Shadow-invariant classification for scenes illuminated by daylight

John A. Marchant and Christine M. Onyango

Image Analysis and Control Group, Silsoe Research Institute, Wrest Park, Silsoe, Bedford MK45 4HS, UK

Received November 23, 1999; revised manuscript received June 14, 2000; accepted June 26, 2000

A physics-based method for shadow compensation in scenes illuminated by daylight is proposed. If the daylight is represented by a simplified form of the blackbody law and the camera filters are of infinitely narrow bandwidth, the relationship between red/blue (r_m) and green/blue (g_m) ratios as the blackbody's temperature changes is a simple power law where the exponent is independent of the surface reflectivity. When the CIE daylight model is used instead of the blackbody and finite bandwidths for the camera are assumed, it is shown that the power law still holds with a slight change to the exponent. This means that images can be transformed into a map of r_m/g_m^A and then thresholded to yield a shadow-independent classification. Exponent A can be precalculated from the CIE daylight model and the camera filter characteristics. Results are shown for four outdoor images that contain sunny and shadowed parts with vegetation and background. It is shown that the gray-level distributions of the pixels in the transformed images are quite similar for a given component whether or not it is in shadow. The transformation leads to bimodal histograms from which thresholds can easily be selected to give good classifications. © 2000 Optical Society of America [S0740-3232(00)00411-7]

OCIS codes: 150.2950, 100.0100.

1. INTRODUCTION

Many attempts to analyze images start with the images as given data. Sometimes in an attempt to improve on this procedure the image analyst will try to control such factors as lighting, camera, and digitizer gains to improve the images. Occasionally the practitioner will use knowledge of the physical processes behind image formation to increase the changes of being able to analyze such images successfully. Although comparatively rare in the whole field of machine vision and image analysis, the latter approach has received some attention.[1,2] An important problem domain for the physics-based approach is in outdoor lighting, where complicated regimes that contain more than one illuminant, variable illumination, interreflections, highlights, shadows, and so on can exist. In these difficult situations the more prior knowledge that can be used, the better the chances of successful analysis will be.

Our study is applicable to many problem domains, but the one in which we are currently interested is detecting vegetation from ground-based vehicles. An important application is in guiding vehicles[3,4] or implements[5] to achieve precise control over weed or plant treatment. A requirement is to classify vegetation and soil correctly and accurately in real outdoor lighting conditions. Although the lighting is always from daylight, its intensity and spectral content may change over time. Also, in any one image there may be unavoidable shadows due to plant leaves shading parts of the scene or due to the vehicle itself, walls, trees, people, and so on, blocking out direct sunlight.

The problem of handling images that are subject to illumination changes was recently addressed by Drew *et al.*[6] Using their technique, one can normalize the image to remove (approximately) the effect of any illumination change when the change is over the whole image. This is a problem different from the one addressed here, in which we wish to compensate for the effect of illumination changes at different parts of the same image. In order to handle the extra problem of spatially variable illumination, we sacrifice the flexibility of being able to handle any illuminant change. In fact, our treatment applies only to blackbody-like illuminants for which the change is in the color temperature. Although this treatment is quite restrictive at first sight, it will be shown to apply to the important practical situation of daylight, including scenes that have shadows.

It is important to recognize that the illumination in *shadows* is not just less intense than that in direct sunlight. In shadows, the illumination is from skylight, which is bluer than that from the Sun. Thus there is an effect in shadows that is caused by there being two illuminants in the same scene, each of which has a different spectral content.[7] This is a different effect from *shading*, which is a phenomenon that can occur when only one illuminant is present. With shading, the relative angles among illumination, surface, and observer change, so the resultant perceived intensity changes.

Shadows are often seen as a nuisance in images and are dealt with by use of image processing techniques. If shadows occupy only a small proportion of the image, their information content can be ignored. With larger areas, a model-based image analysis technique may be used to compensate for the information loss.[8] In schemes that employ change detection, fixed shadows may become part of the background and discounted.[9] Some authors[10] double the number of classes in a classification problem to cope with the possibility that each surface might be in sun or shadow. Although simple in concept, this approach makes the classification problem much more difficult. As

part of a scheme to achieve color constancy in the presence of shadows, Nakauchi *et al.*[11] attempted to detect shadow edges (as opposed to reflectance edges) by noting that the spectral components of reflected light change systematically at shadow edges but in an unconstrained manner at reflectance edges.

In contrast to most previous studies dealing with shadows, we take a physics-based approach. On the basis of a physical model we propose a simple transformation that, given some approximations to the image formation processes, compensates for the shadows formed when scenes are illuminated by daylight. The symbols used in our notation are listed in Appendix A.

2. PHYSICAL BASIS

In the following we make a number of assumptions about the physical processes that apply in our work. Perhaps the most significant of these is that the materials behave as inhomogeneous dielectrics in which incident light enters a pigment colorant layer below the material surface, is scattered, and part of the light emerges from the surface after undergoing a spectral change. We further assume that the ratio between the spectral composition of the light that enters the materials and the light that leaves them is the same for all angles, part of what Tominaga and Wandell[12] call the standard reflectance model. When using our approach for classifying surfaces we also assume that the reflectances of the surfaces are constant over the scene and that reflectances and illumination are isotropic. In addition, we assume that there is no significant illumination from transmitted or interreflected light.

An important model in physics-based vision is the dichromatic reflection model.[13] The model states that the reflected light from an object can be described as the sum of the object's body and surface reflections where each component has an associated reflectance. The body reflectance is the scattered component described above. The spectral composition of the surface reflection component is usually regarded as that of the light source; Lee *et al.*[14] call this the neutral interface model. It is this component that gives rise to highlights or specularities. In this paper we disregard surface reflection. Our justification for doing so is that highlights occupy a small proportion of the total area in any real scene, that the highlights can be detected,[15–17] and that the areas detected as highlights will normally be surrounded by nonhighlight areas and thus can usually be filled in by use of morphological methods.

The physical process of light falling on a surface, its reflection, and interpretation by a video camera can be represented by a sequence of filtering operations; i.e.,

$$C_I = G_I \int S_I(\lambda)\rho(\lambda)E(\lambda)\mathrm{d}\lambda, \qquad (1)$$

where C_I is the output from a color channel (in our case $I \in \{R, G, B\}$), S_I is the spectral sensitivity of each channel, ρ is the reflectance of the surface that is being imaged, E is the illumination, and λ is the wavelength. G_I is a gain factor that is a product of two components. The first component depends on the camera (for example, on the electronics, color balance, aperture, and sensor integration time), and the second component depends on the relative angles of the surface, the illumination, and the observer. The integration is taken over a region to cover the sensitive range of the system. As mentioned above, dichromatic effects have been ignored in this study, and so there is only one component, where ρ is the body reflectance.

We use the JAI-M90 3-CCD color camera (JAI A/S, Glostrup, Denmark), although the theoretical treatment below is valid for any three-band camera. Note that we implicitly assume in the following that the camera outputs are proportional to the incident flux on the sensor, so it is important that the camera's gamma correction be turned off (i.e., set to 1.0). The spectral sensitivity of the M90 camera is shown in Fig. 1, where the vertical scale is arbitrary (any scale can be accommodated within G_I).

Figure 2 shows reflectances of typical vegetation and soil. The vegetation's reflectance was derived from the characteristic in Ref. 18 for maize under normal growth conditions and is in good agreement with characteristics of vegetation reported elsewhere in the literature for other crops.[19–22] The reflectance characteristics of soil in general increase slightly with wavelength, for example, as in the characteristics given in Ref. 23 for sandy loam, clay, and fen soil. Although particular types of vegetation and soil will have characteristics that are different from these, the conclusions from our study do not depend on the exact spectral characteristics, and so those in Fig.

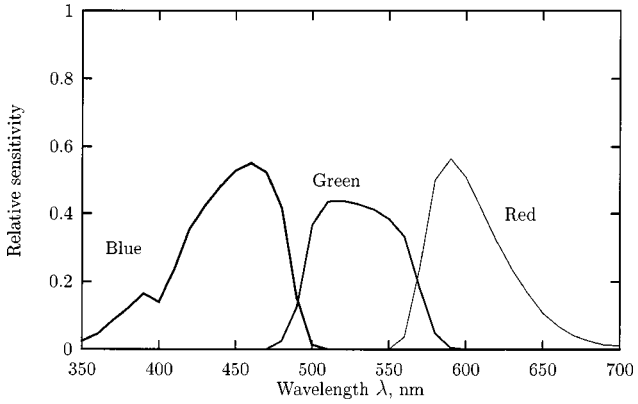

Fig. 1. Spectral sensitivity (arbitrary units) of the three bands of the M90 camera.

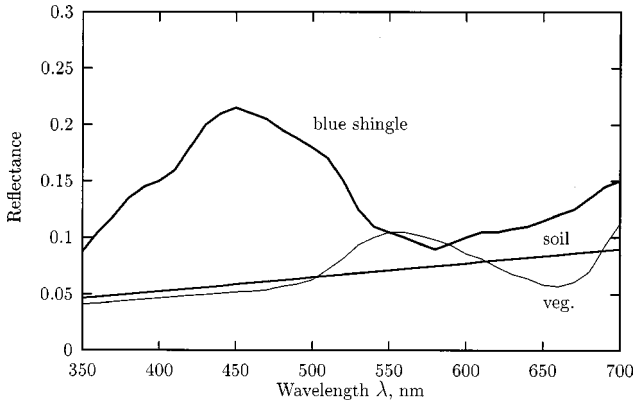

Fig. 2. Spectral reflectance of typical vegetation and soil and of a third material (blue shingle; see Ref. 24).

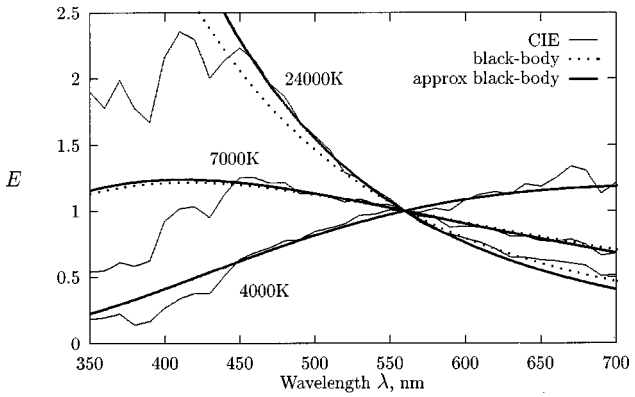

Fig. 3. CIE daylight standard at several CCT's. Also shown are blackbody spectral content and an approximate form of the blackbody equation. Note that the blackbody and approximate forms merge at the lowest CCT.

2 are perfectly adequate for illustration. Also shown is a characteristic for a material outside our particular domain and with a very different spectrum, blue shingle.[24] We shall use this characteristic later in the paper.

Most general-purpose illuminants, e.g., tungsten filament lamps, tubular and compact fluorescent lamps, and daylight, may be characterized according to their correlated color temperatures (CCT's).[25] The CCT is the temperature of a blackbody emitting light of a spectral composition approximately similar to that of the illuminant. The Commission Internationale de l'Eclairage (CIE) has a standard[26] that specifies the spectral composition of daylight at various CCT's on the basis of a study by Judd *et al.*[27] Figure 3 shows data from the CIE standard at a few CCT's, along with two theoretically derived approximations that we shall use below. Three CCT's have been chosen for illustration. The lowest and highest are the extreme values of those usually tabulated and would apply to the sunny component of a very red sky and the shadow component of an unusually blue sky, respectively. The other temperature might typically come from an overcast sky.

3. SHADOW-INVARIANT CLASSIFICATION

Our objective in this section is to show that the color of a surface as seen through a three-band camera can be represented in a way that makes that color nearly invariant to certain illumination changes. Strictly speaking, the illumination should come from an approximation to a blackbody radiator and the allowed changes are of the blackbody temperature. However, it will be shown that if the method is used with the CIE daylight standard at varying color temperatures, the approximations used are still valid. This makes it possible to derive a simple classification scheme, based on sound physical principles, that will deal with the important practical situation of daylight illumination with shadows.

A. Blackbody Illumination

Planck's formula states that the spectral radiant exitance of a blackbody at temperature T, per unit wavelength interval, is given by[24]

$$M_{e\lambda} = c_1\lambda^{-5}[\exp(c_2/T\lambda) - 1]^{-1}, \tag{2}$$

where λ is the wavelength and c_1 and c_2 are constants. As with the CIE daylight standard, this is usually presented in a normalized way, so $M_{e\lambda}$ is unity at 560 nm. The exitance is often approximated as

$$M_{e\lambda} = c_1\lambda^{-5}\exp(-c_2/T\lambda) \tag{3}$$

(see Ref. 24 but note the missing minus there). Figure 3 shows the normalized $M_{e\lambda}$ [Eq. (2)] and the approximate form [Eq. (3)] along with the CIE standard at a number of CCT's. It can be seen that the blackbody and the approximate blackbody formulations represent the CIE standard reasonably well over the visible range, with the two blackbody curves being indistinguishable at the low CCT.

We now introduce a second approximation, which is due to Finlayson *et al.*,[28] that extends the von Kries coefficient rule[29] and allows us to transform the color of a pixel into the color that would have been observed at a different illumination condition. The transformation is a simple scaling on each color channel that is independent of the surface color. A simplified approach to deriving the scalings is given here.

If the passbands of the camera filters were very narrow, it might be possible to represent them by impulse functions centered on the actual filter characteristics. With this (rather gross) approximation, Eq. (1) becomes

$$C_I = g_I\rho(\lambda_{cI})E(\lambda_{cI}), \tag{4}$$

where λ_{cI} is the center frequency of the channel I filter and g_I depends on G_I and the shape of the filter. Thus, to transform from a color measured at a given illumination (suffix m) to a color that would have been observed at a reference illumination (suffix ref), we can use

$$C_{I\text{ref}} = C_{Im}E_{\text{ref}}(\lambda_{cI})/E_m(\lambda_{cI}), \tag{5}$$

which shows that the transformation is a simple scaling for each channel that depends only on the two illuminants (that at the measurement and the reference illuminant) and the camera characteristics. An important feature is that it does not depend on the reflectance properties of the surface.

Our particular problem domain is daylight illumination with shadows. This situation can produce a very large dynamic range in the data if absolute values are used. To reduce potential problems from this source we will follow Barnard *et al.*[30] and use two ratios, which we will call the red and green *band ratios* rather than absolute measurements. These band ratios are defined as $r = C_R/C_B$ and $g = C_G/C_B$. Note that Barnard *et al.* call them *chromaticities*, which most workers define differently, e.g., $r = C_R/[C_R + C_G + C_B]$. Use of the ratios also means that the photometric angle component of the g_I terms cancels in the equations below. This means that effects due to shading will be removed. As was pointed out by Barnard *et al.*,[30] the simple scaling result applies just as well to band ratios as it does to absolute color values, allowing us to write

$$r_{\text{ref}} = r_m/s_r, \qquad g_{\text{ref}} = g_m/s_g, \tag{6}$$

where

$$s_r = \frac{E_{\text{ref}}(\lambda_{cB})E_m(\lambda_{cR})}{E_{\text{ref}}(\lambda_{cR})E_m(\lambda_{cB})}, \qquad s_g = \frac{E_{\text{ref}}(\lambda_{cB})E_m(\lambda_{cG})}{E_{\text{ref}}(\lambda_{cG})E_m(\lambda_{cB})}. \quad (7)$$

If we now substitute $M_{e\lambda}$ for illuminant E from Eq. (3) into Eqs. (7) we obtain

$$s_r = \exp\left[c_2\left(\frac{1}{T_{\text{ref}}} - \frac{1}{T_m}\right)\left(\frac{1}{\lambda_{cR}} - \frac{1}{\lambda_{cB}}\right)\right],$$

$$s_g = \exp\left[c_2\left(\frac{1}{T_{\text{ref}}} - \frac{1}{T_m}\right)\left(\frac{1}{\lambda_{cG}} - \frac{1}{\lambda_{cB}}\right)\right], \quad (8)$$

and so

$$s_r = {s_g}^A, \quad (9)$$

where

$$A = \frac{1/\lambda_{cR} - 1/\lambda_{cB}}{1/\lambda_{cG} - 1/\lambda_{cB}}. \quad (10)$$

Combining Eq. (9) with Eqs. (6), we obtain

$$r_m = F{g_m}^A, \quad (11)$$

where

$$F = r_{\text{ref}}/g_{\text{ref}}^A. \quad (12)$$

Now r_{ref} and g_{ref} can be obtained using Eq. (4) with the appropriate color channels and dividing to form the band ratios as

$$r_{\text{ref}} = \frac{g_R\rho(\lambda_{cR})E_{\text{ref}}(\lambda_{cR})}{g_B\rho(\lambda_{cB})E_{\text{ref}}(\lambda_{cB})}, \qquad g_{\text{ref}} = \frac{g_G\rho(\lambda_{cG})E_{\text{ref}}(\lambda_{cG})}{g_B\rho(\lambda_{cB})E_{\text{ref}}(\lambda_{cB})}. \quad (13)$$

The interpretation of E_{ref} depends on the context. For example, for a blackbody illuminant E_{ref} would be the value of $M_{e\lambda}$ at a temperature chosen as a reference. Thus the factor F in Eq. (11) is a function of the surface (through ρ). It is also a function of the color balance of the camera (which affects g_R/g_B and g_G/g_B) but not of the aperture setting nor of the exposure time, as they do not affect the ratios g_R/g_B and g_G/g_B.

Note that none of the components of F [Eqs. (10), (12), and (13)] depends on E_m, the illumination of the imaged pixel. A and g_I depend only on the camera, ρ depends only on the surface, and E_{ref} depends only on the illumination chosen as reference (which is arbitrary). Therefore a major conclusion is that with any image (fixed camera, fixed color balance, fixed choice of reference illumination), with lighting from an approximate blackbody source, all pixels from the same surface will obey the relationship in Eq. (11). This is true *even if parts of the scene are illuminated with a blackbody at a different temperature*. The conclusion does, of course, depend on the narrow-passband assumption made above. It therefore should be possible to differentiate between surfaces in such a scene by transforming the image to a plot of $F = r_m/{g_m}^A$ and classifying the result as a monochrome image. In a two-component scene one could do this by plotting a histogram of F (a bimodal plot should result) and choosing the threshold on F appropriately. In multicomponent scenes, theoretically it will be possible to choose $n - 1$ thresholds to separate the n components. Of course, just as for any other classification method, its practicality depends on the separation of the components in F and the variability in F in any one component. Note that the method and the exponent are independent of the color balance. If the color balance were changed the histogram might be stretched or compressed, whereupon the appropriate value of F for classification would be different.

B. Effect of Approximations on Scaling Factors

In this section we investigate how well the conclusions above relate to daylight instead of to blackbody illumination and also the effect of finite camera filter bandwidths.

So far, the general conclusions are independent of the particular camera used. To illustrate the effect of using a blackbody instead of daylight, we must fix upon a particular camera. We derive the center frequencies of the M90 camera's filters (which are typical of common commercially available devices) from a weighted average of the sensitivities in each color channel; i.e.,

$$\lambda_{cI} = \frac{\int S_I(\lambda)\lambda\,\mathrm{d}\lambda}{\int S_I(\lambda)\,\mathrm{d}\lambda}, \quad (14)$$

where the integrals are taken over the passbands of the filters. The operation yields the red, green, and blue center frequencies as 605.2, 530.5, and 440.8 nm, respectively. There is no strong reason for using this particular measure of center frequency; however, use of the modes would, for example, make the values rather sensitive to errors in characterizing the filters if the curves had flattish tops.

Some scaling factors for the M90 camera with the CIE standard, the blackbody model [Eq. (2)], and the approximate blackbody model [Eq. (3)] are listed in Table 1 for a reference temperature of 6500 K. Note that these scaling factors, under the assumption of infinitely narrow filter bandwidths, are independent of reflectance and so apply to any surface.[28] It can be seen that the scaling factors derived from both blackbody models are generally within 10% of the CIE standard, with the exception of s_r at the very high temperature.

Table 1. Scaling Factors for Various Illumination Models

Temperature (K)	s_r			s_g		
	CIE	Eq. (2)	Eq. (3)	CIE	Eq. (2)	Eq. (3)
4000	2.49	2.31	2.35	1.82	1.69	1.70
7000	0.91	0.91	0.91	0.94	0.94	0.94
24000	0.43	0.43	0.37	0.57	0.59	0.54

Figure 4 shows s_r plotted against s_g for the three sample CCT's, where the scaling factors have been calculated for the CIE daylight standard. To see how well Eq. (9) describes data calculated in this way, we derived a value for the constant A by using

$$A = \sum \frac{\log s_r}{\log s_g}, \tag{15}$$

where the sum was taken over values tabulated at 1000 K intervals, yielding $A = 1.51$ as opposed to $A = 1.61$ with use of Eq. (10). A curve using $A = 1.51$ is also plotted in Fig. 4 that shows that Eq. (9) with a modified value of A explains the data well. Thus our conclusion that all pixels from the same surface will obey the relationship in Eq. (11) should be just as valid for daylight at different CCT's as it is for an approximate blackbody. All that is required is a slight change in the value of A.

The degree of approximation caused by assuming infinitely narrow camera filters depends on how rapidly E and ρ vary over the passbands. If E were constant over each band, it can be shown that the scaling factors would be the same as in the case of an infinitely narrow bandwidth. It can also be shown that, if ρE varies linearly over the passband and λ_{cI} is defined as in Eq. (14), the scaling factors will be the same again. Although we could possibly claim enough smoothness in ρ and E for this result to apply to some extent, the actual effect depends on the reflectances used. Thus, to estimate the degree of approximation, we use the surfaces that we are primarily interested in, i.e., vegetation and soil, whose reflectances we have characterized in Fig. 2. Note that these curves need not specify precisely any actual vegetation and soil that we use in later tests, as our thesis is that scaling factors are influenced only slightly by the particular surfaces observed (in fact, with our assumptions of approximate blackbody illumination and infinitely narrow camera filters, they are not influenced at all). In order to stretch our thesis we also use a surface outside our problem domain that has a very different characteristic, blue shingle. Table 2 shows scaling factors for the different components with the CIE daylight standard and with the M90 camera. The factors have been calculated by use of Eq. (1) along with the reflectance characteristics, the camera filter characteristics, the CIE standard at the CCT's, and the reference CCT (6500 K). It can be seen that the scaling factors are almost independent of the surface and are very close to the figures for any surface in Table 1. Thus the effect of finite filter bandwidths is small.

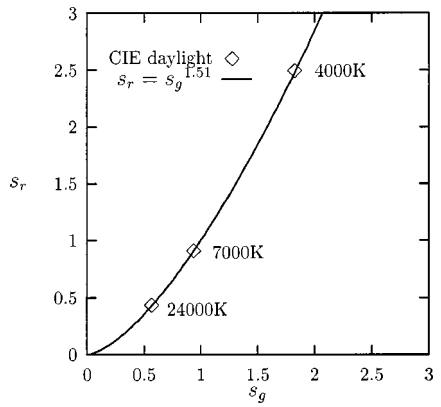

Fig. 4. s_r versus s_g for CIE daylight, assuming infinitely narrow camera filter bands. Also shown is the curve $s_r = s_g^{1.51}$.

It is useful to note that if filter bandwidth did cause a problem, advantage could be taken of spectral sharpening techniques,[31] with which a transformation of the normal color space is used in which the filters are narrower.

C. Potential for Classification

So far we have shown that there should be a simple relationship between band ratios at each pixel in an image [Eq. (11)]. We have also shown that the scaling factors used in the development of this theory should be sensibly independent of the particular surface that is being viewed. This allows us to propose using Eq. (11), for which F is dependent on the surface but not on the illumination, as a way of classifying surfaces in the presence of illumination changes.

To demonstrate the potential for classifying by this approach we calculate plots of r_m versus g_m for daylight at three different CCT's (Fig. 5) for the vegetation and soil components characterized in Fig. 2. Also shown are two curves of Eq. 11 with $A = 1.51$ (the value for daylight and the M90 camera's center frequencies) and $F = 0.75, 1.05$. For any surface in an image, as the illumination changes, the values of r_m and g_m for each pixel will move along curves from this family. The soil and vegetation plots are well represented by the respective curves, even in the presence of the approximations that we have made. This shows that, if values measured from pixels are transformed into $r_m / g_m^{1.51}$, a simple threshold of ~0.9 could be used to separate these two components, whatever the value of the CCT, and in particular independent of whether the components are in sunlight or shadow.

4. EXPERIMENT

A. Offset in the Camera–Digitizer System

The light intensity in shadows is much less than that in direct sun. This makes it difficult to avoid saturation of the camera–digitizer system in the sunny areas while maintaining reasonable pixel values in the shadows.

Table 2. Scaling Factors for Vegetation, Soil, and Blue Shingle for Daylight with the M90 Camera

Temperature (K)	s_r			s_g		
	Vegetation	Soil	Shingle	Vegetation	Soil	Shingle
4000	2.22	2.28	2.28	1.68	1.66	1.61
7000	0.91	0.91	0.91	0.94	0.94	0.95
24000	0.43	0.43	0.43	0.58	0.59	0.62

Tominaga[32] reduced this problem by using more than one exposure, each at a different shutter speed; we may adopt this approach in future research. However, with the possibility of low pixel values we must ensure that any offset in the camera–digitizer is compensated for. The output of the camera is proportional to the integration over time of the light intensity at the sensor.[33,34] Thus, if the light intensity is constant over the exposure time, we can investigate the offset by changing either the light intensity or the integration time. As controlling and measuring the light intensity (while maintaining the same spectral distribution) is not easy, we choose changing the integration time. Note that the offset cannot be established by merely placing the lens cap on. There is a possibility that the offset will be negative, in which case the digitizer will give an output of zero.

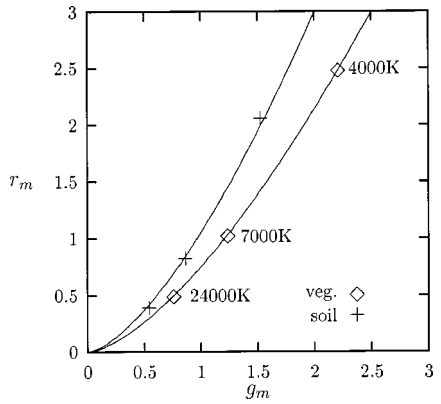

Fig. 5. r_m versus g_m calculated for vegetation and soil with the M90 camera at three CCT's. Also shown are examples of the function $r_m = Fg_m^{1.51}$ for $F = 0.75, 1.05$.

A white reference target was illuminated in a controlled-lighting cabinet. The pixel histograms in each channel were observed and their spread found to be less than 15% of their mean values. Thus it was decided that the illumination was reasonably uniform and that the output of the digitizer could be characterized by the mean value in each channel. The output was plotted against the integration time for five values between 100 and 10 000 μs. The lines were straight ($r^2 > 0.9998$ for all channels), and the offsets were 4, 3, and 3 for the red, green, and blue channels, respectively, with a maximum standard error on the offset of 0.62. As the offsets were found to be positive, placing the lens cap on should give the same result. This was confirmed by experiment. The offsets were subtracted from the pixel values in the following results.

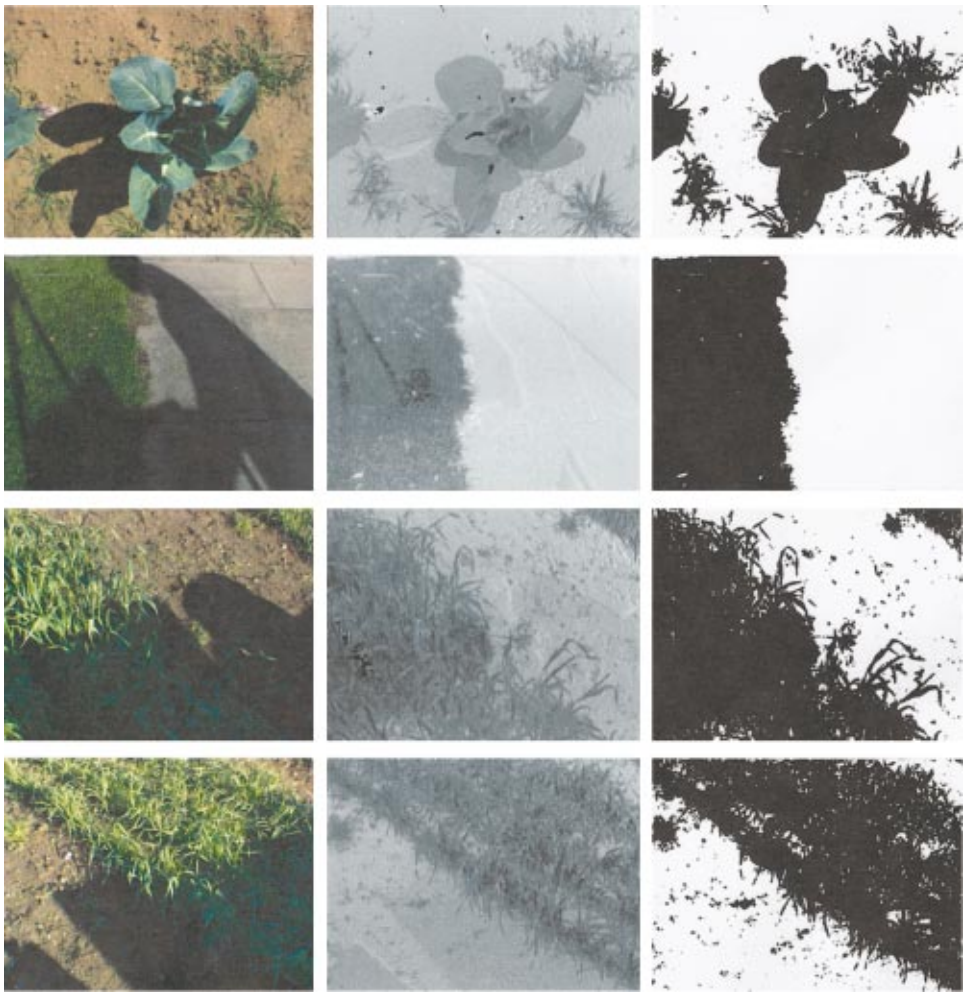

Fig. 6. Images exhibiting sunny and shadow components. Top, $sel5$; upper middle, $sel8$; lower middle $jtse\_000$; bottom $jsse\_000$. Left, original images; middle images of $r_m/g_m^{1.51}$; right, classified images.

B. Classification of Images

Figure 6 shows four images captured on a sunny day. For the moment we consider the first two (the top, known as $sel5$, and the upper middle, known as $sel8$). Each image has two different surfaces (vegetation and background) with shadows. In $sel5$ the vegetation is a cauliflower plant with weeds, and the background is soil. In $sel8$ the vegetation is a grass lawn and the background is a cement slab pathway. Shadows in $sel5$ were cast by the vegetation and in $sel8$ by a computer trolley and the authors.

Ground truth was produced by using the Adobe Photoshop package to assist a human operator to classify the images. The main object of this manual classification (Fig. 7) was to investigate the distribution of intensities in the four classes: (1) vegetation in sun, (2) background in sun, (3) vegetation in shadow, and (4) background in shadow. Because it was not important to classify every pixel, difficult (e.g., the thin straggly weeds in $sel5$) or doubtful (e.g., in deep shadow) areas were classified as "don't know."

Figure 6 (middle column) shows the images transformed such that the gray level represents $r_m/g_m^{1.51}$. The transformation makes the vegetation darker than the background, as the vegetation is greener. Pixels where any of the three components (R, G, or B) was zero have been set to zero to avoid division problems. This was also done with pixels that show saturation in any of the three channels; these areas show as black in the transformed images. It can be seen that the total area in these categories was extremely small. However, dividing small numbers by similarly small numbers gives rise to noise in the transformed image. We partially compensated for this noise by passing a 5×5 median filter over the transformed image. In the $sel5$ transformed image the area of shadowed soil is made more apparent by a bright edge on one side and a dark edge on the other. This effect is due to a camera aberration, possibly a slight misalignment of the color channels, that gives incorrect colors at edges. The rest of the shadowed soil is very close in intensity to the soil in sunlight. In the transformed image of $sel8$ the shadows have virtually disappeared; they are visually apparent only from the aberration mentioned above.

Figure 8 shows the histograms of the various components and the totals of the two images. For $sel5$ the distributions for the sunny and shadowed parts of the same surface are visually quite similar. However, there is a tendency for the shadowed areas to be darker in the transformed image. This tendency is also apparent (to a lesser extent) in the background of $sel8$. Nevertheless, the histograms of the total transformed images show clear bimodal structures, making it easy to choose thresholds for classifying the images into vegetation and background.

If thresholds are chosen at the minima of the troughs between the two modes (0.92 for $sel5$ and 0.82 for $sel8$), the classifications of Fig. 6 (right) result. In $sel5$, with the exception of some very small areas, all the parts actually classified manually (i.e., other than "don't know") have been assigned to the correct components. It is interesting to see how the thresholding of the transformed image has dealt with the areas manually classified as "don't know." For example, the straggly weeds have been classified in a plausible manner, even when they are in shadow toward the lower left of the image. Also, there was a "don't know" region on the left side of the main vegetation area that was so dark in the original image that it could not be classified manually; it has been thresholded as "vegetation." On closer inspection of the image and with some manually assisted image processing, it was found that this was indeed a leaf in shadow. A similarly successful thresholding is seen with $sel8$. Here there

Fig. 7. Manual classification of images $sel5$ (left) and $sel8$ (right). From lightest to darkest: vegetation sun, background sun, vegetation shade, background shade, "don't know."

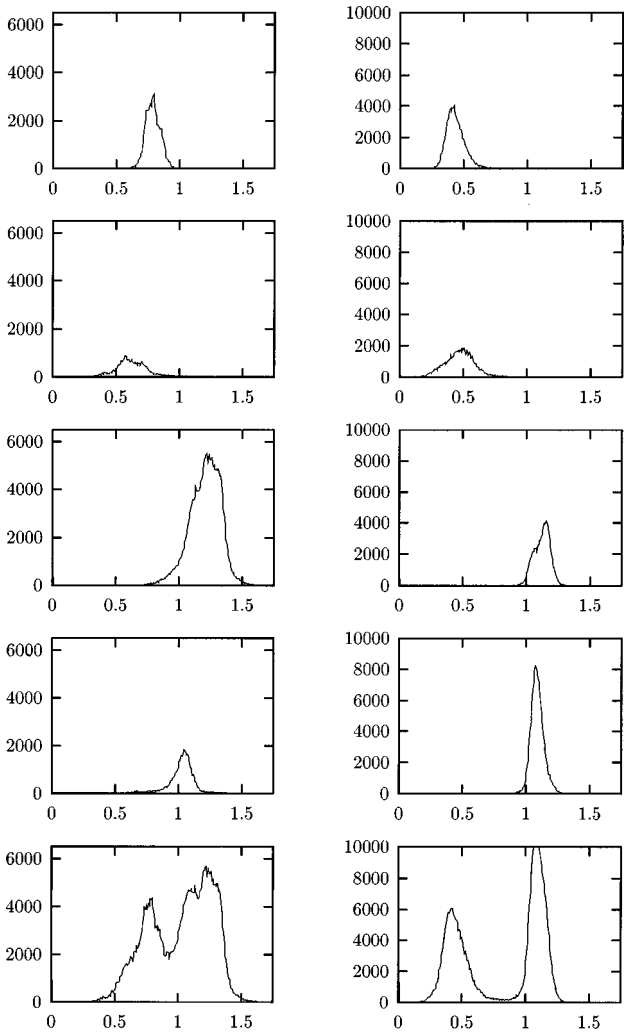

Fig. 8. Histograms of transformed images ($r_m/g_m^{1.51}$). Left, $sel5$; right, $sel8$. Rows from top to bottom: vegetation sun, vegetation shadow, background sun, background shadow, total.

Table 3. Mean Value and Spread of the Function $r_m/g_m^{1.51}$ As the Viewing Direction Was Changed for Images of Tall Wheat Approximately Half in Sun and Half in Shadow

	North		East		South		West	
Variable	Sun	Shadow	Sun	Shadow	Sun	Shadow	Sun	Shadow
Mean	0.41	0.40	0.46	0.42	0.37	0.38	0.40	0.33
Spread	0.31	0.44	0.42	0.35	0.33	0.38	0.53	0.38

Table 4. Spatial Variation of the Mean Value and Spread of the Function $r_m/g_m^{1.51}$ for an Image of Tall Wheat Approximately Half in Sun and Half in Shadow

	Values of F, by Horizonal Region							
Vertical Region 1								
Mean	0.38	0.34	0.37	x^a	0.40	0.35	0.40	0.52
Spread	0.29	0.24	0.30	x^a	0.49	0.39	0.38	0.60
Vertical Region 2								
Mean	0.30	0.30	0.29	x^a	0.39	0.33	0.37	0.33
Spread	0.31	0.30	0.27	x^a	0.59	0.38	0.49	0.39
Vertical Region 3								
Mean	0.27	0.26	0.30	x^a	0.36	0.40	0.36	0.34
Spread	0.25	0.24	0.24	x^a	0.53	0.57	0.51	0.38
Vertical Region 4								
Mean	0.27	0.34	0.34	0.35	x^a	0.45	0.38	0.46
Spread	0.26	0.33	0.29	0.35	x^a	0.49	0.45	0.56

[a] The area contained a mixture of sun and shade so no value is given. The shadow area is to the left of the x and the sunny area to the right.

was a "don't know" region in the lower shadowed section toward the middle. This region has been classified as having a boundary between vegetation and background. Although this classification is not readily apparent from the image (the shadowed area is too dark to allow a decision), it fits in with the prior knowledge that the path edge was straight and extended from the top of the image to the bottom.

C. Violation of Basic Assumptions

In Section 2 we stated the assumptions on which our principle is based. However, in our particular problem domain some of these assumptions will be violated. For example, in crops that form a canopy, light transmitted through and reflected from the leaves will give an illumination greener than that from a blackbody. Also, the small areas of specular reflection that occur from a crop canopy combine to give an apparent reflectance that depends on the relative angles among the Sun, the camera, and the canopy. The macroscopic effect is known as the bidirectional reflectance distribution function, which is influenced by any anisotropy in the canopy.

To check the performance of the function F [Eq. (12)], where the assumptions are obviously violated, we considered images of a wheat canopy. First, images were captured with crop height 0.75 m, camera height 0.85 m above the top of the crop, and camera angle 55° to the vertical. This gave a field of view of approximately 1.2 m width at the top of the image, 0.7 m width at the bottom, and 1.0 m from top to bottom. The wheat was planted in rows of ~0.22 m spacing, and it was just possible to see the soil between the rows from a direct overhead view. The rows were aligned east–west, and the images were captured with the Sun in a direction just a few degrees south from east and at ~30° elevation. Four images were captured with the camera viewing from the north, east, south, and west respectively. The canopy filled the whole of each image, and a shadow was cast over approximately half of the image by use of opaque board.

It is important to realize that we do not expect the small shadows or shading effects within the crop canopy to be eliminated in a transformed image, as these are likely to be influenced by transmission and interreflections. We still expect to see a textured surface in these areas, and we may even wish eventually to use the properties of the texture to help in classification. What we do hope to remove are large shadows caused, for example, by vehicles, equipment, and people near the images. For this to succeed, the value of F must not be influenced too much by macroscopic effects, such as the bidirectional reflectance distribution function, over the extent of the image. Also, F must be similar in sun and shadow in the same image.

Table 3 lists the value of F for each component of each image. The mean value and the spread are listed. The spread is defined as the interval of F that contains 5%–95% of the values in the F histogram. From an inspection of the table, taking into account the spread (which illustrates the fine-scale variation in the transformed canopy image), we see no noticeable change in F caused either by viewing direction or by shadows. We used one of the images (west view), where the camera view was nearly into the Sun, to investigate the spatial variation of F over the image. As well as spatial variation, the images contain viewing angle variations (rays from different parts of the image enter the camera at different angles) of

approximately 34° to 69° along the vertical center line of the image. The image was divided into rectangular regions, eight across the image and four down. Table 4 lists values of F in each of the regions. Once again, we see no consistent difference in F either spatially or between sun and shadow.

To test whether F can be used to classify crop and soil we used two images, known as $jtse\_000$ and $jsse\_000$, of a mixture of crop and soil. $jtse\_000$ is from the crop used in the experiment described above, viewed from the southeast, and $jsse\_000$ is from a shorter crop (height approximately 0.25 m viewed from the same direction. Figure 6 shows the results. Despite the heavy shadowing, the value of F in crop or soil is nearly independent of shadowing. The two images were threshold at the same value of F (0.82), and the resulting classification is excellent. Even slight gaps in the crop canopy where the soil is visible (more apparent with the shorter crop) have been correctly classified. A small area of $jtse\_000$, nearly halfway up toward the left, has been left unclassified because the color components are too low for the transformation to be performed (see Subsection 4.B).

5. CONCLUSIONS

If it is assumed that daylight can be represented by a simplified form of the blackbody law and that the camera filters are of infinitely narrow bandwidth, then the relationship between red and green band ratios of image pixels as the blackbody temperature changes is a simple power law. The exponent A in the relationship is independent of the surface reflectivity and can be precalculated from the camera filter characteristics.

When the CIE daylight model is used instead of the approximate blackbody, and finite bandwidths for the camera are assumed, the power law still holds with a slight change to the exponent. The revised value of the exponent can be precalculated from the CIE daylight model and the camera filter characteristics. This means that images can be transformed to a map of $r_m/{g_m}^A$ and then thresholded to yield a shadow-independent classification.

The results for two outdoor images show that the gray-level distributions of the pixels in the transformed images are very similar for a given component whether or not it is in shadow. This permits bimodal histograms to be formed from which thresholds can easily be selected to give good classifications.

APPENDIX A: SYMBOLS USED

I	Suffix used to denote color channel, R, G, or B
m	Suffix used to denote conditions at a measured value
ref	Suffix used to denote conditions at a reference value
A	Exponent in relationship between red and green band ratios
C_I	Camera output from channel I
c_1, c_2	Constants in Planck's law
E	Spectral power distribution of the illumination
F	Ratio of r_m to ${g_m}^A$ at a pixel
G_I	Gain factor for channel I
g_I	Factor depending on G_I and the shape of the camera filter
g	Band ratio G/B
$M_{e\lambda}$	Spectral radiant exitance of a blackbody radiator
n	Number of components to be classified in a scene
r	Band ratio R/B
S	Sensor spectral sensitivity function
s_g	Scaling factor relating g for different illumination conditions
s_r	Scaling factor relating r for different illumination conditions
T	Color temperature of the illuminant or absolute temperature of a blackbody
λ	Wavelength
λ_c	Wavelength at center frequency of camera filter
ρ	Spectral reflectance function of the surface

ACKNOWLEDGMENT

This research was funded by the Biotechnology and Biological Sciences Research Council, UK.

J. A. Marchant's e-mail address is john.marchant@bbsrc.ac.uk.

REFERENCES

1. G. E. Healey, S. A. Shafer, and L. B. Wolff, eds., *Physics-Based Vision: Principles and Practice* (Jones & Bartlett, Boston, Mass., 1992).
2. "Physics-Based Machine Vision," feature issue, J. Opt. Soc. Am. A **11**, 2922–3100 (1994).
3. R. Brivot and J. A. Marchant, "Segmentation of plants and weeds using infrared images," Proc. Inst. Electr. Eng. **143**, 118–124 (1996).
4. T. Hague, J. A. Marchant, and N. D. Tillett, "A system for plant scale husbandry," in *Proceedings of the 1st European Conference on Precision Agriculture* (BIOS Scientific, Oxford, UK, 1997), pp. 635–642.
5. N. D. Tillett and T. Hague, "Computer-vision-based hoe guidance for cereals—an initial trial," J. Agric. Eng. Res. **74**, 225–236 (1999).
6. M. S. Drew, J. Wei, and Z. Li, "Illumination invariant image retrieval and video segmentation," Pattern Recogn. **32**, 1369–1388 (1999).
7. A. D. Jepson, R. Gershon, and J. K. Tsotsos, "Ambient illumination and the determination of material changes," J. Opt. Soc. Am. A **3**, 1700–1707 (1986).
8. D. Reynard, A. Wildenberg, A. Blake, and J. A. Marchant, "Learning dynamics of complex motions from image sequences," in *Proceedings of the 4th European Conference on Computer Vision* (Springer, Berlin, 1996), pp. 357–368.
9. G. L. Foresti, "Object detection and tracking in time-varying and badly illuminated outdoor environments," Opt. Eng. **37**, 2550–2564 (1998).
10. J. D. Crisman and C. E. Thorpe, "SCARF: a color vision system that tracks roads and intersections," IEEE Trans. Rob. Autom. **9**, 49–57 (1993).
11. S. Nakauchi, K. Takebe, and S. Usui, "A computational model for color constancy by separating reflectance and illuminant edges within a scene," Neural Networks **9**, 1405–1415 (1996).
12. S. Tominaga and B. A. Wandell, "Standard surface-reflectance model and illuminant estimation," J. Opt. Soc. Am. A **6**, 576–584 (1989).
13. S. A. Shafer, "Using color to separate reflection components," Color Res. Appl. **10**, 210–218 (1985).

14. H.-C. Lee, E. J. Breneman, and C. P. Schulte, "Modeling light reflection for color computer vision," IEEE Trans. Pattern Anal. Mach. Intell. **12**, 402–409 (1990).
15. G. Brelstaff and A. Blake, "Detecting specular reflections using Lambertian constraints," in *Proceedings of the 2nd International Conference on Computer Vision* (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 297–302.
16. H. C. Lee, "Method for computing the scene-illuminant chromaticity from specular highlights," J. Opt. Soc. Am. A **3**, 1694–1699 (1986).
17. G. J. Klinker, S. A. Shafer, and T. Kanade, "The measurement of highlights in color images," Int. J. Comput. Vision. **2**, 7–32 (1988).
18. S. J. Maas and J. R. Dunlap, "Reflectance, transmittance, and absorptance of light by normal, etiolated, and albino corn leaves," Agron. J. **81**, 105–110 (1989).
19. C. A. Shull, "A spectrophotometric study of plant reflection of light from leaf surfaces," Bot. Gaz. **87**, 583–607 (1929).
20. W. D. Billings and R. J. Morris, "Reflection of visible and infrared radiation from leaves of different ecological groups," Am. J. Bot. **38**, 327–331 (1951).
21. J. T. Wooley, "Reflectance and transmittance of light by leaves," Plant Physiol. **47**, 656–662 (1971).
22. E. A. Walter-Shea and J. M. Norman, "Leaf optical properties," in *Photon-Vegetation Interactions* (Springer-Verlag, Berlin, 1991), pp. 230–250.
23. A. W. Hooper, G. O. Harries, and B. Ambler, "A photoelectric sensor for distinguishing between plant material and soil," J. Agric. Eng. Res. **21**, 145–155 (1976).
24. G. Wyszecki and W. S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, 2nd ed. (Wiley, New York, 1982).
25. Commission Internationale de L'Eclairage (CIE), "Method of measuring and specifying colour rendering properties of light sources," Tech. Rep. CIE. 13.3 (CIE, Paris, 1995).
26. Commission Internationale de L'Eclairage (CIE) "Colorimetry," Tech. Rep. 2nd ed. (CIE, Paris, 1986).
27. D. B. Judd, D. L. MacAdam, and G. W. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," J. Opt. Soc. Am. **54**, 1031–1040 (1964).
28. G. D. Finlayson, M. S. Drew, and B. F. Funt, "Color constancy: generalized diagonal transforms suffice," J. Opt. Soc. Am. A **11**, 3011–3019 (1994).
29. G. West and M. H. Brill, "Necessary and sufficient conditions for von Kries chromatic adaptation to give color constancy," J. Math. Biol. **15**, 249–258 (1982).
30. K. Barnard, G. Finlayson, and B. Funt, "Color constancy for scenes with varying illumination," Comput. Vision Image Understand. **65**, 311–321 (1997).
31. G. D. Finlayson, M. S. Drew, and B. F. Funt, "Spectral sharpening: sensor transformations for improved color constancy," J. Opt. Soc. Am. A **11**, 1553–1563 (1994).
32. S. Tominaga, "Multichannel vision system for estimating surface and illumination functions," J. Opt. Soc. Am. A **13**, 2163–2173 (1996).
33. J. D. E. Beynon and D. R. Lamb, eds., *Charge Coupled Devices and Their Application* (McGraw-Hill, London, 1980).
34. S. Perrin and T. Redarce, "CCD camera modelling and simulation," J. Intell. Robot. Syst. **17**, 309–325 (1996).