Color Subspaces as Photometric Invariants

Todd Zickler zickler@eecs.harvard.edu Harvard University
Satya P. Mallick spmallick@vision.ucsd.edu
David J. Kriegman kriegman@cs.ucsd.edu
University of California, San Diego
Peter N. Belhumeur belhumeur@cs.columbia.edu Columbia University

Abstract

Complex reflectance phenomena such as specular reflections confound many vision problems since they produce image 'features' that do not correspond directly to intrinsic surface properties such as shape and spectral reflectance. A common approach to mitigate these effects is to explore functions of an image that are invariant to these photometric events. In this paper we describe two such invariants—one invariant to specular reflections, and the other invariant to both specular reflections and diffuse shading—that result from exploiting color information in images of dichromatic surfaces. These invariants are derived from subspaces of RGB color space, and they enable the application of Lambertian-based vision techniques to a broad class of specular, non-Lambertian scenes. Using implementations of recent algorithms taken from the literature, we demonstrate the practical utility of these invariants for a wide variety of applications, including stereo, shape from shading, material-based segmentation, and motion estimation.

1. Introduction

An image is the product of the shape, reflectance and illumination in a scene. For many visual tasks, we require only a subset of this information, and we wish to extract it in a manner that is insensitive to variations in the remaining 'confounding' scene properties. For 3D reconstruction, for example, we seek accurate estimates of shape, and we design systems that are insensitive to variations in reflectance and illumination.

One practical approach to these problems is to compute a function of the input images that is invariant to confounding scene properties but is discriminative with respect to desired scene information. Such functions yield so-called *invariants*, and a number of examples are described in the literature. Perhaps the simplest example for a Lambertian scene is a normalized-RGB image. The normalized RGB color vector at each pixel depends on the spectral reflectance of the corresponding surface patch but not its orientation, which makes it useful for material-based segmentation.

Like normalized-RGB, many existing invariants seek to isolate information about the material properties in a scene and are therefore designed to be invariant to local illumination and viewing geometry. In contrast, we consider a class of invariants that deliberately preserve geometry information in a way that is invariant to specular reflections. These invariants give direct access to surface shape in the form of diffuse shading effects, and since diffuse shading is often well approximated by the Lambertian model, they satisfy the 'constant-brightness assumption' underlying most approaches to stereo reconstruction and structure-from-motion. In addition, these invariants provide access to surface normal information, which can be recovered using Lambertian-based photometric reconstruction methods.

Figure 1. Left: Two frames of an RGB video of a scene with mixed illumination. A blue light on the right and a yellow light on the left induce complex specular effects. Right: Projecting these images onto the one-dimensional subspace orthogonal to the source color vectors in RGB space, yields an invariant to specular reflections that preserves diffuse shading and spectral reflectance information. (The complete video is included as supplemental material.)

This work is motivated in part by the work of Mallick *et al*. [9], who propose a transformation of color space for use in photometric stereo. Their work shows that when surfaces are well-described by the dichromatic model [14], a specular-free image can be computed by projecting the RGB color vector at each image point onto the two-dimensional subspace orthogonal to the illuminant color.

Inspired by these results, we:

1. Derive a general class of specular invariants based on color subspaces. These invariants can be applied to dichromatic surfaces under *mixed* illumination envi-

ronments (see Fig. 1.)

2. Derive a second class of functions that, in addition to specular reflections, are invariant to diffuse shading. They depend only on the spectral reflectance of a dichromatic surface.
3. Demonstrate that these invariants can enhance many existing Lambertian-based vision techniques, vastly expanding their domain of applicability.

2. Related Work

A number of photometric invariants have been proposed for Lambertian scenes without specularity. Normalized-RGB, r-g chromaticity, and hue/saturation images are all examples of representations that are independent of 'diffuse shading' (the geometric relation between a surface normal and the illumination direction) and depend only on the spectral reflectance of the surface and the spectral power distribution (SPD) of the illuminant. Additional invariants to either local geometry or spectral reflectance can be computed when multiple images of a scene are available (*e.g.*, [18]), or when the reflectance of the surface is spatially coherent (*e.g.* [11]); and an invariant to both local geometry and illuminant SPD can be computed from a single image under appropriate imaging conditions [4].

Invariants for scenes with more general reflectance functions are developed by Narasimhan *et al.* [10]. They describe a general model of reflectance consisting of a product of a "material" term (Lambertian albedo, Fresnel coefficient, etc.) and a "geometry" term that encodes the relationship between the surface normal, light-source, and viewing direction. Invariants to both of these terms can be computed from either multiple observations of a single point under variable view or illumination, or from one observation of a spatially-coherent scene. The geometry invariant is of particular interest, since it can be used directly for material-based segmentation [10].

2.1. Invariants for Dichromatic Surfaces

A substantial body of work is devoted to exploiting the dichromatic model of reflection [14] in order to separate diffuse and specular reflection components, which are independent of specular and diffuse reflection effects, respectively. According to the dichromatic model, the observation of a surface point is written

$$I_k = \sigma_d D_k + \sigma_s S_k, \tag{1}$$

where σ_d and σ_s are geometric scale factors that depend on surface shape and material properties and

$$D_k = \int E(\lambda) R(\lambda) C_k(\lambda) d\lambda \tag{2}$$

$$S_k = \int E(\lambda) C_k(\lambda) d\lambda. \tag{3}$$

Here, $E(\lambda)$ is the SPD of the incident illumination, $R(\lambda)$ is the spectral reflectance of the surface, and $C_k(\lambda)$ is the spectral sensitivity of a linear sensor. A typical RGB camera yields three such observations, and in this case we write $\mathbf{I}_{RGB} = \{I_k\}_{k=R,G,B}$ and define $\mathbf{D} = \{D_k\}_{k=R,G,B}$ and $\mathbf{S} = \{S_k\}_{k=R,G,B}$ to be the *diffuse color* and *specular color*, respectively. These are conventionally assumed to be vectors of unit length.

There is practical utility in separating the diffuse and specular components in an image. Since diffuse reflections are typically well-represented by the Lambertian model, this separation allows the application of powerful Lambertian-based vision algorithms to a broad class of non-Lambertian scenes. Unfortunately, computing such a separation is ill-posed. It traditionally requires additional constraints such as texture-less surfaces [6], knowledge about (e.g., segmentation of) diffuse colors, constraints on the neighborhood of a pixel [8, 15] or specific parametric models for specular reflectance [13].

When the source color is known and constant over a scene, one can compute invariants to specular reflections that are based on transformations of RGB color space and do not require explicit specular/diffuse separation. Tan and Ikeuchi [15] obtain such a specular invariant using a non-linear combination of the RGB values $\mathbf{I}_{RGB}$ at a pixel and the RGB color vector of the source $\mathbf{S}$. The transformation is computed independently at each point, and it yields a positive grayscale image that depends only on diffuse reflections (σ_d and $\mathbf{D}$) and is independent of specular effects (σ_s). Park [12] defines an alternative, linear transformation providing two color channels that, while not pure invariants, are highly insensitive to specular reflections. In this transformation, one of the coordinate axis of color space is aligned with the source color $\mathbf{S}$, leaving the remaining two channels to be predominantly diffuse.

A third transformation is proposed by Mallick *et al.* [9], who define the *SUV color space* using $\mathbf{I}_{SUV} = \mathbf{R}\mathbf{I}_{RGB}$, where $\mathbf{R}$ is any rotation of $\mathbb{R}^3$ satisfying $\mathbf{RS} = (1,\ 0,\ 0)$. Like Park's transformation, this rotation aligns one of the color axes (the S-axis) with the source color $\mathbf{S}$, but unlike Park's transformation, the components along the remaining two axes are indeed invariant to specular reflections. This is easily seen by applying the transformation to Eq. 1 and verifying that the U and V components are given by [9]

$$I_U = \sigma_d \mathbf{r}_2^\top \mathbf{D}, \quad I_V = \sigma_d \mathbf{r}_3^\top \mathbf{D}, \tag{4}$$

where $\mathbf{r}_2^\top$ and $\mathbf{r}_3^\top$ denote the 2^{nd} and 3^{rd} rows of the rotation matrix $\mathbf{R}$. Since they are independent of the specular geometric scale factor σ_s, these components constitute a specular invariant. An important feature of this transformation is that it preserves and isolates the diffuse shading information (σ_d).

3. Color Subspaces

The SUV color transformation can be viewed as a projection of RGB color vectors onto the two-dimensional subspace orthogonal to the source color $\mathbf{S}$. (See left of Fig. 2.) This interpretation provides the main motivation for this paper, and in this section we show that it: 1) can be generalized to mixed illuminants and hyper-spectral images; and 2) leads naturally to a notion of generalized hue.

The SUV color transformation is based on Eq. 1, which in turn is premised on the assumption that the illuminant SPD is constant over the incident hemisphere of a surface point (*i.e*., that the illuminant 'color' is the same in all directions.) Notationally, if $L(\boldsymbol{\omega}_i, \lambda)$ represents the incident radiance at a surface point, where $\boldsymbol{\omega}_i = (\theta_i, \phi_i) \in \Omega$ parameterizes the hemisphere of incident directions, the model requires that this input radiance field can be factored (with a slight abuse of notation) as $L(\boldsymbol{\omega})E(\lambda)$. To relate this to the terms in Eq. 1, we let $f(\boldsymbol{\theta}, \lambda)$ with $\boldsymbol{\theta} = (\theta_i, \phi_i, \theta_o, \phi_o, \lambda)$ denote the BRDF of the surface, and we write the image formation equation as

$$I_k = \int_\lambda \int_\Omega f(\boldsymbol{\theta}, \lambda) L(\boldsymbol{\omega}_i, \lambda) C_k(\lambda) \cos\theta_i \; d\boldsymbol{\omega}_i d\lambda.$$

According to the dichromatic model, the BRDF of the surface can be decomposed according to

$$f(\boldsymbol{\theta}, \lambda) = f_d(\boldsymbol{\theta}) S(\lambda) + k_s f_s(\boldsymbol{\theta}), \tag{5}$$

where k_s is a constant, and this yields the expressions

$$\begin{aligned} \sigma_d &= \int_\Omega f_d(\boldsymbol{\theta}) L(\boldsymbol{\omega}_i) \cos\theta_i \; d\boldsymbol{\omega}_i \\ \sigma_s &= k_s \int_\Omega f_s(\boldsymbol{\theta}) L(\boldsymbol{\omega}_i) \cos\theta_i \; d\boldsymbol{\omega}_i \\ D_k &= \int S(\lambda) E(\lambda) C_k(\lambda) \; d\lambda \\ S_k &= \int E(\lambda) C_k(\lambda) \; d\lambda. \end{aligned}$$

To generalize the model, we consider a mixed illumination environment whose spectral content can be written in terms of a finite linear basis:

$$L(\boldsymbol{\omega}_i, \lambda) = \sum_{j=1}^{N} L_j(\boldsymbol{\omega}_i) E_j(\lambda). \tag{6}$$

An example with $N = 2$ is an office environment where the illumination can be described as a solid angle of daylight in the direction of the window and a distinct solid angle of fluorescent light in the direction of the ceiling. When the input radiance field can be decomposed in this manner, the BRDF decomposition of Eq. 5 yields

$$I_k = \sum_{j=1}^{N} \sigma_d^{(j)} D_k^{(j)} + \sigma_s^{(j)} S_k^{(j)}, \tag{7}$$

with

$$\begin{aligned} \sigma_d^{(j)} &= \int_\Omega f_d(\boldsymbol{\theta}) L_j(\boldsymbol{\omega}_i) \cos\theta_i \; d\boldsymbol{\omega}_i \\ \sigma_s^{(j)} &= k_s \int_\Omega f_s(\boldsymbol{\theta}) L_j(\boldsymbol{\omega}_i) \cos\theta_i \; d\boldsymbol{\omega}_i \\ D_k^{(j)} &= \int S(\lambda) E_j(\lambda) C_k(\lambda) \; d\lambda \\ S_k^{(j)} &= \int E_j(\lambda) C_k(\lambda) \; d\lambda. \end{aligned}$$

Equation 7 suggests the existence of a specular invariant that is analogous to the two-dimensional subspace used by Mallick *et al*. [9]. In their formulation, the illuminant color is assumed constant over the input hemisphere (which corresponds to $N = 1$ in Eq. 7) and the specular invariant subspace computed from a three-channel RGB image is two-dimensional. In general, given an M-channel (possibly hyper-spectral) image and an N-dimensional spectral basis $\{E_j(\lambda)\}_{j=1\ldots N}$ for the incident illumination, there exists a subspace of dimension $(M - N)$ that is independent of all $\sigma_s^{(j)}$ and therefore invariant to specular reflections. Letting $\{\mathbf{r}_l\}_{l=1\ldots(M-N)}$ represent an orthonormal basis for this specular invariant subspace, the l^{th} component (or 'channel') of the specular invariant image is given by

$$J_l = \sum_{j=1}^{N} \sigma_d^{(j)} \mathbf{r}_l^\top \mathbf{D}^{(j)}. \tag{8}$$

A specular invariant image with $(M - N)$ channels defined by this equation can be treated as an image with $(M - N)$ 'colors', but these 'colors' can assume negative values. In some cases it is more convenient to use a grayscale specular invariant given by

$$J_{\text{inv(M−N)}} = \left(\sum_{l=1}^{M-N} I_l^2 \right)^{\frac{1}{2}}, \tag{9}$$

where the subscript $J_{\text{inv(u)}}$ is used to indicate that the grayscale invariant is derived from a u-dimensional invariant subspace.

Since the vast majority of cameras record three (RGB) channels, the most interesting case to consider is $N = 2$. An example is shown in Fig. 1, where light comes from two sources with different SPDs. These SPDs induce two source color vectors $\mathbf{S}^{(1)}$ and $\mathbf{S}^{(2)}$ in RGB space, and by projecting the RGB color vectors of the input image onto the one-dimensional subspace orthogonal to these vectors, we create an image that is void of specular reflection effects.

An essential feature of the specular invariants of Eqs. 8 and 9 is that they *preserve diffuse reflection effects* encoded in the geometric scale factors $\sigma_d^{(j)}$. For many surfaces,

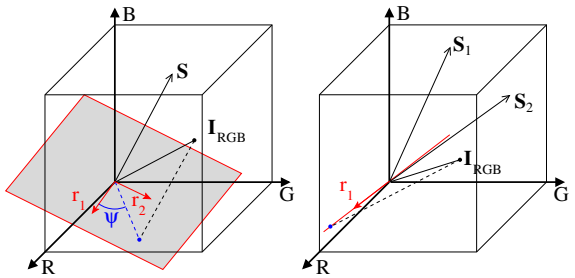

Figure 2. Left: Projecting RGB color vector $\mathbf{I}_{RGB}$ onto the 2D subspace orthogonal to the source color $\mathbf{S}$ results in a specular invariant that preserves diffuse shading. The ratio between the channels in this subspace represents *generalized hue* (ψ), which provides a second invariant depending only on spectral reflectance. Right: For two source colors, the specular invariant subspace is one-dimensional. By projecting RGB color vectors onto this line, a specular invariant can still be computed (see Fig. 1.)

the diffuse component is well approximated by the Lambertian model, meaning that the term $f_d(\boldsymbol{\theta})$ in Eq. 7 is a constant function of $\boldsymbol{\theta}$ and the geometric scale factors $\sigma_d^{(j)}$ do not change with viewpoint. This implies that the specular invariant images defined by Eqs. 8 and 9 often: 1) satisfy the 'constant-brightness assumption' underlying most stereo and structure-from-motion systems; and 2) provide access to surface normal information through Lambertian-based photometric reconstruction methods such as shape-from-shading. As a result, by computing these invariants as a pre-processing step, we can expand the domain of applicability of many Lambertian-based algorithms to include a much broader class of specular, non-Lambertian surfaces.

Applications are explored in Sect. 4. Next, we define a second class of invariants that can be computed from the color subspaces defined above.

3.1. Generalized Hue

Returning to the case of uniform source color ($N = 1$ in Eq. 6), we derive a second invariant by taking the ratio between specular invariant channels in Eq. 8. The result,

$$J_1/J_2 = \mathbf{r}_1^\top \mathbf{D}/\mathbf{r}_2^\top \mathbf{D},$$

is independent of both the diffuse and specular geometric scale factors σ_d and σ_d. As shown in Fig. 2, it is instructive to interpret this ratio as an angle and define

$$\psi = \tan^{-1}\left(J_1/J_2\right) = \tan^{-1}\left(\mathbf{r}_1^\top \mathbf{D}/\mathbf{r}_2^\top \mathbf{D}\right). \qquad (10)$$

We refer to ψ as the *generalized hue*, since it reduces to the standard definition of hue when the source color $\mathbf{S}$ is white.

Examples of generalized hue images are shown in Fig. 3 for a specular globe under two different source colors. In each case, the known source color is used to compute a two-channel subspace image according to Eq. 8, and the ratio between the two channels is used to compute ψ. Since it

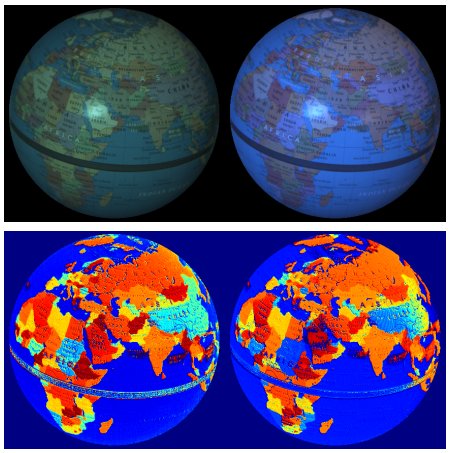

Figure 3. Generalized hue images (bottom), each computed from a single RGB image (top) of a globe. Generalized hue is invariant to both specularities and diffuse shading, and is discriminative only with respect to the spectral reflectance of the surface.

depends only on $\mathbf{D}$, each country on the globe appears 'flat' and is free of both specular reflections and diffuse shading.

It is interesting to note that this isolation of spectral reflectance can be generalized to mixed illumination and hyper-spectral images, albeit at the expense of requiring either multiple images, spatial coherence, or both. As an illustration, consider an illumination environment with two different source colors ($N = 2$ in Eq. 6), and suppose we acquire two RGB images with altered illumination directions. (The SPDs of the sources remain the same.) In this case, the specular invariant subspace is one-dimensional, and the specular invariant images J and $\bar{J}$ of a surface point under the two lighting configurations are given by (see Eq. 8)

$$\begin{aligned} J &= \sigma_d^{(1)} \mathbf{r}^\top \mathbf{D}^{(1)} + \sigma_d^{(2)} \mathbf{r}^\top \mathbf{D}^{(2)} \\ \bar{J} &= \bar{\sigma}_d^{(1)} \mathbf{r}^\top \mathbf{D}^{(1)} + \bar{\sigma}_d^{(2)} \mathbf{r}^\top \mathbf{D}^{(2)}. \end{aligned}$$

Here, we have suppressed the subscript corresponding to the invariant channel index, since there is only a single channel.

Now, suppose we have identified two additional image points (p_2 and p_3) that correspond to surface points having the same geometric configurations (*e.g.*, the same surface normal under distant lighting) but distinct spectral reflectance. This yields a total of six specular invariant observations—three points under two illuminations—that can be assembled into a 3×2 observation matrix. As shown by Narasimhan *et al.* [10], such a matrix can be factored as

$$\begin{bmatrix} J_{p_1} & \bar{J}_{p_1} \\ J_{p_2} & \bar{J}_{p_2} \\ J_{p_3} & \bar{J}_{p_3} \end{bmatrix} = \begin{bmatrix} \mathbf{r}^\top \mathbf{D}_{p_1}^{(1)} & \mathbf{r}^\top \mathbf{D}_{p_1}^{(2)} \\ \mathbf{r}^\top \mathbf{D}_{p_2}^{(1)} & \mathbf{r}^\top \mathbf{D}_{p_2}^{(2)} \\ \mathbf{r}^\top \mathbf{D}_{p_3}^{(1)} & \mathbf{r}^\top \mathbf{D}_{p_3}^{(2)} \end{bmatrix} \begin{bmatrix} \sigma_d^{(1)} & \bar{\sigma}_d^{(1)} \\ \sigma_d^{(2)} & \bar{\sigma}_d^{(2)} \end{bmatrix},$$

from which it follows that the ratio of determinants of its two 2×2 sub-matrices satisfies

$$\frac{J_{p_1}\bar{J}_{p_2} - \bar{J}_{p_1}J_{p_2}}{J_{p_2}\bar{J}_{p_3} - \bar{J}_{p_2}J_{p_3}} = \frac{\mathbf{D}_{p_1}^{(1)\top}\mathbf{D}_{p_2}^{(2)} - \mathbf{D}_{p_1}^{(2)\top}\mathbf{D}_{p_2}^{(1)}}{\mathbf{D}_{p_2}^{(1)\top}\mathbf{D}_{p_3}^{(2)} - \mathbf{D}_{p_3}^{(2)\top}\mathbf{D}_{p_2}^{(1)}}, \qquad (11)$$

which depends only on the spectral reflectance of the surface points and is invariant to the geometric scale factors.

While this example relies on the identification of three surface points (p_1, p_2, p_3), a similar invariant can be computed using a single surface point and multiple specular invariant channels collected from a hyper-spectral image. Determinant-based invariants of this form have been well studied by Narasimhan *et al*. [10], who apply them directly to RGB images and obtain invariants for relatively general reflectance functions under monochromatic environments. The analysis presented here can be viewed simultaneously as an extension of their approach to mixed-illumination environments as well as a specialization to the dichromatic model. The latter is an important difference, because it enables the distinction between specular and diffuse material properties in the resulting invariants.

3.2. Practical Limitations

The utility of the proposed invariants relies on the angular separation between the diffuse and source colors ($\mathbf{D}^{(j)}$ and $\mathbf{S}^{(j)}$) in color space. When this separation is small, the signal-to-noise ratio (SNR) in the subspace image can be prohibitively low. This is evident, for example, in the generalized hue image of the globe in the bottom-right of Fig. 3, where the hue variation within the People's Republic of China is seen to be large. In practice, this can be improved using high dynamic range images. Additionally, surface points with low SNR can be detected by monitoring the angle between the source colors $\mathbf{S}^{(j)}$ and the input color vectors $\mathbf{I}_{RGB}$, and this information can be incorporated into any robust vision algorithm (see, *e.g*., [17]).

It is also important to note that in order to compute the invariants described in the previous sections, we require knowledge of the source colors $\mathbf{S}^{(j)}$. In a controlled setting, these colors can be measured by imaging a 'white' surface under the given illuminants; and estimates of the illuminant colors in an uncontrolled settings can be obtained using existing methods (*e.g*., [1, 16]). From a practical standpoint, it is difficult to provide a meaningful quantitative evaluation of the sensitivity of these invariants to errors in the source colors since it depends on the spectral reflectance and illuminant SPD of a particular scene. We leave a statistical evaluation for future work, and instead, in the next section we assess the practical utility of these invariants by evaluating their performance in a very broad range of applications.

4. Applications and Evaluation

This section demonstrates the utility of the proposed invariants using RGB images for a number of vision algorithms and compares the results to those obtained using standard grayscale images $I_g = (R + G + B)/3$. For RGB images, when the illumination is a mixture of two known colors, the two-color specular invariant J_1 from Eq. 8 is grayscale and is equal to $J_{\mathrm{inv}(1)}$ from Eq. 9. On the other hand, a single-color specular invariant computed from an RGB image includes two diffuse channels $\{J_1, J_2\}$, which can be combined into a grayscale invariant $J_{\mathrm{inv}(2)}$ using Eq. 9. (This is equivalent to the representation of Mallick *et al*. [9].) The results in this section show that $J_{\mathrm{inv}(2)}$ is effective in situations where the source color is uniform, and that it performs much better than I_g. In situations where the scene illumination is a mixture of two colors, however, $J_{\mathrm{inv}(2)}$ is not invariant to all specular reflections, and significantly better results can be obtained using the invariant $J_{\mathrm{inv}(1)}$, which is derived from the one-dimensional subspace orthogonal to both source colors.

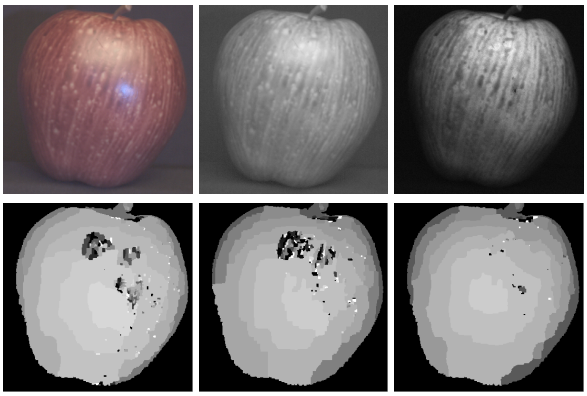

Figure 4. Stereo reconstruction under mixed illumination. Top left: One image of an input stereo pair with blue and yellow illumination. Top center: Single-color invariant image $J_{\mathrm{inv}(2)}$ from Eqs. 8 and 9 with $\mathbf{S}$ in the direction of the blue source. Top right: Two-color invariant $J_{\mathrm{inv}(1)}$ obtained by projecting to the 1D subspace orthogonal to both sources. Bottom row: depth map obtained using the stereo algorithm of Boykov *et al*. [3] in each case.

This section also includes an application of generalized hue to the problem of material-based segmentation.

4.1. Stereo

The vast majority of stereo algorithms are based (either explicitly or implicitly) on the assumption that surfaces are Lambertian. Since specular reflections violate this assumption, stereo reconstructions of specular surfaces are often inaccurate. In cases of significant specular reflections and complex illumination conditions, we can improve the accuracy of existing stereo algorithms by computing the specular invariant as a pre-process. This is demonstrated in Fig. 4, which compares binocular stereo results obtained using conventional grayscale images I_g, the single-illuminant (2D subspace) invariant $J_{\mathrm{inv}(2)}$, and the two-color (1D subspace) invariant $J_{\mathrm{inv}(1)}$. In this figure, the grayscale and invariant images are computed from a rectified RGB stereo pair (top of Fig. 4) and are used as input to the binocular

stereo algorithm of Boykov *et al*. [3]. The original RGB image includes two specular highlights caused by blue and yellow illuminants. The blue highlight is largely eliminated in the single-color invariant $J_{\mathrm{inv}(2)}$, while image $J_{\mathrm{inv}(1)}$ is invariant to specular reflections of both colors. As expected, the results from the grayscale and single-color invariant images are poor in specular regions, and the depth map obtained using $J_{\mathrm{inv}(1)}$ is significantly improved.

4.2. Optical Flow

Motion estimation through the computation of optical flow is another example of an application that can benefit from specular invariance. Recovering dense optical flow relies on the 'constant-brightness assumption', which is violated when an observer moves relative to a static, specular scene. As demonstrated by the results in Fig. 5, optical flow in the presence of specular reflections in a complex illumination environment can be improved by computing a specular invariant as a pre-processing step.

In Fig. 5, an RGB image sequence is captured by a camera translating horizontally relative to a static scene. The sequence is used to compute a conventional grayscale sequence $I_g(t)$, a single-color invariant sequence $J_{\mathrm{inv}(2)}(t)$, and a two-color invariant sequence $J_{\mathrm{inv}(1)}(t)$. These three videos are used as input to Black and Anandan's algorithm for robust optical flow [2]. The left of Fig. 5 shows a single image from each sequence, and the right shows the recovered flows in the indicated window. Since the camera undergoes pure translation, the 'ground-truth' flow lies along parallel horizontal lines. The flow recovered using the conventional grayscale and single-color invariant sequences are shown in green and blue, respectively; and as expected, these flows are severely corrupted by specular highlights. In contrast, the flow computed from the mixed-illuminant invariant (shown in red) is close to the ground truth and is largely unaffected by these non-Lambertian effects.

4.3. Shape from Shading

The previous two sections demonstrate the utility of the specular invariant for stereo matching and optical flow, both of which benefit from the fact that the specular invariant does not change with viewpoint. Here we show that since it also preserves diffuse (ideally Lambertian) shading information, these specular invariants can also be used to enhance photometric reconstruction methods.

As an example, we consider the special case of a single point light-source in direction $\hat{\mathbf{l}}$, so the specular invariant image of Eq. 9 reduces to

$$J_{\mathrm{inv}(2)} = f_d \left((\mathbf{r}_1^\top \mathbf{D})^2 + (\mathbf{r}_1^\top \mathbf{D})^2 \right)^{\frac{1}{2}} \hat{\mathbf{n}} \cdot \hat{\mathbf{l}}.$$

It is the image formation equation for a Lambertian surface with an effective albedo given by the first two terms, and it suggests that the specular invariant can be used directly as input to Lambertian-based shape from shading algorithms.

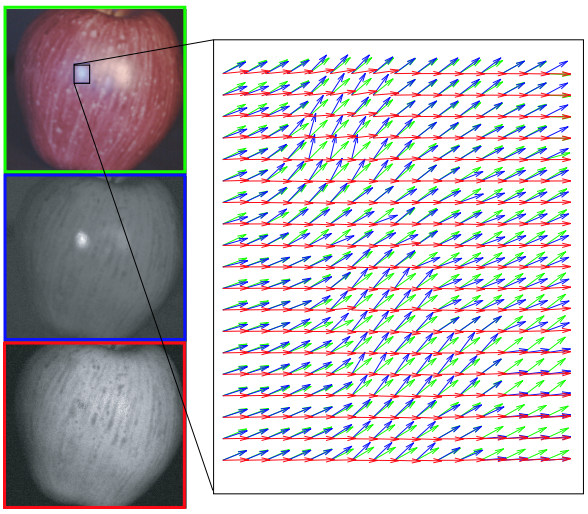

Figure 5. Optical flow comparison. An RGB image sequence (top left) is captured by a camera translating left relative to a specular apple under yellow and blue illumination. Derived conventional grayscale $I_g(t)$, yellow-invariant $J_{\mathrm{inv}(2)}(t)$ (left middle), and two-color invariant $J_{\mathrm{inv}(1)}(t)$ (left bottom) sequences are computed and used as input to Black and Anandan's robust optical flow algorithm [2]. Right: flows obtained in the three cases. Green and blue flows are from grayscale and yellow-invariant sequences, respectively, and both are corrupted by specular reflections. Red flow is computed from the two-color invariant and is much closer to ground truth, which is horizontal and to the right.

The benefit of this approach is demonstrated in Fig. 6, where we assess the performance of Zheng and Chellappa's shape from shading algorithm [20] for both a conventional grayscale image I_g and a single-color invariant image $J_{\mathrm{inv}(2)}$. The top of the figure shows grayscale and specular invariant images computed from an RGB image of a pear, and the middle row shows the surfaces that are recovered by applying the same algorithm in the two cases. The solid blue profile in the bottom graph shows that specular reflections cause severe artifacts when the algorithm is applied to the grayscale image. In contrast, as shown by the dashed red profile, one can obtain vastly improved results using the same algorithm by computing the specular invariant as a pre-processing step.

4.4. Photometric/Geometric Reconstruction

More generally, the specular invariant can be used to improve the performance of a broad class of Lambertian-based reconstruction systems in the presence of specular, non-Lambertian surfaces. This includes, for example, methods that combine both geometric and photometric constraints to obtain accurate surface shape [5, 7, 19]. To provide an example, we use the passive photometric stereo algorithm described by Lim *et al*. [7]. This method begins with an

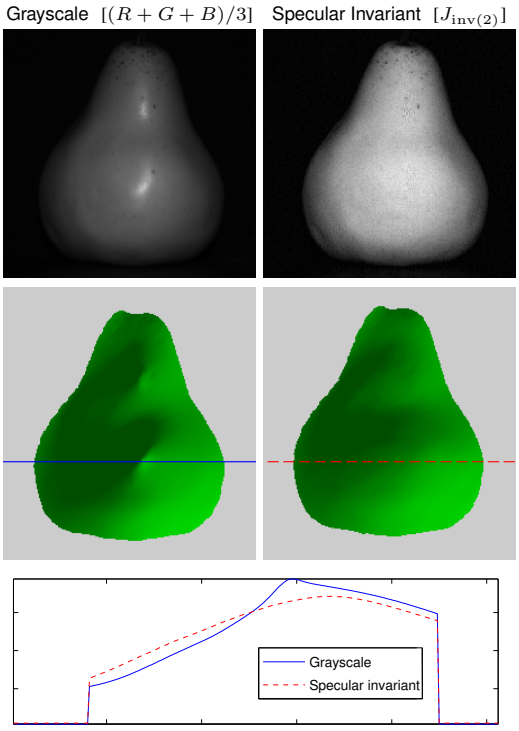

Figure 6. Shape from shading comparison. An RGB image of a pear is used to compute conventional grayscale (top-left) and specular invariant (top-right) images, and these are input to Zheng and Chellappa's shape from shading algorithm [20]. Middle row: surfaces recovered in both cases. Bottom row: cross-sections of the recovered surfaces along the indicated horizontal lines.

approximate, piece-wise planar reconstruction obtained by tracking a small number of features across a video sequence under (possibly varying) directional illumination. Then, an iterative method based on uncalibrated Lambertian photometric stereo simultaneously refines the reconstruction and estimates the unknown illumination directions.

Figure 7 compares the results obtained from an image sequence that consists of a moderately specular cylinder moving under fixed illumination and viewpoint. The shape is estimated by applying the same algorithm to both the conventional grayscale sequence ($I_g(t)$) and the specular invariant sequence ($J_{\mathrm{inv}(2)}$) computed from the same RGB data. The right-most surface in Fig. 7 shows that the reconstruction obtained using the specular invariant is nearly cylindrical, while that computed from the conventional grayscale sequence is severely corrupted by specular reflections.

4.5. Material-based Segmentation

Sections 4.1–4.4 demonstrate the utility of the first class of specular invariants for a variety of visual tasks. In this section, we demonstrate an applications of the second class of invariants, which is independent of both the specular reflections and diffuse shading in an image. Potential applications of this invariant include tracking and recognition. Here, we

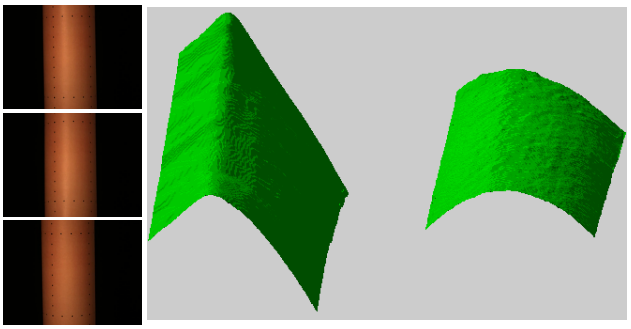

Figure 7. Comparison of shape from combined photometric and geometric constraints. Left: three RGB frames of a specular cylinder moving under fixed view and illumination. Right frame: result of simultaneous tracking and photometric reconstruction (as described by Lim *et al.* [7]) using both the conventional grayscale (left) and specular invariant (right) sequences.

consider the application of generalized hue to the problem of material-based segmentation.

Figure 8 shows an RGB image of a dichromatic scene under uniform source color ($N = 1$) as well as a number of pseudo-colored representations related to the invariants presented in Sect. 3. The top row shows conventional grayscale and specular invariant images, and in the latter, the specular effects (most notably on the green apples, the pumpkin, and the red pepper) are largely eliminated. The bottom-right of Fig. 8 shows the generalized hue image given by Eq. 10, which is invariant to diffuse shading in addition to specular reflections, and therefore depends only on the spectral reflectance. The fact that the generalized hue within each region is relatively constant suggests that it is a useful representation for segmentation. The same is not true for the conventional hue image (shown on the bottom-left) because the illuminant is not white.

5. Conclusion

This paper presents two classes of photometric invariants that are derived from color subspaces. They are efficiently computed from a single image of a dichromatic scene and are valid in cases of mixed (*i.e.*, spectrally-varying) illumination environments. The invariants are computed pointwise and therefore place no restriction on scene texture.

Computation of these invariants requires that the source color(s) be known *a priori*, but in the future, we plan to investigate methods that exploit these representations to recover this information from the data. For scenes such as that in Fig. 8, for example, it is possible that the entropy of the generalized hue image provides an indicator of the accuracy of the estimated source color.

The practical utility of these invariants is demonstrated by their ability to improve the performance of a wide variety of vision algorithms, including those for stereo and motion estimation. As a result, they provide a means for extending the applicability of existing Lambertian-based algorithms to

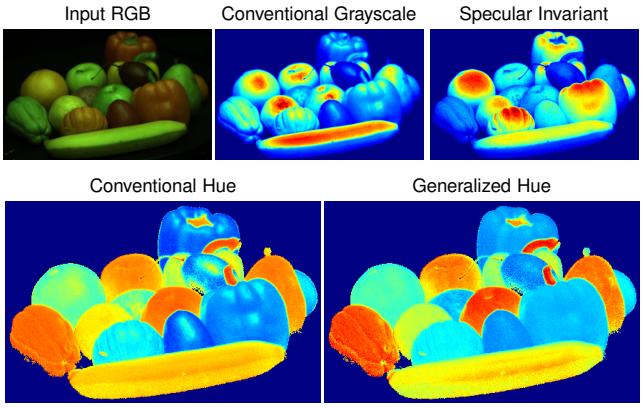

Figure 8. Generalized hue for material-based segmentation. Each panel shows a pseudo-colored representation that is computed from the RGB image on the top-left. The generalized hue image on the bottom-right is useful for segmentation because it depends only on the spectral reflectance of the surfaces. The same is not true for a conventional hue image (bottom-left) unless the illuminant is white.

a more general class of non-Lambertian scenes.

Acknowledgment

We would like to acknowledge the authors of the algorithms used in Sect. 4 who generously released their code on the Internet—Vladimir Kolmogorov (stereo) and University of Central Florida vision group (shape from shading)—or provided an implementation to us upon our request—-Jongwoo Lim (passive photometric stereo) and Michael J. Black (robust optical flow).

This work was supported in part by the National Science Foundation. T. Zickler was funded under IIS-05-41173, S. Mallick and D. Kriegman were funded under IIS-03-08185 and EIA-02-24431, and P. Belhumeur was funded under IIS-03-08185, EIA-02-24431, and ITR-00-85864.

References

[1] K. Barnard, V. Cardei, and B. Funt. A comparison of computational color constancy algorithms–part I: Methodology and experiments with synthesized data. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 11(9), 2002.

[2] M. J. Black and P. Anandan. A framework for the robust estimation of optical flow. In *Proc. IEEE Int. Conf. Computer Vision*, pages 231–236, 1993.

[3] Y. Boykov, O. Veksler, and R. Zabih. Markov random fields with efficient approximations. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 648–655, 1998.

[4] S. D. Hordley, G. D. Finlayson, and M. S. Drew. Removing shadows from images. In *Proc. European Conf. Computer Vision*, pages 823–836, 2002.

[5] H. Jin, D. Cremers, A. Yezzi, and S. Soatto. Shedding light on stereoscopic segmentation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2004.

[6] G. Klinker, S. Shafer, and T. Kanade. The measurement of highlights in color images. *Int. Journal of Computer Vision*, 2(1):7–32, 1988.

[7] J. Lim, J. Ho, M.-H. Yang, and D. Kriegman. Passive photometric stereo from motion. In *Proc. IEEE Int. Conf. Computer Vision*, 2005.

[8] S. P. Mallick, T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman. Specularity Removal In Images and Videos: A PDE Approach. In *Proc. European Conf. Computer Vision*, 2006.

[9] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur. Beyond Lambert: Reconstructing specular surfaces using color. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005.

[10] S. G. Narasimhan, V. Ramesh, and S. K. Nayar. A class of photometric invariants: Separating material from shape and illumination. In *Proc. IEEE Int. Conf. Computer Vision*, volume 2, pages 1387–1394, 2003.

[11] S. K. Nayar and M. Bolle. Reflectance based object recognition. *Int. Journal of Computer Vision*, 17(3):219–240, 1996.

[12] J. B. Park. Efficient color representation for image segmentation under nonwhite illumination. In *SPIE, Volume 5267*, pages 163–174, 2003.

[13] H. Ragheb and E. Hancock. Separating lambertian and specular reflectance components using iterated conditional modes. In *Proc. British Machine Vision Conference*, pages 541–522, 2001.

[14] S. Shafer. Using color to separate reflection components. *COLOR research and applications*, 10(4):210–218, 1985.

[15] R. T. Tan and K. Ikeuchi. Separating reflection components of textured surface using a single image. In *Proc. IEEE Int. Conf. Computer Vision*, pages 870–877, 2003.

[16] R. T. Tan, K. Nishino, and K. Ikeutchi. Color constancy through inverse-intensity chromaticity space. *J. Optical Society of America A*, 21(3):321–334, 2004.

[17] J. van de Weijer and T. Gevers. Robust optical flow from photometric invariants. In *Proc. IEEE Int. Conf. Image Processing*, pages 1835–1838, 2004.

[18] L. Wolff and E. Angelopoulou. Three-dimensional stereo by photometric ratios. *J. Optical Society of America A*, 11:3069–3078, 1994.

[19] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. In *Proc. IEEE Int. Conf. Computer Vision*, pages 618–625, 2003.

[20] Q. Zheng and R. Chellappa. Estimation of illuminant direction, albedo, and shape from shading. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(7):680–702, 1991.