A Standardized Workflow for Illumination-Invariant Image Extraction

Mark S. Drew, Muntaseer Salahuddin, and Alireza Fathi; School of Computing Science, Simon Fraser University, Vancouver, British Columbia, Canada V5A 1S6 {*mark,msalahud,alirezaf*} *@cs.sfu.ca*

Abstract

The illumination-invariant image is a useful intrinsic feature latent in colour image data. The idea in forming an illumination invariant is to postprocess input image data by forming a logarithm of a set of chromaticity coordinates, and then project the resulting 2-dimensional data in a direction orthogonal to a special direction, characteristic of each camera, that best describes the effect of lighting change. Lighting change is approximately simply a straight line in the log-chromaticity domain; thus, forming a greyscale projection orthogonal to this line generates an image which is approximately independent of the illuminant, at every pixel. One application has been to effectively remove shadows from images. But a problem, addressed here, is that the direction in which to project is camera-dependent. Moreover, preprocessing with a spectral sharpening transform to linearly transform the sensor curves to more narrowband ones greatly improves shadow attenuation, but sharpening is also camera-dependent and we may not have information on the camera. So here we take a simpler approach and assume that every input image consists of data in the standardized sRGB colour space. Previously, this assumption has led to the suggestion that the built-in mapping of sRGB to XYZ tristimulus values could be used by going on to sharpen the resulting XYZ and then seeking for an invariant. Instead, here we sharpen the sRGB directly and show that performance is substantially improved this way. This approach leads to a standardized sharpening matrix for any input image and a fixed projection angle as well. Results are shown to be satisfactory, without any knowledge of camera characteristics.

1. Introduction

The illumination invariant image is formed from image data by taking the logarithm of band-ratio chromaticity colour coordinates, and then projecting in a certain direction [1]. The input colour data is 3-dimensional RGB, and the chromaticity is effectively 2D colour. Projecting in a 2-space direction generates a 1D, greyscale image. If the direction is chosen with care, the resulting greyscale image is quite independent of the lighting at each pixel, therefore forming an illumination invariant. The cleverness of the invariant is that it is formed at each pixel independently, with no global image processing required.

The special direction for projection is that which is orthogonal to the direction that is followed along, in the 2-space, as the lighting changes (within a simplified model). Since lighting is thus removed, as a particular case shadows are also removed, or at least greatly attenuated [2, 3].

Since in fact we are projecting onto a line through the origin in a colour *2-space*, we need not think of the result of projection as merely a 1D, greyscale image: we do have as well, after all, a 2D coordinate position on that line, so we could state the projection answer as a 2D chromaticity [4]. Projection removes the lighting, but this can then be partially added back, by shifting the

chromaticity projection line so as to make the chromaticity for bright pixels match that for bright pixels in the input image. So projection does not have to completely remove colour.

Once we indeed have an invariant image, we can go on to remove shadows by comparing edges in the original to edges in the invariant image. Removing or blending edges where these two edge maps disagree provides a strategy for re-integrating the edge map back into a full-colour, shadowless RGB image [2, 5].

While the idea for finding the invariant is fairly simple, carrying out finding the proper direction in which to project is not necessarily as straightforward. In [2], the camera itself was calibrated, in this invariant image sense, by utilizing a set of images of a colour target, under different illuminants, to find the best 2D direction characterizing lighting change. In [3], evidence in the image itself was used to discover the correct direction orthogonal to the lighting change direction. There, it was argued that such a projection direction is best described as that leading to a minimum-entropy distribution in greyscale values.

Here, we would like to argue that it is possible to do a good enough job in finding the invariant image by simply assuming the input data to live in the standardized sRGB colour space and sharpening that space. In this way, we are not tied to finding the invariant projection for a particular camera, or using the data in a particular image, and can develop a standardized workflow that can be applied directly to any input image. Of course, deriving an invariant that is sensor- or image-adaptive instead, as originally conceived, will likely work better than a one-size-fits-all approach, but here we show that results are indeed adequate using an approach applying the same transform to any image – e.g., shadows are principally attenuated, no matter what the input image.

Previously, a similar approach was tried by recognizing that the sRGB standard [6] contains not only a mapping from nonlinear to linearized colour values, but also a relationship from the sRGB gamut to corresponding XYZ values via matrixing. Thus it was proposed [7] that input images could be assumed to be in nonlinear sRGB colour space, linearized to linear-sRGB, and then transformed to XYZ. Then, in XYZ, the XYZ curves themselves could be sharpened. The results for shadow removal were indeed better than simply removing gamma-correction. However, the invariant direction still was found using the entropy method of [3], so a fully standardized method was not developed.

We would argue, moreover, that going from linear-sRGB to XYZ is in itself counterintuitive, in that the sRGB colour-matching functions are close to sharpened colour-matching curves already [8], so that going over to XYZ curves is a kind of broadening transform. Following this with a further sharpening of the XYZ curves is not as direct as simply sharpening the sRGBs themselves, as proposed here. And, we show that the new idea, of sharpening the sRGB data, provides better performance for producing an illumination invariant.

Moreover, once we decide that all data is sRGB, and we provide a sharpening transform for sRGB, we can in fact find the best projection direction simply using synthetic data and then apply the same transform and projection once and for all to any input image. We show that this simple strategy produces usable results, within the application of shadow removal.

In §2 we compare the strategy of sharpening XYZ data, for input nonlinear-sRGB images, to the new approach of sharpening the sRGB data directly, and show that a better invariant (more invariant to lighting change) arises from the latter approach. In §3.1 we extract the illumination invariant from measured, nonlinear input data for images of the Macbeth chart across 105 different illumination environments. Applying the standardized illumination invariant extraction scheme presented here produces images much more independent of lighting change. And in §3.2 we apply the new method to the problem of reducing or removing shadows from imagery, by generating an invariant image from input colour images, making use of the new standardized method.

2. Sharpening XYZ versus Sharpening sRGB

Sharpening XYZ

The first approach to a standardized sharpening and projection scheme was to sharpen XYZ values arising from input nonlinear-sRGB images [7]. Here we propose re-examining this approach, but here we sharpen colour-patch data directly rather than sharpen the XYZ colour-matching curves — that is, we take a maximum-prescience approach rather than a maximum-ignorance one. But what colour data should we utilize? As a set of fairly generic inputs, suppose we simply use the 24 patches of a Macbeth ColorChecker [9], with synthetic values for tristimulus values under Planckian lights [10]. However, here we are aiming at the idea of starting with sRGB data; therefore we first transform the resulting XYZ values back to linear-sRGB colour space, and thence back to XYZ again. The thought here is that the transform from XYZ to sRGB [6] may involve clipping to the range [0,1], and we wish to take that into account.

Therefore we generate a set of synthetic images of the Macbeth chart, formed under 9 Planckian lights for temperatures T=2,500°–10,500° in 1,000° intervals. We define the synthetic data in XYZ coordinates rather than in sRGB so that we have meaning and generality for the data. Taking the resulting XYZ triples to linear-sRGB colour space does turn out to involve some clipping. Then we take the data back to XYZ space.

Finally, we wish to consider an invariant in a colour 2-space, and here we make use of log-chromaticities formed as the logarithm of ratios of the XYZ to their geometric mean [11]:

$$\log x_k \;=\; \log\left[\{X,Y,Z\}/(X \cdot Y \cdot Z)^{1/3}\right] \tag{1}$$

This generates 3-vector quantities but, in fact, in the log space every such 3-vector lies in the plane orthogonal to the unit vector $\boldsymbol{u} = (1/\sqrt{3})(1,1,1)^T$; thus only two coordinates are independent. We can rotate into that plane (cf. [4]) by forming a 2-vector $\boldsymbol{\chi}$ by making use of the 2×3 rotation matrix $\boldsymbol{U}^T$ equal to the orthogonal matrix factorizing the projector onto the subspace perpendicular to $\boldsymbol{u}$:

$$\begin{aligned} P^\perp &= I - \boldsymbol{u}\,\boldsymbol{u}^T, \\ P^\perp &= \boldsymbol{U}\,\boldsymbol{U}^T, \quad \boldsymbol{U} \text{ is } 3 \times 2 \\ \boldsymbol{\chi} &= \boldsymbol{U}^T \log \boldsymbol{x} \end{aligned} \tag{2}$$

A plot of the resulting 2D colour coordinates in Fig. 1(a) shows that, rather than forming straight lines as expected, we see some curvature in the plots as lighting changes. If we center the data by subtracting the mean χ vector for each colour patch, we would like to see as close as possible to a single straight line through the origin, for the purposes of forming a lighting invariant: a single straight line would indicate that, in Fig. 1(a) we could simply project in a direction orthogonal to the direction of that line and effectively eliminate the influence of lighting on the feature. I.e., we could generate a 1D greyscale illumination invariant.

However, in Fig. 1(b), for mean-subtracted data, we instead see that the data is fairly spread out. We can discount the effect of outliers to a degree by finding the best slope using a robust statistical method [12], but still, we find the data has a correlation coefficient R of only 0.605 — not an excellent indicator of straight-line behaviour.

Therefore we consider *sharpening* the colour-patch data [13], in order to make the illumination-invariant image formation model [1] more applicable, since the theory behind the model requires quite sharp camera sensors. We thus make use of the data-based sharpening method [13] to determine a sharpening matrix T; we choose the synthesized data under the most red and the most blue lights, and find the best least-squares matrix transforming one into the other. The sharpening matrix T is the set of eigenvectors of the least-squares transform. Fig. 1(c) shows that sharpening does indeed straighten out the log-chromaticity plots; for mean-subtracted data in Fig. 1(d), we now find a correlation coefficient R=0.764, a much improved value.

Sharpening sRGB

The objective of this paper is to determine whether sharpening sRGB values themselves can produce a better illumination invariant than can sharpening XYZ values. Therefore now we compare how sRGB log-chromaticities fare under lighting change — can we sharpen analogues to eq. (1) constructed from linear-sRGB values and arrive at a better invariant?

Firstly, we examine how sRGB itself does in forming an invariant. We plot linear-sRGB for the synthetic images under 9 different lights, with results shown in Fig. 2(a). We see that sRGB coordinates do indeed form straighter lines than do XYZ coordinates (in Fig. 1(a)). For mean-subtracted values, in Fig. 2(b), we find a correlation coefficient R=0.837, already better than sharpened XYZ notwithstanding outliers created by clipped values. (We use a generalized logarithm [7], not a logarithm, to avoid the log of zero.)

Data-based sharpening in this case actually makes the correlation coefficient worse: Table 1 shows that applying sharpening results in R=0.630 (the mean-subtracted data is somewhat spread out).

However, if we make use of a white-point preserving data-based sharpening [14], then R is improved: R=0.877, the highest value found so far.

Optimized sRGB Sharpening Transform

The result above is encouraging, since it indicates that sharpening sRGB does indeed produce the best illumination invariant, the result we argue for in this paper. However, while the result is good, it could be better.
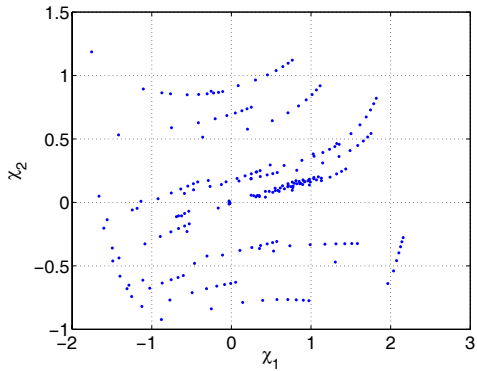
Sensor sharpening simply has the objective of concentrating energy in each sensor in its associated colour band. However, here we have a specific objective: producing the best invariant coordinate. Therefore we adopt the optimization strategy in [15],
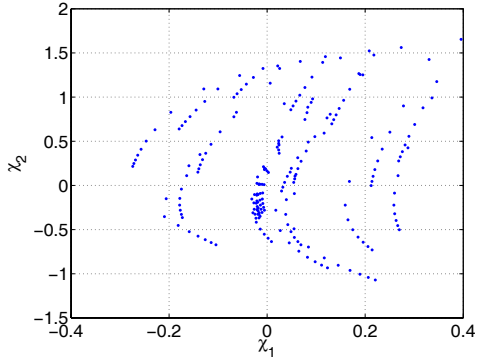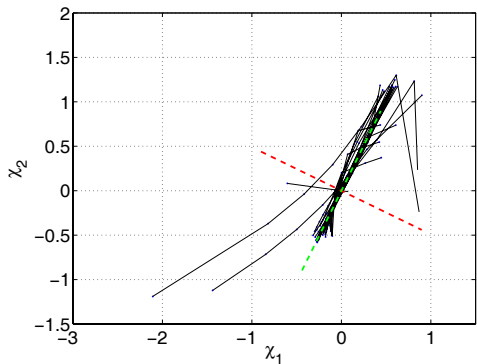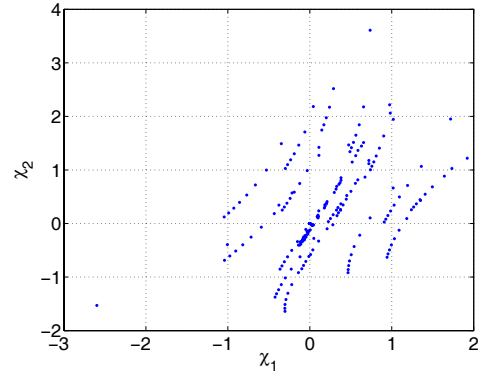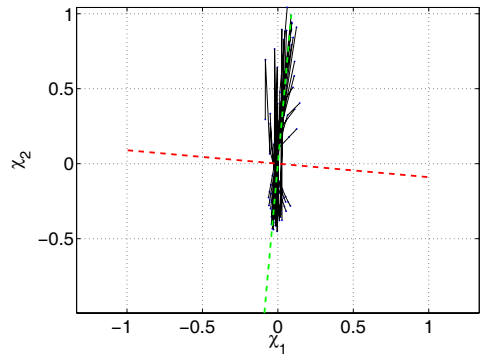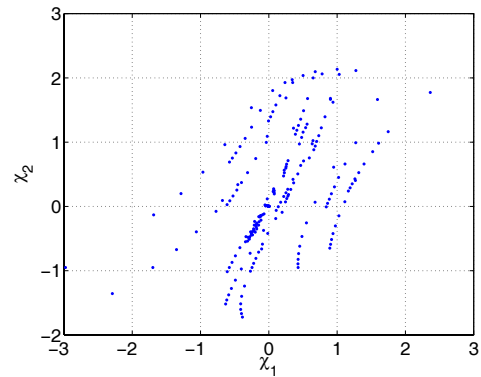
(a)
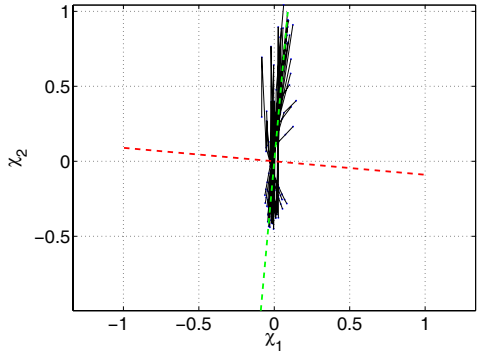
(b)

(c)

(d)

Figure 1. *Log-chromaticity XYZ coordinates for Macbeth patches, as light changes. (a): χ vectors; (b): Mean-subtracted values: best (robust) direction in green, orthogonal direction in red, R=0.605. Black lines joining data points are for each colour patch, as lighting changes. (c): Sharpened XYZ; (d): Sharpened, mean-subtracted: R=0.764.*
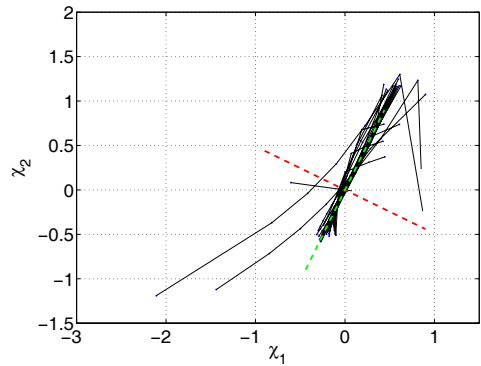
Figure 2. *Log-chromaticity sRGB coordinates for Macbeth patches, as light changes. (a): χ vectors; (b): Mean-subtracted values: R=0.837. (c): Sharpened sRGB; (d): Sharpened, mean-subtracted: R=0.877.*

which aims specifically at finding the best sensor transform T that minimizes the spread of the lines plotted in a mean-subtracted log-chromaticity space. The optimization also insists on non-negative results, after applying the colour transform T.

Applied to the sRGB data, we find the following transform

$$T = \begin{pmatrix} 0.9968 & 0.0228 & 0.0015 \\ -0.0071 & 0.9933 & 0.0146 \\ 0.0103 & -0.0161 & 0.9839 \end{pmatrix} \quad (3)$$

which is post-multipled times row vectors. The sRGB data forms quite straight lines, now, in the transformed space, as shown in Fig. 3(a).
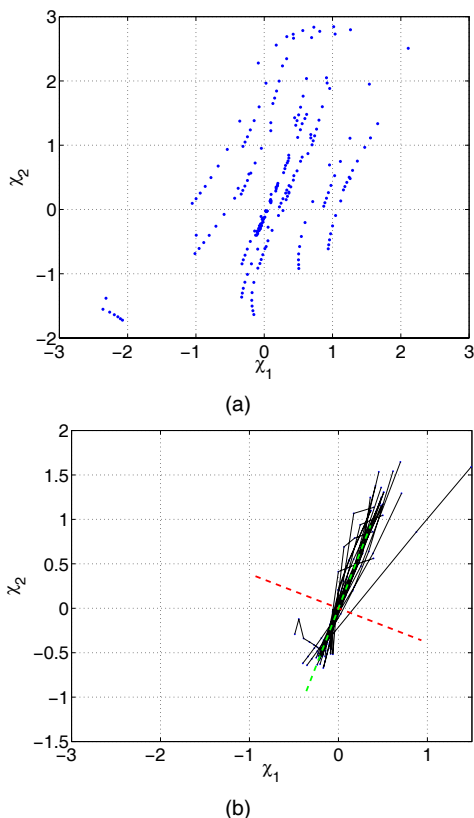

(a)

(b)

Figure 3. *Optimized log-chromaticity sRGB coordinates. (a): χ vectors; (b): Mean-subtracted values: R=0.920.*

For the mean-subtracted data in Fig. 3(b), we now find the improved correlation coefficient value: 0.920. Thus we suggest adopting the linear-sRGB colour space transform matrix (3) as a *standard* colour transform. The direction for orthogonal projection found by a robust regression, shown in red in Fig. 3(b), is given by the 2D vector $e^{\perp}$ orthogonal to the lighting-change direction:

$$e^{\perp} = (0.9326, -0.3609)^T \quad (4)$$

Thus, overall, we argue here that as a standardized workflow for producing an illumination invariant image from an input colour image we proceed as follows:

Scheme	R
XYZ	0.605
XYZ$^{\#}$	0.764
sRGB	0.837
sRGB$^{\#}$	0.630
sRGB$^{\#}_{WPP}$	0.877
sRGB$_{T\_OPT}$	0.920

Table 1: Correlation coefficient values R for projection of mean-subtracted log-chromaticity, formed according to the method in columns "Scheme": from XYZ coordinates, and from sharpened XYZ; from sRGB, from sharpened sRGB, and from white-point preserving sharpened sRGB; and finally using an optimized tranform T from eq.(3) on sRGB coordinates.

Transform input image nonlinear sRGB to linear-sRGB.
Transform to sharpened colour space. I.e., if linear sRGB
   values are ρ, then $\rho^{\#} = T\rho$, where T is given by (3).
Form 2D log-chromaticity coordinates χ as in eq. (2) for
   sRGB values.
   I.e., form 2-vectors r via
   $r = \log \rho^{\#} - (1/3)\sum_{k=1}^{3} \log \rho_k^{\#}$,
   $\chi = U^T r$, using sharpened $\rho^{\#}$ values.
   E.g., use $U = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/\sqrt{6} \end{pmatrix}$
Project onto line perpendicular to lighting-change direction,
   using vector $e^{\perp}$ in (4).
Form 2D-colour from projected point by rotating back
   to a 3-vector using 3×2 matrix U.
Exponentiate to go back to non-log coordinates.
Move to chromaticity in an L_1 norm by dividing
   by $(R + G + B)$.

The above algorithm generates 3D colour, but only from values projected onto the projection line, so effectively 2D. Nonetheless, the 2D colour information can still be useful [4].

In the next section, we apply this algorithm to empirical images of the Macbeth chart, situated in various illumination environments, and show the efficacy of such a generic sharpening plus projection scheme.

3. Experiments
3.1 Invariant from Measured Chart Data

A set of various images that include a Macbeth chart in the scene were acquired under 105 different lighting conditions. [1] Fig. 4(a,b) shows two if these images, which we treat as nonlinear-sRGB. Forming the mean-subtracted log-chromaticities, we find that without any colour space transform the correlation coefficient is only R=0.775. Thus we would not expect to achieve a reliable illumination invariant without transforming the colour space.

Now if we apply the algorithm given above, applying transform T, we then achieve an R value of R=0.809. I.e., while an optimization applied to this data would do better, the pre-defined transform derived from synthetic data already does quite well. In

---

[1]These images are due to Prof. Graham Finlayson. The images are nonlinear; the camera used was a Nikon D70.
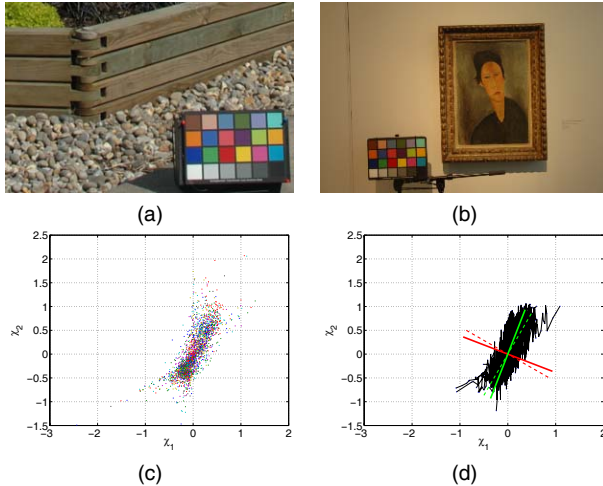
Figure 4. *Log-chromaticity sRGB coordinates for measured empirical Macbeth patches, for 105 different lighting conditions. (a,b): Samples of images. (c): Mean-subtracted values: R=0.775. (d): Sharpened, mean-subtracted: R=0.809.*

Fig. 4(d), we show the pre-determined projection line as a solid line, and the best-fit one for the actual data in a dashed line — the two are not far apart.

In the next section, we apply the standardized algorithm to ordinary images, with a view to testing the efficacy with respect to shadow removal in an invariant image.

3.2 Shadow Attenuation

Here we apply the standardized algorithm to a set of images acquired under a variety of illumination environments. We form the 2D chromaticity, both without and then with invariant image processing applied as described above. If the standardized approach to extraction of an illumination invariant does indeed work, we expect shadows to be attenuated, compared to in the original chromaticity image.

Fig. 5 shows several images that contain shadows. The effect in every case, over the cameras utilized, is to remove or at least reduce the effect of shadows. This is shown by displaying the chromaticity images with their edge-map overlaid: edges for shadows appear in the original, but not in the invariant version of the chromaticity. This can then be used to go on to remove shadows from images (see [16] and [5] for approaches to this task).

5. Conclusions

We have outlined a simple, standardized method to generate a reasonable illumination invariant, from input colour images. The method is based on the idea of simply treating every input image as inhabiting sRGB colour space, and transforming that space. The transformation is found by optimizing the lighting invariance for generic, synthetic data when taken to log-chromaticity space and projected into a 1D invariant. The invariant image itself can be understood as a 2D-colour chromaticity image. Experiments show that applied the standardized invariant extraction method generates reasonable independence to lighting, across conditions and cameras. The approach set out here might be usefully employed in place of a more rigorous, camera- and image-dependent method for extraction of an illumination invariant.

References

[1] G.D. Finlayson and S.D. Hordley. Colour constancy at a pixel. *J. Opt. Soc. Am. A*, 18(2):253–264, Feb. 2001. Also, UK Patent #2360660, "Colour signal processing which removes illuminant colour temperature dependency".

[2] G.D. Finlayson, S.D. Hordley, and M.S. Drew. Removing shadows from images. In *ECCV 2002: European Conference on Computer Vision*, pages 4:823–836, 2002. Lecture Notes in Computer Science Vol. 2353.

[3] G.D. Finlayson, M.S. Drew, and C. Lu. Intrinsic images by entropy minimization. In *ECCV 2004: European Conference on Computer Vision*, pages 582–595, 2004. Lecture Notes in Computer Science Vol. 3023.

[4] M.S. Drew, G.D. Finlayson, and S.D. Hordley. Recovery of chromaticity image free from shadows via illumination invariance. In *IEEE Workshop on Color and Photometric Methods in Computer Vision, ICCV'03*, pages 32–39, 2003.

[5] C. Fredembach and G.D. Finlayson. Hamiltonian path based shadow removal. In *16th British Machine Vision Conf, BMVC2005*, pages 970–980, 2005.

[6] International Electrotechnical Commission. Multimedia systems and equipment – colour measurement and management – part 2-1: Colour management – default RGB colour space – sRGB. IEC 61966-2-1:1999.

[7] G.D. Finlayson, C. Lu, , and M.S. Drew. Invariant image improvement by sRGB colour space sharpening. In *AIC 2005: The Tenth Congress of the International Colour Association*, 2005.

[8] G.D. Finlayson and S. Süsstrunk. Performance of a chromatic adaptation transform based on spectral sharpening. In *Eigth Color Imaging Conference: Color, Science, Systems and Applications.*, pages 49–55. Society for Imaging Science & Technology (IS&T)/Society for Information Display (SID) joint conference, 2000.

[9] C.S. McCamy, H. Marcus, and J.G. Davidson. A color-rendition chart. *J. App. Photog. Eng.*, 2:95–99, 1976.

[10] G. Wyszecki and W.S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulas*. Wiley, New York, 2nd edition, 1982.

[11] 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.

[12] P. J. Rousseeuw and A. M. Leroy. *Robust Regression and Outlier Detection*. Wiley, 1987.

[13] 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.

[14] G.D. Finlayson and M. S. Drew. Constrained least–squares regression in color spaces. *Journal of Electronic Imaging*, 6:484–493, 1997.

[15] M.S. Drew, C. Chen, S.D. Hordley, and G.D. Finlayson. Sensor transforms for invariant image enhancement. In *Tenth Color Imaging Conference: Color, Science, Systems and Applications.*, pages 325–329. Society for Imaging Science & Technology (IS&T)/Society for Information Display (SID) joint conference, 2002.

[16] G.D. Finlayson, S.D. Hordley, C. Lu, and M.S. Drew. On the removal of shadows from images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 28:59–68, 2006.
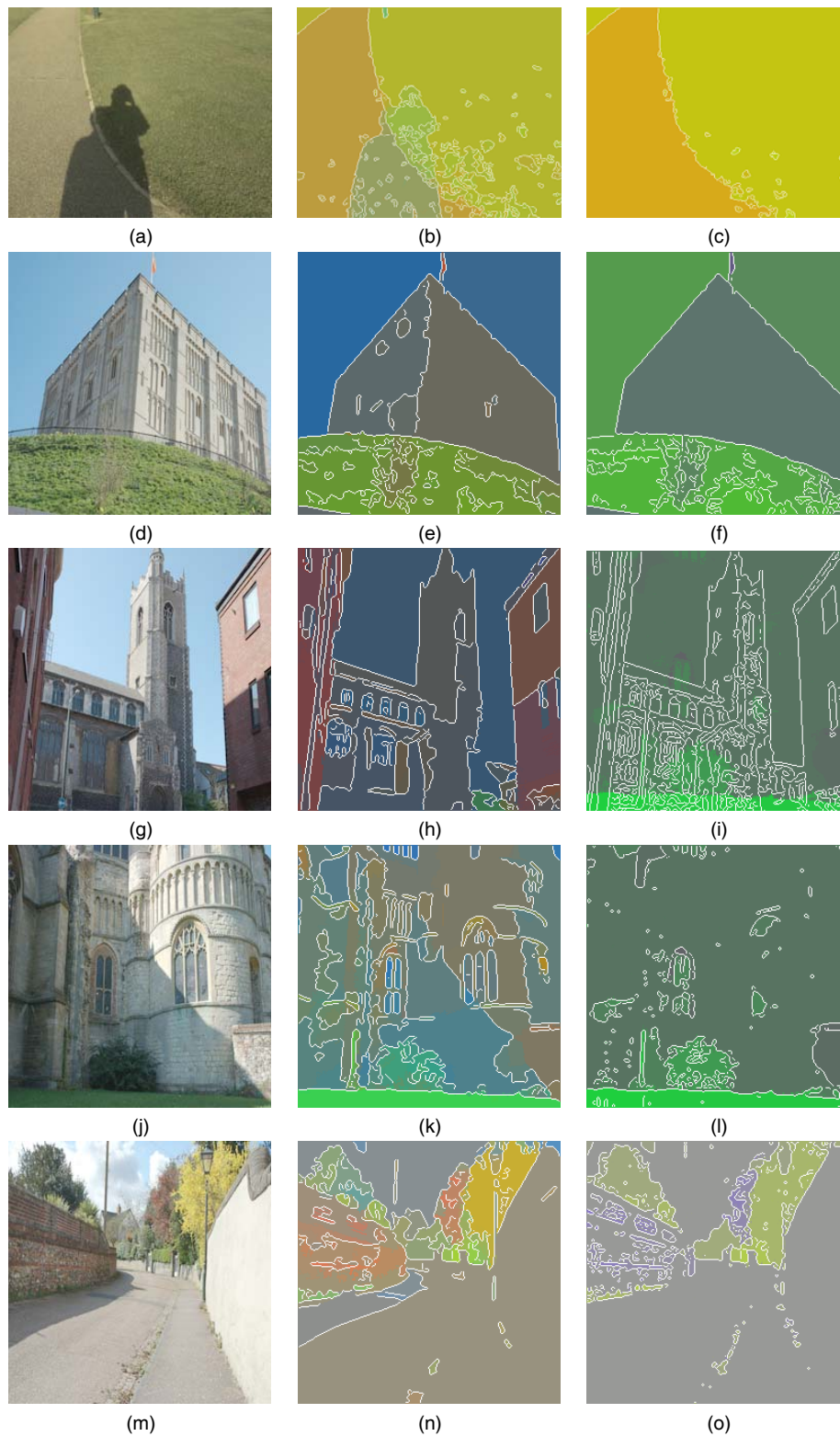
Figure 5. *Input colour images, their chromaticity, and the chromaticity images for an extracted illumination invariant. Here, the Mean-Shift algorithm has been applied to generate a cleaner image, and edge-detection overlaid — the illumination invariant has fewer edges on shadow boundaries. Cameras used were an HP 912 and a Nikon D70.*