

# Colour Indexing Across Illumination

Graham D. Finlayson and Gui Yun Tian

Colour & Imaging institute, University of Derby, Kingsway  
Derby DE22 3HL, UK

## Abstract

Because the colours in an image convey a lot of information, almost all image database systems support colour content queries. Unfortunately colour based queries do not always return the images that were sought even although there is a good colour match. Such failures are easily explained. We as human observers do not see raw image colours but rather make an interpretation of the colours in an image. Our interpretation allows us to decouple the intrinsic colour of the objects and surfaces, captured in an image, from the colour of the illumination. An indoor picture with a yellowish colour cast is interpreted as just that, we do not think that all the objects in the scene are more yellow than they usually are. In contrast, image database systems generally make no such comparable interpretation.

In this paper we set forth an experimental study that attempts to quantify the nature and magnitude of the illumination colour problem. We are interested in measuring how image colours depend on illumination and how this dependency might be removed. Our work based on a small, but accurately calibrated, image database comprising 11 colourful objects imaged under 4 typical household lights. Because illumination colour impacts so dramatically on image colours, querying this dataset by colour-content delivers very poor indexing. To improve indexing performance, the illumination bias in images needs to be removed. This is done by applying an appropriate mapping to the image colours (e.g. a reddish cast can be removed by reducing the redness at each pixel). We found that mapping image colours based on a measure of the illuminant results in good indexing. However, when the mapping depends on both scene content and illumination together, the indexing performance is even better still. This is a surprising result since it is accepted doctrine that a change in illuminants should result in a systematic change in image colours and this change should effect all images equally (scene content should not add any useful information). Of course if illumination colour depends on scene content then it will be difficult to measure since the spectral statistics of the scene are also unknown. If measurement is difficult, estimation (using a colour constancy algorithm) must be more difficult still.

## 1 Introduction

Because the colours in an image convey a lot of information, almost all image database systems support colour content queries: colour as an important cue for image matching and retrieval. Colour queries often take the form of the distribution of colours based in an image. For example, Swain and Ballard[1] developed a system called colour indexing for matching images based on the similarity of colour histograms. Their technique has been incorporated in many image database systems including IBM's QBIC (Query by image content[2]). Part of the appeal of Swain and Ballard's method (and other derivative techniques) is that a colour histogram is independent of many imaging conditions: e.g. the orientation of a scene, the relative position of particular scene elements and the absence (or occlusion) of some of the colours.

However, early on it was argued that colour indexing would have strictly limited application because image colours depend on the lighting condition [1]. Indeed small variations in lighting can lead to complete indexing failure. To overcome the lighting dependency problem, three solutions have been proposed. First, in the computer vision field, it is common to control the lighting conditions to remove the dependency [3]. Clearly, this solution cannot be applied in general imaging environment where capture conditions cannot be controlled. The second solution is to use colour-constancy algorithms as a preprocessing step before colour indexing [1]. Colour constancy algorithms attempt to estimate the colour of the prevailing illuminant through image analysis with the colour cast due to the estimated illuminant being removed at a second stage. Unfortunately, existing colour constancy algorithms cannot deliver sufficiently accurate estimates to deliver good indexing[4]. In the third approach, colour invariant features are extracted from images and these invariants are used for indexing. Funt and Finlayson [5] showed that the ratio of adjacent colours is independent of illumination. Histograms of colour ratios sufficed to support good indexing across illumination. In another study Healey and Slater [6] derived functions of colour distribution moments that were invariant to illumination. These too can support image indexing. Many other studies have shown similar results: information rich colour invariants exist and support illuminant independent indexing[7].

Despite the strength of the invariant approach (quantitative data demonstrating better indexing compared with the direct use of raw colours) colour indexing is commonly used but invariant indexing is not. Part of the explanation for this circumstance is that colour is a feature that we find intuitive but invariants are

counterintuitive. Moreover, lighting conditions are sometimes key to our search: in searching for images taken at sunset one is looking for the lighting conditions!

In this paper we seek to re-examine the colour vs. invariant question. For example, it could be that there exists a colour constancy algorithm sufficient to support cross illumination indexing, but that we have not found it yet. Thus, we set forth an experimental study that attempts to quantify the nature and magnitude of the illumination colour problem. We are interested in measuring how image colours depend on illumination and how this dependency might be removed. Our work is based on a very small, but accurately calibrated, image database comprising 11 colourful objects imaged under 4 typical household lights. For each image the illumination is *measured* and this measurement is used to colour correct the image. By measuring the illuminant we are basically assuming a prescient colour constancy algorithm: one that, regardless of scene content, can always estimate the light colour correctly. To correct the image colours, we calculate three scalars, one for each of the R, G, and B colour channels, which corresponds to the redness, greenness and blueness of the illuminant. The reciprocals of these scalars are then applied to every pixel and in so doing the colour bias is removed.

This paper reports a very interesting and quite unexpected discovery. Even when we apply this optimal colour constancy pre-processing; we still find problems indexing the colour-corrected images; indexing, though much improved, is still not good enough. This surprising circumstance might be explained in two ways. First, the scaling model of illumination change might be too simplistic to model illumination change and second, our measurement of the illuminant could be wrong. The former was found not to be the case. Modelling illumination change by a 3x3 linear transform (theoretically and experimentally quite adequate [6, 17]) did not improve indexing. However, we did find evidence that our illumination measurements were inaccurate. Instead of calculating scaling factors (or linear correction transform) based on the illuminant measurement we calculated the best mapping taking image colours recorded under one light to those recorded under known reference conditions. Of course, this approach is only useful if the image content recorded across images is in registration. This was the case for our data set: the same object was always imaged in exactly the same position. We found that for different objects imaged across the same pair of lights that different correction transforms (3 scalars or 3x3 matrices) deliver the best colour correction; that is, colour correction depends both on the light colour and scene content. The implication of this result is that the ‘effective’ illumination depends both on the light source **and** the scene content. But, why should the effective illuminant depend on the scene content?

We believe that part of the answer lies in the very nature of illumination. Contrary, to what we may think the illumination in a scene does not just depend on the light source but rather depends on the scene content itself. Light strikes surfaces and is reflected and this reflected light then strikes other surfaces and so forth. So, the scene content itself has an effect on the illumination. However, measuring this effect is difficult. Interreflection is a function of light and surface, and the colour of the surfaces is not generally known. Of course, if illumination is hard to measure it follows it must be even harder to estimate using a colour constancy algorithm.

So if illumination by itself is insufficient for colour correction then should we use scene statistics as well? Pragmatically, if we wish to get the best indexing then the answer is yes. However, when we base colour correction on scene statistics then we are entering the realm of colour image normalization [9]. Colour normalization removes bias due to illumination but does not recover the true colours of surfaces. Rather, normalized colours are a function of the true colour and the context in which they are scenes. Unfortunately, this means that the same object viewed in different scenes must have different normalized colours and so colour normalization itself may not deliver good indexing. Applying colour normalization in local image regions can solve the context problem. Local colour normalizations are called **colour invariants**. Thus, the fact that illuminant colour is so difficult to measure, endorses, in a round about sort of way, the colour invariant approach to image indexing. More detailed discussion of colour invariant indexing is presented in [18].

Before embarking further on the paper we wish to comment a little further on our experiments. In particular, we are the first to admit that our image database is very small. However, here are two very good reasons why such a small database was used. First, to build a calibrated data set where lighting conditions are rigorously controlled is a very time consuming process. The calibrated set we used (44 images in total) involved scores of man-hours to compile (good calibration is hard). Yet, if we are to understand the nature of illumination change an accurately calibrated dataset is a prerequisite. Second, the problems of illumination colour and estimating the illumination (the technical focus of this paper) manifest themselves very clear in this small set. Because such problems exist in such a small database implies they must exist in the large scale as well.

In section 2 of this paper we review colour image formation and how colour depends on illumination. The experiments and test images and indexing results are reported in section 3. The paper finishes with a short conclusion.

## 2 Colour Image Formation and Modeling Colour Correction

The light reflected from a surface depends on the spectral properties of the surface reflectance and of the illumination incident on the surfaces:

$$\underline{\rho}_k = \int_{\omega} E(\lambda) S(\lambda) F_k(\lambda) d\lambda \quad (1)$$

Where  $\lambda$  is wavelength,  $\underline{p}$  is the k-vector of sensor response (RGB pixel values),  $F(\lambda)$  is the k-vector of sensor spectral sensitivities (the red, the green and the blue sensing channels),  $E(\lambda)$  is the spectral power distribution of the illumination (assumed constant across the scene),  $S(\lambda)$  is the spectral reflectance function for a surface. The integral range  $\omega$  is over the visible spectrum  $\omega \in [380, 780nm]$ .

Let us approximate the spectral reflectance:

$$S(\lambda) \cong \sum_{j=1}^3 \sigma_j s_j(\lambda) \quad (2)$$

Where  $s_j(\lambda)$  are a set of  $n$  fixed basis functions. Let  $\underline{p}$  denote the column vector of sensor measurement  $\underline{p}=(R,G,B)^T$ , and let  $\underline{\sigma}=(\sigma_1, \sigma_2, \sigma_3)^T$  of spectral reflectance function coefficients. We can write

$$\underline{p}^E = \Lambda^E \underline{\sigma} \quad (3)$$

Where  $\Lambda^E$  is the 3x3 matrix with entries:

$$\Lambda^E = \int_{\omega} E(\lambda) s_j(\lambda) F_k(\lambda) d\lambda \quad (4)$$

The colours of a single surface viewed under different illuminants  $E(\lambda)$  and  $E'(\lambda)$  is described by:

$$\underline{p}^E \cong \Lambda^E \underline{\sigma} \quad \underline{p}^{E'} \cong \Lambda^{E'} \underline{\sigma}$$

The vector  $\underline{\sigma}$  describes the distribution of spectral reflectance of the object, and is independent of illumination. Because camera response is a linear transform of surface weight vector  $\underline{\sigma}$ , it follows that camera responses are related across illumination by a linear transform:

$$\underline{p}^E \cong M \underline{p}^{E'} \quad (5)$$

$$M = \Lambda^E [\Lambda^{E'}]^{-1}$$

Equation (5) is a very important result. It tells us that colours matched across illumination are a linear transform apart. Of course the validity of (5) rests on our assumption that we could describe surface reflectance using a 3 parameter linear model. In fact this is only approximately true. Three parameter models do model reflectances quite accurately, but it has been argued that 6 or 7 parameter models are needed for exact reconstruction[10]. However, in reality we are not too bothered about reconstruction of reflectances but rather, are interested in how surfaces interact with light and sensor in forming an RGB response. By modeling reflectances by a basis that captures the important variance of surfaces in terms of how they interact with lights, Marimont and Wandell[11] have shown that a 3 dimensional model of reflectance is very accurate indeed.

Equation (5) plays a central role in our study of illumination change since it effectively places a lower bound on the complexity of the illuminant colour problem. If we can recover the 9 parameters of the lighting matrix  $\Lambda$ , then we can discount colour bias due to illumination. However, nine dimensional problems are hard to visualize and solve. Indeed, to our knowledge there exists no implemented colour constancy algorithm that tackles this problem (one has been suggested[12] but never implemented). To simplify matters  $M$  in Equation (5) is often taken to be a 3-parameter diagonal matrix; thus reducing colour constancy to a 3-parameter problem. But, can such a simplifying step really be made?

In fact it has been shown[13,14] that if the sensor spectral sensitivities respond only to a single wavelength of light (they are delta functions:  $F_k = \delta(\lambda - \lambda_k)$  ( $k=1,2, 3$  or  $r, g, b$ ), then a diagonal matrix is a perfect model of illumination change. Of course camera sensitivities are not delta functions so we must consider how this rather nice theoretical observation manifests itself in practice. First, Worthey and Brill[14] have shown that because light and surface tend to change slowly over narrow bands of wavelengths, that a diagonal transform also describes illumination change for sensors that are only somewhat narrow-band. Almost all 3-chip colour cameras have sufficiently narrow support to render the diagonal model of light change very accurate indeed. Even for cameras that have broadband sensitivities, they can generally be made to behave like narrow-band sensors[15] and so a diagonal model is valid again.

### 3 Experiments of Colour Based Object Recognition

In our experiments we used the Simon Fraser calibrated image set comprising 11 different colour objects viewed under 4 typical colours of light. (Figure 1 shows the objects)<sup>1</sup>. The pictures were taken with a Sony SXC-930 3-CCD colour video camera balanced for 3200K lighting with the gamma correction turned off so that its response is essentially a linear function of luminance. The images were captured under 4 different illuminants: Macbeth Judge II illuminant A, a Sylvania Cool White Fluorescent, Philips Ultralume Fluorescent and the Macbeth Judge II 5000 Fluorescent. The illuminant spectra are plotted in Figure 2.

<sup>1</sup> The authors are grateful to Prof. Brian Funt for making these images available.

Importantly, the position of objects was held fixed when the illumination colour was changed. Thus, across illumination (for a single object) we have registered images and so we can compare image colours recorded under one light with corresponding colours under another light. The images recorded under one of the illuminants forms the model image database. The other 3 groups of images are used as ‘test’ datasets.



Figure 1: The 11 colourful objects for colour indexing (Cool White Fluorescent)

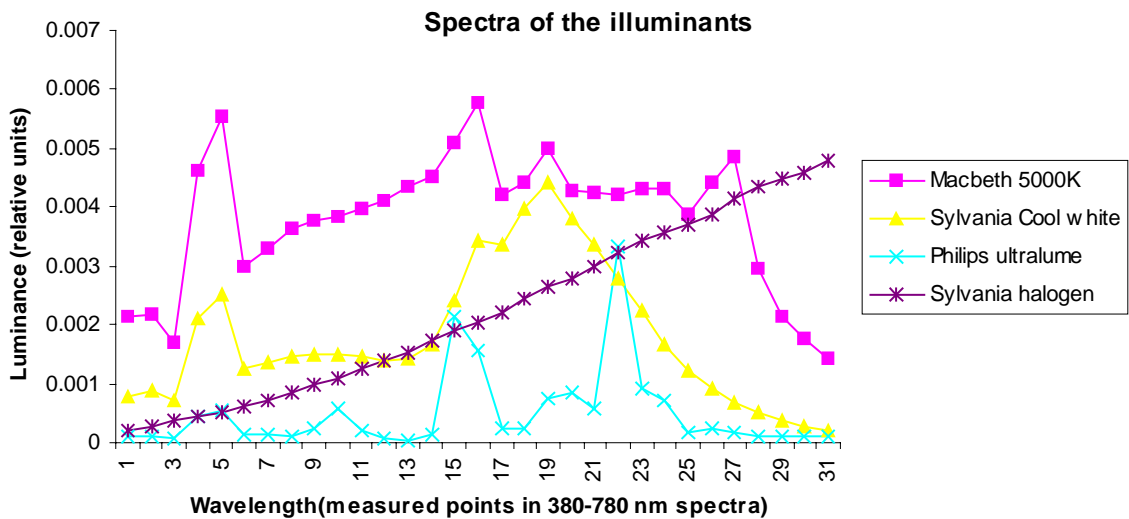


Figure 2: Spectra of the five test illuminants as measured by a photospectrometer

Like other authors (e.g. [16]) we found that the brightness of individual pixels caused a problem. In particular the relative position of object surface and light source changed from image to image and so the shading field (brightness of individual pixels) also changed. To factor out brightness we mapped RGB colours to chromaticities prior to indexing. The  $(r, g)$  chromaticity is defined to be  $(R/(R+G+B), G/(R+G+B))$ . To index our database we use the colour indexing method of Swain and Ballard. If  $I$  represents the chromaticity histogram of a test image with  $n^2$  bins (chromaticity space is divided into  $n$  bins along each of the  $r$  and  $g$  dimensions) and  $M$  a model histogram then the closeness of the pair of histograms (and hence images) is defined to be:

$$d_1 = \sum_{i=1}^n \sum_{j=1}^n \left| \frac{I(i, j)}{N_I} - \frac{M(i, j)}{N_M} \right| \quad (6)$$

Where  $I(i, j)$  and  $M(i, j)$  are the number of pixels in each bin of test image and model image.  $N_I$  and  $N_M$  are respectively the total number of pixels in the test and model histograms. Clearly, for any images,

$\sum_{i=1}^n \sum_{j=1}^n \frac{I(i, j)}{N_I} = 1$  or  $\sum_{i=1}^n \sum_{j=1}^n \frac{M(i, j)}{N_M} = 1$  and so (6) is not affected by image size. The measure calculated in

(6) is directly related to the histogram intersection measure of Swain[1] and is also a distance metric (it is the L1 or city-block distance). In our experiments, we found that setting  $n=16$  compromised the need to capture the shape of the distribution of chromaticities with the requirement that the histogram should be insensitive to quantization artifacts (e.g. the same chromaticity being mapped to adjacent bins in the presence of image noise).

Taking each illuminant in turn we built our model histogram database. We then took the remaining 33 images (for the 3 other lights) as our test set and tried to identify these by matching the test image histograms to the database set. Without correcting for the illumination, the match success was very poor. Table 1 records the indexing performance. For all choices of model illumination, there is very poor indexing of test images. Over all choices of model illumination indexing by chromaticity histogram comparison delivered a recognition rate of only 36%. This performance is extraordinarily bad in such a small database. Moreover, the illuminants that we have used are not extreme but rather are typical indoor lights with similar colour temperatures.

Test group by illuminant	No colour correction	Linear transform for colour correction	Diagonal transform for colour correction
Halogen	5	31	32
Macbeth 5000K	13	29	27
Philips ultralume	16	32	32
Sylvania cool white	14	31	32
Correct match ratio	48/132=36%	123/132=93%	123/132=93%
Clusterability (1-Interclass/betweenclass)	N/A (see text)	16%	16%

**Table 1: Matching ratio and cluster test by colour correction based on illuminant light**

Test group by illuminant	No colour correction	Linear transform for colour correction	Diagonal transform for colour correction
Halogen	5	33	32
Macbeth 5000K	13	31	30
Philips ultralume	16	32	32
Sylvania cool white	14	32	32
Correct match ratio	48/132=36%	128/132=97%	126/132=95%
Clusterability (1-Interclass/betweenclass)	No clustering	41%	31%

**Table 2: Match ratio and cluster test by colour correction based on image data**

To improve the colour indexing across illumination, a colour correction based on the measured illuminant is used. Given an illuminant measurement we calculate the best transform that takes image RGB responses to the reference lighting conditions (the same light that the database object was captured under). For example if our database is recorded under Halogen and a test image is recorded under cool white fluorescent then the transform that best maps fluorescent RGBs to Halogen counterparts is calculated. The mapping is based on the regression of corresponding RGBs for a Macbeth colour chart imaged under a pair of lights. The regression was either the best 3x3 linear transform relating illuminant pairs (since the linear model is theoretically sufficient) or the best 3x3 diagonal transform.

We now in a position to repeat the indexing experiment. Again, the model database is generated for each of the four illuminants. Test images are colour corrected to the model illumination using the diagonal and linear regressions that map Macbeth colour checker colours recorded under the test light to corresponding colours under the model illumination. It is clear from looking at Table 1 that correcting for illumination has dramatically increased indexing performance. The correct model image is recovered 93% of the time. The 7% of objects that didn't match were, in all cases the second best answer and so are 'soft' failures. These failures illustrate the limitation of indexing based on a 256 number chromaticity distribution representation (histograms for visually different scenes can look similar). The indexing experiment was repeated for a diagonal model of

illumination change. As might be expected (from the discussion at the end of the last section) a diagonal matrix delivers performance similar to the general 3x3 linear model.

Of course recognition rate is a single figure of merit that does not tell us much about the confidence in the match. It could be that the more accurate linear model of illumination change delivers much more reliable matches. To test this hypothesis we need to measure how clustered the data is (post colour correction). If histograms of the same object are very close to one another but those of different objects are very far apart then we can be very confident about the indexing. The closeness (actually the variance) of the 4 histograms (for each of the 11 objects) is calculated post-colour correction. This is divided by the variance of each cluster to the mean overall. The ratio of withinclass by betweenclass variance is a standard measure of how well data is clustered. The better the clustering the closer this ratio is to zero. Subtracting the cluster ratio from 1 gives us a % measure of clustering. This % clusterability is reported in the final row of Table 1. We see again that diagonal and linear models are equally good in terms of the clustering metric; 16% is recorded in both cases. Unfortunately, 16% clustering is quite small, informing us that within class variance is only marginally smaller than between class variance. For indexing without colour correction the clustering was so poor that the cluster ratio was bigger than 1 and so the % clusterability was less than 0 (evidence that the data is not clustered).

We repeated the indexing experiment but now calculated correction transforms based on the images themselves. For example, to calculate the mapping that takes images, of a particular object, recorded under cool white fluorescent and Halogen, we find the mapping that minimizes the error between all the RGBs in the two images. Such a regression can be performed because we have registered images. Indexing performance is assessed as before and is reported in Table 2 below. Notice there are fewer false matches. However, much more importantly the clustering of the data is increased. For the linear model of illumination change we have 41% clustering compared with 16% before. The diagonal model is similarly improved (though, relatively a little poorer). Clustering at a 31% level is indicative of confident indexing performance.

## 4 Conclusions and Further Works

Quantitative experimental results indicate that in order to correct for illumination bias in images it is necessary to know both the colour of the light and the colour statistics of surfaces present in a scene. But, surely, one can only account for the colour bias due to illumination by measuring the illumination? After all scene surface statistics are not known (indeed if they were then indexing across illumination would not be so problematic). In fact, illumination does sometimes depend on scene content since light that strikes a surface is reflected and this reflected light then strikes other surfaces and so on. Interreflection effects (and other factors) can constitute a significant part of the illuminant. However, these effects are difficult to measure and so the *effective* illumination is also difficult to measure. Unfortunately, if illumination is hard to measure it follows it must be even harder to estimate using a colour constancy algorithm and so colour constancy preprocessing may not, by itself, suffice to deliver image indexing across illumination (to date, it cannot [4]).

One way forward would be to index images based on image features that do not depend on illumination (the invariant indexing approach). These invariants relate image colours to other colours in a local image context where the nature of the relationship is independent of the illumination. Indeed, it is possible to view the illuminant+scene statistics results presented in this paper from the invariant viewpoint. When we correct colours based on the light colour and the scene content we are effectively calculating scene dependent invariants. A more detailed comparison of colour indexing and colour invariant indexing is presented in [18].

## Acknowledgements

This work was funded by EPSRC grant GR/L60852 (Colour and Texture Selective Indexing).

## References

- [1] Swain M. J. and Ballard D. H. Colour indexing. *International Journal of computer vision* 1991; 7(1):11-32.
- [2] Flickner M. et al, Query by image and video content - the QBIC system, *Computer*, 1995; 28(9): 23-32.
- [3] Swain M. J. Colour indexing. PhD thesis, University of Rochester, Department of computer science, 1990.
- [4] Funt B. V., Barnard K. and Martin L. Is machine colour constancy good enough. *ECCV'98*, 1998, 1:446-459.
- [5] Funt B. V. and Finlayson G. D. Colour constant colour indexing. *IEEE transactions on pattern analysis and machine intelligence* 1995; 17(5):522-529.
- [6] Healey G. and Slater D. Global colour constancy: recognition of objects by use of illumination invariant properties of colour distributions. *J. Opt. Soc. Am. A* 1995; 11(11):3003-3010.
- [7] Gevers T. Colour image invariant segmentation and retrieval, PhD thesis, ISBN 90-74795-51-X, University of Amsterdam, The Netherlands, 1996.

- [8] Maloney L.T. and Wandell B.A., Colour constancy: a method for recovering surface spectral reflectance, *J. Optical Society of America*, 1986; 3(1):29-33
- [9] Finlayson G.D., Schiele B and Crowley L J. Comprehensive colour image normalization, *Computer vision ECCV'98 1998; Volume I: 475-489*.
- [10] Maloney L T. Evaluation of linear models of surfaces spectral reflectance with small numbers of parameters. *J. Opt. Soc. Am. A* 1986; 3(10):1673-1683.
- [11] Marimont D. and Wandell B.A., Linear models for surface and reflectance, *J. Opt. Soc. Am. A* 1992; 9(11): 1905-13.
- [12] Forsyth D.A., A novel algorithm for colour constancy, *International Journal of Computer Vision*, 1990; 5(1):pp.5-36.
- [13] Land E. H. The retinex theory of colour constancy. *Scientific american*, 1977:108-129.
- [14] Worthey J. A. and Brill M. H., Heuristic analysis of Von Kries colour constancy. *J. Opt. Soc. Am. A* 1986;3(10):1708-1712.
- [15] Finlayson G.D., Drew M.S. and Funt B.V., Spectral sharpening: sensor transformations for improved colour constancy, *J. Optical Society of America*, 1994; 11(5): 1553-63.
- [16] Lin S. and Lee S.W., "Using chromaticity distributions and eignspace analysis for pose-, illumination- and specularly-invariant recognition of 3d object. *CVPR97*, 1997;426-431.
- [17] Hung T.W. R. and Ellis T. J., Spectral adaptation with uncertainty using maching, *IEE int. conf. On image processing*, Edinburgh, July, 1995:780-790.
- [18] Finlayson G. D. and Tian G. Y. "Perfect colour constancy vs. colour normalization for object recognition", *Colour & Imaging Institute, Technical note FAL/1*, 1998.