On the Removal of Shadows from Images

Graham D. Finlayson, *Member, IEEE*, Steven D. Hordley, Cheng Lu, and Mark S. Drew, *Member, IEEE*

Abstract—This paper is concerned with the derivation of a progression of shadow-free image representations. First, we show that adopting certain assumptions about lights and cameras leads to a 1D, gray-scale image representation which is illuminant invariant at each image pixel. We show that as a consequence, images represented in this form are shadow-free. We then extend this 1D representation to an equivalent 2D, chromaticity representation. We show that in this 2D representation, it is possible to relight all the image pixels in the same way, effectively deriving a 2D image representation which is additionally shadow-free. Finally, we show how to recover a 3D, full color shadow-free image representation by first (with the help of the 2D representation) identifying shadow edges. We then remove shadow edges from the edge-map of the original image by edge in-painting and we propose a method to reintegrate this thresholded edge map, thus deriving the sought-after 3D shadow-free image.

Index Terms—Shadow removal, illuminant invariance, reintegration.

✦

1 INTRODUCTION

O NE of the most fundamental tasks for any visual system is that of separating the changes in an image which are due to a change in the underlying imaged surfaces from changes which are due to the effects of the scene illumination. The interaction between light and surface is complex and introduces many unwanted artifacts into an image. For example, shading, shadows, specularities, and interreflections, as well as changes due to local variation in the intensity or color of the illumination all make it more difficult to achieve basic visual tasks such as image segmentation [1], object recognition [2], and tracking [3]. The importance of being able to separate illumination effects from reflectance has been well understood for a long time. For example, Barrow and Tenenbaum [4] introduced the notion of "intrinsic images" to represent the idea of decomposing an image into two separate images: one which records variation in reflectance and another which represents the variation in the illumination across the image.

Barrow and Tenenbaum proposed methods for deriving such intrinsic images under certain simple models of image formation. In general, however, the complex nature of image formation means that recovering intrinsic images is an ill-posed problem. More recently, Weiss [5] proposed a method to derive an intrinsic reflectance image of a scene given a sequence of images of the scene under a range of illumination conditions. Using many images ensures that the problem is well-posed, but implies that the application of the method is quite restricted. The Retinex and Lightness algorithms of Land [6] and others [7], [8], [9], [10] can also be seen as an attempt to derive intrinsic reflectance images under certain restrictive scene assumptions. Specifically, those algorithms

are founded on the premise that scenes are 2D planar surfaces constructed from a tessellation of uniform reflectance patches. In addition, the intensity of illumination across the scene is assumed to vary only slowly and is assumed to be spectrally constant. Under these conditions, it is possible to distinguish changes in reflectance from changes in illumination and to factor the latter out, thus deriving an intrinsic reflectance image referred to as a lightness image.

Estimating and accounting for the color of the prevailing scene illumination is a related problem which has received much attention [11], [12], [13], [14], [15], [16], [17], [18], [19], [20]. In this body of work, the focus is not on deriving intrinsic reflectance images, but rather on obtaining a rendering of a scene as it would appear when viewed under some standard illumination. Often, these color constancy algorithms as they are called, are derived under the same restrictive conditions as the lightness algorithms and factors such as specularities, shading, and shadows are ignored. A different approach to this problem is the so-called illuminant invariant approach [21], [22], [23], [24], [25], [26], [27]. Instead of attempting to estimate the color of the scene illuminant, illuminant invariant methods attempt simply to remove its effect from an image. This is achieved by deriving invariant quantities —algebraic transformations of the recorded image values— which remain constant under a change of illumination. Methods for deriving quantities which are invariant to one or more of illumination color, illumination intensity, shading, and specularities have all been proposed in the literature.

In this paper, we consider how we might account for shadows in an imaged scene: an illumination which has so far largely been ignored in the body of work briefly reviewed above. That accounting for the effect of shadows on color constancy in images has not received more attention is somewhat surprising since shadows are present in many images and can confound many visual tasks. As an example, consider that we wish to segment the image in Fig. 2a into distinct regions each of which corresponds to an underlying surface reflectance. While humans can solve this task easily, identifying two important regions corresponding to the grass and the path, such an image will cause problems for a segmentation algorithm, which will quite likely return at least three regions corresponding to shadow, grass, and path. In

● *G.D. Finlayson and S.D. Hordley are with the School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ UK. E-mail: {graham, steve}@cmp.uea.ac.uk.*
● *C. Lu and M.S. Drew are with the School of Computing Science, Simon Fraser University, Burnaby, British Columbia, V5A 1S6 Canada. E-mail: {clu, mark}@cs.sfu.ca.*

fact, identifying shadows and accounting for their effects is a difficult problem since a shadow is in effect a local change in both the color and intensity of the scene illumination. To see this, consider again Fig. 2a. In this image, the nonshadow region is illuminated by light from the sky and also by direct sunlight, whereas in contrast, the shadow region is lit only by light from the sky. It follows that to account for shadows, we must be able, in effect, to locally solve the color constancy problem—that is, identify the color of the scene illuminant at each pixel in the scene.

We propose three different shadow-free image representations in this paper. We begin by summarizing previous work [28], [29] which showed that given certain assumptions about scene illumination and camera sensors it is possible to solve a restricted color constancy problem at a single image pixel. Specifically, given a single triplet of sensor responses it is possible to derive a 1D quantity invariant to both the color and intensity of the scene illuminant. This in effect provides a 1D reflectance image which is, by construction, shadow-free. Importantly, results in this paper demonstrate that applying the theory to images captured under conditions which fail to satisfy one or more of the underlying assumptions, still results in gray-scale images which are, to a good approximation, shadow-free. Next, we consider how to put some of the color back in to the shadow-free representation. We show that there exists an equivalent 2D representation of the invariant image which is also locally illuminant invariant and, therefore, shadow free. Furthermore, we show that given this 2D representation we can put some illumination back into the scene. That is, we can relight all image pixels uniformly (using, e.g., the illumination in the nonshadow region of the original image) so that the image remains shadow-free but is closer in color to a 2D representation of the original image. This 2D image representation is similar to a conventional chromaticity [30] representation (an intensity invariant representation) but with the additional advantage of being shadow-free.

Finally, we show how to recover a full-color 3D image representation which is the same as the original image but with shadows removed. Here, our approach is similar to that taken in lightness algorithms [6], [7], [8], [10]. In that work, the effects of illumination are factored out by working with an edge representation of the image, with small edges assumed to correspond to the slowly changing illumination while large changes correspond to a change in reflectance. Under these assumptions, small changes are factored out and the resulting edge-map is reintegrated to yield an illumination-free lightness image. In our case, we also work with an edge-map of the image, but we are concerned with separating shadow edges from reflectance edges and factoring out the former. To do so, we employ the 2D shadow-free image we have earlier derived. We reason that a shadow edge corresponds to any edge which is in the original image but absent from the invariant representation and we can thus define a thresholding operation to identify the shadow edge. Of course, this thresholding effectively introduces small contours in which we have no edge information. Thus, we propose a method for in-painting edge information across the shadow edge. Finally, reintegrating yields a color image, equal to the original save for the fact that it is shadow-free.

Before developing the theory of shadow-free images it is useful to set out some initial assumptions and limitations of our approach. The derivation of a one-dimensional image representation, invariant to both illumination color and intensity, is founded on a Lambertian model of image formation. That is, we assume that image pixel values are linearly related to the intensity of the incident light, and that images are free of effects such as specularities and interreflections. Furthermore, the theory is developed under the assumption of an imaging device with perfectly narrow-band sensors (sensors responsive to just a single wavelength of light), and we also assume that our scenes are lit by Planckian illuminants. Of course, not all of these assumptions will be satisfied for an image of an arbitrary scene, taken with a typical imaging device. However, the theory we develop can be applied to any image, and we discuss, in Section 2, the effect that departures from the theoretical case have on the resulting 1D invariant representation. A more detailed discussion of these issues can also be found in other works [28], [31]. It is also important to point out that, for some images, the process of transforming the original RGB representation to the 1D invariant representation might also introduce some undesirable artifacts. Specifically, two or more surfaces which are distinguishable in a 3D representation, may be indistinguishable (that is, metameric) in the 1D representation. For example, two surfaces which differ only in their intensity will have identical 1D invariant representations. The same will be true for surfaces which are related by a change of illumination (as defined by our model). Similar artifacts can be introduced when we transform an image from an RGB representation to a 1D gray-scale representation since they are a direct consequence of the transformation from a higher to lower dimensional representation. The two and three-dimensional shadow-free representations we introduce are both derived from the 1D invariant. This implies that the assumptions and limitations for the 1D case also hold true for the higher dimensional cases. The derivation of the 3D shadow-free image also includes an edge detection step. Thus, in this case, we will not be able to remove shadows which have no edges or whose edges are very ill-defined. In addition, we point out that edge detection in general is still an open problem, and the success of our method is therefore limited by the accuracy of existing edge detection techniques. Notwithstanding the theoretical limitations we have set out, the method is capable of giving very good performance on real images. For example, all the images in Fig. 5 depart from one or more of the theoretical assumptions and, yet, the recovered 1D, 2D, and 3D representations are all effectively shadow-free.

The paper is organized as follows: In Section 2, we summarize the 1D illuminant invariant representation and its underlying theory. In Section 3, we extend this theory to derive a 2D representation, and we show how to add illumination back in to this image, resulting in a 2D shadow-free chromaticity image. In Section 4, we present our algorithm for deriving the 3D shadow-free image. Finally, in Section 5, we give some examples illustrating the three methods proposed in this paper and we conclude the paper with a brief discussion.

2 ONE-DIMENSIONAL SHADOW FREE IMAGES

Let us begin by briefly reviewing how to derive one-dimensional shadow-free images. We summarize the analysis given in [28] for a three-sensor camera but note that the same analysis can be applied to cameras with more than three
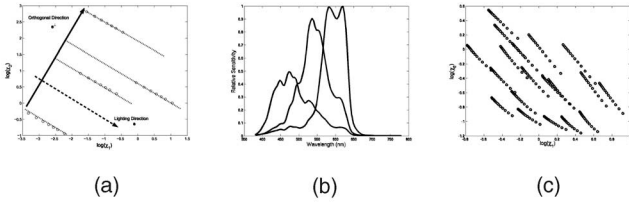
Fig. 1. (a) An illustration of the 1D invariant representation, for an ideal camera and Planckian illumination. (b) The spectral sensitivities of a typical digital still camera. (c) The log-chromaticities calculated using the sensitivities from (b) and a set of daylight illuminants.

sensors, in which case, it is possible to account for other artifacts of the imaging process (e.g., in [32] a four-sensor camera was considered and it was shown that, in this case, specularities could also be removed).

We adopt a Lambertian model [33] of image formation so that if a light with a spectral power distribution (SPD) denoted $E(\lambda, x, y)$ is incident upon a surface whose surface reflectance function is denoted $S(\lambda, x, y)$, then the response of the camera sensors can be expressed as:

$$\rho_k(x, y) = \sigma(x, y) \int E(\lambda, x, y) S(\lambda, x, y) Q_k(\lambda) d\lambda, \qquad (1)$$

where $Q_k(\lambda)$ denotes the spectral sensitivity of the kth camera sensor, $k = 1, 2, 3$ and $\sigma(x, y)$ is a constant factor which denotes the Lambertian shading term at a given pixel—the dot product of the surface normal with the illumination direction. We denote the triplet of sensor responses at a given pixel (x, y) location by $\underline{\rho}(x, y) = [\rho_1(x, y), \rho_2(x, y), \rho_3(x, y)]^T$.

Given (1), it is possible to derive a 1D illuminant invariant (and, hence, shadow-free) representation at a single pixel given the following two assumptions: First, the camera sensors must be exact Dirac delta functions and, second, illumination must be restricted to be Planckian [34]. If the camera sensitivities are Dirac delta functions, $Q_k(\lambda) = q_k \delta(\lambda - \lambda_k)$. Then, (1) becomes simply:

$$\rho_k = \sigma E(\lambda_k) S(\lambda_k) q_k, \qquad (2)$$

where we have dropped for the moment the dependence of ρ_k on spatial location. Restricting illumination to be Planckian or, more specifically, to be modeled by Wien's approximation to Planck's law [34], an illuminant SPD can be parameterized by its color temperature T:

$$E(\lambda, T) = I c_1 \lambda^{-5} e^{-\frac{c_2}{T\lambda}}, \qquad (3)$$

where c_1 and c_2 are constants and I is a variable controlling the overall intensity of the light. This approximation is valid for the range of typical lights $T \in [2,500, 10,000]^o K$. With this approximation, the sensor responses to a given surface can be expressed as:

$$\rho_k = \sigma I c_1 \lambda_k^{-5} e^{-\frac{c_2}{T\lambda_k}} S(\lambda_k) q_k. \qquad (4)$$

Now, let us form band-ratio two-vector chromaticities $\underline{\chi}$:

$$\chi_j = \frac{\rho_k}{\rho_p}, \quad k \in \{1, 2, 3\}, \quad k \neq p, \quad j = 1, 2, \qquad (5)$$

e.g., for an RGB image, $p = 2$ means $\rho_p = G$, $\chi_1 = R/G$, $\chi_2 = B/G$. Substituting the expressions for ρ_k from (4) into
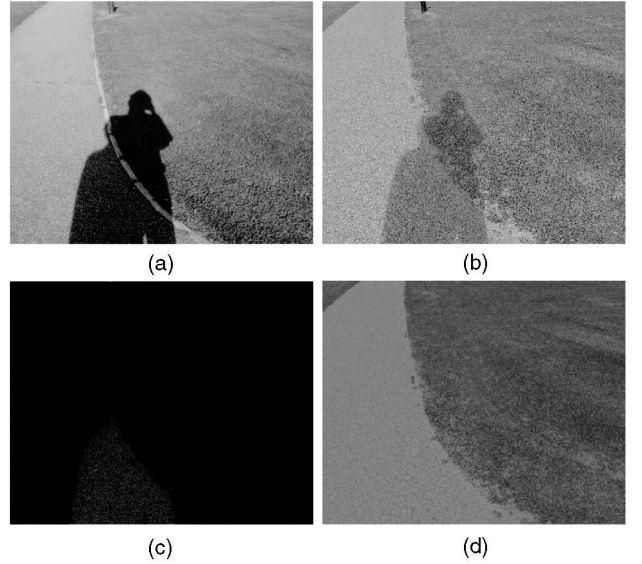

Fig. 2. An example of the 1D illuminant invariant representation. (a) The original image. (b) and (c) Log-chromaticity representations (χ_1' and χ_2'). (d) The 1D invariant $\mathcal{I}$.

(5), we see that forming the chromaticity coordinates removes intensity and shading information:

$$\chi_j = \frac{\lambda_k^{-5} e^{-\frac{c_2}{T\lambda_k}} S(\lambda_k) q_k}{\lambda_p^{-5} e^{-\frac{c_2}{T\lambda_p}} S(\lambda_p) q_p}. \qquad (6)$$

If we now form the logarithm $\underline{\chi}'$ of $\underline{\chi}$, we obtain:

$$\chi_j' = \log \chi_j = \log\left(\frac{s_k}{s_p}\right) + \frac{1}{T}(e_k - e_p), \quad j = 1, 2, \qquad (7)$$

where $s_k \equiv \lambda_k^{-5} S(\lambda_k) q_k$ and $e_k \equiv -c_2/\lambda_k$.

Summarizing (7) in vector form, we have:

$$\underline{\chi}' = \underline{s} + \frac{1}{T} \underline{e}, \qquad (8)$$

where $\underline{s}$ is a two-vector which depends on surface and camera, but is independent of the illuminant, and $\underline{e}$ is a two-vector which is independent of surface, but which again depends on the camera. Given this representation, we see that as illumination color changes (T varies) the log-chromaticity vector $\underline{\chi}'$ for a given surface moves along a straight line. Importantly, the direction of this line depends on the properties of the camera, but is independent of the surface and the illuminant.

It follows that, if we can determine the direction of illuminant variation (the vector $\underline{e}$), then we can determine a 1D illuminant invariant representation by projecting the log-chromaticity vector $\underline{\chi}'$ onto the vector orthogonal to $\underline{e}$, which we denote $\underline{e}^\perp$. That is, our illuminant invariant representation is given by a gray-scale image $\mathcal{I}$:

$$\mathcal{I}' = \underline{\chi}'^T \underline{e}^\perp, \qquad \mathcal{I} = \exp(\mathcal{I}'). \qquad (9)$$

Without loss of generality, we assume that $\|\underline{e}^\perp\| = 1$. Fig. 1a illustrates the process we have just described. The figure shows log-chromaticities for four different surfaces (open circles), for perfect narrow-band sensors under a range of Planckian illuminants. It is clear that the chromaticities for

each surface fall along a line (dotted lines in the figure) in chromaticity space. These lines have direction $\underline{e}$. The direction orthogonal to $\underline{e}$ is shown by a solid line in Fig. 1a. Each log-chromaticity for a given surface projects to a single point along this line regardless of the illumination under which it is viewed. These points represent the illuminant invariant quantity $\mathcal{I}'$ as defined in (9).

Note that to remove any bias with respect to which color channel to use as a denominator, we can divide by the geometrical mean $\rho_M = \sqrt[3]{RGB}$ in (5) instead of a particular ρ_p and still retain our straight line dependence. Log-color ratios then live on a plane in three-space orthogonal to $\underline{u} = (1, 1, 1)^T$ and form lines exactly as in Fig. 1a [35].

We have derived this 1D illuminant invariant representation under quite restrictive conditions (though the conditions on the camera can be relaxed to broad-band sensors with the addition of some conditions on the reflectances [36]) and it is therefore reasonable to ask: In practice, is the method at all useful? To answer this question, we must first calculate the orthogonal projection direction for a given camera. There are a number of ways to do this but the simplest approach is to image a set of reference surfaces (we used a Macbeth Color Checker Chart which has 19 surfaces of distinct chromaticity) under a series of n lights. Each surface produces n log-chromaticities which, ideally, will fall along straight lines. Moreover, the individual chromaticity lines will also be parallel to one another. Of course, because real lights may be non-Planckian and camera sensitivities are not Dirac delta functions, we expect there to be departures from these conditions. Fig. 1b shows the spectral sensitivities of a typical commercial digital still camera and, in Fig. 1c, we show the log-chromaticity coordinates calculated using these sensitivity functions, the surfaces of a Macbeth Color Checker and a range of daylight illuminants. It is clear that the chromaticity coordinates do not fall precisely along straight lines in this case. Nevertheless, they do exhibit approximately linear behavior and, so, can we solve for the set of n parallel lines which best account for our data in a least-squares sense [28]. Once we know the orthogonal projection direction for our camera, we can calculate log-chromaticity values for any arbitrary image. The test of the method is then whether the resulting invariant quantity $\mathcal{I}$ is indeed illuminant invariant.

Fig. 2 illustrates the method for an image taken with the camera (modified such that it returns linear output without any image postprocessing) whose sensitivities are shown in Fig. 1b. Fig. 2a shows the color image as captured by the camera (for display purposes, the image is mapped to sRGB [37] color space)—a shadow is very prominent. Figs. 2b and 2c show the log-chromaticity representation of the image. Here, intensity and shading are removed but the shadow is still clearly visible, highlighting the fact that shadows represent a change in the color of the illumination and not just its intensity. Finally, Fig. 2d shows the invariant image (a function of Figs. 2b and 2c) defined by (9). Visually, it is clear that the method delivers very good illuminant invariance: The shadow is not visible in the invariant image. This image is typical of the level of performance achieved with the method. Fig. 5 illustrates some more examples for images taken with a variety of real cameras (with nonnarrow-band sensors). We note that in some of these examples, the camera sensors were unknown and we estimated the illumination direction using an automatic

procedure described elsewhere [35]. In all cases, shadows are completely removed or greatly attenuated.

In other work [28], we have shown that the 1D invariant images are sufficiently illuminant invariant to enable accurate object recognition across a range of illuminants. In that work, histograms derived from the invariant images were used as features for recognition and it is notable that the recognition performance achieved was higher than that obtained using a color constancy approach [38]. It is also notable that the images used in that work were captured with a camera whose sensors are far from narrow-band, and under non-Planckian illuminants. An investigation as to the effect of the shape of camera sensors on the degree of invariance has also been carried out [31]. That work showed that good invariance was achieved using Gaussian sensors with a half bandwidth of up to 30 nm, but that the degree of invariance achievable was somewhat sensitive to the location of the peak sensitivities of the sensors. This suggests that there is not a simple relationship between the shape and width of sensors and the degree of invariance, so that the suitability of sensors is best judged on a camera by camera basis. In other work [39], it has been shown that it is possible to find a fixed 3×3 linear transform of a given set of sensor responses so that the 1D image representation derived from the transformed sensors has improved illuminant invariance. In addition, we also note that, for any set of camera sensors, it is possible to find a fixed 3×3 linear transform which when applied to the sensors brings them closer to the ideal of narrow-band sensors [40]. Finally, we point out that in our studies to date, we have not found a set of camera sensors for which the 1D representation does not provide a good degree of illuminant invariance.

3 TWO-DIMENSIONAL SHADOW FREE IMAGES

In the 1D invariant representation described above we removed shadows but at a cost: We have also removed the color information from the image. In the rest of this paper we investigate how we can put this color information back in to the image. Our aim is to derive an image representation which is shadow-free but which also has some color information. We begin by observing that the 1D invariant we derived in (9) can be equally well be expressed as a 2D log-chromaticity. Looking again at Fig. 1, we see that an invariant quantity is derived by projecting 2D log-chromaticities onto the line in the direction $\underline{e}^\perp$. Equally, we can represent the point to which a pixel is projected by its 2D coordinates in the log-chromaticity space, thus retaining some color information. That is, we derive a 2D color illumination invariant as:

$$\underline{\tilde{\chi}}' = P_{\underline{e}^\perp} \underline{\chi}', \tag{10}$$

where $P_{\underline{e}^\perp}$ is the 2×2 projector matrix:

$$P_{\underline{e}^\perp} = \underline{e}^\perp \underline{e}^{\perp T}. \tag{11}$$

$P_{\underline{e}^\perp}$ takes log-chromaticity values onto the direction orthogonal to $\underline{e}$ but preserves the resulting quantity as a two-vector $\underline{\tilde{\chi}}'$. The original 1D invariant quantity $\mathcal{I}'$ is related to $\underline{\tilde{\chi}}'$ by:

$$\mathcal{I}' = \underline{\tilde{\chi}}' \cdot \underline{e}^\perp. \tag{12}$$

To visualize the 2D invariant image, it is useful to express the 2D chromaticity information in a 3D form. To do so, we write the projected chromaticity two-vector $\underline{\tilde{\chi}}'$ that lies in a plane
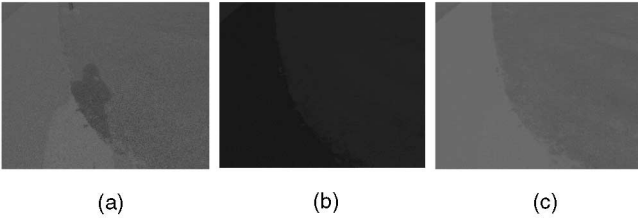
Fig. 3. (a) A conventional chromaticity representation. (b) The 2D invariant representation ($\underline{\tilde{\chi}}$). (c) The 2D invariant with lighting added back in.

orthogonal to $\underline{u} = (1, 1, 1)^T$ in its equivalent three-space coordinates $\underline{\tilde{\rho}}'$. We do this by multiplying by the 3×2 matrix U^T which decomposes the projector onto that plane:

$$\underline{\tilde{\rho}}' = U^T \underline{\tilde{\chi}}', \qquad (13)$$

where $UU^T = I - \underline{u}\underline{u}^T / \|\underline{u}\|^2$ and the resulting $\underline{\tilde{\rho}}'$ is a three-vector. Note, this transformation is not arbitrary: Any 2D log-chromaticity coordinates are othogonal to $(1, 1, 1)$ (intensity) and, so, we must map 2D to 3D accordingly. Finally, by exponentiating (13), we recover an approximation of color:

$$\underline{\tilde{\rho}} = \exp(\underline{\tilde{\rho}}'). \qquad (14)$$

Note that (14) is a three-dimensional representation of 2D information: $\tilde{\rho}$ contains no brightness or shading information and, so, is still effectively a chromaticity representation. The usual way to derive an intensity independent representation of 3D color is to normalize a 3D sensor response $\underline{\rho}$ by the sum of its elements [30]. We take our 3D representation into this form by applying an L_1 normalization:

$$\underline{\rho} = \{\tilde{\rho}_1, \tilde{\rho}_2, \tilde{\rho}_3\}^T / (\tilde{\rho}_1 + \tilde{\rho}_2 + \tilde{\rho}_3). \qquad (15)$$

This representation is bounded in $[0, 1]$ and we have found that it has good stability.

An illustration of the method is shown in Fig. 3. Fig. 3a shows the L_1 chromaticity representation $\underline{r}$ of an image, with intensity and shading information factored out: $\underline{r} = \{R, G, B\} / (R + G + B)$. It is important to note that in this representation the shadow is still visible—it represents a change in the color of the illumination and not just its intensity. Fig. 3b shows the illumination invariant chromaticity representation derived in (10), (11), (12), (13), (14), and (15) above. Now, the shadow is no longer visible, indicating that the method has successfully removed the shadow, while still maintaining some color information. Comparing Figs. 3a and 3b, we see that the colors in the two images are quite different. This is because the representation in Fig. 3b has had all its illumination removed and, thus, it is in effect an intrinsic reflectance image. To recover a color representation closer to that in Fig. 3b, we must put the illumination back into the representation [41]. Of course, we don't want to add illumination back on a pixel-by-pixel basis since this would simply reverse what we have just done and result in an image representation which once again contains shadows. To avoid this, we want to relight each pixel uniformly by "adding back" illumination. To see how to do this, consider again the 2D chromaticity representation defined in (10). In this representation, illumination is represented by a vector of arbitrary magnitude in the direction $\underline{e}$:

$$\text{illumination} = \underline{\chi}'_E = a_E \underline{e}. \qquad (16)$$

We can put this light back into the illuminant invariant representation defined in (10) by simply adding the chromaticity of the light to the invariant chromaticities:

$$\underline{\tilde{\chi}}' \to \underline{\tilde{\chi}}' + \underline{\chi}'_E = \underline{\tilde{\chi}}' + a_E \underline{e}. \qquad (17)$$

The color of the light we put back in is controlled by the value of a_E. To determine what light to add back in, we observe that the pixels in the original image that are brightest, correspond to surfaces that are not in shadow. It follows then that, if we base our light on these bright pixels, then we can use this light to relight all pixels. That is, we find a suitable value of a_E by minimizing

$$\|\underline{\chi}'_b - (\underline{\tilde{\chi}}'_b + a_E \underline{e})\|, \qquad (18)$$

where $\underline{\chi}'_b$ and $\underline{\tilde{\chi}}'_b$ correspond to the log-chromaticity and the invariant log-chromaticity of bright (nonshadow) image pixels. Once we have added the lighting back in this way, we can represent the resulting chromaticity information in 3D by applying (15).

Fig. 3c shows the resulting chromaticity representation with lighting added back in. Here, we found a_E by minimizing the term in (18) for the brightest 1 percent of pixels in the image. The colors are now much closer to those in the conventional chromaticity image (Fig. 3a) but are still not identical. The remaining difference is due to the fact that when we project chromaticities orthogonally to the illuminant direction we remove illumination, as well as any part of a surface's color which is in this direction. This part of the object color is not easily put back into the image. Nevertheless, for many surfaces the resulting chromaticity image is close to the original, with the advantage that the representation is shadow-free. Fig. 5 shows this shadow-free chromaticity representation for a variety of different images. In all cases, shadows are successfully removed.

4 THREE-DIMENSIONAL SHADOW-FREE IMAGES

The 2D chromaticity representation of images is often very useful. By additionally removing shadows from this representation, we have gained a further advantage and increased the value of a chromaticity representation. However, there is still room for improvement. Chromaticity images lack shading and intensity information and are also unnaturally colored. In some applications, an image which is free of shadows, but which is otherwise the same as a conventional color image would be very useful. In this section, we consider how such an image might be obtained.

4.1 The Recovery Algorithm

Our method for obtaining full-color shadow removal has its roots in methods of lightness recovery [8], [9], [7], [10], [6]. Lightness algorithms take as their input a 3D color image and return two intrinsic images: one based on reflectance (the lightness image) and the other based on illumination. Lightness computation proceeds by making the assumption that illumination varies slowly across an image whereas changes in reflectance are rapid. It follows then that by thresholding a derivative image to remove small derivatives, slow changes (due, by assumption, to illumination)

can be removed. Integrating the thresholded derivative image results in the lightness intrinsic image.

Importantly, a lightness scheme will not remove shadows since, although they are a change in illumination, at a shadow edge the illumination change is fast, not slow. Given their assumptions, lightness algorithms are unable to distinguish shadow edges from material edges. However, in our case, we have the original image which contains shadows and we are able to derive from it 1D or 2D images which are shadow-free. Thus, by comparing edges in the original and the shadow-free images, we can identify those edges which correspond to a shadow. Modifying the thresholding step in the lightness algorithm leads to an algorithm which can recover full-color shadow-free images. There are two important steps which must be carefully considered if the algorithm is to work in practice. First, the algorithm is limited by the accuracy with which we can identify shadow edges. Second, given the location of the shadow edges, we must give proper consideration to how this can be used in a lightness type algorithm to recover the shadow-free image.

Let us begin by defining the recovery algorithm. We use the notation $\rho_k(x,y)$ to denote the gray-scale image corresponding to a single band of the 3D color image. Lightness algorithms work by recovering an intrinsic image from each of these three bands separately and combining the three intrinsic images to form a color image. We observe in (4) that under the assumption of Dirac delta function sensors, sensor response is a multiplication of light and surface. Let us transform sensor responses into log space so that the multiplication becomes an addition:

$$\rho_k'(x,y) = \sigma'(x,y) + E'(\lambda_k,x,y) + S'(\lambda_k,x,y) + q_k'. \quad (19)$$

In the original lightness algorithm, the goal is to remove illumination and as a first step toward this, gradients are calculated for the log-image:

$$\nabla_x\rho_k'(x,y) = \frac{\partial}{\partial x}\rho_k'(x,y)$$

$$\nabla_y\rho_k'(x,y) = \frac{\partial}{\partial y}\rho_k'(x,y). \quad (20)$$

These gradients define edge maps for the log image. Next, a threshold operator $T(\cdot)$ is defined to remove gradients of small magnitude:

$$T(\nabla_i\rho_k'(x,y)) = \begin{cases} 0 & \text{if } \|\nabla_i\rho_k'(x,y)\| < \tau \\ \nabla_i\rho_k'(x,y) & \text{otherwise,} \end{cases} \quad (21)$$

where $i \in \{x,y\}$ and τ is the chosen threshold value.

In our case, the goal is not to remove illumination per se (the small values in (21) above) but rather we wish only to remove shadows. In fact, we actually want to keep the illuminant field and rerender the scene as if it were captured under the same single nonshadow illuminant. To do this, we must factor out changes in the gradient at shadow edges. We can do this by modifying the threshold operator defined in (21). In principle, identifying shadows is easy: We look for edges in the original image which are not present in the invariant representation. However, in practice, the procedure is somewhat more complicated than this. For now, let us assume that we have identified the shadow edge and leave a

discussion of how we find it to the next section. Let us define a function $q_s(x,y)$ which defines the shadow edge:

$$q_s(x,y) = \begin{cases} 1 & \text{if } (x,y) \text{ is a shadow edge} \\ 0 & \text{otherwise.} \end{cases} \quad (22)$$

We can then remove shadows in the gradients of the log image using the threshold function $T_S(\cdot)$:

$$T_S(\nabla_i\rho_k', q_s(x,y)) = \begin{cases} 0 & \text{if } q_s(x,y) = 1 \\ \nabla_i\rho_k' & \text{otherwise,} \end{cases} \quad (23)$$

where again $i \in \{x,y\}$. That is, wherever we have identified that there is a shadow edge we set the gradient in the log-image to zero, indicating that there is no change at this point (which is true for the underlying reflectance). After thresholding, we obtain gradients where sharp changes are indicative only of material changes: There are no sharp changes due to illumination and, so, shadows have been removed.

We now wish to integrate edge information in order to recover a log-image which does not have shadows. We do this by first taking the gradients of the thresholded edge maps we have just defined to form a modified (by the threshold operator) Laplacian of the log-image:

$$\nabla_{T_S}^2\rho_k'(x,y) = \begin{aligned} &\nabla_x T_S(\nabla_x\rho_k'(x,y), q_s(x,y)) \\ &+\nabla_y T_S(\nabla_y\rho_k'(x,y), q_s(x,y)). \end{aligned} \quad (24)$$

Now, let us denote the shadow-free log-image which we wish to recover as $\tilde\rho'(x,y)$ and equate its Laplacian to the modified Laplacian we have just defined:

$$\nabla^2\tilde\rho_k'(x,y) = \nabla_{T_S}^2\rho_k'(x,y). \quad (25)$$

Equation (25) is the well-known Poisson equation. The shadow-free log-image can be calculated via:

$$\tilde{\rho_k}'(x,y) = (\nabla^2)^{-1}\nabla_{T_S}^2\rho_k'(x,y). \quad (26)$$

However, since the Laplacian is not defined at the image boundary, we must specify boundary conditions for uniqueness. Blake [8] made use of Neumann boundary conditions, in which the normal derivative of the image is specified at its boundary. Here, we use homogeneous Neumann conditions: The directional derivative at the boundary is set to zero.

There are two additional problems with recovering $\tilde{\rho_k}'(x,y)$ according to (26) caused by the fact that we have removed shadow edges from the image. First, because we have modified the edge maps by setting shadow edges to zero, we can no longer guarantee that the edge map we are integrating satisfies the integrability condition. For the edge map to be integrable, the following condition should be met (cf. [42]):

$$\nabla_y\nabla_x\rho_k'(x,y) = \nabla_x\nabla_y\rho_k'(x,y). \quad (27)$$

The second problem is caused by the fact that to ensure shadows are effectively removed, we must set to zero, edges in quite a large neighborhood of the actual shadow edge. As a result, edge information pertaining to local texture in the neighborhood of the shadow edge is lost and the resulting (shadow-free) image is unrealistically smooth in this region. To avoid this problem, rather than simply setting shadow edges to zero in the thresholding step, we apply an iterative diffusion process which *fills in* the derivatives across shadow

edges, bridging values obtained from neighboring nonshadow edge pixels. We also deal with the problem of integrability at this stage by including a step at each iteration to enforce integrability, as proposed in [43].

This iterative process is detailed below where t denotes artificial time:

1. Initialization, $t = 0$, calculate:

$$\left(\nabla_x \rho'_k(x,y)\right)^t \rightarrow T_S\left(\nabla_x \rho'_k(x,y), q_s(x,y)\right)$$
$$\left(\nabla_y \rho'_k(x,y)\right)^t \rightarrow T_S\left(\nabla_y \rho'_k(x,y), q_s(x,y)\right).$$

2. Update shadow edge pixels (i,j):

$$\left(\nabla_x \rho'_k(i,j)\right)^t \rightarrow$$
$$\left(\nabla_x \rho'_k(i-1,j)\right)^{t-1} + \left(\nabla_x \rho'_k(i,j-1)\right)^{t-1}$$
$$\left(\nabla_x \rho'_k(i+1,j)\right)^{t-1} + \left(\nabla_x \rho'_k(i,j+1)\right)^{t-1},$$

$$\left(\nabla_y \rho'_k(i,j)\right)^t \rightarrow$$
$$\left(\nabla_y \rho'_k(i-1,j)\right)^{t-1} + \left(\nabla_y \rho'_k(i,j-1)\right)^{t-1}$$
$$+ \left(\nabla_y \rho'_k(i+1,j)\right)^{t-1} + \left(\nabla_y \rho'_k(i,j+1)\right)^{t-1}.$$

3. Enforce integrability by projection onto integrable edge map [43], and integrate:

$$F_x(u,v) = \mathcal{F}[\nabla_x \rho'_k], \quad F_y(u,v) = \mathcal{F}[\nabla_y \rho'_k],$$

$$a_x = e^{2\pi i u/N} - 1, \quad a_y = e^{2\pi i v/M} - 1,$$

$$Z(u,v) = \frac{a_x^* F_x(u,v) + a_y^* F_y(u,v)}{|a_x|^2 + |a_y|^2}, \quad \rho'(0,0) = 0,$$

$$(\nabla_x \rho')^t = \mathcal{F}^{-1}[a_x Z], \quad (\nabla_y \rho')^t = \mathcal{F}^{-1}[a_y Z],$$

where image size is $M \times N$ and $\mathcal{F}[\cdot]$ denotes the Fourier Transform. Here, we use forward-difference derivatives $\{-1,1\}^T$, $\{-1,1\}$ corresponding to the a_x, a_y above in the Fourier domain: i.e., the Fourier transform of a derivative $\nabla_x Z$ in the spatial domain corresponds to multiplication by $a_x(u)$ in the Fourier domain—this result simply follows by writing $\rho'(n+1) - \rho'(n)$ in terms of Fourier sums in the Discrete Fourier Transform (DFT). The projection step deriving $Z(u,v)$ follows [43], but for a forward-difference operator.

4. If $\|(\nabla_x \rho')^t - (\nabla_x \rho')^{t-1}\| + \|(\nabla_y \rho')^t - (\nabla_y \rho')^{t-1}\| \geq \epsilon$, $t \rightarrow t+1$, go to 2.
   where ϵ defines the stopping criterion.

Finally, we then solve the Poisson equation (26) using a final round of enforcing integrability by projection as above, with the reintegrated image given by

$$\tilde{\rho}'_k(x,y) = \mathcal{F}^{-1}[Z(u,v)]. \tag{28}$$

We actually operate on an image four times the original size, formed by symmetric replication in x and y, so as to enforce periodicity of the data for the DFT and homogeneous Neumann boundary conditions.

Equation (28) recovers $\tilde{\rho}'_k(x,y)$ up to an unknown constant of integration. Exponentiating $\tilde{\rho}'_k(x,y)$, we arrive
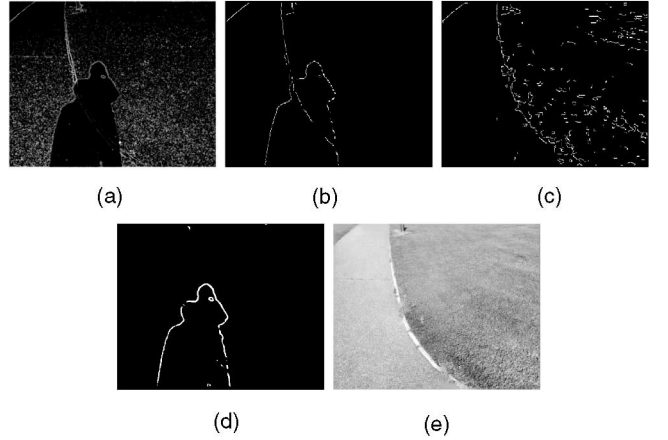

Fig. 4. (a) An edge-map obtained using simple finite differencing operators. (b) Edges obtained using the Canny operator on the Mean-Shifted original image. (c) Edges obtained using the Canny operator on the Mean-Shifted 2D invariant image. (d) The final shadow edge. (e) The recovered shadow-free color image.

at the reconstructed gray-scale image $\tilde{\rho}_k(x,y)$ (up to an unknown *multiplicative* constant). Solving (26) for each of the three color bands results in a full color image $\tilde{\underline{\rho}} = \{\tilde{\rho}_1 \ \tilde{\rho}_2 \ \tilde{\rho}_3\}^T$, where the shadows are removed.

To fix the unknown multiplicative factors, we apply a mapping to each pixel which maps the brightest pixels (specifically, the 0.005-percentile of pixels ordered by brightness) in the recovered image to the corresponding pixels in the original image.

4.2 Locating Shadow Edges

To complete the definition of the recovery algorithm, we must specify how to identify shadow edges. The essential idea is to compare edge maps of the original image to those derived from an invariant image, and to define a shadow edge to be any edge in the original which is not in the invariant image. We could start by calculating edge maps as simple finite difference approximations to gradients,

$$\nabla_x \rho_I(x,y) = \rho_I(x,y) \otimes \{-1, \ 0, \ 1\}^T/2,$$

$$\nabla_y \rho_I(x,y) = \rho_I(x,y) \otimes \{-1, \ 0, \ 1\}/2, \tag{29}$$

where $\rho_I(x,y)$ is the intensity image, taken here as the L_1 norm of the original image: $\rho_I = (1/3)(\rho_1 + \rho_2 + \rho_3)$. Unfortunately, as Fig. 4a illustrates, finite differencing produces nonzero values at more locations than those at which there are true edges. Thus, while in the example in Fig. 4a the edges of the road and the shadow are clear, so too are many edges due to the texture of the imaged surfaces as well as noise in the image. Obtaining the true edges in which we are interested from these edge maps is nontrivial, as evidenced by the large literature on edge detection (see [44] for a review).

For a more careful approach, we begin by applying a smoothing filter (specifically the Mean-Shift algorithm proposed in [45]) to both the original image and the 2D invariant image derived by exponentiating the invariant log image. This has the effect of suppressing features such as noise and high frequency textures so that in subsequent processing fewer spurious edges are detected. Then, we replace simple differencing by the Canny edge detector [46],
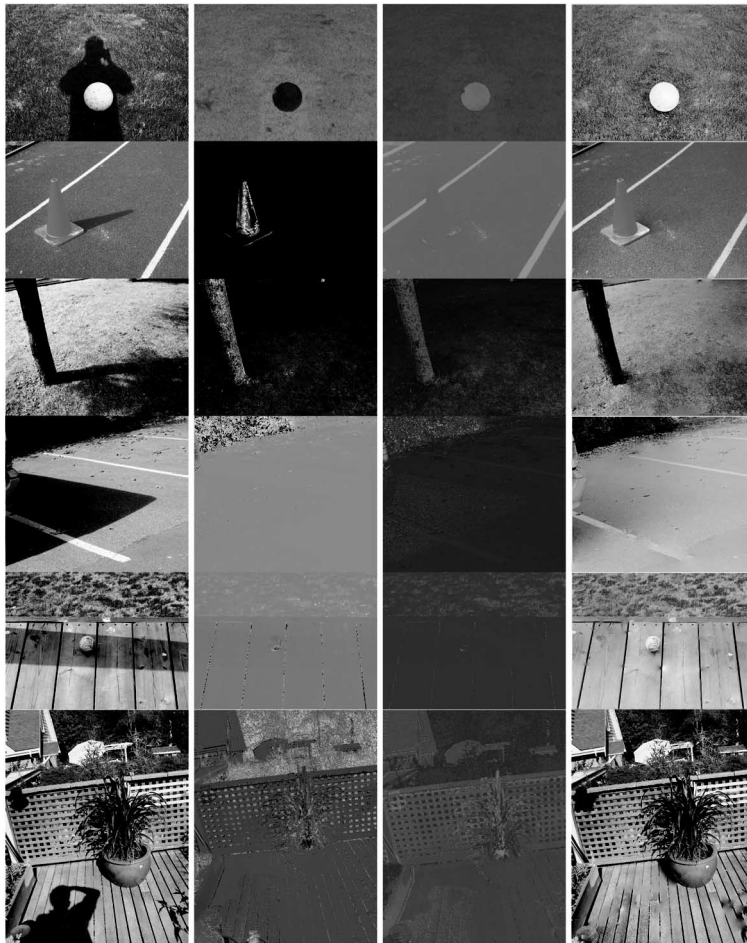
Fig. 5. Some example images. From left to right: original image, 1D invariant representation, 2D representation, and recovered 3D shadow-free image.

returning estimates for the strength of horizontal and vertical edges at each image location:

$$\|\tilde{\nabla}_x \rho_i(x,y)\| = C_x[\rho_i(x,y)]$$
$$\|\tilde{\nabla}_y \rho_i(x,y)\| = C_y[\rho_i(x,y)], \tag{30}$$

where $C_x[\cdot]$ and $C_y[\cdot]$ denote the Canny (or any other well-behaved) operators for determining horizontal and vertical edges, respectively.

We determine an edge map for the invariant image in a similar way, first calculating horizontal and vertical edge strengths for each channel of the 2D invariant image:

$$\|\tilde{\nabla}_x \chi_k(x,y)\| = C_x[\chi_k(x,y)]$$
$$\|\tilde{\nabla}_y \chi_k(x,y)\| = C_y[\chi_k(x,y)]. \tag{31}$$

The edge maps from the two channels are then combined by a max operation:

$$\|\tilde{\nabla}_x \underline{\chi}(x,y)\| = \max(C_x[\tilde{\chi}_1(x,y)], C_x[\tilde{\chi}_2(x,y)])$$
$$\|\tilde{\nabla}_y \underline{\chi}(x,y)\| = \max(C_y[\tilde{\chi}_1(x,y)], C_y[\tilde{\chi}_2(x,y)]), \tag{32}$$

where $\max(\cdot, \cdot)$ returns the maximum of its two arguments at each location (x,y). Figs. 4b and 4c show the resulting edge maps for the original image (calculated by (30)) and the invariant image (calculated by (31)-(32)). While still not perfect, the real edges in each image are now quite strong and we can compare the two edge maps to identify shadow edges.

We use two criteria to determine whether or not a given edge corresponds to a shadow. First, if at a given location the original image has a strong edge but the invariant image has a weak edge, we classify that edge as a shadow edge. Second, if both the original image and the invariant image have a strong edge, but the orientation of these edges is different, then we also classify the edge as a shadow edge. Thus, our shadow edge map is defined as:

$$q_s(x,y) = \begin{cases} 1 & \text{if } \|\tilde{\nabla}\rho_i\| > \tau_1 \quad\quad \& \quad \|\tilde{\nabla}\underline{\chi}\| < \tau_2 \\ & \text{or } \left| \frac{\|\tilde{\nabla}_x \rho_i\|}{\|\tilde{\nabla}_y \rho_i\|} - \frac{\|\tilde{\nabla}_x \underline{\chi}\|}{\|\tilde{\nabla}_y \underline{\chi}\|} \right| > \tau_3 \\ 0 & \text{otherwise,} \end{cases} \tag{33}$$

where τ_1, τ_2, and τ_3 are thresholds whose values are parameters in the recovery algorithm. As a final step, we employ a morphological operation (specifically, two dilations) on the binary edge map to "thicken" the shadow edges:

$$q_s(x,y) \rightarrow (q_s(x,y) \oplus D) \oplus D, \tag{34}$$

where $\oplus$ denotes the dilation operation and D denotes the structural element, in this case, the eight-connected set. This dilation has the effect of filling in some of the gaps in the shadow edge. Fig. 4d illustrates a typical example of a recovered shadow edge map $q_s(x,y)$. It is clear that even after the processing described, the definition of the shadow edge is imperfect: there are a number of spurious edges not removed. However, this map is sufficiently accurate to

allow recovery of the shadow-free image shown in Fig. 4e based on the integration procedure described above.

5 DISCUSSION

We have introduced three different shadow-free image representations in this paper: A 1D invariant derived from first principles based on simple constraints on lighting and cameras, a 2D chromaticity representation which is equivalent to the 1D representation but with some color information retained and, finally, a 3D full color image. Fig. 5 shows some examples of these different representations for a number of different images. In each example, all three representations are shadow-free. The procedure for deriving each of the three representations is automatic, but there are a number of parameters which must be specified. In all cases, we need to determine the direction of illumination change (the vector $\underline{e}$ discussed in Section 2). This direction can be found either by following the calibration procedure outlined in Section 2 above or, as has recently been proposed [35] by using a procedure which determines the illuminant direction from a single image of a scene having shadow and nonshadow regions. The examples in Fig. 5 were obtained based on the latter calibration procedure. In addition to the calibration step, in the 2D representation we also have a parameter to control how much light is put back in to the image. We used the procedure described in Section 3 to determine this parameter for the examples in Fig. 5.

Recovering the 3D representation is more complex and there are a number of free parameters in the recovery algorithm. As a first step, the original full-color images were processed using the mean shift algorithm which has two free parameters: a *spatial bandwidth* parameter which was set to 16 (corresponding to a 17×17 spatial window), and a *range* parameter which was set to 20. The process of comparing the two edge maps is controlled by three thresholds: τ_1, τ_2, and τ_3. τ_1 and τ_2 relate to the edge strengths in the original and the invariant image, respectively. We chose values of $\tau_1 = 0.4$ and $\tau_2 = 0.1$ after the gradient magnitudes have been scaled to a range $[0,1]$. Our choice for these parameters is determined by the hysteresis step in the Canny edge detection process. τ_3 controls the difference in the orientation between edges in the original image and those in the invariant. Edges are classified into one of eight possible orientations, but by taking advantage of symmetry we need consider only four of them. So, τ_3 is set equal to $\pi/4$. These parameters were fixed for all images in Fig. 5 and, although the recovered shadow edge is not always perfect, the resulting shadow-free image is, in all cases, of good quality. We note however, that the algorithm in its current form will not deliver perfect shadow-free images in all cases. In particular, images with complex shadows, or diffuse shadows with poorly defined edges will likely cause problems for the algorithm. However, the current algorithm is robust when shadow edges are clear, and we are currently investigating ways to improve the algorithm's performance on the more difficult cases. In addition, it is possible for the method to misclassify some edges in the original image as shadow edges. For example, if two adjacent surfaces differ in intensity, an edge detector will find an edge at the border of these two surfaces. However, in the 1D invariant image intensity differences are absent, and so no edge will be found in this case. Thus, the edge between the two surfaces will

wrongly be classified as a shadow edge. Indeed, the fifth example in Fig. 5 exhibits such behavior: The boundary between the painted white line on the road surface, and the road surface itself, is not fully recovered, because the two surfaces (paint and road) differ mainly in intensity. A similar problem can arise if adjacent surfaces are related by a color change in the direction in which illumination changes. Here again, an edge will be found in the original image, but will be absent from the invariant images. The examples in Fig. 5 (and the many other images we have processed) suggest that such problems arise only infrequently in practice. However, in future work, we intend to investigate ways to overcome these problems.

In summary, we conclude that the approach to shadow removal proposed in this paper yields very good performance. In all three cases (1D, 2D, and 3D), the recovered images are of a good quality and we envisage that they will be of practical use in a variety of visual tasks such as segmentation, image retrieval, and tracking. As well, the method raises the possibility of enhancing commercial photography such as portraiture.

REFERENCES

[1] G.J. Klinker, S.A. Shafer, and T. Kanade, "A Physical Approach to Color Image Understanding," *Int'l J. Computer Vision,* vol. 4, pp. 7-38, 1990.

[2] M.J. Swain and D.H. Ballard, "Color Indexing," *Int'l J. Computer Vision,* vol. 7, no. 1, pp. 11-32, 1991.

[3] H. Jiang and M. Drew, "Tracking Objects with Shadows," *Proc. Int'l Conf. Multimedia and Expo,* 2003.

[4] H. Barrow and J. Tenenbaum, "Recovering Intrinsic Scene Characteristics from Images," *Computer Vision Systems,* A. Hanson and E. Riseman, eds., pp. 3-26, Academic Press, 1978.

[5] Y. Weiss, "Deriving Intrinsic Images from Image Sequences," *Proc. Int'l Conf. Computer Vision,* pp. II: 68-75, 2001.

[6] E.H. Land, "Lightness and Retinex Theory," *J. Optical Soc. Am., A,* vol. 61, pp. 1-11, 1971.

[7] B.K. Horn, "Determining Lightness from an Image," *Computer Graphics and Image Processing,* vol. 3, pp. 277-299, 1974.

[8] A. Blake, "On Lightness Computation in the Mondrian World," *Proc. Wenner-Gren Conf. Central and Peripheral Mechanisms in Colour Vision,* pp. 45-59, 1983.

[9] G. Brelstaff and A. Blake, "Computing Lightness," *Pattern Recognition Letters,* vol. 5, pp. 129-138, 1987.

[10] A. Hurlbert, "Formal Connections between Lightness Algorithms," *J. Optical Soc. Am., A,* vol. 3, no. 10, pp. 1684-1693, 1986.

[11] E.H. Land, "The Retinex Theory of Color Vision," *Scientific Am.,* pp. 108-129, 1977.

[12] L.T. Maloney and B.A. Wandell, "Color Constancy: A Method for Recovering Surface Spectral Reflectance," *J. Optical Soc. Am., A,* vol. 3, no. 1, pp. 29-33, 1986.

[13] G. Buchsbaum, "A Spatial Processor Model for Object Colour Perception," *J. Franklin Inst.,* vol. 310, pp. 1-26, 1980.

[14] D. Forsyth, "A Novel Algorithm for Colour Constancy," *Int'l J. Computer Vision,* vol. 5, no. 1, pp. 5-36, 1990.

[15] G.D. Finlayson, S.D. Hordley, and P.M. Hubel, "Color by Correlation: A Simple, Unifying Framework for Color Constancy," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 11, pp. 1209-1221, Nov. 2001.

[16] D.H. Brainard and W.T. Freeman, "Bayesian Method for Recovering Surface and Illuminant Properties from Photosensor Responses," *Proc. IS&T/SPIE Symp. Electronic Imaging Science and Technology,* vol. 2179, pp. 364-376, 1994.

[17] V. Cardei, "A Neural Network Approach to Colour Constancy," PhD dissertation, School of Computing Science, Simon Fraser Univ., 2000.

[18] K. Barnard, "Practical Colour Constancy," PhD dissertation, School of Computing Science, Simon Fraser Univ., 2000.

[19] S. Tominaga and B.A. Wandell, "Standard Surface-Reflectance Model and Illuminant Estimation," *J. Optical Soc. Am. , A,* vol. 6, no. 4, pp. 576-584, 1996.

[20] H.-C. Lee, "Method for Computing Scene-Illuminant Chromaticity from Specular Highlights," *Color,* G.E. Healey, S.A. Shafer, and L.B. Wolff, eds., pp. 340-347, Boston: Jones and Bartlett, 1992.

[21] Th. Gevers and H.M.G. Stokman, "Classifying Color Transitions into Shadow-Geometry, Illumination Highlight or Material Edges," *Proc. Int'l Conf. Image Processing,* pp. 521-525, 2000.

[22] B.V. Funt and G.D. Finlayson, "Color Constant Color Indexing," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 5, pp. 522-529, May 1995.

[23] G. Finlayson, S. Chatterjee, and B. Funt, "Color Angular Indexing," *Proc. Fourth European Conf. Computer Vision,* vol. II, pp. 16-27, 1996.

[24] T. Gevers and A. Smeulders, "Color Based Object Recognition," *Pattern Recognition,* vol. 32, pp. 453-464, 1999.

[25] M. Stricker and M. Orengo, "Similarity of Color Images," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases III,* vol. 2420, pp. 381-392, 1995.

[26] D. Berwick and S.W. Lee, "A Chromaticity Space for Specularity, Illumination Color- and Illumination Pose-Invariant 3-D Object Recognition," *Proc. Sixth Int'l Conf. Computer Vision,* 1998.

[27] G. Finlayson, B. Schiele, and J. Crowley, "Comprehensive Colour Image Normalization," *Proc. European Conf. Computer Vision,* pp. 475-490, 1998.

[28] G. Finlayson and S. Hordley, "Color Constancy at a Pixel," *J. Optical Soc. Am. A,* vol. 18, no. 2, pp. 253-264, UK Patent application no. 0000682.5, under review, British Patent Office, 2001.

[29] J.A. Marchant and C.M. Onyango, "Shadow Invariant Classification for Scenes Illuminated by Daylight," *J. Optical Soc. Am., A,* vol. 17, no. 12, pp. 1952-1961, 2000.

[30] R. Hunt, *The Reproduction of Colour,* fifth ed. Fountain Press, 1995.

[31] J.L.N.J. Romero, J. Hernandez-Andres, and E. Valero, "Testing Spectral Sensitivity of Sensors for Color Invariant at a Pixel," *Proc. Second Computer Graphics, Imaging, and Vision Conf.,* Apr. 2004.

[32] G. Finlayson and M. Drew, "4-Sensor Camera Calibration for Image Representation Invariant to Shading, Shadows, Lighting and Specularities," *Proc. Int'l Conf. Computer Vision,* pp. 473-480, 2001.

[33] B.K. Horn, *Robot Vision.* MIT Press, 1986.

[34] G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulas,* second ed. New York: Wiley, 1982.

[35] G.D. Finlayson, M.S. Drew, and C. Lu, "Intrinsic Images by Entropy Minimisation," *Proc. European Conf. Computer Vision,* 2004.

[36] M.H. Brill and G. Finlayson, "Illuminant Invariance from a Single Reflected Light," *Color Research and Application,* vol. 27 pp. 45-48, 2002.

[37] M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta, "A Standard Default Color Space for the Internet—SRGB," http://www.color.org/sRGB.html, 1996.

[38] B. Funt, K. Barnard, and L. Martin, "Is Machine Colour Constancy Good Enough?" *Proc. Fifth European Conf. Computer Vision,* pp. 455-459, June 1998.

[39] M. Drew, C. Chen, S. Hordley, and G. Finlayson, "Sensor Transforms for Invariant Image Enhancement," *Proc. 10th Color Imaging Conf.: Color, Science, Systems, and Applications,* pp. 325-329, 2002.

[40] G.D. Finlayson, M.S. Drew, and B.V. Funt, "Spectral Sharpening: Sensor Transformations for Improved Color Constancy," *J. Optical Soc. Am., A,* vol. 11, no. 5, pp. 1553-1563, 1994.

[41] M.S. Drew, G.D. Finlayson, and S.D. Hordley, "Recovery of Chromaticity Image Free from Shadows via Illumination Invariance," *Proc. Workshop Color and Photometric Methods in Computer Vision,* pp. 32-39, 2003.

[42] B. Funt, M. Drew, and M. Brockington, "Recovering Shading from Color Images," *Proc. Second European Conf. Computer Vision,* G. Sandini, ed., pp. 124-132, May 1992.

[43] R. Frankot and R. Chellappa, "A Method for Enforcing Integrability in Shape from Shading Algorithms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 10, pp. 439-451, 1988.

[44] R. Jain, R. Kasturi, and B. Schunck, *Machine Vision.* McGraw-Hill, 1995.

[45] D. Comaniciu and P. Meer, "Mean Shift Analysis and Applications," *Proc. Seventh Int'l Conf. Computer Vision,* pp. 1197-1203, 1999.

[46] J. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 8, pp. 679-698, 1986.

Graham D. Finlayson received the BSc degree in computer science from the University of Strathclyde, Glasgow, Scotland, in 1989. He then pursued his graduate education at Simon Fraser University, Vancouver, Canada, where he was awarded the MSc and PhD degrees in 1992 and 1995, respectively. From August 1995 until September 1997, He was a lecturer in computer science at the University of York, United Kingdom, and from October 1997 until August 1999, he was a reader in color imaging at the Colour and Imaging Institute, University of Derby, United Kingdom. In September 1999, he was appointed a professor in the School of Computing Sciences, University of East Anglia, Norwich, United Kingdom. He is a member of the IEEE.

Steven D. Hordley received the BSc degree in mathematics from the University of Manchester, United Kingdom, in 1992 and the MSc degree in image processing from Cranfield Institute of Technology, United Kingdom in 1996. Between 1996 and 1999, he studied at the University of York, United Kingdom, and the University of Derby, United Kingdom, for the PhD degree in color science. In September 2000, he was appointed a lecturer in the School of Computing Sciences at the University of East Anglia, Norwich, United Kingdom.

Cheng Lu received his undergraduate education in computer applications from the University of Science and Technology in Beijing, China, and he received the master's degree in the School of Computing Science at Simon Fraser University. He is a PhD candidate in the School of Computing Science at Simon Fraser University, Vancouver, Canada. His research interests include image and video indexing, computer vision, and color processing for digital photography.

Mark S. Drew is an associate professor in the School of Computing Science at Simon Fraser University in Vancouver, Canada. His background education is in engineering science, mathematics, and physics. His interests lie in the fields of multimedia, computer vision, image processing, color, photorealistic computer graphics, and visualization. He has published more than 80 refereed papers in journals and conference proceedings. Dr. Drew is the holder of a US Patent in digital color processing. He is a member of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.