAPACITY

1. Independent Channel Model

In order to include the effect of color in our information theory approach to image descriptors, we must utilize a model to describe the processing of chrominance information by the human visual system. There is some fairly strong evidence elucidating the probable action of the color mechanism in humans. It is known that, at the retina, light is received by individual "red," "green," and "blue" color receptors (cones), as stipulated by the Young-Helmholtz theory. However, the output of these receptors appears to be coded as the sum of a luminance signal, which adds contributions from all three receptors, and two color-difference signals [39,40]: Red vs Green and Yellow vs Blue. Thus, the total signal is transmitted to the brain in three independent channels - one for luminance and two for chrominance.

The model of three independent processing channels allows us to formulate, in a straightforward way, an extension of the theory presented in Section V.C. to include the effect of chrominance information. We seek an expression for the total channel capacity H of the display-observer system. To obtain the total channel capacity, we first write the total number of distinguishable states W of the display-observer system as

$$W = W_L W_{C1} W_{C2} \tag{30}$$

where W_{C1} and W_{C2} represent the contributions of the two chrominance channels, and W_L is the luminance contribution discussed above (Section V.C). The simple product form of Eq. (30) is the result of the hypothesis of independent luminance and chrominance channels in the visual system. If the channels did not carry on their respective processing functions independently, mixing between the channels would occur, and the resulting expression for W would be far more complicated than the simple form of Eq. (30).

As a result of the product form of Eq. (30), the total channel capacity is simply the sum of contributions from the three independent channels,

$$\begin{aligned} H &= \log_2 W \\ &= H_L + H_{C1} + H_{C2} \end{aligned} \tag{31}$$

The contributions H_{C1} and H_{C2} of the two chrominance channels to the total channel capacity are simply

$$H_{C1} = \log_2 W_{C1}$$
$$H_{C2} = \log_2 W_{C2} \tag{32}$$

The luminance contribution H_L was previously discussed in detail (see Section V.C).

The technique of Section V.C.1, employed for the computation of the luminance channel capacity, may be used to express the quantities H_{Ci} (i = 1,2) as integrals over spatial frequency of the base-2 logarithm of the number of perceivable chrominance levels n_{Ci} at a given spatial frequency:

$$H_{Ci} = w \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \log_2 n_{Ci}(\omega) \text{ where } i = 1,2 \tag{33}$$

The meaning of the quantities n_{Ci} is as follows: For the luminance case, the quantity $n(\omega)$, appearing in Eq. (24) for H_L, represents the number of perceivable brightness levels for a given spatial frequency and for constant chrominance. The quantities $n_{Ci}(\omega)$ represent the number of perceivable colors for a given spatial frequancy and luminance value, along the Red/Green and Yellow/Blue axes. Perceptually, $n(\omega)$ and $n_{Ci}(\omega)$ appear to be entirely different quantities. In the sense of information theory, however, all of these quantities appear on an equal footing.

2. Chrominance Contrast-Sensitivity Functions and Equivalent rms Chrominance Modulation

It now remains to obtain expressions for the quantities $n_{Ci}(\omega)$. In principle, we should require the same kind of experimental information for the chrominance case that the work of Nachmias and Sansbury [31] provided for the luminance case. That is, given a chrominance sine-wave grating of given modulation depth (color excursion) and spatial frequency, how much additional modulation must be supplied in order for a perceivable color difference to be detected? Unfortunately, we know of no such detailed measurements. However, the experiments of MacAdam [41] do give considerable information regarding the distribution of chrominance levels in the low-frequency limit. These experiments

indicate that, in the vicinity of white, the required chrominance change for a jnd in color varies relatively slowly along any direction in the CIE chromaticity diagram. This observation indicates that, as long as we deal only with small excursions from white (relatively desaturated colors), a model based on a linear distribution of perceivable chrominance levels with fractional color saturation as the relevant psychophysical coordinate may be appropriate. The results of Section III show that, for natural scenes, color excursions are indeed small on a statistical basis. Therefore, in order to compute the chrominance contribution to the total channel capacity, we shall utilize a linear model for the distribution of perceived chrominance levels.

We proceed by analogy to the luminance case and postulate two threshold modulation functions for chrominance $c_{Ti}(\nu)$ [i = 1,2]. The functions $c_{Ti}(\nu)$ represent the required fractional saturation modulation for an observer to perceive a color change when presented with a chrominance sine-wave grating centered about white. Figures 54 and 55 reproduce the experimental data of van der Horst and Bouman [42] for chrominance gratings along the respective Red/Blue-Green and Yellow/Blue axes of the chromaticity diagram. The modulation was centered about the equal-energy white point. Also shown in these figures are the solid curves representing simple empirical fits to the experimental data. At a retinal illumination of 160 td ($\sim$ 90-mL source brightness), the approximate expressions for the $c_{Ti}(\nu)$, where ν is in cycles/degree-of-vision, are*

$$c_{T,R/Bl-G}(\nu) = 0.0049[1 + (\nu/4.5)^3]^{1/2}$$
$$c_{T,Y/B}(\nu) = 0.0108[1 + (\nu/3)^3]^{1/2} \qquad (34)$$

Several important remarks should be made regarding these functions. First, it should be noted that, in contrast to the properties of the luminance-sensitivity functions [cf. Fig. 48], there is no evidence of a low-frequency roll-off in the chrominance-threshold-sensitivity functions; the low-frequency behavior of

*The sensitivity scale for the $c_{Ti}(\nu)$ is defined as the required distance in the CIE diagram for an observer to perceive a jnd of purity, in units of the fraction of the distance between the equal-energy white point and the dominant wavelength. The dominant wavelength was 492 nm for Red/Blue-Green modulation and 573 nm for Yellow/Blue modulation.

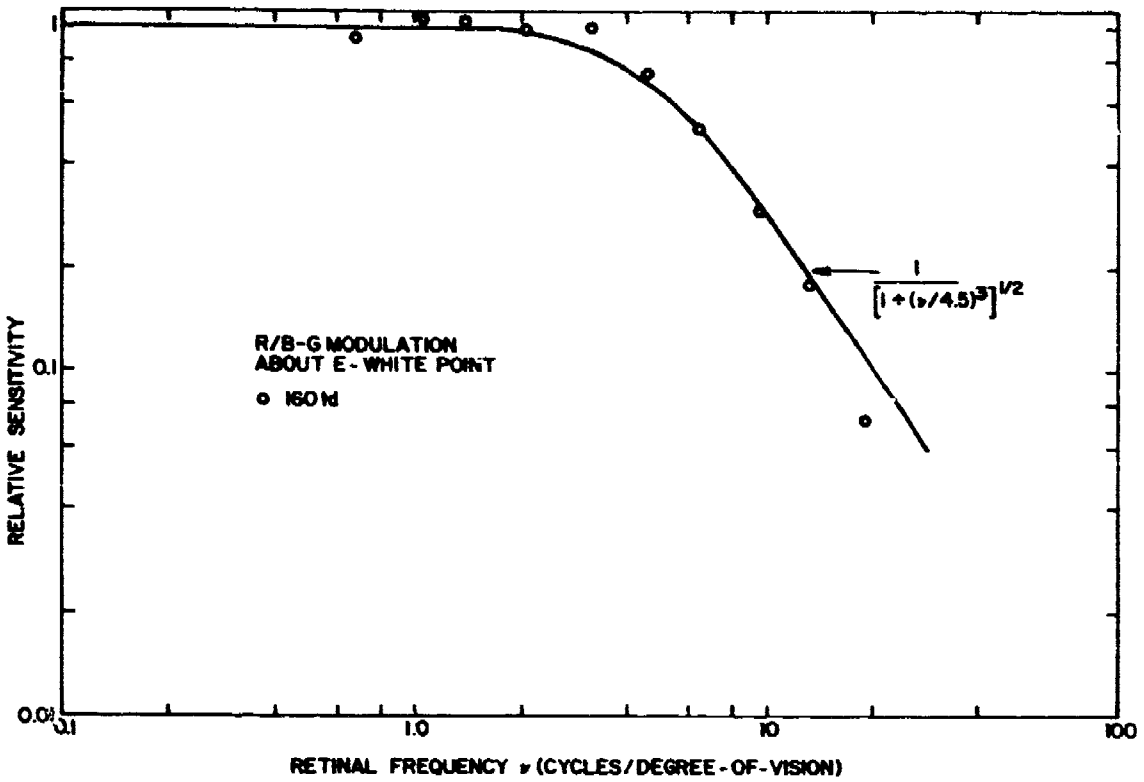


Figure 54. Relative threshold sensitivity for chrominance gratings with amplitude directed along the Red/Blue-Green line in the CIE chromaticity diagram. Zero-amplitude corresponds to the equal-energy (E) white point. The experimental points are taken from van der Horst and Bouman [42]. The value of the retinal illumination was 160 td. A simple empirical fit is indicated by the solid curve.

the $c_{T1}(\nu)$ is flat. Second, the maximum sensitivity of the Red/Blue-Green axis is larger by more than a factor of 2 than that for the Yellow/Blue axis. Finally, the frequencies at which the two threshold-sensitivity functions roll off differ from each other by less than a factor of 2, but neither function rolls off faster than the luminance-threshold function given in Fig. 48.* These properties should be borne in mind, for they will play an important part in understanding the results presented later in this section.

*It is likely that, above about 20 cyles/degree-of-vision, the chrominance-threshold-sensitivity functions roll off more rapidly than would be predicted by Eq. (34).

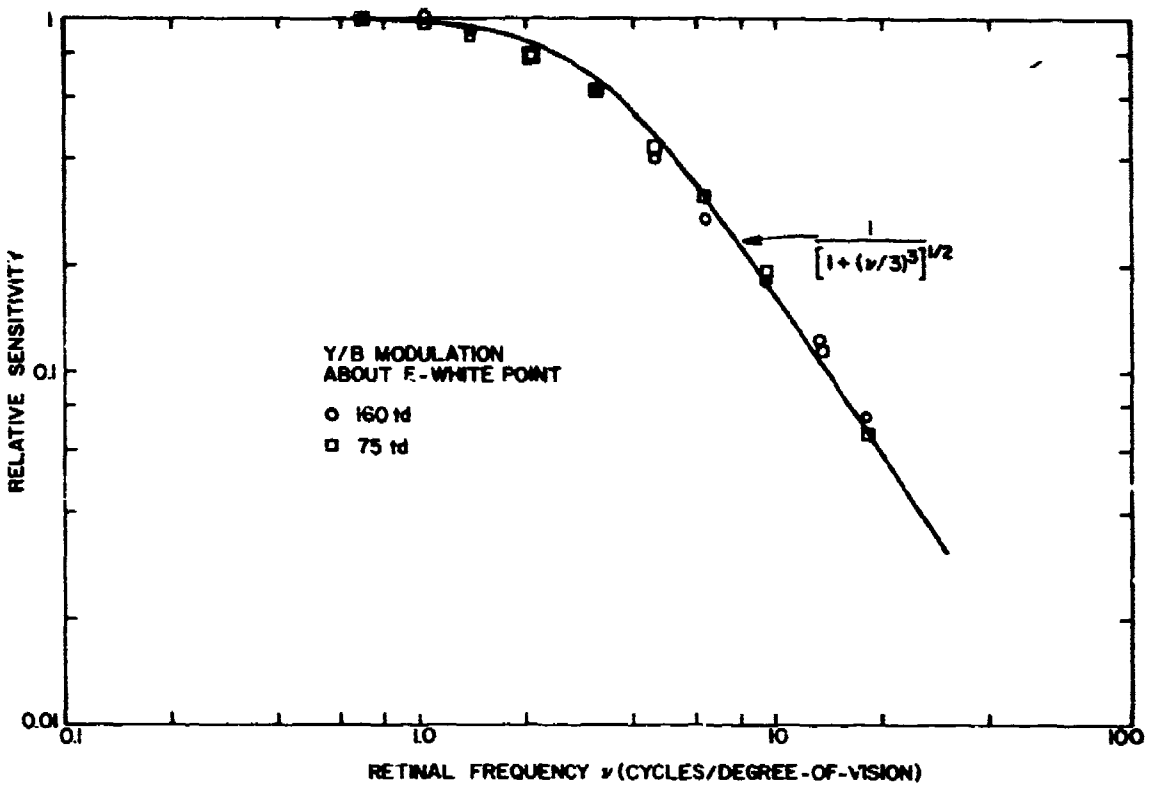


Figure 55. Relative threshold sensitivity for chrominance gratings with amplitude directed along the Yellow/Blue line in the CIE chromaticity diagram. Zero-amplitude corresponds to the equal-energy (E) white point. The experimental points are taken from van der Horst and Bouman [42]. The values of the retinal illumination were 160 and 75 td. A simple empirical fit is indicated by the solid curve.

Next, as was done in the derivation of the luminance channel capacity, we represent the statistically averaged chrominance power spectrum by an equivalent rms chrominance sine-wave modulation $c_{eq,i}$ for each of the two chrominance channels $i = 1,2$ of the visual system. This representation rests on the assumption of the existence of frequency-specific independent channels for the processing of chrominance information, just as is observed for luminance sine-waves [34]. Whereas there is considerable experimental evidence regarding the existence of such channels for the luminance case, we know of no experimental work that has probed for indications of such channels in chrominance spatial-frequency processing. Nevertheless, we shall assume that such channels do in fact exist, an assumption justified, at this point, only by indirect experimental evidence

[43] and by the physiological and psychophysical similarities of luminance and chrominance processing in the visual system.

With the assumption of frequency-specific chrominance channels, the derivation of an expression for $c_{eq,i}$ proceeds in exactly the same manner as that which led to Eq. (20) for the equivalent contrast m_{eq} for the luminance case. We make use of the experimental observation (Section III) that the chrominance power spectrum can be represented by an inverse-square-frequency characteristic above a lower cutoff frequency $\omega_L \tilde{=} 2\pi/w$ with vanishing average chrominance. The expression for $c_{eq,i}$ is directly analogous to Eq. (20):

$$c_{eq,i}^2 = (4/\pi)(r/w)\Delta\nu_{Ci}\,\overline{c_{m,i}^2}/\nu^2 \quad \text{where } i = 1,2 \tag{35}$$

Here the quantity $\Delta\nu_{Ci}$ represents the width of the frequency-specific channel for axis i, and $\overline{c_{m,i}^2}$ is the corresponding mean-square chrominance modulation.

We shall treat only the case of displays with no chrominance noise. Then, for a linear distribution of perceived chrominance levels, the number of levels n_{Ci} is given by the expression

$$n_{Ci} = c_{eq,i}/c_{Ti} + 1 \quad \text{where } i = 1,2 \tag{36}$$

Just as in the luminance case, the frequency response characteristics of the display are accounted for by multiplying the right-hand side of Eq. (35) for $c_{eq,i}^2$ by $|R_{eff,Ci}|^2$, the square of the effective overall chrominance MTF of the display. Then substituting Eq. (35) into Eq. (36) and expressing the result in terms of the display frequency coordinate ω, we have

$$n_{Ci}(\omega) = [8\omega_L\Delta\nu_{Ci}|R_{eff,Ci}|^2/r]^{1/2}\sqrt{\overline{c_{m,i}^2}}/\omega c_{Ti}(\omega r/2\pi) + 1$$
$$\text{at } \omega_L = 2\pi/w$$

Equations (33) and (37) constitute our result for the contribution of the chrominance channels to the total channel capacity.

3. Application of the Chrominance Channel Capacity

In Fig. 56 is shown the computed number of perceivable chrominance levels n_{Ci} as a function of retinal frequency for chrominance sine-wave gratings of

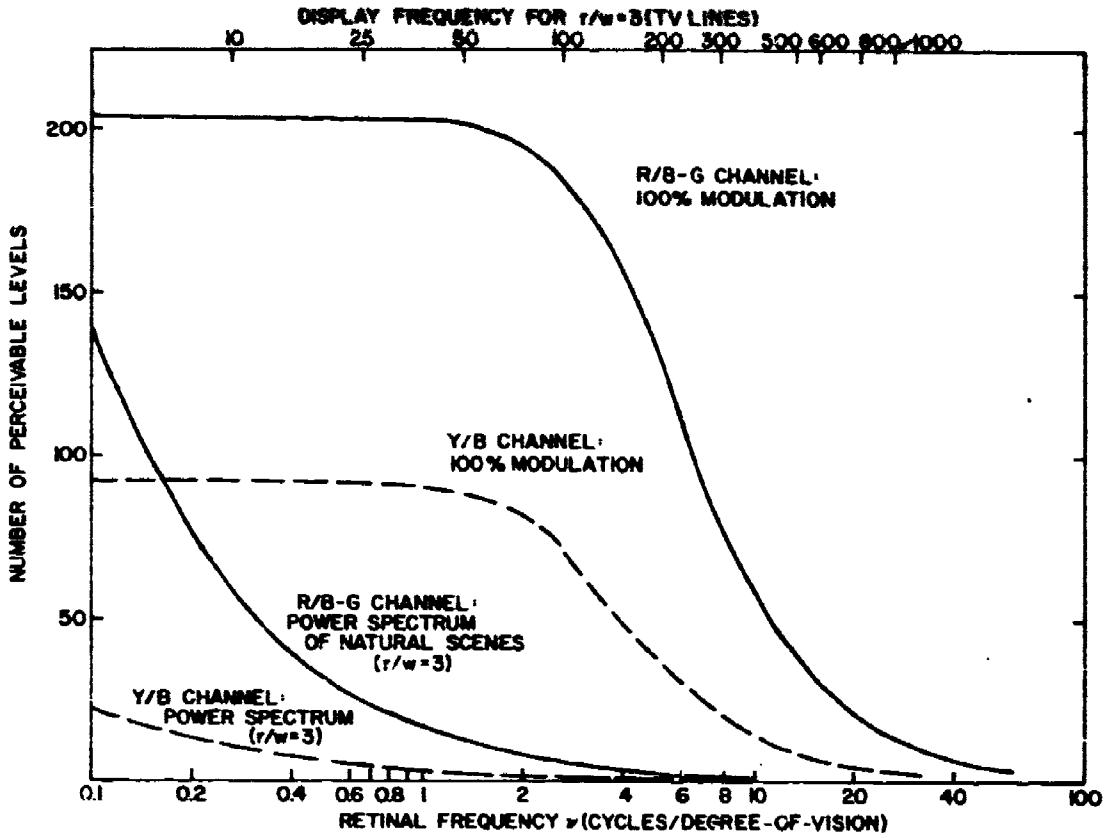


Figure 56. Computed number of perceivable chrominance levels as a function of retinal frequency ν for gratings of 100% modulation (full saturation) and for the chrominance power spectrum of natural scenes at the viewing distance r/w = 3. The corresponding display frequencies at r/w = 3 are indicated at the top of the figure. Curves are shown for the Red/Blue-Green and Yellow/Blue chrominance channels.

100% modulation and for the chrominance power spectrum of natural scenes at the viewing distance r/w = 3. For computations based on the power spectrum of natural scenes, Eq. (37) was employed; a perfect display $R_{eff,Ci} = 1$ was assumed, and the width of the frequency-specific channels $\Delta\nu_{Ci}$ was taken to be 1 cycle/degree-of-vision, the same value as was used in the luminance case [33]. The value of the rms chrominance modulation was taken to be 0.30 for the Red/Blue-Green channel and 0.10 for the Yellow/Blue channel. These values are typical of the largest measured values for the I_C and Q_C axis, respectively*

*We neglect the small angular rotation ($\sim 20°$) between the I_C, Q_C axes in the chromaticity diagram, along which measurements of Section III were taken, and the Red/Blue-Green and Yellow/Blue axes, for which the perceptual data [42] are available.

(see Section III). The use of smaller values of the rms chrominance modulation will reduce the computed n_{Ci} accordingly.

From Fig. 56, it is seen that the Red/Blue-Green channel contributes far more perceivable chrominance levels than does the Yellow/Blue channel. This observation is particularly true for the curves calculated for the case of the power spectrum of natural scenes. It is due to the combination of the higher inherent sensitivity of the Red/Blue-Green channel, as indicated by Eq. (34), and the larger observed rms chrominance modulation for the I_C axis, as discussed in Section III.

In comparing the results shown in Fig. 56 with those shown in Fig. 49 for the luminance case, the following important differences are noted: first, at the lowest retinal frequencies, the chrominance channels provide a far greater number of perceivable levels than the luminance channel; second, for the power spectrum case, the chrominance levels roll off rapidly with increasing retinal frequency, whereas the luminance levels are spread over a relatively wide band of frequencies. This fundamentally different behavior is due to the absence of a low-frequency roll-off in the chromaticity-sensitivity functions. The good low-frequency sensitivity of the chromaticity channels, combined with the inverse-square frequency power spectra, concentrates most of the available perceivable levels at very low frequencies. It should be emphasized that this result does not depend on the relative passband widths of the various channels; that is, the acuity or high-frequency cutoff of the sensitivity functions is not involved in determining this fundamental behavior.

In Fig. 51 are shown the computed chrominance channel capacities H_{Ci} as a function of display bandwidth. Equations (33) and (37) were employed in the calculations. The display chrominance passbands were represented by an ideal low-pass filter. It is seen from this figure that the computed H_{Ci} rise rapidly at first, but soon bend over and approach an asymptotic value. Indeed, the contribution of the Red/Blue-Green channel alone is larger than the luminance contribution for display bandwidths less than about 40 TV-lines. The sum of the contributions of the two chrominance channels is larger than the luminance contribution for display bandwidths below about 140 TV-lines. The sharp initial rise of the chrominance channel capacities is due to the concentration of the perceivable chrominance levels at low retinal frequencies, as discussed above.

As a practical application of the channel capacity formalism, we have computed the optimum bandwidth allocation among the three information channels - luminance and chrominance - for a display whose total passband is constrained to a particular value. In the example considered, the sum of the widths of the luminance and chrominance passbands was constrained to the value of 450 TV-lines, and the viewing distance was once again taken to be three picture widths. These values were chosen because they approximate the bandwidth and typical viewing conditions for U.S. commercial television. We sought the optimum tradeoff, from the information theory point of view, between the passbands for the luminance, Red/Blue-Green and Yellow/Blue channels, subject to the constraint that the total available bandwidth* is 450 TV-lines. Mathematically, the total channel capacity H, given in Eq. (31), was maximized for the assumed total bandwidth constraint.

The result of the calculation is given in the first two entries of Table 10. It is seen that by devoting about 140 TV-lines of the available passband to chrominance signals, the total channel capacity is significantly enhanced over the case where the total passband is devoted to luminance signals alone. This result is in very good agreement with the actual apportionment utilized in current U.S. color receivers. The computed 5:1 ratio of the passband widths assigned to the Red/Blue-Green and Yellow/Blue channels is to be compared with the actual 3:1 ratio for the bandwidths associated with the I_C and Q_C axes of

TABLE 10. BANDWIDTHS AND CHANNEL CAPACITIES OF DISPLAYS AT r/w = 3

	Bandwidth (TV-lines)			
Display	Luminance	Red/Blue-Green	Yellow/Blue	H (bits)
Optimum Color†	310	118	22	1850
450-TV-Line Monochrome	450	0	0	1520
Perfect Color	∞	∞	∞	2700
Perfect Monochrome	∞	0	0	1600

†Maximum H for total passband of 450 TV-lines.

*In the example, it was assumed that the passband is continuous and cannot be sampled at discrete frequencies. Thus, interspersing the luminance and chrominance signals is not possible, as it is in the case of color television.

the U.S. color (NTSC) system. The computed ratio can be reduced by taking into account the slight misalignment of the perceptual chrominance axes and the I_C and Q_C axes. Considering the crude nature of the calculations, the overall agreement with the empirically established color-television passband allocation is considered very satisfactory.

The last two entries in Table 10 give the computed results for the total channel capacity of a "perfect" color display (one with infinite passbands for all three channels) and a "perfect" monochrome display (one with an infinite luminance passband but no color capability). Comparing these results with the computed H-value for the optimized color display, it is seen that the channel capacity of the optimized, but band-limited, color display exceeds that of the perfect monochrome display, even though only 140 TV-lines are devoted to chrominance signals. This result is a testimony to the effectiveness of color as a low-frequency information channel and arises, as discussed earlier, from the excellent low-frequency color sensitivity of the visual system.

Table 10 indicates that significantly greater channel capacity can be achieved in wide-band color systems. We suspect that the strength of this conclusion is somewhat exaggerated by our extension of the chrominance-sensitivity functions to retinal frequencies above those actually measured [42], resulting in a probable overestimate of the values of H_{Ci} for large chrominance bandwidths. However, we consider the qualitative nature of our conclusion to be valid.

REFERENCES

1. R. W. Cohen, I. Gorog, and C. R. Carlson, "Image Descriptors for Displays," Technical Report to the Office of Naval Research, Contract No. N00014-74-C-0184, Jan. 1975. (Referred to in the text as TR1.)

2. R. W. Cohen, C. R. Carlson, and G. D. Cody, "Image Descriptors for Displays," Technical Report to the Office of Naval Research, Contract No. N00014-74-C-0184, Jan. 1976. (Referred to in the text as TR2.)

3. A. Rose, *Vision: Human and Electronic* (Plenum Press, New York, 1973).

4. O. H. Schade, Sr., *Image Quality, A Comparison of Photographic and Television Systems* (RCA Laboratories, Princeton, N.J., 1975).

5. L. M. Biberman, ed., *Perception of Displayed Information* (Plenum Press, New York, 1973).

6. E. M. Crane, "An Objective Method for Rating Picture Sharpness: SMT Acutance," *J. SMPTE* 73, 643 (1964).

7. E. H. Linfoot, *Fourier Methods in Optical Image Evaluation* (Focal Press Ltd., London, 1964).

8. R. G. Gendron, "An Improved Objective Method for Rating Picture Sharpness: CMT Acutance," *J. SMPTE* 82, 1009 (1973).

9. C. A. Padgham and J. C. Saunders, *The Perception of Light and Colour* (Academic Press, New York, 1975), pp. 74-77.

10. G. B. Wetherill and H. Levitt, "Sequential Estimation of Points on a Psychometric Function," *Brit. J. Math. Stat. Psychol.* 18, 1 (1965).

11. C. W. Eriksen and H. W. Hake, "Multidimensional Stimulus Differences and Accuracy of Discrimination," *J. Exp. Psychol.* 50, 153 (1955).

12. W. R. Garner, *Uncertainty and Structure as Psychological Concepts* (John Wiley & Sons, Inc., New York, 1962).

13. F. Smith, D. Lott, and B. Cronnell, "The Effect of Type Size and Case Alternation on Word Identification," *Am. J. Psychol.* 82, 248 (1969).

14. G. M. Reicher, "Perceptual Recognition as a Function of Meaningfulness of Stimulus Material," *Am. J. Exp. Psychol.* 18, 274 (1969).

15. E. E. Smith and K. T. Spoehr, "The Perception of Printed English," in B. Kantowitz, ed., *Human Information Processing: Tutorials in Performance and Cognition* (Halsted Press, New York, 1974).

16. J. J. Mezrich, "The Word Superiority Effect in Brief Visual Displays: Elimination by Vocalization," *Perception and Psychophysics* 13, 45 (1973).

17. J. C. Johnston and J. L. McClelland, "Visual Factors in Word Perception," *Perception and Psychophysics* 14, 365 (1973).

18. R. Tyte, J. Wharf, and B. Ellis, "Visual Response Times in High Ambient Illumination," Intern. Symp. Society for Information Display, Washington, D.C., April 22-24, 1975, *SID Digest* of Technical Papers, p. 98.

19. W. R. Garner, *The Processing of Information and Structure* (Halsted Press, New York, 1974).

20. W. B. Pillsbury, "A Study in Apperception," *Am. J. Psychol.* 8, 315 (1897).

21. J. Nachmias, "Effect of Exposure Duration on Visual Contrast Sensitivity with Square-Wave Gratings," *J. Opt. Soc. Am.* 57, 421 (1967).

22. J. J. Mezrich, "Spatial Frequency Processing for Threshold Visibility and for Picture Recognition," doctoral dissertation, University of Michigan, Ann Arbor, Mich., 1975.

23. B. G. Breitmeyer, "Simple Reaction Time as a Measure of the Temporal Properties of Transient and Sustained Channels," *Vision Res.* 15, 1411 (1975).

24. D. R. J. Laming, *Mathematical Psychology* (Academic Press, London, 1973).

25. R. G. Pachella, "The Interpretation of Reaction Time in Information Processing Research," in B. Kantowitz, ed., *Human Information Processing: Tutorials in Performance and Cognition* (Halsted Press, New York, 1974).

26. J. J. Mezrich, "Structure Visibility Change with Contrast Reversal," in preparation (1977).

27. J. S. Bendat and A. G. Piersol, *Measurement and Analysis of Random Data* (John Wiley & Sons, Inc., New York, 1966), Chapter 1.

28. K. McIlwain and C. E. Dean, eds., *Principles of Color Television* (John Wiley & Sons, Inc., New York, 1956).

29. D. G. Fink, ed., *Television Engineering Handbook* (McGraw-Hill Book Co., New York, 1957), Chapter 10.

30. M. W. Baldwin, Jr., "The Subjective Sharpness of Simulated Television Images," *Bell Syst. Tech. J.* 19, 563 (1940).

31. J. Nachmias and R. V. Sansbury, "Grating Contrast: Discrimination May Be Better Than Detection," *Vision Res.* 14, 1039 (1974).

32. M. A. Georgeson and G. D. Sullivan, "Contrast Constancy: Deblurring in Human Vision by Spatial Frequency Channels," *J. Phys. London* 252, 627 (1976).

33. R. F. Quick and T. A. Reichert, "Spatial-Frequency Selectivity in Contrast Detection," *Vision Res.* 15, 637 (1975).

34. M. B. Sacks, J. Nachmias, and J. G. Robson,, "Spatial-Frequency Channels in Human Vision," *J. Opt. Soc. Am.* 61, 1176 (1971).

35. See, for example, Mischa Schwartz, *Information Transmission, Modulation, and Noise*, 2nd ed. (McGraw-Hill Book Co., New York, 1969); J. C. Dainty and R. Shaw, *Image Science* (Academic Press, London, 1974).

36. H. Nyquist, "Certain Factors Affecting Telegraph Speed," *Bell Syst. Tech. J.* 3, 324 (1924).

37. R. V. L. Hartley, "Transmission of Information," *Bell Syst. Tech. J.* 7, 535 (1928).

38. R. W. Cohen and I. Gorog, "Visual Capacity - An Image Quality Descriptor for Display Evaluation," *Proc. SID* 15, 53 (1974).

39. R. L. DeValois and K. K. DeValois, "Neural Coding of Color," in Vol. 5, F. C. Carterette and M. P. Friedman, eds., *Handbook of Perception* (Academic Press, New York, 1975).

40. P. L. Walraven, *On the Mechanisms of Color Vision* (University of Utrecht, Netherlands, 1962).

41. D. L. MacAdam, "Visual Sensitivities to Color Differences in Daylight," *J. Opt. Soc. Am.* 32, 247 (1942).

42. G. J. C. van der Horst and M. A. Bouman, "Spatiotemporal Chromaticity Discrimination," *J. Opt. Soc. Am.* 59, 1482 (1969).

43. B. N. Timney et al., "Chromatic Grating Thresholds and the McCollough Effect," *Vision Res.* 14, 1033 (1974).

DISTRIBUTION LIST

Chief of Naval Research
800 N. Quincy St.
Arlington, VA 22217
Attn: Codes 212 — 6
455 — 1

Defense Documentation Center
Cameron Station
Alexandria, VA 23314 — 12

Director, Naval Research Laboratory
Washington, DC 20390
Attn: Code 2627 — 1

Office of Naval Research
Branch Office
New York Area Office
715 Broadway (5th Floor)
New York, NY 10003 — 1

Director
Office of Naval Research
Branch Office
1030 East Green Street
Pasadena, CA 91106 — 1

Office of Naval Research
Branch Office
495 Summer Street
Boston, MA 02210 — 1

Director
Office of Naval Research
Branch Office
536 Clark Street
Chicago, IL 60605 — 1

Office of the Chief of Naval Operations
Department of the Navy
Washington, DC 20350
Attn: OP-098T — 1

Headquarters
Department of the Navy
Naval Material Command
Washington, DC 20360
Attn: Systems Effectiveness Branch MAT 034 — 1

Commander
Naval Air Systems Command
Washington, DC 20360
Attn: AIR 5335 — 1
5315 — 1
340D — 1
340F — 1
360F — 1
03P — 1

Commander
Naval Sea Systems Command
Washington, DC 20360
Attn: NSEA 0341 — 1

Commander
Naval Electronic Systems Command
Washington, DC 20360
Attn: ELEX 320 — 1
330 — 1

Commanding Officer
U. S. Naval Air Development Center
Warminster, PA 13974
Attn: Codes 3042 — 1
207 — 1
5413 — 1

Commander
Naval Ocean Systems Center
San Diego, CA 92152 — 1

Commander
Naval Weapons Center
China Lake, CA 93555
Attn: Code 3175 — 1

Commander
Naval Surface Weapons Center
Dahlgren Laboratory
Dahlgren, VA 20910
Attn: Technical Library — 1

DISTRIBUTION LIST (Cont.)

Commander
Naval Avionics Facility
6000 E. 21st Street
Indianapolis, IN 46218
Attn: Technical Library 1

Naval Training Equipment Center
Orlando, FL 32813
Attn: Technical Library 1

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940 1

Commander
Naval Underwater Systems Center
Department SB 324
Newport, RI 1

Director, U.S. Army Research Institute
1300 Wilson Boulevard
Arlington, VA 22209 1

Commanding General
U.S. Army Electronics Command
Fort Monmouth, NJ 07703
Attn: AMSEL-VL-E 1
AMSEL-TL-BD 1
AMSEL-VL-I 1

Commandant, U.S. Marine Corps
Headquarters, U.S. Marine Corps
Washington, DC 20591
Attn: RD-1 1

Commandant
U.S. Coast Guard Headquarters
400 7th Street, NW
Washington, DC 20591
Attn: GDST/62 TRPT 1

Commanding General
U.S. Army Material Command
Washington, DC 20315
Attn: AMCRD-HA 1

Director
Human Engineering Labs
Aberdeen Proving Grounds, MD 21005
Attn: AMXRD-HEL 1

Air Force Avionics Laboratory
Air Force Systems Command
Wright-Patterson AFB, OH 45433
Attn: AFAL/RWI 1

Aerospace Medical Research Laboratory
Wright-Patterson AFB, OH 45433
Attn: AMRL/HEA

Aeronautical Systems Division
Air Force Systems Command
Wright-Patterson AFB, OH 45433
Attn: ASD/RW 1

Air Force Office of Scientific Research
Bolling Air Force Base
Washington, DC 20332 1

Headquarters, Rome Air Development Center
Air Force Systems Command
Griffiss Air Force Base, NY 13441
Attn: RBRAC 1

Federal Aviation Agency
NAFEC Bldg. 10
Atlantic City, NJ 1

Defense Advanced Research Project Agency
1400 Wilson Blvd.
Arlington, VA 22209 1

University of Illinois
2-113 Coordinated Sciences Laboratory
Urbana, IL 61801
Attn: Dr. R.L. Johnson 1

DISTRIBUTION LIST (Cont.)

Virginia Polytechnic Institute
Dept. of Industrial Engineering
Blacksburg, VA 24061
Attn: Dr. H.L. Snyder 1

Honeywell, Inc.
Systems and Research Division
2600 Ridgway Parkway
Minneapolis, MN 55413
Attn: Dr. A. Kanarick 1

Boeing Aerospace Company
Research and Engineering Division
P.O. Box 3999
Seattle, WA 98124
Attn: Crew Systems 1

Zenith Radio Corporation
1900 North Austin Avenue
Chicago, IL 60639
Attn: A. Sobel 1

General Electric
Research and Development
Box 43
Schenectedy, NY 12301
Attn: J.E. Bigelow 1

Magnavox Company
Advanced Technology Group
Fort Wayne, IN 46804
Attn: Dr. C. Craighead 1

Kaiser Aerospace and Electronics
Corporation
1651 Page Mill Road
P.O. Box 11275 Sta. A
Palo Alto, CA 94306
Attn: G. Carroll 1

IBM Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598
Attn: Ifay Chang 1

North Hills Electronics
Alexander Place
Glen Cove, New York 11542
Attn: S. Sherr 1

Xerox Corporation
Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304
Attn: B. Kazan 1

Tektronix, Inc.
P.O. Box 500
Beaverton, OR 97005
Attn: A. Silzars 1

Sperry Flight Systems
Phoenix, AZ
Attn: J.R. Trimmier 1