Method for estimating scene parameters from color histograms

Carol L. Novak

Siemens Corporate Research, 755 College Road East, Princeton, New Jersey 08540

Steven A. Shafer

School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Received July 22, 1993; revised manuscript received June 6, 1994; accepted June 8, 1994

One of the key tools in applying physics-based models to machine vision has been the analysis of color histograms. In the mid-1980's it was recognized that the color histogram for a single inhomogeneous surface with highlights will have a planar distribution in color space. It has since been shown that the colors do not fall randomly in a plane but form clusters at specific points. Physics-based models of reflection predict that the shape of the histogram is related not only to the illumination color and the object color but also to such noncolor properties as surface roughness and imaging geometry. We present an algorithm for analyzing color histograms that yields estimates of surface roughness, phase angle between the camera and the light source, and illumination intensity. These three scene parameters are related to three histogram measurements. However, the relationship is complex and cannot be solved analytically. Therefore we developed a method for estimating these properties that is based on interpolation between histograms that come from images of known scene properties. We present tests of our algorithm on simulated data, and the results compare well with the known simulation parameters. We also test our method on real images, and the results compare favorably with the actual parameters estimated by other means. Our method for estimating scene properties is very fast and requires only a single color image.

1. INTRODUCTION

Color histograms have long been used by the machine vision community in image understanding. Color is usually thought of as an important property of objects and is often used for segmentation and classification. Unfortunately, color is not uniform for all objects of a given class or even across a single object. Color variation has come to be expected in real images, and vision researchers have been working on modeling this variation.

The earliest uses of color histograms modeled the histogram as a Gaussian cluster in color space.' For example, pixels that correspond to grass were modeled as having a canonical color of green with some possible deviation from this color. The color variation was modeled as a probability distribution, so that the further a pixel was from the characteristic color, the less likely that it corresponded to grass. Figure 1 shows a diagram of this idea.

In 1985, Shafer showed that, for dielectric materials with highlights from a single light source, the color histogram associated with a single object forms a plane.[2] This plane is defined by two color vectors: a body-reflection vector and a surface-reflection vector. Every pixel's color is a linear combination of these two colors. In a diagram accompanying this idea (Fig. 2), Shafer presented his visualization that the histogram would fill out a parallelogram. Reference 2 does not contain a description of how to determine these two vectors. In fact, for any planar histogram, there are many sets of two vectors that could define it.

In 1987, Klinker and Gershon independently observed that the color histogram for a dielectric illuminated by a single source forms a T shape or a dogleg in color space.[3,4] Figure 3 shows that this color histogram is composed of two linear clusters, one corresponding to pixels that exhibit mostly body reflection and one corresponding to pixels that exhibit mostly surface reflection. This T shape made it possible to identify characteristic body reflection and illumination colors. In 1988, Healey showed that the number of dimensions occupied by the histogram may be used to distinguish metals from dielectrics.[5]

In a previous study we showed that color histograms have identifiable features that depend in a precise mathematical way on such noncolor scene properties as surface roughness and imaging geometry.[6] In this paper we show that three scene properties, namely, the illumination intensity, the roughness of the surface, and the phase angle between camera and light source, may be recovered from three measurements of the histogram shape.

However, the functions that relate the scene properties to the histogram measurements are interdependent and highly nonlinear, involving trigonometric and exponential functions. Because an analytic solution is not possible, we developed a method that interpolates between data from a lookup table. The lookup table covers a range of scene parameters and records measurements of the histograms that result from each combination of these scene parameters. Then a polynomial function is fitted to the data to approximate the relationships between scene parameters and histogram shape.

Our research also shows how the colors observed in a highlight depend not only on the scene colors but also on the surface roughness and the imaging geom-

0740-3232/94/113020-17$06.00 © 1994 Optical Society of America

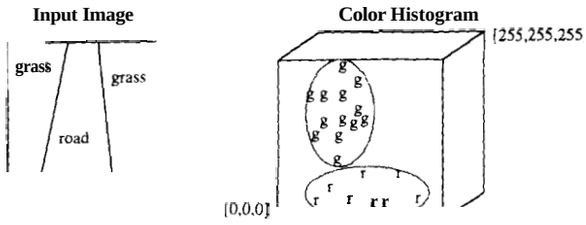

Fig. 1. Gaussian histogram

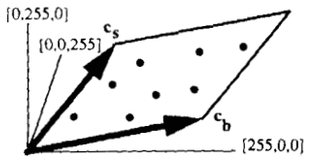

Fig. 2. Planar histogram

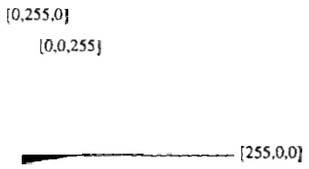

Fig. 3. T-shaped histogram

etry. Our estimates of these scene parameters enable us to improve our initial estimate of the illumination color. This capability, along with the estimate of the illumination intensity, allows us to discount the effect of the illuminant to recover estimates of the object's reflectance. Section 2 explains the relationship between the color histogram features and the various scene parameters. Section 3 presents an algorithm for computing estimates of these parameters from the histogram. We also describe the development of our algorithm for an ideal camera. Section 4 shows how the algorithm has been extended to accommodate real cameras. Section 5 presents our results from experiments on real images.

2. UNDERSTANDING COLOR HISTOGRAMS

When we use the term color histogram, we are talking about a distribution of colors in the three-dimensional red–green–blue space. For a typical imaging system with 8 bits for each color band, there are 256^3 bins into which a pixel may fall. In this study we consider only whether a bin is full or empty. We do not use a fourth dimension to display the number of pixels that have a particular red–green–blue value. A fourth dimension would be difficult to visualize but, more significantly, would also be dependent on such things as object size and shape.

In our research we use the dichromatic reflection model, which states that the light L reflected from a dielectric object is the sum of two components: a surface component L_s and a body component L_b (Ref. 2):

$$L(\lambda, \theta_i, \theta_r, \theta_p) = L_b(\lambda, \theta_i, \theta_r, \theta_p) + L_s(\lambda, \theta_i, \theta_r, \theta_p). \quad (1)$$

Each of the two components is a function of the wavelength of light (λ) and the angles of incidence (θ_i), reflection (θ_r), and phase angle (θ_p). The dichromatic reflection model further states that each component in turn may be separated into a color term c that depends only on (λ) and a magnitude term m that depends only on (θ_i), (θ_r), and (θ_p):

$$L(\lambda, \theta_i, \theta_r, \theta_p) = c_b(\lambda)m_b(\theta_i, \theta_r, \theta_p) + c_s(\lambda)m_s(\theta_i, \theta_r, \theta_p). \quad (2)$$

In a color camera the light is typically sampled by sensors with red, green, and blue filters to yield a color pixel vector $\boldsymbol{c}$. If the sensors have responsivity $s^f(\lambda)$ (where f is one of red, green, or blue), then the elements of the color vector are formed:

$$c^f = \int L(\lambda, \theta_i, \theta_r, \theta_p)s^f(\lambda)\mathrm{d}\lambda. \quad (3)$$

Because tristimulus integration is a linear transformation, the color vector $\mathbf{c}$ may also be separated into body-reflection and surface-reflection components and may be decomposed into color vectors $\mathbf{c}_b$ and $\mathbf{c}_s$ and magnitudes m_b and m_s:

$$\mathbf{c} = m_b\mathbf{c}_b + m_s\mathbf{c}_s. \quad (4)$$

We require that the color vectors $\mathbf{c}_b$ and $\mathbf{c}_s$ be of length 1; as a result, the magnitude term must account for the overall brightness of the light coming from the surface, as well as for the gain of the sensor in transforming the light into pixel values.

Figure 4 contains a sketch of a typical color histogram for a dielectric surface illuminated by a single light source. As labeled, the histogram has two linear clusters of pixels: the body-reflection cluster and the highlight cluster. The first of these clusters extends from the black corner of the cube (point a) to the point of maximum body reflection (point b). The other cluster starts somewhere along the body-reflection cluster (point c) and extends to the highlight maximum (point d).

If the object has a complete blackbody color, there will be no body-reflection component and hence no body-reflection cluster. If the body-reflection color is the same hue as the surface-reflection color, the body-reflection cluster and the highlight cluster will be collinear. This is the case for white and gray objects because neither their body color nor their surface color imparts any hue to reflected light. Therefore objects of these colors cannot be analyzed by this type of method.

A. Body-Reflection Cluster

The linear cluster that we call the body-reflection cluster corresponds to pixels that exhibit mostly body reflection

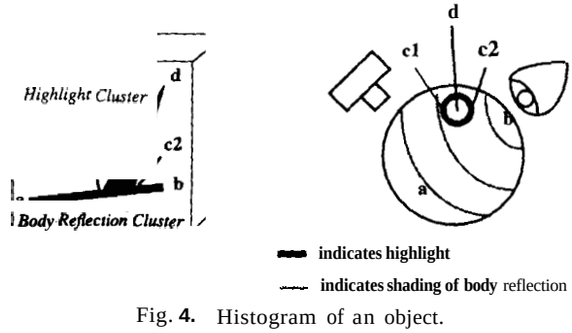

Fig. 4. Histogram of an object.

with very little surface reflection. If there is no ambient illumination in the scene, this cluster begins at the black point of the color cube (point a in Fig. 4), corresponding to points on the surface whose normal is 90° or more away from the direction of the illumination. The point at the other extreme of the body-reflection cluster (point b) corresponds to the largest amount of body reflection seen anywhere on the object. If we assume that the body-reflection component is Lambertian, the magnitude term will obey the equation

$$m_b = \gamma B_b \cos(\theta_i), \tag{5}$$

where θ_i is the angle of illumination incidence. The gain of the camera in converting photons measured by the CCD array into pixel values is represented by γ. The brightness of the body reflection is represented by the term B_b. This term factors in both the reflectance of the object (albedo) and the intensity of the light.

In the Lambertian model the magnitude of the body reflection is proportional to the cosine of the incidence angle, so pixels located halfway along the body-reflection cluster correspond to surface points with normals that are $\arccos(1/2)$ or 60" away from the illumination direction. If the object exhibits all possible surface normals the body-reflection cluster will be full length and densely filled. If the object is composed of a small number of flat surfaces there will be gaps in the body-reflection cluster. For this paper we assume that the objects that we are viewing have a broad, continuous distribution of surface normals.

A vector fitted to the body-reflection cluster (from point a to point b in Fig. 4) will point in the direction of the body-reflection color, which is the product of the object color and the illumination color. Once the illumination color has been determined by analysis of the highlight, the object color alone may be estimated, as proposed in some color constancy methods.[7-10]

If we assume that there is on the object some point that is visible to the camera and that has a surface normal pointing directly at the light source, then at that point $\cos(\theta_i) = 1$. This means that the length of the fitted vector (the magnitude $|\overline{ab}|$) corresponds to the gain γ times the object's apparent brightness B_b. If the intensity of the illumination can also be recovered by highlight analysis, then the albedo of the object can be separated out from the object's apparent brightness (assuming that the gain of the camera had been calibrated). Otherwise it will be impossible to distinguish a bright light shining on a dark surface from a dim light shining on a bright surface. Fortunately, highlights provide an invaluable clue in distinguishing between these cases.

B. Highlight Cluster

The cluster of pixels that we call the highlight cluster corresponds to pixels that show a nonnegligible amount of surface reflection. This cluster corresponds exactly to the area of the image that we would call the highlight. In the histogram, the highlight cluster starts where this cluster intersects with the body-reflection cluster (point c in Fig. 4) and extends upward from there to the brightest point of the highlight (point d).

In this presentation we use the Torrance–Sparrow model of scattering." This theory models a surface as a collection of tiny facets, each of which may have a local surface normal that is different from the global surface normal. The distribution of facet normals is modeled as Gaussian, with σ describing the standard deviation. Smooth surfaces will have a small standard deviation, whereas rougher surfaces will have a larger standard deviation. The facets are larger than the wavelength of visible light but are too small to be seen as texture. We refer to roughness on this small scale as optical roughness. We assume that the facet size is a constant for the surfaces in which we are interested.

The equation that we use for scattering gives the amount of surface reflection as

$$m_s = \gamma B_s \frac{FG(\theta_i, \theta_r, \theta_p)\alpha}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s^2}{2\sigma^2}\right), \tag{6}$$

where θ_s is the off-specular angle and θ_r is the angle of reflection. B_s is the intensity of the illumination, y is the camera gain, and α is a constant that includes the facet size (a variable in the original Torrance–Sparrow model). G is an attenuation factor that depends on geometry and that comes into play at grazing angles. G is a complicated function of incidence angle, reflection angle, and phase angle, and we do not reproduce it here; see Ref. 11 for details.

F is the Fresnel coefficient that describes the percentage of the light that is reflected at the interface; this coefficient is a function of geometry, wavelength, polarization state, and index of refraction of the material in question. However, it is very weakly dependent on incident angle and wavelength (over the visible range), so we will follow the neutral interface reflection model and assume that F is constant for a given material." Furthermore, for a wide range of plastics and paints, the indices of refraction are very nearly identical. Henceforth we assume that materials have an index of refraction of 1.5, corresponding to 4.0% Fresnel reflection.

1. *Length of Highlight Cluster*

When looking at highlights on a variety of surfaces, we can quickly observe that highlights are brighter and sharper on some surfaces and dimmer and more diffused on other surfaces. (It is the fact that highlights may be diffused that causes us to eschew the term diffuse reflection and to use the term body reflection instead.) Very smooth surfaces exhibit only a tiny amount of scattering of the surface reflection, whereas very rough surfaces have a great deal of scattering. This scattering of surface reflection is a result of the optical roughness of the surface.

We can see from Eq. (6) that the sharpness of the peak is determined by the standard deviation of facet angles σ and that the height of the peak is inversely proportional to σ. Intuitively this result makes sense because surface reflection scattered over a very small area will be more concentrated. A smooth object will have a small standard deviation of facet slopes, σ, resulting in a long highlight cluster. A rough object will have a large σ and thus will exhibit a shorter cluster. Figure 5 shows a plot of the length of the highlight cluster versus the object's roughness for simulated images where all other factors have been held constant. The horizontal axis indicates the standard deviation of facet angles (in degrees),

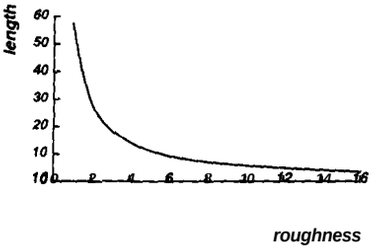

Fig. 5. Dependence of highlight-cluster length on object roughness.

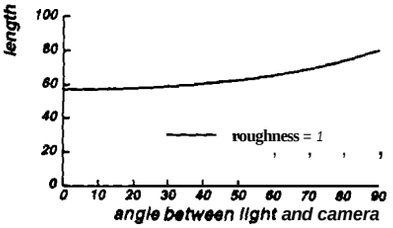

Fig. 6. Dependence of cluster length on imaging geometry.

which is our measure of roughness. The vertical axis indicates the Euclidean distance in red–green–blue space from point c to point d (see Fig. 4).

Equation (6) indicates that the intensity of the light source B_s also affects the magnitude of the surface reflection and thus the length of the highlight cluster. It is obvious from **Eq.** (6) that the length is directly proportional to this brightness. It will be assumed that the gain of the camera γ has been calibrated and that the facet size σ is known. If not, then only the overall gain $\gamma\alpha B_s$ could be recovered from the histogram.

The graph presented in Fig. 5 was calculated for the imaging geometry in which the light source and the camera are separated by a phase angle of 0" with respect to the object. However, **Eq. (6)** predicts that the imaging geometry will have an effect on highlight magnitude, as indicated by the $\cos(\theta_r)$ term in the denominator and the attenuation term G in the numerator. Figure 6 shows how the length of the highlight cluster changes as the camera and the light source are separated by different angles with respect to the object (while the roughness is held constant at 1). This figure demonstrates that the length changes slowly as the imaging geometry changes. The effect is small but noticeable, so imaging geometry must be considered to make an accurate estimate of surface roughness.

2. *Width* of *Highlight* Cluster

Another difference between histograms for smooth and rough surfaces is the width of the highlight cluster where it meets the body-reflection cluster (the distance from point c_1 to point c_2 in Fig. 4). The highlight cluster will be wider for rougher surfaces and narrower for smoother surfaces. This is because rougher objects will scatter surface reflection more widely, over a larger number of reflection angles.

In the color histogram a noticeable amount of surface reflection results in pixels that are displaced from the body cluster in the direction of the illumination color. If we take any highlight pixel and project along the surface-color vector onto the body-reflection vector, we can tell how much body reflection is present in that pixel. If we consider all the pixels in the highlight area of the image and look at how much body reflection is in each of them, we will obtain some range of body-reflection magnitudes. If the surface is very smooth with a sharp, bright highlight, this range will be small. However, if we consider a rougher object with a more diffused highlight, the range of body-reflection magnitudes will be larger because the highlight is spread over a larger number of surface normals.

This property is independent of object size or shape. It simply shows the variation in surface normals over the area of the highlight. We do not have to fit a surface shape to the image to know how much scattering the object exhibits. It is encoded right there in the histogram.

We simulated objects with different roughness values and identified those pixels that showed surface reflection. We calculated the body reflection for each of these points and computed the variation. This variation was divided by the overall length of the body-reflection vector to yield a fraction (the length of $\overline{c_1c_2}$ divided by the length of $\overline{ab}$ in Fig. 4). A fraction of 0.5 would mean that the highlight cluster's base extended across half the length of the body-reflection cluster. Figure 7 shows how the highlight-cluster width varies with the surface roughness if a point light source is used.

The brightness of the illumination B_s will also have some effect on the width of the highlight cluster. As the light intensity is increased, pcints on the surface that had amounts of surface reflection too small to be noticed may become bright enough to be included with the highlight pixels. Clearly the width will grow as the light intensity grows. However, the growth is not linear because width measures the variation in body reflection amounts over the area of the highlight rather than the actual intensity of the highlight. In simulations we found that the highlight width grows very slowly as the illumination intensity increases.[13] Consequently, although roughness affects both the length and the width of the highlight cluster, changes in illumination intensity primarily affect the length. Therefore it should be possible to distinguish a bright source illuminating a rough object from a dim source illuminating a shiny one.

Although the width of the highlight cluster does not depend on the object's size and shape, it does depend on the imaging geometry. To understand why this is so, imagine a highlight that spreads 15" in every direction from its maximum. If the camera and the light source are separated by 30" the perfect specular angle will be at 15° with respect to the illumination direction. The highlight will spread over points with surface normals ranging from 0° to 30". [For ease of explanation, we ignore the influence of the $1/\cos(\theta_r)$ term.] The amount

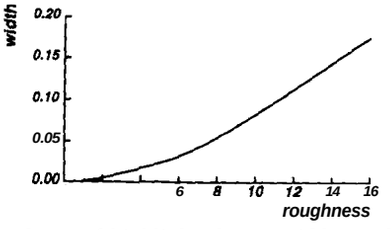

Fig. 7. Dependence of highlight-cluster width on object roughness.

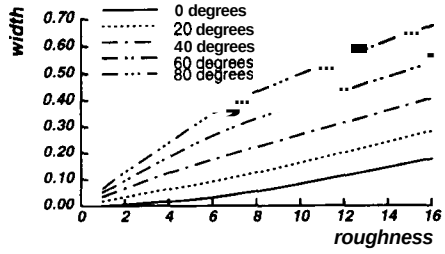

Fig. 8. Dependence of highlight width on roughness for different imaging geometries.

of body reflection at these points will vary from $\cos(0) = 1.0$ to $\cos(30) = 0.87$, a width of 0.13. If the camera and the light source are separated by 90°, the perfect specular angle will be at 45°, with the highlight spreading from 30° to 60". Then the amount of body reflection will vary from $\cos(30) = 0.87$ to $\cos(60) = 0.50$, a width of 0.37.

Figure 8 shows how the width of the highlight cluster varies with roughness for a variety of imaging geometries. The angle label is a measure of the angle (in degrees) that separates the light source and the camera (with respect to the object). We assume that the object is small enough and far enough from the camera that this angle is the same for all points on the surface.

For the case of highlight-cluster width, the measurement is very sensitive to different viewing geometries, so the phase angle between the camera and the light source must be known or estimated somehow. Thus it is particularly fortuitous that such an estimate can be made directly from the histogram itself, as we now describe.

3. *Intersection of Clusters*

When we introduced the diagram in Fig. **4,** we described the highlight cluster as beginning somewhere along the body-reflection cluster. Klinker derived the 50% heuristic, which stated that for a large range of viewing geometries the highlight cluster would start somewhere in the upper 50% of the body-reflection cluster.[9] Now we show how to pinpoint the location.

The distance along the body-reflection cluster where the two clusters intersect (the length of $\overline{ac}$ divided by the length of $\overline{ab}$ in Fig. **4)** shows the amount of body reflection at those points on the surface that are highlighted. Assuming that body reflection is Lambertian, the amount of body reflection is proportional to the cosine of the incidence angle. If the two clusters intersect at the maximum point on the body-reflection cluster, then the highlight occurs at those points that have the maximum amount of body reflection, where surface normals point directly at the light source. If the two clusters meet halfway along the body-reflection cluster, the highlight must occur at points with surface normal pointing $\arccos(1/2)$ or 60" away from the illumination direction.

If the body reflection is Lambertian, it does not depend in any way on the angle from which it is viewed. Thus the body reflection does not tell us anything about the camera direction. However, the surface reflection is dependent on both the illumination and the camera directions. If we ignore for a moment the $1/\cos(\theta_r)$ term in **Eq.** (6), we can see that the maximum amount of surface reflection will occur at those points on the surface where the angle of incidence equals the angle of reflection. Thus, if the highlight occurs at a point at which the surface normal faces 10" away from the light-source direction, the light source and the camera must be 20" apart with respect to that point on the surface.

Figure 9 graphically illustrates this phenomenon. The histograms have been projected into the plane defined by the body-reflection and the surface-reflection colors, so that the horizontal direction indicates increasing amounts of body reflection and the vertical direction indicates increasing amount of surface reflection. This time the amount of roughness has been held constant at $\sigma = 1$ while the angular separation of the light source and the camera has been varied from 0" to 80". This graph shows how the meeting point decreases as the angle separating the camera and the light source increases. Incidentally, this graph also shows how the length and the width of the cluster are affected by imaging geometry as described in Subsections 2.B.1 and **2.B.2.**

It does not matter whether the object has one highlight or many. If the object is small compared to the distance to the light source and the camera, the highlight(s) will always occur at points with the same surface normal for a given imaging geometry. Figure 10 shows what happens when we plot intersection versus imaging geometry. The horizontal axis shows the phase angle between the light source and the camera with respect to the object. The vertical axis is the intersection ratio $|\overline{ac}|/|\overline{ab}|$ (see Fig. 4).

The presence of the $1/\cos(\theta_r)$ term in **Eq.** (6) indicates that the maximum amount of surface reflection will not always occur precisely at the perfect specular angle. This is particularly true of rougher surfaces, on which the highlight is spread over a wide range of reflection angles so that $1/\cos(\theta_r)$ varies significantly. This causes the off-specular peaks described in Ref. 11. The result is that the intersection is very slightly dependent on the surface roughness (see Ref. 13 for more details).

4. *Direction* of *Highlight Cluster*

The highlight cluster is usually long and narrow in shape, and a vector can be fitted to it (from point c to point d in Fig. 4). Klinker argued that this vector will usually cor-

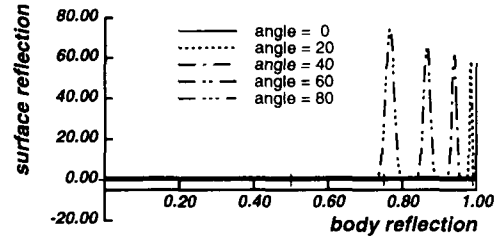

Fig. 9. Illustration of how highlight cluster varies with changing imaging geometry.

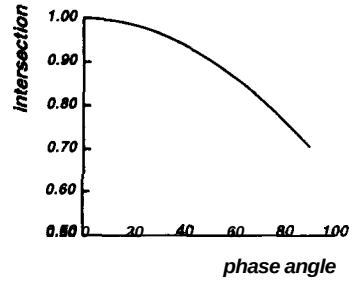

Fig. 10. Dependence of intersection point on imaging geometry.

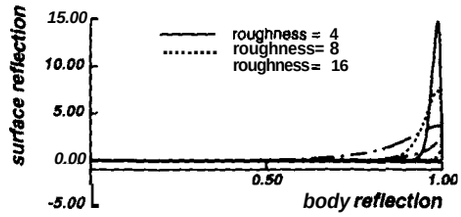

Fig. 11. Dependence of highlight-cluster direction on object roughness.

respond closely to the surface-reflection color.[9] This is true for smooth objects, where the highlight has a small area, and for imaging geometries, where the body reflection changes slowly over that area. In this case the amount of body reflection at the base of the highlight cluster and the amount at the tip varies by a small amount.

In contrast, if the object is optically rough and the highlight occurs on a part of the object where the cosine of the incidence angle changes more rapidly, then the amount of body reflection at the base of the highlight cluster may vary significantly from the amount at the tip. This condition has the effect of skewing the highlight cluster away from the direction of the illumination color, toward the body-reflection color. The estimate of the illumination color made from fitting a vector to this cluster will be somewhat inaccurate.

We can visualize this phenomenon by projecting the histogram onto the plane defined by the body-reflection color and the surface-reflection color. We simulated dichromatic reflection for a series of objects with increasing roughness values, but with body color held constant. Figure 11 shows a cross section of the histograms that result. The horizontal direction is the direction of increasing amounts of body reflection; all three histograms fall exactly on this line for body-reflection pixels. The vertical direction is defined by increasing amounts of surface reflection. For surfaces with a standard-deviation of facet angles (σ) equal to 4°, a vector fitted to the highlight cluster will point exactly in this direction. A vector fitted to the highlight cluster when $\sigma = 8$ will deviate slightly from the vertical direction. In the extreme case in which $\sigma = 16$, the vector will deviate dramatically.

In the color histogram the vectors describing the body-reflection color and the illumination color are not generally perpendicular. The angle between them depends on the hue difference between the two colors, which is not known in advance. If the vector fitted to the highlight cluster does not point exactly in the direction of increasing amounts of surface reflection, the estimate of illumination color will be inaccurate by some amount. This effect will in turn bias the estimate of the object color, which we obtain by dividing the body-reflection color by the illumination color.

Furthermore, skewing of the highlight cluster will also affect measurements of the cluster's length, width, and intersection. Figure 12 shows an enlargement of the histogram for the roughest surface, with key points labeled. The estimate of illumination color that we obtain by fitting a vector to this highlight cluster will give the direction of the light color as $\overrightarrow{cd}$, whereas the correct illumination color is defined by the vertical direction $\overrightarrow{c'd}$ (because the histogram was projected into the plane defined by the body- and the surface-reflection colors).

When the highlight cluster is skewed away from the direction of the illumination color, the cluster will appear longer than it would otherwise be. The amount of surface reflection at the brightest point is the vertical displacement in the graph (the distance from point c' to point d in Fig. 12) rather than the distance along the highlight cluster (the distance from point c to point d). Similarly, the width of the cluster is determined by projecting all pixels along the illumination-color vector (straight down for this graph). Thus the correct measure of highlight width is the variation in body reflection, which is the distance from c_1 to c_2' rather than the distance from c_1 to c_2. It follows, then, that the correct intersection ratio is given by $|\overline{ac'}|/|\overline{ab}|$ rather than $|\overline{ac}|/|\overline{ab}|$.

The vector fitted to the highlight cluster (from point c to point d in Fig. 4) is a good first estimate of the illumination color, but we now know that it may be skewed. If we know the surface roughness and the imaging geometry, we can calculate the amount of skewing and can compensate for it.

Recently Wolff showed that the Lambertian model is somewhat inaccurate for body reflection from many real surfaces, particularly at extreme reflection angles; he proposed a different model to replace it.[14] Also, some researchers prefer to use the Beckmann–Spizzichino model[15] or a combination of the Beckmann–Spizzichino and the Torrance–Sparrow models[16] to describe surface scattering. Although a choice of different models would change the exact positions of the graphs in this section, the general relationships would remain the same. These dependencies of the color histogram on scene parameters are a consequence of the different geometric distributions as well as the different hues of the two components of reflection.

In Subsections 2.B.1 and 2.B.2 we showed that the roughness of the object affects the length and the width of the highlight cluster but that there is some dependence on imaging geometry. Then in Subsection 2.B.3 we showed that the imaging geometry determines the intersection of the two clusters but that there is some dependence on roughness. Furthermore, the intensity of the illumination affects the length and the width of the highlight's cluster as well, although in different ways. The degree of dependence of each histogram measurement on the scene parameters is characterized in Table 1.

In this section we have shown that the direction of the highlight cluster can be skewed away from the direction of the illumination color, depending on both the amount of roughness and the imaging geometry. The highlight-cluster length, width, and intersection are defined with respect to the cluster direction, so measurements of these features also depend on the estimate of the illumina-

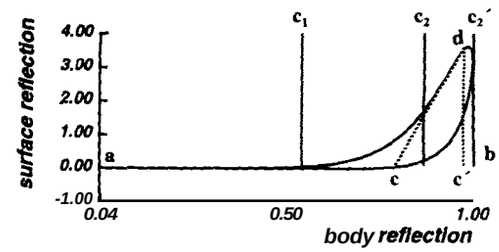

Fig. 12. Skewed highlight.

Table 1. Dependence of Histogram Features on Scene Parameters

Feature	Roughness	Phase Angle	Illumination Intensity
Length	Strong	Weak	Strong
Width	Strong	Strong	Weak
Intersection	Weak	Strong	None

tion color. If that direction changes, the other histogram measurements will also change.

Obviously these factors are all interdependent. Therefore we propose to solve for roughness, phase angle, and illumination intensity simultaneously, based on the measurements of the histogram.

3. ANALYZING COLOR HISTOGRAMS

Section 2 described the relationship between the shape of the color histogram and the various scene properties. Understanding the relationship is the first step in analyzing color histograms. The next step is determining how to exploit the histogram to recover quantitative measures of the scene properties.

In this section we assume an ideal sensor that has highly desirable properties. This kind of camera, which is unobtainable in practice, shows the limit of what can be obtained under the best possible imaging conditions. Specifically, we assume a noise-free camera with a linear response. We also assume that the camera is linear over the range of all possible pixel values that might occur in our images, so that we do not have to worry about clipping at some maximum value. In Section 4 we show how the program has been adapted for application to images obtained with real cameras.

A. Knowns and Unknowns

The known image parameters that can be measured from the color histogram are the body-reflection cluster's length and direction; the highlight cluster's length, width, and direction; and the intersection point of the two clusters. The result is four scalar values and two vector quantities, which are denoted by the following variables:

- l, length of highlight cluster;
- w, width of highlight cluster;
- i, intersection of two clusters;
- b, length of body-reflection cluster;
- $\mathbf{d}_s$, direction of highlight cluster;
- $\mathbf{d}_b$, direction of body-reflection cluster.

The unknown scene parameters that we would like to recover from the histogram can also be divided into scalar values and vector quantities. The scalar values are the surface roughness, the phase angle between the light-source and the camera directions, the illumination intensity, and the surface albedo. The vector quantities are the chromaticity of the illumination and of the object. These variables are

- σ, optical roughness;
- θ_p, phase angle;
- B_s, illumination intensity;
- B_o, object albedo;
- $\mathbf{C}_s$, chromaticity of light source;
- $\mathbf{C}_o$, object chromaticity under white light.

For convenience, we have separated the illumination color and the object reflectance into intensity components and chromatic components. The object's chromaticity $\mathbf{C}_o$ may be estimated in a straightforward way from the direction of the body-reflection cluster $\mathbf{d}_b$ and from the color of the light source $\mathbf{C}_s$. The red, the green, and the blue components of the body-reflection vector are divided by the red, the green, and the blue components, respectively, of the light-source color:

$$\mathbf{C}_o = [d_b{}^R/C_s{}^R,\ d_b{}^G/C_s{}^G,\ d_b{}^B/c_s{}^B]. \tag{7}$$

The result is normalized to length 1. This estimate assumes that the illumination and the object-reflectance functions are well sampled by the three sensor functions and that these sensor functions form an approximately orthogonal measurement space. Fortunately the color filters used in many color cameras are nearly orthogonal, and the reflectance functions of many common objects are smooth functions of wavelength.[17] In some situations (such as that of fluorescent lights, which have sharp peaks in the spectrum) the assumption of smooth spectral functions will not hold, and then the estimate may be less accurate.

The albedo of the object B_o is also recovered in a straightforward manner from the length of the body-reflection cluster b and the illumination intensity B_s. From Eq. (5) the body-reflection magnitude is

$$m_b = \gamma B_b \cos(\theta_i), \tag{8}$$

where B_b is the object's apparent brightness. The body-cluster length b is equal to the maximum value of m_b, when $\cos(\theta_i) = 1$. Therefore $b = \gamma B_b$. Because the apparent brightness of the object is the product of the object albedo B_o and the illumination intensity B_s, the albedo is

$$B_o = B_b/B_s = b/(\gamma B_s) \tag{9}$$

For the remaining unknowns the situation is not so simple. From Section 2 we know that the length l is related to surface roughness and illumination intensity but is also dependent on imaging geometry. The remaining knowns (l, w, i, $\mathbf{d}_s$) and unknowns (σ, θ_p, B_s, $\mathbf{C}_s$) are examined in detail in the next few subsections.

B. Exact Solution

Equations (5) and (6) describe the amounts of body and surface reflection m_b and m, as a function of imaging geometry and light intensity B_s. Equation (6) also shows how the amount of surface reflection at a given point varies with the roughness of the surface σ.

Unfortunately, a direct solution of these scene parameters from the histogram measurements is not possible. The length of the highlight cluster indicates the maximum amount of surface reflection seen anywhere on the object. For given values of σ, θ_p, and B_s, the length l may be calculated from

$$l = \max\left[\gamma B_s \frac{FG(\theta_i, \theta_r, \theta_p)\alpha}{\sigma \cos(\theta_r)} \exp\left(-\frac{\theta_s{}^2}{2\sigma^2}\right)\right] \quad \text{over all the values of } \theta_i, \theta_r, \theta_s. \tag{10}$$

This equation cannot be inverted to yield a direct solution of the scene parameters σ, θ_p, and B_s from the length.

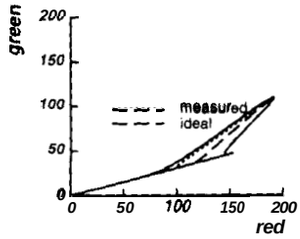

Fig. 13 Skewed length measurement.

Although **Eq.** (10) has no direct solution, it might be possible to solve it iteratively, through some sort of search (e.g., by gradient descent). However, such an approach assumes that the length l can be accurately measured from the histogram. In Subsection 2.B.1, the length is measured from the tip of the highlight cluster to its base along the direction of the highlight color. Unfortunately, the highlight color is not typically known in advance. As described in Subsection **2.B.4,** one may estimate the highlight color by fitting a vector to the pixels that form the highlight cluster. However, this cluster may be skewed.

This effect is shown in Fig. **13,** which is the histogram of a simulated rough object illuminated with white light. The dotted, measured line shows the direction calculated for the best-fit vector to the highlight cluster. (In this case the best-fit line to the cluster would not pass through the brightest point in the highlight cluster.) The position of the dotted line in Fig. **13** shows the projection of the brightest highlight pixel onto the body-reflection vector along the best-fit vector. The ideal line indicates the direction of the actual illumination color. The skewing causes the measured length to be longer than it would have been if the correct illumination color had been known.

This illustration contrasts the concept of ideal length that we would like to obtain from the histogram with the measured length that can actually be recovered without *a priori* knowledge about the illumination color. We obtained all the graphs presented in Section 2 that relate histogram measurements to scene parameters by having knowledge of the illumination color.

Figure 14 shows how both the ideal length and the measured length vary with roughness for the case in which the camera and the light source are separated by 20". For smooth surfaces both values show an inverse relationship between length and roughness. As the roughness increases, the deviation between the two curves grows; eventually the measured length begins increasing again. This is because increasing the roughness also causes an increase in the skew angle of the highlight cluster, as discussed in Subsection **2.B.4.** As the skew of the highlight cluster increases, so does its apparent length. If the measured length were the only value available, it would be difficult to tell a very smooth surface from a very rough one.

A similar story applies to the measurement of width and intersection. When the highlight is skewed, the values measured from the histogram will be different from the ideal values shown in the graphs in Section 2. In the absence of *a priori* information, we can measure only what is available in the histogram. How do we derive the ideal values from those that are measured? Once that task is accomplished, how do we recover the image parameters that interest us?

C. Approximate Solution

Our approach is to recover scene parameters by an approximate method, directly from the initial histogram measurements. Therefore we do not need to recover ideal histogram values from the measured ones. The ideal values of the histogram are a useful abstraction because their relationship to scene parameters is easy to explain. However, the ideal values cannot be obtained from the histogram without knowledge of the illumination color.

Figure **15** shows the variation in the measured length as roughness and phase angle are changed. Each value of length describes a contour within the space of roughness and phase angle. Given a length measurement from a histogram, the associated scene parameters must lie somewhere on that contour. When illumination intensity is considered along with roughness and phase angle, these three scene parameters form a three-dimensional parameter space. A length measurement would then describe a two-dimensional surface within this space, showing the possible roughness, phase angle, and light-intensity values that could give rise to a histogram with this highlight-cluster length.

The intersection and width measurements will also describe surfaces within the parameter space. Ideally each of the histogram measurements will intersect at a single point in the parameter space, making it possible to recover unique values for surface roughness, phase angle, and illumination intensity. Hence the obvious questions

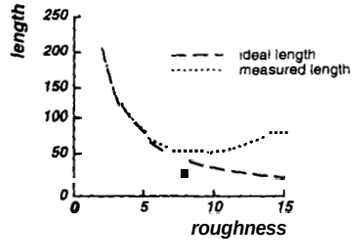

Fig. 14. Ideal versus measured length

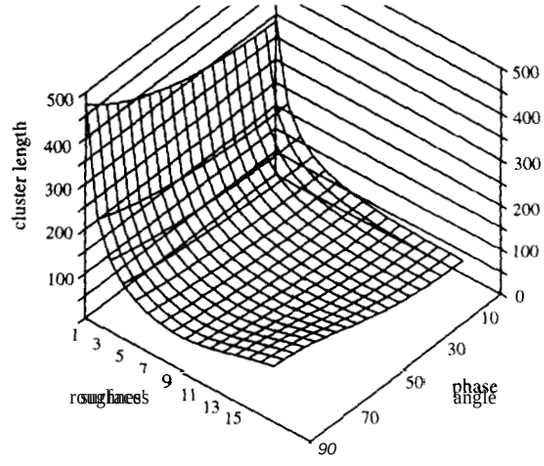

Fig. 15. Variation in cluster length with changes in roughness and phase angle.

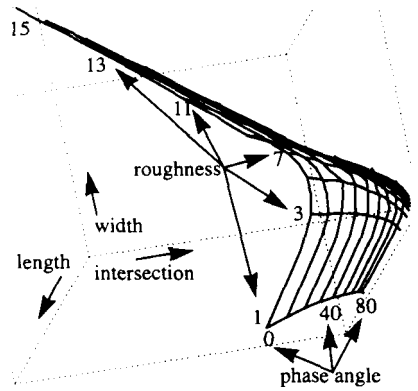

Fig. 16. Surface of roughness and phase angles within length–width–intersection (LWI) space.

are: how can we generate these contours of equal length, equal width, and equal intersection; and do these contours intersect to give unique solutions?

Subsection 3.B pointed out that there is no analytic solution to generate the contours. However, the graph in Fig. 15 shows how the highlight-cluster length varies with the roughness and the phase angle at discrete points. These values come from simulating an object with these parameters and then measuring the length, the width, and the intersection of the highlight cluster in the resulting color histogram. By simulating a large range of roughness values, phase angles, and illumination intensities, we create lookup tables of length, width, and intersection measurements. A program can then search through the lookup tables to find the scene parameters that correspond to a given set of histogram measurements.

The more interesting question is whether a unique solution exists for a given triple (l, w, i). If a particular triple has more than one solution, then different combinations of scene parameters can give rise to identical histogram measurements. The existence of more than one solution for a triple also means that a search through the contours in parameter space cannot be guaranteed to converge.

To visualize the distribution of possible (l, w, i) triples, Fig. 16 shows the space of all possible values for length, width, and intersection. The axis coming out of the page encodes the range of highlight-cluster lengths; the vertical axis shows the range of widths; and the horizontal axis shows intersection measurements. The surface within this cube is defined by lines of equal roughness in one direction and lines of equal phase angle in another direction. This surface shows which combinations of LWI are possible for histograms within the simulated range. Points in the space not falling on the surface correspond to histograms that do not make sense.

Although the surface curves around in the LWI space, it does not intersect itself anywhere. Thus any triple of LWI that falls on the surface is associated with a unique set of surface-roughness and phase angle values. The only remaining problem is to determine from these histogram measurements where on the surface these values will lie.

D. Generating Lookup Tables

A large range of roughness values, phase angles, and illumination intensities are used to create the lookup tables. The ranges are shown in Table 2. The roughness value is the standard deviation of facet angles with respect to the global surface normal, in degrees. The phase angle is the angle between the camera and the light source with respect to the object, also in degrees. The light intensity is a percentage of a hypothetical light's maximum output.

For each set of roughness, phase angle, and light-intensity values, a simulated object is generated with the noise-free camera model. The histogram associated with the object is automatically separated into body-reflection and highlight clusters (see Ref. 13 for details). Vectors are fitted to each of the clusters.

Once the direction of the highlight cluster has been measured, the vector $\mathbf{d}_s$ is used to project all highlight pixels onto the body-reflection vector $\mathbf{d}_b$. These projections determine the relative contributions of the vectors in each pixel. Each color pixel $\mathbf{p}$ in the histogram can then be defined as

$$\mathbf{p} = m_s\mathbf{d}_s + m_b\mathbf{d}_b . \tag{11}$$

This is essentially the dichromatic equation, although the highlight-cluster direction $\mathbf{d}_s$ may differ from the actual highlight color. The histogram measurements are then defined simply as

$$I = \max(m_s) \quad \text{over all } \mathbf{p}, \tag{12}$$

$$b = \max(m_b) \quad \text{over all } \mathbf{p}, \tag{13}$$

$$i = m_b/b \quad \text{for that } \mathbf{p} \text{ with maximum } m_s, \tag{14}$$

$$w = [\max(m_b) - \min(m_b)]/b \quad \text{over all } \mathbf{p} \text{ for which } m_s > T . \tag{15}$$

The threshold T can be set to any arbitrary small value for the case of a noise-free sensor. The value will be set more carefully in Section 4, where we discuss real cameras.

E. Calculating Roughness, Phase Angle, and Illumination Intensity

The surface shown in Fig. 16 shows contours of equal roughness and phase angle within the LWI space. Once length, width, and intersection measurements have been obtained from a histogram, the problem is to determine which contours they fall on. It is unlikely that an arbitrary histogram will have measurements that fall exactly on the contours shown. Therefore some sort of interpolation is required.

In our research we use a polynomial approximation to the surface in LWI space. We assume that the roughness

Table 2. Range of Parameters

Parameter	Minimum	Maximum	Increment	Total Used
Roughness	1"	15"	2"	8
Phase angle	0"	90°	10°	10
Intensity	50%	100%	10%	6
Overall				480

Table 3. Fit of Polynomial Functions

Function	R squared
$\sigma = f_n(l, w, i)$	0.877
$\theta_p = g_n(l, w, i)$	0.939
$B_s = h_n(l, w, i)$	0.712

can be approximated as a polynomial function of length, width, and intersection measurements of the histogram

$$\sigma \approx f_n(l, w, i) \approx A + Bl + Cw + Di + El^2 + Fw^2 + Gi^2 + Hlw + Ili + Jiw + \ldots . \quad (16)$$

The lookup table provides the means for calculating the coefficients of the polynomial. This table also provides almost 500 sets of histogram measurements and the associated roughness values. Least-squares estimation is used to calculate the best-fit nth-degree polynomial to the entire data set. A fourth-degree polynomial is used in our experiments.

Similarly, the phase angle and the illumination intensity are also approximated as polynomial functions of the histogram length, width, and intersection:

$$\theta_p \approx g_n(l, w, i), \quad (17)$$

$$B_s \approx h_n(l, w, i). \quad (18)$$

Least squares is used to calculate the coefficients for these polynomials also. The R-squared fit of the fourth-degree polynomials to the data is shown in Table 3. The R-squared value of 0.877 for σ indicates that a fourth-degree polynomial can account for 87.7% of the variance of this variable.

Generation of the lookup table is obviously very time consuming (taking approximately 8 h on a SPARC II) because it involves calculation of almost 500 graphics simulations. However, the table generation and coefficient calculation must be done only once and can be done ahead of time. At run time our system takes a histogram from an image with unknown parameters, automatically separates it into two clusters, and measures their dimensions. The polynomial equations are then applied to yield a quick estimate of the roughness, the phase angle, and the illumination intensity. The run-time portion is very quick, taking less than 3 s for a histogram corresponding to an image region of 3600 pixels. If the histogram has already been split into clusters in the process of segmenting the image,' the time required for calculating the scene parameters is less than 1 s.

To test the polynomial approximations, 100 test images were simulated and were then analyzed by our method. These test images were noise free, generated with the same idealized camera model used to generate the lookup table. The surface-roughness, the phase angle, and the illumination intensity values used in the test images were chosen by a pseudorandom-number generator. The test values were constrained to lie within the ranges used in the lookup table (see Table 2).

The calculated values of σ, θ_p, and B_s were compared with the original values used to generate the image. In almost all cases the calculated values were near the original ones. However, for 2% of the cases, the values were obviously wrong. For example, a negative value of roughness or illumination intensity is clearly unreasonable. Fortunately, we can detect bad values automatically by checking to see whether recovered values are within the allowable range. Recovered values that fall outside this range indicate that a different method should be used to recover the scene parameters. For example, a lower-degree polynomial approximation could be used for these cases. Although the fourth-degree polynomial more effectively approximates the function over its whole range, it may deviate somewhat at the extremes of the function. In fact, the problem cases (2%) occur when the roughness is very low (between 1° and 2°) and the phase angle is very large (greater than 75°). The problem disappears if a third-degree polynomial is used instead, although the overall error in all cases is slightly higher.

Table 4 shows the results for the remaining 98 cases, in which the fourth-degree polynomial produced reasonable estimates. The average error in recovering the parameter is shown, and the step sizes used in the table are reiterated. The errors are smaller than the table resolution, which shows that the interpolation method is fairly effective.

The results for calculating roughness and phase angle are very good. They show that these noncolor parameters may be calculated with reasonably high accuracy just by consideration of the shape of the color histogram. The error in calculating illumination intensity is a bit higher, although it still provides a useful estimate. This error is not too surprising because the R-squared fit of the function for calculating illumination intensity is somewhat worse than for the other two functions (see Table 3).

F. Calculating Illumination Chromaticity

As we pointed out in Subsection 2.B.4, the highlight cluster may be somewhat skewed from the direction of the highlight color. The skew is particularly pronounced at certain imaging geometries and when the surface is rough. These two factors determine how much the body reflection changes over the area of the highlight.

Therefore, if we know, or can calculate, the surface roughness and the imaging geometry, we can in turn calculate the amount of highlight skewing. Once the skew is known, its effect can be subtracted from the direction of the highlight cluster to give the true color of the illumination.

In Subsection 3.E we showed how to estimate the roughness and the phase angle from the color histogram. These estimates are now used to estimate the skewing. A similar lookup-table approach is used. When the simulations are performed to calculate length, width, and intersection for the range of scene parameters given in

Table 4. Results on Simulated Data

Parameter	Average Error	Table Resolution
Roughness	1.20"	2"
Phase angle	4.40"	10°
Intensity	8.18%	10%
Cases considered		98/100

Table 5. Results in Calculating Skew

Average error	1.73"
Average skew	8.63"
Minimum value	0.01"
Maximum value	27.5"
Number of skews > 1°	69
Cases considered	98/100

Table **2,** the skewing of the highlight is also calculated. In the graphics simulation the correct illumination color is obviously known in advance, so the angle between this color and the vector fitted to the highlight cluster is calculated and is stored in the lookup table along with the values for length, width, and intersection. Then a polynomial function is used to calculate the skew angle as a function of roughness, phase angle, and illumination intensity:

$$\text{skew} \approx \Lambda_n(\sigma, \theta_p, B_s). \tag{19}$$

The coefficients of the nth-degree polynomial function are calculated with a least-squares fit to the data in the lookup table. **A** third-degree polynomial was tested and gave an R-squared fit of 0.993.

Once the skew has been calculated, the highlight color $\mathbf{C}_s$ may be calculated from the measured direction $\mathbf{d}_s$ with the calculated skew angle. Obviously, if the polynomial functions described in Subsection 3.E produce bogus estimates of roughness, phase angle, or illumination intensity, there is little point in plugging them into the equation for calculating skew. In those 2% of the 100 test cases, the program did not attempt to calculate the illumination color. For the remaining 98% test cases, the skew angle was used to calculate the illumination color. The results are shown in Table 5. The error in estimating skew is the difference between the correct skew angle and the skew angle calculated by our method. The correct skew angle is easily calculated from the illumination color used to generate the test picture. The table shows the average error over the 98 cases considered. It also shows the minimum, the maximum, and the average of the actual skew values. For 69 of the test images, the scene parameters were such that the highlight was skewed by 1" or more.

As an example, we now describe the performance of our algorithm on one of the test simulations. The test image, shown in Plate 15, is a red cylinder under white light ($\mathbf{C}_s = [0.58, 0, 58, 0.581$). The cylinder position is aligned such that its vertical axis is parallel to the vertical axis of the simulated imaging plane. The surface roughness is 12.06" standard deviation of the facet angles, the phase angle between the camera and the light source with respect to the object is 63.30°, and the illumination intensity is 90% of maximum.

The histogram associated with this image is shown in Fig. 17. The graph shows the projection of the color histogram into the red–green plane. The histogram forms a smooth curve because it corresponds to a noise-free simulation of an idealized cylinder. The program automatically divided the histogram into body-reflection and highlight clusters. The unit vector fitted to the direction of the highlight cluster is $\mathbf{d}_s = [0.81, 0.45, 0.371$. The length of the highlight cluster was measured as 74.1, the width of the highlight cluster was measured as 0.47 (the highlight cluster extends over slightly less than half the body cluster), and the intersection of the two clusters was measured as 0.51 (the brightest point in the highlight cluster was projected onto the body cluster just above the halfway point). These measurements were all obtained with the vector **d,** to calculate the amount of body reflection and surface reflection.

The direction fitted to the highlight cluster is significantly skewed away from the direction of the actual illumination color. It represents a much redder color than the white illumination color and hence would be a poor estimate of the illumination color. It would also yield an inaccurate estimate of the object color when the influence of the illumination color is divided out of the body-reflection color.

With the application of polynomial relations (16)–(18) to the length, the width, and the intersection measurements, the program estimated the roughness value as 11.99°, phase angle as 67.05°, and the illumination intensity as 89% of maximum. We then applied relation (19) to these estimates of σ, θ_p, and B, and estimated the skew between the highlight-cluster direction and the actual illumination color to be 18.75". Application of this skew to the cluster direction **d,** produced an estimate of the light-source chromaticity as $\mathbf{C}_s = [0.59, 0.57, 0.57]$. This is very close to the original white color.

A picture made by simulation of the cylinder with the recovered parameters is shown in Plate 16. The difference between this picture and the original is very subtle. The results for this example are summarized in Table **6.** The full algorithm is diagramed in Fig. 18.

4. EXTENDING THE ANALYSIS TO REAL IMAGES

In Section **3** we described how to estimate scene parameters from a color histogram if an ideal sensor is used. Specifically, we assumed that the camera used for viewing objects is noise free and linear over the range of all possible intensity values. The possibility of clipping at some maximum camera value was not considered in Section 3.

However, these assumptions are not very realistic. Some forms of noise exist in all real-camera systems. And although linearity of response over a limited range is common in many cameras, all real cameras have some maximum value. Clipping at the maximum value is particularly common in pictures with highlights because highlights are so much brighter than other parts of the scene. Other problems that are common in real images

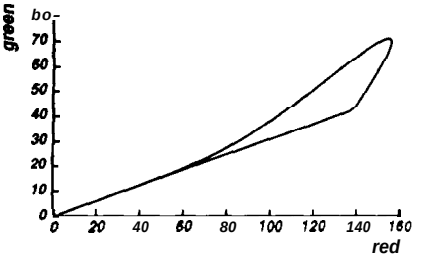

Fig. 17. Histogram of the simulated image

Table 6. Example Results

Simulated Image	Histogram Measurements	Recovered Parameters
$\mathbf{C}_s = [0.58, 0.58, 0.58]$	$\mathbf{d}_s = [0.81, 0.45, 0.37]$	$\mathbf{C}_s = [0.59, 0.57, 0.57]$
$\sigma = 12.06°$	$l = 74.1$	$\sigma = 11.99°$
$\theta_p = 63.30°$	$w = 0.47$	$\theta_p = 67.05°$
$E_s = 90\%$	$i = 0.51$	$B_s = 89\%$

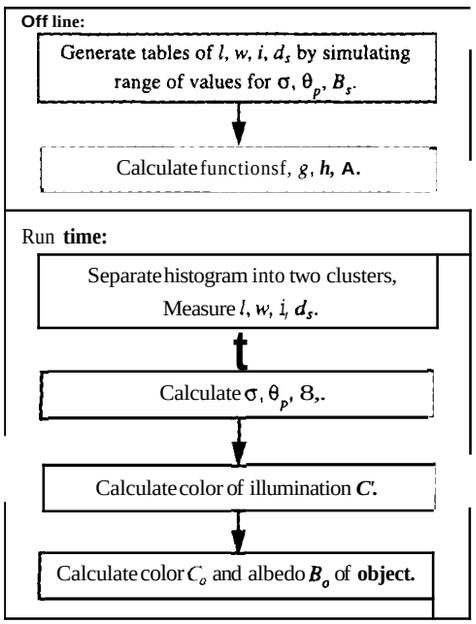

Fig. 18. Algorithm for calculating scene parameters.

are nonlinearity, color imbalance, and chromatic aberration. The next few subsections briefly describe how we adapted the algorithm to deal with these issues.

A. Camera Noise

Real cameras present many limitations for taking accurate images. CCD cameras, which are used by most vision researchers, present many problems such as dark current, shot noise, amplifier noise, and line jitter.[11] Some of these problems, such as shot noise, are an inherent property of CCD cameras and cannot even be removed with high-quality manufacturing.

In dealing with noise in the pixel values, we take a two-pronged attack: we explicitly account for the noise in generating the lookup tables, and we try to remove some noise when analyzing each histogram.

We modeled camera noise as having a Gaussian distribution with a mean of zero. The standard deviation of camera noise, σ_c, was modeled as 1.275. The camera is assumed to measure 8 bits/pixel, giving a range from 0 to 255.

In developing the algorithm for ideal cameras, we required a threshold for distinguishing highlight pixels from pixels on the body-reflection cluster, although this threshold could be arbitrary (see Subsection **3.D).** With an explicit noise model, the threshold should be based on the standard deviation of the camera noise. We chose to set the threshold T at $4\sigma_c$, so more than 99% of the noisy body-reflection pixels will fall within the body-reflection cluster

When we generated the lookup tables for Section 3, we simulated noise-free images and measured the resulting histograms. Now, for the noisy-camera model, we simulate images again, this time adding pseudorandom Gaussian noise. The resulting histograms are analyzed, and new values of length, width, and intersection are recorded in the lookup table. These measurements reflect the changes in the overall histogram shape caused by the presence of noisy pixel values. (Ref. 13 discusses these changes in more detail.)

Furthermore, when measuring the histograms, we use local smoothing. Each pixel from the simulated image is averaged with its four nearest neighbors before it is plotted in the histogram. For pixels that fall in the highlight or at its edge, this procedure will have the effect of smoothing out the highlight somewhat but will also tend to average out the noise. The effect of highlight smoothing on the measurements of length, width, and intersection will also be recorded in the lookup table.

B. Camera Clipping

Like noise, clipping is an inherent property of real cameras because all real cameras have some maximum value that they can measure. Many real images with highlights suffer from clipping because highlights are often very bright. When taking real pictures, researchers can often adjust the aperture, the light level, or the exposure time so that the brightest point in the image is within the dynamic range of the camera. However, highlights on shiny objects can be an order of magnitude brighter than surrounding nonhighlight areas, so that, if the camera is adjusted to measure highlights properly, the nonhighlight areas will be so dim as to be lost in the camera noise.

Therefore any algorithm that is meant to work on real pictures with highlights should take clipping into account. The effect of clipping on the color histogram is easy to model. Any pixels that would have exceeded the maximum value **(255** for an 8-bit camera) will be held to that value. **As** a result, the highlight cluster will not be so long as it otherwise would have been. Clipping usually occurs when the surface roughness has a standard deviation of facet angles of **3°** or less.

In analyzing color histograms, clipping is easy to detect. If a pixel has the maximum value in any band, it was almost certainly clipped. However, dealing with the clipping presents a greater problem because the effect on the length measurement may be dramatic. The length measurement is an important parameter in calculating the scene parameters, along with the width and the intersection measurements. For calculating the three unknown values of surface roughness, phase angle, and illumination intensity, three known histogram measurements are required. If we cannot trust the length measurement, we need to replace it with some other measurement.

When clipping occurs, we replace the length measurement with another measurement of the cluster shape. The width measurement and the length measurement are a way of characterizing the shape of the highlight cluster. We want to capture this shape because it is related to the relative distributions of body and surface reflection, which in turn are determined by the scene parameters that we want to estimate.

When the length measurement is untrustworthy because of clipping, we use a second width measurement. The first width measurement indicates how wide the highlight cluster is at the threshold distance T from the body-reflection cluster. The second width measurement shows how wide the highlight cluster is at a distance of $3T$ from the body-reflection cluster.

When the lookup tables are calculated, we record both the length and whether it is clipped. We also record the second width measurement, as well as the first width and intersection measurements previously described. After the lookup table is generated two sets of polynomial functions are fitted: one set for analyzing histograms that are not clipped, and one set for histograms that are. If the second width measurement is referred to as w_2 then the scene parameters may be calculated with

$$\left.\begin{aligned}\sigma &= f_1(l, w, i)\\ \theta_p &= g_1(l, w, i)\\ B_s &= h_1(l, w, i)\end{aligned}\right\} \quad \text{if histogram is not clipped,} \tag{20}$$

$$\left.\begin{aligned}\sigma &= f_2(w_2, w, i)\\ \theta_p &= g_2(w_2, w, i)\\ B_s &= h_2(w_2, w, i)\end{aligned}\right\} \quad \text{if histogram is clipped.} \tag{21}$$

The second width measurement does not provide quite so much predictive power as the length measurement, so we use the first set of equations whenever possible.

C. Nonlinearity and Color Imbalance

The real images used in this paper were obtained in the Calibrated Imaging Laboratory (CIL) at Carnegie Mellon University. The CIL provides a controlled setting for taking high-quality images. It also permits a careful analysis of camera issues that may cause problems for vision algorithms. The equipment and capabilities are described in Ref. 19.

The methods that we have developed for analyzing histograms require that the camera have a linear response. This is because these methods measure specific distances in the color space and assume that these distances are proportional to the amounts of body reflection and surface reflection observed on the object. Fortunately, linear CCD cameras are commonly available. In fact, CCD cameras are inherently linear because they produce a voltage that is proportional to the number of photons; the problem is that some models are designed to drive displays and have added circuitry that makes them nonlinear. In this case the nonlinearity may be measured with a standard reference chart and may be corrected with a lookup table.[9,20] The camera used to obtain the real images used in this paper already has a linear response, so no correction is necessary.

Another common problem with CCD cameras is color imbalance. CCD's are typically much less sensitive to blue wavelengths than they are to green or red wavelengths. In addition, incandescent lights are often used for taking pictures. Incandescent lights are also strong in the long wavelengths and weak in the short wavelengths. Therefore, when incandescent lights are used with CCD cameras, the blue response is extremely poor. If color pictures are taken without some form of color correction, blue objects will be extremely dark, and white objects will appear yellowish.

In the CIL, color pictures are obtained with a gray-scale camera and a filter wheel to produce the red, the green, and the blue bands. This allows adjustments to be made individually for each band. A white test card is used to determine camera exposure times for each band. In the CIL the exposure time for the blue band is typically approximately ten times longer than the red or the green exposures.

D. Chromatic Aberration

Another common problem in real images is chromatic aberration. This is an inherent property of lenses because the index of refraction of optical components varies as a function of wavelength. Chromatic aberration can be partially compensated for during manufacturing; however, experience in the CIL has shown that even high-quality lenses suffer from chromatic aberration that is readily apparent in histogram analysis.

Chromatic aberration is most pronounced where there are sharp edges in the image. For smoother surfaces the transition between highlight and nonhighlight areas will be more sudden, and thus the chromatic aberration will be more pronounced. The effect of chromatic aberration on the histogram is that pixels are displaced from their expected values, causing a scattering effect. The degree of this scattering depends not only on the sharpness of the transition but also on the characteristics of the lens, the color of the object, and the location of the highlight in the image. (Chromatic aberration is most pronounced in the periphery of the image.) This makes it very difficult to model in a general way.

In our research we attack the problem of chromatic aberration in two ways. The first step is to eliminate as much of the chromatic aberration as possible at the time the image is obtained. This is done by active lens compensation. Because the CIL permits fine control of the lens position and because the color bands are imaged separately, the lens is refocused, rezoomed, and repositioned for each of the red, the green, and the blue color bands. This approach typically reduces the chromatic aberration by an order of magnitude.[20,21]

Unfortunately this method does not eliminate all chromatic aberration in images with highlights. Also, this method is time consuming to calibrate and is not generally applicable to camera setups that do not have a finely controlled lens. Therefore the second approach is to take some chromatic aberration into account when measuring the dimensions of the color histogram.

Figure 19 shows a histogram for a very shiny red pail. The original image is shown in the upper right-hand corner of the figure. The top of the highlight cluster is fanned out, which makes it wider than the bottom of the highlight cluster.

This fanning of the highlight cluster is particularly

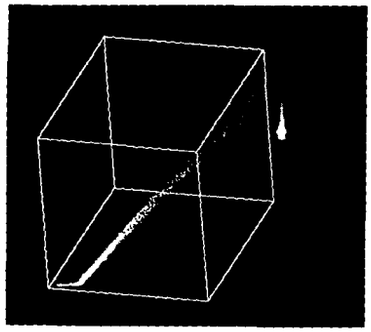

Fig. 19. Histogram for a very shiny pail.

troublesome because it runs counter to the model of histogram behavior. The model says that there will be more variation of body-reflection amounts at the base of the highlight cluster than at the upper end. A highlight cluster that is wider at the top implies that there are very bright highlight pixels occurring at several different reflection angles but that, for some reason, pixels with lesser amounts of highlight occur at only a few reflection angles. If the histogram is to be believed, then the amount of highlight does not fall off as the off-specular angle is increased. This effect would indicate a rather unusual distribution of facet angles. A more plausible explanation of the fan-out of the highlight cluster is that the chromatic aberration causes large hue shifts at points where the highlight is increasing very rapidly.

Because the chromatic aberration makes the highlight cluster wider in regions in which the highlight increases rapidly, this would imply that the width measurement is not very trustworthy at the top end of the highlight cluster. A program analyzing the histogram shown in Fig. 19 will measure a very wide highlight cluster, even though it corresponds to a very shiny object. This is because the original algorithm examines the amount of body reflection present in all highlight pixels to compute the width measurement. A pixel is classified as a highlight if the amount of surface reflection exceeds the threshold $4\sigma_c$. The surface-reflection model predicts that the widest part of the highlight cluster is at the base, but this property is not explicitly tested by the program described in Section 3.

Although the highlight cluster is plagued by chromatic aberration at the top end, the base of the cluster is still well behaved. Now that chromatic aberration is known to be a problem, the algorithm measures width over all the highlight pixels near the threshold point instead of over all the pixels that exceed the threshold. Equation (15), which gives the width measurement in terms of the computed amounts of body reflection m_b and surface reflection m_s, is modified to

$$w = [\max(m_b) - \min(m_b)]/b$$

over all $\mathbf{p}$ for which $T < m_s < T + \sigma_c$. (22)

Similarly, the second width measurement, used when the highlight cluster is clipped, becomes

$$w_2 = [\max(m_b) - \min(m_b)]/b$$

over all $\mathbf{p}$ for which $3T < m_s < 3T + \sigma_c$. (23)

The process for measuring length and intersection remain unchanged.

5. EXPERIMENTAL RESULTS

In the previous section we described some modifications made to the analysis method along with some techniques for obtaining high-quality images. Taken together, they allow the algorithm to work on real images. The next few subsections describe the performance of the algorithm in estimating phase angle, illumination intensity, and roughness from real images.

A. Estimating Phase Angle

An experiment was set up in the CIL to test the histogram analysis algorithm at estimating phase angle from real images. A series of images was taken for which the camera and the light source were separated by different phase angles. The angle was measured with a large protractor and strings to indicate the direction of the camera and the light source. A diagram of the setup is shown in Fig. 20. The angles measured by this method were estimated to be accurate to within 5°. The light was a 250-W spotlight; it was estimated to appear a few degrees across at the distance used in the experiment [approximately 5 ft (1.5 m)]. Therefore it is only a crude approximation of a point source.

The first image in the sequence is shown in Plate 17. This picture was taken when the camera and the light source were approximately 10° apart. The phase angle was then increased by 10° between each picture. The last image in the sequence is shown in Plate 18. This picture was taken when the phase angle between the camera and the light source was 90°.

The program automatically split the color histograms of the objects into two clusters, fitted lines to those vectors, and calculated the values of length, width, and intersection. The color histogram of the image in Plate 18 is shown in Plate 19. The superimposed white lines show the lines fitted by the program.

This process was repeated for each image in the sequence. Figure 21 shows the measured intersection value for each image, plotted against the phase angle that was in effect at the time the image was taken. A comparison of Fig. 21 with Fig. 10 in Section 2 shows that, as predicted, the intersection decreases with increasing phase angle. The sole exception is at $\theta = 90°$. A slight error in setting up the light could explain this problem: if the phase angle were actually slightly larger than 90°, the brightest point on the object would be out of view of the camera, throwing off the intersection ratio.

The program described in Sections 3 and 4 was used to calculate the phase angle from the length, the width, and the intersection measurements obtained from each image. The results are shown in Fig. 22. The dotted line shows

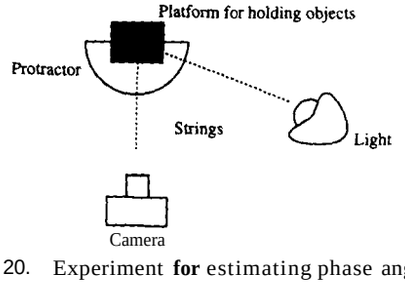

Fig. 20. Experiment for estimating phase angle.

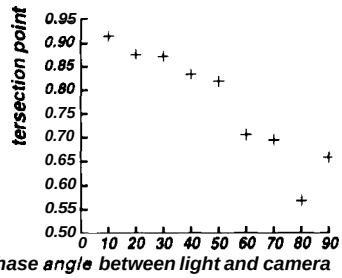

Fig. 21. Plot of intersection versus phase angle.

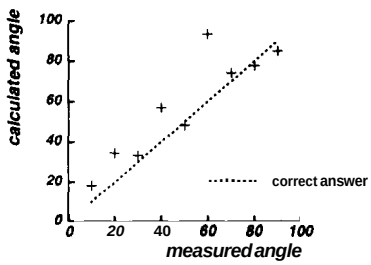

Fig. 22. Results for calculating phase angle from real images.

the correct answer, with the phase angle being measured by the protractor as ground truth. The average error in estimating the angle is 9.96".

As the light was moved around, the distance from the object did not remain constant because of the layout of the laboratory. Therefore the intensity of the illumination was not assumed to be constant. For two of the images in the sequence, the illumination was close enough that the highlight saturated the camera, and clipping occurred. This effect was seen at $\theta_i = 60°$ and $\theta_i = 80°$. In these two cases the second set of equations [see Eqs. (21)] for clipping was used. Although this technique produced a reasonable answer for the case in which $\theta_i = 80°$, it did not do so well for $\theta_p = 60°$. If this data point is eliminated from consideration, the average error in estimating phase angle is 7.04".

Overall, the method developed for estimating phase angle from analyzing color histograms works fairly well, especially considering that the ground-truth measurement of the phase angle is fairly crude. Also, the lookup tables in Section **4** were calculated without calibrating the simulated images to the conditions in the CIL. In particular, the noise of the camera was not measured precisely, and the light source used in the experiments was not a point source, as was the case in the simulations. Moreover, the Torrance–Sparrow roughness model is an idealized model that assumes isotropic, Gaussian scattering and may not effectively describe the roughness of the real object used in these experiments.

B. Estimating Illumination Intensity

A second experiment was performed in the CIL to test the performance of the algorithm at estimating illumination intensity. The spotlight was plugged into a variable voltage supply with a manually operated dial. A sequence of images was taken under increasing levels of illumination while the imaging geometry and the target object were kept constant. (Varying voltage will also cause some change in the illumination color, as discussed below.) Altogether six images were taken. The illumination level was measured with a luminance spot meter aimed at a white card. The spot measurements were estimated to be repeatable to within *5%.*

Again the program analyzed the histograms to produce measurements of length, width, and intersection for each image. The polynomial equation for calculation of illumination intensity was then applied to these measurements. The results are shown in Fig. **23.** The horizontal axis shows the luminance (candles/square meter) measured by the spot meter, whereas the vertical axis shows the intensity estimated by the histogram analysis. The gain of the camera has not been calibrated, so the program gives a relative estimate of intensity. The dotted line shows the best linear fit to the data. If the slope of this line is considered to be the gain of the camera, then the average error in estimating illumination intensity is 5.07%.

We computed the results shown in Fig. **23** by using only the length, the width, and the intersection measurements of the highlight cluster. The graph in Fig. **24** shows the roughness estimate for each image in the sequence, computed at the same time as the intensity estimates. These measurements are fairly stable across the sequence, as would be expected, because the imaging geometry and the target object were the same in every case. This result is particularly important because both increasing smoothness and increasing illumination intensity will lengthen the highlight cluster. We also found that the estimates of the phase angle made at the same time were reasonably consistent. This result shows that, for the most part, the algorithm does not confuse changes in the histogram shape that are due to increasing intensity levels with the effects of roughness and phase angle.

The ground-truth measurements for this experiment came from a luminance meter and were compared with estimates of intensity. Unfortunately, these measurements are somewhat different. A luminance meter integrates with the human luminous efficiency function $V(\lambda)$. **A** luminance meter shows how bright a spot will appear

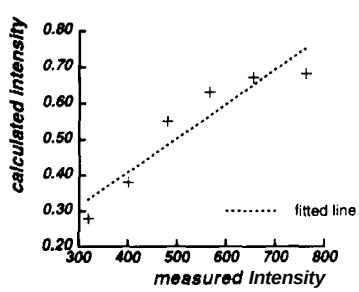

Fig. **23.** Results **for** calculating illumination intensity from real images.

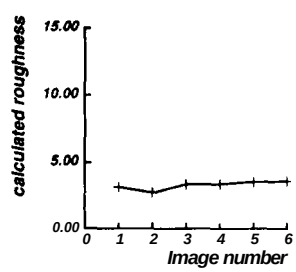

Fig. **24.** Calculating roughness when intensity changes.

Table 7. Results for Estimating Roughness

Object	Calculated Roughness (degrees)
Alligator	10.07
Pumpkin	8.93
Terra-cotta ball	3.61
Red ball	0.40
Red pail	0.10

to humans, but not necessarily how bright it will appear to a color camera. A narrow-band red illumination will have a much lower luminance than a narrow-band green one, yet both may have the same intensity as measured by a color camera.

The difference between luminance and intensity is relevant to the results presented here because, as the voltage fed to the spotlight is decreased, the light becomes not only dimmer but also noticeably redder in hue. This effect indicates that luminance may not correlate well with the intensity in this experiment. Although the histogram analysis did correctly calculate increasing illumination intensities as the light level was increased, the calculated values do not appear as linear as one might hope. The use of a luminance meter might be causing some of the problems.

C. Estimating Roughness

A third experiment was performed to show how the system estimates surface roughness from color histograms. Plate 20 shows a composite of five images of different objects. The objects are a green plastic pool toy in the shape of an alligator, an orange plastic pumpkin for trick-or-treating, a terra-cotta ball (also shown in Plate 17), a red plastic ball, and a red plastic pail (also shown in Fig. 19).

Table 7 shows the roughness calculated by the system for each of these objects. The objects are listed in order of decreasing roughness, as estimated by human observation. The calculated roughness number is an estimate of the standard deviation of facet angles, in degrees. There is no error measure for these results because there are no ground-truth data far the actual roughness values. Nevertheless, the roughness ranking from the program agrees with that produced by a human observer.

The results presented in this section demonstrate that the method developed here can be applied to real images. The algorithm produces reasonable estimates of phase angle, illumination intensity, and surface roughness by analyzing the dimensions of color histograms. The method is able to succeed even with such camera problems as noise, clipping, and chromatic aberration. It is likely that the method would perform even better if the lookup tables used measurements from real histograms rather than from simulated ones. Nevertheless, the method works reasonably well, even without prior reference to real data.

6. CONCLUSIONS

The color histogram of an image is a rich source of information, but in the past it has not been fully exploited. We have shown that the color histogram of a dielectric object may be characterized by a small number of measurements, which relate directly to many scene properties. We have shown how these histogram measurements may be used to recover estimates of surface roughness, imaging geometry, illumination intensity, and illumination color. These estimates may in turn be used to estimate object color and albedo.

We have shown that the resulting algorithm can be applied to real images and produces good estimates of phase angle, illumination intensity, and surface roughness. The method is independent of the shape of the object and works on shapes ranging from that of a pumpkin to that of an alligator. The model used to develop the lookup tables is fairly general and was not calibrated to match the actual imaging conditions such as light-source extent or camera noise characteristics. This kind of analysis may be applied to such varied tasks as surface inspection and object recognition

ACKNOWLEDGMENTS

The authors thank Reg Wilson, Jim Moody, and Bill Ross, who developed much of the hardware and the software that makes picture taking possible in the Calibrated Imaging Laboratory at Carnegie Mellon University. We also thank the reviewers, who made many useful suggestions for improving the paper. This research was sponsored by the Avionics Lab, Wright Research and Development Center, Aeronautical Systems Division (AFSC), U.S. Air Force, Wright-Patterson AFB, Ohio 45433-6543, under contract F33615-90-C-1465, ARPA order 7597, and by the Air Force Office of Scientific Research under contract F49620-86-C-0127.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government.

REFERENCES

1. R. M. Haralick and G. L. Kelly, "Pattern recognition with measurement space and spatial clustering for multiple images," in Proc. IEEE *57,* 654–655 (1969).
2. S. A. Shafer, "Using color to separate reflection components." Color Res. Appl. 10,210–218 (1985)
3. G. J. Klinker, S. A. Shafer, and T. Kanade, "Using a color reflection model to separate highlights from object color," in *Proceedings of the International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, New York, 1987), pp. 145–150.
4. R. Gershon, "The use of color in highlight identification," in *Proceedings of the 10th International Joint Conference on Artificial Intelligence* (Morgan Kaufman, Los Altos, Calif., 1987), pp. 752–754.
5. G. Healey, "A color reflectance model and its use for segmentation," in *Proceedings of the International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 460–466.
6. C. L. Novak and S. A. Shafer, "Anatomy of a color histogram," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1992).
7. M. D'Zmura and P. Lennie, "Mechanisms of color constancy," J. Opt. Soc. *Am.* A **3,** 1662–1672 (1986).
8. G. Healey and T. O. Binford, "The role and use of color in a general vision system," in *Proceedings of the ARPA Image Understanding Workshop* (Defense Advanced Research Projects Agency, Arlington, Va., 1987), pp. 599–613.

9. G. J. Klinker, "A physical approach to color image understanding," Ph.D. dissertation (Carnegie Mellon University, Pittsburgh, Pa., 1988).
10. S. Tominaga and B. A. Wandell, "Standard surface-reflectance model and illuminant estimation," J. Opt. Soc. Am. A 6, 576–584 (1989).
11. K. Torrance and E. Sparrow, "Theory for off-specular reflection from roughened surfaces," J. Opt. Soc. Am. **57,** 1105–1114 (1967).
12. H. C. Lee, E. J. Breneman, and C. P. Schulte, "Modeling light reflection for color computer vision," IEEE Trans. Pattern Anal. Mach. Intell. 12, 402–409 (1990).
13. C. L. Novak, "Estimating scene properties by analyzing color histograms with physics-based models," Ph.D. dissertation (Carnegie Mellon University, Pittsburgh, Pa., 1992).
14. L. B. Wolff, "Diffuse reflection," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition* (Institute of Electrical and Electronics Engineers, New York, 1992).
15. P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waoes from Rough Surfaces* (Macmillan, New York, 1963).
16. S. K. Nayar, K. Ikeuchi, and T. Kanade, "Surface reflection: physical and geometrical perspectives," IEEE Trans. Pattern Anal. Mach. Intell. 13, 611–633 (1991).
17. L. T. Maloney, "Evaluation of linear models of surface spectral reflectance with small numbers of parameters," J. Opt. Soc. Am. A 3, 1673–1683 (1986).
18. G. Healey and R. Kondepudy, "Modeling and calibrating CCD cameras for illumination insensitive machine vision," in *Optics, Illumination and Image Sensing for Machine Vision VI*, D. J. Svetkoff, ed., Proc. Soc. Photo-Opt. Instrum. Eng. 1614, 279–290 (1991).
19. R. G. Willson and S. A. Shafer, "Precision imaging and control for machine vision research at Carnegie Mellon University," in *High-Resolution Sensors and Hybrid Systems*, M. M. Blouke, W. Chang, R. P. Khosla, L. J. Thorpe, eds., Proc. Soc. Photo-Opt. Instrum. Eng. 1656, 297–314 (1992).
20. C. L. Novak, S. A. Shafer, and R. G. Wilson, "Obtaining accurate color images for machine vision research," in *Perceiving, Measuring and Using Color*, M. M. Brill, ed., Proc. Soc. Photo-Opt. Instrum. Eng. 1250, 222–235 (1990).
21. R. G. Wilson and S. A. Shafer, "Active lens control for high precision computer imagmg," in *Proceedings of the International Conference on Robotics and Automation* (Institute of Electrical and Electronics Engineers, New York, 1991).

Physics-Based Machine Vision

Color Plates

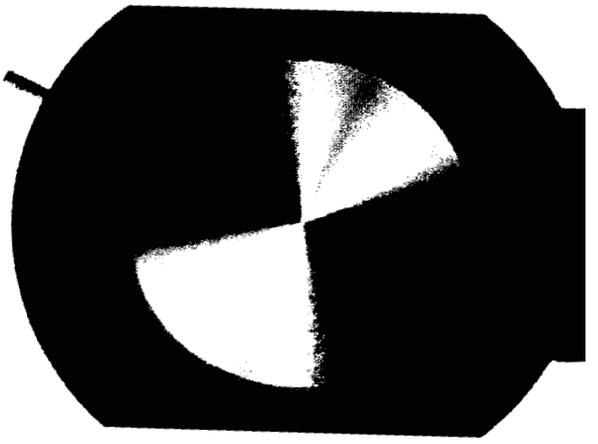

Plate 1. Polarization image of a radial-polarization filter taker by a two CCD polarization camera with a polarizing beam splitter ar

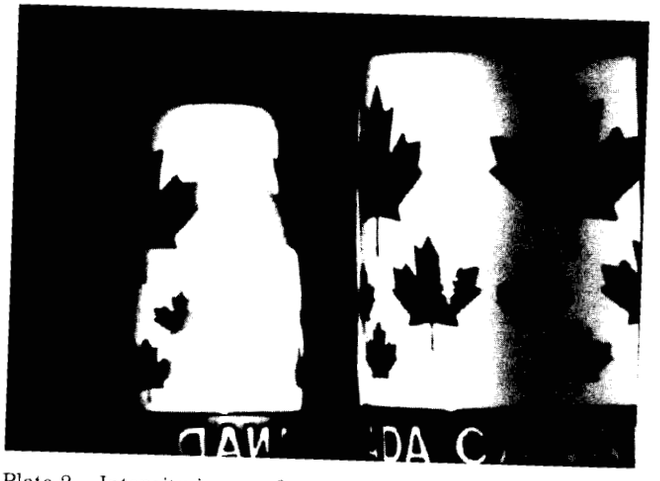

Plate 2. Intensity image of a ceramic cup and its reflection in a glass mirror.

Plate 3. Partial-polarization–intensity image produced by two-CCD polarization camera with a polarizing beam splitt Partial polarization is represented as the

Plate 4. Polarization image produced by a two-CCD polarization camera with a polarizing beam splitter and a TN liquid crystal.

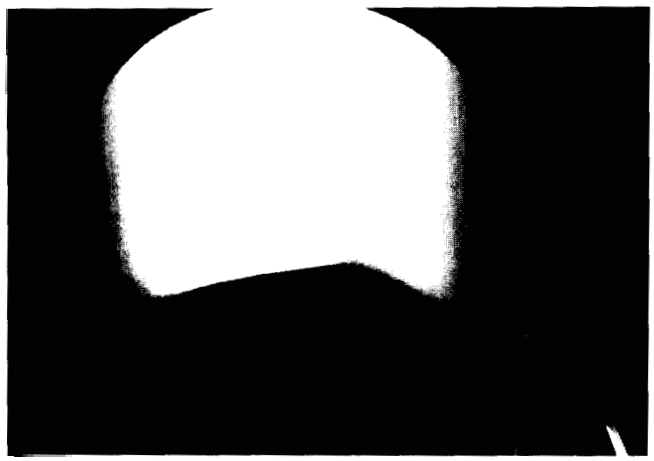

Plate 5. Intensity image of a ceramic cup under extended illumination.

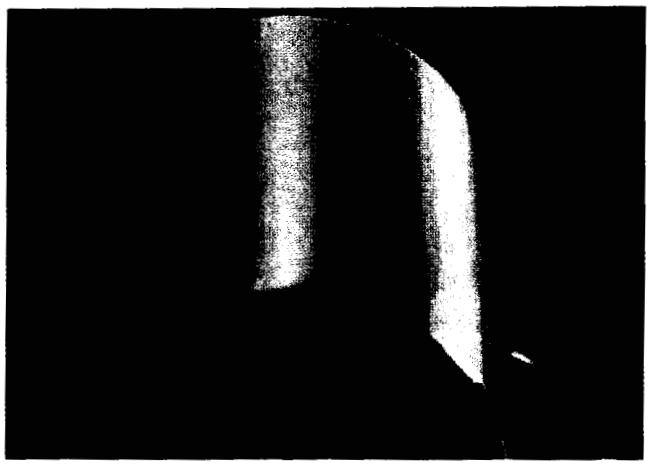

Plate 6. Polarization image produced by a two-CCD polarization camera with a polarizing beam splitter and a TN liquid crystal.

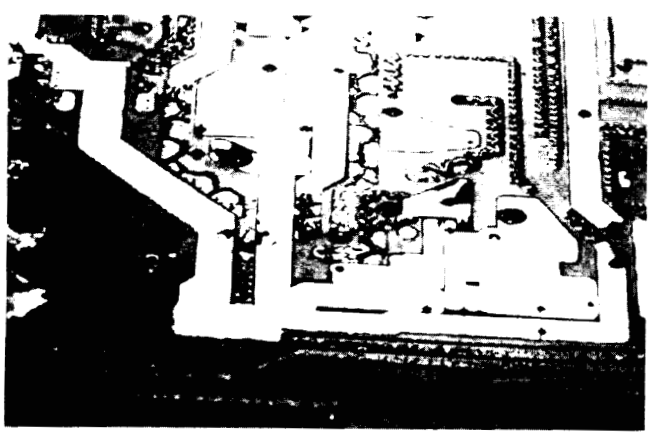

Plate 7. Intensity image of a circuit board with metal solder, plastic dielectric substrate, and translucent dielectric coating on solder metal.

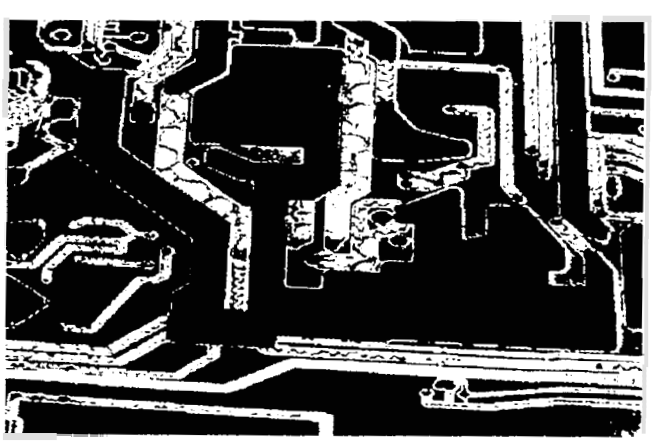

Plate 8. Material segmentation produced by a two-CCD polarization camera with a polarizing beam splitter. Blue corresponds to dielectric. red corresponds to metal. and yellow corresponds to translucent dielectric coating on solder metal. Some red shows where translucent dielectric coating on solder is crinkled.

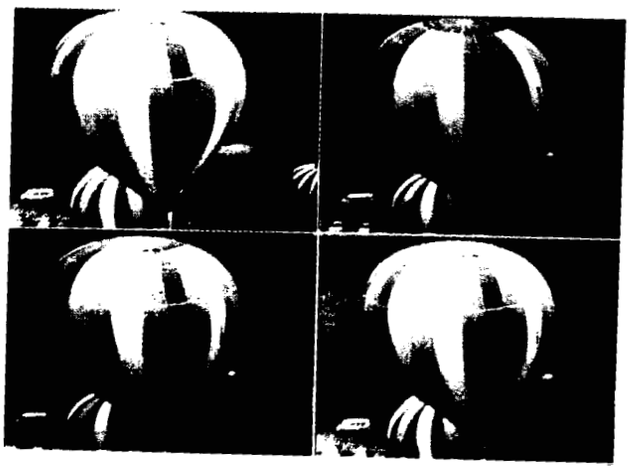

Plate 9. Balloon.

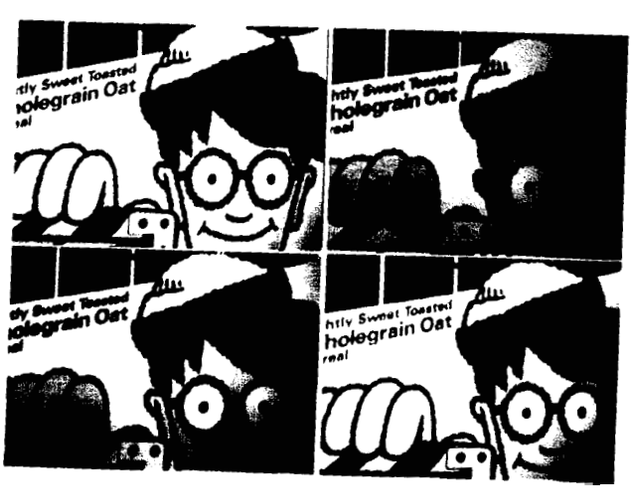

Plate 10. Cereal box.

Plate 11. Chalk box.

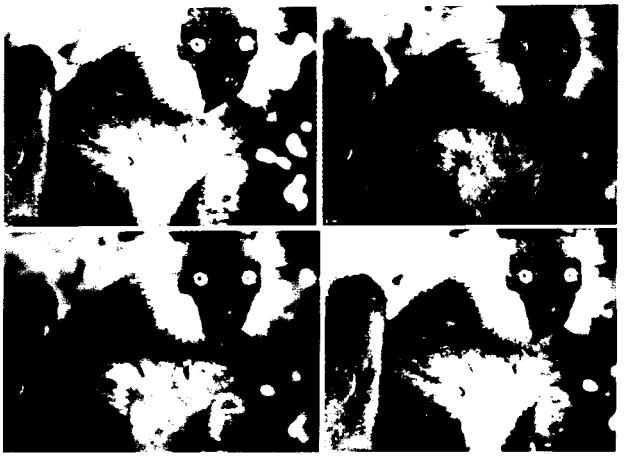

Plate 13. Lemur.

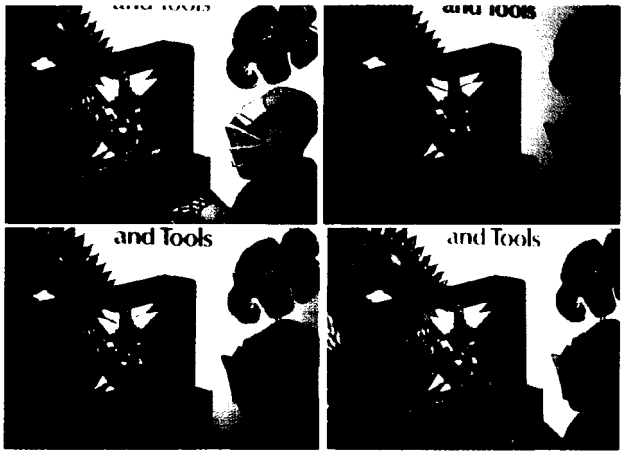

Plate 12. Dragon.

Plate 14. Tiger

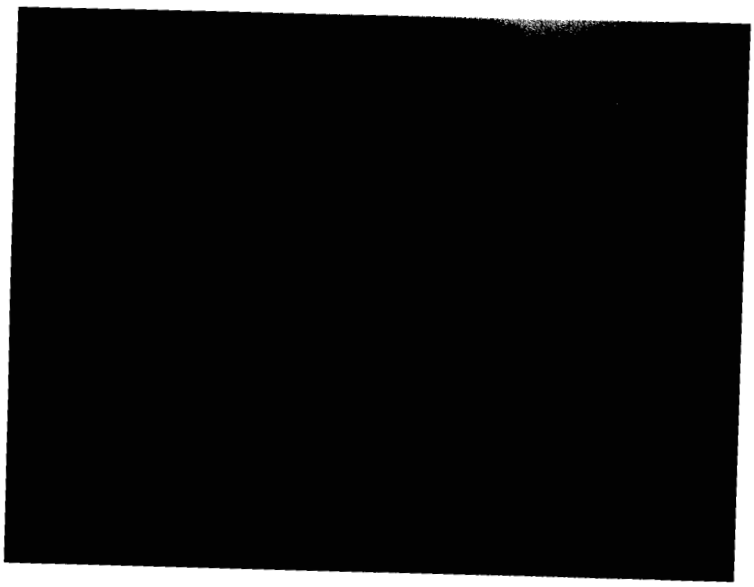

Plate 15. Simulated test image.

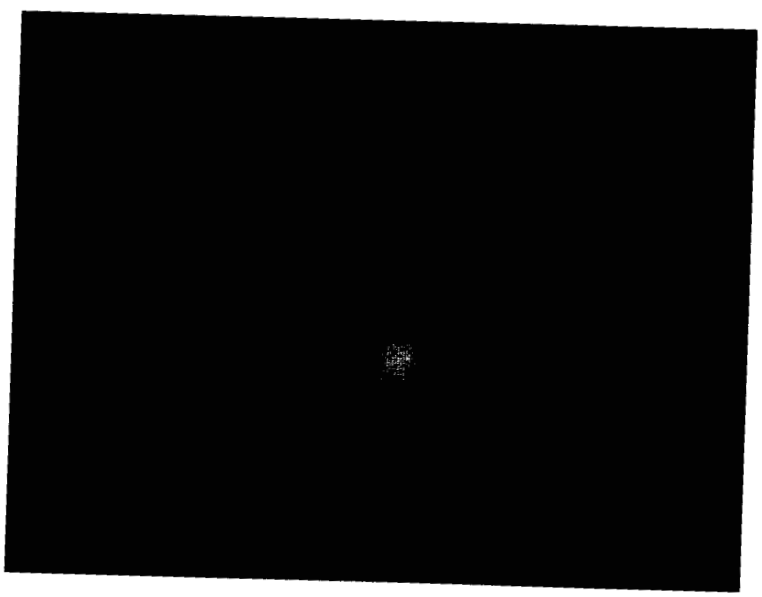

Plate 17. Real image with phase angle of 10°.

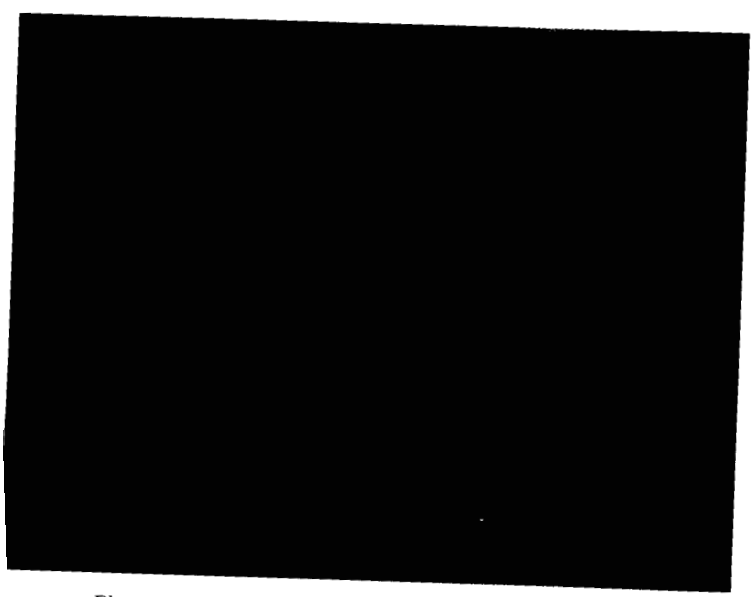

Plate 16. Simulation from recovered parameters.

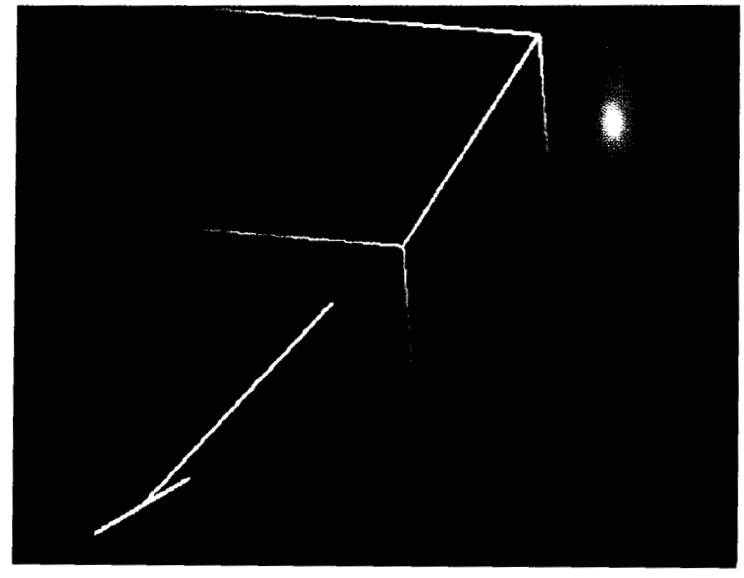

Plate 19. Histogram for phase angle of 90°

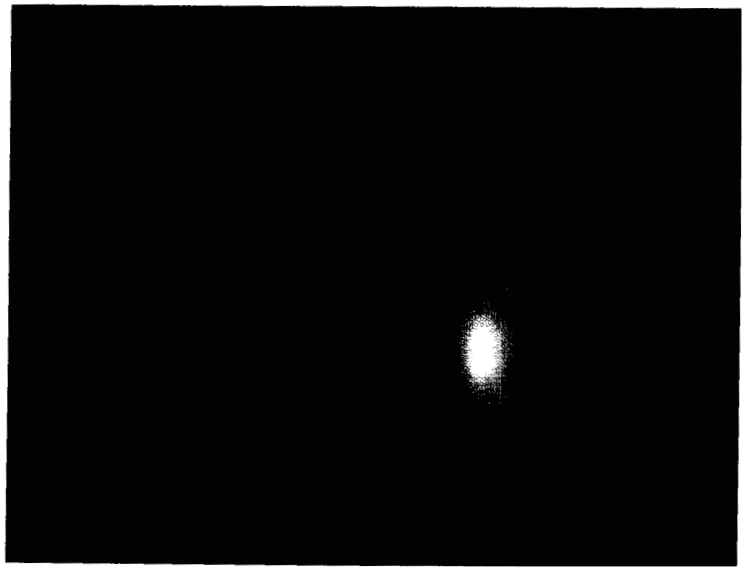

Plate 18. Real image with phase angle of 90"

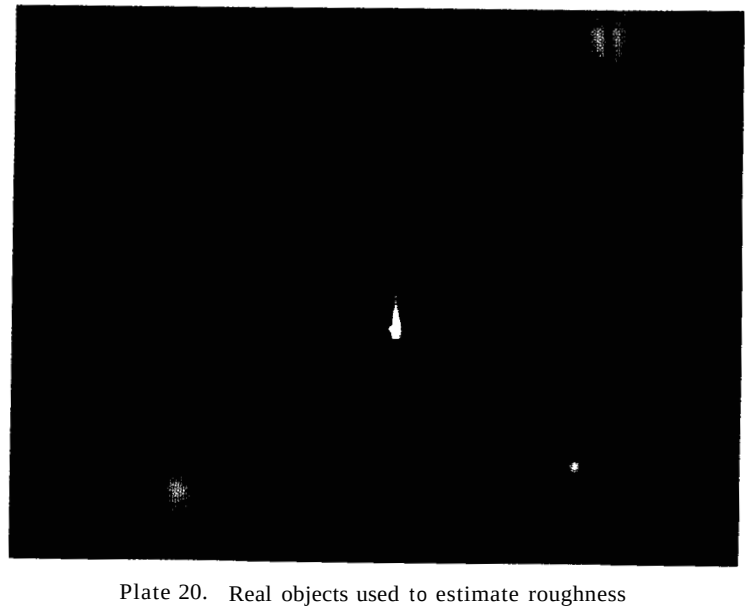

Plate 20. Real objects used to estimate roughness

Plate 21. Example of color-encoded gradient (p, q) as produced in the near-real-time (15-Hz) implementation of photometric stereo. The inset (lower right) shows the color rosette used to encode the gradient.

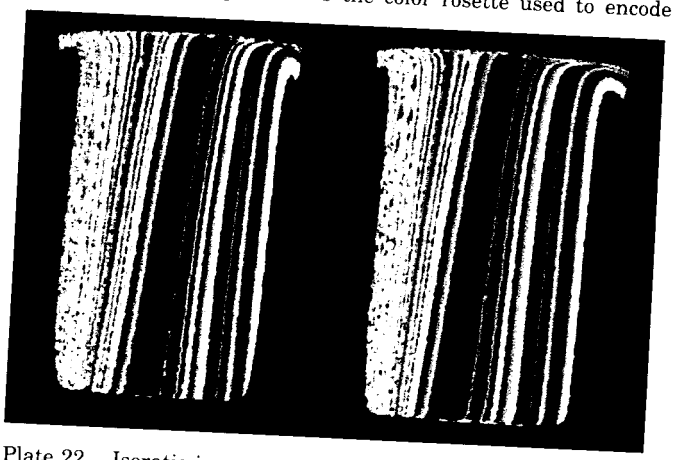

Plate 22. Isoratio image curves for (a) the left-hand and (b) the right-hand camera views of the cylinder.

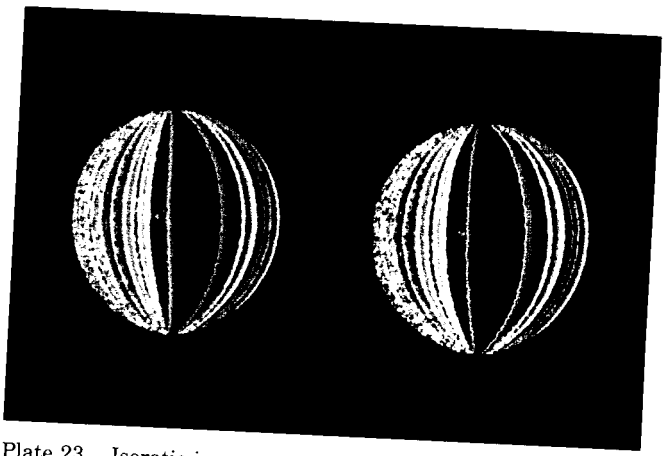

Plate 23. Isoratio image curves for (a) the left-hand and (b) the right-hand camera views of the sphere.

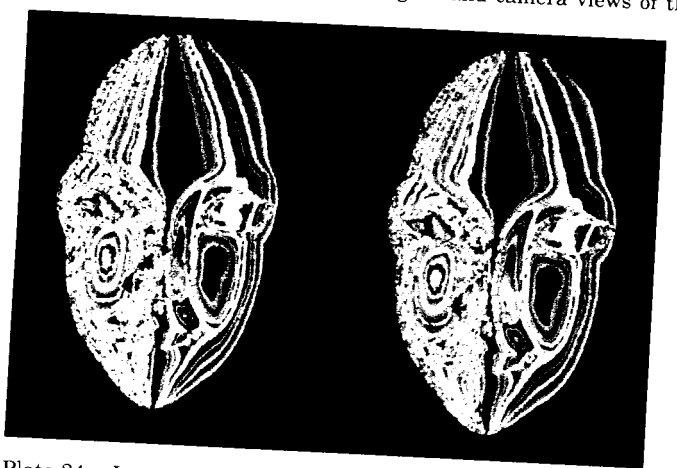

Plate 24. Isoratio image curves for (a) the left-hand and (b) the right-hand camera views of the face mask.