V. Summary

Motivated by the need to develop an instructional system that can help humans learn associations efficiently, we considered the random-trial increments model and for a quite general formulation of the instructional objective, we derived optimal instructional policies which are very simple to implement. Key to the derivation of optimal instructional policies was the representation of the instructional problem as a decomposable Markov decision process. We also showed how this representation can be used to investigate the optimality of the instructional policies derived from the random-trial increments learning model for more complex learning models.

Acknowledgment

The author wishes to thank S. Engelbrecht and D. Fisher for their helpful comments.

References

[1] M. F. Norman, "Incremental learning on random trials," *J. Math. Psych.*, vol. 1, pp. 336–350, 1964.
[2] R. C. Atkinson and J. A. Paulson, "An approach to the psychology of instruction," *Psycol. Bull.*, vol. 78, pp. 49–61, 1972.
[3] R. R. Bush and F. Mosteller, *Stochastic Models for Learning*. New York: Wiley, 1955.
[4] G. H. Bower, "Application of a model to paired associate learning," *Psychometrika*, vol. 26, pp. 255–280, 1961.
[5] W. Karush and R. E. Dear, "Optimal stimulus presentation strategy for a stimulus sampling model of learning," *J. Math. Psych.*, vol. 3, pp. 19–47, 1966.
[6] R. D. Smallwood, "The analysis of economic teaching strategies for a simple learning model," *J. Math. Psych.*, vol. 8, pp. 285–301, 1971.
[7] R. Bellman, *Dynamic Programming*. Princeton, NJ: Princeton Univ. Press, 1957.
[8] R. A. Howard, *Dynamic Programming and Markov Processes*. Cambridge, MA: MIT Press, 1960.
[9] N. Meulau, M. Hauskrecht, K. Kim, L. Peshkin, L. P. Kaebling, T. Dean, and C. Boutilier, "Solving very large weakly coupled markov decision processes," presented at the Proc. 15th National Conf. Artificial Intelligence, 1998.
[10] K. V. Katsikopoulos, "Characterizing and Optimizing the Performance of Younger and Older Adults in Paired Associate Tasks: A Markov Modeling Approach," Unpublished doctoral dissertation, Univ. Massachusetts, Cambridge, MA, 1999.
[11] T. D. Wickens, *Models for Behavior: Stochastic Processes in Psychology*. San Fransisco, CA: Freeman, 1982.
[12] A. R. Cassandra, L. P. Kaebling, and M. L. Littman, "Algorithms for solving partially observable Markov decision processes," Brown Univ., Providence, RI, Tech. Rep. 94-14, 1994.

Color Blindness and a Color Human Visual System Model

Curtis E. Martin, John G. Keller, Steven K. Rogers, and Matthew Kabrisky

***Abstract*—A physiologically motivated human color visual system model which represents visual information with one brightness component (A) and two chromatic components (C_1 and C_2) is used to create stimuli for testing the color perception of deuteranomalous trichromats. Two experiments are performed. Using simple ramp patterns, the first experiment finds that three deuteranomalous trichromat test subjects can distinguish variations only in the C_2 component of the color vison model. This finding is further tested in the second experiment, a set of paired comparison preference tests. Two altered versions of each of three natural color images are prepared by setting either one of the color components to a constant over the full image. Pairs of an original and a distorted image are presented to the test subjects, and they are asked to indicate which image they prefer. Constant C_1 images are frequently preferred over originals, while constant C_2 images are described as appearing monochrome to the deuteranomalous trichromats. These experimental results indicate that the C_1 channel is severely attenuated in the deuteranomalous trichromat test subjects, and that nearly all their color sensation is mediated by the C_2 channel of the color vision model.**

I. Introduction

Recent experiments with deuteranomalous trichromat observers and a physiologically-motivated model of the human color vision system (color HVS) have produced interesting results. Using the color HVS model, RGB images are transformed into the perceptual space of the model, where they are manipulated by setting one of the two chromatic components of the model equal to a constant (the third component, brightness, is never changed). The resulting images are transformed back to the original RGB primary space for observation on a color monitor. Setting one of the two chromatic components in the perceptual space equal to a constant produces images that deuteranomalous trichromats cannot distinguish from the original full-color images, despite clear differences apparent to normal trichromats. Setting the other chromatic channel to a constant produces images that the deuteranomalous trichromats perceive as monochromatic, but normal trichromats report as having a range of different colors. These results provide insight into the operation of the color HVS and support the use of the model in image processing applications. This paper describes the experiments more fully and discusses the implications of their results.

Section II provides a brief review of normal and defective human color vision and the color HVS model used in the experiments. Section III describes the experiments performed, Section IV presents the results obtained, and Section V discusses the implications of these results.

II. Background

General agreement has been reached on a number of basic facts about the physiology and psychophysics of the color HVS. Results used in the development of the color HVS model are outlined in the first part of this section, and the model itself is described in the second part. The section

Manuscript received November 19 1998. This work was supported in part by the Defense Dissemination Program Office. This paper was recommended by Associate Editor M. S. Obaidat.

C. E. Martin and J. G. Keller are with the Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433 USA.

S. K. Rogers and M. Kabrisky are with Qualia Computing, Inc., Beavercreek, OH 45431 USA.

Publisher Item Identifier S 1083-4427(00)01181-4.

concludes with a brief overview of color blindness which motivates the experiments described in Section III.

A. Human Color Vision

The most obvious starting place for consideration of the human visual system is the eye itself. At the back of the eyeball, the visual signal is sensed by a closely packed array of photo-receptive neurons (rods and cones) in the retina [1]. Rods are principally involved in vision at low (scotopic) light levels and do not convey color information, while cones respond at higher (photopic) light levels and provide the means for color vision. Three types of cone receptors are identified on the basis of their spectral absorption characteristics. The three absorption curves have approximately the same shape, but peak at three different wavelengths: 445, 540, and 570 nm. These three types of cones are identified here as S cones (short wavelength peak), M cones (medium wavelength peak), and L cones (long wavelength peak). Three types of cones provide a physiological basis for trichromatic theories of color vision—the representation of colors as an additive mixture of three suitably chosen primary colors with fixed spectral characteristics.

The physiological response of the cones has been found to be proportional to the logarithm of the stimulus intensity over a large range of intensities [2]–[4]. This physiological measurement has a correlate in psychophysical measurements, known as Fechner's law, which states that a just noticeable brightness difference is proportional to the logarithm of the stimulus intensity [5]. The signals produced by the cones are passed up through several layers of neurons in the retina, ending with ganglion cells, which send the signals to the lateral geniculate nucleus (LGN) via the optic nerve.

The LGN appears to perform a colorimetric transformation of the visual information from the retina. Six types of cells have been identified in the LGN [6]–[8]. The first two of these, called spectrally nonopponent cells, respond primarily to brightness information, and are denoted here as +Wh-Bl and +Bl-Wh cells. The activity of these cells increases (for +Wh-Bl cells) or decreases (for +Bl-Wh cells) as a weighted sum of the outputs of the L and M cones (and possibly S cones) increases. The other four types of LGN cells, called spectrally opponent cells, respond differentially to the color, rather than brightness differences. These cells are grouped in two systems: red-green, receiving inputs from L and M cones, and yellow-blue, most likely receiving inputs from L and S cones [9]. The cells in these systems are maximally excited (inhibited) by light of one wavelength, and maximally inhibited (excited) by light of another wavelength. +G-R cells are excited by 500 nm and inhibited by 630 nm, +Y-B cells are maximally excited by 600 nm and maximally inhibited by 440 nm, and the responses of +R-G and +B-Y cells are roughly mirror images of those of the +G-R and +Y-B cells. Thus, the LGN receives input from three cone types and produces three different output channels: one achromatic (brightness) channel which adds log outputs L, M, and S cones, and two different chromatic channels which subtract log outputs of L and M (red-green) or L and S (yellow-blue) cones.

B. The Faugeras Color HVS Model

The Faugeras color HVS model is based on both psychophysical and neurophysiological data, and models three stages of the HVS, as illustrated in Fig. 1. The first stage models the sensing process in the retina. This is accomplished by transforming the image from its original representation space to a cone pigment absorption space developed by Stiles [11], and then applying a logarithmic nonlinearity to the resulting three channels to model the nonlinear response of the cone receptors. For images already expressed in a tristimulus space (e.g., CIE XYZ or RGB), the transformation into the cone absorption space is performed by a linear transformation matrix U. In agreement with the

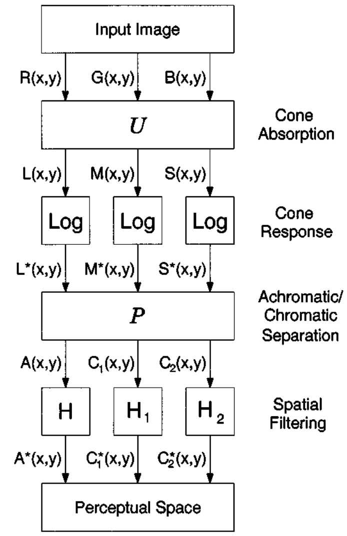

Fig. 1. Faugeras model of the human color vision system [10].

notation suggested earlier, the outputs of this first stage are designated as $L^*(x, y)$, $M^*(x, y)$, and $S^*(x, y)$, using the asterisk as a reminder that the logarithmic nonlinearity has been applied. There is some disagreement among researchers as to the true functional form for modeling the nonlinear response of the cones (see, for example, Cornsweet [4]). Faugeras uses the logarithmic nonlinearity to provide both computational simplicity and a homomorphic, automatic gain control characteristic to his model [10]. For a monitor with a D_{6500} white point and RGB primaries, the cone-space transformation matrix U is given by Faugeras as [10]

$$U = \begin{bmatrix} 0.3634 & 0.6102 & 0.0264 \\ 0.1246 & 0.8138 & 0.0616 \\ 0.0009 & 0.0602 & 0.9389 \end{bmatrix}. \tag{1}$$

The second stage of the Faugeras model accounts for the transformation which occurs in the LGN. This stage applies a linear colorimetric transformation $\mathcal{P}$ to the outputs of the cone absorption stage to produce an achromatic channel, $A(x, y)$ and two chromatic channels, $C_1(x, y)$ and $C_2(x, y)$. While the theoretical basis for this stage arises from physiological measurements of the LGN, the actual parameters in this transformation are derived from psychophysical brightness matching and color matching data [10]. The achromatic channel parameters are chosen to produce an approximation to $V(\lambda)$ (the photopic luminous efficiency function), while the chromatic channel parameters are obtained by finding a transformation which maps MacAdam ellipses (describing nonuniformities in human color matching [12]), onto circles of equal radius, producing an approximately uniform chromaticity space. This yields the transformation matrix [10]

$$\mathcal{P} = \begin{bmatrix} 13.8312 & 8.3394 & 0.4294 \\ 64 & -64 & 0 \\ 10 & 0 & -10 \end{bmatrix}. \tag{2}$$

The final stage of the Faugeras model accounts for spatial effects of the HVS using linear, space invariant filters. These spatial effects include both high and low frequency attenuation attributed to various mechanisms in the visual system. High frequency effects model the optics of the eye, the finite density of receptors in the retina, and spatial summation of neural activity along the visual path. Low frequency attenuation is generally attributed to lateral inhibition between receptors and neurons at various stages along the visual path, some possibly even at the level of the visual cortex. Although these effects actually occur

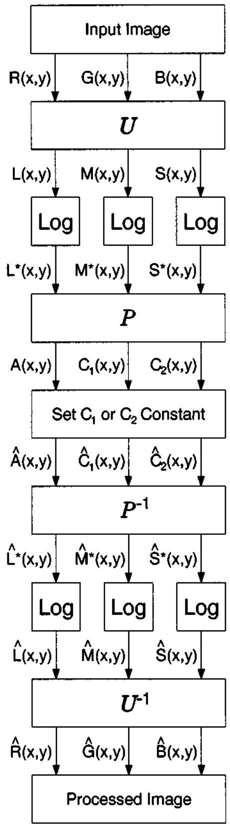

Fig. 2. Image processing in the perceptual space of the Faugeras color HVS model (adapted from [15]).

all along the visual path, Faugeras assumes that they may be successfully modeled by three linear, spatially invariant filters [10]. Because the present experiments are only concerned with color effects, these filters were omitted in the processing of test images.

HVS models provide the ability to manipulate images in a so-called perceptual space—a space which describes a sort of inner representation of what is seen—sometimes achieving great advantages. Stockham [13] pioneered the idea of measuring image quality based upon an HVS model, and Hall [14] obtained remarkable results compressing images using their perceptual space representation. Faugeras used his color HVS model to demonstrate the effects of several different manipulations of images in perceptual space [15]. The basic approach used by Faugeras, illustrated in Fig. 2, was used in preparing stimuli for this paper. By applying the HVS model, the input image is transformed into perceptual space, where the desired manipulations are performed. Then, the manipulated image is transformed back into "real world" space by passing it through the inverse of the HVS model.

C. Color Blindness

Something of a misnomer, the term "color blindness" refers to an inherited condition affecting certain individuals whose perception of color varies significantly from the majority. The purpose of this section is to summarize the generally accepted understanding of color blindness [1], [16], [17].

Hereditary color defects afflict around 8% of all males. About a quarter of these are dichromats, while the remainder (6% of all males) are referred to as anomalous trichromats. In dichromats, one of the cone types is almost completely unresponsive. These observers are able to match all spectral hues in color matching experiments using just two spectrally fixed primaries. There are three types of dichromats: protanopes, who typically confuse reds with greens, deuteranopes, who confuse purples and greens, and tritanopes, who only have sensation of red and green.

Anomalous trichromacy fills the gap between dichromacy and normal trichromacy. These observers require three spectrally fixed primaries to match all spectral hues in the color matching experiments, but their ability to distinguish colors varies significantly from that of normal trichromats, because the response of one set of cones is weak. In deuteranomalous trichromats, which are the most commonly occurring color blinds (roughly 5% of all males), the middle wavelength cones are the weak set. These observers characteristically confuse light shades of purple with light shades of green. Protanomalous trichromats (1% of males) have weak long wavelength cones and confuse pinks with light blue-greens. Tritanomalous trichromats, which are very rare, have weak short wavelength cone response, and confuse yellows with blues.

The majority of current research into color vision defects (dichromacy and anomalous trichromacy) involve physiological and psycho-physical studies relating the optical density of visual photopigments of the abberant cones to the type/severity of colorblindness (e.g., [18]–[20]). Color matching studies are used to investigate and quantify the perceptual qualities of color appearance ([21]–[23]), but though a perceptual color space is inferred, the actual studies generally focus on the tri-chromatic retinal outputs. Little work is being done in modeling these forms of color-blindness with respect to the bi-chromatic perceptual channels output from the human LGN, a lack the Faugeras vision model addresses.

III. Procedure

A. Subjects

Three male subjects, herein referred to as A, B, and C, participated in the experiment. All three were aware that they had color deficiencies, but these were not limiting to their engineering occupations. In conjunction with the experiments reported here, the three were given the 1944 Dvorine color perception test [24], which consists of a color naming portion and a color contrast portion. In the color naming portion, which uses a light and a dark tint of eight basic colors, all three subjects named more than half of the light tints incorrectly, and identified the dark violet tint as blue. In the color contrast portion of the test (the familiar "numbers" test), none of the three could distinguish between the color pairs of red/brown, green/orange, yellow/orange, green/yellow, or blue/violet; subject C had further difficulty distinguishing between red/gray, orange/brown, and gray/brown color pairs. Because the majority of color blinds are deuteranomalous trichromats [17], and all three subjects manifested largely the same difficulties, it is believed that the three subjects are all deuteranomalous trichromats.

B. Chromatic Variations

The current experiment was conducted as a follow-up to an incidental result in a related study of the color HVS. In an experiment to determine the existence of color Mach bands, a stimulus consisting of a Mach band type stimulus (a ramp between two plateaus) was constructed, with the only variation in the C_1 component [25]. On observing this stimulus, subjects A and B reported seeing a square of uniform color and brightness, while subject C indicated a change in the perceived brightness (but not of the color) of the stimulus. When the variation was switched to the C_2 component, all three the subjects reported different colors occurring in the stimulus. These results led to the conjecture that the perception of color for these subjects is mediated by only the C_2 channel. That is, these subjects seem to be unable to detect color differences when they are differences in C_1, but they do appear to be able to detect changes in C_2. So, if the C_1 component of an image is altered (by, for example, setting it equal to a constant),

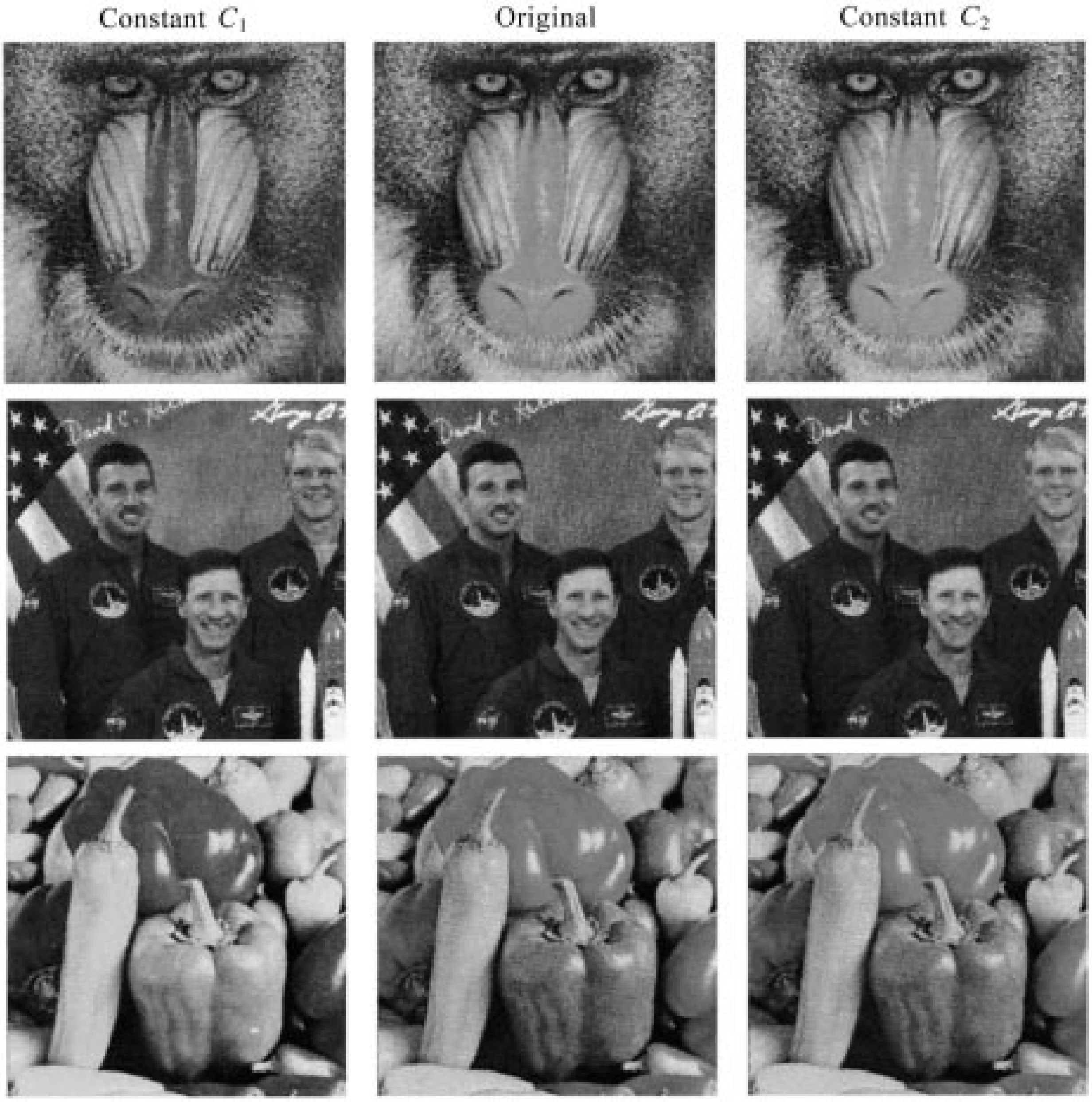

Fig. 3. Images used in the image preference experiments. Top: Mandrill, Center: Astronauts, Bottom: Peppers. The Mandrill and Peppers images are from the USC-SIPI image database and are used by permission (for a color version of this image, see www.qualia-computing.com/pubs/2000).

the subjects should be unable to distinguish the altered image from the original. Conversely, if the C_2 variation is removed from an image, the resulting image should appear to have no color variation whatever for these subjects. The image preference experiment was designed to test this proposition.

C. Image Preference

Three test images were chosen, and color-distorted versions of each of the three images were prepared using the approach described in the previous section and depicted in Fig. 2. To generate each distorted image, the original test image was transformed into (A, C_1, C_2) space using the HVS model, and either the C_1 or C_2 components were set to a constant. The inverse of the HVS model was then applied to the result to transform it back into the monitor's (R, G, B) primary space. Any values falling outside the closed interval $[0, 1]$ after this operation were assigned the nearest endpoint value (0 or 1).

Fig. 3 shows the images used throughout the study. While it is recognized that a perfect match between colors displayed on the monitor and the printed page is virtually impossible to achieve, this figure can still give the reader an idea of the distortions introduced by this process. Hard copies of the images were produced by converting them to color PostScript format and printing them on a Tektronix thermal wax color PostScript printer. When shown these hard copies, subject A found little difference between the displayed and the printed images—his color perceptions were largely the same for both media.

In each trial, both an original and a distorted version of an image were displayed side by side on a Sun SPARC-20 workstation 24-bit color monitor. The images occupied roughly 7° of the subject's field of view. The subjects were not informed of the color differences between the two images. In order to avoid any reference to color, each

TABLE I
SUMMARY OF IMAGE PREFERENCE EXPERIMENTS

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

subject was simply asked which image he preferred. After he indicated his preference, the subject was asked to describe the differences in the two images qualitatively, and to justify his choice.

IV. RESULTS

Table I summarizes the results of the image preference experiments. In the C_1 experiments, the subjects had a difficult time distinguishing between the original and the constant C_1 images, choosing the color-distorted image as identical to or better than the original in half of the trials. Particularly interesting was the consistent description of the nose in the constant C_1 MANDRILL picture as darker *red*, which normal trichromats characterize as dark *green*. All three subjects indicated that there was little perceptible difference between the original and the constant C_1 ASTRONAUT image. While they were clearly able to identify the distorted version of the PEPPERS image, the subjects still had difficulty naming the colors in it.

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

V. DISCUSSION

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

The results of the experiments with the constant C_2 images also largely agree with the predictions, particularly those involving the ASTRONAUT image. Without being solicited, the immediate reaction of subjects A and B to the constant C_2 image was that it looked monochrome, or black and white. In the constant C_2 MANDRILL image, neither subject A nor subject C was able to identify more than one colored area, while at least three different colors (pink on the nose, light green on the sides of the nose, and red in the eyes) are evident to normal trichromats.

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

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

It is remarkable to note that both the color contrast and image preference experiments seem to indicate that the color deficit (the "lost" or severely reduced axis) seems to be in (A, C_1, C_2) space, one step beyond the stage where the physiological measurements indicate it would be [the (L, M, S) cone absorption space]. Why this should happen is not immediately apparent.

VI. Summary

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

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

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

Significantly, the images used in the experiments were neither created nor manipulated in the portion of the model corresponding to cone outputs. Rather, the operations were performed in the space corresponding to one stage later in the optic tract—the outputs of the LGN. Since deuteranomalous trichromacy is attributed to a weak M cone output, it is remarkable that manipulating the model analogs of the LGN outputs accounts so well for the deficiency in the color blind subjects.

The experiments described herein largely agreed with our predictions, providing a novel insight into the perception of color by anomalous trichromats (and perhaps dichromats as well). Though only a limited number of test subjects and images were used here, and thus no statistically significant claims can be made, the results are indeed striking. Planned investigations will use a larger pool of subjects and a more comprehensive collection of images and will add to the current understanding of how color is perceived by the human brain.

References

[1] B. A. Wandell, *Foundations of Vision*. Sunderland, MA: Sinauer Assoc., Inc., 1995.

[2] T. Tomita, "Electrical response of single photo receptors," *Proc. IEEE*, vol. 56, pp. 1015–1023, 1968.

[3] F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. San Francisco, CA: Holden-Day, 1965.

[4] T. N. Cornsweet, *Visual Perception*. New York: Academic, 1970.

[5] L. M. Hurvich and D. Jameson, *The Perception of Brightness and Darkness*. Boston, MA: Allyn and Bacon, 1966.

[6] R. L. DeValois, C. J. Smith, S. T. Kital, and A. J. Karcly, "Responses of single cells in different layers of the primate lateral geniculate nucleus to monochromatic light," *Science*, vol. 127, pp. 238–239, 1958.

[7] R. L. DeValois, I. Abramov, and G. H. Jacobs, "Analysis of response patterns of lgn cells," *J. Opt. Soc. Amer.*, vol. 56, pp. 966–977, 1966.

[8] T. N. Wiesel and D. H. Hubel, "Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey," *J. Neurophysiol.*, vol. 29, pp. 1115–1156, 1966.

[9] R. L. De Valois, "Central mechanisms of color vision," in *Central Processing of Visual Information A*. New York: Springer-Verlag, 1973, vol. VII/3, ch. 3, pp. 209–253.

[10] O. D. Faugeras, "Digital color image processing and psychophysics within the framework of a human visual model," Ph.D. dissertation, Univ. Utah, Salt Lake City, June 1976.

[11] W. L. Stiles, "A modified Helmholtz line element in brightness-color space," *Proc. Phys. Soc. (U.K.)*, vol. 58, pp. 41–65, 1946.

[12] W. R. J. Brown and D. L. MacAdam, "Visual sensitivities to combined chromaticity and luminance differences," *J. Opt. Soc. Amer.*, vol. 39, no. 10, pp. 808–834, 1949.

[13] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, July] 1972.

[14] C. F. Hall, "Digital color image compression in a perceptual space," Ph.D. dissertation, Univ. Southern California, Los Angeles, 1978.

[15] O. D. Faugeras, "Digital color image processing within the framework of a human visual model," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 380–393, Aug. 1979.

[16] U. Ophthalmology Branch, Clinical sciences division, *Aerospace Ophtalmology*, 1988, vol. II, ch. 15.
[17] G. Wyszecki and W. S. Stiles, *Color Science*. New York: Wiley, 1975.
[18] J. Neitz, M. Neitz, and P. M. Kainz, “Visual pigment gene structure and the severity of color vision defects,” *Science*, vol. 274, pp. 801–803, 1996.
[19] S. K. Shevell and J. C. He, “The visual photopigments of simple deuteranomalous trichromats inferred from color matching,” *Vis. Res.*, vol. 37, no. 9, 1997.
[20] J. Lutze, N. J. Cox, V. C. Smith, and J. Pokorny, “Genetic studies of variation in Rayleigh and photometric matches in normal trichromats,” *Vis. Res.*, vol. 30, no. 1, 1990.
[21] H. Scheibner and T. Kremer, “Deuteranomaly studied with four perceptual criteria,” *Vision Research*, vol. 36, no. 19, 1996.
[22] C. M. Cicerone, A. L. Nagy, and J. L. Nerger, “Equilibrium hue judgements of dichromats,” *Vis. Res.*, vol. 27, no. 6, 1987.
[23] A. B. Poirson and B. A. Wandell, “Appearance of colored patterns: Pattern-color separability,” *J. Opt. Soc. Amer.*, vol. 10, no. 12, pp. 2458–2470, 1993.
[24] I. Dvorine, *The Dvorine Color Perception Test*. Baltimore, MD: Waverly, 1944.
[25] C. E. Martin, S. K. Rogers, and M. Kabrisky, “Digital production of color Mach bands using a color human visual system model,” IEEE Trans. Syst., Man, Cybern., Aug. 1996, submitted for publication.

Integrating Knowledge Sources in Devanagari Text Recognition System

Veena Bansal and R. M. K. Sinha

***Abstract*—The reading process has been widely studied and there is a general agreement among researchers that knowledge in different forms and at different levels plays a vital role. This is the underlying philosophy of the Devanagari document recognition system described in this work. The knowledge sources we use are mostly statistical in nature or in the form of a word dictionary tailored specifically for optical character recognition (OCR). We do not perform any reasoning on these. However, we explore their relative importance and role in the hierarchy. Some of the knowledge sources are acquired *a priori* by an automated training process while others are extracted from the text as it is processed.**

A complete Devanagari OCR system has been designed and tested with real-life printed documents of varying size and font. Most of the documents used were photocopies of the original. A performance of approximately 90% correct recognition is achieved.

***Index Terms*—Devanagari document processing, knowledge-based systems, optical character recognition.**

I. Introduction

Segmentation and classification are the two primary phases of a text recognition system. The segmentation process extracts recognition units from the text. The recognition unit is usually a character. The classification process computes certain features for each isolated character and each character is classified to a class which may be the true class (correct recognition), wrong class (substitution error), or an unknown class (rejection error). Based on the output of the classification process, resegmentation of selected units is attempted [7], [14] and reclassification is done.

However, some classification errors still remain which many researchers have tried to correct using contextual knowledge [12], [16], [22]–[25].

More than one classifying feature have been employed for classification to take care of different discriminating features under varying real-life situation [11], [13]. In other words, the contextual knowledge is used during the classification phase also [1].

In this paper, our concern is Devanagari script. It is the script for Hindi which is official language of India. It is also the script for Sanskrit, Marathi, and Nepali languages. The script is used by more than 450 million people on the globe. The algorithms which perform well for other scripts can be applied only after extensive preprocessing which makes simple adaptation ineffective. Therefore, the research work has to be done independently for Devanagari script. Sinha *et al.* [15], [20], [21] and some other researchers [17], [18] have reported various aspects of Devanagari script recognition. However, none of these works have considered real-life documents consisting of character fusions and noisy environment.

An optical character recognition (OCR) for Devanagari and Bangla (an Indian language script) printed script has been described by Chaudhuri and Pal [9], [10]. The emphasis is on recognizing the vowels and consonants (basic characters) which constitute 94–96% of the text. The basic characters are described in terms of nine types of strokes. A stroke is described in terms of its length, height and expected angle. These angles are very much font specific. Moreover, the fading of some black pixels may mislead the stroke extraction process. Even the hand-crafted character specific features may not sustain their utility across various fonts.

Our approach to Devanagari script recognition is to use a number of knowledge sources and integrate them in hierarchical manner. There is no reasoning performed with the knowledge sources. These are mostly statistical in nature, or in the form of word dictionary tailored specifically for OCR. The character classification is based on a hybrid approach. The decision for further segmentation of an image box is based on the outcome of the classification process and the statistical analysis of width of the image box.

We integrated the knowledge sources using the blackboard model primarily because it facilitates easy modifiability of the system. Some of the knowledge sources are acquired by a separate automated training phase. Whereas, the transient knowledge sources are constituted from the knowledge extracted from the text as it is processed. The performance is studied and presented in Section V. The overall system architecture and the knowledge sources are presented in Section III. Section IV describes the implementation details. We introduce Devanagari script from the OCR viewpoint in the next section.

II. Devanagari Script from OCR View Point

Devanagari script is a logical composition of its constituent symbols in two dimensions. It is an alphabetic script. Devanagari has 11 vowels and 33 simple consonants. Besides the consonants and the vowels, other constituent symbols in Devanagari are set of vowel modifiers called *matra* (placed to the left, right, above, or at the bottom of a character or *conjunct*), pure-consonant (also called half-letters) which when combined with other consonants yield conjuncts.

A horizontal line called *shirorekha* (a headerline) runs through the entire span of work. Some illustrations are given in Figs. 1 and 2.

Devanagari script is a derivative of ancient Brahmi script which is mother of almost all Indian scripts. Word formation in Indian scripts

Manuscript received July 21, 1998; revised August 2, 1999. The review of this paper was recommended by Associate Editor S. Lakshmivarahan.

V. Bansal is with the Department of Indistrial and Management Engineering, Indian Institute of Technology, Kanpur 208 016 India (e-mail: veena@iitk.ac.in).

R. M. K. Sinha is with the Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur 208 016, India (e-mail: rmk@iitk.ac.in).

Publisher Item Identifier S 1083-4427(00)01173-5.