Illuminant invariance at a single pixel

Graham D. Finlayson[1], Steven D. Hordley[1]

Michael H. Brill[2]

[1] School of Information Systems
University of East Anglia
Norwich NR4 7TJ
United Kingdom

[2] Sarnoff Corporation,
CN5300
Princeton, NJ 08543-5300
USA

Email:{graham,steve}@sys.uea.ac.uk mbrill@sarnoff.com

Abstract

This paper addresses the question of what can be said about the colors in images that is independent of illumination. We make two main assumptions: Firstly, the illumination can be characterized as Planckian (a realistic assumption for most real scenes). Secondly, the camera behaves as if it were equipped with narrow band sensors (true for a large number of cameras). An alternative set of assumptions (with broad-band sensors) leading to the same invariant expression is also presented to show the robustness of the invariant. The resulting physics-based method results in a transformation of the original color image to a grey-scale one which does not vary with illumination. Experiments demonstrate that the distribution of grey-scale invariants in an image is a reliable cue for illumination independent object recognition.

1. Introduction

The light reaching our eye is a function of surface reflectance and illuminant color. Yet, the colors that we perceive depend almost exclusively on surface reflectance; the dependency due to illuminant color is removed through color constancy computation. As an example, the white page of a book looks white whether viewed under blue sky or under artificial light.

In digital photography the color constancy problem is usually posed as a two step process. First an estimate of the illuminant is derived through image analysis. Second this estimate is 'subtracted' from the image (the image is color corrected to remove any color cast due to illumination). The success or failure of the color constancy algorithm depends mostly on the validity of the illuminant estimate made. After many years of research, there are now good solutions to color constancy, solutions that deliver good estimates of illumination most of the time[1]. However, even the best algorithms can and do fail. The failures that occur are almost always due to images of scenes with color diversity is low (e.g. a scene of nature containing only shades of green). This is probably not surprising since in the pathological case of a scene containing a single surface it is not possible to separate light from reflectance. A pink surface viewed under white light is physically indistinguishable from a white surface viewed under pink light.

In this paper we look at how one might usefully analyse and work with color deficient scenes. What can we say about the colors in the image that is independent of illumination? That is, we do not aim to solve the classical color constancy problem but rather seek to identify that part of the problem that might easily be solved. As an example it may be possible that we can say that an image region corresponds to a pink or purple surface but that its not green. This weak conclusion is useful since it may ultimately help us to solve the color constancy problem: if a surface is pink then it might belong to a Caucasian face and face color can be used as a reference cue for illumination estimation.

Since we are no longer aiming to estimate the illuminant but rather are pulling out scene information independent of illumination we are really talking about color invariance (though we will make a strong link to classical color constancy later). Invariants are algebraic functions of a small number of proximal pixels which have the property that they are independent of (by construction they cancel out) dependency due to illumination. The key insight that is usually exploited is that, assuming linear models of illumination, and linear device response, that RGBs across an illumination are linearly or bilinearly related. Interestingly, many color invariants[2, 3, 4, 5, 6, 7] already exist and their value has been shown in many applications. Unfortunately, existing invariant approaches suffers from four intrinsic problems. First, the linear illumination assumption is over general. The non-linear Planckian locus (the region of color space where typical illuminants lie) is parameterizable by two numbers: intensity and temperature. Yet, to contain the Planckian locus using a linear model requires 3 degrees of freedom. This increase in dimensionality, necessary because linear models are being used, results in an illumination model that can describe many lights which cannot occur in nature (it is over general). This in turn reduces the amount of invariant information that can be extracted from an image. Second, invariant computation to date is only possible given spatial context (many pixels are required) and so is sensitive to occlusion. Third, invariants can be calculated only assuming there are two or more colors adjacent to one another (not true for objects such as bananas and oranges). Fourth, invariants can be calculated after color constancy computation but the converse is not true[8]: color constancy adds more information if it can be computed.

In this paper, we bridge the gap between the classical color constancy computation and the invariant approach. We begin with the premise that the chromaticities of typical illuminants fall along a [Planckian] non-linear locus parameterizable by two numbers. This is true for most color cameras and most illuminants. With this premise in hand, we ask: 'does there exist exist a 1-dimensional color coordinate, expressed as a function of the RGB or chromaticity, for which the color constancy problem can be solved?' The idea here is that we take our RGB image, convert it in some way to a grey-scale image, and then attempt to map the image grey-values to those observed under reference lighting conditions. We show that if a camera is equipped with narrow-band sensors (or behaves as if this were the case) then there exists a color coordinate where color constancy computation is trivial: there exists a coordinate where no computation actually needs to be done. By construction the grey-scale image factors out all dependencies due to light intensity and light color. We go on to generalize this result to the case of non-narrow-band sensors.

To validate our new invariant theory, images of different objects under different lights are taken with a SONY-DXC 930 camera (a camera that has an effectively narrow-band response). Each image is converted to a corresponding invariant grey-scale images. We find that the grey-scale invariant images of the same object viewed under different lights are almost the same. Indeed, the distribution of grey-scales in an invariant image is shown to be a useful cue for content based image indexing.

In section 2 of this paper we discuss color image formation and image variation due to Planckian illumination. The invariant coordinate transform, for cameras with narrow band sensitivities, is derived in section 3. The result is generalized to non-narrow-band sensors in section 4. Experimental results are presented in section 5.

2. Background

An image taken with a linear device such as a digital color camera is composed of sensor responses that can be described by

$$p_k = \int_\omega E(\lambda)S(\lambda)R_k(\lambda)d\lambda \;\; (k = R, G, B) \qquad (1)$$

where λ is wavelength, p_k is sensor response $k = R, G, B$ (red, green and blue sensitivity), E is the illumination and S is the surface reflectance and R_k is a camera sensitivity function. Integration is performed over the visible spectrum ω.

Let us assume the $R_k(\lambda) = \delta(\lambda - \lambda_k)$: it is a Dirac delta function with sensitivity only at some wavelength λ_k. Dirac delta functions have the well known sifting property:

$$p_k = \int_\omega E(\lambda)S(\lambda)\delta(\lambda - \lambda_k)d\lambda = E(\lambda_k)S(\lambda_k) \qquad (2)$$

Clearly, under $E_1(\lambda)$ and $E_2(\lambda)$ the RGB response for a particular surface can be related:

$$\begin{bmatrix} E_1(\lambda_R)S(\lambda_R) \\ E_1(\lambda_G)S(\lambda_G) \\ E_1(\lambda_B)S(\lambda_B) \end{bmatrix} = \begin{bmatrix} \frac{E_1(\lambda_R)}{E_2(\lambda_R)}E_2(\lambda_R)S(\lambda_R) \\ \frac{E_1(\lambda_G)}{E_2(\lambda_G)}E_2(\lambda_G)S(\lambda_G) \\ \frac{E_1(\lambda_B)}{E_2(\lambda_B)}E_2(\lambda_B)S(\lambda_B) \end{bmatrix} \qquad (3)$$

To ease the notation we rewrite (3) as:

$$\begin{bmatrix} R_1 \\ G_1 \\ B_1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} R_2 \\ G_2 \\ B_2 \end{bmatrix} \qquad (4)$$

Remarkably, even Equation (4), which is really quite simple, is an over general model of image formation. Illumination color is not arbitrary (we say nothing about the shape of $E(\lambda)$ in Equation (1)) and so the scalars α, β and γ in Equation (4) are not arbitrary either. To see that this is so, let us suppose that illumination might be modeled as a black-body radiator using Planck's famous equation[9]:

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

Equation (5) defines the spectral concentration of radiant exitance, in Watts per square metre per wavelength interval as a function of wavelength λ (in meters) and temperature T in Kelvins. The constants c_1 and c_2 are equal to 3.74183×10^{-16} Wm^2 and 1.4388×10^{-2} mK respectively. Equation (5) does not account for varying illuminant power. To model varying power we add an intensity scalar I to Planck's formula:

$$E(\lambda, T) \;=\; I * c_1 \lambda^{-5} \left(e^{\frac{c_2}{T\lambda}} - 1 \right)^{-1} \tag{6}$$

While the shape of daylights and Planckian radiators is similar, this is not true for fluorescents (which tend to have highly localized emission spikes). But, even here Equation (6) can be used. This can be done because we are not really interested in spectra per se but rather in how they combine with sensor and surface in forming RGBs. For almost all daylights and typical man-made lights, including fluorescents, there exists a black-body radiator, defined in (6), which, when substituted in (1), will induce very similar RGBs[9]. Interestingly, if such a substitution cannot be made, the color rendering index (broadly, how good surface colors look under a particular light) is poor[9]. Indeed, the lighting industry strives to manufacture lights with chromaticities close to the Planckian locus.

3. Invariance at a pixel: narrow band sensors

In this section we take our model of image formation together with Planck's equation and show that there exists one coordinate of color, a function of RGB, that is independent of light intensity and light color (where color is defined by temperature). However, to make the derivation cleaner we first make a small (and often made[9, 10]) simplifying alteration to (6). In Planck's equation λ is measured in metres; thus, we can write wavelength $\lambda = y * 10^{-7}$ where $y \in [1, 10]$ (the visible spectrum is between 400 and 700 nanometers 10^{-9}). Temperature is measured in thousands of Kelvin or equivalently $t * 10^3$ (where $t \in [1, 10]$). Substituting into the exponent of Equation (6) we see that:

$$\frac{c_2}{T\lambda} = \frac{1.4388 \times 10^{-2}}{y \times 10^{-7} * v * 10^3} = \frac{1.4388 \times 10^2}{y * t} \tag{7}$$

Because t is no larger than 10 (10000K) and there is no significant visual sensitivity (for humans or most cameras) after 700nm, $y \leq 7$, it follows that $e^{\frac{1.4388 \times 10^2}{yt}} >> 1$ and so:

$$E(\lambda, T) \;\approx\; I c_1 \lambda^{-5} e^{-\frac{c_2}{T\lambda}} \tag{8}$$

Substituting (8) in (2) we see that:

$$p_k = \int_\omega E(\lambda) S(\lambda) \delta(\lambda - \lambda_k) d\lambda = I c_1 \lambda_k^{-5} e^{-\frac{c_2}{T\lambda_k}} S(\lambda_k) \tag{9}$$

Taking natural logarithms of both sides of (9),

$$\ln p_k = \ln I + \ln(S(\lambda_k)\lambda_k^{-5} c_1) - \frac{c_2}{T\lambda_k} \tag{10}$$

That is, log-sensor response is an additive sum of three parts: $\ln I$ (depend on the power of the illuminant but is independent of surface and light color), $\ln(S(\lambda_k)\lambda^{-5}c_1)$ (depends on surface reflectance but not illumination) and $-\frac{c_2}{T\lambda}$ (which depends on illumination color but not reflectance).

Remembering that in (10), $k = R, G, B$; we have 3 relations which exhibit the same structure: each of the ln R, ln G and ln B sensor responses are an additive sum of intensity, surface and illumination components. By canceling common terms, we show below that we can derive two new relations which are intensity independent (but depends on illumination color) and from these a final relation which depends only on reflectance.

We begin by introducing the following simplifying notation: let $S_k = \ln(S(\lambda_k)\lambda_k^{-5}c_1)$ and $E_k = -\frac{c_2}{\lambda_k}$ ($k = R, G, B$ sensor). The following two relations, red and green, and blue and green log-chromaticity differences (or **LCD**s), are independent of light intensity:

$$\begin{aligned} p'_R &= \ln p_R - \ln p_G = S_R - S_G + \tfrac{1}{T}(E_R - E_G) \\ p'_B &= \ln p_B - \ln p_G = S_B - S_G + \tfrac{1}{T}(E_B - E_G) \end{aligned} \tag{11}$$

Now using the usual rules of substitution it is also a simple matter to derive a relation that is independent of temperature:

$$p'_R - \frac{(E_R - E_G)}{(E_B - E_G)} p'_B = S_R - S_G - \frac{(E_R - E_G)}{(E_B - E_G)}(S_B - S_G) \tag{12}$$

where all S_k and E_k are independent of illuminant color and intensity. Equation (12) informs us that there exists a weighted combination of LCDs that is independent of light intensity and light color.

4. Invariance at a pixel: non-narrow-band sensors

4.1. A Physical Argument

Here we introduce an analytic framework in which the illuminant-invariance at a pixel is exact for non-narrowband sensors. Let the reflectance spectrum be a Gaussian in some monotonic function x of wavelength:

$$S(x) = a e^{-\frac{(x-v)^2}{2s^2}} \tag{13}$$

Also, let the illuminant SPD be approximated by an exponential in x (a tolerable approximation derived above from the black-body spectrum, and also derived earlier from a similar argument and used to create color-correction filters[10]).

$$E(x) = ce^{fx} \quad (14)$$

In (13) and (14), a, v, s, c, and f are constants.

Then given a measured response vector $\underline{p} = [R \;\; G \;\; B]^t$, it is plausible that one can use the three response equations to eliminate the two illuminant parameters (c and f) and be left with a function of the parameters of the reflectance Gaussian. Indeed, we show there is a closed-form solution if the visual spectral sensitivities $R_k(x)$ are equal-spread Gaussians in x:

$$R_k(x) = e^{-\frac{(x-w_k)^2}{2}} \quad (15)$$

where w_k and the variable x are expressed in units of the width (standard deviation) of this Gaussian.

The response values p_k can be expressed as

$$p_k = \int S(x)R_k(x)E(x)dx$$

$$= ac[2\pi/V]^{0.5} e^{[-(\frac{\frac{v^2}{s^2}+w_k^2}{2})+\frac{(f+\frac{v}{s^2}+w_k)^2}{2V}]} \quad (16)$$

where $V = (1 + \frac{1}{s^2})$.

[**NOTE:** The steps in performing the integral between infinite limits of $e^{(Ax^2+Bx+C)}$ are as follows:

1. Start with the well known result that the standard normal probability density function in x is $(2\pi)^{-0.5}e^{-x^2/2}$, whose integral is 1. Then it follows that the integral of e^{-x^2}, over infinite limits is $\pi^{0.5}$.

2. Complete the square of the quadratic in e^{Ax^2+Bx+C} so it becomes $e^{-[(-A)^{0.5}x-0.5\frac{B}{(-A)^{0.5}}]^2} e^{C+\frac{B^2}{4A}}$.

3. Change the variable in the integral to $[x((-A)^{0.5}) - 0.5\frac{B}{(-A)^{0.5}}]$, so that performing the integration will bring out the factor $[\frac{\pi}{-A}]^{0.5}$. Of course, the constant factor $e^{C+\frac{B^2}{4A}}$ propagates into the final evaluation.]

Starting with Eq.16, there is much cancellation (and removal of dependence on the illuminant parameter c) when one evaluates a ratio of two components of $\underline{p}$:

$$log(\frac{p_k}{p_j}) = (1/2)[-\frac{w_k^2 - w_j^2}{1+s^2} + (f+\frac{v}{s^2})(w_k - w_j)] \quad (17)$$

We now reveal a cyclic combination of quantities such as in the above equation that eliminates any dependence on the remaining illuminant parameter f. Given measured sensor values R, G, B, and sensor peak-wavelengths $w_1 = r$, $w_2 = g$, $w_3 = b$, one way of expressing the invariant is:

$$P = (g-b)log(R) + (b-r)log(G) + (r-g)log(B) \quad (18)$$

The quantity P is the following function of the reflectance standard deviation s:

$$P = -0.5[(g-b)r^2 + (b-r)g^2 + (r-g)b^2]\frac{1}{s^2+1} \quad (19)$$

which is independent of the illuminant and also (incidentally) of the reflectance peak wavelength v. Evaluated pixel-by-pixel, P comprises the sought-after illuminant-invariant map. The first expression for P is the same as derived earlier[11] using von-Kries-adapted tristimulus values in place of R, G, B. However, the earlier invariant depends on a reference-white von-Kries denominator, and this one does not. Therefore, the current description represents an advance over the older theory.

4.2. A statistical argument

Though the conditions set forth in Section 4.1 seem rather severe, they conjoin with the assumptions set forth in Section 3 to make a larger domain of invariance for Eq. (18). This simpler picture has the advantage that no constraint is placed on the shape of reflectance spectra but the disadvantage, by definition, of requiring narrow-bandedness. Narrow-band sensors are blind to large areas of the visible spectrum and so are not useful for general imaging applications. However, in practice narrow-bandness need not be implemented physically. Rather, it suffices that a camera which has non-narrow band sensors behaves as if it had narrow-band sensitivities. This is in fact true for many cameras. Even when cameras do not behave as if they were equipped with narrow-band sensitivities this condition can often be enforced. All that is required is that the initial sensitivities are transformed to a special 'sharp' basis which behave like narrow-band sensors. Such a transformation exists for the broad band sensitivities of the human cones[12] and the spectrally broad band Kodak DCS 460 camera[13]. A grey scale invariant can be calculated for almost all imaging devices[14].

5. Experiments

Empirically we found that a SONY DXC-930 camera (sensitivities shown in Figure 1) behaved as if it had narrow-band sensitivities. This enabled us to calculate all the terms in (12) and so to calculate invariant images. We took 10 SONY DXC-930, RGB images (from the Simon Fraser dataset[15]) of two colorful objects (a beach ball and a detergent package) under the 5 illuminants: Macbeth fluorescent color temperature 5000K (with and without blue filter), Sylvania Cool white fluorescent, Philips Ultralume, Sylvania Halogen. These illuminations constitute typical everyday lighting conditions: yellowish to whitish to bluish lights. The luminance grey scale images, calculated by summing $R + G + B$, are shown in the first and third columns of Figure 2. It is clear that the simple luminance grey scale is not stable across illumination. In columns 2 and 4 the corresponding invariant grey-scale images are calculated (we map RGBs to scalars using (12) and code these as a grey values). It is equally clear that the grey-scale pixels in these images do not change significantly as the illumination changes. Also notice that qualitatively the invariant images maintain good contrast: not only have we obtained illumination invariance but the images that result are visually salient.

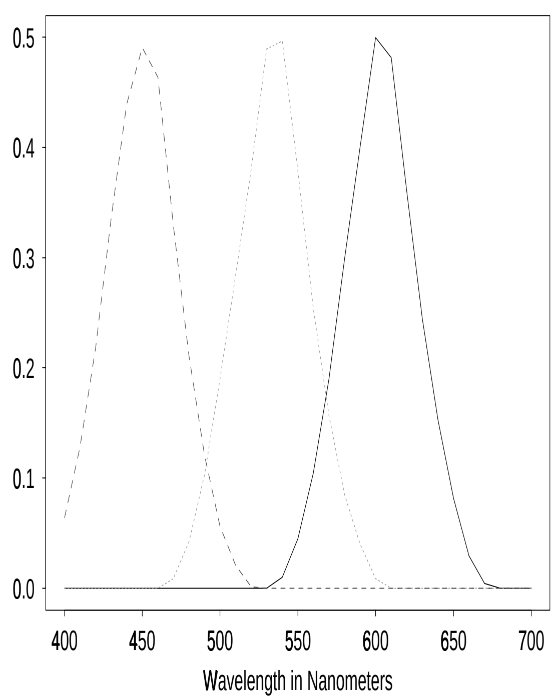

Figure 1: *The spectral sensitivities for a SONY DXC-930 camera*

As a more concrete test of the utility of our calculated illuminant invariant we carried out a set of object recognition experiments. To the beach ball and detergent packages we added 9 other colorful objects. These too were imaged under all 5 lights[15]. For all 55 images we calculated their respective grey-scale invariant histograms and then used these as an index for object recognition. Specifically, we took each light in turn and used the corresponding 11 object histograms as feature vectors for the object database. The remaining 44 object histograms were matched against the database; the closest database histogram being used to identify the object.

We found that a 16 bin invariant grey-scale histogram, matched using the Euclidean distance metric[16], delivers near perfect recognition. Almost 96% of all objects were correctly identified. Moreover, those incorrectly matched, were all found to be the second best matching image. This performance is really quite remarkable. Funt et al[15] measured the illuminant using a spectra-radiometer and then corrected the image colors based on this measurement (so called perfect color constancy). They then indexed objects by matching corrected chromaticity histograms. Surprisingly they found that they could achieve only 92.3% recognition. Moreover, at least one object was matched in 4th place (the correct matching histogram was the fourth best answer). These results indicates how difficult it is to correct image colors across illumination even when the light is known. In contrast, the invariant calculated here appears to be more stable (or at least more salient) and this is reflected in better recognition results.

Funt et al also used a variety of color constancy algorithms, including max RGB, grey-world and a neural net method[17, 18, 19], as a preprocessing step in color distribution based recognition. All methods tested performed significantly worse than the perfect color constancy case. No algorithm delivered supported more than a 70% recognition rate.

Other color invariant based methods, predicated on functions of many image pixels, have also been tried on the same data set[8]. None delivered results better than the 96% recognition rate reported here.

6. Conclusions

In this paper we looked at image formation under Planckian illumination. For the special case of cameras equipped with narrow-band sensors or cameras that behave as if they had narrow band sensitivities, we showed that it is possible to synthesise a grey-scale image which does not vary with illumination. This result is verified by experiment: grey-scale invariant histograms are used as a cue for recognizing objects viewed under different illuminants. The result is generalized to the case of non-narrow-band sensors.

Acknowledgements

Research at the University of East Anglia was supported by EPSRC grant GR/L60852 and Lightseer Ltd.

References

[1] G.D. Finlayson, S.D. Hordley, and P.M. Hubel. Colour by correlation: a simple unifying theory of colour constancy. In *IEEE international conference on computer vision*, pages 835–842, 1999.

[2] B.V. Funt and G.D. Finlayson. Color constant color indexing. *PAMI*, 17(5):522–529, May 1995.

[3] S.K. Nayar and R. Bolle. Computing reflectance ratios from an image. *Pattern Recognition*, 26:1529–1542, 1993.

[4] G. Healey and L. Wang. The illumination-invariant recognition of texture in color images. *Journal of the optical society of America A*, 12(9):1877–1883, 1995.

[5] G. Healey and D. Slater. "Global color constancy: recognition of objects by use of illumination invariant properties of color distributions". *Journal of the Optical Society of America, A*, 11(11):3003–3010, November 1994.

[6] G.D. Finlayson, S.S. Chatterjee, and B.V. Funt. Color angular indexing. In *The Fourth European Conference on Computer Vision (Vol II)*, pages 16–27. European Vision Society, Springer Verlag, 1996.

[7] G.D. Finlayson, B. Schiele, and J.L. Crowley. Comprehensive colour image normalization. In *The Fifth European Conference on Computer Vision*. European Vision Society, Springer Verlag, 1998.

[8] G.D. Finlayson and G.Y. Tian. Color normalization for color object recognition. *International Journal on Pattern Recognition and Artificial Intelligence*, pages 1271–1285, 1999.

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

[10] C. McCamy. A nomograph for selecting light-balancing filters for camera exposure of color films. *Photog. Sc. Eng.*, 3:302–304, 1959.

[11] M.H. Brill and G. West. Spectral conditions for color constancy via von kries adaptation. In *AIC COLOR 81*, 1981. paper J10.

[12] G.D. Finlayson and B.V. Funt. Coefficient channels: Derivation and relationship to other theoretical studies. *COLOR research and application*, 21(2):87–96, 1996.

[13] G.D. Finlayson and M.S. Drew. Positive bradford curves through sharpening. In *IS&T and SID's 7th Color Imaging Conference*. 1999.

[14] G.D. Finlayson and S.D. Hordley. Uk patent application no. 0000682.5. Under review, British Patent Office.

[15] B.V. Funt, K. Barnard, and L. Martin. Is machine colour constancy good enough. In *The Fifth European Conference on Computer Vision (Vol II)*, pages 445–459. European Vision Society, 1998.

[16] J. Hafner, H.S. Sawhney, W. Equitz, M. Flickner, and W. Niblack. Efficient color histogram indexing for quadratic form distance functions. *IEEE transactions on Pattern analysis and Machine Intelligence*, 17(7):729–735, 1995.

[17] E.H. Land. The retinex theory of color vision. *Scientific American*, pages 108–129, 1977.

[18] R.W.G. Hunt. *The Reproduction of Color*. Fountain Press, 5th edition, 1995.

[19] V. Cardei B.V. Funt and K. Barnard. Learning color constancy. In *4th IS&T and SID Color Imaging Conference*. 1996.

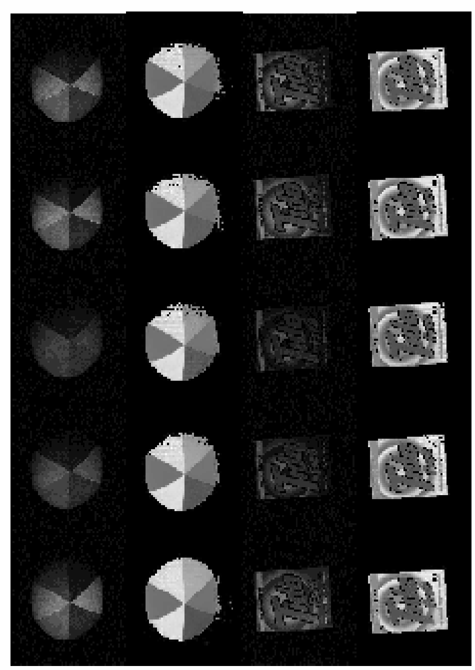

Figure 2: *Raw luminance images of a beach ball box and Tide detergent box calculated across 5 coloured lights (cols 1 and 3) are compared with the corresponding invariant grey-scale images (cols 2 and 4) .*