Removing Shadows from Images

Graham D. Finlayson[1], Steven D. Hordley[1], and Mark S. Drew[2]

[1] School of Information Systems
The University of East Anglia
Norwich
England NR4 7TJ
{graham, steve }@sys.uea.ac.uk
[2] School of Computing Science
Simon Fraser University
Vancouver, British Columbia
Canada V5A 1S6
mark@cs.sfu.ca

Abstract. Illumination conditions cause problems for many computer vision algorithms. In particular, shadows in an image can cause segmentation, tracking, or recognition algorithms to fail. In this paper we propose a method to process a 3-band colour image to locate, and subsequently remove shadows. The result is a 3-band colour image which contains all the original salient information in the image, except that the shadows are gone.

We use the method set out in [1] to derive a 1-d illumination invariant shadow-free image. We then use this invariant image together with the original image to locate shadow edges. By setting these shadow edges to zero in an edge representation of the original image, and by subsequently re-integrating this edge representation by a method paralleling lightness recovery, we are able to arrive at our sought after full colour, shadow free image. Preliminary results reported in the paper show that the method is effective.

A caveat for the application of the method is that we must have a calibrated camera. We show in this paper that a good calibration can be achieved simply by recording a sequence of images of a fixed outdoor scene over the course of a day. After calibration, only a single image is required for shadow removal. It is shown that the resulting calibration is close to that achievable using measurements of the camera's sensitivity functions.

Keywords. Texture, shading, & colour, shadow removal, lightness recovery, illuminant invariance.

1 Introduction

Illumination conditions can confound many algorithms in vision. For example, changes in the colour or intensity of the illumination in a scene can cause problems for algorithms which aim to segment the image, or to track or recognise,

A. Heyden et al. (Eds.): ECCV 2002, LNCS 2353, pp. 823–836, 2002.

objects in the scene. One illumination effect which can cause particular problems for these algorithms is that of shadows. The disambiguation of edges due to shadows and those due to material changes is a difficult problem and has a long history in computer vision research [2]. In addition, the investigation of shadows as cues for image understanding has an even older lineage [3]. Recently, the importance of understanding shadows has come to the fore in digital photography applications including colour correction[4] and dynamic range compression[5].

One possible solution to the confounding problems of shadows is to derive images which are shadow free: that is to process images such that the shadows are removed whilst retaining all other salient information within the image. Recently, a study [6] aimed at lightness computation [7] set out a clever method to attenuate the effect of shadows in an image. Unfortunately however, this method requires not just a single image, but rather a sequence of images, captured with a stationary camera over a period of time such that the illumination in the scene (specifically the position of the shadows) changes considerably.

The example used by the author was a sequence of grey-scale images of a fixed outdoor scene, captured over the course of a day. Assuming that material changes are constant in the scene and that shadows move as the day progresses, it follows that the *median* edge map (for the sequence) can be used to determine material edges (shadow edges since they move are transitory and so do not effect the median). Given the material edge-map it is possible to create an intrinsic image that depends only on reflectance. This reflectance map might then be compared against the original sequence and an intrinsic illuminant map for each image recovered. While this method works well a major limitation of the approach is that the illumination independent (and shadow free) image can only be derived from a sequence of time varying images.

In this paper we propose a method for removing shadows from images which in contrast to this previous work requires only a single image. The approach is founded on an application of a recently developed method for eliminating from an image the colour and intensity of the prevailing illumination [1,8]. The method works by finding a single scalar function of image an RGB that is invariant to changes in light colour and intensity i.e. it is a 1-dimensional invariant image that depends only on reflectance. Because a shadow edge is evidence of a change in only the colour and intensity of the incident light, shadows are removed in the invariant image. Importantly, and in contrast to antecedent invariant calculations, the scalar function operates at a pixel and so is not confounded by features such as occluding edges which can affect invariants calculated over a region of an image.

Fig. 1(a) shows a 3-band colour image, taken in outdoor lighting using an experimental Hewlett-Packard HP912 digital still camera modified to produce raw output. For purposes of display, the raw camera RGB values have been converted to the standard sRGB[9] colour space. The image shows two people photographing a lawn — their shadows are prominent. It is important to realise that the shadows in this image, and shadows in outdoor scenes in general, represent a change in both intensity and colour of the prevailing illumination. In the image

Fig. 1. (a): Original image. (b): Grey-scale illuminant-invariant image. (c): Grey-scale non illuminant-invariant image. (d): Edge map for invariant image. (e): Edge map for non-invariant image. (f): Recovered, shadow-free image.

in Figure 1, the region of the footpath which is not in shadow is illuminated by a mixture of sunlight and skylight, whereas the shadowed region is lit only by skylight. Thus, there is a change in the effective correlated colour temperature of the illumination in the two regions. Fig. 1(b) shows the invariant, reflectance only, image. There are two features of this image which are worthy of comment. First is the fact that the shadows present in the original image are effectively removed in the invariant image. Second, it is important to notice that the invariant image is grey scale — in removing the effect of the scene illumination at each pixel in the image, information is lost. Shadows are removed but we have moved from a rich colourful RGB image to a shadowless grey-scale representation.

In this paper we address this problem and set out to recover an RGB colour image from which the shadows are removed. To achieve this we focus on the derivative images of both the original image and the illumination invariant image. More specifically we look at the differences in the edge maps of the two images. We reason that material edges should occur in both RGB and invariant images. In effect, we can use the invariant edge map as a mask to locate non-shadow edges in the RGB edge map. Re-integrating the material-edge-only map should result in a full colour image where shadows are removed.

Before proceeding it is useful to make a couple of caveats. First, that the generation of an invariant image requires calibration. In [1] the calibration was performed off-line using measurements of a camera's sensor sensitivity functions. Here we investigate what can be done when these measurements are unavailable. In particular we show that the calibration step can be achieved using a sequence of images of a fixed scene, taken under a variety of daylight conditions. This offers the possibility that both the method presented here, and indeed, the method to

derive the invariant grey-scale image can be used with uncalibrated cameras in, for example, tracking or surveillance applications. The second caveat we make is about the nature of cast shadows. Implicit in our technique is the idea that shadow edges appear in the edge-maps of RGB images. For very diffuse shadow boundaries this may in fact not be the case. However, in this worst case when shadows are not found, the method will basically reintegrate to the original image. However, experiments demonstrate that our method strongly attenuates shadows in a variety of images even when shadow boundaries are not sharp. In all images we have tried shadows are less prominent after processing than before.

In Section 2, we set out the method used for forming an invariant image, and in Section 3 we detail how we use this invariant image to remove shadows from a colour image. Section 4 considers the issue of camera calibration and Section 5 shows further results and states some conclusions.

2 Invariant Image Formation

Suppose we image a set of coloured surfaces under a particular Planckian light, in a controlled light box, say. If surfaces are Lambertian, then for each pixel the log of the chromaticity $\{r, g\}$ appears as a dot in a 2-d plot; for flat or curved surfaces every pixel in each patch is approximately collapsed into the same dot. Fig. 2(b) illustrates the log-chromaticities for the 24 surfaces of the Macbeth Color Checker Chart shown in Fig. 2(a). The plot shows 19 distinct clusters of points — each cluster corresponds to chromaticities from a single patch, there are 19 clusters rather than 24 since the patches in the last row of the chart are all neutral in colour and so have the same chromaticity.

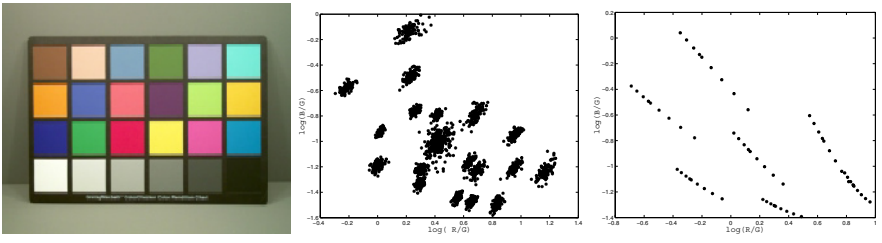

Fig. 2. (a): Macbeth Color Checker Chart image under a Planckian light with an HP912 Digital Still Camera. (b): Log-chromaticities of the 24 patches of the imaged chart. (c): Chromaticities for 6 different patches, imaged under a set of different Planckian illuminants.

Now if sensors are fairly narrow-band (and if they are not they can be made more so via a spectral sharpening procedure [10]) then, for Planckian lights, changing the lighting colour simply amounts to moving the log-chromaticity colour vector along a direction multiplying a term which depends on the illuminant temperature T and which is independent of the magnitude and direction

of the lighting. Fig. 2(c) illustrates this for 7 patches of the colour chart shown in Fig. 2(a); the figure shows the same 7 patches imaged under a range of different Planckian illuminants. Here, because the camera sensors are not exactly narrow band, the linear shift of chromaticities is only approximate. As we will see later though, this approximate linearity is sufficient for our needs. Assuming that the change with illumination is indeed linear, projecting colours perpendicular to this direction of change produces a 1-d greyscale image that is invariant to lighting change.

This direction is in principle different for each camera, and thus must be recovered from a calibration. While lighting may in fact not be truly Planckian, most lights, and certainly all daylights, fall very near to the Planckian locus in a chromaticity plot, and in practice the Planckian assumption is not crucial.

To see how this linear behaviour with lighting change is derived in the ideal case we recapitulate the invariant image calculation here. Consider the RGB colour formed at a pixel from illumination with spectral power distribution $E(\lambda)$, impinging on a surface with surface spectral reflectance function $S(\lambda)$. If the three camera sensor sensitivity functions form a set $\boldsymbol{R}(\lambda)$ then the RGB colour $\boldsymbol{\rho}$ at any pixel results from an integral over the visible wavelengths

$$\rho_k \;=\; \int E(\lambda)S(\lambda)Q_k(\lambda)d\lambda \;,\; k = R, G, B \;. \tag{1}$$

If camera sensitivity $Q_k(\lambda)$ is exactly a Dirac delta function $Q_k(\lambda) = q_k\delta(\lambda - \lambda_k)$, with q_k the strength of the sensor $q_k = Q_k(\lambda_k)$, then Eq. (1) reduces to the simpler form

$$\rho_k \;=\; E(\lambda_k)S(\lambda_k)q_k \;. \tag{2}$$

Now suppose lighting can be approximated by Planck's law.

$$E(\lambda, T) = \; I\, c_1\lambda^{-5}\left(e^{\frac{c_2}{T\lambda}} - 1\right)^{-1} \tag{3}$$

Constants c_1 and c_2 equal 3.74183×10^{-16} Wm^2 and 1.4388×10^{-2} mK, respectively. The variable I controls the intensity of the incident light. For illuminants in the temperature range 2500K to 10000K (reddish through whitish to bluish) the term $e^{\frac{c_2}{T\lambda}} >> 1$ and Wien's approximation can be used:

$$E(\lambda, T) \simeq \; I\, c_1\lambda^{-5}e^{-\frac{c_2}{T\lambda}} \;. \tag{4}$$

Both the above equations generate functions which are very smooth functions of wavelengths. In contrast, daylights have many high frequency components. The approximations set forth above make sense only because we integrate over all wavelengths to form RGB. From the point of view of most cameras daylight spectra are effectively smooth (because camera sensitivities cannot see high frequency spectral components).

Returning to the narrow-band sensor response equation, RGB colour ρ_k, $k = 1 \ldots 3$ is simply given by

$$\rho_k \;=\; Ic_1\lambda_k^{-5}e^{-\frac{c_2}{T\lambda_k}}S(\lambda_k)q_k \;. \tag{5}$$

Now suppose we first form band-ratio chromaticities from colour values $\boldsymbol{\rho}$ given by eq. (1):

$$r_k \;=\; \rho_k/\rho_p \tag{6}$$

where p is one of the channels and k indexes over the remaining responses. In our experiments $p = 2$ (i.e. we divide by green) and so we calculated R/G and B/G. As in all chromaticity operations, we effectively remove intensity information. The intensity variable I is absent from the chromaticity coordinates. To isolate the temperature term (so we might later remove it) in (4) we take the log of (6).

$$r'_k \;\equiv\; \log(r_k) \;=\; \log(s_k/s_p) + (e_k - e_p)/T\,, \tag{7}$$

where we define $s_k = c_1\lambda_k^{-5}S(\lambda_k)q_k$ and $e_k = -c_2/\lambda_k$. As temperature changes, 2-vectors $r'_k, k = R, B$, will form a straight line in 2-d colour space. Equation (7) is the vector equation for a line. Calibration then amounts to determining the 2-vector direction $(e_k - e_p)$ in the space of logs of ratios. We discuss the practical aspects of this calibration in more detail in Section 4.

The invariant image is that formed by projecting 2-d colours into the direction $\boldsymbol{e}^{\perp}$ orthogonal to the vector $(e_k - e_p)$. The result of this projection is a single scalar which we then code as a grey-scale value. Here, and henceforth, the grey-scale invariant is defined as:

$$gs = c_1 r'_R - c_2 r'_B \tag{8}$$

Where c_1 and c_2 are constants such that the vector $[c_1\ c_2]$ is in the direction $[e_B - e_R]$ (it is orthogonal to the lighting direction). The grey-scale invariant image is denoted $gs(x, y)$.

Experiments have shown that images of the same scene containing objects of different colours illuminated by any complex lighting field (including lights of different colours and intensities) will map to the same invariant image. Most importantly for this paper, shadows which occur when there is a change in light but not surface will disappear in the invariant image.

Of course by definition we expect the illuminant invariance properties. We have carefully shown by considering the physics of image formation how light intensity and temperature are cancelled out. But, we draw the reader's attention to a possible problem. Specifically, the invariant is designed to work for Planckian lights. Additive combinations of Planckians (which might result indoors when there is mixed light from a Tungsten source and outdoor illumination through a window) is non-Planckian. However, because the Planckian locus is a very shallow crescent shape, additive combinations of light tend to fall close to the locus. Experimentally, the invariant image factors out the light even for additive combinations of Planckian illuminants[1].

3 Method for Shadow Removal

Our method for shadow removal has its roots in methods of lightness recovery. Lightness is a term usually used to mean that part of a photometric signal that

depends only on reflectance. An RGB image is input and two intrinsic images are output: one based on reflectance (the lightness intrinsic image) and the other based on illumination. Lightness computation proceeds by making assumptions about the world. In particular, it is usually assumed that illumination varies slowly across an image. In contrast changes in reflectance are rapid. It follows then that by thresholding a derivative image to remove small derivatives, slow changes (due to, by assumption, illumination) can be removed. Integrating the thresholded derivative image results in the lightness intrinsic image. Clearly, we wish to adopt a similar strategy here. However, our assumptions must at the outset be different. Shadows are evidence of a sharp change in illumination and this will lead us to a different kind of thresholding operation.

Let us begin by recapitulating the standard lightness recovery algorithm. The algorithm works by finding the intrinsic image in each of the separate R-, G- and B-channels separately. Let us use the notation $\rho(x, y)$ to represent one of the three channel images. We are going to use thresholded derivatives to remove the effect of illumination. We observe in Equation (2) that sensor response is a multiplication of light and surface. The gradient differential operator takes differences of adjacent image pixels. Assuming locally constant illumination, the difference between log colour responses removes the effect of the illumination. Denoting the log channel response as $\rho'(x, y)$ we write the gradient (the vector of the x and y derivatives as:

$$\text{gradient of channel response } \|\nabla\rho'(x, y)\| \tag{9}$$

Given the log-image edge map $\nabla\rho'(x, y)$ we can define a threshold operator T to remove effects due to illumination. In the original lightness formulation[11], T thresholds out gradients of small magnitude:

$$T(\nabla\rho'(x, y)) = \begin{cases} 0 & \text{if } \|\nabla\rho'(x, y)\| < \text{threshold} \\ \nabla\rho'(x, y) & otherwise \end{cases} \tag{10}$$

Here our goal is not to remove illumination per se (we are not solving for colour constancy) but rather we wish only to remove shadows. In fact we actually want to keep the illuminant field and re-render the scene as if it were captured under the same single non-shadow illuminant. To do this we must factor out changes in the gradient at shadow edges. Let us denote

$$\text{gradient of greyscale invariant image } \|\nabla gs(x, y)\| \tag{11}$$

Since the invariant image is a function of reflectance, shadow edges must disappear. Thus, we can remove shadows in the gradient of the log response image using the following threshold function $S()$:

$$S(\nabla\rho'(x, y), gs(x, y)) = \begin{cases} 0 & \text{if } \|\nabla\rho'(x, y)\| > \text{thresh}_1 \text{ and } \|\nabla gs(x, y)\| < \text{thresh}_2 \\ \nabla\rho'(x, y) & otherwise \end{cases} \tag{12}$$

That is if the magnitude of the gradient in the invariant image is close to zero where the gradient of the log response is larger than zero then this is evidence

of a shadow edge. At such places the gradient of the log response image is set to zero indicating that there is no change at this point (which is true for the underlying reflectance). We point out that a similar 'rule-based' approach to determining the semantics of edges (e.g. highlight edges vs material edges) has been proposed by Gevers and Stockman[12]. Though, that approach fails to account for shadows.

After thresholding we should now have a gradient image 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 the gradient in order to recover a log response image which does not have shadows. To do this we must solve a Poisson equation of the form:

$$\nabla^2 q'(x,y) \;=\; \nabla \cdot S(\nabla \rho'(x,y), gs(x,y)) \tag{13}$$

On the left hand-side of Equation (13) we have the Laplacian $(\frac{d^2\rho'}{dx^2} + \frac{d^2\rho'}{dy^2})$ of the image we wish to recover. On the right-hand side we have the Laplacian of the input image where the Laplacian is computed in two steps. First the thresholded (shadow free) gradient is calculated (12). Second, the Laplacian is calculated from the gradient. However, the Laplacian by itself is not sufficient to allow the Poisson equation to be solved (the Laplacian is not defined at the boundaries of an image). Rather we must make some assumption about what is happening at the image boundary. Here we assume Neumann boundary conditions: the derivative at the boundary is zero. Subject to this constraint we can recover $q(x,y)$ uniquely up to an unknown additive constant. Exponentiating q', we arrive at the reconstructed greyscale image (up to an unknown multiplicative constant). Solving (13) for each of the three colour bands results in a full colour image where the shadows are removed.

However, to obtain "realistic" image colours we must deal with the unknown multiplicative constants. To fix these constants, we consider the top 1-percentile of pixels in the recovered image in each band. We use the average of these pixels as a measure of the maximum response in the images. Mapping the maximum value of the RGB image to (1 1 1) effectively removes the unknown constants. Moreover, adjusting the white-point of an image in this way is a simple way of discounting the colour of the (in this case non-shadow) illuminant.

Note that reasonably accurate shadow edges are required for the method to work well. Obtaining them depends on how we define the thresholding operator S. We have experimented with a variety of methods for defining S and what works best depends to some extent on image content. The method outlined here then is not necessarily optimal and the results should thus be treated as a convincing proof in principle.

The difficulty in defining S is illustrated by the images in Figs. 3(a,b). Fig. 3(a) shows the edge map for the intensity image $(\frac{1}{3}(R+G+B))$ of Fig. 1(a) while that in Fig. 3(b) is the edge map for the corresponding illuminant-invariant image. The edges in these images are calculated thus:

$$\left(\|\rho \star \{-1,\ 0,\ 1\}^t\|^2 + \|\rho \star \{-1,\ 0,\ 1\}^t\|^2\right)^{1/2} \tag{14}$$

where ρ represents the relevant image and $\star$ denotes convolution. Simple edge operators of this kind produce non-zero values at more locations than those at which there are true edges. In the examples in Fig. 3 the edges of the road and (in the case of the intensity image) the shadow are clear, but so too are many edges due to the texture of the imaged surfaces and also noise in the images. Obtaining the edges in which we are interested from these edge maps is non-trivial as evidenced by the large literature on edge detection (see [13] for a review).

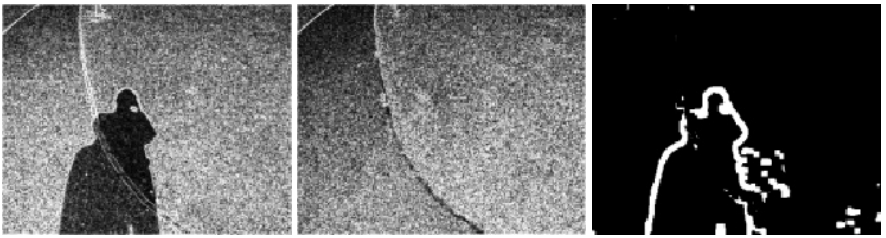

Fig. 3. (a): Edges in the intensity image. (b): Edges in the invariant image. (c): Final recovered shadow edge.

One simple approach to determining the true edges is to threshold the edge maps such that weak edges are set to zero. We found however that this still does not produce edge maps which are clean enough for our purposes. Instead we have employed more sophisticated edge detection algorithms such as that proposed by Canny [14] and the SUSAN algorithm proposed in [15]. We have achieved some success with both these methods; the results obtained here use the SUSAN edge detector. In this algorithm the image to be edge detected is first smoothed by convolution with a kernel function. Next there is an edge detection step after which the resulting edges are thresholded to produce strong edges.

Employing this approach we find shadow edges by looking for edges which are in the log red-, green- or blue- response image but are not in the invariant image. As a final step we employ a morphological operation (specifically an opening) on the binary edge map to "thicken" the shadow edges. This thickening of the edges was found to be necessary to ensure that the shadows are properly removed in the re-integration step. The resulting shadow edge (shown in Fig. 3(c)) is used to guide the operator S in the thresholding step of the recovery algorithm outlined above. Even after this processing 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 1(f).

Even after outlining the difficulties presented above, we have found that a very simple workflow produces good results for most images that have pronounced shadow edges. Moreover, the worst case performance for our algorithm is when shadow edges are not recovered. In this case the reintegration returns im-

ages which are closer than we would like to the original image (i.e. with shadows still present).

4 Uncalibrated Cameras

In order for the method outlined above to succeed, the invariant log lighting direction $e_k - e_p$ must be known. Obtaining this knowledge amounts to performing a calibration of the camera. Camera calibration is simple if its sensor sensitivity functions are known[1]. The calibration is more difficult when the camera sensors are unknown.

Here we investigate how it is possible to calibrate a camera given only a set of images of the same scene taken at different times of the day. We adopt the following reasoning: over the course of the day, the height of the sun, and general weather conditions change and so the effective illumination must change too. It follows then that the plotted log-chromaticities for a single pixel must lie on a line. Moreover, as discussed earlier the orientation of the lines discovered for all pixels will be the same. The orientation that best describes all lines defines the illuminant direction.

To test whether this reasoning is sufficient to obtain an accurate calibration of the camera we captured 14 images of the scene shown in Fig. 4(a) at different times throughout the day. Fig. 4(b) shows the chromaticities of the 24 colour patches of the Macbeth Color Checker which was placed in the scene. The change of log-chromaticity with time (and hence with a change of daylight illumination) is approximately linear. We used this data to derive the invariant direction for the camera and the results are illustrated in Fig. 4(c). Also shown in this figure are the recovered vectors for a calibration based on knowledge of the camera sensors (solid line) and that based on the sequence of daylight images (dashed line). It is clear that the directions are very close — the angle between the two lines is 2.4^o. Figs. 4(d) and (e) shows the invariant images derived using the two different calibrations. The resulting invariant images are very similar for identification of edge content, and, importantly, the shadows are greatly attenuated in both cases.

This simple experiment suggests that an accurate camera calibration can be obtained without knowledge of the camera's sensors, thus broadening the potential application of both the method reported here and the original method set out in [1]. A final step to generalise (and automate) the calibration procedure would be to derive the invariant based on the whole images rather than just the patches from the test chart as was done here. To do this, however, it is necessary for the images to be registered — such a set of images was unavailable at this time. However, given that registered images are available, and that the scene contains a range of different colours, good calibration should be possible.

Finally, an experimental calibration has two main advantages over a calibration based on known spectral sensitivities. First, RGBs in camera are often gamma corrected (R, G and B are raised to some power γ) prior to storage. Indeed most images viewed on a computer monitor are (roughly) the square root of

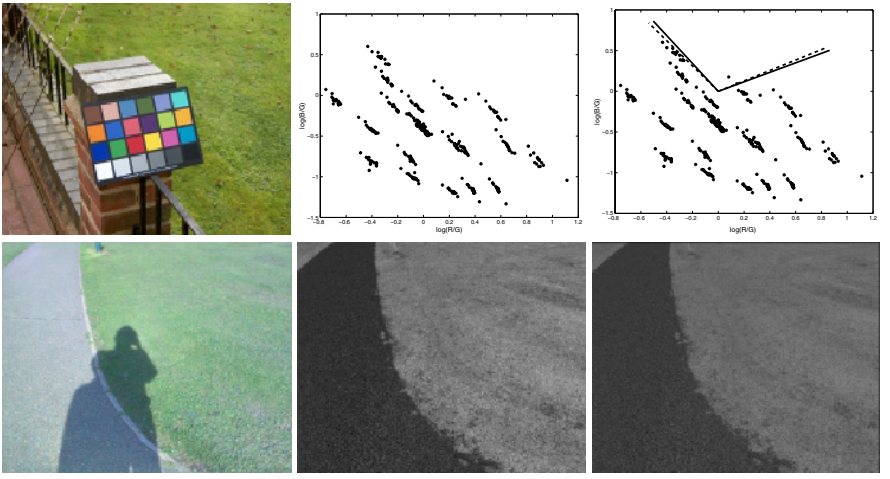

Fig. 4. (a): The scene used for the camera calibration. (b): Chromaticities of the 24 colour patches for the set of 14 images. (c): Invariant directions recovered using camera sensors (solid line) and the image sequence (dashed line). (d): Image taken with the same camera. (e): Invariant image based on a calibration using camera sensors. (f): Invariant image based on a calibration using the image sequence.

the linear signal. This is because monitors have a squared transfer function and so the squaring of the monitor cancels the square root of the camera resulting in the required linear signal. However, for the calibration set forth above, the gamma is simply an unknown multiplier in the recovered parameter and does not change the direction of the lighting direction.

For consider the effect of a gamma correction on the invariant image calculation:

$$\begin{aligned} \rho_k &\rightarrow \gamma(\rho_k) \\ r'_k &\rightarrow \gamma \log(s_k/s_p) + \gamma(e_k - e_p)/T \,. \end{aligned} \tag{15}$$

Clearly, we simply deduce a different vector $e_k - e_p$ than that we would have calculated using linear signals; but the effect on images is the same: $\boldsymbol{e}^{\perp}$ produces an invariant image.

The second advantage of an experimental calibration is that the camera sensitivity may change as a function of time and temperature. A continuous adaptive calibration would support shadow removal even if the current state of the camera differed from manufacturer specifications.

5 Results

Fig. 5 shows some further results of the shadow removal method set out here. In all cases the shadows are removed quite effectively. There are, however, a number of artifacts introduced into the images. These artifacts are due to the

fact that the determination of the shadow-edge is imperfect, resulting in edge pixels being set to zero when they shouldn't be. Pragmatically then the results of this paper do not deliver photo-quality images. However, the technique might usefully subserve tasks such as tracking[16,17] or object recognition[18,19] where shadows are known to cause problems.

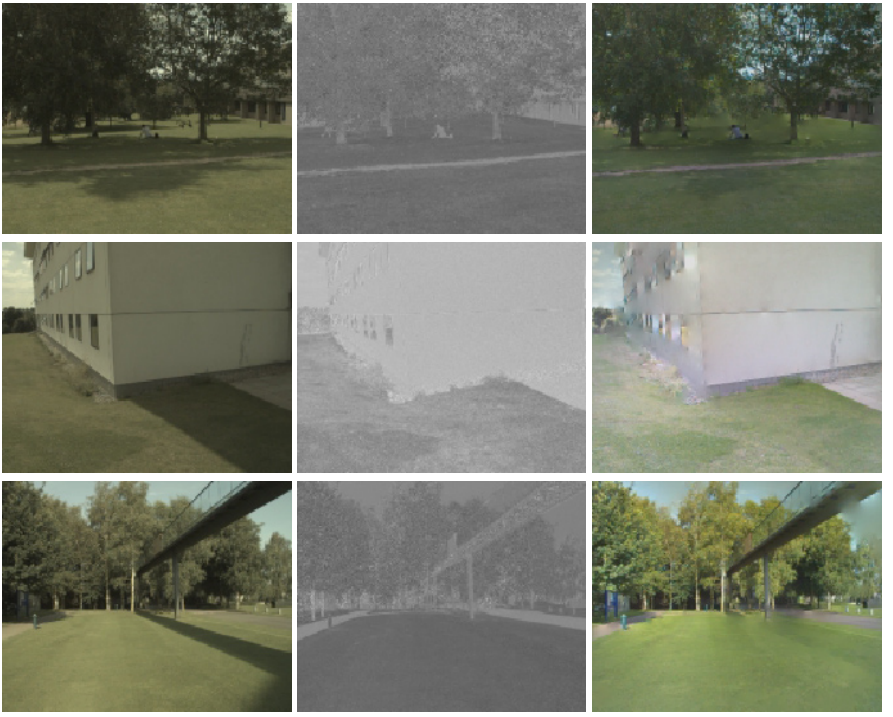

Fig. 5. Some example images. Each row shows results for a different image. The first column shows the original image, with shadow. The second column is the shadow-free illuminant-invariant greyscale image. The final column shows the recovered 3-band colour images which are shadow-free.

6 Conclusions

In this paper we have presented a method for finding and removing shadows from images. Our method builds on two strands of prior research: lightness algorithms and a recently developed light colour and light intensity invariant intrinsic image. Lightness algorithms attempt to disambiguate light from surface using the assumption that illumination varies slowly over an image. relative to this assumption, small gradients in images are due to illumination and large gradients

due to reflectance changes. Thresholding out small changes in gradient images and re-integrating the result yields an illumination-free intrinsic reflectance image. Of course this method fails when shadows are present because shadow edges have large gradients.

In this paper we modified the threshold function using the light intensity invariant image. As the name suggests this image depends only on reflectance. More specifically, it was shown that under the assumption of Planckian lights there exists a single scalar function of R, G and B that factors out illumination. By direct implication, shadows vanish in this image. It follows then that edges in an input image which do not appear in the invariant image are evidence of the presence of a shadow edge. Thresholding out image gradients that are due to shadows and re-integrating delivers full colour shadow free images.

Several examples of the method operating on real images are included in the paper. Shadows are always removed or very strongly attenuated. The paper includes a full disclosure on the necessary (but simple) steps required to calibrate a camera to support shadow removal.

References

1. G.D. Finlayson and S.D. Hordley. Color constancy at a pixel. *J. Opt. Soc. Am. A*, 18(2):253–264, Feb. 2001. Also, UK Patent application no. 0000682.5. Under review, British Patent Office.
2. R. Gershon, A.D. Jepson, , and J.K. Tsotsos. Ambient illumination and the determination of material changes. *J. Opt. Soc. Am. A*, 3:1700–1707, 1986.
3. D. L. Waltz. Understanding line drawings of scenes with shadows. In P.H. Winston, editor, *The Psychology of Computer Vision*, pages 19–91. McGraw-Hill, 1975.
4. P.M. Hubel. The perception of color at dawn and dusk. In *IS&T and SID's 7th Color Imaging Conference*, pages 48–51. 1999.
5. J.J. McCann. Lessons learned from mondrians applied to real images and color gamuts. In *IS&T and SID's 7th Color Imaging Conference*. 1999.
6. Y. Weiss. Deriving intrinsic images from image sequences. In *ICCV01*, pages II: 68–75. IEEE, 2001.
7. H.G. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics from images. In A.R. Hanson and E.M. Riseman, editors, *Computer Vision Systems*, pages 3–26. Academic Press, 1978.
8. G.D. Finlayson and M.S. Drew. 4-sensor camera calibration for image representation invariant to shading, shadows, lighting, and specularities. In *ICCV'01: International Conference on Computer Vision*, pages II: 473–480. IEEE, 2001.
9. http://www.srgb.com.
10. G.D. Finlayson, M.S. Drew, and B.V. Funt. Spectral sharpening: sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563, May 1994.
11. A. Blake. Boundary conditions for lightness computation in Mondrian world. *Comp. Vision, Graphics, and Image Proc.*, 32:314–327, 1985.
12. Th. Gevers and H. M. G. Stokman. Classifying color transitions into shadow-geometry, illumination highlight or material edges. In *International Conference on Image Processing*, pages 521–525, 2000.
13. R. Jain, R. Kasturi, and B.G. Schunck. *Machine Vision*. McGraw-Hill, 1995.

14. John Canny. A computational approach to edge detection. *IEEE Trans. Patt. Anal. Mach. Intell.*, 8:679–698, 1986.
15. S.M. Smith and J.M. Brady. SUSAN – A new approach to low level image processing. *Int. J. Comp. Vis.*, 23:45–78, 1997.
16. Yogesh Raja, J. McKenna, and Shaogang Gong. Colour model selection and adaptation in dynamic scenes. In *The Fifth European Conference on Computer Vision*. European Vision Society, 1998.
17. Mark S. Drew, Jie Wei, and Ze-Nian Li. Illumination-invariant image retrieval and video segmentation. *Pattern Recognition*, 32:1369–1388, 1999.
18. B.V. Funt and G.D. Finlayson. Color constant color indexing. *PAMI*, 17(5):522–529, May 1995.
19. M.J. Swain and D.H.. Ballard. Color indexing. *International Journal of Computer Vision*, 7(11):11–32, 1991.