AFAMRL-TR-84-23

AD-A143 246

DEVELOPMENT OF COLOR CRITERIA FOR ADVANCED DISPLAYS

FRANK WARD
DENISE WILSON

SCIENCE APPLICATIONS INCORPORATED

DONALD L. WALLQUIST, 1Lt, USAF

AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY

MAY 1984

AIR FORCE AEROSPACE MEDICAL RESEARCH LABORATORY
AEROSPACE MEDICAL DIVISION
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from Air Force Aerospace Medical Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

TECHNICAL REVIEW AND APPROVAL

AFAMRL-TR-84-23

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

CHARLES BATES, JR.
Director, Human Engineering Division
Air Force Aerospace Medical Research Laboratory

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED

1b. RESTRICTIVE MARKINGS:

2a. SECURITY CLASSIFICATION AUTHORITY:

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE:

3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S): SD/L: 84-02-157

5. MONITORING ORGANIZATION REPORT NUMBER(S): AFAMRL-TR-84-023

6a. NAME OF PERFORMING ORGANIZATION: Science Applications, Inc.

6b. OFFICE SYMBOL *(If applicable)*:

7a. NAME OF MONITORING ORGANIZATION: Air Force Aerospace Medical Research Laboratory

6c. ADDRESS *(City, State and ZIP Code)*: 1010 Woodman Drive, Suite 200, Dayton OH 45432

7b. ADDRESS *(City, State and ZIP Code)*: Wright-Patterson Air Force Base OH 45433

8a. NAME OF FUNDING/SPONSORING ORGANIZATION:

8b. OFFICE SYMBOL *(If applicable)*: AFAMRL/HEA

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F33615-82-C-0500

8c. ADDRESS *(City, State and ZIP Code)*:

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
62202F	7184	11	44

11. TITLE *(Include Security Classification)*: Development of Color Criteria for Advanced Displays (U)

12. PERSONAL AUTHOR(S): Ward, Frank, Wilson, Denise and Wallquist, Donald L.*

13a. TYPE OF REPORT: Technical

13b. TIME COVERED: FROM Sep 82 TO Sep 83

14. DATE OF REPORT *(Yr., Mo., Day)*: 1984 May

15. PAGE COUNT: 97

16. SUPPLEMENTARY NOTATION: *Visual Display Systems Branch, Human Engineering Division

17. COSATI CODES

FIELD	GROUP	SUB. GR.
05	05	
05	08	

18. SUBJECT TERMS *(Continue on reverse if necessary and identify by block number)*: Color Discrimination; Color Vision; Image Enhancement; Image Processing; Image Quality; Synthetic Aperture Radar; Pseudocolor

19. ABSTRACT *(Continue on reverse if necessary and identify by block number)*

The research described herein involves color calibration of a CRT, a color discrimination study, and a color-coding system using Synthetic Aperture Radar (SAR) imagery. The results of the color calibration effort showed that, with appropriate regression coefficients color output, both on a CRT and on film could be predicted. The color discrimination study showed that color discrimination is poorest for "red" colors. These data agree with our previous studies on color matching and discrimination, and we recommend caution when using the UCS representation as the index of color difference. The radar imagery color-coding study found, as have other investigators, that black and white coding is superior to color coding. If imagery is to be color-coded, hue coding is better than hue-brightness coding.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐

21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL: Gilbert G. Kuperman

22b. TELEPHONE NUMBER *(Include Area Code)*: 513-255-3727

22c. OFFICE SYMBOL: AFAMRL/HEA

DD FORM 1473, 83 APR — EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

PREFACE

The work summarized in this report was performed by Science Applications, Inc. (SAI) under Contract No. F33615-82-C-0500 from the U.S. Air Force. Work Unit No. 71841144 was monitored by Mr. Gilbert Kuperman who was also the technical manager. The work supported projects funded under LDF 83-3.

Special thanks are due Fran Greene of AFAMRL, who supplied many of the references from European sources. Gilbert Kuperman provided many insightful suggestions for the false-color study as did Mike Syracuse of the SAI Falls Church office.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

GLOSSARY OF TERMS AND ABBREVIATIONS

Additive color primaries	Primaries (red, green, blue) whose radiant energies are summed by the observer's visual system to produce a sensation of color; color CRTs use additive primaries.
Bit-value	One of the 256 possible intensity levels a pixel may be assigned in a digital image processor incorporating 8-bit digital-to-analog convertors.
Brightness	The subjective correlate of luminance
Chromaticity chart	A diagram upon which colors are plotted
CIE	Commission Internationale de'Eclairage, the international body that sets lighting and color standards
CIE Standard Observer	Tabulated color functions used to define typical color perception
Color coding	Use of colors to represent different values or dimensions or in color coding of radar intensity returns. Also referred to as false color and pseudo color.
Color discrimination	The ability to perceive chromaticity differences based on hue and saturation alone.
Color primaries	Colors, usually red, green, blue, which when mixed together produce the desired color or color match.
CRT	Cathode ray tube
dB	Decibel, defined by the equation: $dB = 20 \log_{10}\left(\frac{\text{maximum scene luminance}}{\text{minimum scene luminance}}\right)$
Dynamic range	The range of radar intensity returns from the minimum background and thermal noise to the maximum return expressed in dB.
D-65 white	One of the standard daylight whites defined by the CIE and having a correlated color temperature of $6500^{o}K$.

GLOSSARY OF TERMS AND ABBREVIATIONS (continued)

Ft-L	Foot-Lambert, a unit of luminance
II	Image Interpreter
Luminance	Radiant energy weighted for the sensitivity of the human eye; defined by the 1924 CIE standard observer.
Misconvergence	Failure of the three electron beams of a color shadow mask CRT to strike the appropriate phosphors in a triad.
Multiple regression	A mathematical technique to use several variables to predict a value of a criterion variable.
Pixel	Picture element; one of the n x n elements that comprises a given CRT image. Many digital graphics systems use 512 x 512 pixel arrays.
Psychophysics	Techniques used to measure subjective entities such as color, brightness, loudness, etc.
SAR	Synthetic aperture radar
Shadow mask	A type of CRT tube employing a matrix of embedded phosphor dots in triads. Each triad contains red, green, and blue phosphors which are selectively energized by each of the electron beams.
Spectroradiometer	An instrument that measures the energy content of a part of the electromagnetic spectrum.
Subtractive color primaries	The process of absorbing a portion of the energy in a radiant spectrum, leaving the desired spectral distribution for rendering color; subtractive primaries yellow, magenta, cyan are employed in color film transparencies.
Tristimulus values	Color matching functions that describe the amount of each color primary required to match visually a particular spectral distribution.
VIPER	Visual Image Processing, Enhancement, and Reconstruction

1.0 TASK OBJECTIVES

The overall goal of the Air Force Aerospace Medical Research Laboratory (AFAMRL) color program is to specify human factors parameters involved in a variety of color display applications. In the past, interest was focused upon color discrimination and color CRT visibility under high ambient illumination (Ward, Greene, and Martin, 1983 a, b).

The present work focuses upon the Visual Image Processing, Enhancement, and Reconstruction (VIPER) facility. There, digital images can be output on a color monitor and/or can be reproduced in hardcopy form using a color graphics camera. For color work, both of these outputs must be calibrated so that chromaticity can be predicted from the software bit assignments. Our first task, therefore, was to develop a methodology for calibrating the color outputs of the VIPER facility. There is also a need to determine ranges of human color discrimination for selected hues of interest. Our second task was to run a study to measure color discrimination in a normal population. Once VIPER outputs were calibrated and some information about color discrimination was available, we were able to combine these data and run a study using pseudocolor coding on digital imagery. Our third task was to develop and test a strategy for color-coding synthetic aperture radar (SAR) imagery.

2.0 INTRODUCTION

The Air Force's interest in color displays has been growing over the years. Initially, full-color displays were limited to ground use because they were not rugged enough to withstand the flight environment. Now, because of recent technological developments, shadow-mask color cathode ray tubes (CRTs) can be manufactured to withstand the demands of military aircraft. (Previously, airborne color displays used rugged beam penetration tubes having a limited color gamut.) The new technology has opened the way for full-color displays in many applications. For example, the Air Force is considering the use of color CRTs for certain displays in the B-1 bomber. Some fighter aircraft may be modified to incorporate color CRT displays. Several of the newest commercial and general aviation aircraft are already using color displays.

Interest in color displays ranges from flight and weapons systems status to navigational and sensor imagery. Besides the obvious aesthetic appeal, there is some evidence to suggest that appropriate color-coding can facilitate information transfer and enhance performance. For example, Stollings (1982) has shown that under high workload conditions, color-coding enables pilots to perform better on a weapons status retention task. Dual task performance such as monitoring engine status while performing target detection also shows faster response times for color-coded displays (Wagner, 1977). The general utility of color-coding is still uncertain as the results of many studies are mixed (Christ, 1975; Teichner, Christ, and Corso, 1977; Tullis, 1981). It may turn out that performance with color-coding excells principally for demanding, high-workload situations.

The rapid development of color CRT hardware has resulted in a proliferation of color displays. In addition to Air Force applications, there are color displays for marine radar (Huddleston, 1983) and the FAA is planning to implement color displays in air traffic control. The potential for use of color display terminals in the business and scientific communities is virtually infinite.

Each new application for color displays has its associated problems. Human operators view the displays and must make decisions based upon what they perceive. The addition of color to a display format therefore brings with it all the problems of color defectives and issues involving the color of ambient lighting, simultaneous color contrast, chromostereopsis, and chromatic aberration.

CRT color is not constant; phosphors age, beams and the shadow mask misconverge, and beam currents may drift. And, of course, when color coding is not consistent across display applications, some confusion may be expected if operators must work with many displays. For color displays to be most effective, consideration must be given to the hardware, to human color perception, and to color-coding. This report will address each of these three areas.

2.1 SOURCES OF COLOR SCIENCE

There are numerous sources that address visual science, color vision, color displays and associated technology. It is beyond the scope of this report to review these in depth; however some citations may be useful. For the individual new to the field, DeMars' (1975) report on color in display systems is a good introduction to color and vision. Human factors considerations are elaborated in many design guides such as Van Cott and Kinkade (1972) and Farrell and Booth (1975). Basic visual psychophysics is treated in great detail in Graham (1965).

Color perception and the psychophysics leading to international color standards are described by Wright (1969); Boynton (1979) also described some of the historic development of color science. Color calculations and practical examples relating to color measurement may be found in Judd and Wyszecki (1975). Extensive data relating to color and optical considerations are given by Wyszecki and Stiles (1982). Color coding of display information has been reviewed by Christ (1975; 1976) and a color display design guide by Krebs, Wolf, and Sandvig (1978) addresses many color-coding situations.

3.0 HARDWARE

3.1 BACKGROUND: THE COLOR CRT

Development of the color shadow-mask CRT actually began in 1949 when RCA decided to develop a full-color tube. Technology has come a long way from the earliest tubes that were first demonstrated publically in 1950. Herold (1974) gives an extensive discussion of these historical details.

A typical shadow-mask color CRT has three electron guns each driven by its own electronics. The beam from each gun is focused such that it strikes one of the three primary phosphor dots. The phosphor dots are grouped in threes to form triads. Each phosphor dot emits light energy having its own characteristic spectral emission. Thus, the output from the CRT is typically a mixture of three relatively narrowband photic emissions. It is the human eye that "integrates" these three "primaries" into the array of colors normally perceived. Consequently, the perceiver cannot be divorced from the display, for without the perceiver there are no colors, only distributions of energy. For this reason the science of color measurement has had to develop procedures to take the observer "into account." Bartleson (1968) presents a very lucid summary of color perception as it relates to CRT viewing.

Assuming a consistent observer, there are many ways in which a CRT display can affect perceived color. Differential aging of the three phosphors can alter the color balance. For example, the P22 "green" phosphor ages more quickly than the "blue" and "red" P22 phosphors (Martin, 1977).

The choice of appropriate phosphors is a very complicated process involving color, persistence, brightness, aging, and resolution. Woodcock and Leyland (1979) have summarized many of these issues in modern phosphor technology. The light output of phosphors is not linear with the input signal. In black and white monitors, the gamma (luminance versus intensity function) can be optimized by an iterative process to produce a set of equally discriminable contrast steps (Catmull, 1979; Briggs, 1981). For color monitors, the problem is much more complicated because there is no single gamma function that would be suitable for all colors. Each beam affects the color balance and, for particular ratios of currents, even small variations can dramatically alter the perceived color. At high beam currents the shadow mask may heat up and expand causing a misregistration between the phosphor dots and their beams. This is one way misconvergence may occur, altering color purity. Color purity may also be influenced by the environment. Shock, vibration, and ambient magnetic fields decrease convergence and hence, purity. In the past, these problems were lessened by using beam penetration tubes rather than shadow mask CRTs; however the color gamut was restricted to red, orange, yellow and greens. It has not been practical for beam penetration tubes to produce blues. The two CRT technologies have been compared by Brun and Martin (1980).

Calibration of a CRT monitor is not a simple process because so many parameters are subject to changes outside of the operator's control. We will describe our calibration methodology in detail in a later section of this report. At this point it will suffice to point out that:

> "If CRTs are used to display images for evaluation purposes, the image chain analysis must consider their characteristics, and strict quality control measures must be applied during their use. Even when operating as designed, almost all CRTs introduce a variety of contrast and geometric errors into the displayed image. A further source of error is found in conventional color CRTs where registration problems are common. Regardless of the type of CRT installation in use, luminance, contrast response, color, and geometric characteristics must be tested at least daily during evaluation sessions" (Booth and Schroeder, 1977, p. 17).

3.2 THE VIPER FACILITY

The AFAMRL VIPER facility consists of an International Imaging Systems (I2S) Model 70E digital imagery hardware with a Digital Equipment Corporation PDP-11/34 host computer. The system can digitize imagery from hardcopy and video cassettes, create graphics imagery, or can read in imagery from digital tape. Once digitized, images can be transformed and enhanced in a variety of ways including the application of false or pseudocolor.

A picture (in VIPER) is defined as one refresh memory (image channel) which ultimately drives the R, G, B electron guns of a color monitor or the single gun of the color graphics camera. The image channel is composed of 8 planes of 512 x 512 pixel arrays, together forming pixels which equate to 8 bits of possible intensity (values 0-255). The information in each pixel is fed through three look-up tables (R, G, B), three output function memories (R, G, B), three D/A converters (R, G, B), three video generators and then drives the three R, G, B guns of the monitor or the single gun of the color camera.

The actual colors and luminances to be produced are determined by assigning the appropriate intensity values to each of the R, G, B look-up tables. These intensities or bit-values are assigned under software control. Thus, the operator can specify intensity over the range 0.0 to 1.0 (minimum to maximum) and obtain bit-values from 0 to 255 by a linear transformation for each of the R, G, B tables.

For these studies, an Aydin Model 8025 R, G, B high resolution monitor provided softcopy display. Hardcopy output was via film transparencies produced by a Matrix Instruments Co. Model 4007 color graphics camera. The camera system is built around a high-resolution black and white CRT that exposes photographic media through each of three color separation filters. Color film was exposed once through each of the three filters. CRT controls for contrast, brightness and exposure duration for each color separation filter were set to achieve an acceptable color balance. That is, we desired film exposures that would yield a white (2500°K-6500°K) for tungsten projection when the bit settings were approximately equal (Greenberg, Marcus, Schmidt, and Gortney, 1982). Once settings that produced an acceptable white were obtained, no further adjustments were made on the hardware.

In theory, both outputs can be nearly equivalent. Clapper, Gendron, and Browstein (1973) have argued that the color gamuts of additive color systems (e.g., CRTs) and subtractive systems (e.g., films) are alike if mechanisms associated with the photographic system are taken into account. That is, modern film dyes have crosstalk inhibition across the layers to reduce unwanted absorptions. Also, film density can never be zero; there is always some absorption. At equal luminances, the CRT and film systems are similar in color gamut. In practice, this was not the case with the VIPER system because the same video signal fed both the CRT monitor and the graphics camera's CRT. It would have been an impossible task to have found a single set of color bit-values such that the same colors would have been produced simultaneously on both outputs. This is due to the fact that the camera-film-projection system has a gamma (luminance/transmissivity versus intensity) function different from the CRT monitor.

4.0 COLOR THEORY AND CALIBRATION

In this section we will review some of the aspects of color theory relevant to our task objectives and will indicate why there are still problems with color calculations based on theory. Our treatment is brief and non-mathematical; detailed explanations are available in the references cited.

4.1 BACKGROUND

The development of a framework for the specification of colors has had a long history extending back to Maxwell's color mixing curves of the 1860's (Maxwell, 1860). Yet, even today, color specification is not complete and, in fact, current specifications contain substantial inaccuracies that lead to shortcomings for our applications. Problems with color specification have a serious impact on the visual display domain when color CRTs are incorporated into a system design. Currently, one cannot specify the design criteria for chromatic displays and be confident that the gamut of hues and saturations chosen will be visible and discriminable if particular display luminances and ambient lighting conditions are employed. Not only is this state of affairs frustrating for design engineers, it represents a crucial gap in our ability to assure that military applications will actually benefit from the latest technology; it remains for the human factors specialist to delineate the appropriate display requirements to render the use of color beneficial.

Research on chromatic displays has been reviewed in some detail over the past decade. In general, the research has been directed at color coding, visual search, and symbology size and density (Cook, 1974; Christ, 1975; Barker and Krebs, 1977; Shurtleff, 1980). Most of the data describe color coding effects on human performance and little human factors research has been directed at problems of simultaneous color contrast and hue discrimination. Human factors design guides for graphic displays address the color contrast and ambient illumination issues only in a qualitative manner (Krebs, et. al., 1978; Smith, 1978). Even the Boeing group's extensive treatment on color does not include any quantitative detail on chromatic discrimination under conditions relevant for the cockpit environment (Farrell and Booth, 1975). In fact, Boeing had to conduct their own color display studies to obtain design criteria for the color displays in the 757-767 commercial airplanes (Silverstein and Merrifield, 1981).

One may ask, rightfully, why theory and data are so deficient in this area. Part of the answer lies in the technological advances alluded to earlier. Due to the rapid development of higher luminance shadow mask CRTs, a data base different from that typically employed to support low luminance work is required. The original, low luminance work was done with broadband light sources under precise experimental control. Modern CRT displays employ relatively narrowband light sources. These sources may produce a quality of color perception different from that suggested by the standard data obtained with broadband sources. In addition, the precisely controlled experimental data may over-estimate performance under field conditions. The other part of the problem lies within the intrinsic nature of color itself. The manner in which color specification was originally formulated was dictated by the fact that color is not a physical phenomenon but is a subjective, psychological one. Colors are the result of human detection and perception of electromagnetic energy over a wavelength range of approximately 400-700 nm. Since color resides in the brain, no physical instrument can measure color directly; rather, spectral radiometric measurements are made and then multiplied by functions believed to represent <u>average</u> human sensitivity to light or color.

Current color specifications produce some disturbing anomalies (Kinney, 1982). For example, one might think that a blue and a white light of equal photopic luminance should appear equally bright; they do not. The blue will be brighter. If two colors match under moderate illumination, the match may fail under high illumination, especially if the degree of metamerism is high (Wright, 1936). Metamerism refers to the subjective match of colors which in fact have different spectral distributions. (A precise discussion of metamerism is given by Robertson (1983).) These anomalies can be very frustrating but are not surprising given the nature of the development of chromatic specification.

4.2 THE CIE

The current method for specifying photopic luminance was established in 1924 by the Commission Internationale de l'Eclairage (CIE). The CIE approved a visibility curve, $V\lambda$, which is supposed to represent the average sensitivity of a fully light-adapted human fovea to the visible spectrum of radiant energy. $V\lambda$ was derived by combining data from two principal techniques of heterochromatic photometry: the cascade, or step-by-step method, and flicker photometry. Unlike flicker photometry, the cascade method compares only similar spectral regions and avoids heterochromatic matches across large spectral separations (e.g., blue vs. red). Under the conditions employed in determining $V\lambda$, the laws of color additivity appear to hold. When additivity is valid, the weighted energy at each wavelength can be summed to arrive at a photometric luminance value. Flicker, small step-by-step and minimum border techniques usually obey additivity laws (Wagner and Boynton, 1972). Heterochromatic brightness matches across large spectral separations often yield additivity failures; this is why, for example, blue may appear brighter than its luminance would suggest. Judd (1951) proposed a correction to $V\lambda$ to minimize this problem. Clearly, the shape of the $V\lambda$ curve depends upon the method used in collecting the data. In 1924 when spectral luminosity was specified, it was adequate for a science of photometry based on tungsten sources where differences between measurement and visual perceptions are minor (Kinney, 1982). Now that there are numerous narrowband CRT sources, many researchers have suggested that $V\lambda$ be revised, especially in the short wavelength region (Cowan, 1982). No doubt there will be considerable reluctance to do this because color specifications are also tied to $V\lambda$.

The luminosity function $V\lambda$ is valid only under the conditions for which the original data were collected. Most researchers used a 2- or 3-degree field and averages were used to compute $V\lambda$. Actually, there is a substantial amount of individual variability in the data (see, for example, Committee on Colorimetry, 1963, p. 226 and Wright's telling comment in Boynton, 1979, p. 403, and for a recent discussion, Vienot, 1980).

In 1931, the CIE sanctioned the "standard observer" for color. The color-matching functions of Guild (1931) and Wright (1928-29) were combined to produce a set of average color-matching functions based on 17 subjects. Wright has summarized the details of the process in a transcript reprinted in Boynton (1979). The functions resulted from the experiments on color matching with either mixtures of spectral primaries of red, green and blue (Wright, 1928-29) or R, G, and B filters (Guild, 1931). When produced with real primaries, color-matching functions always contain some negative numbers. The CIE therefore decided to transform the primary system from the real R, G, B values to a new <u>imaginary</u> set, X, Y, Z, that had certain desirable features. First, a positive coordinate system was produced so that all color-matching functions based on real primaries will contain only positive values. Second, X and Z were located on the alchyne, i.e., the locus of colors in chromaticity space that has zero luminance. The color-matching function for Y was made identical to $V\lambda$ after some adjustment of the R, G, B stimulus luminances. The assumption that color and luminance are independent of each other was a clever notion. Strictly speaking, they are not. Nevertheless, $V\lambda$ was incorporated into the definition of the 1931 CIE standard observer. When the X, Y, Z color-matching functions are each multiplied by the spectral radiance of a color to be specified and weighted appropriately, the 1931 CIE X, Y, Z tristimulus values are produced. These are normalized to 1.0 and the familiar 1931 chromaticity chart is obtained by plotting X and Y.

The 1931 CIE system is a mathematical abstraction of the original data and the fact that the X, Y, Z primaries are imaginary (not physically realizable) must be accepted as a consequence of the constraints imposed in a linear projection from one coordinate system (R, G, B) to another (X, Y, Z). It must be remembered that the 1931 CIE chromaticity diagram is simply a method to represent color in a standardized manner. The diagram was derived from averaged data and should not be used to infer differences in color appearance and perception without some care. For example, the 1931 chromaticity diagram was based on a 2-degree field size. For fields above 4-degrees, the CIE has approved the 1964 CIE large field standard observer. Stiles (19[illegible]) has described the history of color matching and compared the large- and small-field data. To date, there is no standard observer for fields smaller than 1-degree.

The 1931 and 1964 chromaticity diagrams are similar in shape and function. They both allow specification of a color's chromaticity independent of its luminance. Neither preserves the perceptual spacing among colors. This means that the chromaticity diagrams cannot be used to infer perceived color difference.

4.3 COLOR DIFFERENCES

Color discrimination has been extensively investigated by MacAdam (1942). Although his original work involved only low luminances (13.9 ft-L) and one subject, later studies with many subjects confirmed the basic data set (Brown and MacAdam, 1949; Brown, 1951, 1957). MacAdam (1942) did color matches to determine color discrimination and calculated the standard deviations for the matches. He found that, when represented on the 1931 CIE chromaticity diagram, the standard deviations plotted as ellipses. This meant that discrimination (and hence, color perception) was not equal throughout the 1931 CIE diagram.

MacAdam and his coworkers have thoroughly researched methods to represent color differences. They have shown that any two-dimensional representation will involve some compromises as the basic space of chromaticity differences is three-dimensional (Silberstein, 1943). Nevertheless, MacAdam has developed a way to represent color differences in a space such that distances among colors are as perceptually equal as possible. These data have become the basis of the 1960 CIE Uniform Chromaticity Scale (UCS). The UCS diagram is a linear projection of the 1931 CIE diagram such that any color difference is taken to be equally perceptible from any other based on a 2-degree field size. If the diagram were perfect and there were no experimental error, MacAdam's ellipses would plot as circles (see MacAdam, 1942); in actuality, they only approximate this ideal situation.

Like the 1931 CIE representation, the 1960 UCS diagram does not address the luminance issue. The absolute appearance of hues does depend upon their luminances (Purdy, 1931; Savoie, 1973) but color discrimination does not change much over moderate changes in luminance (adaptation levels) of 1-100 ft-L (Brown and MacAdam, 1949; Brown, 1951; Siegel, 1969). Field size and chromatic adaptation are more important determinants of color discrimination (MacAdam, 1959; Bartleson, 1979).

4.4 COLOR DISCRIMINATION FORMULAE

An obvious extension of the 1960 UCS data is to include the luminance factor. This has been done in a variety of proposed color-difference formulae (Wyszecki and Stiles, 1982). Recently the CIE sanctioned two revised color spaces (CIE Colorimetry Committee, 1974; Robertson, 1977). The CIE L*a*b* and L*u*v* spaces can be used to calculate an index of color difference. Whether or not a given difference will be discriminable under some particular set of circumstances is another matter.

Galves and Brun (1975) have proposed a formula which produces a "Detection Index" that includes factors for chrominance, luminance, and contrast thresholds. They are sketchy on the sources of some of their data but the concept is an interesting one because it attempts a synthesis of many of the factors affecting visual detection.

A different approach has been taken by John Laycock and his coworkers in England (Laycock and Viveash, 1981; Laycock, 1982). They have developed computer routines to delineate lines of constant hue and saturation at various luminances. Their representation is a convenient extension of the CIE data and can even take chromatic adaptation into account. The approach derives from theory and its merits will have to be demonstrated in an applied environment.

Recent work on chromatic discrimination seems to emphasize an integration of data and theory. Boynton and his associates have developed equations to describe cone sensitivity and have measured chromatic differences to test an opponent-color model (Boynton and Kambe, 1980; Boynton and Wisowaty, 1980). Post, Costanza, and Lippert (1982) have compared chromatic and achromatic contrast in terms of the CIE color-difference formulae. Carter (1982) has developed a computer routine to ascertain the optimal set of colors to be displayed depending on number of targets, luminance, and CRT phosphor primaries. All of these efforts are directed to take color science closer to sound quantitative predictions with relevance for the applied environment. However, until sufficient research exists, some care must be taken in using the results of calculations too literally. We will argue later that the models of color vision developed to date are not accurate enough to permit good prediction of color discrimination for applied uses.

4.5 CALIBRATION AND COLOR THEORY

4.5.1 Color CRTs

According to the rules of color mixing, the chromaticity of a color on a CRT can be specified by knowing the relationship between the intensity bits set in the software of the imaging system and the R, G, B gun outputs. That is, each CRT gun and its associated phosphor can be measured separately for luminance and chromaticity and the coordinates of the display colors can be derived by applying Grassman's color mixing equations (Wyszecki and Stiles, 1982). A necessary and usually valid assumption is that the CRT guns operate independently of each other. The concept is that the tristimulus values for each source (gun) can be summed to compute the tristimulus values for the resultant color. The mathematical summation is meant to describe the additive color mixture processed by the eye of the CIE standard observer.

To obtain individual R, G, B tristimulus values, it is necessary to be able to predict the chromaticities and luminances of each of the CRT primaries. In practice, the chromaticity is not constant as the intensity of a given CRT gun is varied, nor is the luminance output linear with the software bit assignment. Some typical bit assignment versus luminance functions are shown in Figure 1. Clearly, a large number of measurements must be made to specify color and luminance accurately. Since CRT monitors drift, some procedure for measurement correction is also necessary if very high accuracies are required.

One approach that has been used to calibrate a color CRT monitor employs an automatic data collection and analysis procedure (Farley and Gutmann, 1980). A spectroradiometer, under computer control, measures the CRT output. Input to the CRT is also computer generated. With appropriate software, a large number of data points can be collected and the CRT primaries characterized throughout their intensity range.

According to Farley and Gutmann (1980), quadratic models can be derived to fit the luminance/bit-value relationship. In practice though, they found that the luminance predictions were not sufficient for their display system applications. Rather, they chose to linearize the luminance output of each CRT primary. They measured the luminance of each R, G, B gun at each of 16 equally spaced intervals over the range of the intensity bit settings. Values between the 16 measured luminances were obtained by interpolation and linear bit-value/luminance tables were constructed to map the linear-nonlinear relationship. This extra step simplified calculations because the CIE tristimulus values were linearly related to the new bit-value/luminance tables. Once the tristimulus values were determined for each bit-value, the inverse function could be found which related chromaticity and luminance to bit-value.

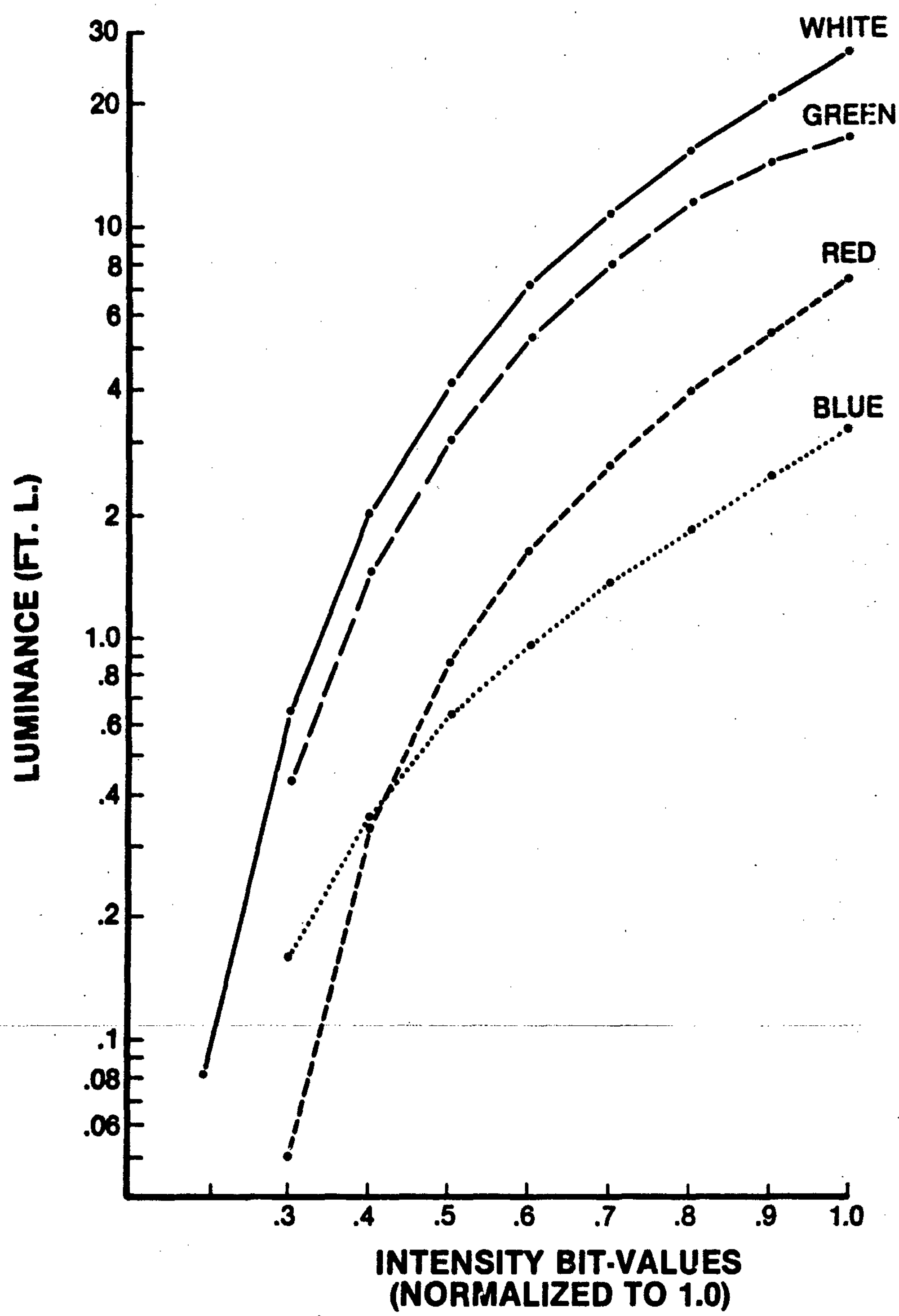


Figure 1. Color CRT luminance versus intensity bit-values for the R, G, B primaries plus white. (These data are from the VIPER Aydin color monitor.)

Farley and Gutmann (1980) also designed an automated recalibration system so that the spectroradiometer could be calibrated periodically as the color monitor was being characterized. Clearly, such an exact characterization of a color CRT is only possible with the aid of computer control, as their entire calibration procedure required many hours (Post, 1983).

4.5.2 Color Films

Color CRT monitors are but one possible output device for a digital image processing system. With a suitable color graphics camera, color transparencies can be produced. Naturally, there is a desire to relate the chromaticities produced by illumination of the film to the digitally assigned bit values. In this case the characterization process is more complicated than for a CRT because the system is open loop; the film must be processed and then the chromaticities determined as referenced to a particular source of white light.

The color film output is also more difficult to characterize because film transparency is a subtractive color system. Rather than adding three primaries to produce a given color as a CRT does, subtractive systems selectively remove the unwanted portions from a white illuminant to leave the desired colors. Each of the dye layers of color film acts as a notch filter when exposed and developed. The resulting colors depend upon all three dyes because there is considerable "crosstalk" in addition to the nonlinearities in the density that results from exposure.

Wallis (1975) has described a film calibration procedure in his dissertation. In general, the approach is similar to that described earlier for CRT characterization except that no attempt was made to linearize bit values and luminance (density, for film). This made Wallis' (1975) calculations very complicated.

If a CRT image is photographed, the resulting colors will not be a faithful reproduction of the original colors. The color distortions could be eliminated if suitable predistortions were introduced such that the mapping were the exact inverse of the expected distortions. As Wallis (1975) points out, conceptually this is a simple process but computationally it is definitely nontrivial. In his research he describes two general approaches. The direct method requires a mathematical model that describes the interactions of the display and film. The model is then inverted to produce the predistortion needed to give faithful film recording.

Wallis' (1975) second approach is indirect in that formal mathematical inversion is not required. He argues that "even with these approximate numerical procedures, the dominant sources of subjective errors in the colorimetric corrections are due to inadequacies in the visual model upon which the calculations are based, and not on the numerical errors themselves" (1975, p. 154). Wallis (1975) actually settled on the indirect procedure for mapping color to film transparency. He determined the ensemble of input and output vectors empirically and implemented his predistortion calculations. Wallis (1975) found that chromatic adaptation was very important. Simple normalization of the X, Y, Z tristimulus values, as is normally done for reflecting objects, was not adequate to correct for adaptation to tungsten illumination. Adaptation to other chromaticities can be even more dramatic; with blue-green adaptation, whites can look yellow-red (Kinney and Cooper, 1967).

Wallis (1975) also found that attempts to add neutral density to the film did not appear neutral when projected. Apparently the eye will not completely adapt to the "white" of the projector. Film is nonlinear and the color balance of the image scene affects the film. The particular white chosen is critical; exposure variations of less than 1/2 f-stop resulted in noticeable color shifts. Surprisingly, Wallis (1975) found that the film processing itself was quite consistent.

Both the CRT characterization and the color film calibration just described applied the formal rules of color science to produce "exact" solutions in terms of color calibration. In both cases the investigators had applications that required precise color control -- control that was only achieved with extensive computer processing and constant recalibration. Does this mean that no control over color is possible without these elaborate procedures? Not at all. We suggest that the following methodology will suffice if a facility's requirements are not too stringent.

4.6 AN ALTERNATIVE APPROACH

Theoretical approaches to color characterization via models of the imaging systems offer mathematical elegance that many investigators find desirable. From the practical side of color science, however, it must be realized that there are insufficient data to permit the development of a perfect model. And some of the more successful models are nonlinear (MacAdam, 1963; Bartleson, 1979; Burns, Smith, Pokorny, and Elsner, 1982). There is no value in obtaining precision that exceeds needs. We have therefore chosen to take a very basic approach to the task of color characterization of the CRT and film outputs of the AFAMRL VIPER facility. This approach is in keeping with the constraints imposed by not having an automated data collection facility. Since we lacked the means to collect large amounts of data easily, we could not determine colors by following the theory of color science and summing the tristimulus values for each primary. We needed an alternative approach to predict what software bit assignments would produce the CRT colors of most interest to the VIPER facilty. In addition, we wanted to develop color calibration procedures that were simple and did not require continued use of spectroradiometric equipment.

We chose to use multiple regression analysis to predict chromaticity from the software bit-value assignments. The procedure had the advantage of treating the entire system as a black box and was simple to implement. The procedure is adequate for basic false-color assignments and is easy to apply with a minimum of measurements. It also allowed us to avoid confusing the system's many operators as we did not have to alter the linear transform between software intensity assignments and the actual bit-values that determined each beam's intensity.

5.0 VIPER CRT COLOR CALIBRATION

5.1 INITIAL SETUP

The Aydin color monitor was adjusted to factory specifications by SAI technicians and the access to the controls was locked, leaving only the overall brightness control available to the operators.

Preparatory to each measurement session, the room lights were extinguished and a black cloth hood was placed over the color monitor and the radiometer head (Photo Research Model 710) to insure that only CRT-produced light was measured. Test colors were always displayed in the center of the screen in a square patch covering an area of 1.25 in. x 1.25 in. The radiometer head was positioned orthogonal to the CRT tube face at a distance of 3.5 ft. We made all of our spectroradiometric readings by using the automatic exposure mode wherein the instrument integrated over a time sufficient for accurate readings. The radiometer was also set to automatically obtain and average 10 individual readings.

At the beginning of each measurement session, we displayed an 8 x 8 matrix grey scale (64 levels) over the full CRT face and adjusted the brightness control by eye such that the lowest two intensities were not discriminable from the background of the CRT screen (about 0.8 ft-L). This adjustment was done with the room lights off and the hood over the monitor and radiometer. The luminance of the brightest grey shade (bit value 255) was 30 ft-L. This same grey shade adjustment procedure was used before each session; the minimum luminance was 0.8 ft-L and the maximum luminance always varied between 27-30 ft-L. The grey shade scale produced a useful dynamic range of about 30 dB.

The CRT R, G, B primaries were measured for chromaticity and luminance as a function of the software bit-value assignments. The relationship between the software assignment and the actual bit-value was always linear so that a software value of 1.0 represented the maximum bit-value (255) and 0.0 gave the lowest value (0). Figure 1 (shown earlier) reveals that the luminance/bit-value relationships are not linear. Table 1 gives the luminances and chromaticities (1960 UCS) as a function of the

Table 1. Luminance and Chromaticity as a Function of Software Bit-Value Assignment

Color	Bit-Value (Normalized to 1.0)	Luminance (ft.L)	Chromaticity Coordinates (1960 UCS)	
			u	v
Red	1.0	7.43	.4045	.3566
	.9	5.43	.4113	.3564
	.8	4.00	.4091	.3558
	.7	2.62	.4125	.3562
	.6	1.65	.4120	.3553
	.5	.86	.4210	.3560
	.4	.33	.4445	.3553
	.3	.06	.4954	.3504
Green	1.0	16.7	.1151	.3637
	.9	14.4	.1085	.3664
	.8	11.5	.1047	.3683
	.7	8.11	.1039	.3691
	.6	5.35	.1029	.3701
	.5	3.02	.0993	.3721
	.4	1.46	.0923	.3747
	.3	.43	.0806	.3785
Blue	1.0	3.22	.1787	.1003
	.9	2.46	.1814	.0949
	.8	1.85	.1814	.0925
	.7	1.37	.1817	.0934
	.6	.95	.1821	.0915
	.5	.63	.1823	.0915
	.4	.35	.1843	.0875
	.3	.16	.1872	.0825

software intensity assignment. Note that chromaticity varies as a function of beam intensity. The color gamut is larger for lower beam intensities. This effect is shown in Figures 5 and 6. The spectral distributions of each primary are shown in Figures 2, 3, and 4. As is usual in many color CRTs using P22 phosphors, blue and green have relatively broad peaks whereas red has several spikes.

Spectroradiometric measurements of all three primaries, combined at their maximum intensities, showed that the white produced was a rather blueish 15,000 K. We adjusted the software assignments to 1.0, 0.92, 0.73 for the R, G, B guns, respectively, to produce a D65 white (6500 K), one of the "standard" daylight whites. Figure 5 shows D65 plotted along with the CRT primaries in the CIE 1931 chromaticity space. The same data are plotted in Figure 6 in the CIE 1960 UCS diagram. (The remainder of the chromaticity plots in this report will use the 1960 UCS diagram. The 1976 UCS diagram is intended primarily for use with the L*a*b and L*u*v color difference formulae.) The triangle formed by the R, G, B primaries defines the CRT color gamut for relatively high beam intensities. The chromaticity shifts for the lower beam intensities and are plotted as "X's" in the two figures.

5.2 DISPLAY VARIABILITY

During each measurement session, the three primaries, the bit-values which had earlier approximated D65 white, and the white produced by combining the three primaries at their maximum intensities were measured for chromaticity, luminance, and color temperature (where applicable) as a function of the software bit-value assignments. Each of these measurements (an average of 10 separate determinations) was made at least once in each of 5 sessions extending over a one-month period. In most cases, each measurement was made 3 or 4 times during a given session. The means and standard deviations were computed and are given in Table 2.

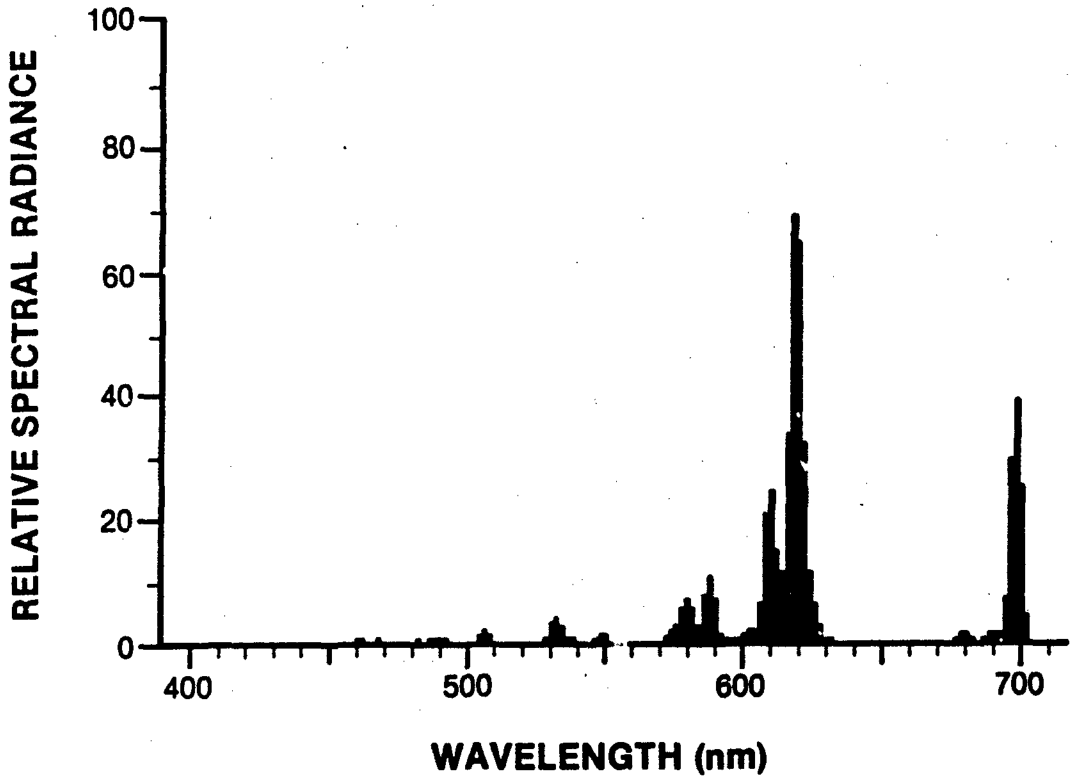


Figure 2. Spectral Distribution of the Aydin Red Primary

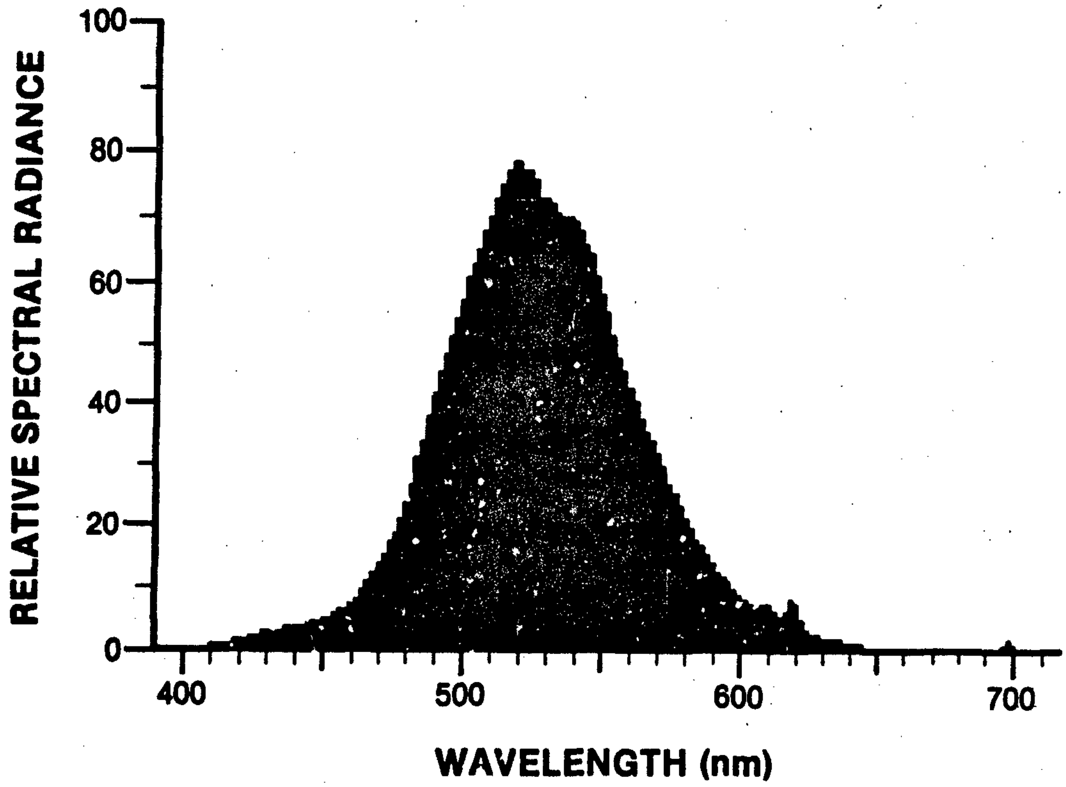


Figure 3. Spectral Distribution of the Aydin Green Primary

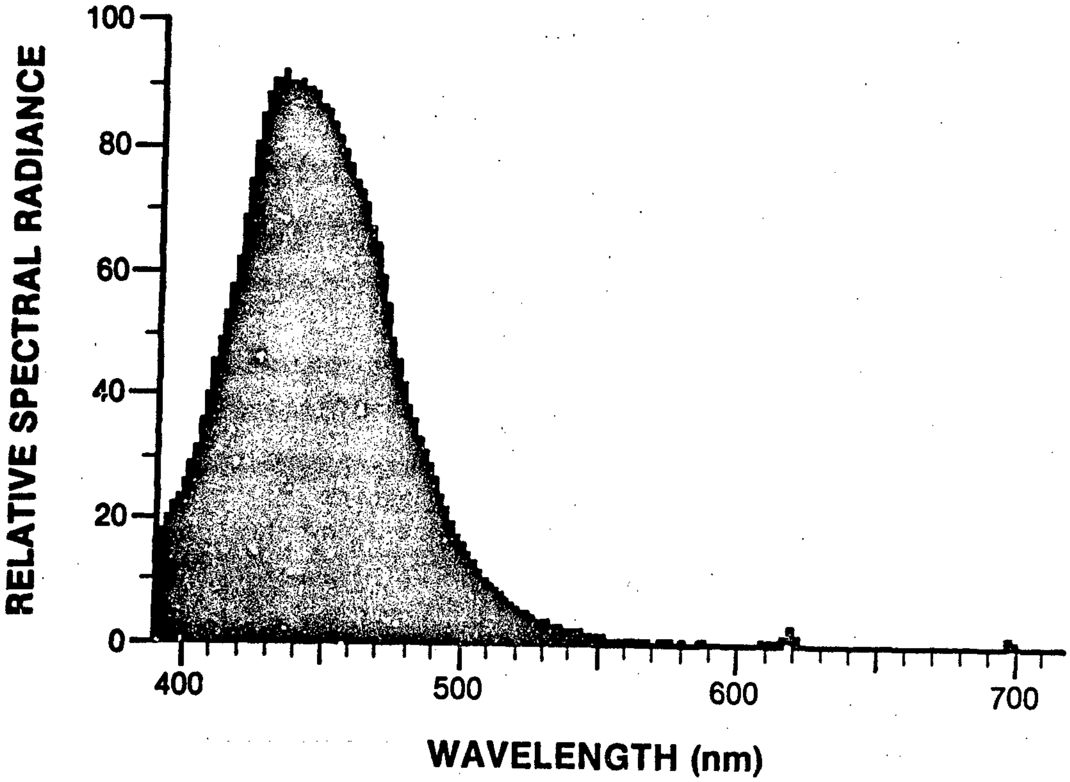


Figure 4. Spectral Distribution of the Aydin Blue Primary

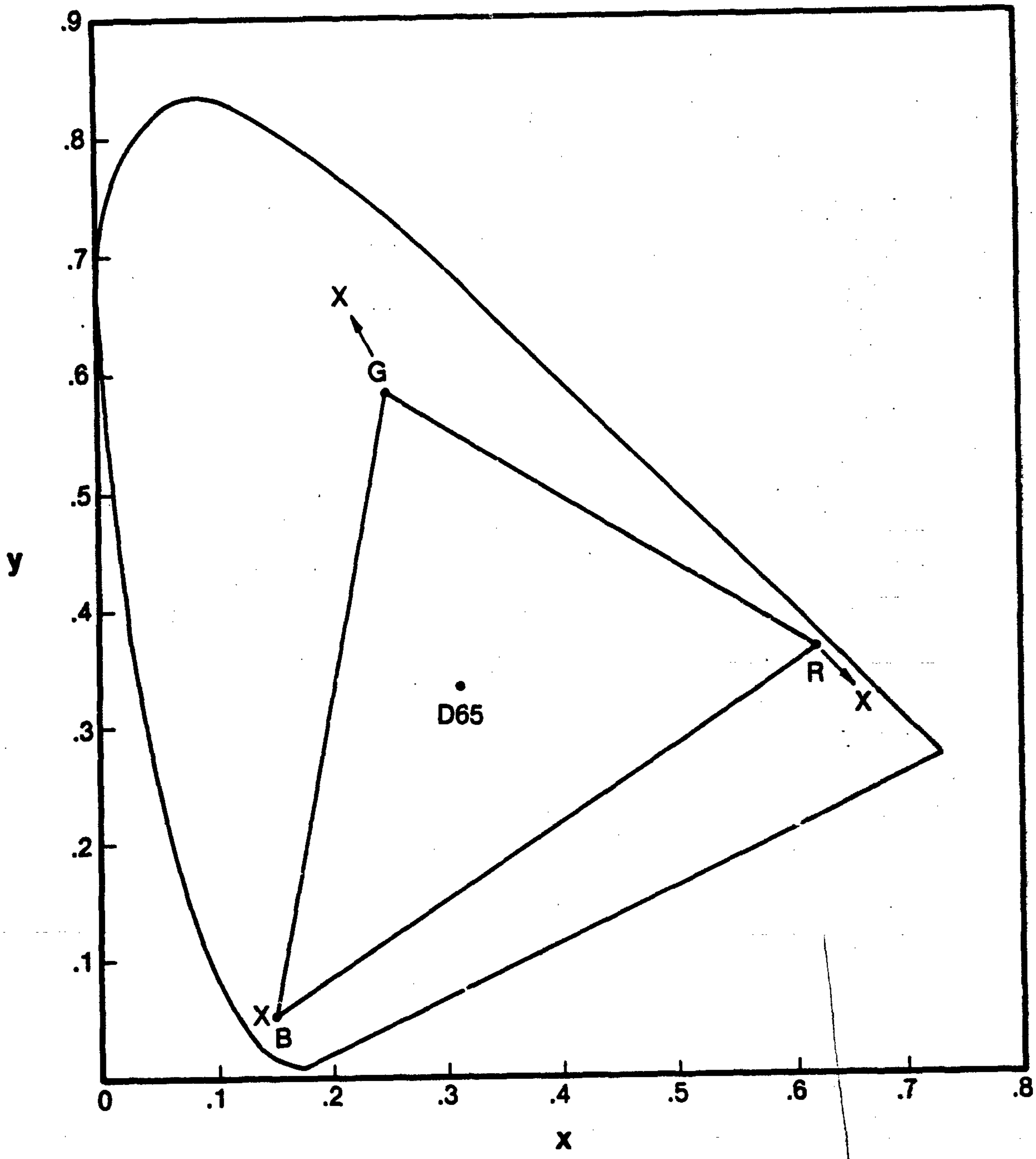


Figure 5. Aydin color monitor primaries in 1931 CIE space

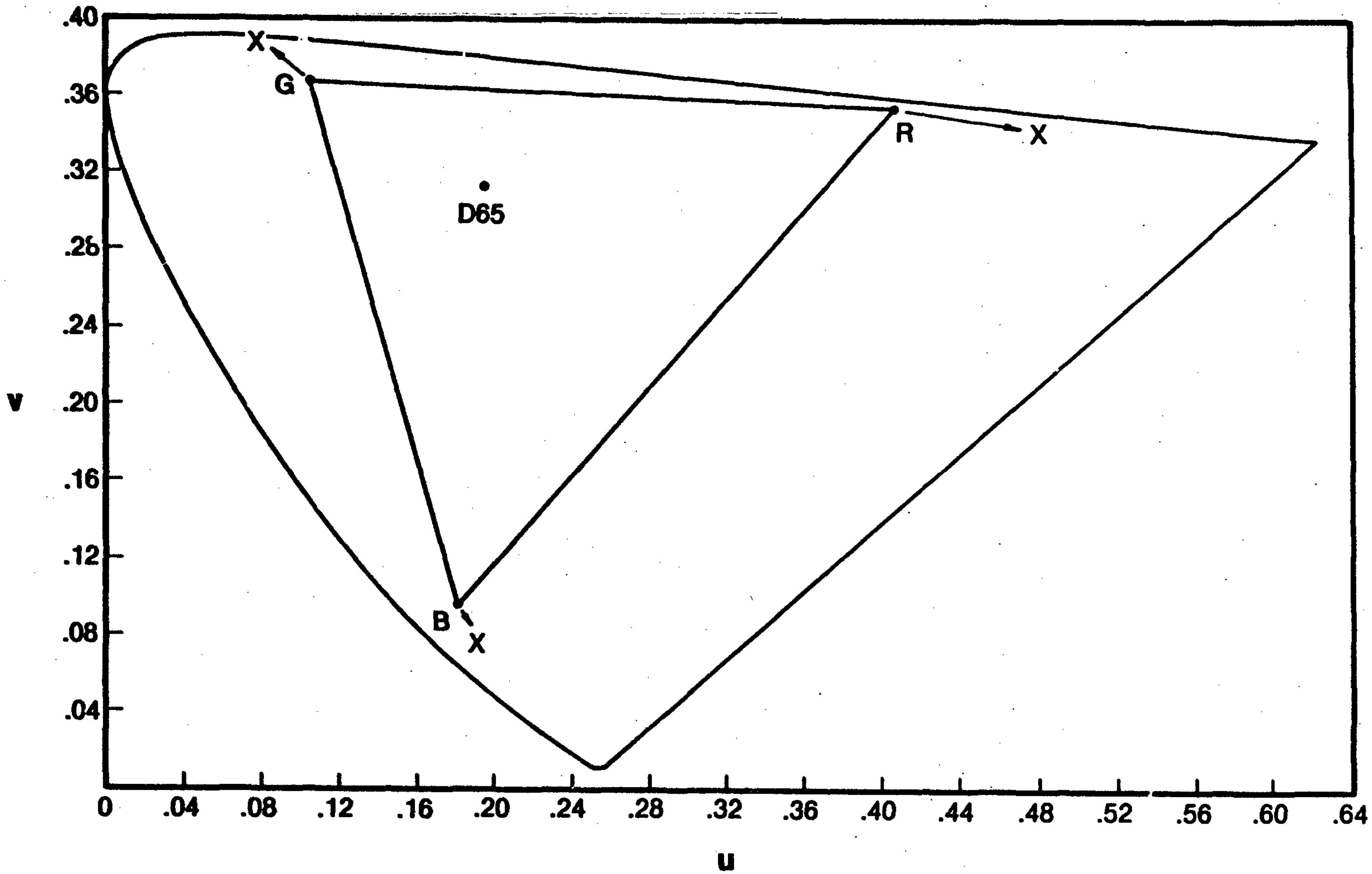


Figure 6. Aydin color monitor primaries in 1960 UCS diagram

Table 2. CRT Display Variability of Luminance, Chromaticity, and Color Temperature Over a One-Month Period

		Luminance (ft.L)	u	v	°K
MAXIMUM INTENSITY WHITE	mean	27.8	.1837	.2833	14,854
	s.d. n=65	.76	.0015	.0014	975
D-65 WHITE	mean	22.7	.1956	.3128	6,623
	s.d. n=9	.41	.0010	.0007	91
RED	mean	6.34	.4069	.3562	-
	s.d. n=15	.31	.0050	.0003	-
GREEN	mean	16.36	.1090	.3665	-
	s.d. n=22	1.84	.0073	.0028	-
BLUE	mean	2.72	.1798	[illegible]983	-
	s.d. n=16	.20	.0031	.0064	-

An examination of Table 2 shows that the luminance of the CRT monitor was quite stable. The green primary appears to exhibit the most variability both in terms of luminance and chromaticity. The color temperature is not very stable; small changes in green may have been responsible. Given that the measurements took place over an extended period, it appears that the monitor was reasonably consistent for our purposes.

5.3 COLOR PREDICTION

The accuracy of predictions of the chromaticities of CRT colors based on multiple regression of the software bit-value assignments rests strongly on the choice of colors chosen for deriving the regression coefficients. To test the efficacy of the regression strategy, we created three data sets of measured CRT chromaticities.

Data Set 1 was derived from points chosen from the limits of the CRT color gamut. When the chromaticities of the three CRT primaries are joined by lines, the limits of possible CRT colors are defined by the space enclosed within the resulting triangle. Figure 7 shows the triangular color gamut and the 61 chromaticities that comprise Data Set 1.

Data Set 2 consisted of 27 points beginning at D65 white and radiating outwards towards the midpoint of each side of the triangle that defines the limits of the CRT color gamut. Figure 8 shows the chromaticities of this data set.

Data Set 3 was formed from 63 color points that had been measured over a period of time. The points were somewhat randomly scattered throughout the color space and are shown in Figure 9.

For each of the 3 data sets, multiple regressions were run to predict u and v from the software bit-value assignments for the R, G, B guns. Regressions were also run including luminance as a predictor.

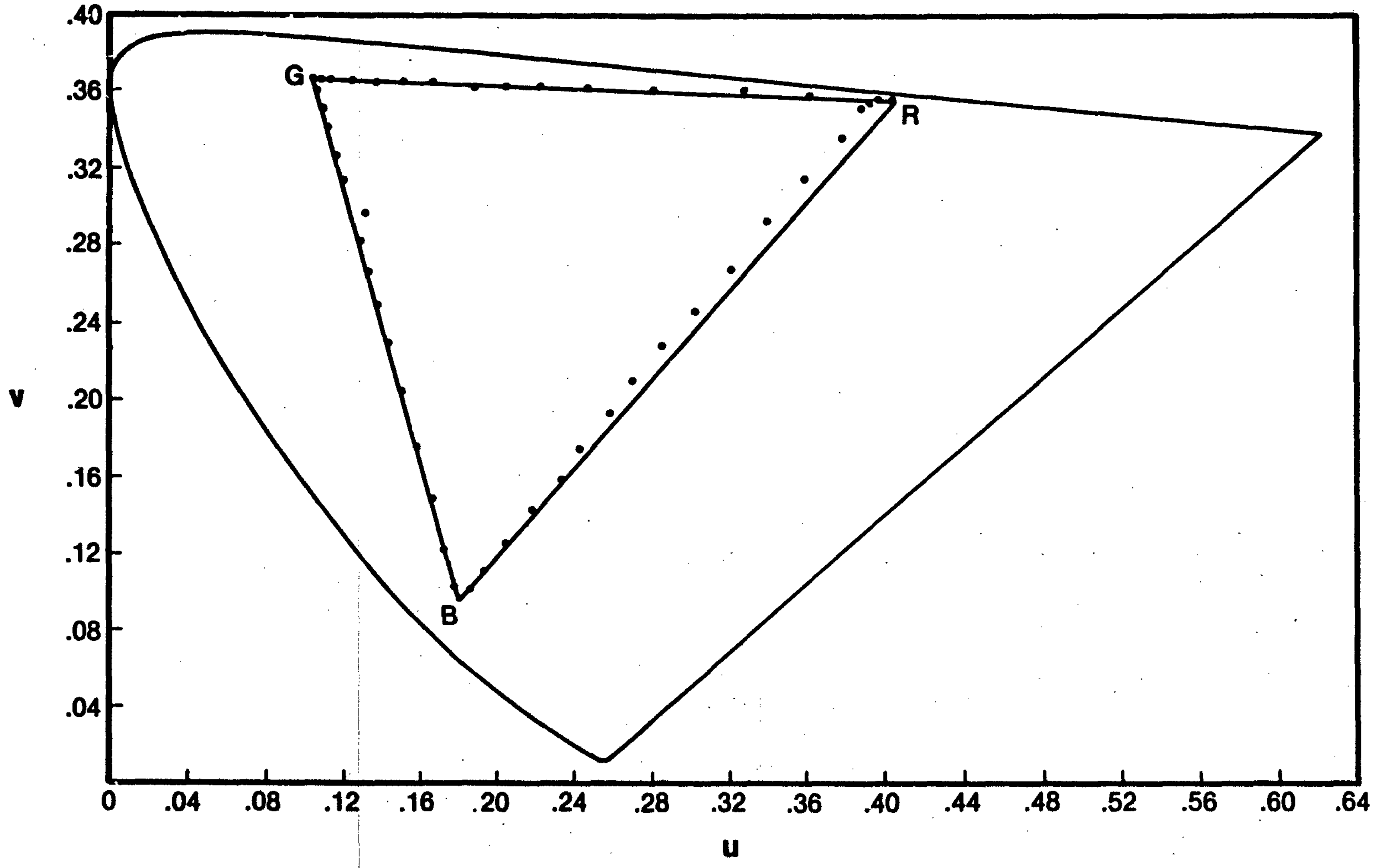


Figure 7. Data Set 1 Chromaticities

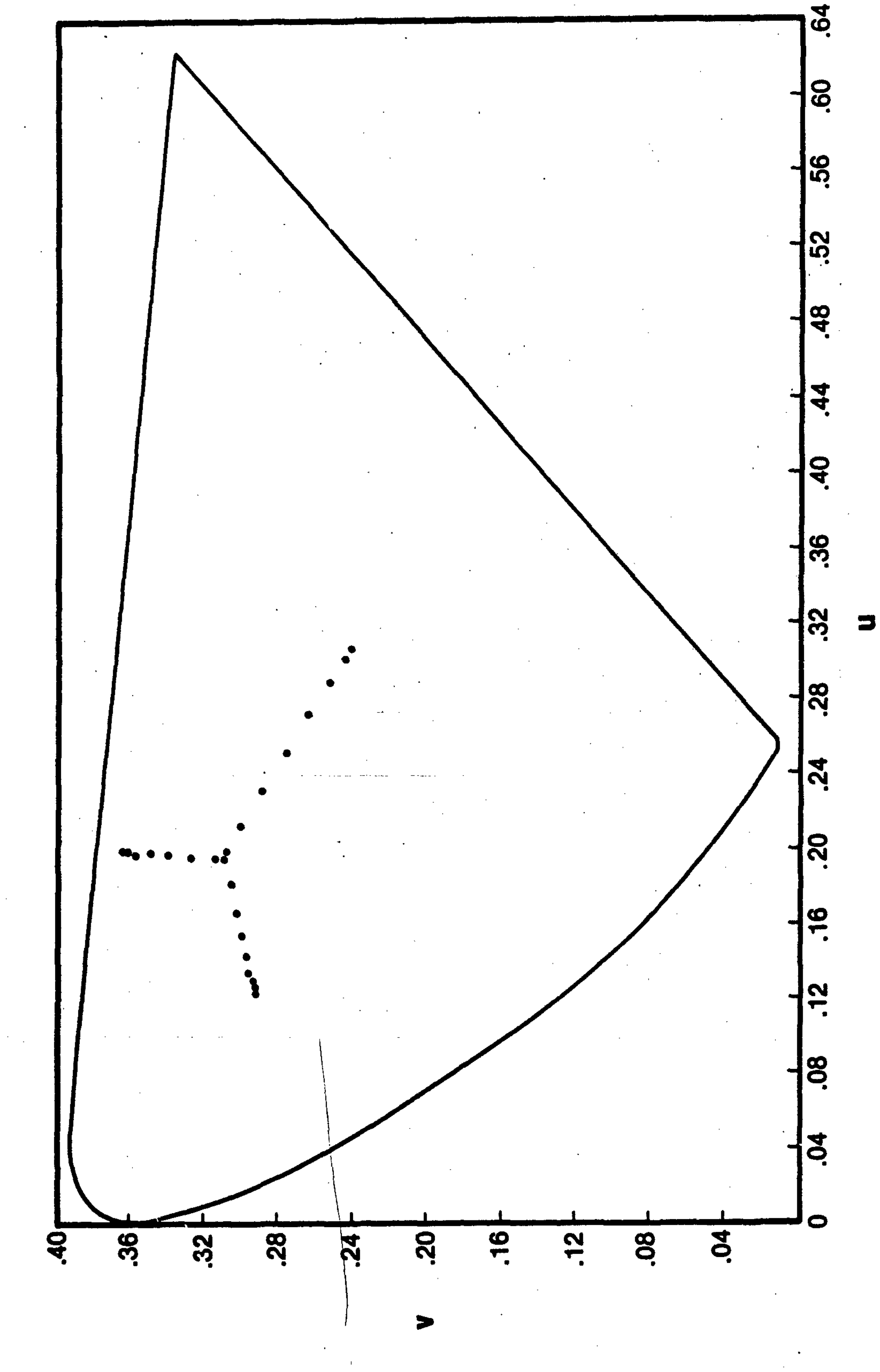


Figure 8. Data Set 2 Chromaticities

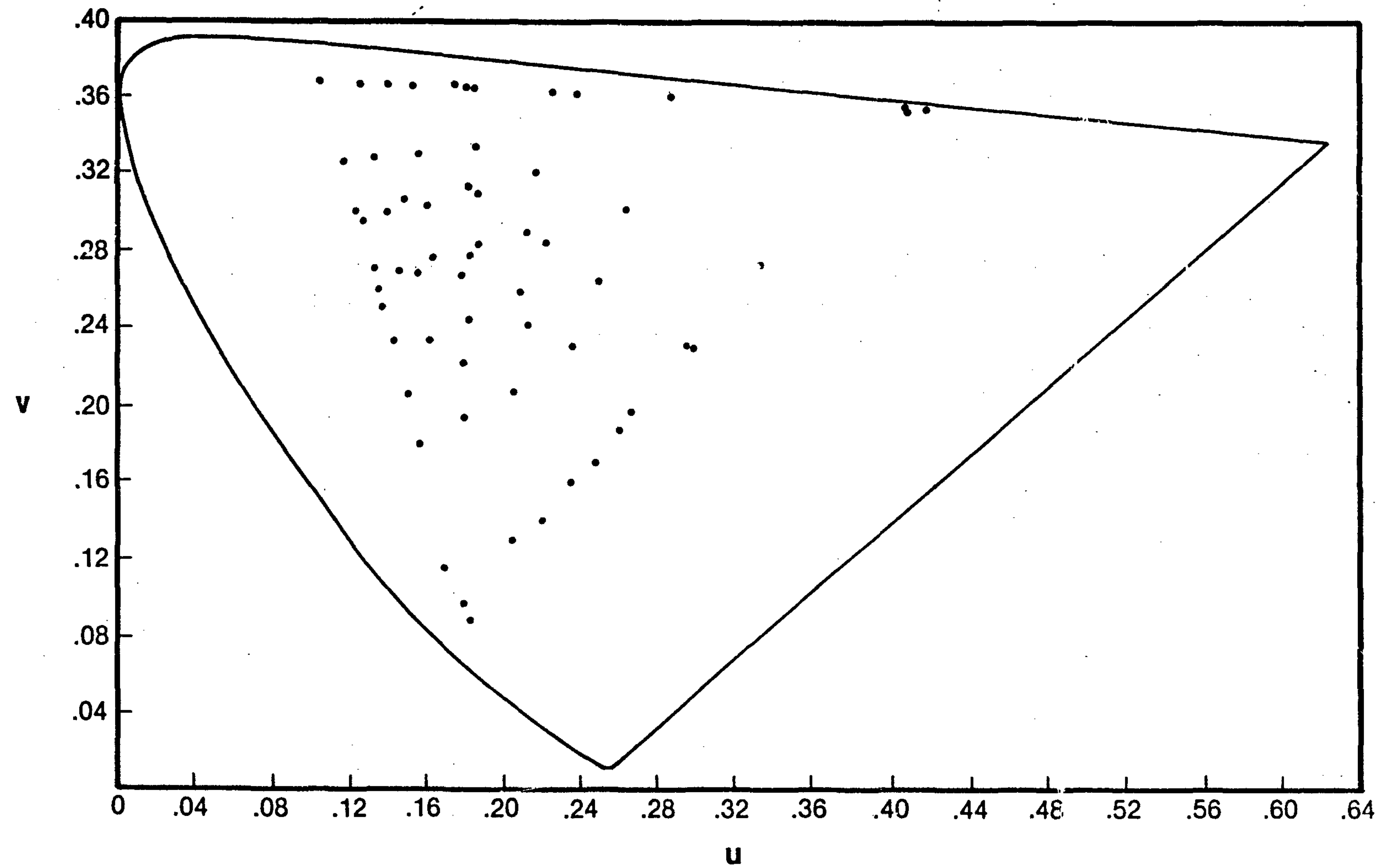


Figure 9. Data Set 3 Chromaticities

The quality of the predictions for Data Set 1 was determined in two ways. First, the Pearson product-moment correlation between the predicted and actual chromaticity coordinates was computed. Then, the regression coefficients were used to predict the chromaticity coordinates of Data Set 3. The predicted and actual coordinates were then correlated. In addition to the correlations, we calculated the mean and standard deviation of the Euclidean distance (on the 1960 UCS diagram) from the predicted to the actual points. Table 3 summarizes the results.

The prediction equations are of the form:

$$u = \text{intercept} + rR + gG + bB$$

$$v = \text{intercept} + r'R + g'G + b'B, \text{ where}$$

R, G, B are the software bit-values normalized to 1.0 and r, g, b are the regression coefficients.

When luminance is added to the predictors the equations become:

$$u = \text{intercept} + rR + gG + bB + lL$$

$$v = \text{intercept} + r'R + g'G + b'B + l'L, \text{ where}$$

L is the luminance in ft-L and l is the coefficient.

As is apparent from Table 3, points along the periphery of the color gamut do not predict points inside the color triangle very successfully. The difficulty is that all predictor points were composed of only two primaries. The net effect on multiple regression is that the dimensionality of the input variable space is reduced and the regression is not accurate. Other regressions not reported here also reinforce the idea that the best predictions occur when the colors comprising the regression data set are composed of all three primaries.

The predictions of data sets based on Data Set 2 are shown in Table 4. In Table 4 we show predictions when the data set from which the regression coefficients are derived is reduced from 27 to 18 to 9 points. The regression predictions are still acceptable even when the data base is

Table 3. Multiple Regression Predictions Derived From Data Set 1 and Tested Against Data Set 3

RESULTS FROM REGRESSIONS USING DATA SET 1						RESULTS FOR PREDICTING DATA SET 3			
DATA SET	REGRESSION PARAMETERS	REGRESSION COEFFICIENTS FOR PREDICTING: u	v	CORRELATION u	v	CORRELATION u	v	MEAN DISTANCE ERROR	S.D. OF DISTANCE ERROR
61 Points	Intercept	.2837	.265	.90	.94	.85	.93	.04	.02
	R	.0783	.0583						
	G	-.1499	.0993						
	B	-.0743	-.1446						
61 Points	Intercept	.2747	.274	.90	.95	.81	.91	.05	.03
	R	.1128	.0235						
	G	-.0513	.0002						
	B	-.0598	-.1593						
	L	-.0057	.0057						

Table 4. Multiple Regression Predictions Derived From Data Set 2 and Tested Against Data Set 3

RESULTS FROM REGRESSIONS USING DATA SET 2						RESULTS FOR PREDICTING DATA SET 3			
DATA SET	REGRESSION PARAMETERS	REGRESSION COEFFICIENTS FOR PREDICTING: u	REGRESSION COEFFICIENTS FOR PREDICTING: v	CORRELATION u	CORRELATION v	CORRELATION u	CORRELATION v	MEAN DISTANCE ERROR	S.D. OF DISTANCE ERROR
27 Points	Intercept R G B	.2586 .092 -.1625 -.012	.271 .023 .0962 -.0985	.99	.99	.86	.92	.04	.03
27 Points	Intercept R G B L	.2517 .1027 -.1357 -.0124 -.0013	.2826 .0052 .0514 -.0979 .0022	.99	.99	.86	.92	.04	.03
18 Points	Intercept R G B	.2623 .0959 -.1705 -.0128	.2597 .0259 .1024 -.0933	.99	.99	.86	.92	.04	.03
18 Points	Intercept R G B L	.2609 .0975 -.1665 -.0127 -.0002	.2721 .0118 .0677 -.0942 .0016	.99	.99	.86	.92	.04	.03
9 Points	Intercept R G B	.2628 .101 -.1765 -.0124	.2593 .0271 .1063 -.0995	.99	.99	.86	.92	.04	.03
9 Points	Intercept R G B L	.2791 .0863 -.2117 -.0145 .0015	.2677 .0196 .0883 -.1006 .0008	.99	.99	.85	.92	.04	.04

reduced. The addition of luminance as a predictor variable does not seem to help in the predictions. Based on these data, one could argue that only a few measurements in the interior of the color gamut need be made to achieve "reasonable" chromaticity predictions.

"Reasonable" needs some clarification because it is true that our mean distances are substantially larger than MacAdam's (1942) color-discrimination ellipses. It might be inferred, therefore, that our predictions are too imprecise to be useful for imagery display. We have several reasons to believe that this is not the case. First, MacAdam's (1942) ellipses, which represent one standard deviation in color-matching data, were produced under very carefully controlled experimental conditions. Data collected under conditions relevant to CRT viewing show that color discrimination is much poorer than the MacAdam's (1942) data would suggest (Ward, et al., 1983 a, b). The Ward et al. (1983 a, b) data show that ease of color discrimination is very hue dependent even when the 1960 UCS representation is used; reds are the most difficult to discriminate from one another. Furthermore, it should be remembered that the VIPER facility uses color principally for imagery pseudocoloring. This means that, normally, there would not be more than 15 or 20 simultaneous colors used at any one time and that their luminances would probably be similar. Clearly, our predictions are more than sufficient to insure that 20 discriminable colors can be assigned and be reproduced on the CRT monitor.

The points of Data Set 3 were also subjected to regression analysis to determine if a large number of points would produce better regressions. As Table 5 shows, the correlations are not nearly as high as those of Data Set 2. Data Set 3 was our criterion data set so naturally we had no comparison for mean distance-error.

5.4 CONCLUSIONS

Our overall conclusion is that the regressions on Data Set 2 constitute the best predictions (given our methodology) for the chromaticities that can be expected on the Aydin CRT monitor as a result of given software bit-value assignments. Our methodology demonstrates that only a few points inside the triangular color gamut need be measured for reasonable prediction if interest is restricted to the higher luminances. It should also be noted that we employed only first-order equations for predicting. Our approach throughout this work is to use the simplest procedure to obtain the desired result.

Table 5. Multiple Regression Predictions Derived From Data Set 3.

RESULTS FROM REGRESSIONS OF DATA SET 3					
DATA SET	REGRESSION PARAMETERS	REGRESSION COEFFICIENTS FOR PREDICTING: u	v	CORRELATION u	v
63 Points	Intercept R G B	.2319 .0989 -.1243 -.0323	.2661 .0442 .1097 -.1406	.87	.93
63 Points	Intercept R G B L	.2363 .0906 -.1466 -.0361 .0014	.2691 .0386 .0944 -.1432 .0009	.87	.93
27 Points	Intercept R G B	.1784 .1449 -.1297 -.0056	.2459 .0565 .146 -.164	.97	.98

6.0 VIPER COLOR FILM CALIBRATION

6.1 INITIAL SETUP

The exposure, contrast, and brightness settings of the Matrix Instruments color graphics camera Model 4007 were initially adjusted as suggested by the factory manual. A Konica FS-1 35mm camera back was used with Kodak Ektachrome ASA 64 daylight color film. The film was processed by Kodak-recommended standard chemistry.

A test series of slides in which 64 grey shades (luminance gradations) were assigned one of the three primary colors was produced. From this series, we determined the best relative exposure settings for each primary (color separation filter). Attempts to reproduce the pure grey shades by combining the three primaries failed initially. We found white and grey shades to be unmistakably red. The excess red exposure also gave the colors a muddy appearance. We began a systematic variation of the red and blue exposures until an acceptable white could be produced along with the best possible dynamic range of color. When this had been achieved, the overall exposure was varied systematically and adjusted.

The slides were projected onto a screen (Knox Series 300 Spectator) by a Kodak Ektagraphic projector (Model AF-2) at a distance of 13 feet. The zoom lens (Ektanar, 4-6 in, f/3.5) of the projector was set to produce the largest image. The chromaticities were measured using the spectroradiometer at a distance of 12 ft. from the screen with the room lights extinguished.

6.2 DISPLAY VARIABILITY

A number of readings of luminance, chromaticity, and color temperature (where applicable) were taken from slides made in five different sessions during a three-week period. Table 6 presents the means and standard deviations of the three measures for the maximum intensity white; luminance and chromaticities are also shown for the three primaries.

Table 6. Color Slide Variability for Luminance, Chromaticity, and Color Temperature Over a Three-Week Period

		LUMINANCE (FT-L)	u	v	°K
MAXIMUM INTENSITY WHITE	mean	20.5	.2230	.3484	3758
	s.d. n=23	2.00	.0020	.0024	89
RED	mean	2.15	.5056	.3491	-
	s.d. n=11	.66	.0245	.0022	-
GREEN	mean	7.92	.1586	.3754	-
	s.d. n=10	.94	.0059	.0014	-
BLUE	mean	.16	.1841	.0832	-
	s.d. n=11	.20	.0038	.0284	-

An examination of Table 6 reveals that the white is very stable. This stability is an indication that the primaries must also be consistent as small changes in the balance of primaries will usually alter the color temperature of the white.

6.3 COLOR PREDICTION

As was done for the CRT colors, multiple regression of the software bit-value assignments was used to predict chromaticities of the slide colors. Again, the index of prediction quality was the correlation between predicted and actual chromaticity coefficients. In the case where previously derived regression coefficients were used to predict to a different set of points, the mean Euclidean distance from the actual to predicted points was used as a quality index. Two data sets of measured slide chromaticities were used. Data Set 1 consisted of 60 points. Four intensity levels from each primary were combined in all possible combinations, resulting in a set of 60 stimuli with useable densities. Figure 10 shows the 60 chromaticities in this data set. Note that these chromaticities represent the yellow-green region of the film color gamut well. Saturation ranges are also well-represented. The figure does not show many reds or blues because slides made with these chromaticities had very little transmittance and were therefore not useful for our purposes. Our interest in color discrimination and pseudocolor requires film imagery with high transmittances.

All of the points in Data Set 2 were produced by restricting one or two of the primaries to zero intensity. In effect, this produced points with chromaticities near the extremes of the film color gamut. Actually seven basic or reference colors were originally produced (see Figure 11) and then small variations were introduced into the principal primary to generate a graduated series of chromaticities about the reference color. Only the reference chromaticities are plotted in Figure 11 because each chromaticity series is indistinguishable from the reference point on the scale of the figure. In all, there were 62 points in Data Set 2.

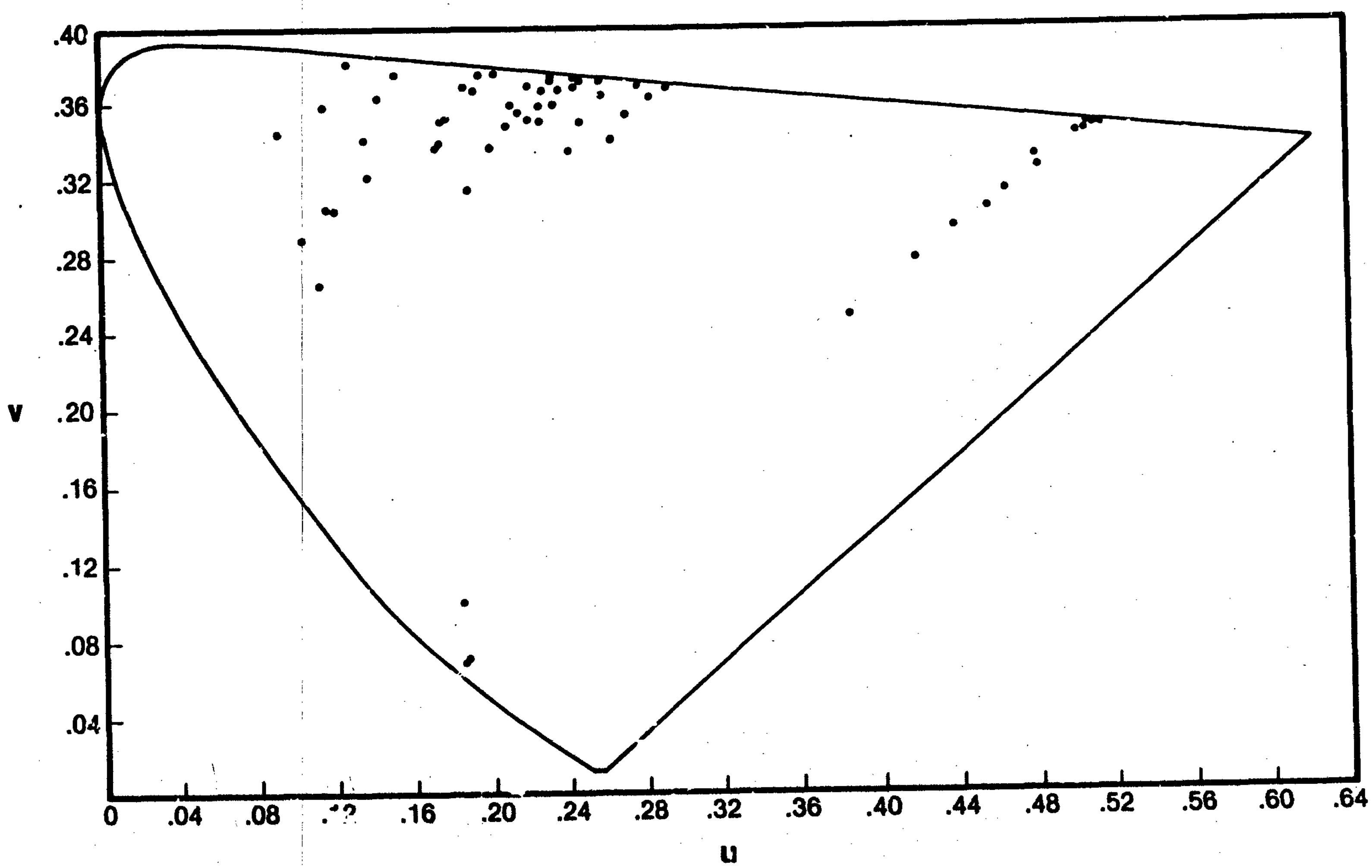


Figure 10. Data Set 1 Chromaticities

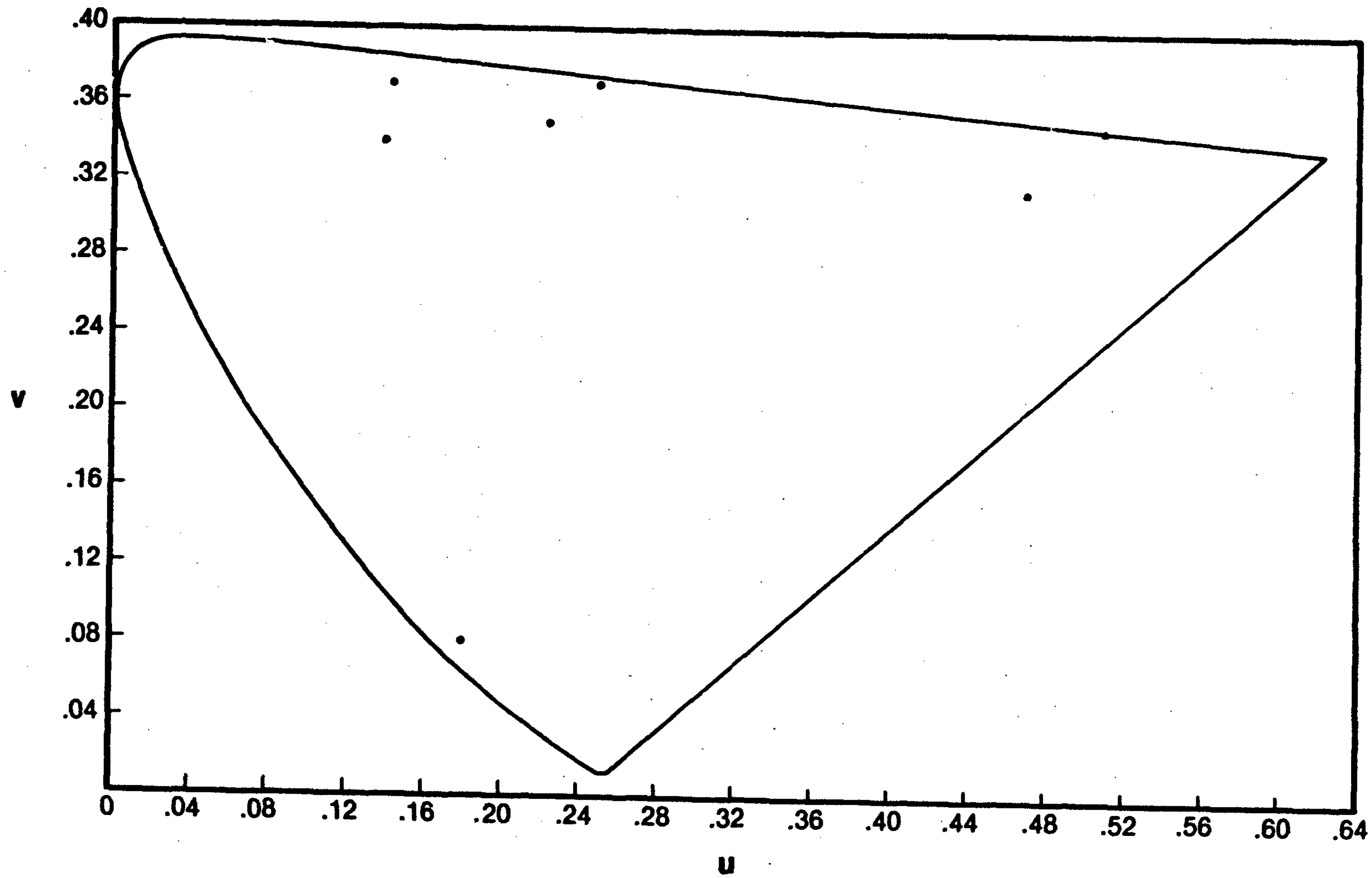


Figure 11. Data Set 2 Reference Chromaticities

The results of the regression analysis on both data sets are shown in Table 7. Preliminary analysis showed that regression predictions for film were inferior to those for the CRT colors. The correlation coefficients were somewhat improved by the addition of second order variables to the regression equation. The prediction equations are of the form:

$$u = \text{intercept} + r1R + g1G + b1B + r1'R^2 + g1'G^2 + b1'B^2$$

$$v = \text{intercept} + r2R + g2G + b2B + r2'R^2 + g2'G^2 + b2'B^2$$

As was the case for the CRT hues, the addition of luminance as a predictor variable did not seem to improve the predictions.

Data Set 1 was reduced to 49 points by eliminating those points with u or v residuals above 0.05 in the original regression. This resulted in considerable improvement in the correlation coefficients for the regressions. Figure 12 shows the chromaticities of the 49 points retained. This data set was then systematically reduced further to determine the minimum number of points necessary for reasonable prediction.

The 49 point data set was reduced progressively to 33, 22, 15, and 10 points. Each reduction was made by deleting every third observation from the next larger data set.

To test the efficacy of the regression coefficients determined for each subset above, we predicted selected points of Data Set 2. The points chosen were those for which not more than one of the primaries was set at zero intensity. The results of these analyses are shown in Table 8. It is clear that the mean error distance of prediction does not increase dramatically until our regression data base is reduced to 10 points. Based on the individual u and v correlations, 22 points may constitute the minimum regression data base.

Table 7. Regression Data for Complete Data Sets 1 and 2

RESULTS FROM REGRESSIONS USING DATA SETS 1 AND 2						RESULTS FOR PREDICTING DATA SET 2			
DATA SET	REGRESSION PARAMETERS	REGRESSION COEFFICIENTS FOR PREDICTING: u	v	CORRELATION u	v	CORRELATION u	v	MEAN DISTANCE ERROR	S.D. OF DISTANCE ERROR
Data Set 1 60 points	Intercept R G B	.3244 .1644 -.2316 -.0499	.2586 .0777 .0941 -.0506	.92	.76	.94	.89	.05	.03
Data Set 1 60 points	Intercept R G B R^2 G^2 B^2	.3316 .226 -.4923 -.0333 -.0615 .2778 -.0128	.2437 .1483 .1668 .0098 -.0742 -.0766 -.0638	.95	.77	.96	.91	.04	.03
Data Set 2 62 points	Intercept R G B	.3081 .1642 -.1789 -.0683	.2354 .0981 .1125 -.0804	.88	.82	Did not predict, as the set being predicted is a subset of Data Set 2			
Data Set 2 62 points	Intercept R G B R^2 G^2 B^2	.3206 .4498 -.5985 -.3276 -.2962 .4293 .2626	.2343 .254 .1062 -.1522 -.1607 .0065 .0779	.90	.82				

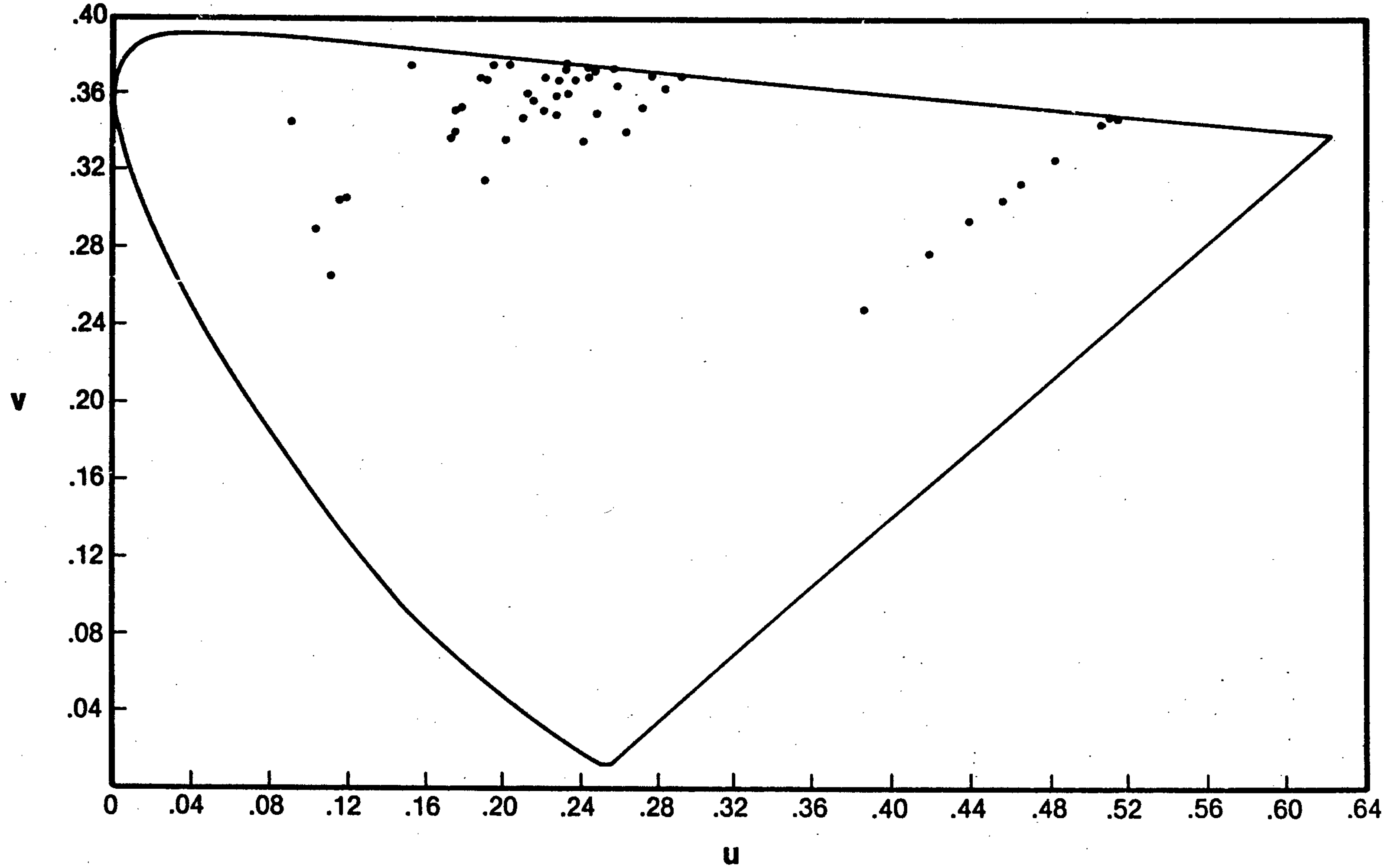


Figure 12. Chromaticities of Reduced Data Set 1

To ensure that the above results were not due to the particular selection procedure, we subdivided the 49-point data in three different ways. A 16-point data set was formed consisting of those points deleted when the data set was reduced from 49 to 33 points. A 10-point data set was formed by selecting intensity bit-values for single primaries and grey shade combinations. A 12-point data set was formed by selecting colors representing the entire useful color space with each color approximately equidistant from its nearest neighbors (see Figure 17).

The analyses of the regressions of these three additional data sets are also shown in Table 8. The 16- and 12-point data sets are vastly superior to the 10-point set based on mean error distance. The poor prediction associated with the 10-point set is probably due to the fact that these points fall at the extremes of the color space. When the 16-point data set was compared to the previous subdivision of the 49 points, we found that this set was equally effective.

When the 12- and 16-point data sets are compared, the 12-point set is a superior predictor. This is not surprising because the 12-point set was chosen systematically to represent the entire useful color film gamut.

6.4 CONCLUSION

The data presented above indicate that regression analysis can be applied to predict chromaticities resulting from a subtractive color process (film) as well as additive color processes (CRT). As was the case for the CRT color predictions, film color predictions are best when the predicting data base is restricted to points from the interior of the available color gamut. Insuring a systematic representation of the color gamut produces the best data base.

Table 8. Regression Data for Subsets of Data Set 1

RESULTS FROM REGRESSIONS USING DATA SET 1						RESULTS FOR PREDICTING DATA SET 2			
DATA SET	REGRESSION PARAMETERS	REGRESSION COEFFICIENTS FOR PREDICTING: u	v	CORRELATION u	v	CORRELATIONS u	v	MEAN DISTANCE ERROR	S.D. OF DISTANCE ERROR
49 points eliminate high residuals from original 60 points	Intercept	.3714	.3009	.99	.91	.99	.94	.02	.02
	R	.1382	.0487						
	G	-.5293	.0789						
	B	-.0296	.0237						
	R^2	.0031	-.0083						
	G^2	.2842	-.0295						
	B^2	-.0111	-.0681						
33 points delete every 3rd observation from 49 points	Intercept	.3657	.3053	.99	.91	.99	.95	.02	.02
	R	.1454	.0445						
	G	-.531	.0752						
	B	-.0146	.0241						
	R^2	-.0063	-.01						
	G^2	.2881	-.0264						
	B^2	-.0222	-.0658						
22 points delete every 3rd observation from 33 points	Intercept	.3717	.2856	.99	.94	.99	.87	.02	.02
	R	.1248	.0804						
	G	-.5094	.0934						
	B	-.0068	.014						
	R^2	.0138	-.0245						
	G^2	.2588	-.0437						
	B^2	-.0311	-.0567						
15 points delete every 3rd observation from 22 points	Intercept	.3573	.2659	.99	.95	.98	.76	.03	.03
	R	.1598	.1295						
	G	-.497	.1057						
	B	-.0351	-.0081						
	R^2	-.005	-.0504						
	G^2	.2429	-.0582						
	B^2	-.004	-.0349						
10 points delete every 3rd observation from 15 points	Intercept	.175	.2985	.99	.90	.81	.87	.09	.08
	R	.6303	.0454						
	G	-.0549	.0849						
	B	-.019	-.0887						
	R^2	-.297	.0217						
	G^2	.2528	-.0492						
	B^2	-.0167	.0432						
16 points using every 3rd observation from 49 points	Intercept	.3911	.2906	.99	.89	.98	.90	.03	.02
	R	.1032	.0585						
	G	-.5713	.0913						
	B	-.1098	.0249						
	R^2	.0377	-.007						
	G^2	.3375	-.0415						
	B^2	.0596	-.0718						
10 points primaries and gray shades	Intercept	.3097	.2242	.99	.99	.81	.87	.3	.07
	R	.5158	.3127						
	G	-.4706	.3392						
	B	-.3115	-.3117						
	R^2	-.3167	-.1888						
	G^2	.31	-.1852						
	B^2	.1886	.1581						
12 points selected for false color study	Intercept	.3641	.3121	.99	.80	.98	.95	.03	.02
	R	.0274	.0278						
	G	-.6083	.028						
	B	-.1685	.0965						
	R^2	.1172	.0072						
	G^2	.3879	.0264						
	B^2	.1533	.1416						

7.0 COLOR DISCRIMINATION EXPERIMENT

In the previous section we have shown that it is possible to predict adequately film color from software bit-values. These predictions were all done using the 1960 UCS diagram. Equal distances on this diagram are supposed to represent equal perceptual distances or steps. As was discussed earlier, there is some uncertainty about the adequacy of the UCS representation for certain color applications. Since our interest is in pseudocolor of film imagery, it is important to establish the suitability of the UCS representation for this application. The following experiment was designed to determine the limits of film color discrimination for a color-normal population.

7.1 METHOD

7.1.1 Stimuli

Our stimuli were designed to present two juxtaposed patches of color of equal or slightly different chromaticities. Seven chromaticities were chosen as reference colors and always appeared on the left side of the stimulus field. The reference colors chosen were the three primaries, the three hues produced by combining equal-intensity bit-values of two primaries, and a white produced by combining equal-intensity bit-values of the three primaries. (The chromaticities of these reference colors were shown in Figure 11.)

The white data set consisted of three subsets. Each subset was formed by systematically reducing the intensity bit-value of one of the three primaries. This resulted in a yellow-tinted set, a purple-tinted set, and a cyan-tinted set; all of these sets had the same white reference hue. Table 9 lists the reference white and its three subsets of comparison hues. The comparison hue, which was sometimes identical to the reference hue, always appeared on the right side of the stimulus field.

For each of the reference colors other than white, a single data set was formed consisting of hues which were identical to the reference hue

Table 9. Chromaticities of Desaturated Stimuli Used in Color Discrimination Experiment

REFERENCE HUE	COMPARISON HUE u	v	Luminance (Ft-L)	UCS-Distance from Reference
Yellow-Tinted White	.2243	.3496	19.7	.002
Luminance = 22.1 Ft-L	.2200	.3507	19.8	.002
u = .2223	.2221	.3535	19.1	.007
v = .3334	.2279	.3577	20.7	.010
	.2299	.3597	19.3	.014
Purple-Tinted White	.2235	.3485	19.1	.0001
Luminance = 22.1	.2234	.3475	19.1	.0007
u = .2223	.2238	.3475	19.3	.001
v = .3334	.2251	.3495	19.8	.001
	.2257	.3477	20.1	.002
Cyan-Tinted White	.2221	.3474	19.8	.002
Luminance = 22.1	.2216	.3480	19.7	.002
u = .2223	.2192	.3479	20.6	.006
v = .3334	.2180	.3531	23.7	.010
	.2141	.3485	19.1	.011

or varied slightly in chromaticity from it. These hues, which always appeared in the right side of the stimulus field, were produced by systematically reducing the intensity bit-value of one of the primaries. Table 10 lists these six reference colors and their respective data sets. There were 60 slides in all.

7.1.2 Subjects

The subjects were 63 students who were enrolled in an introductory Psychology class at a midwestern university. They received extra credit for participating in the study. No subjects reported having any color vision problems when questioned.

7.1.3 Testing Procedure

Subjects were tested in groups numbering 6 to 10. They were seated 8 to 15 feet from the projection screen on which slides were projected by a Kodak Ektagraphic projector from a distance of 13 feet. Each block of color measured 15 in. wide x 25 in. long on the projection screen. The two adjacent blocks to be discriminated were separated by a vertical line 0.5 in. wide. The stimulus pattern subtended a minimum of 9.3 degrees of visual angle for subjects furthest from the screen. All lights in the room were extinguished with the exception of a single 60-watt bulb providing indirect lighting behind the projector and the subjects.

The 60 slides were presented in random order for approximately 10 seconds each. The subject's task for each slide was to indicate on his answer sheet whether he perceived the two blocks of color to be the same or different.

7.2 RESULTS

Figures 13 and 14 show the proportions of persons perceiving the two blocks of color as different as a function of the UCS-distance between the hues. Where hues are principally white, such as those in Figure 13,

Table 10. Chromaticities of Saturated Stimuli Used in Color Discrimination Experiment

REFERENCE HUE	COMPARISON HUE			UCS-Distance from Reference
	u	v	Luminance (Ft-L)	
GREEN Luminance = 7.76 Ft-L u = .1556 v = .3753	.1492 .1423 .1343 .1277	.3765 .3780 .3796 .3800	6.26 5.72 4.71 4.05	.007 .016 .024 .032
BLUE Luminance = 1.02 u = .1831 v = .0744	.1853 .1834 .1837 .1808 .1745	.0768 .0745 .0841 .0785 .0915	.0890 .0731 .0741 .0582 .0415	.001 .002 .009 .006 .017
RED Luminance = 1.73 u = .5160 v = .3480	.5150 .5129 .5143 .5107 .5023	.3483 .3483 .3485 .3480 .3497	1.76 1.71 1.63 1.55 1.72	.0009 .004 .003 .004 .003

Table 10. (Continued)

REFERENCE HUE	COMPARISON HUE			UCS-Distance From Reference
	u	v	Luminance (Ft-L)	
YELLOW	.2467	.3713	17.7	.0003
Luminance = 16.2 Ft-L	.2466	.3720	16.7	.002
u = .2444	.2482	.3721	15.1	.005
v = .3712	.2524	.3721	14.8	.008
	.2564	.3720	16.3	.011
PURPLE	.4549	.3117	3.39	.005
Luminance = 3.34	.4612	.3107	2.88	.008
u = .4506	.4732	.3189	2.83	.019
v = .3093	.4758	.3231	2.77	.030
	.4842	.3299	2.66	.041
CYAN	.1393	.3330	10.3	.003
Luminance = 10.1	.1361	.3334	9.58	.006
u = .1425	.1377	.3381	9.39	.011
v = .3301	.1369	.3441	8.96	.018
	.1389	.3519	8.84	.026

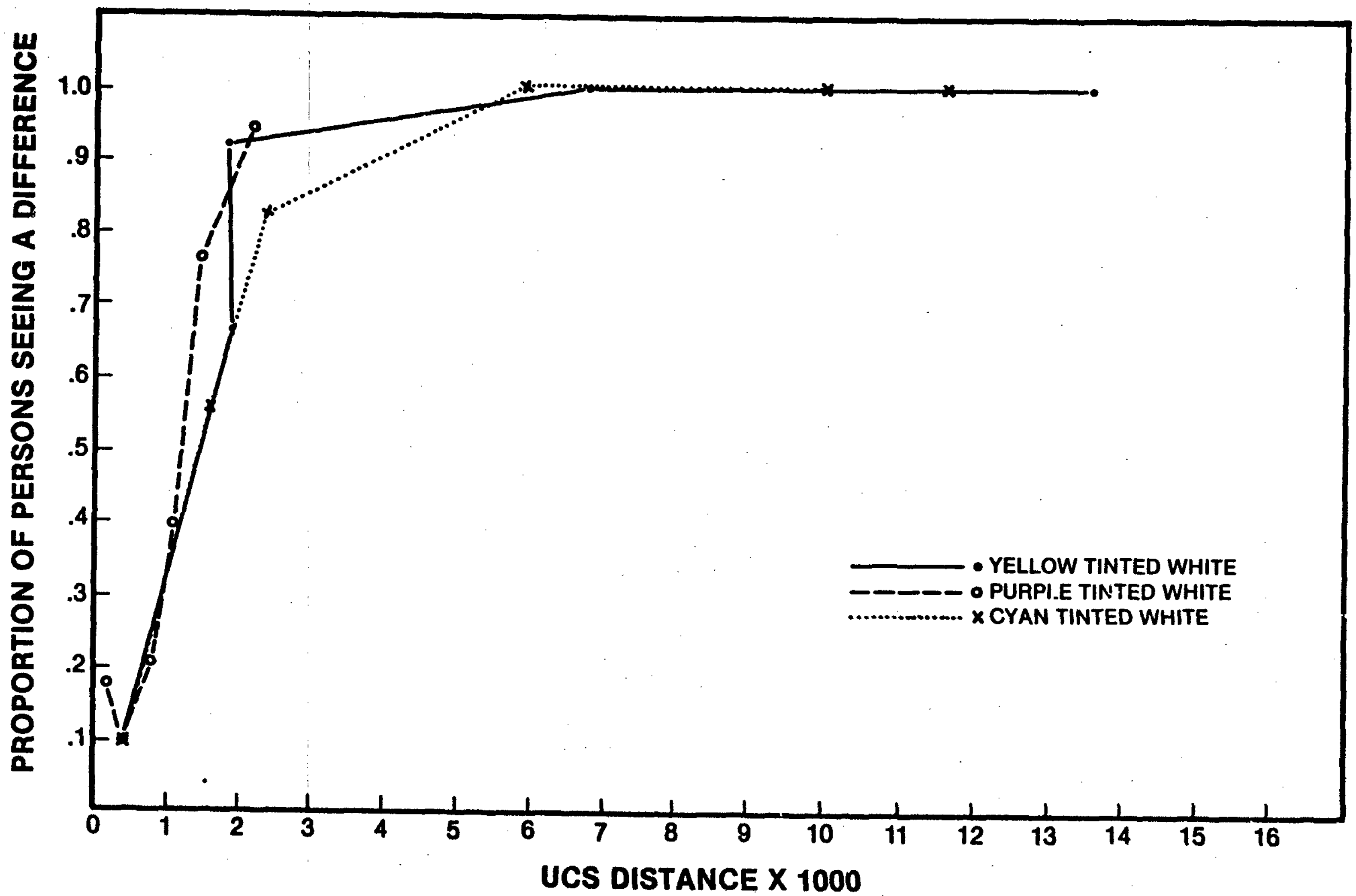


Figure 13. Detection of Color Difference for Desaturated Colors as a Function of UCS Distance

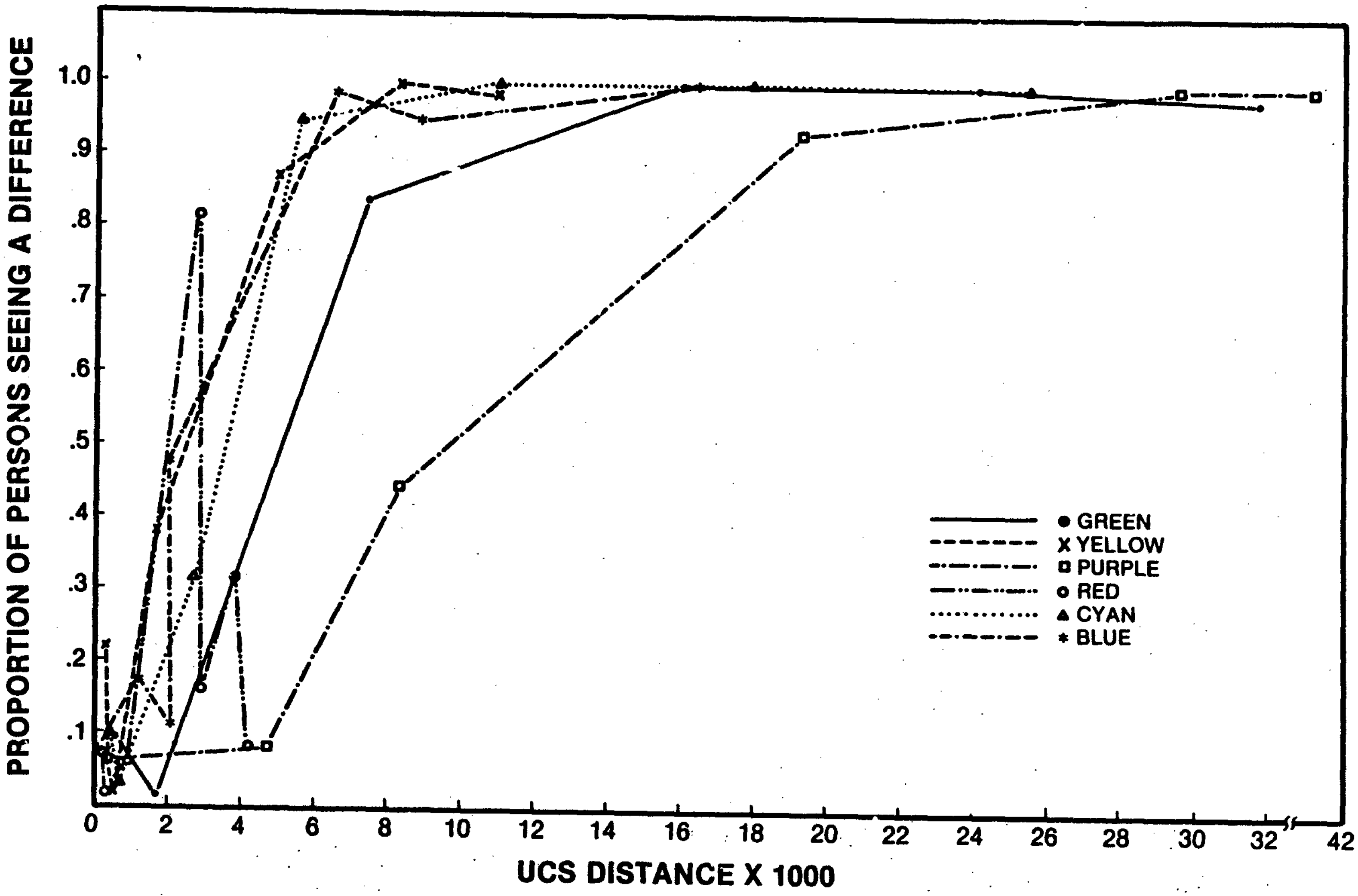


Figure 14. Detection of Color Difference for Saturated Colors as a Function of UCS Distance

most subjects can discriminate hue differences of very small UCS distances. As the UCS-distance between hues approaches 0.006, nearly 100 percent of subjects perceived a color difference. Discrimination of highly saturated hues, as shown in Figure 14, is almost as good for cyan, blue, yellow, and greens as for the tinted whites. At a UCS distance of 0.008, color discrimination is nearly 100 percent for cyan, blue, and yellow. Discrimination for green reached 100 percent at approximately 0.015 UCS distance-units. However, discrimination of the red and purple hues is much poorer. In fact, there are some reversals for very small UCS distances. Presumably, these are due to experimental error.

One might question whether reduced luminance is the cause of poor discrimination of the purples and reds. However, looking back at Table 10, we can see that the luminances do not differ greatly.

We must therefore conclude that persons cannot discriminate chromaticity differences of reds and purples as well as those of other hues. Ward, et al. (1983 a, b) reported the same result for hue discrimination in high luminance conditions.

7.3 CONCLUSION

These data and those of others indicate that discrimination is poor for red hues or those hues composed of a high red content. Therefore, we recommend that reds, purples, and oranges be avoided or their chromaticities be widely spaced to insure adequate color discrimination.

8.0 COLOR-CODING

8.1 INFORMATION DISPLAY WITH COLOR

The use of color to code information displays goes back to William Playfair who, in 1801, began color-coding maps and graphical materials. The use of color in statistical maps has been in widespread use since the mid 19th century (Fienberg, 1979). Only recently, however, have researchers begun to study the perceptual effects of color-coded informational displays. Some of these coding effects can be quite subtle. In map color-coding, red areas tend to be judged larger than green areas if the colors are highly saturated; there is no difference for low saturation colors (Cleveland and McGill, 1983). A similar effect has been reported with Munsell color chips and, in general, warm colors (reds and yellows) appear larger than cool colors (greens and blues) (Tedford, Bergquist, and Flynn, 1977).

The use of color can be overdone. Just because the eye can discriminate over a million hues (Halsey and Chapanis, 1951) does not mean that <u>any</u> arrangement of color coding will be useful. The Bureau of Census has used hue and saturation to color-code two dimensions in a bivariate map. Their approach quickly leads to visual confusion (Trumbo, 1981). Tufte (1983, p. 153) has commented that "the complexity of multifunctioning elements can sometimes turn data graphs into visual puzzles, cryptographical mysteries for the viewer to decipher."

8.2 COLOR AND RADAR IMAGERY

The use of color to represent dimensions has been extended to radar images with only limited success. Harney, Martin, and Sullivan (1978, cited by Huddleston, 1983) coded infrared radar strength returns as 13 levels of brightness and range information as hue, with red for close, yellow for intermediate, and blue for far. There were four levels of each hue. The utility of the coding was restricted by the limited CRT brightness range and variations in ambient lighting.

Most work with radar imagery has been limited to a univariate representation; the intensity of the radar return is coded as brightness in a monochrome representation. The number of grey shades available for the coding depends upon both the sensor and display media. In general, most systems are display limited. Ordinary CRTs seem to be capable of 4-6 bit (16 to 64 grey shades) resolution (Briggs, 1981) which is well below the dynamic range of modern Synthetic Aperture Radar (SAR) systems, for instance. Lamonica (1977) reports a dynamic range of 60 dB for SAR systems and an ideal range of only about 30 dB for CRT and film display media. Obviously all of the information gathered cannot be displayed at once. Actually, there is no advantage to having a better display range as the instantaneous monochrome dynamic range of the human visual system is about 25-35 dB under ideal conditions (Graham, 1965, p. 26). To circumvent the display limits, typically the sensor signal is compressed at the high or low end or the entire range is compressed, with the result that the fine granularity needed for target recognition is sacrificed.

As color display systems have become more common, researchers have become interested in exploiting color as a coding vehicle. This may be done by using color to extend the "effective" dynamic range of the display or the dynamic range may be compressed and color-coded to declutter the imagery. Radar information is usually coded in the intensity (brightness) dimension, but the information could be coded as hue and/or saturation.

Many radar color-coding schemes use some type of density (bit) slicing technique (Gonzalez and Wintz, 1977). The available density range of the image is divided into a number of intervals and a color is assigned for each interval. Thus, if a digital image had 256 intensity levels, the range 0-85 might be represented by blue, 86-170 by green, and 171-255 by red. In this example there are only three effective levels of information. Color and intensity are often combined to extend the range. Only a few colors (often 3) are used and the input grey level or "density" range may be divided as is shown in Figure 15.

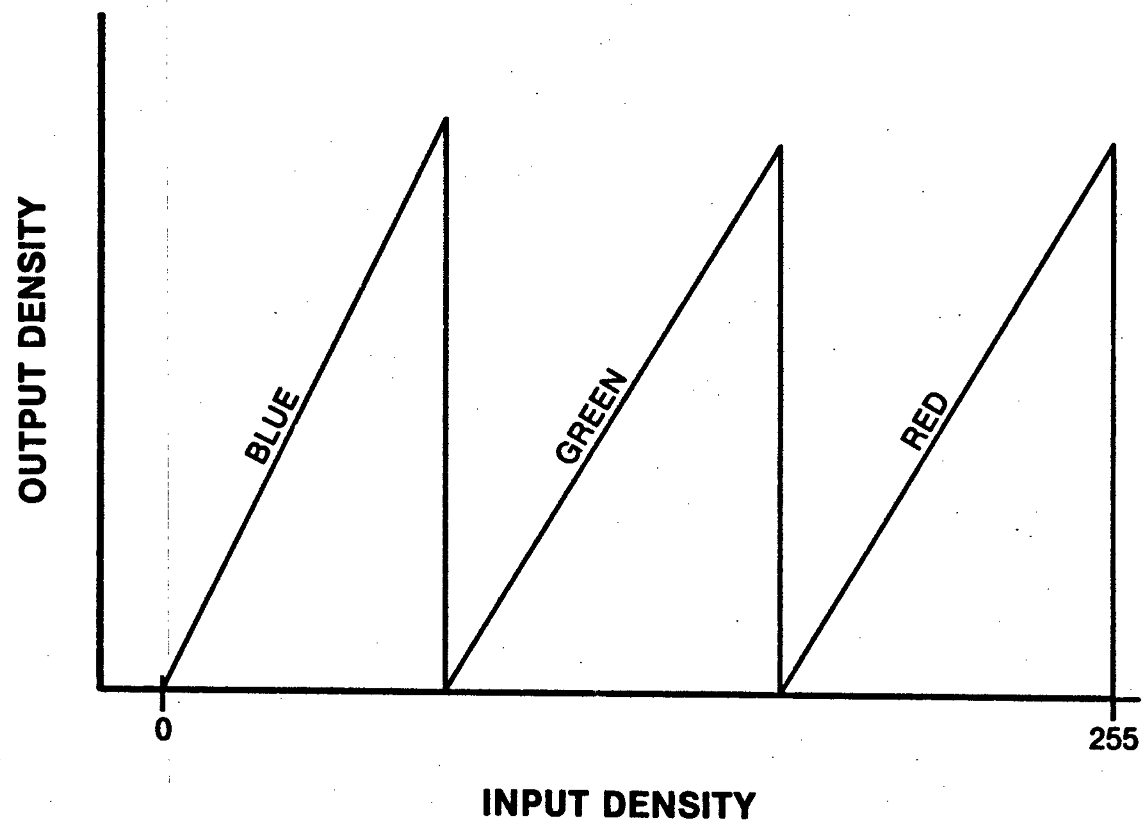


Figure 15. A Color-coding Strategy That Combines Hue and Brightness

The problem with combining hue and intensity is that interspersed colors throughout the image create false boundaries that can be interpreted as changes in terrain types or they may mask interesting features such as military vehicles (Dragavon, Hershberger, and Whitt, 1977). A modification of the bit-slice approach is to retain the color coding and also keep a brightness residual as the colors are transitioned. Thus, the image brightness always correlates with increases in input signal intensity, but colors transition as signal intensity increases. Color becomes a redundant code with this approach. Unfortunately, the colors seem to mask some targets; the transitions are unlikely to occur at the critical densities to accentuate every target type.

Saturation can also be used as a coding strategy. With film and CRT displays there seems to be a limit of about 5 saturation levels for each of the R, G, B primary colors. More levels just blend together as indistinguishable white. Dragavon, et al., (1977) found the results of saturation coding to be "very disappointing" subjectively; they did not collect any objective data.

Lamonica (1977) has performed extensive studies of color-coded SAR imagery. His studies measured observer performance on three kinds of mission tasks: strategic, tactical, and reconnaissance. He generated a variety of color codes, based principally upon terrain rendition, and settled upon two codes for detailed study. Both codes combined color and brightness, with color redundant. There were two different codes because one was thought appropriate for tactical and reconnaissance work and the other was deemed better for ground painting and the larger targets of strategic work. Lamonica (1977) also used a black-white intensity code. In general, he found that pseudocolor did not help strategic mission tasks. Color coding did improve performance for certain tactical and reconnaissance tasks, especially when the target returns were above the background clutter.

In general, it appears that color coding is viewed as an interesting coding strategy for radar imagery, albeit of limited practical value. Dragavon, et al. (1977), felt that color coding led to "ambiguity and difficulty of interpretation." Even though Lamonica (1977) found some value in color coding, the performance differences with respect to black and white were not dramatic.

There is another color-coding procedure which does not appear to have been investigated. The bit-slices could be coded by hue only, keeping the brightness approximately constant. Figure 16 shows that the image input density is divided such that each interval is coded with a different color. Obviously, the number and chromaticities of the hues must be chosen so that all chromatic intervals are discriminable from each other.

The last point above is nontrivial for two reasons. First, there is no indication that past work has attempted to ensure that the colors used in coding would be equally discriminable. Second, even if they had been, there is some doubt about using the UCS representation without considering the difficulty in red discriminations. Our discrimination study showed that the UCS color-difference representation must be used with care. Because of these uncertainties in the past work on color coding, we decided to run a controlled experiment using the same SAR imagery for several types of coding strategies.

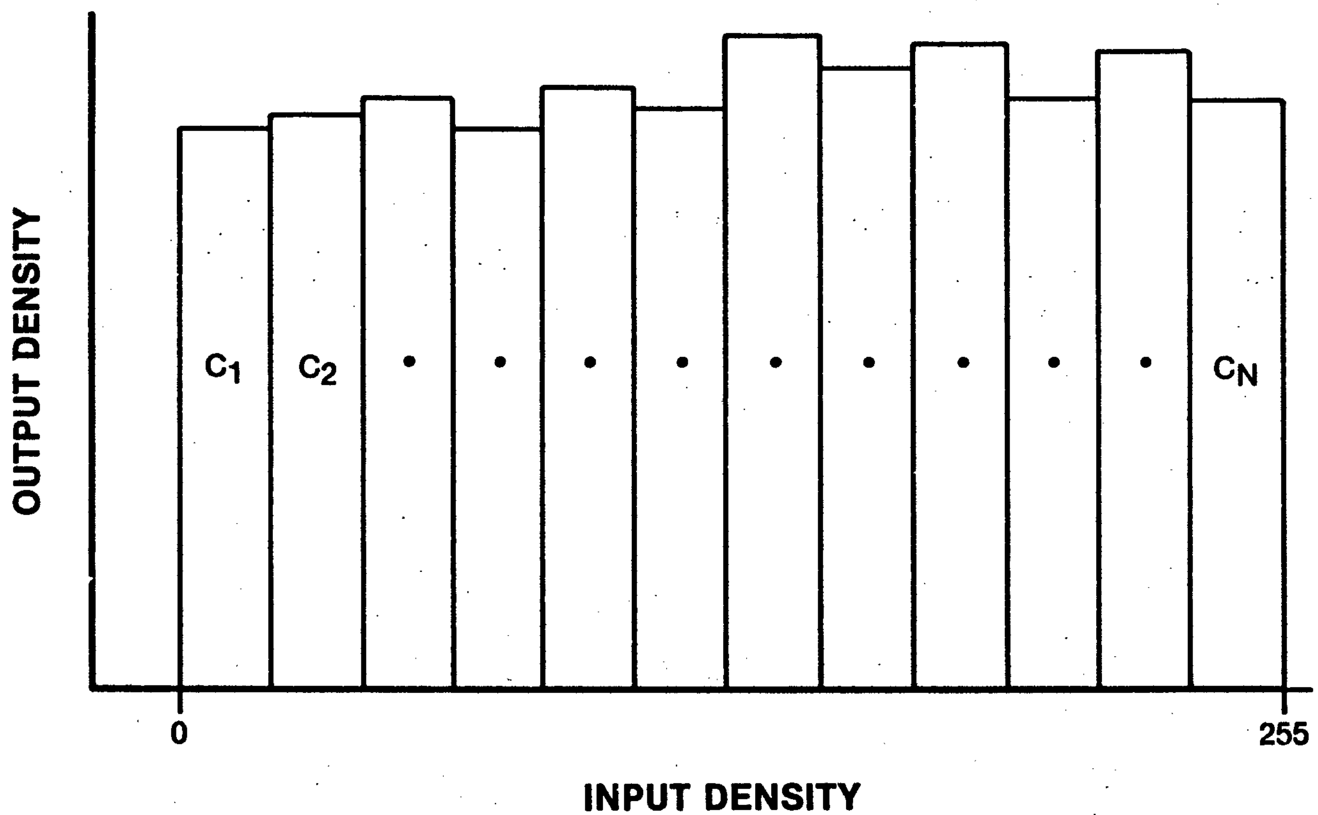


Figure 16. A Color-coding Strategy Based on Hue

9.0 SAR IMAGE INTERPRETATION EXPERIMENT

9.1 METHOD

9.1.1 Stimuli

We took SAR hardcopy (photographic positive print) that had a dynamic range of 23 dB and digitized the imagery (512 x 512 pixels with 8 bit intensity resolution) in the VIPER facility to produce 10 different images each encompassing approximately 3.75 square miles of terrain. The original imagery included an urban area, a seaport, gas and oil refineries, industrial sites, an airfield, and power and gas transmission lines. Ground truth information was taken from standard Geological Survey Maps.

Four coding schemes were applied to the imagery, two black and white and two color. One black and white code was produced by taking the 256 grey shades and outputting them through the Matrix graphics camera onto 35mm Kodak Plus-X Professional film ASA 125. The film dynamic range under our exposure conditions was 19 dB. This was our "normal" black and white code.

The other black and white code was an "enhancement" of the original imagery. The input density for each pixel was normalized over 40 adjacent pixels to produce an image in which the original input range was altered locally. In effect, this produced an image that had a more even distribution of densities, permitting, we thought, better renditions of the input density midranges.

The two color codes tested were those shown in Figures 15 and 16. In the one case, Figure 15, we used blue, green, and red combined with brightness to represent the input density. Each of the three colors was combined with six brightness levels for a total of 18 levels. We designated this condition Color-18 and we have listed the bit-slice range along with the software bit-value assignments (normalized to 1.0) in Table 11. Note that radar returns of low intensity are coded blue and high intensity returns are coded red.

Table 11. Color Codes and Bit-Slice Ranges for Three-Color SAR Coding Strategy

INTENSITY RANGE	BIT-VALUE (R, G, B)	COLOR
0 - 14	0, .2, .5	Blue
- 28	0, .3, .6	
- 42	0, .4, .7	
- 57	0, .5, .8	
- 71	0, .6, .9	
- 85	0, .7, 1	
- 100	0, .5, 0	Green
- 114	0, .6, 0	
- 128	0, .7, 0	
- 142	0, .8, 0	
- 156	0, .9, 0	
- 171	0, 1, 0	
- 185	.5, 0, 0	Red
- 199	.6, 0, 0	
- 213	.7, .1, 0	
- 228	.8, .2, 0	
- 242	.9, .3, 0	
- 256	1, .4, 0	

Our other color code (Color-12) is depicted in Figure 16. For this code, we divided the 256 grey levels into 12 nearly equal intervals and applied the pseudocolors shown in Table 12. The color names given in the table are our descriptions of their appearance. Luminance ratio was measured with a tungsten source and varied somewhat depending upon the illuminant. The UCS chromaticity coordinates (tungsten illumination) of the Color-12 stimuli are shown in Figure 17. Note that, except for the red point, most colors are spaced approximately the same distance from their nearest neighbors. We avoided colors in the deep blue because the density was too great and the range of luminance ratios became excessive. It was our intent that this color code produce imagery where identification would depend upon hue rather than brightness.

9.1.2 Subjects

The five subjects used in this study were all experienced photo-interpreters and they had radar interpretation experience. The subjects' image interpretation experience ranged from 3 to 25 years. All subjects were SAI employees. Their color vision was screened with Dvorine color plates (Dvorine, 1953) and was normal.

9.1.3 Testing Procedure

Our intent was to duplicate the usual detailed reconnaissance image interpretation setting as closely as possible. We furnished subjects with identical light tables (Knox 5000° K) and magnifiers. Photo-reproductions were made of the original SAR hardcopies and each of the 10 images was delineated on these working copies. These reproductions were of poor quality and lacked detail; their sole purpose was to allow the image interpreters (IIs) to orient the film imagery. Each II had his own working copy of the imagery and his own set of 40 test slides. Slide order was random for each subject and IIs were urged to rate each slide on its own merits.

Table 12. Color Codes for 12-Color SAR Coding Strategy

COLOR	RGB BIT VALUE (Normalized to 1.0)	CHROMATICITY u	CHROMATICITY v	LUMINANCE RATIO
PERSIAN BLUE	0, .6, 1.	.1121	.2651	1.1
DESATURATED BLUISH GRAY	.6, .6, 1.	.1906	.3156	1.6
DESATURATED GREEN	.8, .8, 1.	.2106	.3443	3.8
TURQUOISE	0, 1., 1.	.1377	.3222	3.1
HOLLY GREEN	0, .6, .6	.0928	.3453	1.0
GREEN	0, .8, 0	.1276	.3795	1.4
LIME	.6, 1., 0	.1961	.3749	2.6
PURPLE	.8, 0, 1	.4390	.2941	1.1
PINK	1., .6, 1.	.2651	.3409	3.0
YELLOW	1., 8, 0	.2575	.3719	3.8
AMBER	1., .6, 0	.2942	.3690	2.3
RED	1., .4, 0	.3465	.3681	1.1

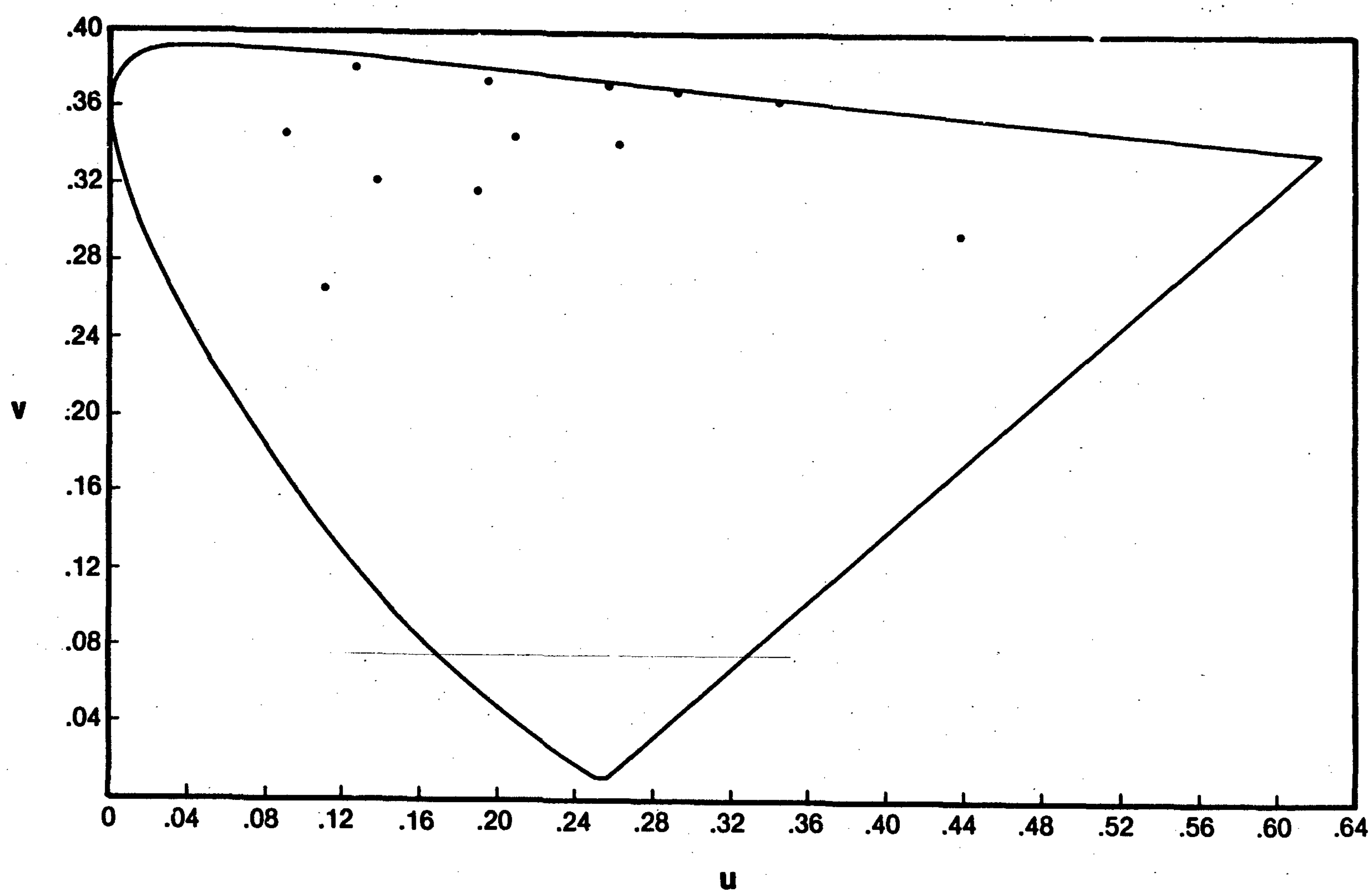


Figure 17. The UCS Chromaticity Coordinates Used in the SAR Color-12 False Color Imagery

Image ratings were accomplished in two ways. For Task 1, we took examples of the National Radar Interpretability Scale (NRIS), a rating scale used to assess synthetic aperature radar system image quality, and rewrote the object descriptors so that they would apply to our imagery. This was necessary because our imagery contained large, cultural features, rather than the detailed military targets for which the scale was originally developed. Tables 13 and 14 give the descriptions of the generic cultural features and the rating scale. Note that our rating scale uses the same detect, classify, and identify scheme that was developed for NRIS and the ratings run from 1 to 9.

Our subjects were briefed on the use of our rating scale and were told that our chief interest was in comparing across the imagery coding types. They were urged to rate the imagery by reference to our scale without making comparisons to imagery or other scales they may have used. They were also told to rate each slide on its own merits, not on what they may have remembered about that image from a previous slide. In addition to the interpretability scale ratings, subjects indicated their confidence in their ratings. They assigned a value from 1 to 3 to indicate "possible, probable, or definite."

Task 2 required that subjects compare the four codes for each of the 10 images and rank-order them in terms of their preference for working with the imagery having that code. They viewed the four slides simultaneously and recorded their rankings. Then, they rated each of the four coding types on a 1-9 scale for preference in terms of ease of image interpretation.

9.2 RESULTS

9.2.1 Task 1: Interpretability/Image Quality Judgement

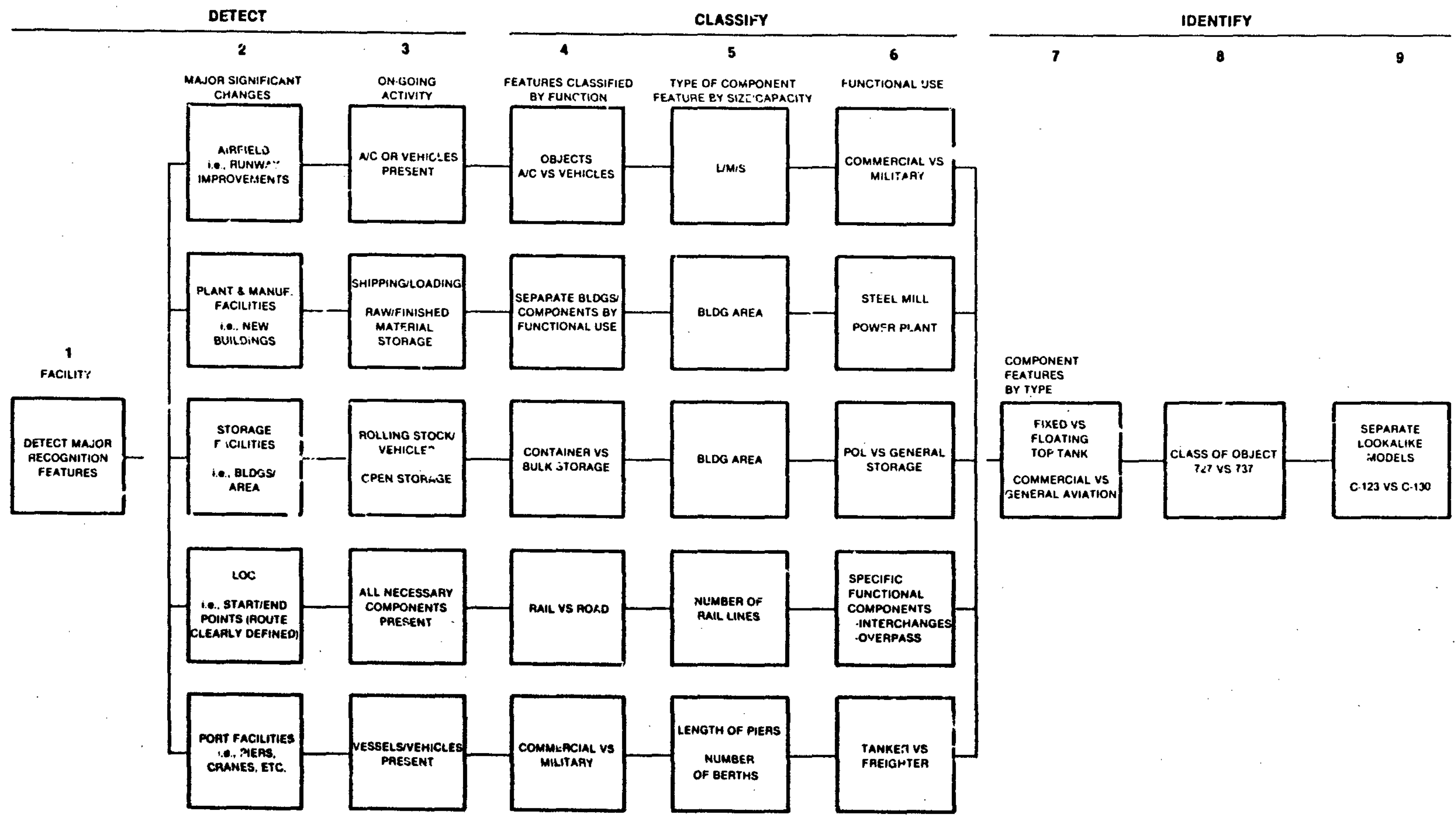
Histograms showing the interpretability scale ratings for our four coding strategies are presented in Figure 18. The frequency plot shows that both black and white codes were approximately equivalent and produced

Table 13. Cultural Features Rated in the SAR Image Coding Study

STORAGE & WAREHOUSING	PLANT & MFG. FACILITIES	AIRFIELD	LINES OF COMMUNICATION	PORT FACILITIES

• Petroleum
• Gas
• Steel
• Scrap
• General Open Storage
	• Power Plant
	• Steel Foundry
	• Cement
	• Asphalt
	• Sewage Treatment
		• Military
		• Commercial
			• Power Line
			• Pipeline
			• Rail
			• Roads
			• Bridges
				• Military
				• Commercial
				  • General Cargo
				  • POL Transshipment
				• Small Boat Anchorage

Table 14. SAR Imagery Rating Scale



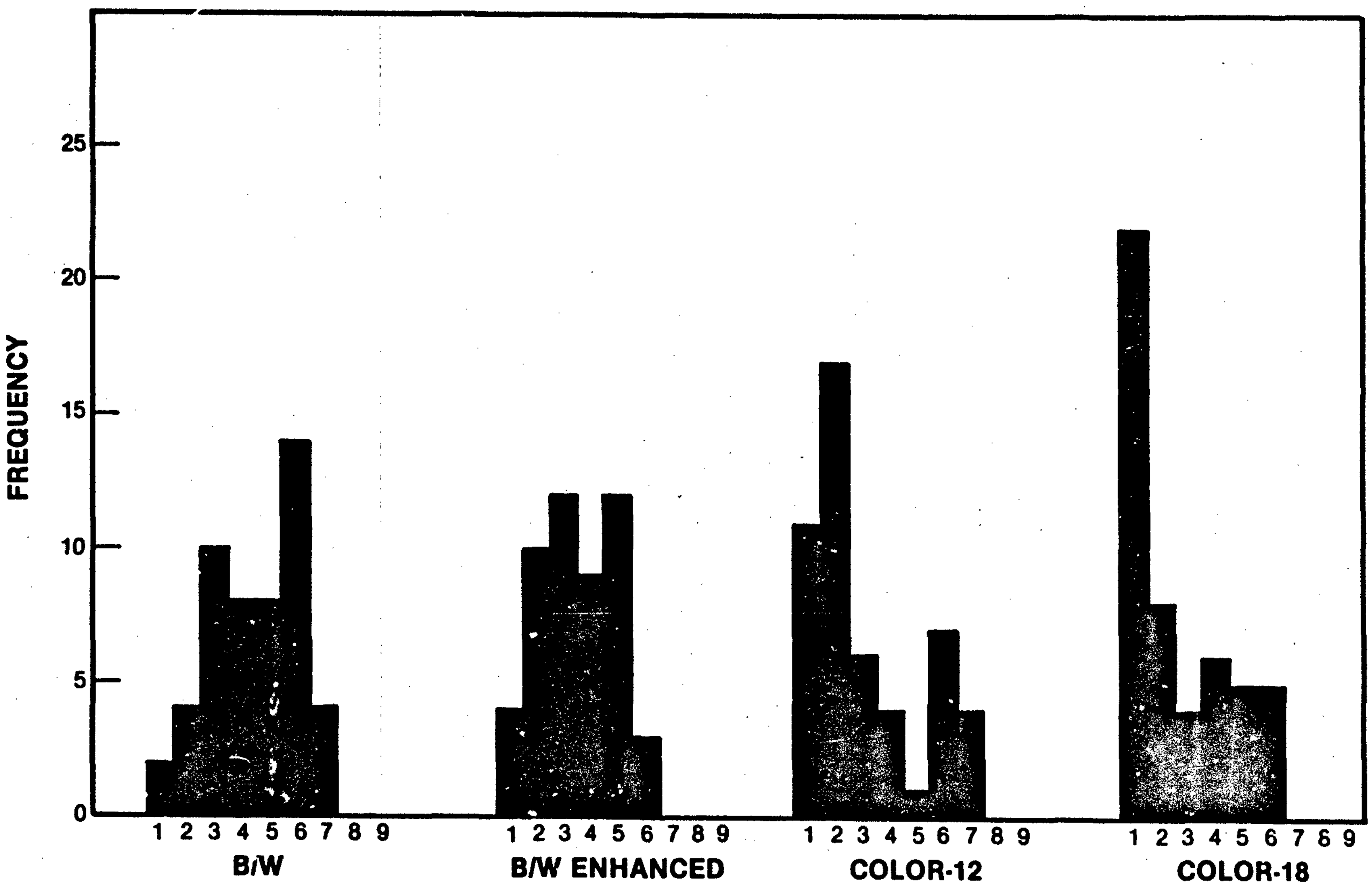


Figure 18. Interpretability Scale Ratings for Four SAR Coding Strategies

higher ratings than the color codes. Color-12 seemed to result in higher ratings than Color-18. A Friedman nonparametric analysis of variance showed a significant difference among the four coding conditions ($p<0.001$). Individual Wilcoxon rank sum tests were performed on the conditions of interest. There was no difference in ratings ($p>0.46$) on the two black and white codes. A Wilcoxon rank sum test on the two color codes showed that Color-12 was significantly better than Color-18 ($p<0.002$). Overall, black and white tested superior to the color codes ($p<0.00003$).

Histograms of Task 1 confidence ratings are plotted in Figure 19. There does not appear to be much difference in confidence accorded the different coding types. Most IIs rated their judgements as "probable" or "definite."

We have also plotted our results on an image-by-image basis. Figure 20 shows median ratings on each image. On only two images, No. 2 -- an oil refinery and No. 6--a port facility, did color produce as high a rating as the black and white codes. In general, it appears that the data shown in Figure 18 are consistent across images.

9.2.2 Task 2: Rank-Order Judgement

The results on the Task 2 preference ratings were very clear. Preferences were in the order: black and white enhanced, black and white normal, Color-12, Color-18. Less than one-fourth of the rankings deviated from this pattern. Color-18 was always ranked third or fourth. The ratings for coding preference are summarized in Figure 21. The ratings show the same pattern as the rankings. Black and white images are rated higher than color in preference; Color-12 was slightly preferred over Color-18.

A comparison of the interpretability scale ratings of Task 1 and the preference ratings of Task 2 shows that these two measures are highly correlated. The relationship is detailed in Figure 22. A Spearman rank-order correlation was highly significant ($r=0.999$, $p<0.001$).

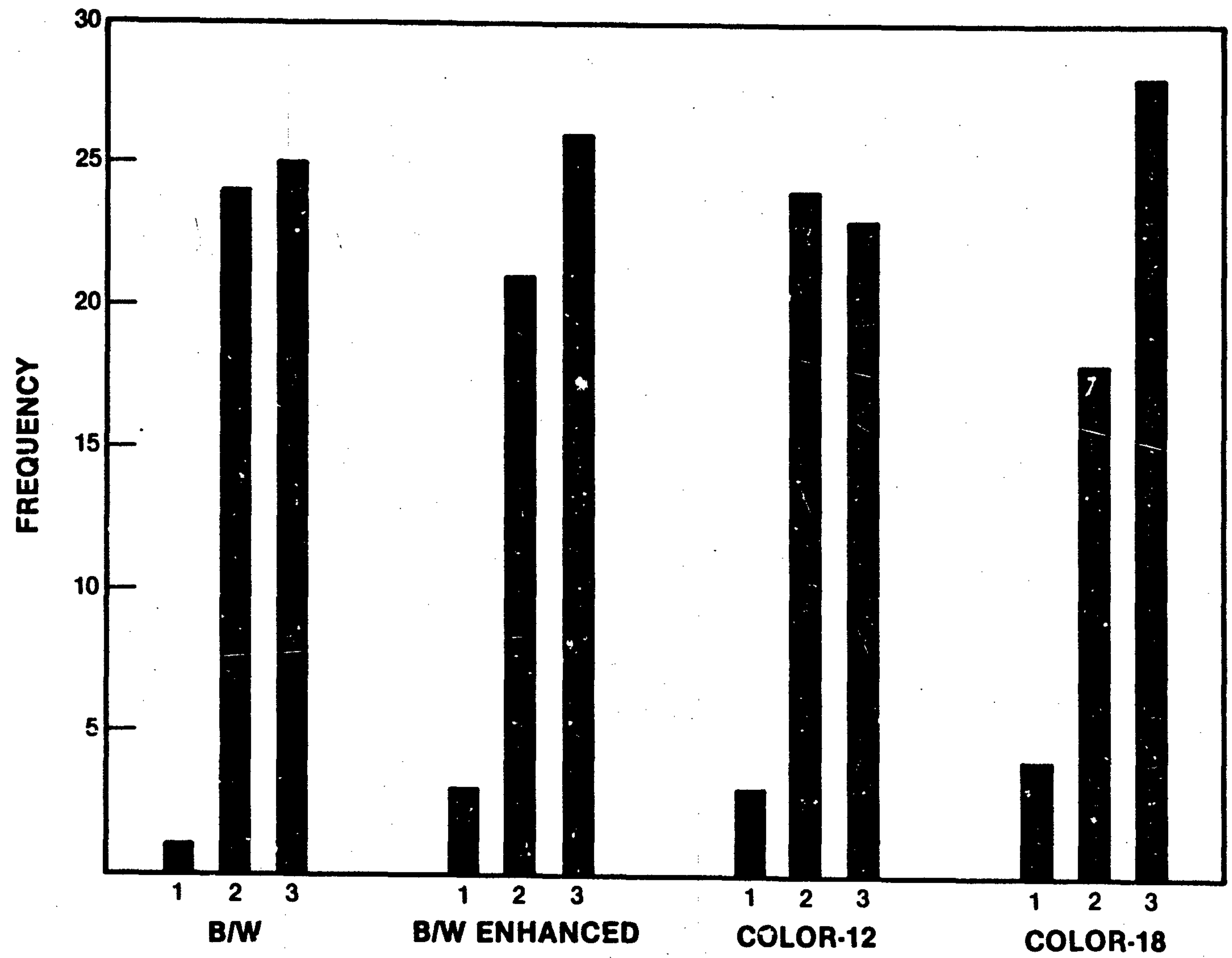


Figure 19. Confidence Ratings for Task 1

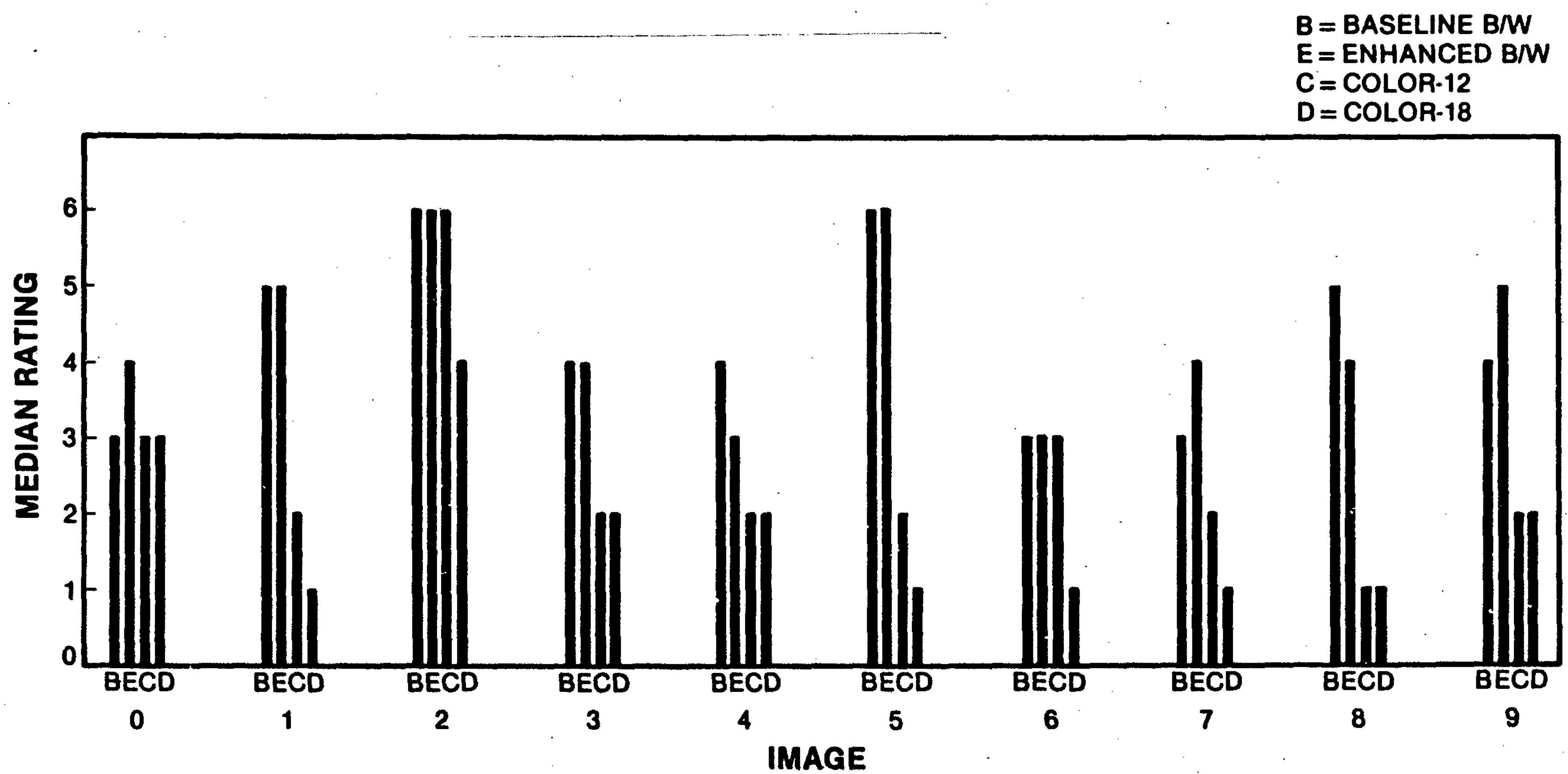


Figure 20. Median NRIS Ratings for Each SAR Image

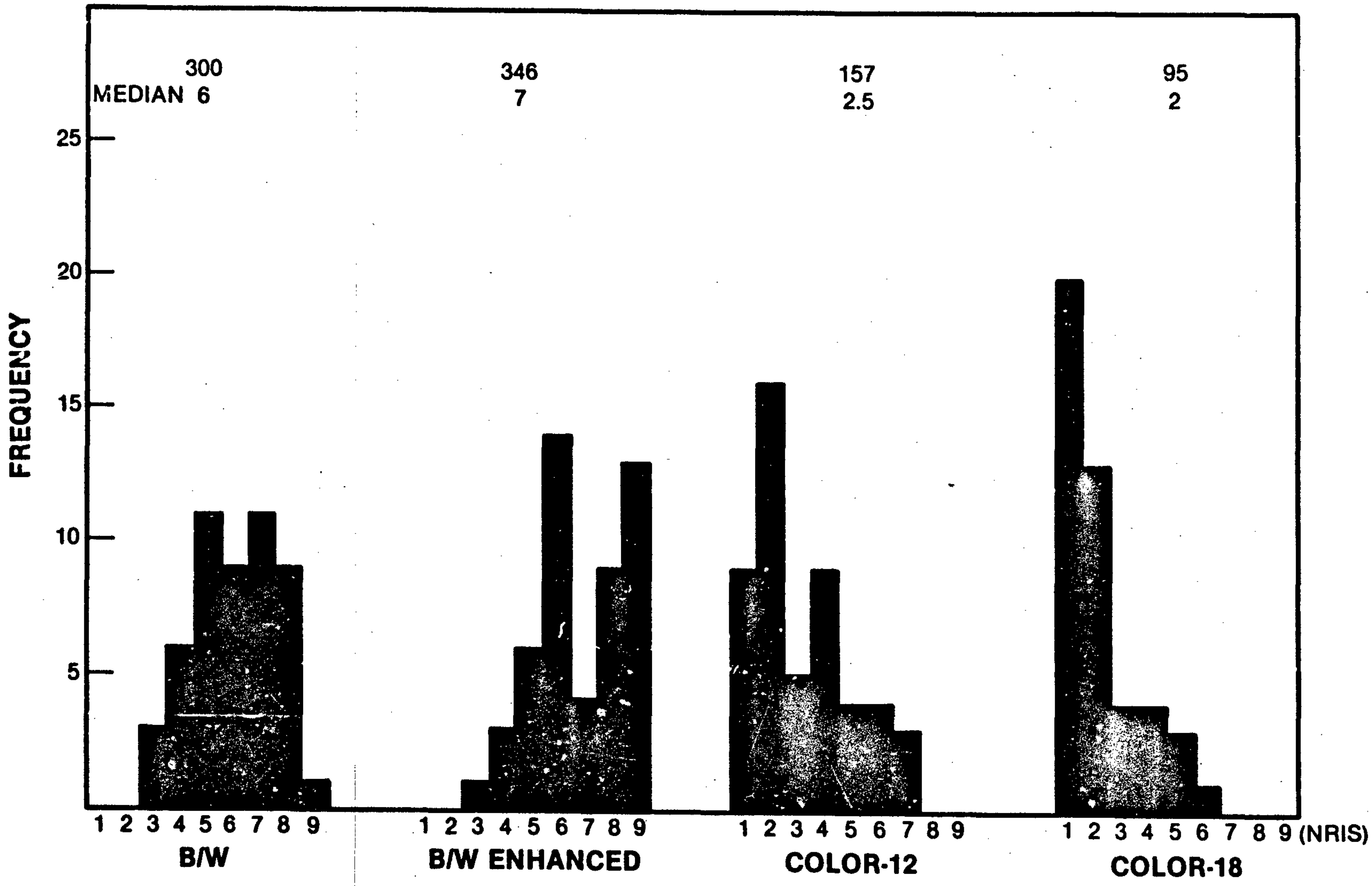


Figure 21. Preference Ratings for the Four SAR Coding Strategies

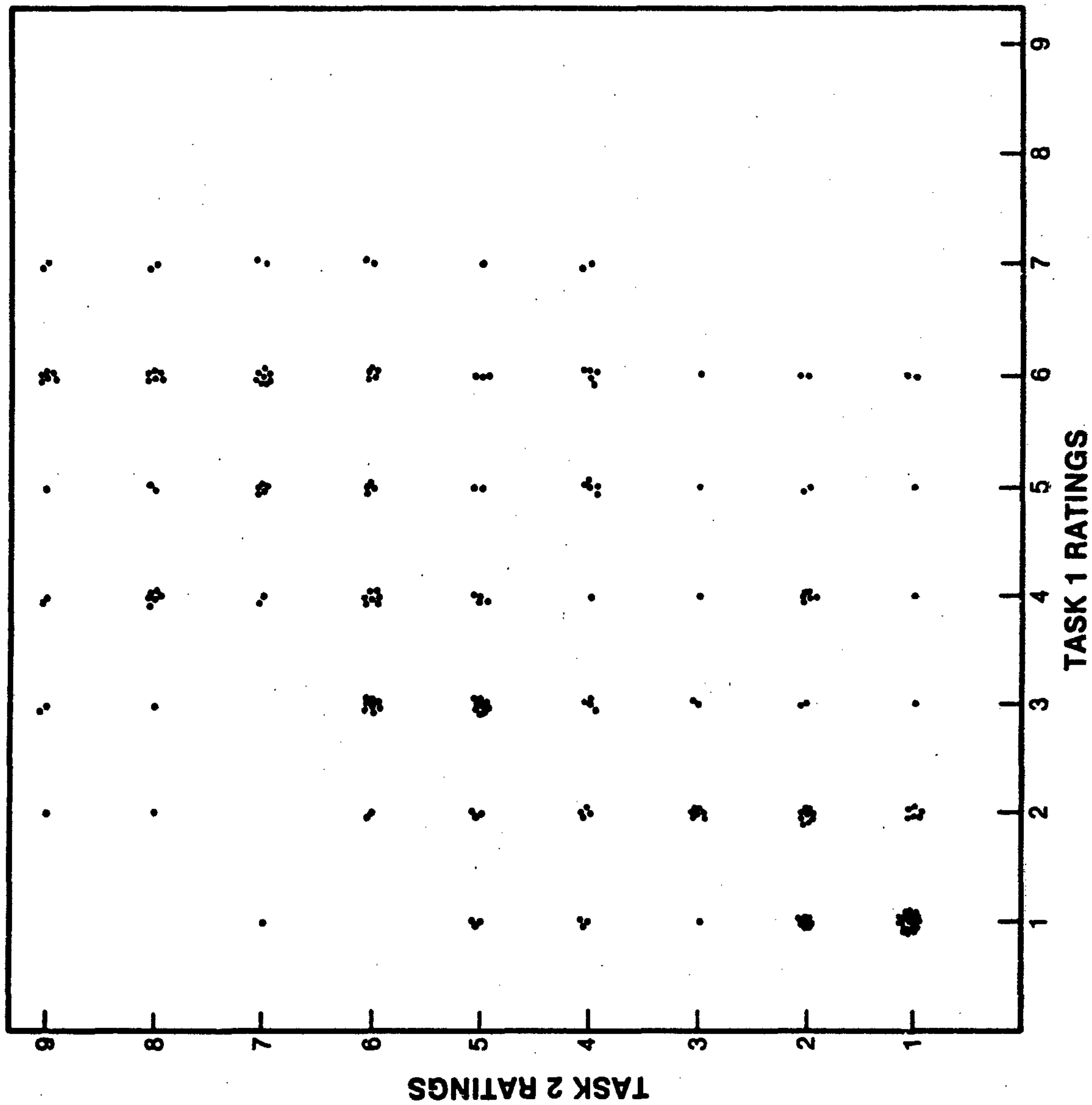


Figure 22. Comparison of Task 1 and Task 2 Ratings

9.3 CONCLUSIONS

The results of the image interpretability study are clear cut. Under our conditions, black and white coding is superior to color coding. Both radar interpretability scale ratings and subjective preference ratings agree that black and white coding is best. It appears that our subjects preferred to work with imagery that they ranked highest in terms of the opportunity to extract information. Informal comments indicated that color was sometimes useful for accentuating features, but that colors tended to "melt" into one another, making estimates of target ground size difficult.

The results of the SAR imagery coding study must be interpreted with reference to the imagery we employed. Our unclassified data base included only large cultural features. It is possible that imagery having better resolution and a wider dynamic range would afford more opportunity for effective color coding. This may be a topic for further work.

10.0 SUMMARY

The task objectives for this work involved color calibration of the CRT and film outputs of VIPER. After suitable color calibration, 35mm color slides were produced for use in a color-discrimination study. Applying the study's results, we then ran a color-coding study using SAR imagery.

The results of the color calibration effort showed that, with appropriate regression coefficients, we could predict colors output both on a CRT and on film. We developed a simplified methodology that can be applied when the needs for color calibration are not too stringent.

The color discrimination study showed that color discrimination is poorest for "red" colors. These data agree with our previous studies on color matching and discrimination and we recommend caution when using the 1960 UCS representation as the index of color difference.

The radar imagery coding study found, as have other investigators, that black and white coding is superior to color coding. If imagery is to be color coded, hue coding appears to be better than hue-brightness coding.

11.0 REFERENCES

Barker, E. and M. Krebs, April 1977, Color Coding Effects on Human Performance, An Annotated Bibliography, Contract No. ONR-CR213-136-1F, Office of Naval Research Code 212, Arlington, VA.

Bartleson, C.J., 1968, "Color Perception and Color Television," Journal of the Society of Motion Picture and Television Engineers, Vol. 77, pp. 1-11.

Bartleson, C.J., 1979, "Predicting Corresponding Colors With Changes in Adaptation," Color Research and Application, Vol. 4, pp. 143-55.

Booth, J.M., and J.B. Schroeder, 1977, "Design Considerations for Digital Image Processing Systems," Computer, Vol. 10, pp. 15-20.

Boynton, R.M., 1979, Human Color Vision, Holt, Rinehart, and Winston.

Boynton, R.M. and N. Kambe, 1980, "Chromatic Difference Steps of Moderate Size Measured Along Theoretically Critical Axes," Color Research and Application, Vol. 5, pp. 13-23.

Boynton, R.M. and J.J. Wisowaty, 1980, "Equations for Chromatic Discrimination Models," Journal of the Optical Society of America, Vol. 70, pp. 1471-76.

Briggs, S.J., 1981, "Photometric Technique for Driving a 'Best Gamma' for Displays," Optical Engineering, Vol. 20, pp. 651-57.

Brown, W.R.J., and D.L. MacAdam, 1949, "Visual Sensitivities to Combined Chromaticity and Luminance Differences," Journal of the Optical Society of America, Vol. 39, pp. 808-34.

Brown, W.R.J., 1951, "The Influence of Luminance Level on Visual Sensitivity to Color Differences," Journal of the Optical Society of America, Vol. 41, pp. 684-88.

Brown, W.R.J., 1957, "Color Discrimination of Twelve Observers," Journal of the Optical Society of America, Vol. 47, pp. 137-43.

Burns, S.A., V.C. Smith, J. Pokorny, and A.E. Elsner, 1982, "Brightness of Equal-Luminance Lights," Journal of the Optical Society of America, Vol. 72, pp. 1225-31.

Brun, J. and A. Martin, April 28, 1980, "Comparative Evaluation of High-Resolution Color CRTs," paper presented at the S.I.D. Symposium, San Diego.

Carter, R.C., 1982, "A Design Tool for Color Displays," Prcceedings of the Human Factors Society, 26th Annual Meeting, p. 589.

Catmull, E., 1979, "A Tutorial on Compensation Tables," SIG. GRAPH Vol. 79, Proceedings, pp. 1-7.

Christ, R.E., 1975, "Review and Analysis of Color Coding Research for Visual Displays," Human Factors, Vol. 17, pp. 542-70.

Christ, R.E., 1976, "Analysis of Color and its Effectiveness in Display Devices," Proceedings of the Third Advanced Aircrew Display Symposium, Patuxent River, MD, pp. 380-92.

CIE Colorimetry Committee, 1974, "Working Program on Color Differences, Technical Note," Journal of the Optical Society of America, Vol. 63, pp. 896-97.

Clapper, F.R., R.G. Gendron, and S.A. Brownstein, 1973, "Color Gamuts of Additive and Subtractive Color-Reproduction Systems," Journal of the Optical Society of America, Vol. 63, pp. 625-29.

Cleveland, W.S. and R. McGill, 1983, "A Color-Caused Optical Illusion on a Statistical Graph," The American Statistician, Vol. 37, pp. 101-05.

Committee on Colorimetry, 1963, The Science of Color, Optical Society of America, Washington, DC.

Cook, T.C., 1974, Color-Coding--A Review of the Literature, Technical Note 9-74, U.S. Army Human Engineering Laboratory, Aberdeen Proving Ground.

Cowan, W.B., 12 April 1982, "The Specification of Hetrochromatic Brightness," Paper presented at the Inter-Society Color Council Symposium on Colorimetry of Self-Luminous Displays, Charlotte, NC.

DeMars, S.A., January 1975, Human Factors Considerations for the Use of Color in Display Systems, NASA Report TM-X-72196, Systems Engineering Division, Directorate of Design Engineering, John F. Kennedy Space Center.

Dragavon, E.J., M.L. Hershberger, and P.J. Whitt, June 1977, Synthetic Aperture Radar Image Processing Techniques Development and Evaluation, Technical Report AFAL-TR-77-113, Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio.

Dvorine, I., 1953, Dvorine Pseudo-Isoduomatic Plates, Harcourt, Brace and World, New York.

Farley, W.W. and J.C. Gutmann, 1980, Digital Image Processing Systems and an Approach to the Display of Colors of Specified Chrominance, Technical Report HFL-80-2, Human Factors Laboratory, Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University.

Farrell, R.J., and J.M. Booth, 1975, Design Handbook for Imagery Interpretation Equipment, Boeing Aerospace Company, Seattle, Washington.

Fienberg, S.E., 1979 "Graphical Methods in Statistics," The American Statistician, Vol. 33, pp. 165-78.

Galves, J-P., and J. Brun, 1975, "Colour and Brightness Requirements for Cockpit Displays: Proposal to Evaluate Their Characteristics," Paper presented at 29th AGARD Avionics Panel Technical Meeting on Electronic Displays.

Gonzalez, R.C., and P. Wintz, 1977, Digital Image Processing, Addison-Wesley, Reading, MA.

Graham, C.G. (ed), 1965, Vision and Visual Perception, John Wiley and Sons, New York.

Greenberg, D., A. Marcus, A. Schmidt, and V. Gortney, 1982, The Computer Image: Applications of Computer Graphics, Addison-Wesley, Reading, MA.

Guild, J., 1931, "The Colorimetric Properties of the Spectrum," Philosophical Transactions of the Royal Society, London, Serial A, Vol. 230, p. 149.

Halsey, R.M., and A. Chapanis, 1951, "On the Number of Absolutely Identifiable Spectral Hues," Journal of the Optical Society of America, Vol. 41, pp. 1057-58.

Harney, R.C., J.S. Martin, and D.R. Sullivan, 1978, Quasi-Three-Dimensional Display of Infrared Radar Images Using Range/Intensity Color Mapping, MIT Technical Note 1978-44, Contract No. F19628-78-C-0002.

Herold, E.W., 1974, "History and Development of the Color Picture Tube," Proceedings of the S.I.D., Vol. 15, pp. 141-49.

Huddleston, J.H.F., May 1983, Color Coding for Computer-Driven Displays, Report 5027, Draft Issue 1, Farranti Computer Systems [an annotated bibliography].

Judd, D.B., 1951, "Report of U.S. Secretariat Committee on Colorimetry and Artificial Daylight," CIE Proceedings, Vol. I, Part 7, p. 11.

Judd, D.B., and G. Wyszecki, 1975, Color in Business, Science and Industry, John Wiley and Sons, New York.

Kinney, J.S., and J.C. Cooper, July 1967, Adaptation to a Homochromatic Visual World, Technical Report No. 497, U.S. Naval Submarine Medical Center, Submarine Base, Groton, CT (AD 659496).

Kinney, J.S., 12 April 1982, "Brightness of Colored Self-Luminous Displays in the C.I.E.," paper presented at the Inter-Society Color Council Symposium on Colorimetry of Self-Luminous Displays, Charlotte, NC.

Krebs, M.J., J.D. Wolf, and J.H. Sandvig, 1978, Color Display Design Guide, Report ONR-CR213-136-2F to Office of Naval Research, Honeywell Systems and Research Center, Minneapolis, MN.

Lamonica, G.L., December 1977, The Effect of Pseudo-Color on the Apparent Displayed Dynamic Ranges of Radar Imagery, Technical Report AFAL-TR-77-186, Air Force Avionics Laboratory, Wright-Patterson Air Force Base, Ohio.

Laycock, J., and J.P. Viveash, May 1981, Calculations for Determining the Perceptibility of Monochrome and Colour Displays Viewed Under Various Illumination Conditions, Technical Report 81067, Royal Aircraft Establishment, Farnborough, Hants, England.

Laycock, J., November 1982, Recommended Colours for Use on Airborne Displays, Technical Report 82110, Royal Aircraft Establishment, Farnborough, Hants, England.

Lehmann, E.L., 1975, Nonparametrics, McGraw-Hill, New York.

MacAdam, D.L., 1942, "Visual Sensitivities to Color Differences in Daylight," Journal of the Optical Society of America, Vol. 32, pp. 247-74.

MacAdam, D.L., 1943, "Specification of Small Chromaticity Differences," Journal of the Optical Society of America, Vol. 33, pp. 18-26.

MacAdam, D.L., 1959, "Small-Field Chromaticity Discrimination," Journal of the Optical Society of America, Vol. 49, pp. 1143-46.

MacAdam, D.L., 1963, "Chromatic Adaptation, II: Nonlinear Hypothesis," Journal of the Optical Society of America, Vol. 53, pp. 1441-45.

Martin, A., June 1977, "The CRT/Observer Interface," Electro-Optical System Design.

Maxwell, J.C., 1860, "On the Theory of Compound Colours and the Relations of the Colours of the Spectrum," Proceedings of the Royal Society (London), Vol. 10, pp. 404-07.

Post, D.L., E.B. Costanza, and T.M. Lippert, 1982, "Expressions of Color Contrast as Equivalent Achromatic Contrast," Proceedings of the Human Factors Society, 26th Annual Meeting, pp. 581-85.

Post, D.L., 1983, personal communication.

Purdy, D. McL., 1931, "Spectral Hue as a Function of Intensity," Journal of Psychology, Vol. 43, pp. 541-59.

Robertson, A.R., 1977, "The CIE 1976 Color-Difference Formulae," Color Research and Application, Vol. 2, pp. 7-11.

Robertson, A.R., 1983, "Critical Review of Definitions of Metamerism," Color Research and Application, Vol. 8, pp. 189-91.

Savoie, R.F., 1973, "Bezold-Brucke Effect and Visual Nonlinearity," Journal of the Optical Society of America, Vol. 63, pp. 1253-61.

Shurtleff, D.A., 1980, How to Make Displays Legible, Human Interface Design, La Miracle, CA.

Siegel, M.H., 1969, "Color Discrimination and Luminance," Perception of Psychophysics, Vol. 6, pp. 163-65.

Siegel, S., 1956, Nonparametric Statistics for the Behavorial Sciences, McGraw-Hill, New York.

Silberstein, L., 1943, "Investigations on the Intrinsic Properties of the Color Domain, II," Journal of the Optical Society of America, Vol. 33, pp. 1-10.

Silverstein, L.D., and R.M. Merrifield, 15-16 September 1981, "Color Selection and Verification Testing for Airborne Color CRT Displays," paper presented at the Fifth Advanced Aircrew Display Symposium, Naval Air Test Center, Patuxent River, MD.

Smith, D.L., December 1978, "Graphic Displays--A Human Engineering Guide for Using Color in CRT Command and Control Displays, Lockheed Missiles and Space Co., Space Systems Division, Sunnyvale, CA.

Stiles, W.S., 1955, "The Basic Data of Colour-Matching," Physical Society Yearbook, pp. 44-67.

Stollings, M.N., 1982, "Information Processing Load of Graphics Versus Alphanumeric Weapon Format Displays for Advanced Fighter Cockpits," Masters Thesis, Wright State University.

Tedford, W.H., S.L. Bergquist, and W.E. Flynn, 1977, "The Size-Color Illusion," Journal of General Psychology, Vol. 97, pp. 145-49.

Teichner, W.H., R.E. Christ, and G.H. Corso, June 1977, Color Research for Visual Displays, Report No. ONR CR213-102-4F, Office of Naval Research (AD A043609).

Trumbo, B.E., 1981, "A Theory for Coloring Divoirate Statistical Maps," The American Statistician, Vol. 35, pp. 220-26.

Tufte, E.R., 1983, The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT.

Tullis, T.S., 1981, "An Evaluation of Alphanumeric, Graphic, and Color Information Displays," Human Factors, Vol. 23, pp. 541-50.

Van Cott, H.P., and R.G. Kinkade (eds), 1972, Human Engineering Guide to Equipment Design, Tamis Books, Inc., Los Angeles, CA

Vienot, F., 1980, "Relations Between Inter- and Intra-Individual Variability of Color Matching Functions. Experimental Results," Journal of the Optical Society of America, Vol. 70, pp. 1476-83.

Wallis, R.H., 1975, Film Recording of Digital Color Images, Report 570, University of Southern California Imaging Processing Institute.

Wagner, D.W., 1977, Experiments With Color Coding on Television, Technical Report No. 5952, Naval Weapons Center.

Wagner, G., and R.M. Boynton, 1972, "Comparison of Four Methods of Heterochromatic Photometry," Journal USA, Vol. 62, pp. 1508-15.

Ward, F.E., F. Greene, and W. Martin, 25-28 April 1983a, "Color Discrimination as a Function of Saturation, Field Size, and Adaptation Level," paper presented at the Second Symposium on Aviation Psychology, Columbus, Ohio.

Ward, F.E., F. Greene, and W. Martin, 1983b, "Color Matches and Discrimination Functions for Selected Hues," Proceedings of the Human Factors Society, 27th Annual Meeting.

Woodcock, S., and J.D. Leyland, July 1979, "The Choice of Phosphor for Modern CRT Display Applications," Displays, pp. 69-82.

Wright, W.D., 1928-29, "A Re-determination of the Mixture Curves of the Spectrum," Transactions of the Optical Society, London, Vol. 30, pp. 141-164.

Wright, W.D., 1936, "The Breakdown of a Color Match With High Intensities of Adaptation," Journal of Physiology, Vol. 87, pp. 23-33.

Wright, W.D., 1969, The Measurement of Color, Van Nostrand Reinhold Company, New York.

Wyszecki, G., and W.S. Stiles, 1982, Color Science (2nd Edition), John Wiley and Sons, New York.

END

☆U.S.Government Printing Office: 1984 — 759-006/223