Multispectral Image Invariant to Illumination Colour, Strength, and Shading

Mark S. Drew and Amin Yazdani Salekdeh
School of Computing Science, Simon Fraser University, Vancouver, B.C., Canada V5A 1S6

ABSTRACT
We present here a method that makes use of multispectral image data and generates a novel “photometric-invariant multispectral image” for this type of data. For RGB, an invariant image has been constructed independent of the colour and intensity of the illuminant and of shading. To generate this image either a set of calibration images is required, or entropy information taken from a single image can be used to develop the parameters necessary to produce the invariant. Nonetheless, generating an invariant image remains a complex and error-prone task for RGB image data. For multispectral images, we show that photometric-invariant image formation is in essence greatly simplified. One of the requirements for forming an invariant is the necessity of narrowband-sensor sensors. Here this is the case, and we show that with the simple knowledge of peak sensor wavelengths we can generate a high-D multispectral invariant. The PSNR is shown to be high between the respective invariant multispectral features for multispectral images taken under different illumination conditions, showing lighting invariance for a per-pixel measure; and the s-CIELAB error measure shows that the colour error between the 3-D colour images used to visualize the output invariant high-D data is also small.

Keywords: multispectral; illumination invariant; lighting invariance; color; color constancy; high-D.

1. INTRODUCTION
This paper presents a novel approach that makes use of hyperspectral or multispectral image data and generates a new “photometric-invariant multispectral image” for this type of data.

While RGB colour is useful in recognizing or identifying objects and surfaces, it is of course dependent on the illumination incident on the surface. This has led researchers to develop a number of features that are to some extent invariant to the colour or intensity of the illuminant, and to shading as well if possible. Most of these invariant features use information from all pixels in the image in their construction. For example, suppose we normalize an image by dividing Red by its mean value, and similarly for Green and Blue: it can be proved that under certain assumptions this will indeed remove the effect of lighting, and if we go over to a chromaticity representation – colour without magnitude – then we can to some extent remove shading as well and arrive at surface colours that provide an intrinsic property of the object itself rather than depending on the sensor and the conditions under which the object is viewed.

Recently, however, a new type of invariant has been constructed that works using only the information at each pixel. To generate this invariant, sensor (camera RGB) information is required in the form of a set of calibration images, taken under different illumination conditions. One important benefit of the invariant is that since illumination is effectively removed, so are shadows. But for a non-laboratory setting, calibration images are not available and instead entropy information taken from a single image can be used to develop the parameters necessary to produce an invariant image. Nonetheless, generating an invariant image remains a complex and error-prone task for RGB image data.

Using multispectral image capture instead of coarsely sampling scene information using broadband camera sensors samples the signal much more finely. The utility of this information is the possibility of spectral reconstruction of an entire scene and consequently more accurate reproduction and other uses such as identification of surface materials. In terms of photometric-invariant image formation, for hyperspectral or multispectral data, although the amount of data is increased substantially the situation is in essence greatly simplified: one of the requirements for forming a strong invariant is the necessity of either making use of a narrowband-sensor camera or of generating an approximately narrowband set

Further author information: (Send correspondence to Mark S. Drew)
Mark S. Drew: E-mail: mark@cs.sfu.ca, Telephone: 1 778 782 4277
Amin Yazdani: E-mail: ayazdani@cs.sfu.ca, Telephone: 1 778 782 4277
of sensors. In this paper, we usually indeed already have narrowband multispectral sensors, and also know the peak wavelength for each of these. This situation simplifies the generation of an invariant and we show that with this simple knowledge of peak wavelengths we can generate multispectral data from the input such that the algorithm output is to a great degree invariant to illumination colour and strength, and has shading removed. In fact, the invariant produced is a good deal more effective in removing the effects of shading and lighting because of the greatly increased information available compared to 3-camera-sensor images. The PSNR is found to be high between the respective invariant multispectral features for multispectral images taken under different illumination conditions, showing lighting invariance for a per-pixel measure; and using the s-CIELAB error measure that takes into account the spatial characteristics of colour error we show as well that difference between the 3-D colour images used to visualize the output invariant high-D data is also small.

2. MULTISPECTRAL IMAGE FORMATION

To begin with, let us briefly state the model we adopt for multispectral image formation.

Let \( \rho \) be a 31-vector for wavelength \( \lambda = 400 \) to 700nm at 10nm intervals, determined by sampling scene spectra using a narrowband spectrophotometer such as the PhotoResearch PR-650 or the SpectraCube.

Let us assume that the scene illuminant is constant across the scene, and is given by Planck’s Law, in Wien’s approximation:

\[
P(\lambda) = c_1 \lambda^{-5} \exp \left( -\frac{c_2}{\lambda T} \right)
\]

with constants \( c_1 \) and \( c_2 \), where temperature \( T \) characterizes the lighting colour. Specifically, \( c_1 = 3.74183 \times 10^{16} \); \( c_2 = 1.4388 \times 10^{-2} \) (note that for these units, \( \lambda \) is in meters).

Let a multiplicative shading term at a pixel be denoted \( \sigma \) (and this could be Lambertian shading — normal vector dotted into normalized lighting direction).

We shall not make use of constant \( c_1 \) above, but instead use an overall multiplicative light intensity scalar \( I \). As well, let us denote by \( S(\lambda) \) the surface spectral reflectance function.

For completeness, we address not only matte reflectance but also a possible specular contribution, in a spectral image formation model. Suppose the strength of a specular term is given by scalar \( \beta \). If we adopt the Neutral Interface Model, here moved into the spectral domain, the spectrum for the specular part of the light reflected from a dielectric is simply given by that of the light itself.

Finally, assuming that our spectrophotometer is sufficiently narrowband, then each sensor is effectively a delta-function, filtering out all wavelengths except \( \lambda_k \), for the peak wavelength of the \( k \)th sensor. Since the measuring instrument may in fact provide different gain factors in each spectral channel \( k = 1..31 \), we should also include in our model an unknown (although in principle measurable) 31-vector \( q_k \) representing the measurement gain in each channel. [Note: We do not need to measure or know this vector, for this work.]

Hence, overall, we model the multispectral signal as:

\[
\rho_k = \sigma I \frac{1}{\lambda_k^5} \exp \left( -\frac{c_2}{\lambda_k T} \right) S(\lambda_k) q_k + \beta I \frac{1}{\lambda_k^5} \exp \left( -\frac{c_2}{\lambda_k T} \right) q_k
\]

Combining both terms, and defining \( \alpha \equiv \beta / \sigma \), we thus have an exponential light-colour term times a factor

\[
\sigma I [S(\lambda_k) + \alpha]
\]

Clearly, we can derive a signal which is invariant to light intensity \( I \) and shading \( \sigma \) by dividing values for every signal channel \( k \) by one particular value of one of the 31 channels, thus forming a type of chromaticity. But to not place undue confidence in any particular channel \( k \), let us instead instead divide by the geometric mean of the \( \rho \) values (cf. Ref.[9]): let us define a chromaticity \( \chi_k \) as dividing by the geometric mean value

\[
\rho_M \equiv \left( \prod_{j=1}^{31} \rho_j \right)^{(1/31)}
\]
so that we generate a modified type of band-ratio chromaticity (colour without magnitude) via

$$\chi_k = \rho k / \rho M, \quad k = 1..31$$  \hspace{1cm} (4)

Then substituting our signal formation model (2), we arrive at the form

$$\chi_k = \left( \frac{S(\lambda_k) + \alpha}{S(\lambda_M) + \alpha} \right) \frac{\lambda_k^5}{\lambda_M^5} \exp \left( - \frac{c_2}{T} \left[ \frac{1}{\lambda_k} - \frac{1}{\lambda_M} \right] \right) q_k / q_M$$  \hspace{1cm} (5)

or

$$\chi_k = \left( \frac{s_k + \alpha}{s_M + \alpha} \right) w_k \exp \left( \frac{1}{T} [e_k - e_M] \right) \phi_k$$  \hspace{1cm} (6)

where we simplify the expressions by defining some short-hand notations as follows:

$$v_k = \lambda_k^5; \quad w_k = v_k / v_M; \quad \phi_k = q_k / q_M; \quad s_k = S(\lambda_k);$$

$$e_k = -c_2 / \lambda_k; \quad v_M = \left\{ \prod_{j=1}^{31} v_j \right\}^{1/31}; \quad q_M = \left\{ \prod_{j=1}^{31} q_j \right\}^{1/31}; \quad e_M = (1/31) \sum_{j=1}^{31} e_j$$

and we define an effective geometric-mean-respecting value $$s_M$$ by setting

$$(s_M + \alpha) = \left( \prod_{j=1}^{31} (s_j + \alpha) \right)^{1/31}$$

Since the light-colour information, controlled by temperature $$T$$, is contained in an exponential, it makes sense to take logarithms. Then rewriting, the log of the chromaticity now becomes

$$\log \chi_k = \log \left( \frac{s_k + \alpha}{s_M + \alpha} \right) + \log w_k + \log \phi_k + \frac{1}{T} (e_k - e_M)$$  \hspace{1cm} (8)

Of course, any version of a chromaticity is meant to hold “colour” information with magnitude normalized away, and hence chromaticity signal values are not independent. Specifically, here the log values above all inhabit a plane orthogonal to the unit vector $$(1, 1, ..., 1)$$ in 31-space. And in fact each separate term in eq. (8) is orthogonal to that unit vector.

In order to remove the effect of light colour, we need to project the spectral information into a subspace orthogonal to the direction of the lighting term $$(e_k - e_M)$$. Previously, in RGB space, we had to calibrate a camera system in order to recover this lighting direction; alternatively we could make use of the lighting-change information in an image to use an entropy measure to determine this direction. But here, in the multispectral domain, we have a very special situation in that every signal channel is narrowband and is assumed given on specific, known wavelengths. Consequently we already know the special direction $$(e_k - e_M)$$ orthogonal to which to project in order to remove light colour information: $$(e_k - e_M)$$ is simply given by $$-c_2 / \lambda_k - c_2 / \lambda_M$$ and is known.

Also, the term $$\log w_k$$ in (8) is also already known, for the same reason. Hence the only unknown sensor term in eq. (8) is the term in sensor gains, $$\log \phi_k$$.

If the neutral-interface model we have adopted worked perfectly, then we could in fact also derive the value of a projection of this sensor-gain 31-vector $$\log \phi_k$$ provided we could correctly identify at least one specularity in the image: For at a pure specularity (i.e., the same as imaging the lighting itself), the first term in eq. (8) goes to the zero 31-vector, as the specular-strength value $$\alpha$$ goes to infinity. So in fact at such a specularity, using the known other terms, we could determine the projected part of the $$\log \phi_k$$ vector, pertaining to the geometric-mean-divided gain chromaticity vector $$q_k / q_M$$. However we found in practice that specular multispectral 31-vectors in a multispectral image are not, in fact, sufficiently free of surface reflectance effects from $$s_k$$ to allow correct identification of the specular point. Therefore we must rest content
with obtaining all terms except \( \log \phi_k \) in eq. (8) (barring discovering this term via a calibration). Nonetheless, we can still obtain a multispectral illumination invariant without knowing the term \( \log \phi_k \) — the effect is that this point becomes the origin in the invariant space.

We have removed the effect of light intensity and shading by going over to a chromaticity space (thus basically descending from a 31-D space to a 30-D one, since chromaticity components \( \chi_k \) are not independent but live on a 31-D plane). To remove the effect of lighting colour, encapsulated in temperature \( T \), we can project the data, written as 31-D values in eq. (8), into a subspace orthogonal to the lighting-change direction \((e_k - e_M)\). For suppose the \((31 \times 31)\) projection matrix (symmetric, idempotent, with a single non-zero eigenvalue) in light-change direction \((e_k - e_M)\) is given by matrix \( P_e \).

Then we can remove the effect of temperature \( T \) by pre-multiplying eq. (8) by the projector onto the orthogonal subspace, \( P_e^\perp = I_{31} - P_e \), where \( I_{31} \) is the \( 31 \times 31 \) identity matrix.

Thus we obtain

\[
P_e^\perp \log \chi_k = P_e^\perp \log \left( \frac{s_k + \alpha}{s_M + \alpha} \right) + P_e^\perp \log w_k + P_e^\perp \log \phi_k
\]

We know the second term on the right because we know the narrow band signal wavelengths being measured. At a non-specular pixel, the first term on the right is the matte characteristic of the surface material itself, and thus forms an illumination invariant. Hence, in this paper, we take the left hand side of eq. (9) to define our illumination-invariant multispectral image. In the following, we shall no longer consider the specular term but instead concentrate on matte components of any image.

We show below that this simple formulation works very well on real, 31-D, multispectral data. To better understand the strengths and limitations of the model (9), first we consider synthetic images, where we know the ground truth, and then in §4 go on to investigate experimental data.

### 3. SYNTHETIC IMAGES

Let us generate a simple synthetic scene, but using full spectra, so as to be able to completely control reflectances and illuminants in the spectral image. Such a synthetically generated 31-band image can then serve as a testing ground for our analysis. Suppose we utilize a set of three shaded spheres as in Ref.[9], with surface colors equal to patches 1, 4, and 9 (dark skin, moderate olive green, moderate red) of the Macbeth ColorChecker.\(^\text{11}\) As illuminant, suppose we take a blueish light, a Planckian with temperature \( T = 10500^\circ K \). For generating RGB images we utilize the set of three camera sensors for the Kodak DCS420 colour still camera, shown in Fig. 1(a), and for the set of 31 spectral gain values \( q_k \) we use those shown in Fig. 1(b) [these are taken to be simply the sum of the camera curves raised to a power]. Let the direction to the light be \( \{0.5, 1.0, 3.0\} \), normalized.

Then the matte image in a Lambertian model is as in top row of Fig. 1(c), displayed in the sRGB (gamma-corrected) colour space\(^\text{12}\) (note the effect of the very blue light and the blue-red mix in the ostensibly blue camera sensor). Note that we show images in RGB, but in fact the synthetic image is generated in 31-D. The Lambertian shading image is shown in the top of Fig. 1(d).

Now since we mean to show that a 31-dimensional illumination invariant image is derivable via eq. (9), in the bottom row of Fig. 1(c) we show the same three surfaces, but illuminated by a reddish Planckian for temperature \( T_2 = 2800^\circ K \). And since we mean to demonstrate that shading can be removed, in this case let the lighting direction be \( \{-1.0, -2.0, 3.0\} \), normalized, so that shading is different than that for the blue light and is as in the bottom of Fig. 1(d).

Now applying the invariant image formulation (9), we derive a 31-D illumination invariant image. In terms of the geometric-mean chromaticity, we can display the resulting image, as in Fig. 1(e). Here we have generated a colour image by applying the human vision colour-matching functions\(^\text{5}\) to generate tristimulus XYZ values and then gone over to the standardized sRGB colour space.

To compare, in Fig. 1(f), we show this chromaticity for the original, spectral, image, and note how differently the spheres appear under different lighting. Since the invariant has in fact removed lighting, we show the chromaticity for the invariant regressed back onto that for the top of Fig. 1(e) [as was done in Ref.[13] in order to put back some lighting]. Naturally, we cannot retrieve any information beyond the terminator for each of these sets of three spheres (where shading \( \sigma \) goes to zero).
Figure 1. (a): Camera sensors; (b): 31-D sensor gains $q_k$; (c): matte image, top row under blue Planckian $T = 10500^\circ K$ bottom under red $T = 2.800^\circ K$; (d): shading under each of two lights; (e): illumination-invariant chromaticity image; (f): chromaticity for original image before invariant processing.
Finally, we can ask how well we do in deriving a 31-D invariant by calculating the mean PSNR for the spectral invariant, in the lighted (shading not zero) areas. We find a mean PSNR value of 73.87 dB for this synthetic example: since there is in fact so much data available, an illumination invariant (at least in this noise-free synthetic example) does better than typical results using only 3-sample RGB data (cf. Ref.[4]).

4. MEASURED MULTISPECTRAL IMAGES

Let us apply the above analysis to measured multispectral image data: suppose that for visualization purposes we make RGB colour out of the spectral database image data collected in Ref.[14]. That is, we generate XYZ tristimulus values from the 31-D data by applying the human vision colour-matching functions and then go over to nonlinear (i.e., gamma-corrected) sRGB colour space pixel values by applying the XYZ to sRGB transformation in the sRGB standard. Let us consider the 31-D image visualized in Fig. 2(a), displayed in sRGB colour space. This image is generated from 31-D spectral data imaged under standard daylight D75.

Now, since we know the actual imaging daylight and in fact we have complete high-D data, it is straightforward to change the lighting by dividing each spectral component by the light-vector value for the original illuminant, D75, and multiplying by a second light. Here we go over to a much more reddish-looking light, standard illuminant D48, whose correlated colour temperature $T$ is 4800°K. Now the visualization appears as in Fig. 2(b), i.e., with colours substantially shifted to the red.

Applying the invariant processing eq. (9), we derive illumination invariant 31-D multispectral images for each of the lights. Figs. 2(c,d) show how the invariant appears in RGB, under each of the lights (where again we have regressed colours back onto those that do not have the lighting removed, Fig. 2(a), as in Ref.[13]).

For the 31-D invariant spectral images derived from each of the lights, we can calculate the PSNR between each of the 31-D pixel spectra, between the two output invariant 31-D images. Here, we find a mean value of 39.09 dB, showing that the invariant produced is excellent: changing the lighting does not greatly affect the invariant result.

We can further characterize the strength of the invariant by examining the error between the colour visualizations in Figs. 2(c,d): we can ask how different is the perceived colour between Fig. 2(c) and Fig. 2(d) by applying the s-CIELAB colour-error metric. Rather than being simply a per-pixel error measure as is PSNR, this colour error metric takes into account human perception of colour difference including the effect of spatial visual acuity in the different opponent colour channels. Here, we find a very small colour difference, with maximum $\Delta E$ value 0.67, i.e., less than 1 just-noticeable difference value.

While invariant images Figs. 2(c,d) appear to retain somewhat the shadowing present in the original images, a close examination shows that this is not in fact the case: in the output images Figs. 2(c,d) the shading inside locations that appeared in shadow is in fact correctly removed, and what now appears as shading is due to specularities that appear in the original data. Fig. 2(e) identifies two points on the original image — the blue circle is in-shadow and the red one is in-light, both for the same surface material. In Fig. 2(f) the solid curves are for the original spectral data, blue for in-shadow and red for in-light. The dashed curves are for the corresponding invariant-spectral-image output — illuminant-invariant processing has brought the two curves into a much closer match, thus showing that that the 31-D invariant does indeed substantially remove the effects of lighting and shading.

Fig. 3 shows both originals and invariant processing for a some of the further 19 images in the database used: the mean PSNR between invariant curves for all 20 spectral images under different lights varies from 37.78dB to 41.06dB, with mean value 39.34, showing that the multispectral invariant processing proposed here is very effective for removing lighting effects for this high-D data. For colour error in RGB visualizations of the invariant between two standard lights, D75 and D48, the maximum spatial-CIELAB $\Delta E$ values range from 0.145 to 1.37 units, i.e. quite a small perceptual difference.

5. CONCLUSION

In this paper we have set out a novel method for producing an illumination-invariant spectral quantity that effectively removes lighting effects from high-D spectral image data. Overall, the method is shown to be quite successful at removing effects of illuminant strength and colour, and of shading.

Specularities are seen to remain a problem for the method proposed. Of course, if we knew or measured the spectral function for the illuminant, then simple subtraction-based methods can in principle remove specularities, customarily
making the assumption that illumination is equi-energy, uniformly “white” across the spectrum.\textsuperscript{17,18} However, making use of the special knowledge of the illuminant spectrum is not in fact generating an illumination “invariant” – i.e., a quantity which is independent of lighting, without knowledge of the lighting; and indeed here we make no assumption about the lighting. In future work we shall address this issue by examining schemes for determining the spectral-gain offset vector $\log \phi_k$, and hence that in effect the specular point in 31-D spectral space is known, so that matte colours might possibly be determined by knowing to what extent specular content is included in spectral image pixel data.

Another interesting direction would be to examine how breaking the assumption of narrowband sensors, and using broader-band sensors, would affect the invariant generated here. Another approach could be to apply spectral sharpening to such broader-band sensors. As well, the model used assumes a Planckian light, whereas the experiments with real data use standard daylights in a light-box. The question remains, then, to what degree departing from the Planckian model affects results in general.

Finally, an intriguing question arises from the present work: whether indeed increasing the number of colour channels helps in estimating the scene illumination in general, and not just in the context of the special narrowband signal measurements used here.

REFERENCES

Figure 2. (a): Visualization of 31-D measured multispectral data, under illuminant D75; (b): visualization under redder illuminant D48; (c): invariant derived from D75 image; (d): invariant derived from D48 image; (e): two points, in-shadow (blue circle) and in-light (red circle); (f): plot of in-shadow (blue) and in-light (red) 31-D pixel data — solid curves are original data and dashed curves are after invariant processing.
Figure 3. Left: Original image; Right: Invariant