To Appear: International Journal of Computer Vision, 1996

Separation of Reflection Components Using Color and Polarization

Shree K. Nayar, Xi-Sheng Fang, and Terrance Boult

CUCS-058-92
November 1993

Department of Computer Science
Columbia University
New York, N.Y. 10027

Abstract

Specular reflections and interreflections produce strong highlights in brightness images. These highlights can cause vision algorithms for segmentation, shape from shading, binocular stereo, and motion estimation to produce erroneous results. A technique is developed for separating the specular and diffuse components of reflection from images. The approach is to use color and polarization information, simultaneously, to obtain constraints on the reflection components at each image point. Polarization yields local and independent estimates of the color of specular reflection. The result is a linear subspace in color space in which the local diffuse component must lie. This subspace constraint is applied to neighboring image points to determine the diffuse component. In contrast to previous separation algorithms, the proposed method can handle highlights on surfaces with substantial texture, smoothly varying diffuse reflectance, and varying material properties. The separation algorithm is applied to several complex scenes with textured objects and strong interreflections. The separation results are then used to solve three problems pertinent to visual perception; determining illumination color, estimating illumination direction, and shape recovery.

1 Introduction

Reflection of light from surfaces can be classified into two broad categories: diffuse and specular[1]. The *diffuse* component results from light rays penetrating the surface, undergoing multiple reflections and refractions, and re-emerging at the surface. It is distributed in a wide range of directions around the surface normal, giving the surface a matte appearance. If the viewing direction of an image sensor is varied, diffuse reflections from scene points change slowly, and in the ideal case of Lambertian surfaces they do not change at all. The *specular* component, on the other hand, is a surface phenomenon. Light rays incident on the surface are reflected such that the angle of reflection equals the angle of incidence. Even for marginally rough surfaces, specular reflections are concentrated in a compact lobe around the specular direction. This concentration of light energy causes strong *highlights* in brightness images of scenes. These highlights can cause vision algorithms for scene segmentation and shading analysis to produce erroneous results. If the sensor direction is varied, highlights shift, diminish rapidly, or suddenly appear in other parts of the scene. This strong directional dependence of specular reflection poses serious problems for vision techniques such as binocular stereo and motion detection. Specularities challenge the robustness of a large variety of vision sensors and algorithms and are most often undesirable in images.

In this paper, we present an algorithm that separates the diffuse and specular components of brightness from images. Separation of reflection components has been a topic of active research in the past few years. Initial work by Brestaff and Blake [6] addressed the problem of detecting specularities in gray-level (black and white) images by invoking the Lambertian constraint. Specularities are detected by identifying rapid brightness variations that violate Lambertian-like behavior. A different approach to the problem is based on color. Most of this work uses the *dichromatic reflectance model* proposed by Shafer [19]. The dichromatic model suggests that, in the case of dielectrics (non-conductors), the diffuse and specular components generally have different spectral distributions. The spectral distribution of the specular component is similar to that of the illumination, while the distribution of the diffuse component is altered by colorants in the surface medium. Hence, the color of an image point can be viewed as the sum of two vectors with different directions in color space. Using this model, Klinker et al. [11] and Gershon [8] independently observed that the color histogram of a surface with uniform diffuse reflectance takes the shape of a "skewed T" or "dog leg" with two limbs. One limb corresponds to purely diffuse points on the surface while the other represents a highlight region. , Klinker et al. [11] proposed an algorithm for automatically identifying the two limbs and used the directions of the limbs to separate the diffuse and specular color vectors at each surface point. Later, an algorithm was proposed that attempts to achieve separation and segmentation simultaneously [11].

[1]These are also referred to as body and surface components (see [19]).

Subsequently, Bajcsy et al. [2] and Novak and Shafer [16] showed that the color histogram of an object could have additional limbs that correspond to highlights caused by interreflections between surfaces. Lee [21] suggested moving a sensor and applying spectral differencing to color histograms of consecutive images to identify specular points in the scene. Recently, Sato and Ikeuchi [18] used local color signatures generated by actively varying the illumination direction to separate reflection components. This work uses the dichromatic model and singular value decomposition to decompose the color signatures into diffuse and specular signatures. These signatures can then be used to estimate surface orientation and local roughness estimates.

All of the above methods rely solely on color information to separate reflection components. Since separation is not possible when an image point is treated in isolation[2], most of the previous methods analyze the anatomy of color histograms. Two major limitations result from this approach. First, real scenes are complex and include objects with texture and at times gradually varying reflectance. Color histograms of such scenes are generally unpredictable and a set of linear clusters such as the skewed T is not guaranteed. Second, all points on a highlight region are assumed to have the same diffuse component. Even for an object with uniform reflectance, this assumption is valid only if the object surface is very smooth. In the case of rough surfaces, highlights spread over a wider range of surface normals with different diffuse components and the specular limb of the skewed T does not have a well-defined direction. This last observation was also made by Novak and Shafer [16] and Lee [21].

Taking a different approach, Wolff and Boult [23] proposed a polarization based method for separating reflection components from gray-level images. A polarization filter is placed in front of the image sensor and the filter is rotated to obtain a sequence of images. In general, the diffuse component tends to be unpolarized while the specular component varies with rotation of the polarization filter. This variation in the specular component is significant only in the case of dielectrics. Wolff and Boult [23] assume the scene consists of dielectric objects and use the variation in image brightness due to polarization to estimate the diffuse and specular components. Though the specular component varies with the position of the polarization filter, it also includes a constant specular offset that does not vary. This specular offset is not separable from the diffuse component which is also constant with respect to polarization. To overcome this problem, Wolff and Boult [23] use a large set of image points that lie on a pre-segmented (using percentage of polarization) highlight region. They assume that the diffuse component is constant over the entire highlight. In addition, the ratio of the Fresnel coefficients [4] is also assumed to be constant over the highlight. Using these assumptions, an estimate for the constant diffuse term is computed.

The above polarization based approach suffers from two problems. First, except

[2]Sato and Ikeuchi [18] rely solely on local color measurements but require a large number of measurements obtained by varying illumination direction.

for very smooth surfaces, highlights typically include a wide range of surface normals and the underlying diffuse component cannot be assumed constant. Furthermore, the diffuse component varies dramatically within a highlight region when the surface is textured. The second problem is related to the constant Fresnel ratio assumption. The Fresnel ratio is sensitive to the angle of incidence. As stated earlier, except for very smooth surfaces, highlights are expected to include points with different surface normals and hence different incidence angles. Therefore, the Fresnel ratio cannot in general be assumed constant within a highlight.

Here, we present a new algorithm for the separation of specular and diffuse reflections from images. This algorithm builds upon previous work on color and polarization. It uses color and polarization *simultaneously* to overcome many of the problems associated with the methods based on either color or polarization. We assume that the scene consists of dielectric objects. This leads to two further assumptions: (a) the dichromatic model is applicable, and (b) the specular component is polarized while the diffuse component is not. The proposed algorithm can estimate specular components that result not only from direct source illumination but also interreflections between points in the scene. We show that by varying the polarization filter, the color of the specular component can be determined independently for each image point. The result is a subspace in color space in which the diffuse vector must lie. This subspace imposes strong constraints on the color of the diffuse component of an image point. Neighboring diffuse colors that satisfy these constraints are used to compute the point's diffuse color vector.

Since the specular color of each image point is computed independently, the present approach has the following advantages over previous methods. (a) It does not assume that image highlights are pre-segmented as most previous methods do. (b) The diffuse component is not assumed to be constant under the highlight region; highlights can be removed from surfaces with strong textures. (c) The Fresnel ratio, which depends on the material properties and the angle of incidence, can vary within the highlights. The algorithm only requires that each image point has a few neighbors that have diffuse components with the same hue, i.e. direction in color space but not necessarily magnitude. Figure 1 shows results produced by the new algorithm for an object that previous algorithms cannot handle. The highlight spreads over a textured region and the color histogram of the highlight region turns out to be rather unstructured. As evidenced by Figure 1(b), the joint use of color and polarization enables accurate highlight removal. In the experimental section, we present several other results obtained by applying our separation algorithm to complex scenes with multiple highlights and interreflections. The algorithm is then used to solve three problems that are fundamental to visual perception: (a) color constancy; (b) estimation of source direction; and (c) shape recovery. The paper is concluded with a critique of the algorithm.

2 Reflection and Interreflection Mechanisms

We begin by describing the mechanisms involved in the processes of reflection and interreflection. Nayar et al. [15] proposed a reflectance framework that includes three primary components: the *diffuse lobe*; the *specular lobe*; and the *specular spike*. The last component exists only in the case of very smooth surfaces. In general, the two specular components can be combined into a single component. We refer to it as simply the *specular component*. Our objective is to remove specularities from images due to both single reflection as well as interreflection.

Figure 2 shows two points, A and B, in a scene. Reflection from the point A has two components, namely, diffuse and specular. The diffuse component arises from scattering of light rays that enter the surface and undergo multiple reflections and refractions due to inhomogeneities in the surface medium. The specular component, on the other hand, is a surface phenomenon and results from single reflection of incident light rays. The surface may be assumed to be composed of several planar elements, or facets, where each facet has its own orientation that differs from the large-scale local orientation of the surface. The result is a specular component that spreads around the specular direction, the width of the distribution depending on the roughness of the surface [22].

Now let us consider the phenomenon of interreflections. Points in the scene receive light not only from light sources but also from other scene points. Assume that point B reflects in the direction of the sensor, light energy that is received only from point A. The resulting image brightness can be viewed as the linear combination of four possible interreflection components[3]: (a) *diffuse-diffuse*; (b) *specular-diffuse*; (c) *diffuse-specular*; and (e) *specular-specular*. In each case, the first term represents the component received from point A and the second is the mechanism causing reflection from point B. In general, B could reflect light due to both direct illumination by sources as well as interreflections from other scene points.

Each brightness value in the image is viewed as the sum of two components, diffuse and specular. The specular component can result either from direct illumination by a light source or due to the diffuse-specular or specular-specular interreflections. It is assumed that the specular reflection received by the sensor from any given scene point, if significant, is either due to source illumination or due to interreflections and not both. In other words, any given scene point is positioned and oriented to specularly reflect from either a light source or another scene point but not both. As we will see shortly, this assumption is used ensure that the polarization characteristics of the specular component are well-behaved and easy to analyze. The above assumption holds well except for very rough surfaces that have wide specular distributions that increase the likelihood of specular components due to both direct illumination and interreflection being dominant

[3] A similar decomposition of interreflection mechanisms is presented by Shafer et al. [20].

(a)

(b)

Figure 1: Separation of reflection components using the proposed algorithm: (a) Image of an object with complex surface texture; (b) Image of the object after removal of specular reflections.

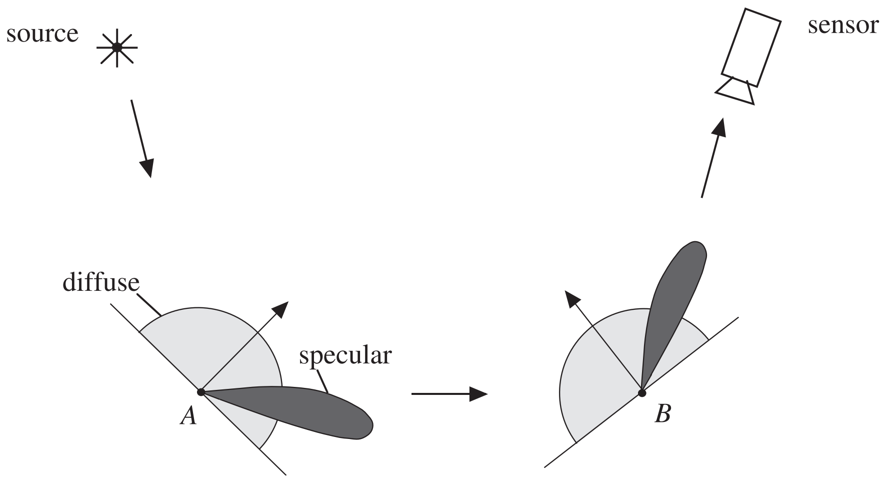


Figure 2: Reflection and interreflection.

in any given sensor direction. On the other hand, surfaces with such extreme roughness do not produce strong highlights and vision tasks such as scene segmentation are less affected by specularities.

3 Polarization

The method presented in this paper uses a polarization filter to determine the color of the specular component. In this section, we present a brief overview of polarization and discuss the type of surfaces for which it provides useful information. Detailed discussions on the theory of polarization can be found in [4]. Polarization methods have been used extensively in the past two decades in the area of remote sensing for segmentation and classification of satellite imagery. A summary of recent results can be found in [7]. In the field of machine vision, polarization methods were introduced by Koshikawa [13] who used ellipsometry for shape interpretation and recognition of glossy objects. More recently, Wolff and Boult [23] have used linear polarization for highlight removal and material classification. Boult and Wolff [5] have also studied the classification of scene edges based on their polarization characteristics.

Figure 3 shows a surface element illuminated by a source and imaged by a sensor. A polarization filter is placed in front of the sensor. From the previous section, we know that the image brightness value corresponding to the surface element can be written as:

$$I = I_{\mathrm{d}} + I_{\mathrm{s}} \tag{1}$$

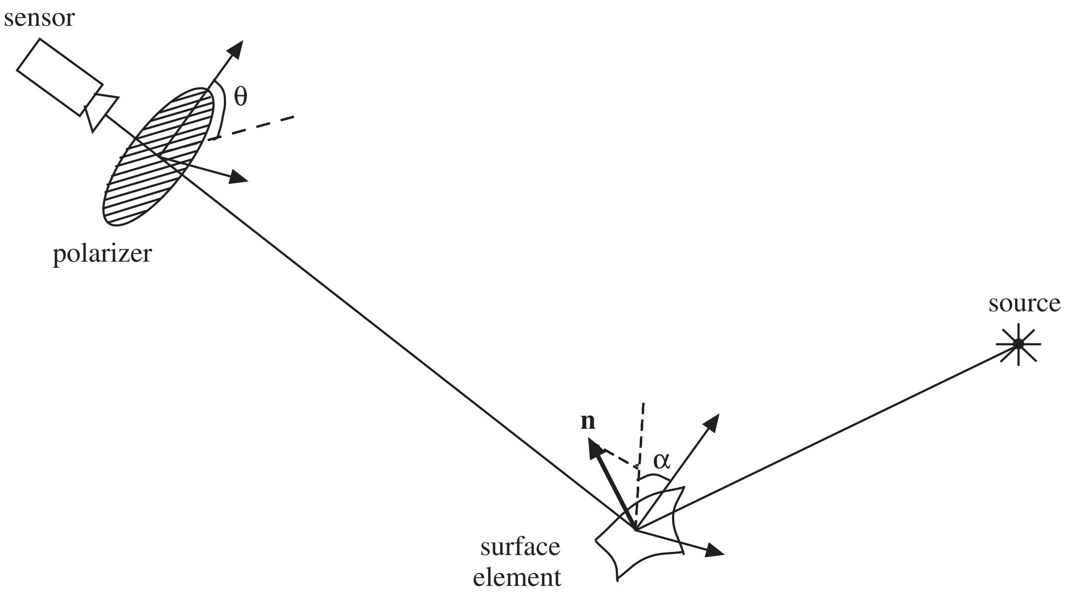


Figure 3: Surface element illuminated by a source and imaged through a polarization filter.

where, I_{d} is the diffuse component and I_{s} is the specular component. The above expression reflects brightness reductions due to the use of the polarization filter. The linear polarization of each light wave reflected by the surface element is determined by the direction of its electric field vector. For an ideal filter, a wave is passed unattenuated when its electric field is aligned with the polarization axis of the filter, and the energy in that wave is attenuated as a cosine function when the filter is rotated. In general, light energy (due to several light waves) reflected by a surface may be partially *polarized*. The extent of polarization depends on several factors including the material of the reflecting surface element, its orientation with respect to the image sensor, and the type of reflection mechanisms (specular or diffuse) at work. The diffuse component tends to be *unpolarized*[4], i.e. if the surface element is perfectly diffuse in reflectance, rotating the polarization filter does not alter image brightness. In contrast, specular reflection tends to be partially polarized. As a result, rotation of the filter varies the specular component as a cosine function, as shown in Figure 4. The specular component therefore can be expressed as the sum of a constant I_{sc} and a cosine function with amplitude I_{sv}:

$$I \;=\; I_{\mathrm{d}} + I_{\mathrm{sc}} + I_{\mathrm{sv}} \cos 2(\theta - \alpha) \tag{2}$$

where, θ is the angle of the polarization filter and α is the phase angle determined by the projection of the surface normal onto the plane of the filter (see Figure 3). The above expression is general in that it describes the variation of the specular component when the incident light is unpolarized, i.e. includes light waves of all polarizations.

[4] Boult and Wolff [5] note that this assumption does not hold near the occluding contour of an object.

In structured environments, such as industrial assembly lines, the illumination of the scene of interest can be controlled. In such cases, removal of specularities is straightforward. A polarization filter can be placed in front of the source to produce incident light that is linearly polarized; all incident light waves have field vectors in the same direction. Once again, the diffuse reflection tends to be more or less unpolarized due to the random nature of the diffuse scattering process that takes place beneath a surface. In contrast, specular reflection at a smooth[5] surface interface preserves the polarization of the incident light. Therefore, by using a polarization filter in front of the sensor with a $\pi/2$ rotation with respect to the filter at the source, the specular component is entirely blocked out producing an image with just the diffuse component. The above procedure is an old trick that has been extensively used in a variety industrial applications (see [1] [14] for examples).

Our purpose is to develop a passive technique for specularity removal that does not rely on controlled illumination. To this end, we shall assume the incident light to be unpolarized and return to expression (2). The exact values of I_{sc} and I_{sv} in (2) depend on the material properties and the angle of incidence ψ. These factors are represented by the *Fresnel reflection coefficients* [4], $F_{\perp}(\eta, \psi)$ and $F_{||}(\eta, \psi)$, that determine the polarization of reflected light waves in the directions perpendicular and parallel to the plane of incidence[6], respectively. The relationship between I_{sc} and I_{sv}, and the Fresnel coefficients is:

$$\frac{I_{\mathrm{sc}} + I_{\mathrm{sv}}}{I_{\mathrm{sc}} - I_{\mathrm{sv}}} = \frac{F_{\perp}(\eta, \psi)}{F_{||}(\eta, \psi)} = q \tag{3}$$

where, q is referred to as the *Fresnel ratio*. The Fresnel coefficients clearly determine the maximum and minimum brightness values obtained by varying the angle θ of the polarization filter. The parameter $\eta = n - i\,k$ is the *complex index of refraction* of the surface medium; n is the *simple index of refraction* and k is the *coefficient of extinction*. The parameter ψ is once again the angle of incidence (and reflection since light rays are specularly reflected). The complex index of refraction is determined by the physical properties of the reflecting material. The Fresnel coefficients for a typical metal and a typical dielectric [9] [10] are shown in Figure 5. For metals, the Fresnel coefficients are nearly equal except for incidence angles close to grazing (when ψ lies between 70 and 90 degrees). In other words, rotating the polarization filter does not alter the measured image brightness sufficiently to be able to compute the specular component or any surface properties. Passive polarization based methods are therefore not effective in the case of metals[7]. In the case of dielectrics (non-conductors), however, the two Fresnel coefficients differ substantially except for near-normal angles of incidence (when ψ is less than 20

[5]Surface roughness that is comparable in dimension to the wavelength of light causes partial *depolarization* of incident light waves (see [3]). Therefore, removal of specular reflection using polarized incident light is most effective for surfaces with low microscopic roughness.

[6]The plane of incidence includes the direction of illumination and the surface normal.

[7]There are a few exceptions to this rule. Some special metals, such as Gold, have Fresnel coefficient plots that resemble those of a dielectric (see [9]).

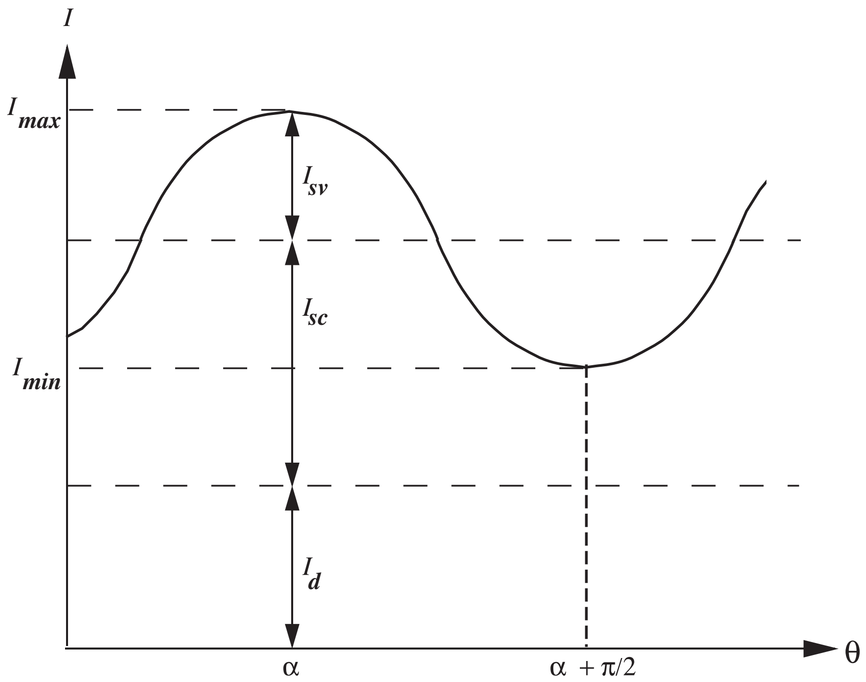


Figure 4: Image brightness plotted as a function of polarization filter angle.

degrees). It is interesting to note that the effect of varying the filter angle for a metal would be similar to that for a dielectric when the incidence angle is small.

Note that the terms I_{d} and I_{sc} in (2) are both constant and can be replaced by a single constant $I_{\mathrm{c}} = I_{\mathrm{d}} + I_{\mathrm{sc}}$:

$$I = I_{\mathrm{c}} + I_{\mathrm{sv}} \cos 2(\theta - \alpha) \tag{4}$$

For any given filter angle θ_i we have:

$$I_{\mathrm{i}} = I_{\mathrm{c}} + I_{\mathrm{sv}} \cos 2(\theta_i - \alpha) \tag{5}$$

The cosine term can be written in the following vector notation:

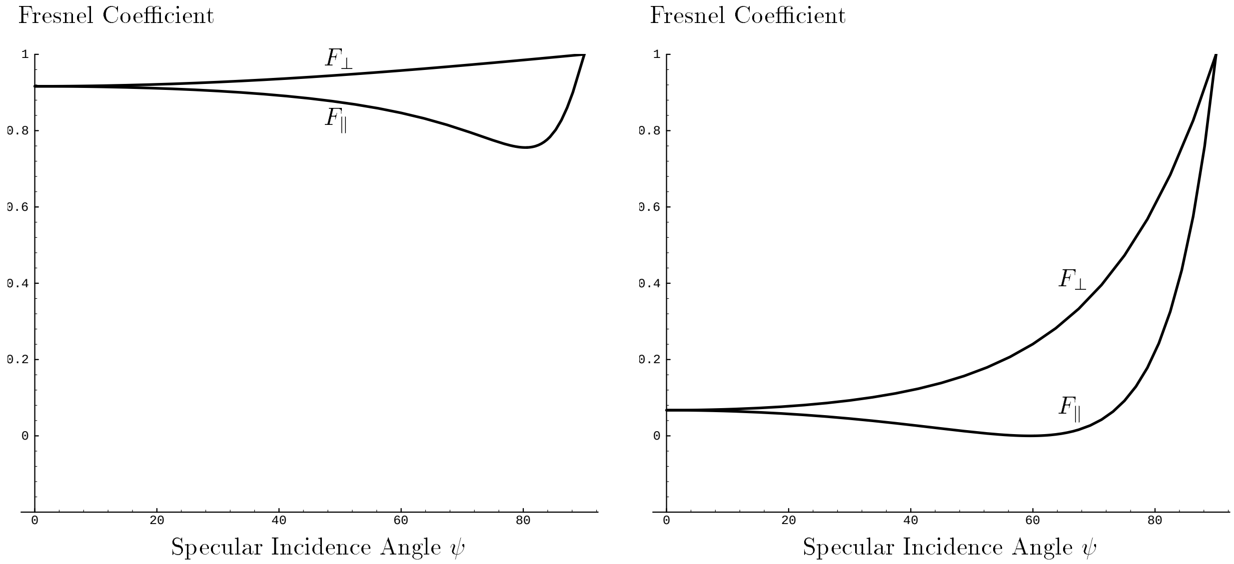
$$\mathbf{f}_{\mathrm{i}} = (1, \cos 2\theta_i, \sin 2\theta_i) \tag{6}$$

$$\mathbf{v} = (I_{\mathrm{c}}, I_{\mathrm{sv}} \cos 2\alpha, I_{\mathrm{sv}} \sin 2\alpha) \tag{7}$$

to obtain:

$$I_{\mathrm{i}} = \mathbf{f}_{\mathrm{i}} \cdot \mathbf{v} \tag{8}$$

This gives us a linear system of equations where $\mathbf{f}_{\mathrm{i}}$ is given for any filter setting and $\mathbf{v}$ is unknown. Let M be the total number of discrete filter angles used to obtain the set of image intensities $\{I_{\mathrm{i}} \mid i = 1, 2, ..., M\}$. If $M = 3$, the linear system can be solved to get $\mathbf{v}$, which in turn yields the parameters I_{c}, I_{sv}, and α. From (6) and Figure 4, one sees



(a) (b)

Figure 5: Typical fresnel coefficient plots for (a) a typical metal and (b) a typical dielectric.

that two solutions to α exit, namely, α and $\alpha + \pi$. This ambiguity cannot be resolved without using additional information about the scene or its illumination. If $M > 3$, we have an over-determined linear system that can be solved to get more robust estimates of I_c, I_{sv}, and α, in the presence of image noise. The above formulation of the problem using vectors saves substantial computations compared to the non-linear formulation of the type $a + b\sin^2(\theta - \alpha)$ used by Boult and Wolff [5] that requires the use of iterative non-linear estimation techniques.

From I_c and I_{sv}, the maximum and minimum values of image brightness are determined as:

$$\begin{aligned} I_{\min} &= I_c - I_{sv} \\ I_{\max} &= I_c + I_{sv} \end{aligned} \tag{9}$$

The *degree of polarization* at a scene point can be estimated as [4]:

$$\rho = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{10}$$

ρ lies between 0 and 1 and can be used during highlight removal to classify points into those that are only diffuse ($\rho \to 0$) and those that include a specular component. However, this measure must be used with caution since both $I_{\min}$ and $I_{\max}$ include the constant specular offset I_{sc} as well as the diffuse term I_d; varying either I_{sc} or I_d has the same effect on ρ.

We conclude this section with a note on previous work on highlight removal using polarization. Wolff and Boult [23] have presented a method for computing I_{d} and I_{s} by rotating the polarization filter. From the above discussion, we know that I_{d} and I_{sc} are both constant. They can be computed from I_{c} only if we know the Fresnel ratio, $q = F_{\perp}(\eta, \psi)/F_{||}(\eta, \psi)$, for the corresponding scene point. As stated earlier, the Fresnel coefficients are determined by the material properties of the scene point as well as the angle of incidence. Neither of these factors are known. To constrain the problem, Wolff and Boult [23] use all points (pixels) on a segmented highlight. They assume that both the diffuse component I_{d} as well as the Fresnel ratio q are constant under the highlight region, and estimate them using all points within the highlight region. This assumption, however, is limiting for two reasons: (a) In real scenes, the diffuse component I_{d} within a highlight is expected to vary due to the curvature of the surface and/or due to the surface being textured. (b) From Figure 5(b), we know that for dielectrics the Fresnel ratio q is sensitive to the angle of incidence ψ. Therefore, we cannot in general assume q to be constant within a highlight as ψ can vary substantially over large highlight regions resulting from extended sources in the scene.

4 Color

In this section, we review the representation of diffuse and specular components in color space and discuss previous work on the removal of highlights using color. In contrast to gray-level images, color images include wavelength (λ) dependence of the light reflected by a scene. Let $x(\lambda)$ be the spectral distribution of the light reflected by a point, and $s(\lambda)$ be the response of the sensor to wavelength. The brightness value measured by the sensor is:

$$I = \int x(\lambda)\, s(\lambda)\, d\lambda \tag{11}$$

Typically, color images are taken using three filters with responses $r(\lambda)$, $g(\lambda)$, and $b(\lambda)$. These response functions have their peaks close to the wavelengths that the human visual system perceives as "red," "green," and "blue." The resulting three brightness values measured by a sensor element constitute the color vector for the corresponding point in scene:

$$\mathbf{I}' = \left[I^{\mathrm{r}},\, I^{\mathrm{g}},\, I^{\mathrm{b}} \right]^T \tag{12}$$

The three components of the color vector are related to the spectral distribution of the reflected light as:

$$\begin{aligned} I^{\mathrm{r}} &= \int x(\lambda)\, r(\lambda)\, s(\lambda)\, d\lambda \qquad (13) \\ I^{\mathrm{g}} &= \int x(\lambda)\, g(\lambda)\, s(\lambda)\, d\lambda \\ I^{\mathrm{b}} &= \int x(\lambda)\, b(\lambda)\, s(\lambda)\, d\lambda \end{aligned}$$

As before, each brightness value includes a diffuse component and a specular component. Hence, in three-dimensional color space we have the following decomposition:

$$\mathbf{I}' \;=\; \mathbf{I}_{\mathrm{d}}{}' + \mathbf{I}_{\mathrm{s}}{}' \tag{14}$$

Shafer [19] proposed the *Dichromatic Reflectance Model* that suggests that for dielectrics, the spectral distribution of the diffuse component is determined by the colorant in the reflecting surface whereas the specular component preserves the spectral distribution of the incident light. Therefore, in general, the two vectors $\mathbf{I}_{\mathrm{d}}{}'$ and $\mathbf{I}_{\mathrm{s}}{}'$ have different directions in color space as illustrated in Figure 6. The two vectors will however have the same direction if, for instance, a gray object is illuminated by white light. In such cases, separation of reflection components using color is not feasible. Klinker et al. [11] and Gershon [8] used the dichromatic reflectance model to remove highlights from images. They observed that the color histogram of a single object with uniform diffuse color has two limbs forming what is referred to by Klinker as a skewed T (also see [12] for a detailed and insightful discussion). One limb results from diffuse reflections from object points that do not have a specular component, while the second limb corresponds to object points that lie on a highlight region and therefore include a specular component. Based on this observation, Klinker et al. [11] and Gershon [8] developed algorithms to automatically identify the two limbs of the skewed T and then separate the diffuse and specular components at each object point. Along similar lines, Bajcsy et al. [2] observed that the color histogram can include additional limbs that correspond to highlights caused by interreflections. The highlight removal algorithms of Klinker and Gershon can therefore be modified to identify additional limbs and remove specularities due to interreflections.

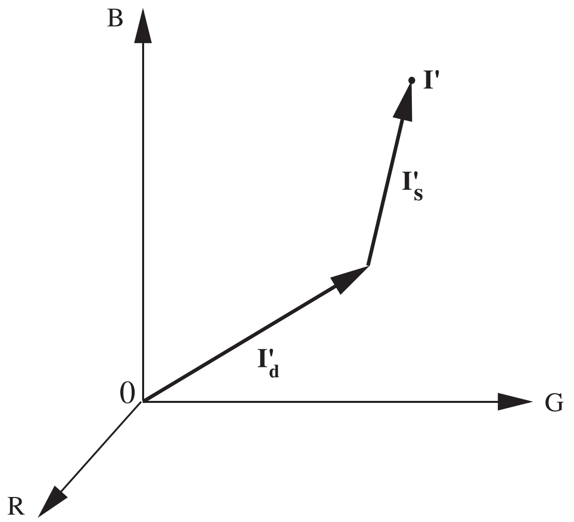


Figure 6: Diffuse and specular components in color space [19].

The above methods are based on two main assumptions. (a) The first assumption is that the object can be segmented away from the scene and has a single uniform diffuse color. Real scenes are complex with objects that may be textured or have patches with different reflectance properties. In either case, the color histogram of the object does not provide neat linear clusters as assumed by Klinker et al. [11] and Gershon [8]. In other words, the skewed T would not be recognizable from the histogram. (b) The second assumption is related to the roughness of the surface. A skewed T would result only if all points on the highlight have nearly equal diffuse components, $\mathbf{I}_{\mathrm{d}}'$. This assumption is valid only if the surface is very smooth. Even for marginally rough surfaces, the highlight is expected to include object points with a range of surface normals. In such cases, the specular limb of the skewed T spreads into a wide cluster in color space that is difficult to separate from the diffuse limb. A detailed discussion on this problem is given by Novak and Shafer [16].

5 Separation of Reflection Components

We now present the algorithm for separating reflection components from images. The following assumptions are made:

(A) The scene consists of dielectric objects. Hence, specular reflections and interreflections are polarized while diffuse reflections are not. Further, the color of incident light at each scene point is different from the color of the material. Consequently, the dichromatic model can be invoked.

(B) Specular interreflections results from either the diffuse-specular mechanism *or* the specular-specular mechanism but not both. In the first case, incident light is unpolarized and the Fresnel ratio is simply $q = F_{\perp}(\eta, \psi)/F_{||}(\eta, \psi)$. In the second case, incident light is already partially polarized and the effective Fresnel ratio for the surface point is $q = aF_{\perp}(\eta, \psi)/bF_{||}(\eta, \psi)$. The parameters a and b account for the partial polarization of the incident light [23].

(C) The Fresnel coefficients $F_{\perp}(\eta, \psi)$ and $F_{||}(\eta, \psi)$ are independent of the wavelength of incident light. This assumption is reasonable (see [10]) since we are operating in a narrow wavelength range, namely, the visible-light spectrum. Assumptions (B) and (C) result in the Fresnel ratio q being constant for all three color bands.

The separation technique is described by focusing on a single image point $\mathbf{x}$. The same procedure is applied independently to all image points. The color vector for $\mathbf{x}$ measured using a polarization filter is:

$$\mathbf{I} = \mathbf{I}_{\mathrm{d}} + \mathbf{I}_{\mathrm{s}} \tag{15}$$

Since the Fresnel ratio q is the same for all three color bands, rotating the filter produces three cosine functions that are in phase. For the filter position θ_i we have the color vector:

$$\mathbf{I}_{\mathrm{i}} = \mathbf{I}_{\mathrm{c}} + \mathbf{I}_{\mathrm{sv}} \cos 2(\theta_i - \alpha) \tag{16}$$

where, as in Section 3, $\mathbf{I}_c = \mathbf{I}_d + \mathbf{I}_{sc}$. Since rotating the polarization filter causes only the specular component to vary, the measurements $\mathbf{I}_i$ lie on a straight line L (a subspace) in color space as shown in Figure 7. For reasons given in Section 3, three filter positions are sufficient to determine the vectors $\mathbf{I}_c$, $\mathbf{I}_{sv}$, and phase α from the above expression. In our experiments, we have used 6 or more polarizer settings. This gives us an over-determined linear system that is solved to obtain robust estimates of $\mathbf{I}_c$, $\mathbf{I}_{sv}$, and α. Next, the vectors corresponding to maximum and minimum polarization are determined using (9):

$$\begin{aligned} \mathbf{I}_{max} &= \mathbf{I}_c + \mathbf{I}_{sv} \\ \mathbf{I}_{min} &= \mathbf{I}_c - \mathbf{I}_{sv} \end{aligned} \tag{17}$$

Then, the degree of polarization $[\rho_r, \rho_g, \rho_b]^T$ is computed from $\mathbf{I}_{max}$ and $\mathbf{I}_{min}$ using (10).

Two tests are used to determine if the image point $\mathbf{x}$ is to be processed any further. Since a point can have any arbitrary diffuse color $\mathbf{I}_d$, the three components of ρ are expected to be different. If $\rho = max\,(\rho_r, \rho_g, \rho_b)$ is less than a threshold T_1, the point is not sufficiently polarized and is assumed to be purely diffuse[8]. In this case, the next image point is examined.

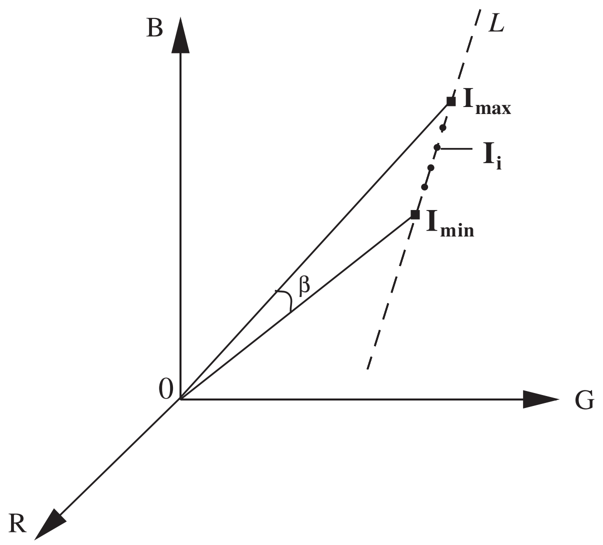


Figure 7: Specular line L in color space computed locally using polarization.

However, if the above test is successful, the angle β (Figure 7) subtended by the vector $2\mathbf{I}_{sv} = \mathbf{I}_{max} - \mathbf{I}_{min}$ from the origin $\mathbf{0}$ is computed:

$$\beta = \cos^{-1} \frac{\mathbf{I}_{max} \cdot \mathbf{I}_{min}}{||\,\mathbf{I}_{max}\,||\;||\,\mathbf{I}_{min}\,||} \tag{18}$$

[8]We are assuming here that the angle of incidence ψ is large enough for all dielectric scene points to be sensitive to polarization.

If β is less than a threshold T_2, the color of the specular component is very similar to that of the diffuse component $\mathbf{I}_d$, and the Dichromatic model cannot be used with confidence. Then, we have no robust means of estimating the relative contributions of the diffuse and specular components of $\mathbf{x}$.

If $\mathbf{x}$ is adequately polarized and its β value is sufficiently large, we proceed to estimate its diffuse component $\mathbf{I}_d$. If $\mathbf{I}_{sc}$ can be estimated, the diffuse component can be computed as $\mathbf{I}_d = \mathbf{I}_c - \mathbf{I}_{sc}$. Once this is done, the specular component $\mathbf{I}_s$ can be separated from any one of the measurements $\mathbf{I}_i$. Note, however, that the specular terms $\mathbf{I}_{sc}$ and $\mathbf{I}_{sv}$ are related as:

$$\mathbf{I}_{sc} + \mathbf{I}_{sv} \;=\; (\,\mathbf{I}_{sc} - \mathbf{I}_{sv}\,)\, q \tag{19}$$

Unfortunately, the Fresnel ratio q is not known as it depends on the local angle of incidence and the material properties of $\mathbf{x}$. Though we have estimates of $\mathbf{I}_c$ and $\mathbf{I}_{sv}$, there is no simple way of determining $\mathbf{I}_s$ and $\mathbf{I}_d$. Alternately, the dichromatic reflectance model requires that the diffuse vector $\mathbf{I}_d$ must lie on the specular line L (Figure 7), but its exact location on the line is not determinable since q is unknown.

Though we are unable to compute $\mathbf{I}_d$ locally, the specular line L provides useful constraints. These constraints are applied to the neighboring pixels of $\mathbf{x}$ to determine $\mathbf{I}_d$. For the purpose of discussion, assume that $\mathbf{I}_d$ corresponds to the point $\mathbf{P}$ on L (see Figure 8). The position of $\mathbf{P}$ on L can be parametrized as:

$$\mathbf{P} \;=\; \mathbf{I}_{min} - p\,\frac{\mathbf{I}_{sv}}{||\,\mathbf{I}_{sv}\,||} \tag{20}$$

where, p is distance of $\mathbf{P}$ from $\mathbf{I}_{min}$. The problem of finding $\mathbf{I}_d$ is therefore equivalent to finding p. The limits of p are $0 \le p \le \tilde{p}$ where $\tilde{p}$ corresponds to the point $\tilde{\mathbf{P}}$ at the intersection of L with one of the three planes of the color space. In the example shown in Figure 8, L intersects the $R-G$ plane. In general, however, L could intersect any one of the three planes that constitute the color space. The appropriate plane is easily found from $\mathbf{I}_{min}$ and $\mathbf{I}_{max}$.

Now, let $\mathcal{P}_1$ be the plane that passes through the points ($\mathbf{0}$, $\mathbf{I}_{min}$, $\mathbf{I}_{max}$). The normal to $\mathcal{P}_1$ is:

$$\mathbf{N}_1 \;=\; \mathbf{I}_{max} \times \mathbf{I}_{min} \tag{21}$$

Since we do not have sufficient constraints to compute the diffuse vector $\mathbf{I}_d$ from the color measurements $\mathbf{I}_i$, we use neighboring image points that satisfy the following constraints:

(1) A neighboring image point $\mathbf{y}$ is used if it has a low degree of polarization ρ and hence can be assumed to be purely diffuse. Alternatively, it can be used if its diffuse component has already been computed. In either case, the diffuse component of $\mathbf{y}$ is referred to as $\mathbf{Q}$.

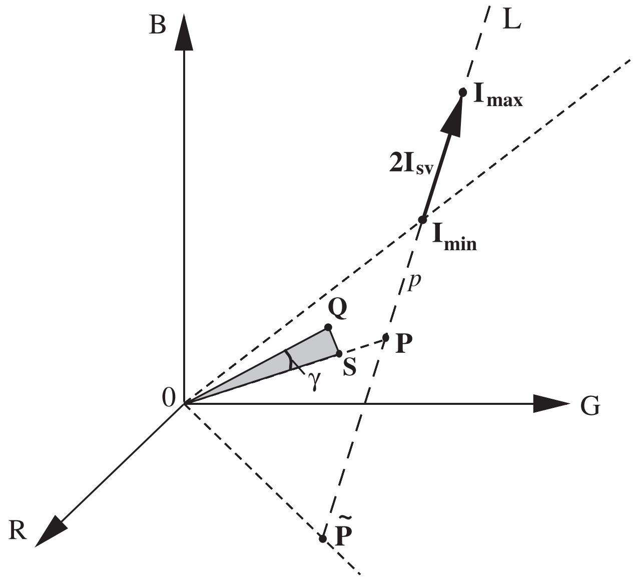


Figure 8: Using neighboring points and the specular line constraint to compute the diffuse component.

(2) In color space, $\mathbf{Q}$ must lie close to plane $\mathcal{P}_1$. Further, it must lie between the vectors $\mathbf{I}_{\mathrm{min}}$ and $\tilde{\mathbf{P}}$ since the diffuse vector $\mathbf{I}_{\mathrm{d}}$ of $\mathbf{x}$ lies on line L between $\mathbf{I}_{\mathrm{min}}$ and $\tilde{\mathbf{P}}$. Note that the magnitude of $\mathbf{Q}$ can differ from that of $\mathbf{I}_{\mathrm{d}}$ since the neighboring point $\mathbf{y}$ could have a different surface normal from $\mathbf{x}$.

If these constraints are satisfied, $\mathbf{y}$ is hypothesized to have the same diffuse color (direction but not necessarily magnitude) as $\mathbf{x}$. Then, the line passing through $\mathbf{Q}$ and the specular line L intersect to give $\mathbf{P}$, an estimate of the diffuse vector $\mathbf{I}_{\mathrm{d}}$. This process is repeated for all neighboring points in a small neighborhood around $\mathbf{x}$. If a large number of $\mathbf{P}$ estimates are obtained (i.e. a large number of neighbors satisfy constraints (1) and (2)), and if these estimates form a compact cluster on L, $\mathbf{I}_{\mathrm{d}}$ is computed as an average of all $\mathbf{P}$ estimates.

Due to noise in the color and polarization measurements, the neighboring diffuse component $\mathbf{Q}$ is not expected to exactly satisfy the above constraints. To accommodate such discrepancies, the angle γ (Figure 8) subtended by $\mathbf{Q}$ with respect to $\mathcal{P}_1$ is computed:

$$\gamma \;=\; \sin^{-1} \frac{\mathbf{N}_1 \,.\, \mathbf{Q}}{||\, \mathbf{N}_1 \,||\; ||\, \mathbf{Q} \,||} \tag{22}$$

If γ is larger than a threshold T_3, $\mathbf{Q}$ does not lie close to $\mathcal{P}_1$ and is not used any further. Note that γ is used rather than the distance of $\mathbf{Q}$ from $\mathcal{P}_1$. This is because a neighboring point may have a very small diffuse component that lies close to the origin and as a result is also close to plane $\mathcal{P}_1$. Such a point could have relatively large errors due to image noise and must not be used in the computation of $\mathbf{I}_{\mathrm{d}}$.

If $\gamma \leq T_3$, $\mathbf{P}$ is found by extending the projection $\mathbf{S}$ of $\mathbf{Q}$ on $\mathcal{P}_1$, to intersect L as shown in Figure 8. This can be done without computing the projection $\mathbf{S}$. Consider the plane $\mathcal{P}_2$ that passes through the points $(\mathbf{0}, \mathbf{Q}, \mathbf{P})$. Its normal is:

$$\mathbf{N}_2 \;=\; \mathbf{Q} \times \mathbf{P} \tag{23}$$

Since the planes $\mathcal{P}_1$ and $\mathcal{P}_2$ are perpendicular, we have:

$$\mathbf{N}_1 \,.\, \mathbf{N}_2 \;=\; 0 \tag{24}$$

Using (21) and (23), we get:

$$(\,\mathbf{I}_{\mathrm{max}} \,\times\, \mathbf{I}_{\mathrm{min}}\,) \,.\, (\,\mathbf{Q} \,\times\, \mathbf{P}\,) \;=\; 0 \tag{25}$$

Substituting the expression for $\mathbf{P}$ given by (20), a direct solution to the line parameter p is obtained:

$$p \;=\; ||\,\mathbf{I}_{\mathrm{sv}}\,|| \;\frac{\mathbf{N}_1 \,.\, (\,\mathbf{Q} \,\times\, \mathbf{I}_{\mathrm{min}}\,)}{\mathbf{N}_1 \,.\, (\,\mathbf{Q} \,\times\, \mathbf{I}_{\mathrm{sv}}\,)} \tag{26}$$

Constraint (2) requires that $\mathbf{P}$ lies below $\mathbf{I}_{\mathrm{min}}$. This is easily verified by checking if $p \geq 0$.

The above process is repeated for all neighboring diffuse components $\mathbf{Q}_{\mathrm{j}}$ that satisfy constraints (1) and (2). The result is a set of estimates $\{p_{\mathrm{j}} \mid j = 1, 2, ..., N\}$. If $N \geq T_4$ (we use $T_4 = 3$ in our implementation), the mean and standard deviation of p_{j} are computed:

$$\overline{p} \;=\; \frac{\sum_{j=1}^{N} \, w_{\mathrm{j}}\, p_{\mathrm{j}}}{\sum_{j=1}^{N} \, w_{\mathrm{j}}} \qquad\qquad \sigma_{\mathrm{p}} \;=\; \frac{\sum_{j=1}^{N} \, w_{\mathrm{j}} \left(\overline{p} \,-\, p_{\mathrm{j}}\right)^2}{(N-1)\,\sum_{j=1}^{N} \, w_{\mathrm{j}}} \tag{27}$$

where, $w_{\mathrm{j}} \;= ||\; \mathbf{Q}_{\mathrm{j}} \;||$. The weights w_{j} are used since neighboring points with larger diffuse components are more robust in the presence of image noise. The mean value $\overline{p}$ is accepted if the standard deviation σ_{p} is less than a threshold T_5, i.e. the estimates p_{j} form a compact cluster on line L. This constraint is used to ensure that different diffuse colors in the neighborhood of $\mathbf{x}$ that happen to lie close to plane $\mathcal{P}_1$, are not used together to produce an erroneous estimate of $\mathbf{I}_{\mathrm{d}}$. Given $\overline{p}$, the diffuse component of $\mathbf{x}$ is determined from (20) as:

$$\mathbf{I}_{\mathrm{d}} \;=\; \mathbf{I}_{\mathrm{min}} \,-\, \overline{p}\,\frac{\mathbf{I}_{\mathrm{sv}}}{||\,\mathbf{I}_{\mathrm{sv}}\,||} \tag{28}$$

$\mathbf{I}_{\mathrm{d}}$ is simply subtracted from any of the images $\mathbf{I}_{\mathrm{i}}$ to get the corresponding specular component $\mathbf{I}_{\mathrm{s}}$.

Figure 9 compares the above algorithm with previous techniques based on color or polarization. Figure 9(a) shows a 30×30 window clipped from the 512×480 image shown in Figure 1. The window includes a highlight that spreads over regions with different diffuse colors. Figure 9(b) shows the histogram of the window in (R,G,B) color space. Note that the anatomy of the histogram is very complex and does not lend itself to the skewed T analysis used in [11]. Figure 9(c) shows the cluster obtained in $(I_{\min},I_{\max})$ space for the green band of the window. The green band was selected in this example as the average degree of polarization was found to be maximum in this band. The polarization based method for highlight removal proposed in [23] is based on the assumption that this cluster in $(I_{\min},I_{\max})$ space is linear. As seen in the Figure 9(c), the cluster does not form a straight line and hence the polarization based method is not viable. Finally, Figures 9(d) and 9(e) show results obtained using the algorithm proposed in this section. Figure 9(d) shows $\mathbf{I}_{\mathrm{min}}$, $\mathbf{I}_{\mathrm{max}}$, and the specular line L, computed for the center pixel of the window in Figure 9(a). The color space constraints (1) and (2) are used to select neighboring diffuse colors and compute the diffuse component $\mathbf{I}_{\mathrm{d}}$ of the center pixel.

The separation algorithm is applied to all pixels in an image. Figure 9(e) shows the final diffuse image obtained after removal of specularities from the image in Figure 9(a). Generally, not all image points end up with an $\mathbf{I}_{\mathrm{d}}$ estimate. A point may lie in the middle of a very large highlight, in which case, it may not have a sufficient number of neighbors with diffuse colors that satisfy constraints (1) and (2), or produce a compact cluster of intersection points on the specular line. Therefore, constraints (1) and (2) are applied repeatedly to the image. This iterative approach is effective in the case of complex scenes; each iteration provides a new set of computed diffuse colors, increasing the likelihood of finding neighboring diffuse colors that satisfy constraints (1) and (2) in the next iteration. A large highlight, for instance, shrinks in size with each iteration. The iterations are discontinued when no new diffuse estimates are obtained.

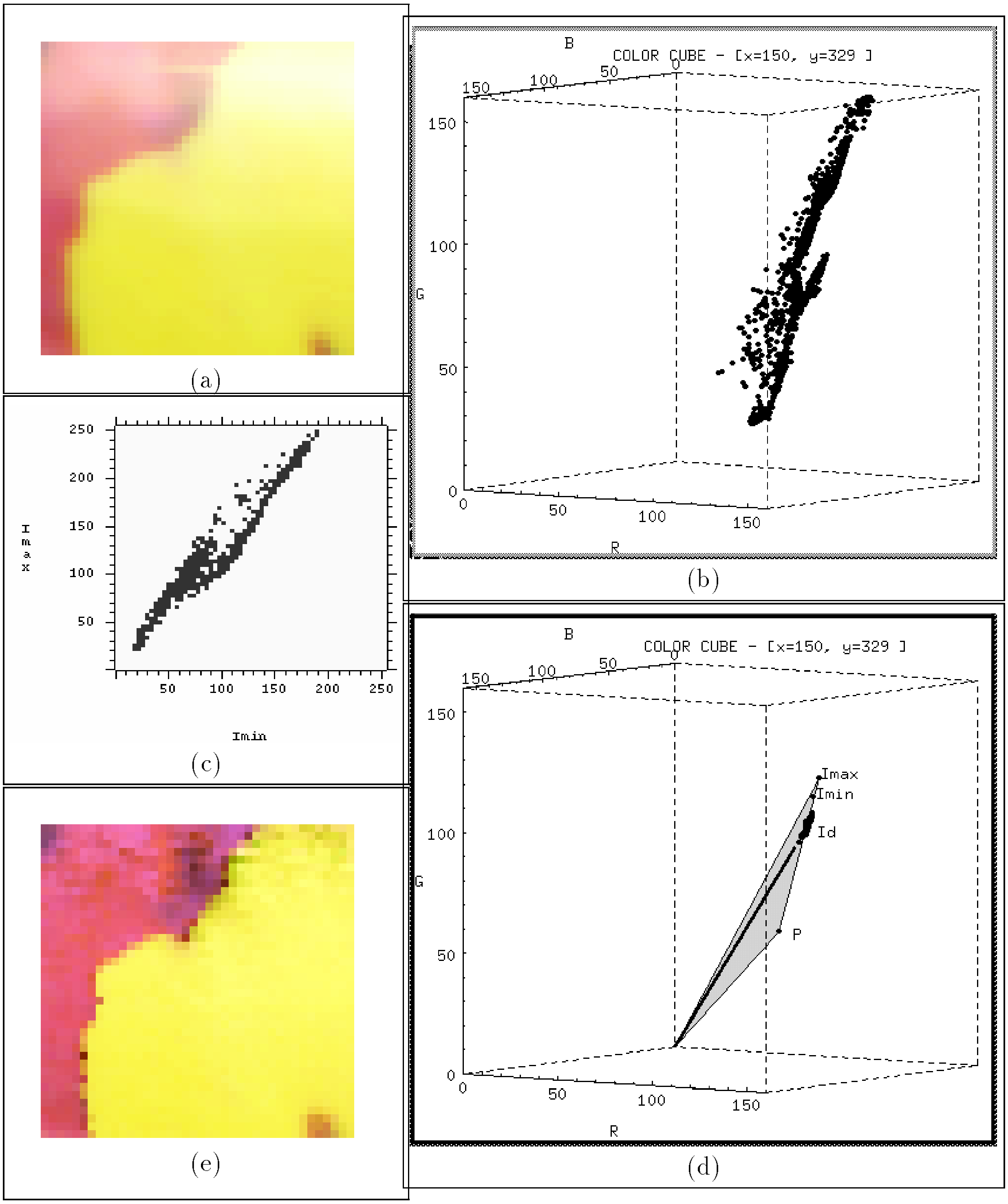


Figure 9: (a) A 30×30 image window that includes a highlight. (b) Color histogram of the image. The anatomy of the histogram is too complex to identify the skewed T used in [11]. (c) Cluster in ($I_{\min}$,$I_{\max}$) space used by the polarization based proposed in [23]. This cluster must approximate a single straight line (which it clearly does not) for the method in [23] to work. (d) Separation of diffuse and specular components of the center pixel of the image using the proposed method based on color and polarization. (e) Diffuse image after removal of specular reflections from the entire image in (a).

6 Implementation

This section describes the implementation of the separation algorithm. First, a brief description of the experimental setup is given. Next, we discuss the selection of algorithm thresholds and the estimation of parameters used by the separation procedure. This is followed by results obtained by applying the algorithm to scenes with textured objects, primary highlights, and interreflections.

6.1 Experimental Setup

The experiments reported here were conducted using two different setups. The primary setup uses a Javelin JE3012 single-chip color camera that produces a video signal in NTSC format. The video signal is digitized using a Sun VideoPix card to obtain a 640x480 RGB image. Due to the low signal-to-noise ratio of this particular camera, 32 frames were averaged to reduce image noise. The automatic gain control on the camera was not completely defeated. In addition, the camera has a non-unity gamma factor. This has the advantage of allowing a larger dynamic range, but reduces brightness resolution. Six "gray" patches on a McBeth color chart were used to calibrate the camera. A simple linear interpolation between the known reflectances of the six patches is used to determine the relation between image brightness and scene radiance. A Fujinon 25mm lens is used to image the scene. No steps were taken to correct the chromatic aberration effects of the lens.

The second experimental setup uses a black and white camera and 3 separate color filters, Tiffen 25(R), 57(G) 47B(B), to acquire RGB images. The camera used is a Sony XC-77 model with a Cosmicar 75mm lens. Again, no chromatic aberration correction was done. 512x480 images were digitized (in this case without temporal averaging) using a Matrox frame grabber. Since the color filters are not matched to the camera's spectral response, the blue channel required a lens aperture that is one F-stop larger than the other two channels. The response of the camera is linear in the range 30-230, which is the range used in the experiments.

In both setups, a linear polarizing filter, mounted in a precision rotation ring, is placed in front of the camera lens. Images are taken at various orientations of the polarizer. We have used 10 polarizer orientations for the first setup and 6 for the second setup. The present setups have a few minor problems that can be overcome using a precision filter mount and simple correction methods. The first is that the movement of the polarizing filter can cause slight spatial shifts of approximately 1 pixel between consecutive images. Secondly, we have not corrected for chromatic aberration of the lens. Finally, we are using CCD technology which is prone to blooming affects near strong highlight regions. All of these problems manifest themselves as errors in polarization fitting (and hence separation) in small neighborhoods of scene boundaries.

6.2 Parameter Estimation and Threshold Selection

For each scene, a set of color images is taken by rotating the polarization filter. The images are first gamma-corrected using the calibration data. Then, the polarization parameters $I_{\min}$, $I_{\max}$, and α are computed for each color channel by fitting. We have experimented with different fitting techniques to improve resilience to noise. For the images taken using the black-and-white camera and separate color filters, we found that a linear least-squares (LS) fitting with just 6 images is sufficient to obtain robust parameter estimates. For the images taken with the color camera, however, stable polarization fitting requires more data. For this setup we compute polarization parameters using sets of 5 pixels at a time. These include the pixel of interest and 4 connected neighbors. Since 10 images are used for this setup, a total of 50 points are involved in each fitting. This greatly over constrains the fitting problem and provides stable polarization parameters at the expense of minor blurring effects in the computed images.

The results of the polarization fitting are 6 images for each color channel: $I_{\min}$, $I_{\max}$, I_{avg} ($= \frac{I_{\max}+I_{\min}}{2}$), ρ (degree of polarization), α (phase angle), and root-mean-square-error (RMSE) in fitting. Of these, only the $I_{\max}$, $I_{\min}$, and ρ images are directly used by the separation algorithm. The rest are used to debug the algorithm and analyze the results. The I_{avg} image is what would be obtained without a polarizer, but instead with a 50% neutral density filter. The RMSE gives the error in fitting image data to expression (2). Most pixels in the image have fitting errors that are less than 0.3 gray levels. If the data is not consistent with the cosine model for polarization, one or more of the assumptions made by the algorithm are violated. Hence, pixels with RMSE values greater than 6 gray levels are marked as outliers and are not used for separation[9].

The separation algorithm requires labeling of points as purely diffuse or partially specular, which is done by using the degree of polarization ρ. Since ρ depends on $I_{\min}$, the noise level in ρ varies with $I_{\min}$. Therefore, rather than applying a fixed threshold T_1 to ρ to identify partially specular points, our implementation uses a threshold that varies with $I_{\min}$; T_1 varies from around 5% for points with $I_{\min} > 200$ to around 10% when $I_{\min} \approx 20$. In general, the selection of different values for T_1 should be done carefully; labeling a specular point as diffuse may cause other pixels to be incorrectly computed. On the other hand, if T_1 is too low, diffuse points may be labeled as specular, reducing the number of pixels that have a sufficient number of neighboring diffuse points. In extreme cases, the iterative application of the algorithm may not help. We have encountered this problem with scenes that include very large highlights on planar regions.

The algorithm has three other thresholds that affect its performance; the threshold T_2 for the angle β in (18), T_3 for the angle γ in (22), and T_4 for the standard deviation σ_{p} in (27). The algorithm is not too sensitive to T_2 and T_4, which are set at 0.08

[9]Due to chromatic aberration and mis-registration of consecutive images, the polarization fitting tends to be unreliable at reflectance edges. This causes errors in separation within 2-3 pixels around edge pixels.

and $cos^{-1}(0.99)$, respectively. The angle threshold T_3, which determines if neighboring diffuse vectors lie close to plane $\mathcal{P}_1$, has a strong affect on separation results as well as computation time. If this number if large, the constraint on the diffuse color of each point is weak. In regions where the diffuse color varies smoothly, or around edges (where blurring, chromatic aberration, and mis-registration of images cause color shifts), T_3 can strongly affect the computed diffuse component. If T_3 is too small, few new diffuse estimates are computed in each iteration. The current implementation starts with a relatively small threshold value, T_3=0.02, and doubles it after every 10 iterations. The algorithm typically converges in about 30 iterations.

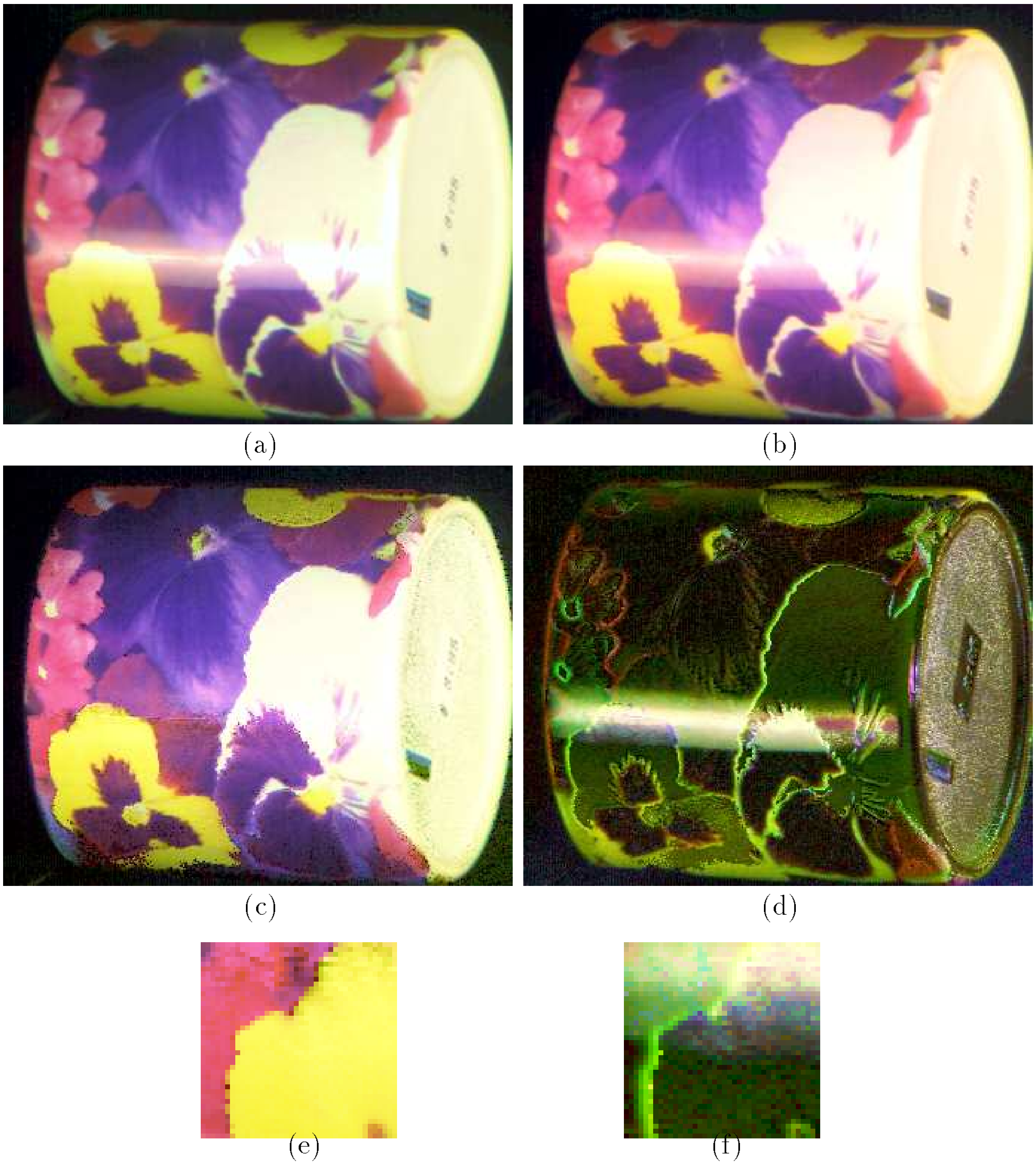
6.3 Separation Results

In each of the following examples, images produced by the polarization fitting are used to compute the diffuse color image $\mathbf{I}_{\mathrm{d}}$. This image is then subtracted from the average color image $\mathbf{I}_{\mathrm{avg}}$ to get the specular color image $\mathbf{I}_{\mathrm{s}}$. The first example shown in Figure 10(a) was briefly discussed in the introduction of the paper. It is the $\mathbf{I}_{\mathrm{avg}}$ image of a mug with a flowered pattern on it. The petals of the flower are of different colors and within each petal there is a moderate amount of diffuse color variation. Along the middle of the mug is a large horizontal highlight. Figure 10(b) shows the $\mathbf{I}_{\mathrm{min}}$ image produced by the polarization fitting process. This image represents the best image (with minimum specular component) that can be obtained by simply rotating the polarization filter. Figure 10(c) and Figure 10(d) show the diffuse image $\mathbf{I}_{\mathrm{d}}$ and the specular image $\mathbf{I}_{\mathrm{s}}$ computed using the separation algorithm. Figure 10(e) and Figure 10(f) show close-up views of the diffuse and specular images for a small section of the mug. Clearly, the algorithm was successful in separating the two reflection components, despite the strong texture underlying the highlight. For reasons given earlier, errors in separation are largely confined to reflectance edges on the surface.

The second example includes strong interreflection effects. Figure 11(a) shows the average image ($\mathbf{I}_{\mathrm{avg}}$) of a scene including a blue plastic plate and a part of a McBeth color chart. Color patches on the chart are reflected by the plate. There are pieces of plastic tape, some dark reddish and others black, stuck on the plate. Also visible is a film canister, which interreflects portions of the color chart as well as the surrounding environment. Figure 11(b) and Figure 11(c) show the diffuse and specular images computed by the algorithm. We see that, despite the strong interreflections, these images are quite accurate. The primary highlight on the left side of plate however is not completely removed. This may have been caused by high brightness values in the highlight region for which the camera calibration is not reliable.

The final example shown in Figure 12(a) is a complex scene for which the algorithm partially fails. Figure 12(b) shows the computed diffuse component. The primary highlights on the blue and red tori are accurately removed. Figure 12(c) and Figure 12(d) show details of the results for the blue torus. The diffuse and specular components are well separated in the highlight region. Errors in separation are however seen on the occluding boundary of the torus. This most likely results from the strong polarization of the diffuse component on the occluding boundary; the assumption that the diffuse component is unpolarized is violated.

In Figure 12(b), most of the interreflections on the left wall of the red box, such as the interreflection of the white cup in the upper left corner of the scene, are successfully removed. However, the results are not as good for the interreflections of the blue torus and the marker pen. This is due to the following reasons. First, the diffuse red color in these interreflection areas is very bright, resulting in a low (3-4 %) degree of polarization. As a result, many of the pixels are incorrectly labeled as "diffuse." Second, because of the geometry of the scene, the walls of the box interreflect between themselves, producing specular reflections that have the same color (red) as the diffuse component. As a result, the specular and diffuse colors could not be separated even when the algorithm thresholds were lowered to force computation. Therefore, it is not that the algorithm has removed too much red in the interreflections of the blue torus and the pen, but rather it was unable to remove the red specular interreflection from the red wall. The dichromatic reflectance assumption breaks down in such cases leading to erroneous separation results.

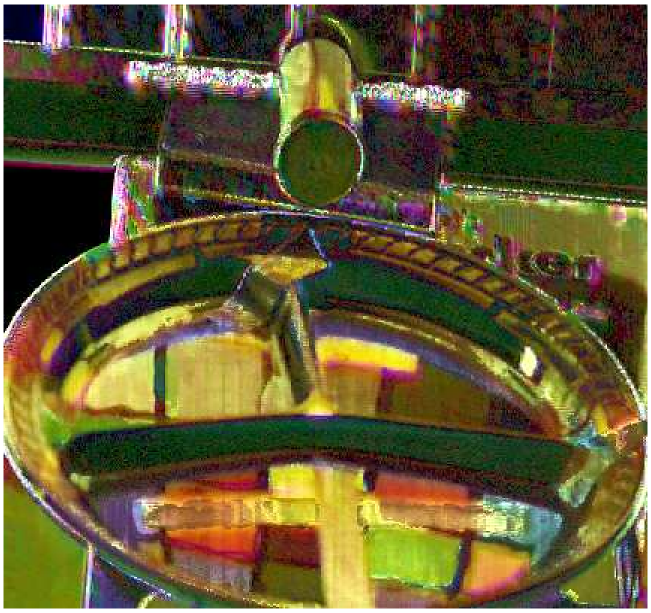


(a) (b) (c) (d) (e) (f)

Figure 10: Results obtained for a mug with diffuse color texture. (a) Original image ($\mathbf{I}_{\rm avg}$) with highlight across texture region. (b) Image expected for the polarizer angle that produces minimum specular reflection ($\mathbf{I}_{\rm min}$). (c) Diffuse component ($\mathbf{I}_{\rm d}$) computed by the algorithm. (d) Specular component ($\mathbf{I}_{\rm s}$) obtained by subtracting $\mathbf{I}_{\rm d}$ from $\mathbf{I}_{\rm avg}$. Errors in the specular component on region boundaries are due to chromatic aberration effects and mis-registration between polarization images. (e) Diffuse component ($\mathbf{I}_{\rm d}$) for a small image window. (f) Specular component ($\mathbf{I}_{\rm s}$) of the image window.

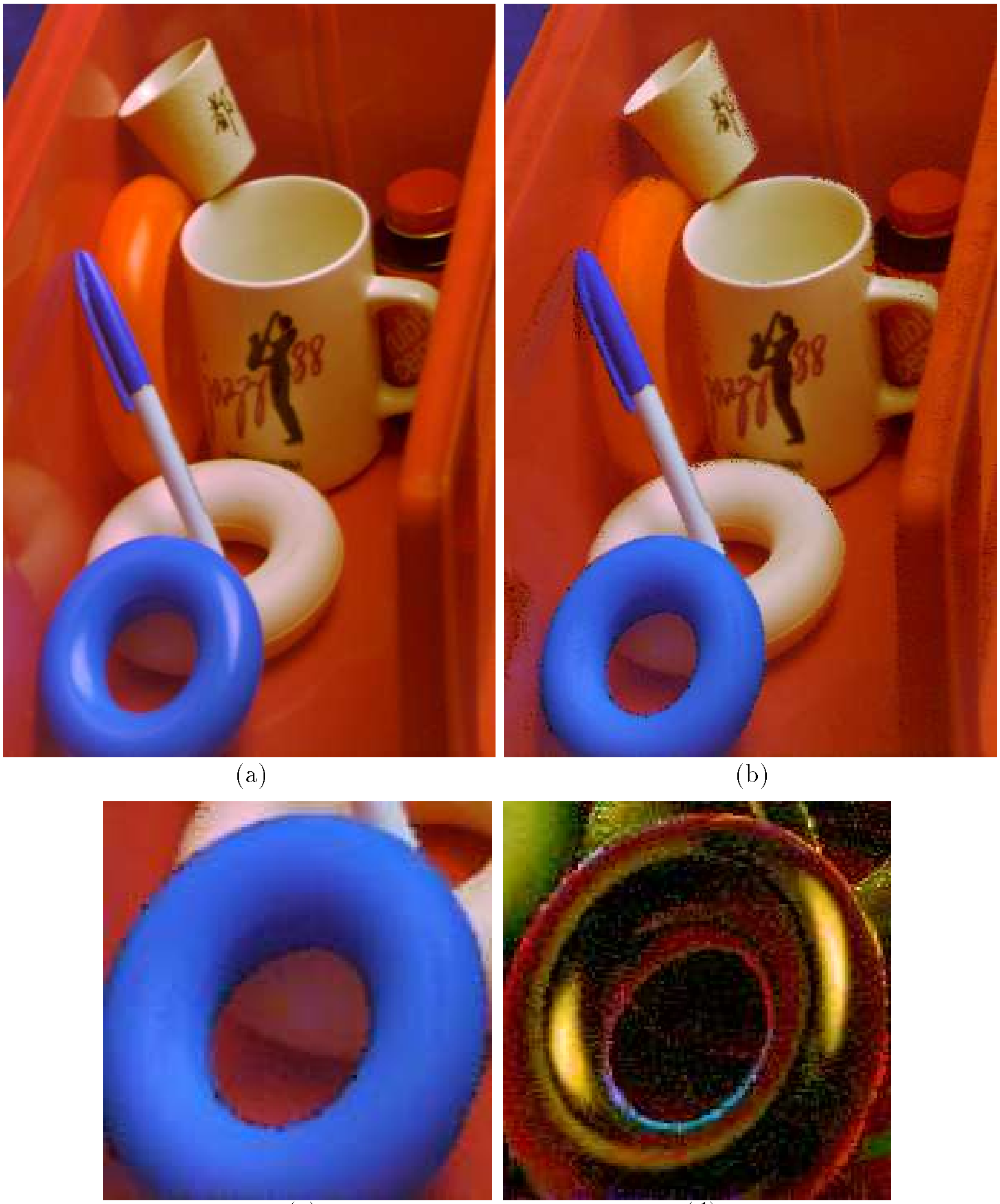
(a)

(b)



(c)

Figure 11: (a) A scene with strong interreflections. The blue plastic plate reflects a part of the McBeth color chart. There are pieces of plastic tape (dark red and black) on the plate. The film canister interreflects from the chart and from the surrounding room. (b) Diffuse and (c) specular components recovered by the algorithm.



(a) (b) (c) (d)

Figure 12: (a) Complex scene with strong primary reflections as well as interreflections. (b) Diffuse component computed by the algorithm (see text for details). (c) Diffuse component of the blue torus. (d) Specular component.

7 Applications

An obvious application of the separation algorithm is simply the removal of specularities from images. Specularities pose severe problems for a variety of vision algorithms, including, segmentation, shape from shading, binocular stereo, and structure from motion. For a given scene, a sequence of polarization images can be taken using a color image sensor and the separation algorithm then applied to recover a diffuse image. The diffuse image is invariant to viewing direction only if the scene is Lambertian. However, even for non-Lambertian surfaces (see [17], [24]), diffuse reflection varies gradually with viewing direction in comparison to specular reflections. The diffuse image is therefore valuable for a variety of vision applications.

Here, we shall explore other applications of the computed diffuse and specular images. The emphasis will be on the explicit use of the specular component to estimate important scene parameters such as illumination color and illumination direction. These parameters are then used in conjunction with the diffuse images to color compensate images and recover the shapes of objects in a scene.

7.1 Illumination Color and Color Constancy

The measured color of a scene point depends on its spectral response to incident light as well as the spectral distribution of incident light. In many vision applications, the spectral content of the illumination is not known and hence it is difficult to determine the spectral characteristics of the reflecting surface without making restrictive assumptions about the scene. The dichromatic model implies that for dielectrics the color of specular reflection is the color of illumination. This fact can therefore be used to determine the color of illumination from the computed specular image and then used to estimate the spectral characteristics of points in the scene. Note that we have represented color using a finite-dimensional (R, G, B) model. As a result, it is not possible to recover the complete continuous spectral response of a scene point. It is only possible to get approximations of images of the same scene under different illuminations.

Figure 13(a) shows the image of a blue torus and the McBeth color chart taken under white-light illumination. We shall use this image to verify our color compensation results. The same scene was illuminated by yellow light to get the image shown in Figure 13(b). The separation algorithm was applied to the scene with yellow-light illumination. The resulting specular image was thresholded to extract image areas with strong specular highlights. The two highlights on the blue torus were detected and the specular vectors $\mathbf{I}_s$ within both highlights were averaged to estimate the color of the illumination. Simple averaging suffices since it gives preference to larger specular components. The computed illumination color was then used to normalize the diffuse color at each pixel in the diffuse image $\mathbf{I}_d$ to get the image shown in Figure 13(c). This image is the one expected under white-light illumination. The accuracy of this color compensation is seen by comparing individual color patches on the McBeth chart in the images in Figure 13(a) and

Figure 13(c). The compensated image appears almost identical (except for the removed highlights) to the one taken under white-light illumination.

7.2 Determining Illumination Direction

By making assumptions related to the material properties of objects in the scene, it is also possible to get estimates of the direction of illumination. The vectors $\mathbf{I}_{\rm d}$, $\mathbf{I}_{\rm min}$, and $\mathbf{I}_{\rm max}$ computed at each pixel are used to determine the two specular components $\mathbf{I}_{\rm sc}$ and $\mathbf{I}_{\rm sv}$. These in turn are used to compute the Fresnel ratio q using (19). As described in Section 3, the Fresnel ratio, for any given surface point, depends on the refractive index η and the angle of incidence ψ. Figure 14(a) shows $q - \psi$ plots for dielectrics with different refractive indices.

As is evident from Figure 14(a), it is not possible to determine ψ from q without approximating the refractive index η. It turns out that a large variety of real-world dielectrics have η values between 1.6 and 1.8, including most plastics, ceramics, and other commonplace synthetic materials (see [10], [9]). Within this range of η values, the $q - \psi$ plots do not vary substantially and thus we assume $\eta = 1.7$. This approximation allows us to determine illumination direction from primary specular highlights. In the case of specular reflection, angle of incidence equals angle of reflection. Even in the case of rough surfaces, only those facets on the surface that are oriented so as to satisfy this specular constraint, produce specular reflections in the image [22]. Therefore, given ψ, the polar angle θ_s between the source and the sensor, subtended by the scene point, is simply $\theta_s = 2\psi$. The azimuth angle ϕ_s of the source is exactly equal to the phase angle α estimated by the polarization fitting. This approach to estimating source direction is applicable only to primary highlights; q estimates for specularities caused by interreflections are not meaningful since the incident light itself is partially polarized and to an unknown extent.

Figure 14(b) shows the scene we have used for this experiment. The separation algorithm is applied to the scene and the strong highlight on the midsection of the vase is detected and used to determine local q estimates within the highlight. An average q value is computed for the entire highlight region and used to estimate the incidence angle ψ from Figure 14(a). Note that each q estimate results in two possible ψ values, on either side of the Brewster angle [4] at which q tends to infinity. For the assumed η value of 1.7, the Brewster angle is approximately 60 degrees. To resolve this two-fold ambiguity in ψ, we assume the source lies within 120 degrees with respect to the sensor direction. This restricts ψ to the range $0 \leq \psi \leq 60$ for which the $q - \psi$ plot is monotonic, giving us a unique mapping from q to ψ. From Section 3, we also know that the phase angle given by the polarization fitting also has a two solutions, α and $\alpha + \pi$. This ambiguity is resolved by specifying a-priori if the source is to the left or the right of the sensor.

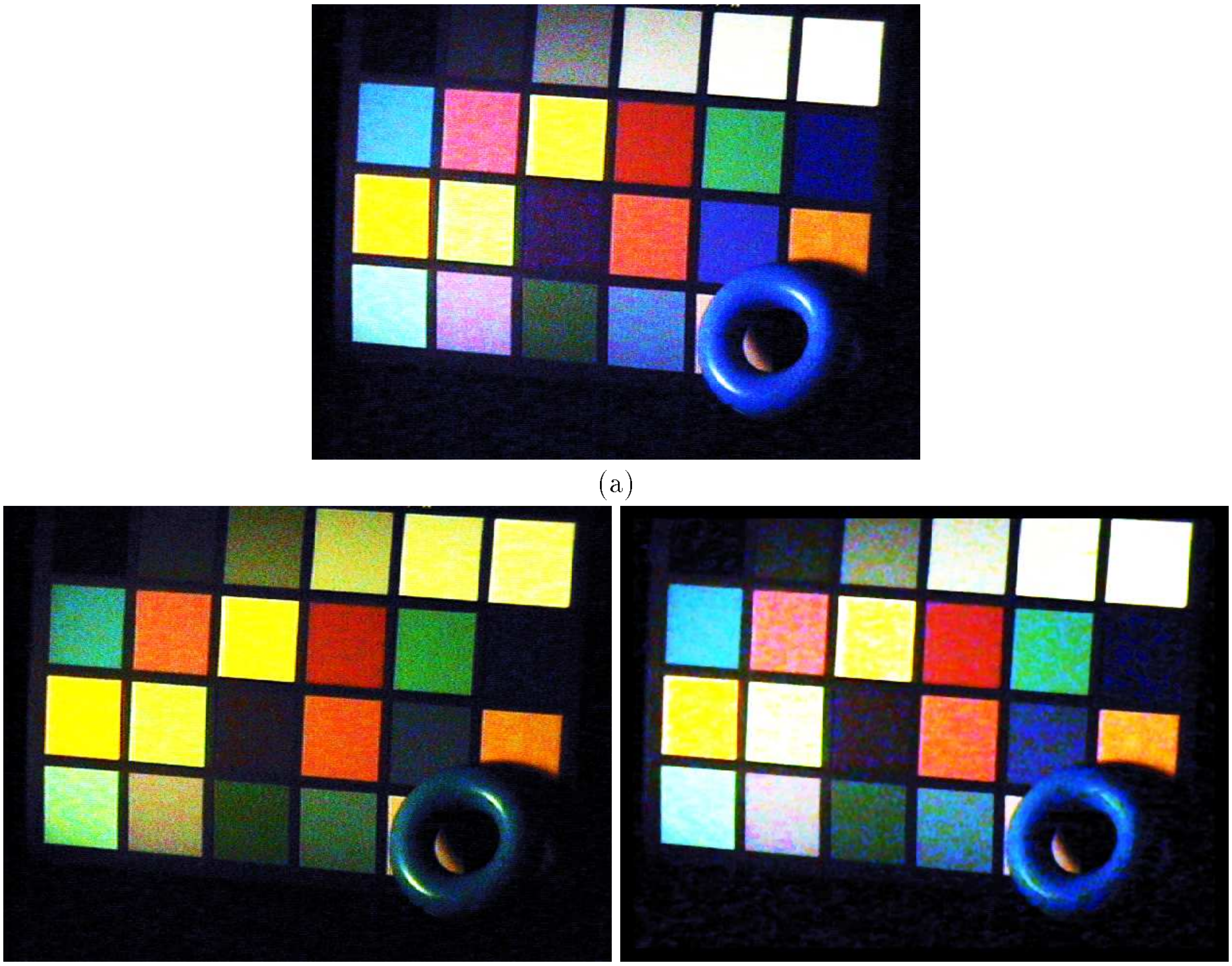
The light source was manually positioned at 4 different known locations on a plane passing through the optical axis of the sensor. The actual and computed source angles are denoted as (θ_s,ϕ_s) and (θ_c,ϕ_c), respectively, and are listed in Figure14(c). Note that the source direction estimates increase in accuracy with the angle of incidence. The reason for this is evident from the plots in Figure 5(b) and Figure 14(a). Even for dielectrics, for small angles of incidence (less than 20 degrees), the Fresnel ratio is close to unity causing the specular component to be insensitive to polarization. Therefore, q estimates computed for small incidence angles are susceptible to noise and are unreliable. The source direction estimation described here is therefore useful only for angles of incidence between 20 and 60 degrees, or polar source angles (θ_s) between 40 and 120 degrees.

7.3 Photometric Stereo with Unknown Sources

The above techniques can be used to recover shape information by shape from shading and photometric stereo without exact prior knowledge of source parameters. Here, we use photometric stereo [25] to demonstrate this idea. Photometric stereo uses multiple source directions and the known reflectance properties of surfaces in the scene to estimate local surface normals. The separation algorithm is used to overcome two problems. First, it is used to automatically estimate the source directions, avoiding the need for precise mechanical fixturing of light sources or the use of any source direction calibration. Of course, this is possible only if objects in the scene produce at least one specular highlight. Next, the separation algorithm is used to remove specular reflections from measured images to facilitate the computation of surface normals from just diffuse reflections. This is advantageous since accurate modeling of specular reflectance is cumbersome as it is closely related to surface roughness.

Figures 15(a) and (b) show two images of an object using different illumination directions. Given the translational symmetry of the object, only two sources are needed to estimate surface normals by photometric stereo. A sequence of polarization images are taken for each source and the separation algorithm is applied. The method described in Section 7.2 is used to estimate the source direction for each illumination from the corresponding highlight on the blue section of the object.[10] The source angles (θ_s, ϕ_s) were found to be (60.8, 61.8) and (64.2, 252.4), respectively. The diffuse color images for the two source directions, and the Lambertian reflectance assumption, were used to estimate local surface normals by photometric stereo. Note that, in principle, any of three color bands can be used to compute normals. To minimize the effect of image noise, the color band that produced the largest intensities at each point was used to estimate its normal. Figure 15(c) shows the computed needle map for the object. Note that the planar and curved sections of the object were successfully recovered. The normal vectors in the highlight regions are also found to be consistent with local surface shape. Some errors in the highlight regions can be attributed to inaccuracies in the separation process and the estimation of source direction using the approximate η value of 1.7.

[10] If the colors of the sources are also unknown, they can be approximated by the method described in Section 7.1 and used to color compensate the diffuse images before shape recovery.

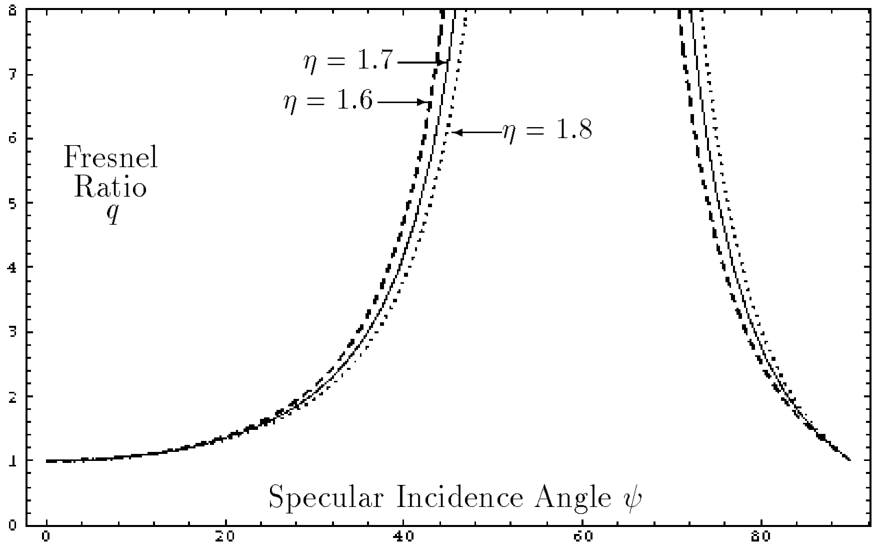


(a)

(b)

(c)

Figure 13: (a) Scene illuminated with white light. (b) The same scene illuminated with yellow light. (c) Color compensated image of the scene in (b). The accuracy of the color compensation is seen by comparing individual color-chart patches in (c) and (a).



(a)

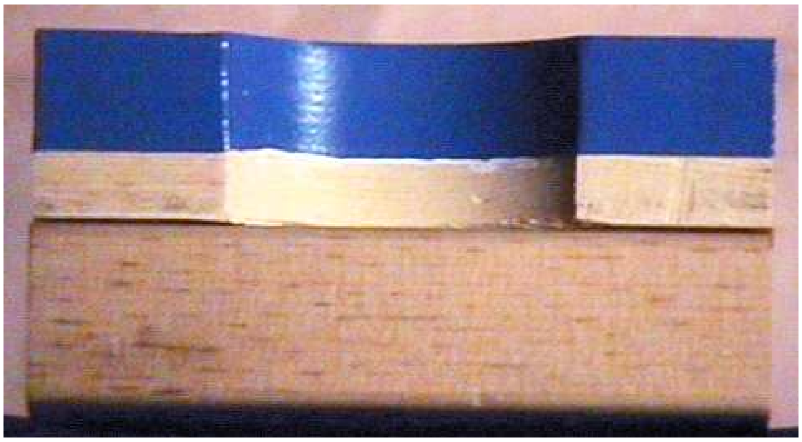
(b)

	ACTUAL DIR. (θ_s, ϕ_s)	COMPUTED DIR. (θ_c, ϕ_c)
SOURCE 1	(30.0, 75.0)	(68.3, 81.7)
SOURCE 2	(45.0, 75.0)	(52.5, 79.7)
SOURCE 3	(60.0, 75.0)	(65.3, 76.1)
SOURCE 4	(75.0, 75.0)	(72.2, 73.3)

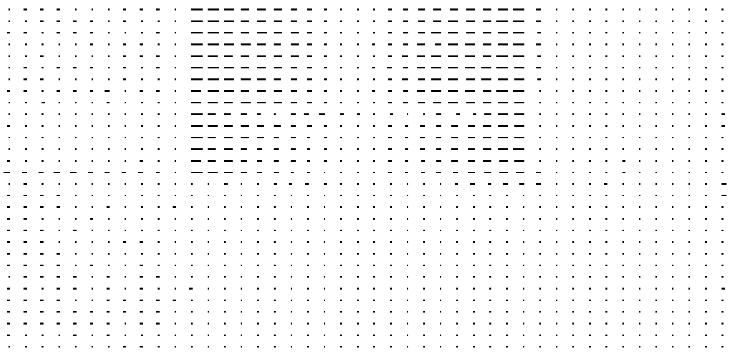
(c)

Figure 14: (a) Relation between Fresnel ratio q and incidence angle ψ for different refractive indices. (b) Source direction is determined from q values computed within image highlights. (c) Comparison between actual and estimated source directions (in degrees) for four different source locations.

(a)



(b)



(c)

Figure 15: Object illuminated by a source on (a) the left and (b) the right of the sensor. The source directions are automatically determined from the highlights. (c) Needle map of the object estimated by applying photometric stereo to the diffuse components of (a) and (b).

8 Discussion

We have presented a novel algorithm for separation of scene images into their reflection components. The algorithm uses the polarization of reflected light to obtain local and independent estimates of the color of the specular component. This constrains the diffuse component at each image point to lie in a linear color subspace. The algorithm exploits a variety of physical constraints that enable it to perform reliably even in the presence of complex textures, strong highlights and interreflections. Our experiments on separation demonstrate the ability of the algorithm to handle scenes of a level of complexity that is far beyond what previous algorithms can cope with. We successfully demonstrated three applications of the separation algorithm, namely, color compensation by determining illumination color, estimation of illuminant direction, and shape recovery using unknown illuminants. In the past, each of the above applications has been viewed as a pertinent research problem.

We now conclude with a critique of our results. It should be noted that the proposed algorithm does require several polarization images to be acquired, the fitting of polarization parameters at each image point, and the iterative application of the separation technique. Given this, it cannot at present be applied to real-time tasks without the use of a parallel machine. It is also worth taking a second look at the assumptions made during the development of the algorithm. The algorithm is not effective for metallic surfaces since polarization does not modulate the specular component for most conductors. Further, the dichromatic model assumed here is not in general applicable to metals as they seldom produce diffuse reflections with spectral distributions that differ from that of the illumination. This confines the present method to dielectrics. Even within the dielectric world, the colors of the diffuse and specular components are not guaranteed to differ; consider for example a red object illuminated by red light, or more likely, a gray object illuminated by white light.

Further problems arise if the illumination is close to the sensor direction. From Figure 5(b) we know that for dielectrics the specular component is not sensitive to the orientation of the polarization filter for angles of incidence less that 20 degrees. This translates to source directions that lie anywhere within 40 degrees with respect to the optical axis of the sensor. This problem is inherent to any polarization based method. Finally, all our experiments have been confined to indoor scenes with man-made objects. In the case of outdoor scenes we have to contend with surfaces that are very rough (sand, for example) and fairly complex textures (grass, for instance). Random textures, for example, make it difficult to find diffuse neighbors for most pixels in the image, a condition that is critical to the separation procedure. Further, in the case of outdoor scenes, the polarization of the incident light varies depending on sky conditions. The above factors limit the applicability of the separation algorithm. These limitations of course are no more severe that those faced by any of the previous separation algorithms. As evidenced by the experiments, the simultaneous use of color and polarization has clear advantages over using either only color or only polarization.

Acknowledgements

This research was supported by ARPA under contract DACA-76-92-C-007. The authors thank Ming-Chao Chiang of Columbia University for his help in optimizing the algorithm.

References

[1] B. G. Batchelor, D. A. Hill, and D. C. Hodgson, *Automated Visual Inspection,* IFS Publications Ltd., Bedford, England, 1985.

[2] R. Bajcsy, S. W. Lee, and A. Leonardis "Color image segmentation with detection of highlights and local illumination induced by interreflections," *Proc. of Intl. Conf. on Pattern Recognition (ICPR),* Atlantic City, NJ, June 1990.

[3] P. Beckmann and A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces,* Pergamon, 1963, New York.

[4] M. Born and E. Wolf, *Principles of Optics,* London:Pergamon, 1965.

[5] T. E. Boult and L. B. Wolff, "Physically based edge labeling," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR),* pp. 656-663, Maui, Hawaii, June 1991.

[6] G. Brelstaff and A. Blake, "Detecting Specular Reflections using Lambertian Constraints," *Proc. of Intl. Conf. on Computer Vision (ICCV),* pp. 297-302, Tampa, Florida, December 1992.

[7] *Polarization and Remote Sensing,* edited by W. G. Egan, SPIE, Vol. 1747, July, 1992.

[8] R. Gershon, *The Use of Color in Computational Vision,* PhD thesis, University of Toronto, 1987.

[9] *Handbook of Laser Science and Technology,* edited by M.J. Marvin, CRC Press Inc., Vol. 4, 1988.

[10] *Handbook of Optics,* edited by W.G. Driscoll, McGraw Hill Inc., 1978.

[11] G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," *International Journal of Computer Vision,* Vol. 2, No. 1, pp. 7-32, 1990.

[12] G. J. Klinker, *A Physical Approach to Color Image Understanding,* A K Peters, Wellesley, Massachusetts, 1993.

[13] K. Koshikawa, "A polarimetric approach to shape understanding," *Proc. of Intl. Joint Conf. on Artificial Intelligence (IJCAI),* pp. 493-495, 1979.

[14] S. Mersch, "Polarized Lighting for Machine Vision Applications," *Proc. of RI/SME Third Annual Applied Machine Vision Conf.*, pp. 40-54, Schaumburg, Feb. 1984.

[15] S. K. Nayar, K. Ikeuchi, T. Kanade, "Surface Reflection: Physical and Geometrical Perspectives," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 611-634, July 1991.

[16] C. Novak and S. Shafer, "Anatomy of a Color Histogram," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR),* pp. 599-605, 1992.

[17] M. Oren and S. K. Nayar, "Seeing Beyond Lambert's Law," *Proc. Third European Conference on Computer Vision,* Vol. 2, pp. 269-280, May, 1994.

[18] Y. Sato and K. Ikeuchi, "Temporal-Color Space Analysis of Reflection," *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR),* pp. 570-576, 1993.

[19] S. Shafer, "Using color to separate reflection components," *Color Research and Applications,* Vol. 10, pp. 210-218, 1985.

[20] S. Shafer, T. Kanade, G. J. Klinker, and C. L. Novak, "Physics-based models for early vision by machine," *Proc. SPIE Conf. on Perceiving, Measuring and Using Color,* Vol. 1250, Feb 1993.

[21] S.W. Lee, *Understanding of Surface Reflection in Computer Vision by Color and Multiple Views,* PhD thesis, University of Pennsylvania, 1991.

[22] K. Torrance and E. Sparrow, "Theory for Off-Specular Reflection from Roughened Surfaces," *Journal of the Optical Society of America*, No. 57, pp. 1105-1114, 1967.

[23] L. B. Wolff and T. E. Boult, "Constraining Object Features using a Polarization Reflectance Model," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 635-657, July 1991.

[24] L. B. Wolff, "A Diffuse Reflectance Model for Dielectric Surfaces," *Proc. SPIE Conf. on Optics, Illumination, and Image Sensing for Machine Vision,* Vol. 1822, pp. 60-73, Nov. 1992.

[25] R. J. Woodham, "Photometric Method for Determining Surface Orientation from Multiple Images," *Optical Engineering,* Vol. 19, No. 1, pp. 139-144, 1980.