Toward Color Image Segmentation in Analog VLSI: Algorithm and Hardware

FRANK PEREZ AND CHRISTOF KOCH
California Institute of Technology, Computation and Neural Systems Program, 216-76, Pasadena, CA 91125
perez@cns.caltech.edu

Abstract

Standard techniques for segmenting color images are based on finding normalized RGB discontinuities, color histogramming, or clustering techniques in RGB or CIE color spaces. The use of the psychophysical variable hue in HSI space has not been popular due to its numerical instability at low saturations. In this article, we propose the use of a simplified hue description suitable for implementation in analog VLSI. We demonstrate that if the *integrated white* condition holds, hue is invariant to certain types of highlights, shading, and shadows. This is due to the *additive/shift invariance* property, a property that other color variables lack. The more restrictive uniformly varying lighting model associated with the *multiplicative/scale invariance* property shared by both hue and normalized RGB allows invariance to transparencies, and to simple models of shading and shadows. Using binary hue discontinuities in conjunction with first-order type of surface interpolation, we demonstrate these invariant properties and compare them against the performance of RGB, normalized RGB, and CIE color spaces. We argue that working in HSI space offers an effective method for segmenting scenes in the presence of confounding cues due to shading, transparency, highlights, and shadows. Based on this work, we designed and fabricated for the first time an analog CMOS VLSI circuit with on-board phototransistor input that computes normalized color and hue.

1 Introduction

The primary goal of color segmentation in machine vision is to determine where changes of material occur in a visual scene. While a material's surface properties are associated with its spectral reflectance signature, it is only the image radiance—spectral reflectance multiplied by the illumination—that is available for information processing (see figure 1). The measurement of surface color from the image radiance is an approximation to obtaining the spectral reflectance signature. Nevertheless, for a standard reference illumination the measurement of relative color differences calculated from the image radiance will relate to material change.

Much work on color segmentation has been based on physically based models. Rubin and Richards (1982) describe an approach based on assigning material differences to spectral crosspoints. Gershon, Jepson, and Tsotsos (1986) incorporate that idea in a double-opponent center surround operator to distinguish material changes from shadow boundaries. Other approaches involve separating image radiance into surface reflection and body reflection. For example, Klinker, Shafer, and Kanade (1988, 1990) discuss a physically based color reflection model—the dichromatic model (Shafer 1985)—that accounts for highlight reflection and matte shading to improve segmentation. Their method is based on evaluating planar clusters in 3-D color space while considering camera limitations. Similarly, Healy and Binford (Healey 1987; Healey & Binford 1989) describe a reflectance model of materials to classify metals and dielectric surfaces.

Nonphysically based segmentation models, in particular those that utilize hue, have been few and far between. Ohlander (1976) uses recursive thresholding on 9 parameter color-space (RGB, HSI, YIQ) histograms for his segmentation algorithm. Ohta (1980) evaluates the performance of 7 different color spaces (21 parameters, 3 of which were HSI) utilizing Ohlander's algorithm. In recent times, Celenk (1990) utilizes the peaks of the 1-D histograms of the equivalent hue-saturation-intensity (HSI) coordinates in the CIE (L*a*b*) uniform-
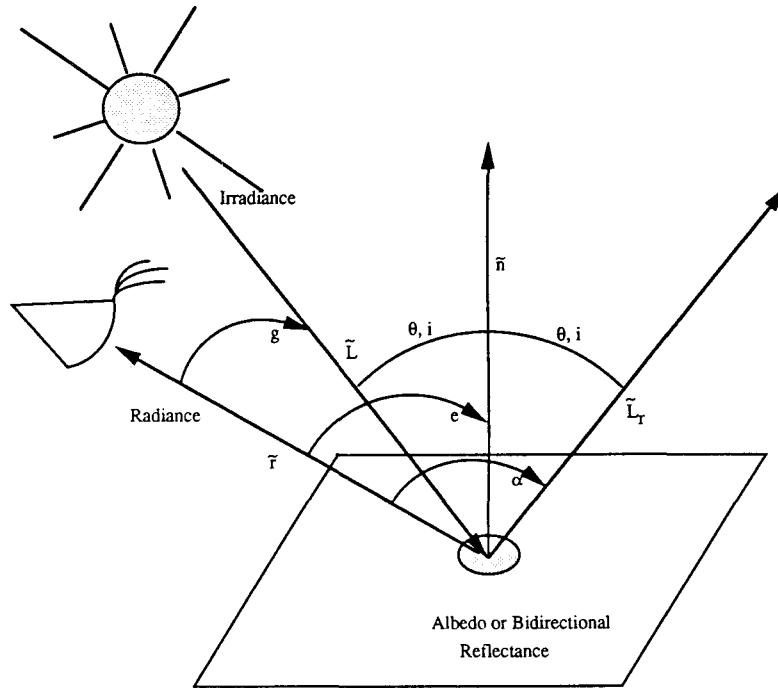
Fig. 1. The geometry of image formation. The surface property, albedo, is contained in the radiance signal—viz. Radiance(λ) = irradiance(λ) * albedo(λ). (Alternate variables i, e, g are utilized in the dichromatic model (Shafer 1985).

color coordinate system to identify cluster regions and then projects the clusters onto a line for one-dimensional thresholding. Similarly, Tominaga (1990) utilizes iterative histogramming on the principal components of the CIE (L*a*b*) color space to identify cluster regions, and then applies a grouping operation dependent on hue difference thresholds.

Bajcsy (1990) uses a physically based model to construct HSI from an orthogonal basis space whereupon shading, highlights, shadows, and interreflections are discounted. Nevertheless, clustering, histogramming, and thresholding are applied to the hue parameter.

All these methods use classical pattern-recognition techniques—histogram thresholding, linear discriminant function, recursive region masking, clustering, etc.—which require heavy computational resources. These techniques represent neither plausible schemes for naturally evolved, efficient, biological systems nor lend themselves very easily to dedicated analog "vision chips." We investigate methods for color segmentation based on a simple hue space, discuss its properties, and show its similarity to other measures of hue. Our research is fueled by the desire to understand the basis of color segmentation in primate cortex as well as to imitate those structures in analog CMOS VLSI circuits

(see for example Harris (1990a), Koch (1991), and Perez (1992a)). In particular, the low accuracy and inhomogeneity of the circuit components we are using (Mead 1989) forces us—similar to biological evolution—to consider simple and robust implementation of vision algorithms. Accordingly, we reevaluate and justify a simplified hue description, propose an algorithm for hue segmentation, and demonstrate its adaptation to the first version of an analog VLSI hue chip.

1.1 Why Hue?

In this study, the psychophysically based parameters hue, saturation, and intensity were used as the starting point for color segmentation. Electrophysiological studies suggest that hue is computed at a high level in the nervous system. An examination of the spectral response curves in area V4 of the monkey visual cortex bears this out (Desimone 1985; Zeki 1983). Also, anthropological studies indicate that hue ordering and color naming are universal and not culturally unique (Berlin 1969).

These results suggest that hue is a high-level variable. The motivation for its selection in image segmentation is that material boundaries correlate more

strongly with hue than with intensity differences. Shadow boundaries are strongly associated with intensity edges, and less so with hue boundaries. The same is true for highlight boundaries and transparency boundaries. This fact is borne out in this study (see also the study of Bajcsy (1990)). Furthermore, segmentation in the 1-D hue space is computationally less expensive than in the 3-D red-green-blue (RGB) space. In comparison to other techniques, the method developed here is simpler, yet as effective in removing confounding cues as the color clustering and histogram thresholding methods developed for color image segmentation.

The history of hue as a computational variable for image segmentation has not been favorable. Kender's 1976 study showed that nonlinear color transforms such as HSI and normalized color have essential singularities, and spurious modes due to the digitizing nature of the nonlinear transforms (Kender 1976). Kender's recommendation was to use linear spaces. Still, researchers have used (Healey 1989; Hurlbert 1989; Nevatia 1977; Poggio 1988; Rubin & Richards 1984) normalized color as the basis of an illumination-independent color space. In section 3 we evaluate the advantages and disadvantages of the hue space, and how it compares to normalized color. But first, we examine the competing color spaces.

2 Color Space Comparison

The different color spaces in use today include: RGB, normalized RGB (Nrgb), YIQ, HSI, Opponent color, Munsell, and various CIE spaces (Daily 1989; Foley 1990; Joblove 1978; Ohta 1980; Schwarz 1987; Smith 1978; Wyszecki 1982). The existence of many different color spaces is mainly a result of color scientists attempting to construct perceptually uniform color organization. The spaces examined in this study are the RGB, Nrgb, CIE (L*a*b*), and HSI color spaces.
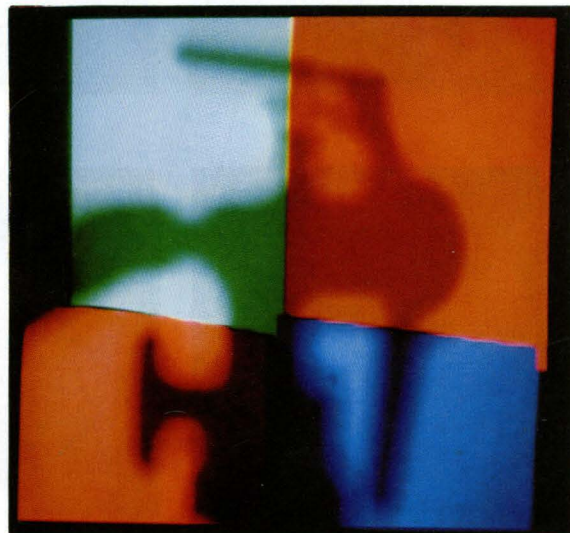
2.1 RGB Space

The red, green, and blue colors represent the tristimulus components (Wyszecki 1982) and define the basic color space. Each of these components corresponds to a filtered spectral mapping from image space to a 3-D sensor space. The equation governing this transformation is

$$C = \int_{\lambda} E(\lambda)S_C(\lambda)\, d\lambda \qquad \text{for } C = (R, G, B) \quad (1)$$

where C are the tristimulus values, $E(\lambda)$ is the incoming light intensity or radiance, and S_C are the three hypothetical color filters.

The RGB images of the original color images in figure 2 are shown in figure 3. These images were chosen because they contained some attributes of shading,

(a)

(b)

Fig. 2. Original color images of (a) shadowed Mondrian and (b) peppers. In both images a projector lamp was used for direct illumination while fluorescent light provided ambient illumination.
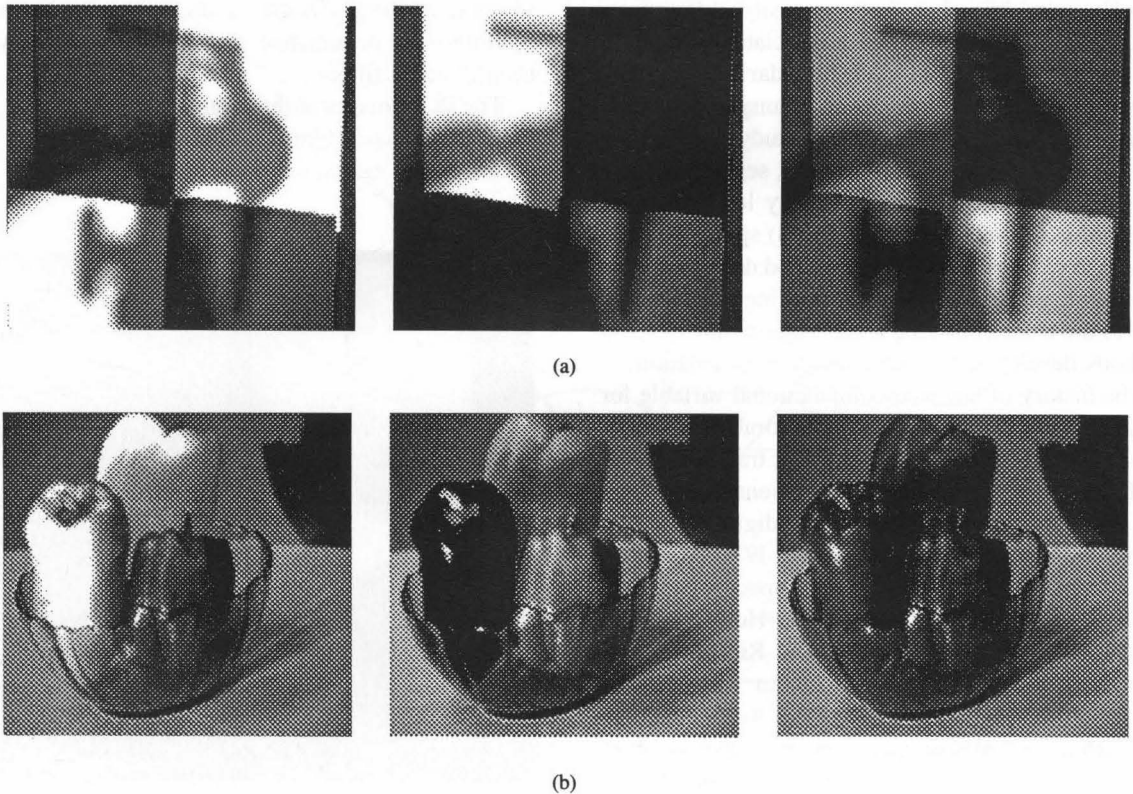
(a)

(b)

Fig. 3. Red, green, and blue components of (a) Mondrian and (b) peppers.

transparency, highlights, and shadowing. Figure 2a is a simplified *Mondrian* consisting of four patches with the shadow of a camera tripod head cast upon it. Figure 2b is a group of peppers in a semitransparent bowl.

The major problem of the RGB space is that segmentation needs to be performed in 3-D space. Interesting attempts to combine information across these spaces (besides intensity) have been made, such as Wright's study of fusing R, G, and B images using Markov random fields (Wright 1989). Still, each of the components in the RGB space are highly correlated and not independent of each other. This result is also confirmed in the principal-component analysis study of Ohta (1980).

2.2 Normalized RGB Space

Nrgb gives a space that is independent of uniformly varying lighting levels. The transformation to normalized colors is given by

$$N_C = \frac{C}{(R + G + B)} \quad \text{for } C = (R, G, B) \quad (2)$$

But, since Nrgb in equation (2) is redundant (viz. $N_B = 1 - N_R - N_G$), the preferred normalized color space is typically formulated (Hurlbert 1989, Kender 1976, Nevatia 1977) as

$$Y = c_1 R + c_2 G + c_3 B$$

$$T_1 = \frac{R}{(R + G + B)}$$

$$T_2 = \frac{G}{(R + G + B)} \quad (3)$$

where c_1, c_2, and c_3 are chosen constants such that $c_1 + c_2 + c_3 = 1$. Y is interpreted as the image luminance of the image pixel and T_1 and T_2 are chromatic variables which are approximately independent of illumination (Nevatia 1977). The Nrgb images of the Mondrian and the peppers are shown in figure 4.

2.3 HSI Space

Of the many similar spaces that achieve hue-saturation-intensity (HSI) characteristics, that is, color-ordering
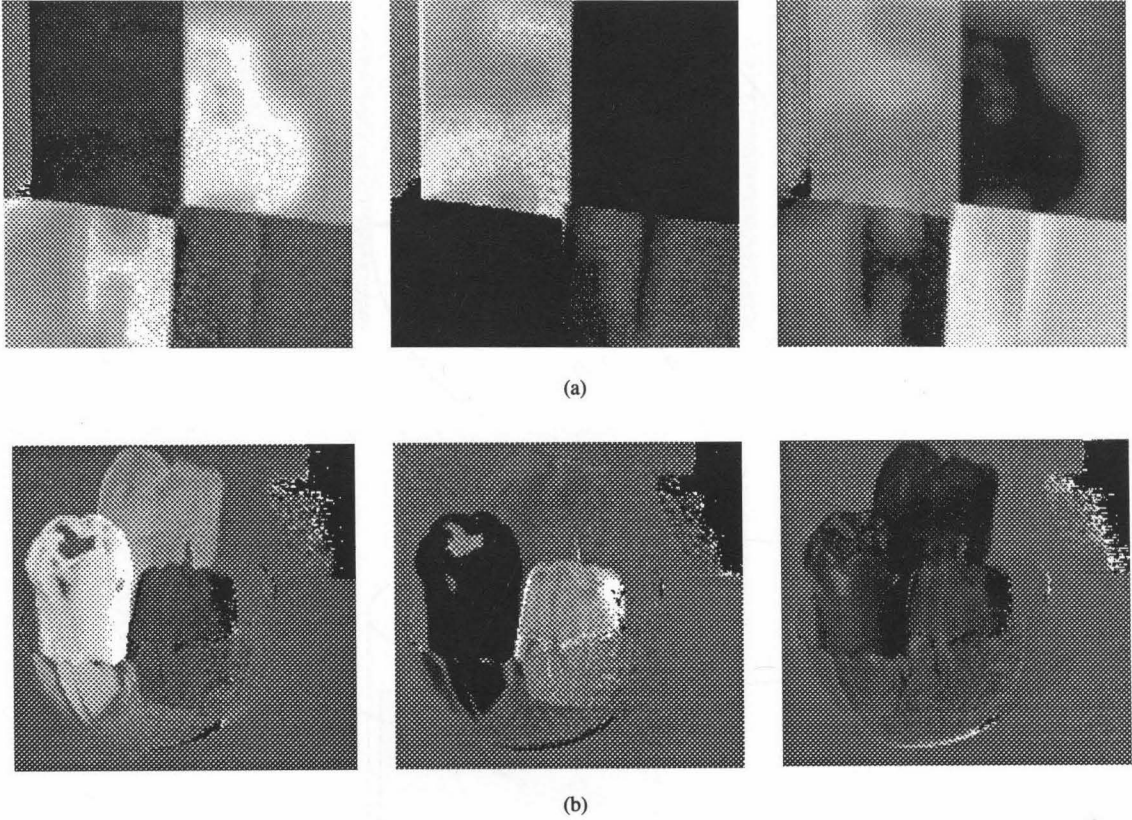
(a)

(b)

Figure 4. Normalized red, green, and blue components of (a) Mondrian and (b) peppers.

systems that are based on human color perception, the Munsell color system is remarkably popular (Wyszecki 1982). This system characterizes color in terms of hue, chroma (or saturation), and value (or intensity) components and has been shown to achieve favorable hue segmentation (Tominaga 1987).

Many transforms from RGB to HSI type spaces have been presented in the computer graphics (Foley et al. 1990; Joblove 1978, Smith 1978) and computer vision (Gershon 1985; Jain 1989; Kender 1976; Tominaga 1987) literature. The transformation to HSI from RGB used in this study is currently implemented with special-purpose digital hardware (Genz 1990) and is given by

$$\text{int} = \frac{(R + G + B)}{3}$$

$$\text{sat} = 1 - \frac{\min (R, G, B)}{\text{Int}}$$

$$\text{hue} = \arctan \left[\frac{\sqrt{3}(G - B)}{(R - G) + (R - B)} \right] \quad (4)$$

where arctan (y/x) utilizes the signs of both y and x to determine the quadrant in which the resulting angle lies. Generally, hue is thought of as the angle between a reference line and the color point in the RGB system. The physical model used to determine the hue angle is based on the diagram shown in figure 5a. If the R, G, B radial basis vectors are equally spaced $\frac{2}{3}\pi$ apart on the unit circle, the x and y component of an arbitrary point can be calculated from basic trigonometry and are given by

$$x = R - \frac{G + B}{2} = \frac{1}{2} [(R - G) + (R - B)]$$

$$y = \frac{\sqrt{3}}{2} (G - B)$$

This results in the hue angle shown in equation (4). Conceptually, one can think of the HSI space as a cylindrical one (figure 5b), where the coordinates r, θ, z respectively correspond to saturation, hue, and intensity. The resulting HSI components for the Mondrian and the peppers are shown in figure 6.
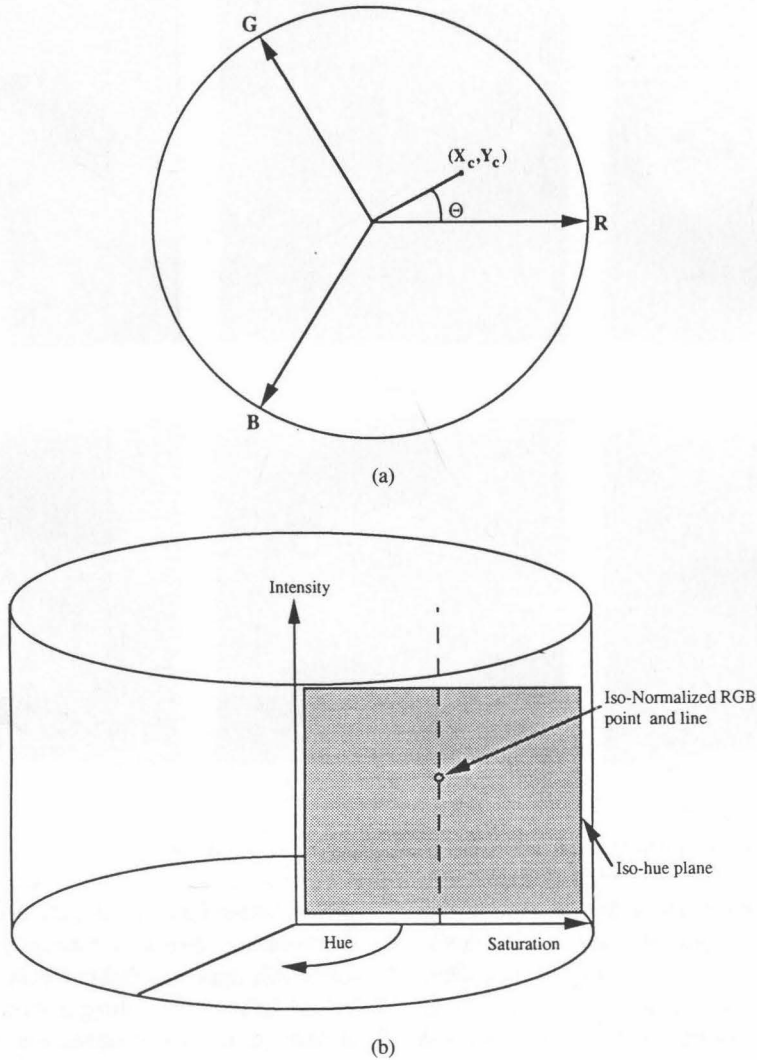
(a)

(b)

Fig. 5. Details of hue space. (a) Physical model for simplified hue based on weighted average of RGB vectors. (b) Comparison of normalized color in HSI space. The nontilted hue planar structure is a simplification of the dichromatic planar hypothesis of Klinker et al. (1990) without the added complexity of color clustering and histogramming.

2.4 CIE Spaces

For completeness we describe the CIE "uniform" perceptual spaces involved in human color perception. The concept of "uniform perceptual distance" is an anthropormorphic one in that these spaces were set up so that traversals of a unit distance in any direction in the space is perceived by human observers to have the same "color difference." Computing the CIE representation of color involves a linear intermediate transformation, followed by a nonlinear transformation. First, the tristimulus values R, G, and B are transformed to another tristimulus set: X, Y, and Z.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.490 & 0.310 & 0.200 \\ 0.177 & 0.813 & 0.011 \\ 0.000 & 0.010 & 0.990 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

For other arbitrary RGB sensor inputs, the transfer matrix must be determined empirically. In particular, the transform for the NTSC receiver primary system is

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (6)$$

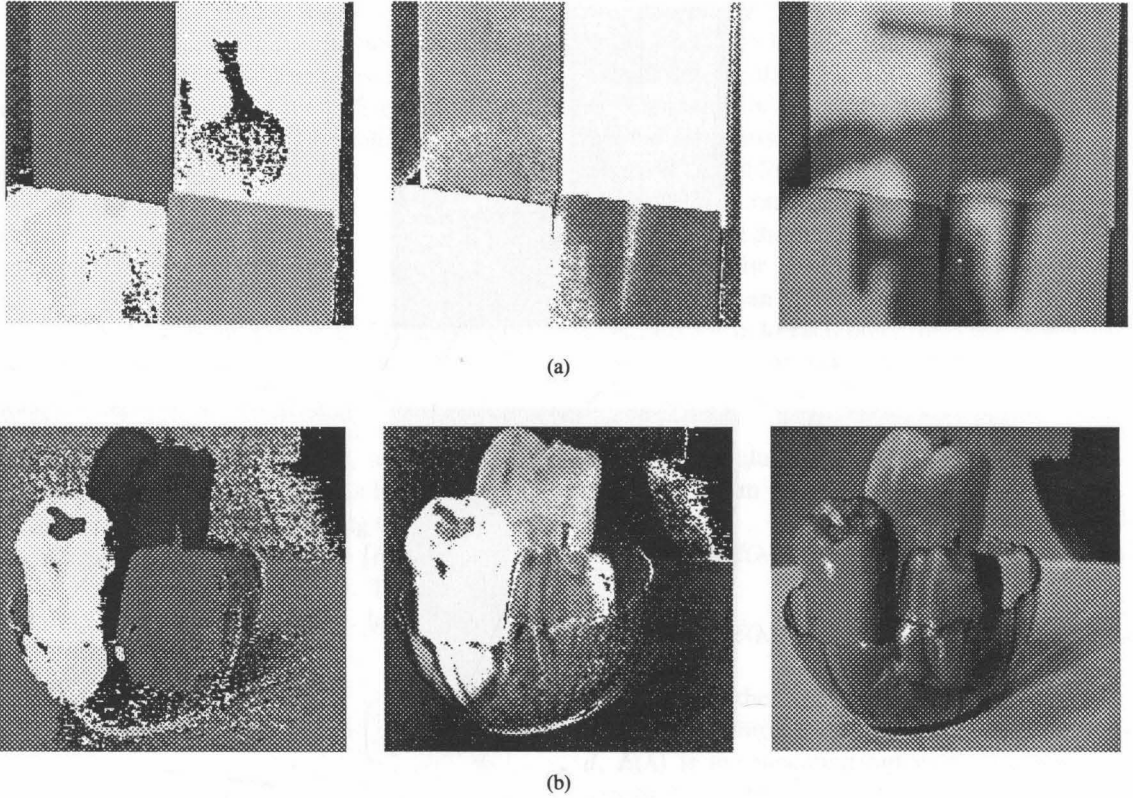Once the XYZ tristimulus coordinates are computed, a number of different CIE spaces can be constructed.

(a)

(b)

Fig. 6. Hue, saturation, and intensity components of (a) Mondrian and (b) peppers.

We use the CIE (L*a*b*) space, because it appears to have more uniform perceptual properties and gives better results than the CIE (L*u*v*) space in segmenting color pictures (Ohta et al. 1980). The CIE (L*a*b*) space is represented by

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{1/3} - \left(\frac{Y}{Y_n} \right)^{1/3} \right]$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right] \quad (7)$$

with the constraint that X/X_n, Y/Y_n, $Z/Z_n > 0.01$. Chroma in this space is defined by

$$\text{chroma} = [(a^*)^2 + (b^*)^2]^{1/2} \quad (8)$$

and hue by

$$\text{hue} = \arctan \left(\frac{b^*}{a^*} \right) \quad (9)$$

The CIE (L*a*b*) hue, chroma, and lightness images of the Mondrian and peppers are shown in figure 7.

Both CIE (L*u*v*) and CIE (L*a*b*) spaces require an intermediate transform to the XYZ system from the system dependent RGB system, and then either a normalization or a cube-root transformation. In comparison, the hue transformation given by equation (4) is substantially simpler. This added complexity is not warranted since we will show that the simplified hue formulation gives satisfactory results. Furthermore, CIE spaces were developed for the psychophysical need to have perceptual uniformity for the standard human observer. We are concerned with the use of color segmentation for analog vision sensors and not in matching human perception.

3 Properties of Hue

In this section we describe the fundamental properties of hue and compare it to normalized color. We show some of its problems, and some of its advantages for locating material changes in images.
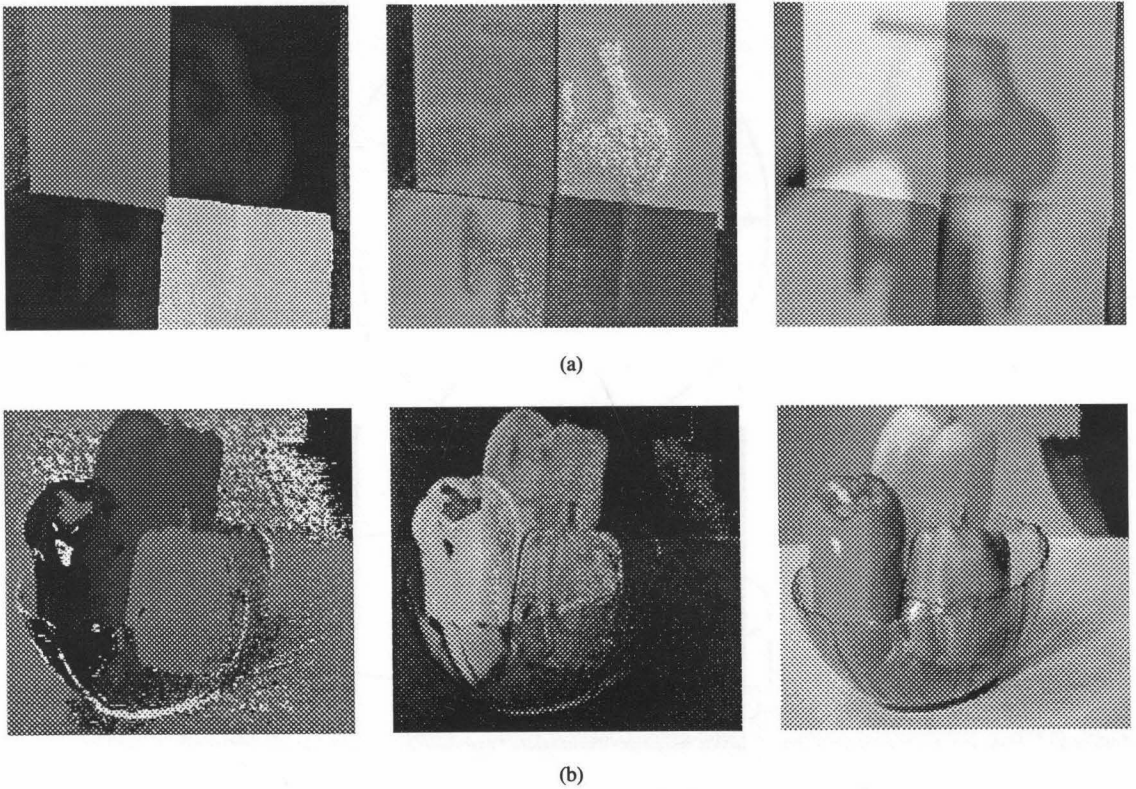
(a)

(b)

Fig. 7. CIE (L*a*b*) hue, chroma, and lightness components of (a) Mondrian and (b) peppers. RGB inputs are based on the NTSC receiver primary system governed by the transform in equation (6). While hue and lightness in this figure are directly related to hue and intensity of figure 6 chroma is unrelated to saturation. Note also that the CIE hue angle is slightly shifted when compared to simplified hue of figure 6.

3.1 Hue Compared to Normalized Color

Both Nrgb and HSI have the desirable property of *multiplicative/scale invariance*, that is, uniform variation of the tristimulus components will not change the measured quantity. This property is illustrated by the following relationships:

$$\text{hue } (R, G, B) = \text{hue } (\alpha R, \alpha G, \alpha B)$$

$$N_C(R, G, B) = N_C(\alpha R, \alpha G, \alpha B) \quad (10)$$

where $\forall \alpha > 0$ and $(\alpha R, \alpha G, \alpha B) \in [0, 1]$. These results can be verified by examining the constituitive equations (2) and (4).

Additionally, hue has *additive/shift invariance* which Nrgb lacks:

$$\text{hue } (R, G, B) = \text{hue } (R + \beta, G + \beta, B + \beta)$$

$$N_C(R, G, B) \neq N_C(R + \beta, G + \beta, B + \beta) \quad (11)$$

where $\forall \beta$ such that $(R + \beta, G + \beta, B + \beta) \in [0, 1]$. This property of *additive/shift invariance* gives isohues

a greater span in the color space. The comparison shown in figure 5b between normalized color and hue in the HSI space illustrates that isohues occupy plane segments while iso-Nrgb occupy line segments. An equivalent statement to equation (11) is that hue is invariant under saturation changes, while normalized RGB is not. Another interpretation is that hue is invariant to white-color vector additions. (We generalize this to the *integrated white* condition in section 3.4.2.) Originally, these *multiplicative/scale* and *additive/shift* invariance properties were used by Kender (1976) to analyze hue instabilities. In this study, we show their advantages.

3.2 Hue in CIE Spaces

Similarly, hue in the CIE (L*a*b*) space defined by equation (9) show *multiplicative/scale* invariance. This property follows from a straightforward evaluation of equation (9) in the invariance relationship of equation (10).

However, the CIE hue lacks strict *additive/shift* invariance. It only approximates *additive/shift* invariance due to its nonlinear cube-root transformation and normalization. Our approach is to avoid the CIE color spaces because of their added complexity in matching human perception.

3.3 Problems of HSI

Kender's 1976 study of the properties of nonlinear color transforms revealed certain problems associated with the use of HSI space. In particular, the HSI transform has the unfortunate property of an unremovable singularity at the axis of the color cylinder, where $R = G = B$ (saturation = 0). We can see this by examining the hues of totally saturated pixels along the red-green segment of the color wheel, that is, hue $\{R = (1 - r)x; G = rx; B = 0\}$, where $0 \leq r \leq 1$. This situation corresponds to int = $x/3$ and sat = 1). It then follows from equation (4) that

$$\lim_{x \to 0} \text{hue } \{(1 - r)x, rx, 0\} = \arctan \left(\frac{\sqrt{3}r}{2 - 3r} \right)$$

Thus, hue varies continuously from 0 (when $r = 0$) to $2\pi/3$ (when $r = 1$). In addition, hue near its singularities is intrinsically unstable. Take for example, pixels whose values are hue $\{x, x, 0\}$, where $x \geq 0$. For 8-bit digital implementation (i.e., hue $\in [0, .255]$) a minimal digital perturbation gives hue $\{x + 1, x, 0\}$ which can result in changes up to $\pi/3$. This characteristic *noise* is prominent in the hue images in figure 6 (and in the hue edges shown in figure 11). Thus, the mapping from RGB to HSI is ill conditioned near the central axis and ill posed at the axis.

The Kender study went as far as recommending not to use nonlinear color transforms such as HSI and normalized color spaces but to use linear transforms such as YIQ and opponent color spaces. This recommendation has been heeded in the applied computer vision literature (Barth 1986; Ohta 1980).

3.4 How Confounding Cues are Discounted in HSI

In this section, we show that assuming reasonable material properties and lighting conditions, hue information will discount intensity edges due to transparency, highlights, shading, and shadowing.

3.4.1 Discounting Transparency. Why should transparency be important? Many robust biological vision systems inhabit the underwater world. Furthermore, water's absorption coefficient shows a nearly uniform spectral transmission throughout most of the visible range (Wolfe 1985), thus, water approximates a neutral density filter. In practice, deep water attenuates red illumination so that the neutral density-filter approximation is only valid for small distances. But a reasonable assumption for an object is that its material points are physically close to each other, thus, the transparency effects for viewing objects in aqueous environments will be negligible because relative transmission distances are small.

The tristimulus equation with and without an absorbing medium is given by

$$X_C = \int_\lambda E(\lambda)S_C(\lambda)d\lambda \qquad \text{for } C = (R, G, B)$$

$$\hat{X}_C(d) = \int_\lambda E(\lambda)S_C(\lambda)\tau(\lambda)^d \, d\lambda \qquad (12)$$

where X_C are the tristimulus values, $\hat{X}_C(d)$ are tristimulus values through transparent medium of distance d, $E(\lambda)$ is the incoming radiance, S_C are the three hypothetical color filters, and $\tau(\lambda)$ is the transmittance per unit distance in the medium.

For a spectrally uniform transmitting medium,

$$\hat{X}_C(d) = (\tau_0)^d \int_\lambda E(\lambda)S_C(\lambda) \, d\lambda = (\tau_0)^d X_C$$

This case has the property of *multiplicative/scale invariance*. Evaluating equation (2) for normalized color and equation (4) for hue results in

$$\text{hue } (X_R, X_G, X_B) = \text{hue } (\hat{X}_R, \hat{X}_G, \hat{X}_B)$$

$$N_C(X_R, X_G, X_B) = N_C(\hat{X}_R, \hat{X}_G, \hat{X}_B)$$

Both normalized colors and hue discount image transparency through a spectrally uniform media.

3.4.2 Discounting Highlights and the Integrated White Condition. In this section we describe the Phong shading model, how it accounts for highlights and how utilizing hue can discount highlights. A similar analysis on other advanced models of shading such as the Cook-Torrance model and the Dichromatic Reflection model gives comparable results (Perez 1992b).

The Phong shading model (Phong 1975) is based on the empirical observation that the radiance from a highlight reflection falls off sharply with increasing α,

where α is the angle between the reflection vector and the viewpoint vector as shown in figure 1. The form adopted by Phong was $\cos^n (\alpha)$, where n varies from 1 to 200 depending on the surface. The full equation for Phong shading is given by

$$I_C = I_{aC} k_{aC} + \frac{I_{pC}}{r + k} [k_{dC} \cos (\theta) + k_s \cos^n (\alpha)]$$

(13)

for $C = (R, G, B)$, where, I_a and k_a are the ambient intensity and ambient reflection coefficient. I_p is the intensity of a point light source, r is the distance from the perspective viewpoint to the surface, k is a constant, k_d is the diffuse reflection coefficient, and k_s is a specular reflection coefficient. Both reflection coefficients k_a and k_d have three components for the tristimulus equations and assume a constant value between 0 and 1. In the Phong model the specular reflection coefficient, k_s, is assumed to be constant and independent of surface color.

In the nonhighlighted (NH) region, the equivalent RGB tristimulus values become

$$C_{NH} = I_{aC} k_{aC} + \frac{I_{pC}}{r + k} [k_{dC} \cos (\theta)]$$

(14)

for $C = (R, G, B)$. Then, if we assume a white light source ($I_{pR} = I_{pG} = I_{pB}$), the equivalent RGB tristimulus for the highlighted (H) region becomes

$$C_H = C_{NH} + \overbrace{\frac{I_{pC}}{r + k} [k_s \cos^n (\alpha)]}^{\beta} = C_{NH} + \beta$$
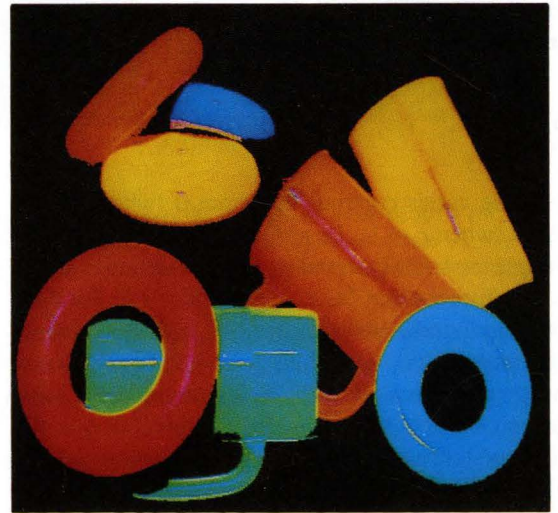
for $C = (R, G, B)$ (15)

We immediately see that this equation satisfies the *additive/shift invariance* condition. Thus, a computation based on hue will discount highlights due to the Phong shading model, while a computation on normalized color, in general, will not. An exception is the condition that the ambient illumination is proportional to the source illumination (i.e., $I_a(\lambda) = c I_p(\lambda)$). Now *multiplicative/scale invariance* holds and both normalized color and hue will discount highlights.

The Phong shading model has been criticized within the computer graphics community because it gives an object a "plastic" appearance. Ironically, the figures used in the study by Klinker, Shafer, and Kanade (1988, 1990) were composed of plastic objects. Furthermore, a simple hue transform on the image in figure 8 gives accurate segmentation while discounting highlights.

(a)

(b)

Fig. 8. (a) Plastic objects with highlights (from Klinker et al. 1988 and 1990), and (b) hue transform of the same image for intensities above 20/255. Note that highlights are removed for most of the plastic objects. Although not considered in this article, we clearly see yellow interreflection on the blue doughnut within the hue space.

The measurements for a horizontal slice through two highlight regions in the subsampled image are shown in table 1. For the central orange cup, a typical slice through the dominant highlight shows that hue varies by 14.9 percent while intensity varies by 53.3 percent of the total range. As we cross the highlight, hue changes 9.4 percent while intensity changes by 32.2 percent.[1]

Table 1.

	Top Left Doughnut ($y = 36$)				Orange Highlight Cup ($y = 57$)		
x	Hue	Sat	Intensity	x	Hue	Sat	Intensity
33	3	126	77	63	7	73	109
34	2	98	88	64	2	19	196
35	0	56	110	65	225	6	245
36	253	57	111	66	249	36	163
37	1	111	78	67	4	64	117
38	3	140	71	68	6	71	112
% span	2.4	32.9	15.7	% span	14.9	27.1	53.3

The hue map shown in figure 8 discounts most of the highlights in the intensity map. The reason for this is that typical plastic has embedded pigment particles within a substrate that is transparent or white (Cook 1981). Thus, light reflected off plastics will have a diffuse colored component and a specular white component.

If we generalize the Phong shading model by assuming that the reflectance coefficients are not constant but vary with wavelength we have

$$I(\lambda) = I_a(\lambda)k_a(\lambda) + \frac{I_p(\lambda)}{r + k}[k_d(\lambda)\cos(\theta) + k_s\cos^n(\alpha)]$$

(16)

This implies that from equation (1) the tristimulus values in the nonhighlighted region becomes

$$C_{NH} = \int_\lambda I_a(\lambda)k_a(\lambda)S_C(\lambda)\,d\lambda +$$

$$\frac{\cos(\theta)}{r + k}\int_\lambda I_p(\lambda)k_d(\lambda)S_C(\lambda)\,d\lambda$$

$$\text{for } C = (R, G, B) \quad (17)$$

For the highlighted region the equivalent tristimulus values become

$$C_H = C_{NH} + \frac{k_s\cos^n(\alpha)}{r + k}\int_\lambda I_p(\lambda)S_C(\lambda)\,d\lambda$$

$$\text{for } C = (R, G, B) \quad (18)$$

We see that hue invariance due to highlights will work exactly if and only if the *integrated white* condition holds,[2] namely if

$$\int_\lambda I_p(\lambda)S_R(\lambda)\,d\lambda = \int_\lambda I_p(\lambda)S_G(\lambda)\,d\lambda$$

$$= \int_\lambda I_p(\lambda)S_B(\lambda)\,d\lambda \quad (19)$$

This follows from an inspection of equation (18) and the fact that the *additive/shift invariance* condition of

equation (11) is reached. The *multiplicative/scale invariance* condition of equation (10) cannot in general be reached for the generalized Phong shading model. The assumption of ambient illumination being proportional to source illumination will not simplify matters much.

But, how reasonable is the *integrated white* assumption? For white light illumination, $I_p(\lambda)$ is independent of λ. This implies that the *integrated white* condition simplifies to

$$\int_\lambda S_R(\lambda)\,d\lambda = \int_\lambda S_G(\lambda)\,d\lambda = \int_\lambda S_B(\lambda)\,d\lambda \quad (20)$$

For artificial systems, this feature may be accommodated by selecting carefully designed spectral filters. Additional compensation of nonwhite illumination to a white standard perhaps utilizing "color constancy" techniques is also necessary. Nevertheless, the general *integrated white* condition, for nonuniform spectral intensities, (equation (19)) must hold in order for hue to discount highlights completely.[3] Anything less will give an approximate hue invariance to highlights. Within the framework of the Phong shading model, normalized color will not generally discount highlights.

3.4.3 Discounting Shading.
Shading or surface orientation change is another confounding cue that will confuse an achromatic vision system. A simple analysis by Rubin (1982) for a single-point light source illuminating a matte surface shows that

$$I_1 = \rho(\lambda)I_p(\lambda)\cos(\theta_1)$$

$$I_2 = \rho(\lambda)I_p(\lambda)\cos(\theta_2) \quad (21)$$

where I_1 and I_2 are the image radiance intensities for two points that differ only by a surface orientation change; $\rho(\lambda)$ is the spectral albedo (which will be similar to the bidirectional reflectance of a lambertian surface); and θ is the angle between the surface normal

and the illumination direction. The tristimulus equations immediately imply that

$$R_1 = \frac{\cos(\theta_1)}{\cos(\theta_2)} R_2$$

$$G_1 = \frac{\cos(\theta_1)}{\cos(\theta_2)} G_2$$

$$B_1 = \frac{\cos(\theta_1)}{\cos(\theta_2)} B_2 \qquad (22)$$

For this simple model, the *multiplicative/scale invariance* condition holds. Thus, both normalized color and hue will discount this model of shading.

If ambient light were added to the model, the *integrated white* condition (such as that defined by equation (19)) must hold for hue to discount confounding shading cues. The analysis for this is similar to the above calculation for highlight invariance for the Phong shading model.

3.4.4 Discounting Shadowing.

For analyzing shadowed regions, we require two sources of illumination: one to cast the shadow on an object and the other to lightly illuminate the darkened region. In practice, the second light source can result from interreflections or scattering from the principle light source. Our starting point is taken from Rubin's analysis for shadowed regions (Rubin 1982). (See also Gershon (1986) for a discussion on distinguishing material from shadow boundaries.) The governing equation for shadowed and nonshadowed regions is given by

$$I_{\text{lit}} = \rho(\lambda)[I_p(\lambda) + I_d(\lambda)]$$

$$I_{\text{shade}} = \rho(\lambda)[I_d(\lambda)] \qquad (23)$$

where I_{lit} and I_{shade} are the radiance intensities in the nonshadowed and shadowed region respectively, ρ is the albedo as defined before, I_p is the illumination intensity, and I_d is the diffuse ambient intensity. The relevant tristimulus equations for the shade region become

$$C_{\text{shade}} = \int_\lambda I_d(\lambda)\rho(\lambda)S_C(\lambda)\, d\lambda \quad \text{for } C = (R, G, B) \qquad (24)$$

While the relevant tristimulus equations for the lit region become

$$C_{\text{lit}} = C_{\text{shade}} + \int_\lambda I_p(\lambda)\rho(\lambda)S_C(\lambda)\, d\lambda$$

$$\text{for } C = (R, G, B) \qquad (25)$$

In the most general case, hue will discount shadows if and only if

$$\int_\lambda I_p(\lambda)S_R(\lambda)\rho(\lambda)\, d\lambda = \int_\lambda I_p(\lambda)S_G(\lambda)\rho(\lambda)\, d\lambda$$

$$= \int_\lambda I_p(\lambda)S_B(\lambda)\rho(\lambda)\, d\lambda$$

This equation seems coincident with the *integrated white* condition of equation (19). However, this condition depends on the surface albedo $\rho(\lambda)$ which depends on material type.[4]

A further simplification is possible if we assume that the ambient lighting and the main lighting are related. In particular, if the ambient lighting is proportional to the main lighting, that is $I_p(\lambda) = \alpha I_d(\lambda)$, then the following tristimulus equations hold

$$C_{\text{lit}} = (\alpha + 1)C_{\text{shade}} \qquad \text{for } C = (R, G, B)$$

For this case, the condition of *multiplicative/scale invariance* holds. Thus, both normalized color and hue will discount this particular shadow cue. In the following table, we compare color values for typical points in the light and shadow regions of the lower right-hand side of the Mondrian of figure 2a within different color spaces.

Color Space	Δ Light to Shade (%)		
R:G:B	80	92	68
nR:G:B	16	68	24
H:S:I	3	55	76
(CIE)H:C:L	4	11	52

Hue varies only 3% across the shadow edge while normalized color varies at least 16%. The performance of the CIE (L*a*b*) hue is similar to the simplified hue (for more details, see section 5.3).

4 Tools for Working in Hue Space

Before we can utilize hue in image segmentation, we need to develop special tools to detect edges in hue space. Further smoothing of the hue parameter—in particular in regions of low saturation—through the Markov random field (MRF) formulation will lead to improved color segmentation.
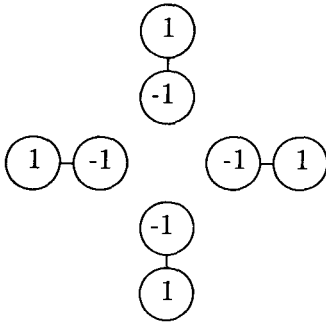
4.1 Finding Edges of a Circular Variable

Standard techniques of intensity edge detection (Canny 1986) by convolving with various masks followed by thresholding or extremum detection such as by Canny's algorithm do not work with hue, because unlike intensity, hue is defined on the ring S^1 rather than on the interval R^1. However, standard convolution techniques can be modified to work with these modulo variables to determine their spatial "edges." Conceptually, this is similar to a 2-D lattice populated with particles whose state is defined by a scalar variable defined on S^1, such as orientation or phase angle θ. In the latter example, "edges" would correspond to all locations on the 2-D lattice across which the phase angle changes maximally.

Traditional edge detection entails finding the *zero-crossings* of the 2-D image convolved with a Laplacian. One of the simplest discrete approximations of a Laplacian is the following kernel:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

In the manner of Terzopoulos (1985), this kernel can be decomposed into the following molecules:

Interpreting these fundamental molecules as distances between nearest neighbors allows the generalization of a modulo Laplacian as the modulo distance of nearest neighbors from a central pixel. Modulo distance or angular distance, $\Phi(y - x)$, can be described as linear saw-tooth pattern with slope 1 and period 2π. It is given by the Fourier series expansion

$$\Phi(y - x) = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin [n(y - x)] \qquad (26)$$

An equivalent form of this equation is given by

$$\Phi(y - x) = (y - x) - 2\pi k \qquad (27)$$

where

$$k = \text{floor} \left[\frac{1}{2} \text{ceiling} \left(\frac{y - x}{\pi} \right) \right]$$

Since the hue values x, y are given in the principal range 0 to 2π, the angular distance formula can be simplified to

$$\Phi(y - x) = y - x + \begin{cases} +2\pi & \text{if } y - x < -\pi \\ -2\pi & \text{if } y - x > \pi \\ 0 & \text{otherwise} \end{cases}$$
$$(28)$$

where positive values of $\Phi(y - x)$ indicate that y is positioned clockwise to x. Therefore, the strategy for finding material boundaries in images is to find the hue *zero crossing* by convolving the hue image with an equivalent mask similar to the Laplacian, the modulo Laplacian.

4.2 Markov Random Field Formulation

Smoothing unstable hue regions is performed by utilization of a regularization technique based on Markov random fields. While other researchers (Hurlbert 1989; Wright 1989) have used the Markov random field formulation in color segmentation, the research of Daily (1989) comes closest in spirit to this work in the selection of hue as a useful measure for image segmentation.

In this study, we use deterministic hue discontinuities to segment hue regions smoothed by a first-order membrane-type stabilizer. This represents a deterministic approximation to the underlying stochastic Markov random field algorithm of Geman (1984); see also Geiger (1990). The advantages of using such a deterministic approach are simplicity, speed, and the fact that hue values are smoothed while hue discontinuities are computed at the same time. Noise is eliminated while discontinuities are preserved. The algorithm utilizes a first-order Tikhonov stabilizing functional (Poggio 1985) by minimizing the following "energy" functional:

$$E_{\text{tot}} = E_{\text{data}} + E_v + E_h + E_{\text{line}}$$

where

$$E_{\text{data}} = \lambda \sum_{ij} [\text{relativedist (hue}_{ij}, D_{ij})]^2$$

$$E_v = \sum_{ij} [\text{relativedist (hue}_{ij+1}, \text{hue}_{ij})(1 - v_{ij})]^2$$

$$E_h = \sum_{ij} [\text{relativedist (hue}_{i+1j}, \text{hue}_{ij})(1 - h_{ij})]^2$$

$$E_{\text{line}} = \alpha \sum_{ij} (h_{ij} + v_{ij})$$

Here, i, j are pixel locations in a rectangular lattice; λ is associated with the data confidence to smoothing ratio; v_{ij} and $h_{i,j}$ are vertical and horizontal hue discontinuity line processes which take on values of 0 or 1; and α is the energy penalty for forming these line processes. The method we used in segmenting scene boundaries was to start with the intensity edge map generated from the Canny edge operator and to gradually eliminate those edges not due to hue differences. Our constraint is that hue discontinuity line processes would only form if pixel saturation values exceeded some

specified minimum value (20% saturation). In doing so, we are discounting edges due to confounding cues of highlights, shading, transparency, and shadows for moderately chromatic scenes.

5 Color Image Segmentation

In this section, we compare conventional edge detection techniques in the RGB and Nrgb spaces to the nonlinear operator developed for hue and intensity in our HSI as well as in the CIE color space.

5.1 Conventional Edge Detection

Standard edge detection in both RGB spaces is performed using Canny's algorithm (Canny 1986). Figure 9a shows the edges obtained from the RGB image of the Mondrian of figures 3a, while figure 9b shows the edges obtained from the RGB image of the peppers of figure 3b. Canny edges for Nrgb images are depicted in figure 10. The computed edges represent intensity
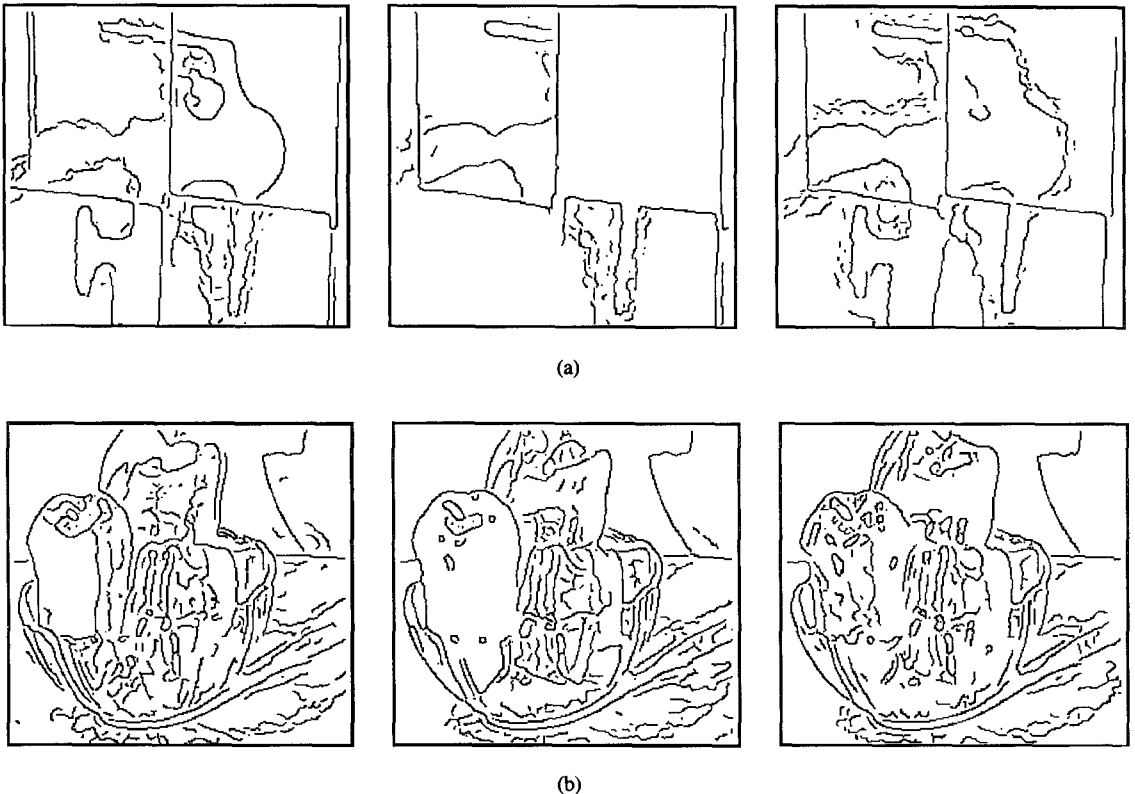

(a)

(b)

Fig. 9. Red, green, and blue Canny edges of (a) Mondrian and (b) peppers.
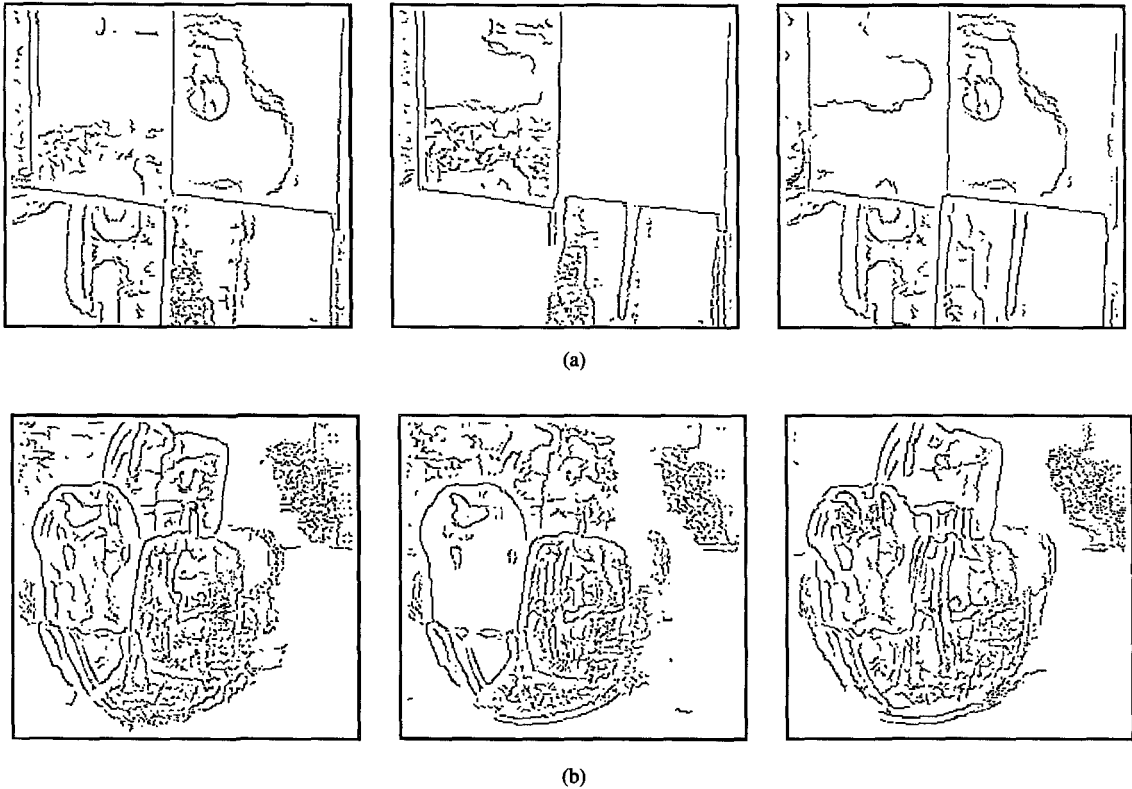
(a)

(b)

Fig. 10. Normalized red, green, and blue Canny edges of (a) Mondrian and (b) peppers.

changes and thus, falsely indicate the shadow as a material boundary. The shadow edges in the Mondrian also show up in the Nrgb space, since only uniform changes in color—what we call *multiplicative/scale invariance—* are factored out.

5.2 Modulo Edge Detection

Based on the modulo algorithm developed in the previous section, we computed "hue edges" associated with both the HSI and the CIE (L*a*b*) color spaces. Figures 11 and 12 display the hue edges as well as the saturation/chroma and intensity/lightness edges of both the Mondrian and the peppers. Comparison of the Mondrian hue edge map with the intensity edge map in both HSI and CIE color spaces show an amelioration of the shadow edge effect. Performance of hue segmentation in this image is discussed in the next section. Additionally, we note that the hue edge map for the peppers in both color spaces is unsatisfactory due to the misbe-

havior of the hue transform. Techniques to improve this performance will be discussed.

5.3 Performance Comparison

Qualitatively, a visual inspection of the Mondrian edges in the four color spaces used illustrates that hue discontinuities in both HSI or CIE spaces appear to be the best measure for detecting material changes, independent of shadow edges. In order for us to form a more quantitative judgement of the relative performance in the different color spaces, we use the figure of merit *F* proposed by Abdou (1979) to compare different edge detection schemes. This is given by

$$F = \frac{1}{\max{(Ni, Na)}} \sum_{i=1}^{Na} \frac{1}{1 + \alpha d^2} \qquad (29)$$

where *Ni* and *Na* represent the number of ideal and actual edge map points and *d* is the closest distance from
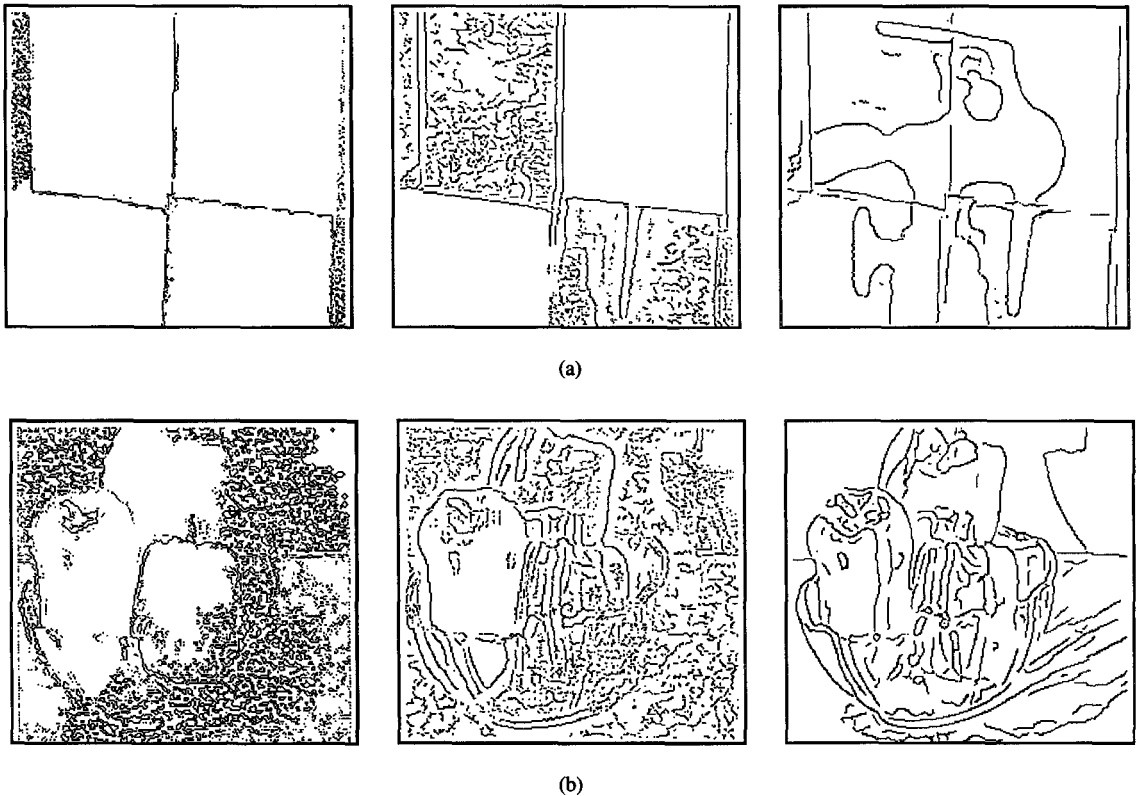
(a)

(b)

Fig. 11. Hue, saturation, and intensity edges of (a) Mondrian and (b) peppers. Excellent performance is achieved in discounting the shadow boundary of the Mondrian hue edge map. The unsatisfactory performance for the peppers is due to the instabilities of the hue transform at low saturations and intensities associated with the shadows and background of figure 2.

an actual edge point to any ideal one.[5] F penalizes both nonlocalized edges and inaccurately positioned ones, with values of F closest to 1 being ideal.

Applying this measure on the saturated portion (the central half) of the Mondrian image gives the following results:

Color Space	Figure of Merit		
R:G:B	0.2596	0.4126	0.2626
nR:G:B	0.1830	0.2509	0.2330
H:S:I	0.9158	0.1765	0.3093
(CIE)H:C:L	0.7421	0.3215	0.2733

In general, normalized color gives spurious edges within low intensity regions. That is why its figure of merit is penalized more than that of the RGB system. For this particular example, hue segmentation in HSI space outperforms the CIE (L*a*b*) hue segmentation.

5.4 Improvements Afforded by MRF Technique

While the hue segmentation works adequately in some areas, it is clear that it performs poorly in other areas. In portions of the image with low saturation and low intensity values, hue values become unstable. This is apparent in the upper left-hand portion of the Mondrian-edge image in figure 11a and in the background regions of the pepper-edge image in figure 11b. This suggests using hierarchical processing based on confidence values that depend on saturation and intensity. That is, hue regions of low saturation and intensity should be smoothed before segmentation is performed. This strategy is the complement of combining all sensory data to find all "true" edges (Ohlander 1976). The rationale behind this algorithm is that a cooperative pixel neighborhood scheme is appropriate for scene segmentation since object pixels will correspond more closely to nearby pixels. This neighborhood correspondence is a fundamental premise in Geman and Geman's MRF formulation.
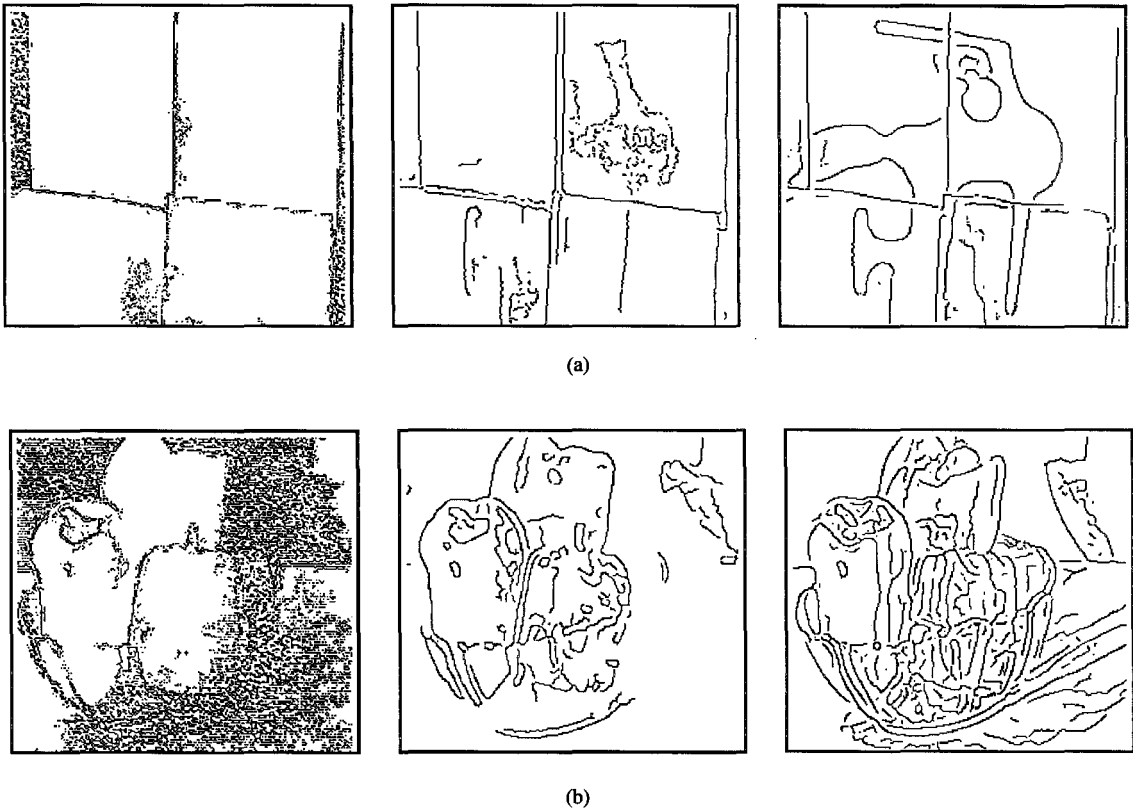
(a)

(b)

Fig. 12. CIL (L*a*b*) hue, chroma, and lightness edges of (a) Mondrian and (b) peppers. Results for the Mondrian are similar to figure 11. Although chromatic edges of the peppers show nearly acceptable scene segmentation, highlight boundaries are confounded.

By utilizing a first-order membrane type stabilizer to smooth low-confidence regions due to low saturation or low intensity, improvement in scene segmentation is obtained. The corresponding hue edges for regions away from unstable saturation are shown in figure 13. The procedure used to generate this map is to start out with an initial edge map identical to the intensity edge map and to minimize the hue energy functional. Edges are created if both critical hue thresholds and critical saturation thresholds are exceeded, while intensity edges are eliminated if hue differences are small.

Notice that edges associated with intensity variations due to shading on the curved bell-pepper surface and intensity filtering through the translucent bowl have been eliminated in the hue edge map. Also, the highlights have been discounted. Thus hue edge map correlates more strongly to the material boundaries than the intensity edge map.
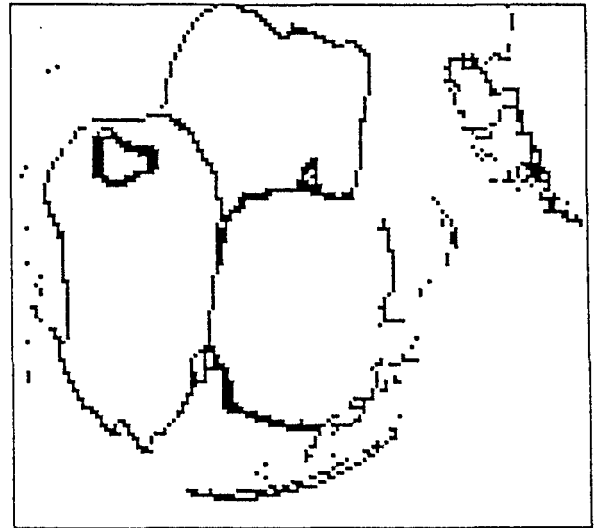

Fig. 13. Smoothed hue edges away from low-saturation regions. The threshold was 20% saturation.

6 Analog VLSI Chip Implementations

Based on the performance of our hue segmentation algorithm, we proceeded to the next step of our research program, implementing a hue sensor using integrated circuit technology. Because the underlying photoreceptor signals are continuous, it seems appropriate to use analog, rather than digital, circuits. Our technology of choice is analog CMOS VLSI technology, as developed and applied to a range of neuromorphic systems by Carver Mead (1989) and his collaborators. A significant number of circuits have been successfully built in this technology, including a *silicon retina* with logarithmic photoreceptors (Sivilotti 1987), resistive networks for smoothing, and "fuses" for detecting discontinuities (Harris 1990a). This will allow us in the future to integrate smoothing and discontinuity detection circuits with our color sensors to build a single *smart sensor* for directly computing hue discontinuities. The alternative CCD camera/color imaging system is a power hungry system requiring analog pixel scanning, A/D conversion, and digital computation. In comparison to CCD camera technologies, analog CMOS circuits operating in the subthreshold regime offer a low power,[6] real-time, illumination independent solution to color segmentation. Our's is the first analog CMOS VLSI circuit that uses on-board photoreceptors responsive to different spectral components.

Although this article successfully demonstrates a number of machine vision algorithms to segment hue, these algorithms are not amenable to analog circuit implementation due to their complexity (witness, for instance, equation (28)). Furthermore, no direct analog circuit can be constructed for hue utilizing equation (4) because the required division operation is not a functional analog computational unit. Our approach fortunately overcomes these obstacles. We report on two color circuits that perform the required computations using simplified algorithms. Because normalization lends itself quite readily to an analog circuit implementation, we first built a circuit for computing normalized color with inputs provided by on-board red, green, and blue photoreceptor sensors. In order to incorporate the *additive/shift* invariance property we utilize an opponency strategy to compute hue components. This final output of the hue circuit are the x and y components of hue rather than the single hue angle. It is interesting that this functional progression from three wavelength-selective signals to opponency to hue imitates the known stages of color computation in the primate visual system

(De Valois 1975; Lenny 1988). These circuits should be viewed as exploratory designs, proving that analog VLSI hue chips are feasible. We first describe some of the practical aspects of chip design and fabrication.

6.1 Chip Background and Experimental Setup

Because all of our chips are fabricated using the government-sponsored silicon foundry service MOSIS, we are restricted to using standard CMOS and BiCMOS processes. Over the visible range of light, the spectral sensitivities of the various kinds of photodiodes and phototransistors available in these processes varies only little (Delbruck 1993). Thus, we are unable to exploit any intrinsic wavelength filtering of silicon. Furthermore, we do not have access to the highly developed colored polymer film-deposition technology found in the majority of modern solid-state commercial video cameras. We therefore had to manually deposit spectral filter over our phototransistors, a quite tedious process which does not lend itself to an extension to one- or two-dimensional arrays of photoreceptors.

We mechanically glued color gelatin filters on top of standard bipolar phototransistors, which were further covered by an IR filter. The area of the phototransistors were approximately sized to accommodate the color filter transmissivity. The phototransistors we use have logarithmic voltage response to over five orders magnitude of intensity change (Mead 1989; Sivilotti 1987). In our application, we use photocurrent output that varies linearly with intensity. All chips were fabricated using the MOSIS 2 μm process. We used the standard 40-pin "Tinychip" die which effectively gives a 1.6 mm by 1.6 mm usable wafer area.

Instead of building our circuits using the prevalent voltage mode of operation (Mead 1989), our circuits operate in the current mode—for an example of building block circuits see the compendium in (Seevink 1988), an approach that is gaining prominent support lately (see Toumazou 1990). Here *charge* rather than *voltage* is the active parameter, resulting in higher usable gain, accuracy, and bandwidth.

To measure spectral tuning curves for our constructed color sensors, we used the calibrated prism monochromator and tungsten incandescent lamp source setup of Delbruck (1993). Although the lamp exhibits nonideal spectral characteristics—intensity is not constant and varies 4 orders of magnitude within its spectrum—our circuits compensates for this variation by

performing a *multiplicative/scale* invariance operation, thus removing intensity contributions from the chromatic computation. We report our results in the 400 to 750 nm wavelength range.

6.2 Normalized Color Sensor

Because of its relative design simplicity, we first designed and built a normalized color sensor (Perez 1992a). The basic circuit design is based on Gilbert's translinear principle (Gilbert 1975) and consists of a current-mode normalization circuit (described by Seevink (1988); see also insert in figure 14b) that has the desired scaling behavior. The input currents I_r, I_g, I_b from the three phototransistors produce normalized output currents N_r, N_g, N_b such that

$$N_r = I_{\text{bias}} \left(\frac{I_r}{I_r + I_g + I_b} \right)$$

$$N_g = I_{\text{bias}} \left(\frac{I_g}{I_r + I_g + I_b} \right)$$

$$N_b = I_{\text{bias}} \left(\frac{I_b}{I_r + I_g + I_b} \right) \tag{30}$$

where I_{bias} is set to operate in the subthreshold regime (Mead 1989). Notice that this equation matches the normalized color equation given in equation (2), and thus offers *multiplicative/scale* invariance.

Figure 14 shows the spectral tuning curves for the normalized color circuit at two different illumination levels: (a) an arbitrary baseline level and (b) 10 times that baseline level. The figure shows that for illumination level increases of one order of magnitude, the circuit output increases no more than 25% for N_r and N_g. That N_b is attenuated by 60% for the baseline illumination is attributed to circuit operation in the nonsaturated regime for this particular experiment. (This glitch can be compensated by circuit redesign.) Since the translinear principle is based on Kirchhoff's voltage law and the transistor voltage-current logarithmic relationship, the circuit normalizing behavior will hold over a large range (4 decades from experiments) of input currents. The advantage of this circuit is that intensity gain control can be achieved electronically rather than via a mechanical iris as done in typical camera systems.

6.3 Hue Sensor

In addition to *multiplicative/scale* invariance, hue has *additive/shift* invariance. To accommodate this extra property, we borrow a concept from biological vision processing, the opponency cell computation. We designed a circuit computing the following *additive/shift* invariant parameters:

$$X^+ = \min (2R - 2G, 0)$$

$$X^- = \min (2G - 2R, 0)$$

$$Y^+ = \min (R + G - 2B, 0)$$

$$Y^- = \min (2B - R - G, 0) \tag{31}$$

Here R, G, and B are the currents from the red, green, and blue phototransistors. Each of the opponency currents X^+ and X^- are half-wave rectified, such that if either one is positive the other one is zero, that is $X = X^+ - X^-$ (the same applies to the opposing pair Y^+ and Y^-). In an earlier version of this chip, we used a single current X that could take on both negative and positive values. However, imbalances due to inherent mismatched properties of the circuit lead to unbalanced output and poor performance. The opponency currents are then processed by the translinear normalization circuit discussed above to give normalized x and y current values that map directly onto a unit diamond on the 4 quadrants of the color plane shown in figure 15. These currents are given by

$$x^+ = I_{\text{bias}} \left(\frac{X^+}{X^+ + X^- + Y^+ + Y^-} \right)$$

$$x^- = I_{\text{bias}} \left(\frac{X^-}{X^+ + X^- + Y^+ + Y^-} \right) \tag{32}$$

with the corresponding equations for y^+ and y^-. Notice that we slightly warp the hue definition given earlier in equation (4) for expedient hardware implementation (R-G is π apart rather than $\frac{2}{3}\pi$).[7] In this case, hue will be defined as

$$\text{hue} = \arctan \left[\frac{R + G - 2B}{2R - 2G} \right]$$

$$= \arctan \left[\frac{y^+ - y^-}{x^+ - x^-} \right] \tag{33}$$

The block diagram circuitry for the hue sensor is shown in figure 15. The output of the chip is the normalized half-wave rectified hue currents x and y, which is shown as a mapping on the unit diamond in the X-Y opponent color space. This mapping is fundamentally different from traditional color space mappings. Rather than [0, 1] mapping onto [0, 1] we have [0, ∞] mapping onto [0, 1]. This is a result of the fundamental differences of the
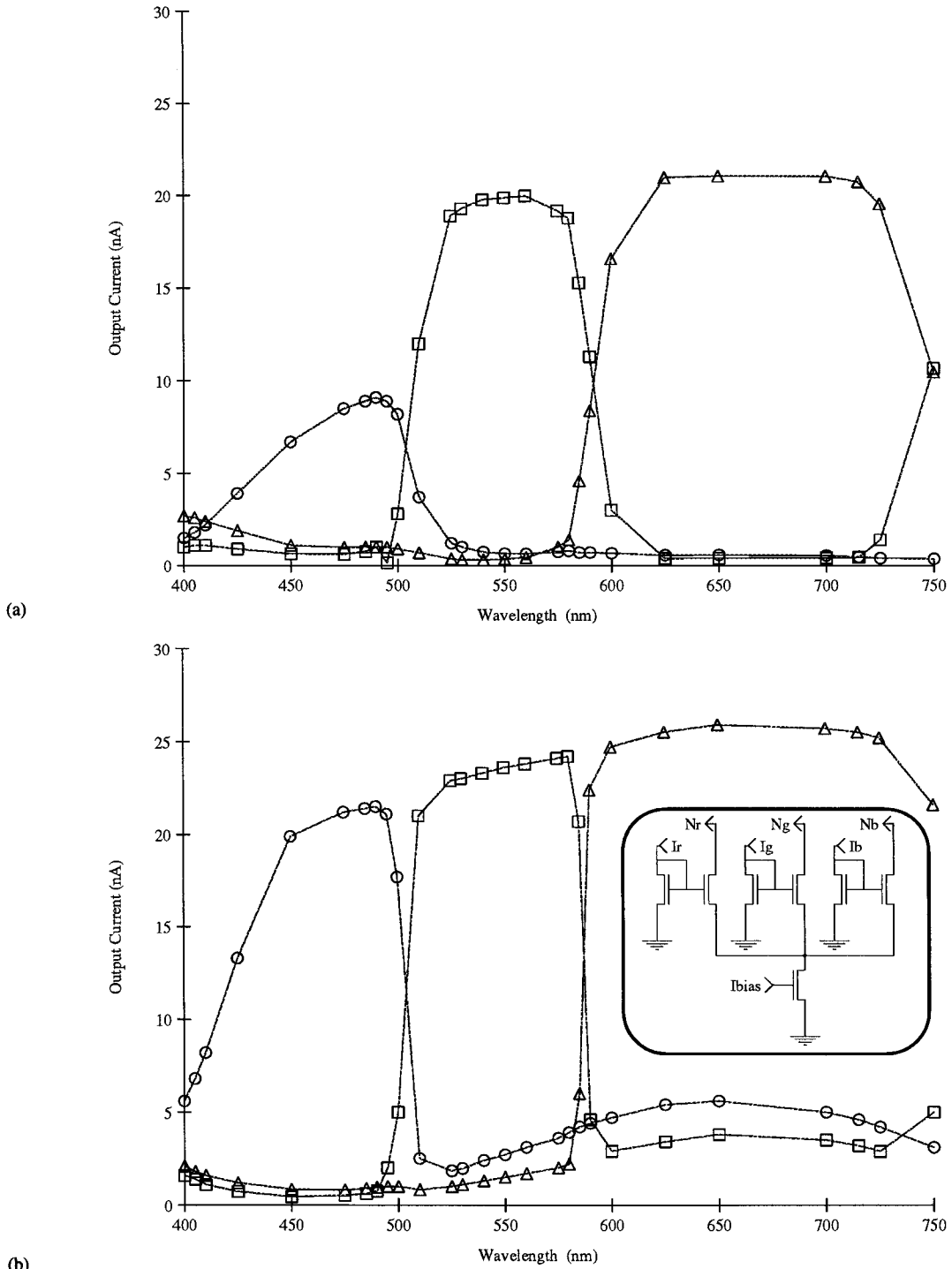
(a)

(b)

Fig. 14. Spectral tuning curves for a normalized color circuit from (a) standard intensity (arbitrarily set) and (b) 10 times standard intensity. The normalization operation is demonstrated on $N_r(\triangle)$ and $N_g(\square)$ because a 10-fold increase in intensity causes a 1.25 increase in circuit output. The 60% attenuation of $N_b(\bigcirc)$ is attributed to low S/N from the experimental setup and the resultant nonsaturation circuit operation. In the central portion of the curve, the sum of the normalized color currents equals a constant I_{bias}. The circuit diagram used for normalized color is shown in the insert.
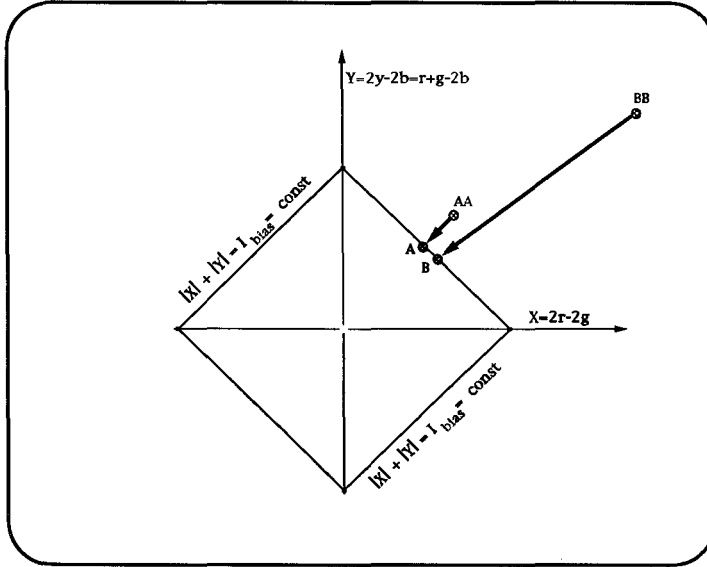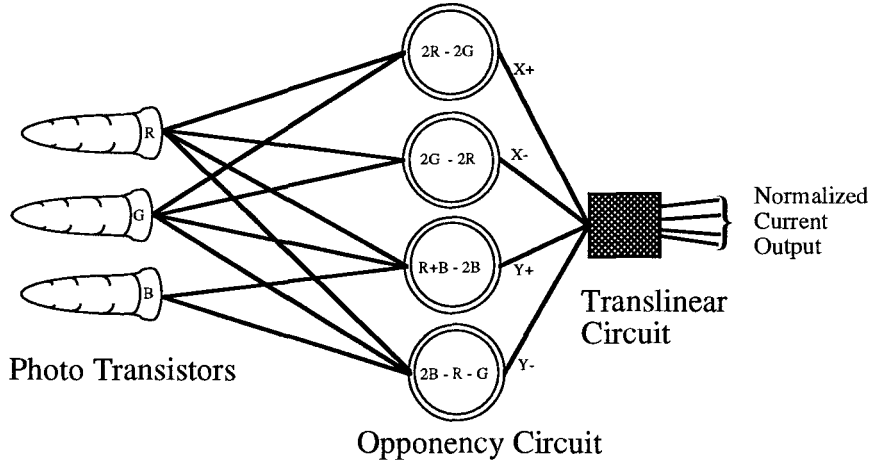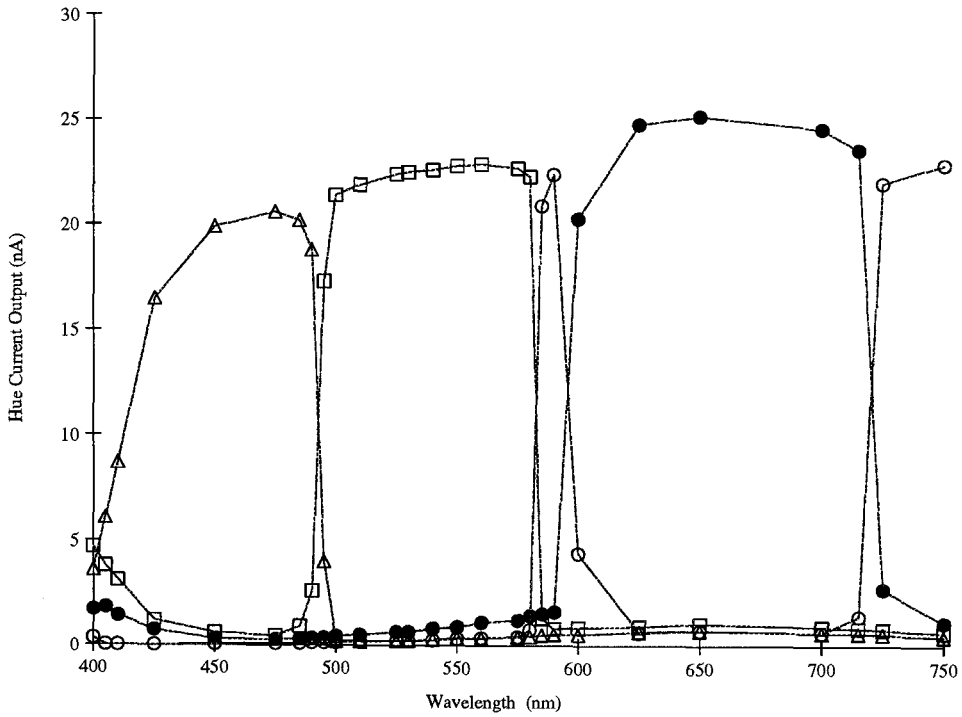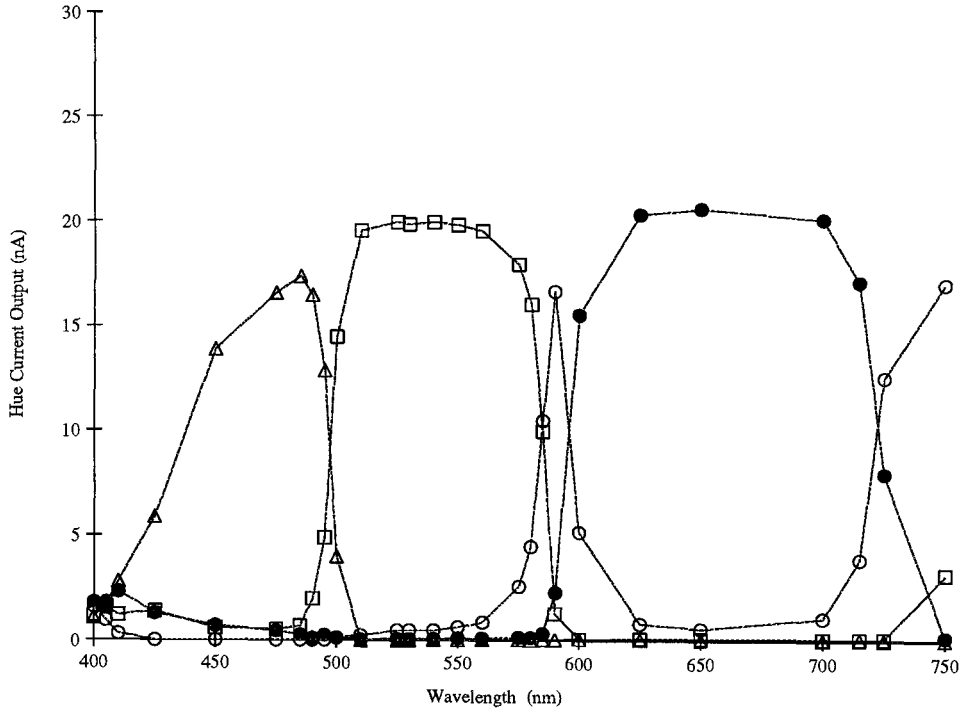
Fig. 15. Block diagram of hue sensor element. RGB current inputs are converted into opponency currents which are then converted into x and y components of hue in the translinear normalization circuit. The opponency circuit gives the *additive/shift* invariance property while the translinear part given *multiplicative/scale* invariance. Opponency X and Y currents in the plane are mapped onto the unit diamond shown in the insert. The unit diamond is defined by $|X| + |Y| = I_{bias} = $ const. In the insert, even though point BB is roughly twice as intense as point AA, their mappings to B and A respectively on the unit diamond indicate their hue similarities.

input signal: R, G, B values are limitless in the analog sensor domain, but are constrained and preprocessed to remain within a fixed domain (usually 8 bits) with traditional imaging systems.

As expected, figure 16 shows that the hue circuit output varies little with changing illumination levels. When the image irradiance is scaled by a factor of 10, the x and y currents scale only by at most 25%, in a manner similar to the normalized color circuit.

Finally, we numerically computed off-chip the hue defined by equation (33) and the measured x and y currents. For comparison, we additionally superimpose the calculated hue based on photosensor input and human visual input in figure 17. The former is computed from

(a)

(b)

Fig. 16. Spectral response of the hue circuit at (a) standard intensity (arbitrarily set) and (b) 10 times standard intensity. Similarities of the half-wave rectified hue currents x^+ (●), x^- (□), y^+ (○), y^- (△) in (a) and (b) indicate that the circuit performs the normalization and thus exhibits *multiplicative/scale* invariance. Here an order of magnitude increase in illumination causes a maximum change of only 25% of current output.
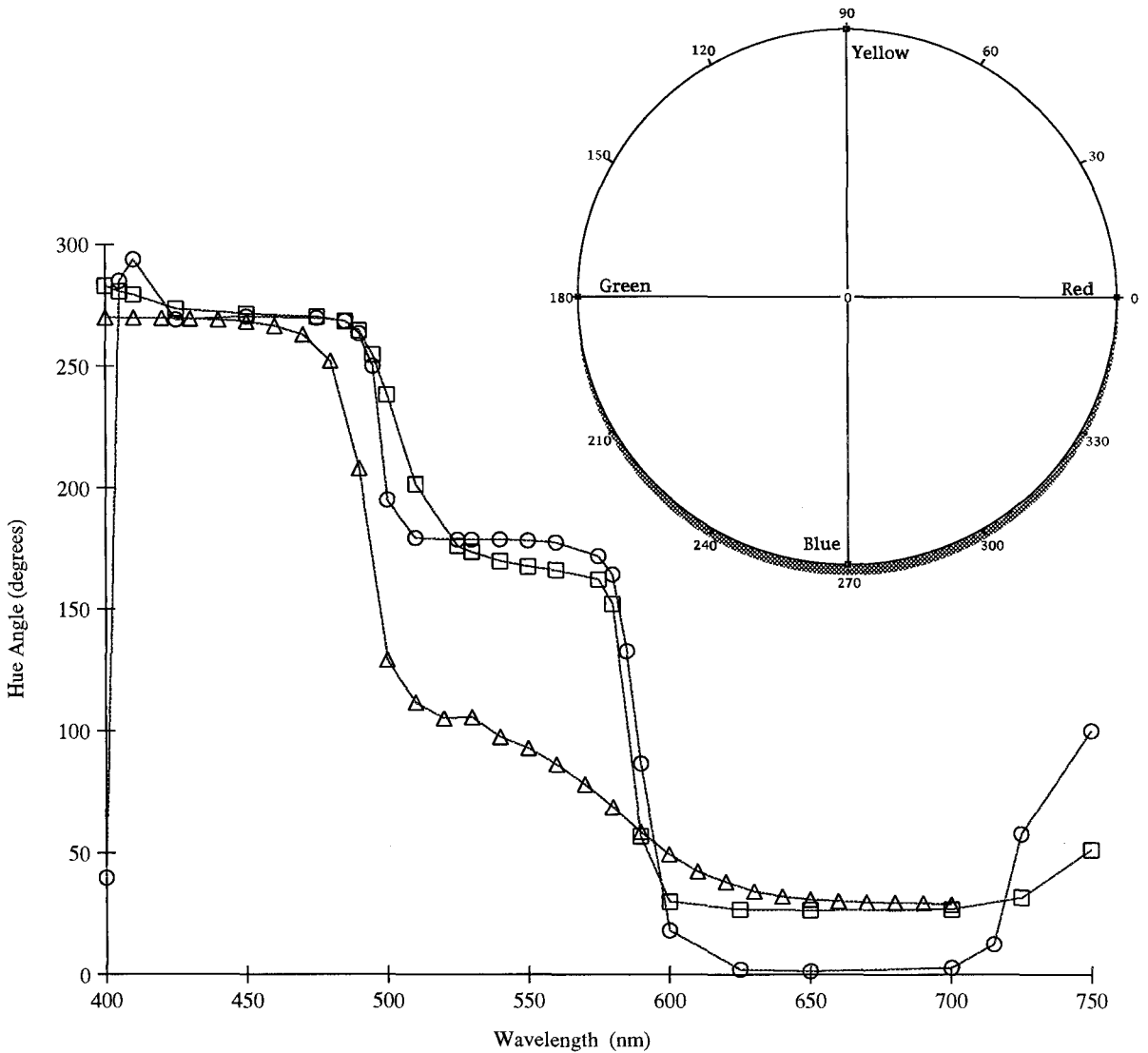
Fig. 17. Comparison of hue as computed using (1) the *x* and *y* output currents of the hue chip (O), (2) the spectral response curves of the R, G, B filtered phototransistors (□), and (3) the spectral response curves (Ingling 1977) from human cones (△). If we discount the borders of the spectral range (near 400 nm) where our experimental setup has low signal content, all curves exhibit monotonically unique hue values in the visible range. The flat range in the band 500 nm to 600 nm for the hue chip is predicted from the color photosensor input. The hue response curve suggests that this sensor would have higher spectral resolution in the transition regions: 475 nm to 525 nm, 575 nm to 625 nm, and upwards of 700 nm. The color wheel in the insert shows the color interpretation of the hue angle.[8]

the measured output photocurrents from R, G, B color filtered phototransistors, while the latter is computed from the equivalent R, G, B spectral-response curves (Ingling 1977) of the human cone system. In both cases, hue is calculated from arctan $[(R + G - 2B)/(2R - 2G)]$. In the visible range (>400 to 700 nm) all curves give monotonic hue values. The hue chip output agrees within experimental error to the calculated hue based on phototransistor current input. Both have a flattened

output spectral response in the 500 to 600 nm band. Because they have different input sensor characteristics, hue calculated from human visual curves does not have this flattened output spectral response.

We select the output of the chip to be the *x* and *y* currents, rather than the single variable hue, because we can easily compute them within analog circuitry. This will enable us in future chips to segment visual scenes based on the *x* and *y* components of hue in a

straightforward and, more importantly, achievable manner in analog VLSI (rather than trying to segment images based on hue using a modular operation on chip).

6.4 Future Hardware Issues

In our current unoptimized circuit design, one pixel, including three (RGB) phototransistors as well as normalized RGB, intensity, opponency, and hue circuits, spans 350 μm by 1600 μm on the die, most of which is inactive silicon real estate. This odd geometrical shape was selected to facilitate color filter placement. Without circuit optimization, a 5 by 5 pixel array can be implemented onto the smallest available chip (using a 2 μm process), while significantly larger array sizes are possible with circuit optimization, bigger die sizes and smaller design rules. Because phototransistors take up 70% of the active silicon real estate area, their design optimization will have the biggest impact on circuit compaction. Still, mechanical spectral filter placement will be difficult to achieve so a method to place the filters easily and accurately is desired. One method is to fabricate color array dyes using a lithographic process and bond to the chip (Dillon 1978). A lower-quality but cheaper version of this method is to use color slide film as the color filter array.

We envision two approaches for designing a hue circuit that locates discontinuities in hue space as discussed in the first part of this article. One is to convert the normalized $x^+x^-y^+y^-$ currents into voltages and utilize available circuits for detecting discontinuities (Harris 1990a). Another approach is to develop current mode versions of discontinuity detectors such as resistive fuses. The advantage of the current mode fuse over voltage mode is less use of silicon real estate for a more compact circuit. Although we do not yet have a complete hue discontinuity sensor, we have all the necessary circuit-building blocks for this task.

7 Conclusion

In this study we argue that edges in *hue* correlate more directly with material boundaries than edges in intensity, RGB, or Nrgb space. In particular, given the properties of *additive/shift* and *multiplicative/scale* invariance, *hue edges* are invariant to particular types of shadows, highlights, and transparency. We illustrate this behavior using video-acquired color images. Segmentation in images containing low saturation and intensity image values can be improved with the help of a smoothing operation. Using an intensity edge map as a starting point and applying a first-order smoothness operator to the hue map results in an edge map that discounts confounding cues. The main advantage of utilizing this operation is the straightforward application to analog VLSI hardware.

We also developed modulo operators to enhance traditional image processing algorithms for segmentation of variables defined on a circular space. By utilizing the HSI space for color segmentation, scene segmentation is enhanced in real images containing high saturation, even in the presence of confounding cues due to shading, transparency, shadows, and highlights. In the shadowed Mondrian of figure 2, hue edges in HSI space corresponded more closely to material boundaries than either hue edges in the CIE space or other edge types. While modulo operators are elegant they are not feasible for analog VLSI implementation. However, they may have utilization in custom digital hardware.

We greatly favor the simple definition of hue employed in the HSI space over the definition of hue in the CIE spaces, since the former is relatively straightforward to implement in analog electronic integrated circuits. On that basis we have proceeded to build electronic test circuits to facilitate color processing. The first electronic integrated structure to be manufactured and tested was the Nrgb color sensor (Perez 1992a). The next electronic structure was the hue color sensor. Tests on this sensor show correct performance to varying illumination. Future analog circuits will be constructed to evaluate hue boundaries for segmentation.

In designing artificial vision systems, there is some benefit to carefully selecting spectral filters such that the *integrated white* condition of equation (20) (or more generally equation (19)) is satisfied. In doing so, the RGB sensors will ultimately construct a hue space that is invariant to highlights, shadows, and surface orientation.

In comparison to physically based models with their associated color clustering and histogramming, the method developed here is comparable in performance yet simpler in concept and execution. It is also more amenable to analog VLSI hardware implementation.

Notes

1. Note that due to the wrap-around nature of hue, low values of hue are near high ones. For example, a hue value of 0 is 3 units removed from 253.
2. Bajcsy (1990) utilizes a similar concept called *illumination whitening*.
3. This is viable in industrial applications where the illumination and the filter characteristics are controlled such that equation (19) is satisfied.
4. The implication of this equation is that the diffuse illumination can be arbitrarily colored and not violate the *integrated white* condition. Segmentation will be achieved for these conditions.
5. In 2-D scenes this definition of d will give an upper bound on F. Also, Pratt chooses $\alpha = \frac{1}{9}$ to penalize offset edges more than smeared edges.
6. The *silicon retina* requires less than 1 mW of power, most of which is used in the photo-conversion stage.
7. In particular, we avoid the $\sqrt{3}$ scaling required from the use of equation (4).
8. Missing representation in the 270 to 360 degree range for hue angles is an affirmation that monochromatic light cannot produce purple hues.

References

Abdou, I.E., and Pratt, W.K. 1979. Quantitative design and evaluation of enhancement/thresholding edge detectors, *Proc. IEEE* 67(5): 753–763.

Bajcsy, R., Lee, S.W., and Leonardis, A. 1990. Color image segmentation with detection of highlights and local illumination induced by inter-reflections. *Proc. 10th Intern. Conf. Patt. Recog. B*, Atlantic City, pp. 785–790.

Barth, M., Parthasarathy, S., Wang, J., Hu, E., Hackwood, S., and Beni, G. 1986. A color vision system for microelectronics: Application to oxide thickness measurement, *Proc. Inter. Conf. Robot. Autom.*, San Francisco, pp. 1242–1247.

Berlin, B., and Kay, P. 1969. *Basic Color Terms: Their Universality and Evolution*, University of California.

Canny, J. 1986. A computational approach to edge detection. *IEEE Tran. Patt. Anal. Mach. Intell.* 8(6): 679–698.

Celenk, M. 1990. A color clustering technique for image segmentation, *Comput. Vis., Graph. Image Process.* 52: 145–170.

Cook, R.L., and Torrance, K.E. 1981. A reflectance model for computer graphics, *Computer Graphics* 15(3): 307–316.

Daily, M.J. 1989. Color image segmentation using Markov random fields, *Proc. Conf. Comput. Vis. Patt. Recog.*, San Diego, pp. 304–312.

Delbruck, T. 1993. Investigations of analog VLSI visual transduction and motion processing, Ph.D. Thesis, California Institute of Technology.

Desimone, R., Schein, S.J., Moran, J., and Ungerleider, L.G. 1985. Contour, color and shape analysis beyond the striate cortex, *Vision Research* 25: 441.

De Valois, R.L., and De Valois, K.K. 1975. Neural coding of color, *Handbook of Perception*, vol. V *Seeing*, Academic Press: San Diego, pp. 117–166.

Dillon, P., Brault, A., Horak, J., Garcia, E., Martin, T., and Light, W. 1985. Fabrication and performance of color filter arrays for solid-state imagers, *IEEE Trans. Electron Devices* 25(2): 97–101.

Foley, J.D., VanDam, A., Feiner, S.K., and Hughes, J.F. 1990. *Computer Graphics: Principles and Practice*, Addison-Wesley: Reading, MA.

Geiger, D., and Girosi, F. 1990. Parallel and deterministic algorithms from MRFs: Surface reconstruction, *IEEE Trans. Patt. Anal. Mach. Intell.* 13(5): 401–412.

Geman, S., and Geman, D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Patt. Anal. Mach. Intell.* 5: 721–741.

Genz, S.E. 1990. Real time chip set simplifies color image processing, *Image Processing Handbook* (8): 56–59, Data Translation™, Marlboro, MA.

Gilbert, B. 1975. Translinear circuits: A proposed classification, *Electronics Letters* 11(1): 14–16.

Gershon, R. 1985. Aspects of perception and computation in color vision, *Comput. Vis. Graph. Image Process.* 32: 245–278.

Gershon, R., Jepson, A., and Tsotsos, J. 1986. Ambient illumination and the determination of material changes, *J. Opt. Soc. Amer. A* 3(10): 1700–1707.

Harris, J.G., Koch, C., and Luo, J. 1990a. A two-dimensional analog VLSI circuit for detecting discontinuities in early vision, *Science* 248: 1209–1211.

Harris, J.G., Koch, C., Staats, E., and Luo, J. 1990b. Analog hardware for detecting discontinuities in early vision, *Intern. J. Comput. Vis.* 4: 211–233.

Healey, G. 1989. Using color for geometry-insensitive segmentation, *J. Opt. Soc. Amer. A* 6(6): 920–937.

Healey, G., and Binford, T.O., 1987. The role and use of color in a general vision system. *DARPA-Image Understanding Workshop*, Los Angeles, pp. 599–613.

Hurlbert, A., and Poggio, T. 1989. A network for image segmentation using color, In *Advances in Neural Information Processing Systems I*, pp. 297–304, Morgan Kaufmann: San Mateo.

Ingling, C.R., and Tsou, B.H. 1977. Orthogonal combination of the three visual channels, *Vision Research* 17: 1075–1082.

Joblove, G.H., and Green, D. 1978. Color spaces for computer graphics, *Computer Graphics* 12(3): 20–25.

Jain, A.K. 1989. *Fundamentals of Digital Image Processing*, Prentice-Hall: Englewood Cliffs, NJ.

Kender, J. 1976. Saturation, hue, and normalized color: Calculation, digitization, and use, Computer Science Technical Report, Carnegie-Mellon University.

Klinker, G.J., Shafer, S.A., and Kanade, T. 1988. The measurement of highlights in color images, *Intern. J. Comput. Vis.* 2: 7–32.

Klinker, G.J., Shafer, S.A., and Kanade, T. 1990. A physical approach to color image understanding, *Intern. J. Comput. Vis.* 4: 7–38.

Koch, C. 1989. Seeing chips: Analog VLSI circuits for computer vision, *Neural Computation*, 1: 184–200.

Koch, C., Moore, A., Bair, W., Horiuchi, T., Bishofberger, B., and Lazzaro, J. 1991. Computing motion using analog VLSI vision chips: An experimental comparison among four approaches, *IEEE Workshop on Visual Motion*, Princeton, October, pp. 312–324.

Lenny, P., and D'Zmura, M. 1988. Mechanisms of color vision, *CRC Crit. Rev. Neurobiol.* 3(4): 333–400.

Mead, C. 1989. *Analog VLSI and Neural Systems*, Addison-Wesley: Reading, MA.

Nevatia, R. 1977. A color edge detector and its use in scene segmentation, *IEEE Trans. Syst. Man, Cyb.* 7(11): 820–826.

Ohlander, R.B. 1976. Analysis of Natural Scenes, Ph.D. thesis, Carnegie Mellon University.

Ohta, Y., Kanade, T., and Sakai, T. 1980. Color information for region segmentation, *Comput. Graph. Image Process.* 13: 222–241.

Perez, F.A., and Koch, C. 1992. Toward color image segmentation in analog VLSI, *Rockwell 4th Annu. Tech. Conf. Cont. Sig. Process.* pp. 246–263.

Perez, F.A., and Koch, C. 1992. Segmenting color images using hue, CNS Memo 20, California Institute of Technology, 16 October.

Phong, B.T. 1975. Illumination for computer generated pictures, *Commu. ACM* 18(6): 311–317.

Poggio, T., Torre, V., and Koch, C. 1985. Computational vision and regularization theory, *Nature* 317(6035): 314–319.

Poggio, T., Gamble, E.B., and Little, J.J. 1988. Parallel integration of vision modules, *Science* 242: 436–440.

Rubin, J.M., and Richards, W.A. 1982. Color vision and image intensities: When are changes material? *Biological Cybernetics* 45: 215–226.

Rubin, J.M., and Richards, W.A. 1984. Color vision: Representing material categories, AI Memo No. 764, Massachusetts Institute of Technology.

Schwarz, M.W., Cowan, W.B., and Beatty, J.C. 1987. Experimental comparison of RGB, YIQ, LAB, and opponent color models, *ACM Trans. Graphics* 6(2): 123–158.

Seevink, E. 1988. Analysis and synthesis of translinear integrated circuits, Elsevier: New York.

Shafer, S.A. 1985. Using color to separate reflection components, *COLOR Res. Appl.* 10(4): 210–218.

Sivilotti, M.A., Mahowald, and Mead, C.A. 1987. Real-time visual computation using analog CMOS processing arrays, *1987 Stanford Conf. VLSI*, MIT Press: Cambridge, pp. 295–312.

Smith, A.R. 1978. Color gamut transform pairs, *Computer Graphics* 12(3): 12–19.

Terzopoulos, D. 1985. Computing visible-surface representations, AI Memo No. 800, Massachusetts Institute of Technology.

Tominaga, S. 1987. Expansion of color images using three perceptual attributes, *Patt. Recog. Lett.* 6: 77–85.

Tominaga, S. 1990. A color classification method for color images using a uniform color space, *10th Intern. Conf. Patt. Recog.*, Atlantic City, pp. 803–807.

Toumazou, C., Lidgey, F., and Haigh, D, eds. 1990. *Analogue IC Design: The Current Mode Approach*, Short Run Press, England.

Wolfe, W.L., and Zissis, G.J. 1985. *The Infrared Handbook*, Environmental Research Institute of Michigan: Ann Arbor.

Wright, W.A. 1989. A Markov random field approach to data fusion and colour segmentation, *Image Vis. Comput.* 7(2): 144–150.

Wyszecki, G., and Stiles, W.S. 1982. *Color Science: Concepts and Methods, Quantitative Data and Formulae*, Wiley: New York.

Zeki, S. 1983. Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours, *Neuroscience* 9(4): 741–765.