Invariant Image Improvement by sRGB Colour Space Sharpening

[1]Graham D. Finlayson, [2]Mark S. Drew, and [2]Cheng Lu

[1]School of Information Systems, University of East Anglia
Norwich, (U.K.)
[2]School of Computing Science, Simon Fraser University
Vancouver (CANADA)
Corresponding author: M.S. Drew (mark@cs.sfu.ca)

ABSTRACT

Reasoning from image formation, we have shown that there exists a greyscale image – the invariant image – that depends only on the reflectances in the scene. Since illumination dependence is removed, one aspect of the invariant image is that shadows are effectively removed. Moreover, given either a calibration, or clean data with good noise statistics, this invariant is easily found. However, we found that the performance was much poorer on ordinary images that include the typical nonlinear processing in cameras. The contribution of this paper is that we can find a good invariant notwithstanding input image nonlinearities. Our strategy is to follow standard colorimetric procedure and convert image RGBs to the appropriate colour space for our method. We do this by converting first to the linear sRGB colour space and then concatenating conversion to XYZ tristimulus values by a spectral sharpening transform. We handle a suite of images which were intractable to the original method and are now able to find a shadow-free intrinsic reflectance image.

1. INTRODUCTION

The light entering a camera is a result of interaction between scene illumination and object surface. If the illuminating light changes in its intensity or colour, so will the RGB values recorded by the camera, and as a result unwanted artefacts, e.g. shadows, are introduced into the image. Separating such illumination effects from surface reflectance has become a fundamental task in computer vision. One approach to this problem is an illumination-invariant method.[1,2] In this body of work the goal is to effectively eliminate the effect of illumination changes from recorded images. This is achieved by deriving a 1-d invariant image, from a 3-band colour image, that is independent of both the colour and intensity of the scene illuminant. Since illumination dependence is removed, one aspect of the invariant image is that shadows are removed or greatly attenuated.

To form such an invariant image, current methods usually employ a multi-step procedure. For a particular colour camera, a target composed of colour surfaces (or, possibly, just a rather colourful scene) is imaged under different lights. The captured RGB values for each pixel are first transformed into a 2D band-ratio chromaticity colour space, e.g., {G/R, B/R}, and then logarithms are taken. Under the assumption of approximately Plankcian lighting, the 2-d log-ratio values for the same surface across different lighting tend to fall on a straight line. For one camera, all such lines are parallel, so that an invariant direction is orthogonal to these lines. The invariant image is thus the projection of the log-chromaticity values onto the invariant direction.

A crucial piece of information is the angle for the invariant direction. The usual strategy[1] requires a preliminary calibration routine, using the camera involved to capture images of a colour target under different lights. The calibration step is avoided in a new strategy[2] in which we show that the invariant direction is that which minimise entropy in the resulting 1-d quantity.

However, we found that performance was much poorer on ordinary images that include typical nonlinear processing in cameras;[3] real cameras also have sensors that are far from Dirac delta functions. As shown in §4, in that case the log-ratio chromaticity values do not form parallel lines and hence some shadows may not completely removed. This motivates this work. To reduce the impact of these limitations, we wish to employ a pre-processing procedure which consists of transforming the input image to a new one that can be considered as having been taken with a linearised and sharpened camera, with the invariant image finding routine carried out in the new colour space. To do so, we first transform the input image to standard linearised sRGB space. Then we can convert pixel values to XYZ using the standard sRGB-to-XYZ conversion. The conversion specifies a gamma of 2.2 and a white point of D65 for XYZ colour-matching functions. The resulting XYZ tristimulus values are then further sharpened by a sharpening transform. We find that this sharpened sRGB colour space does indeed improve the invariant.

2. PRINCIPLE OF INVARIANT IMAGES

Let us begin by summarising previous work[1] which showed that the colour constancy problem at a single pixel can be solved under the assumptions of Planckian scene illumination, Lambertian surfaces, and narrowband camera sensors. At a pixel, if an illuminant with spectral power distribution (SPD) $E(\lambda)$ is incident onto a surface $S(\lambda)$, then the camera sensor response can be expressed as

$$\rho_k = \sigma \int E(\lambda) S(\lambda) Q_k(\lambda) d\lambda \ , \quad k = R, G, B$$

where $Q_k(\lambda)$ denotes the spectral sensitivity of the kth camera sensor, k = 1,2,3, and σ is a constant Lambertian shading factor. To derive the 1-d invariant image, two assumptions must be made: first, the camera sensors are Dirac delta functions: $Q_k(\lambda) = q_k \, \delta(\lambda - \lambda_k)$; second, illumination is restricted to be Planckian. With Wien's approximation, an illuminant SPD can be parameterised by its colour temperature T:

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

where c_1, c_2 are constants and I controls magnitude. Sensor responses then take the form

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

We form band-ratio 2-vector chromaticities, e.g. $r_k = \rho_k / \rho_p$, where k = 1,3 and p = 2 (meaning band ratios R/G and B/G). This chromaticity operation effectively removes intensity and shading information. We then take the log of the chromaticities (denoted with a prime):

$$r_k' \equiv \log(r_k) = \log(s_k / s_p) + (e_k - e_p)/T$$

where we define $s_k = c_1 \lambda_k^{-5} S(\lambda_k) q_k$ and $e_k = -c_2 / \lambda_k$. As temperature changes, the 2-vectors will form a straight line in our 2-d log-ratio chromaticity space. The invariant image is formed by projecting 2-vectors onto the direction orthogonal to the line. The main issue in finding an invariant to illuminant thus becomes deciding on direction $e_k - e_p$. This is solved either by a calibration routine[1] or the minimum-entropy method.[2]

3. A SHARPENED SRGB COLOUR SPACE

The invariant image finding technique relies on the notion that camera sensors can be at least roughly approximated by narrowband sensors. Of course, this clearly does not hold in most cases. Therefore, a study[6] was carried out to determine an optimum matrixing scheme for linearly transforming camera RGB values to an intermediate, sharpened, colour space in which the assumptions of the invariant image algorithm held more accurately. It turned out in that study that a straightforward spectral sharpening transform[5] was not in fact the optimum linear transform.

Here, we explore how a standardised sharp colour space might be utilised to replace an optimization – i.e., we aim at a standard workflow to transform input RGB values into sharpened ones such that finding the invariant image proceeds directly, without any need for calculation of an unknown linear transform. To this end, we begin by making use of the standardised sRGB colour space.[4] The sRGB ("Standard") space, developed by Hewlett-Packard and Microsoft, is meant to fulfill the role of a default RGB colour space, forming a single recommendation for an intermediate linear colour space. The sRGB standard includes an sRGB-to-XYZ conversion matrix M. The first step transforms nonlinear sRGB to a linear sRGB colour space, essentially via gamma correction of input image pixels (but using a linear transform on the portion of values close to zero). The resulting linear sRGB is then converted to XYZ with a D65 white point: i.e., XYZ tristimulus values for image pixels, under standard illuminant D65. The XYZ D65 colour-matching functions derived from sRGB are in fact already quite sharp. However, we show here that making use of a further sharpening transform allows the invariant image finding routine to work correctly.

Denote the transformation from sRGB to linear sRGB space as a function $S(\rho)$ where ρ is an input RGB triple, and the transform matrix for taking linear sRGB values into XYZ-D65 values as *M*. Hence we capture the transform from the image to XYZ-D65 as:

$$\rho \rightarrow M\,S(\rho)$$

In order to sharpen the sRGB sensors, we wish to determine a further *sharpening transform matrix*[5] *T*. The so-called "database-sharpening" variant of spectral sharpening usually consists of producing matrix *M* directly from two sets of patch images formed with the camera under two different lights, such that the transformed images perform best in regard to diagonal colour constancy; i.e., they can be well converted to each other using a diagonal matrix transform. Here, we would compute eigenvectors of the least-squares transform from colour patches under one light to colours under another light. But in fact we know that what we would like to sharpen are the XYZ colour-matching functions themselves under D65. Therefore we first form the least-squares transform *P* from XYZ-D65 to XYZ-D50 (or to any other standard light — it does not greatly matter what light the XYZ-D65 is converted to), and then diagonalise this:

$$XYZ_{D65} \approx P\;XYZ_{D50}\,, \quad P = T\;diag\,(D)\,T^{-1}\,.$$

The resulting eigenvector matrix is our desired sharpening transform *T*. The XYZ colour-matching functions under D65 after the sharpening transform have narrower curves.

Hence we can form a matrix *Q* via a concatenation of these two matrices: the 3×3 matrix *M* taking us from linear sRGB colour space to XYZ, and the matrix *T* that sharpens the XYZ colour-matching curves: $Q = M\;T$. Now we can perform the invariant image finding routine on an input image by first matrixing the linearised sRGB values with *Q* and then operating within that new colour space, equivalent to a sharpened sRGB space.

The procedure for finding an invariant image involves using logarithms of colour ratios. Since the log involves a singularity, this may create mathematical problems at small values. To correct this, we as well change the previous procedure[1] by making use of a generalised logarithm function:

$$g(x) = \alpha(x^{1/\alpha} - 1)\,,$$

In the limit $\alpha \rightarrow \infty$ this approaches a log function. This step also improves the resulting, shadow-free images.

4. RESULTS

Figure 1 shows output of the invariant image finder, applied to real, nonlinear images. In all cases, the shadows are removed quite effectively. To compare, simply using gamma to linearise the input colour image and then finding the invariant does not do well for an input image that is indeed nonlinear RGB. But transforming to sRGB and then to sharpened sRGB, and then finding the invariant, does indeed do well.

5. CONCLUSION

We have presented a scheme for linearising and sharpening real image data, such that finding the invariant has better performance. The method utilises an sRGB colour space for linearising the input image points and then transforming to XYZ tristimulus values. To make use of sharper XYZ colour-matching curves, a database-sharpening method is used and the sharpening matrix obtained is then concatenated with the sRGB-XYZ transform to form a sharpened sRGB colour space. We have shown that the results of finding an invariant image benefit greatly from making use of this sharpened sRGB transform.

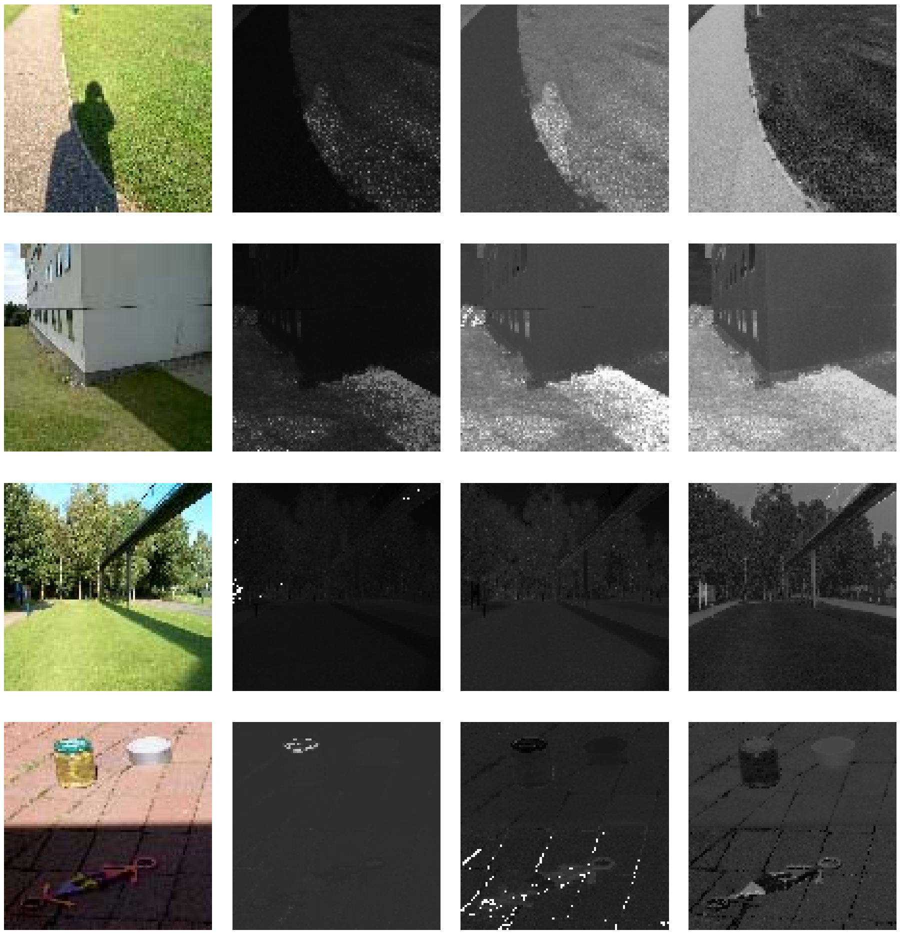

Figure 1 (a): Input; (b): invariant image using gamma linearisation; (c): invariant image using linear sRGB colour space; (d): invariant image using sharpened sRGB.

References

1. G.D. Finlayson and S.D. Hordley, “Colour constancy at a pixel,” J. Opt. Soc. Am. A, 18, 253-264 (2001).
2. G.D. Finlayson, M.S. Drew, and C. Lu, “Intrinsic images by entropy minimization,” in ECCV: European Conf. on Computer Vision (Springer-Verlag, Berlin 2004) pp. 582-595.
3. R. Ramanath, W.E. Snyder, Y.F. Yoo, and M.S. Drew, “Color Image Processing Pipeline in Digital Still Cameras,” IEEE Signal Processing, 22, 34-43 (2005).
4. http://www.srgb.com.
5. G.D. Finlayson, M.S. Drew, and B.V. Funt, “Spectral sharpening: Sensor transformations for improved color constancy,” J. Opt. Soc. Am. A, 11, 1553-1563 (1994).
6. M.S. Drew, C. Chen, S.D. Hordley, and G.D. Finlayson, “Sensor transforms for invariant image enhancement,” 10th Color Imaging Conference (IS&T, Springfield, VA, 2002), pp.325-329.
7. G.D. Finlayson and M.S. Drew, “Constrained least-squares regression in colour spaces,” J. of Electronic Imaging, 6, 484-493 (1997).