HOW USEFUL ARE COLOUR INVARIANTS FOR IMAGE RETRIEVAL?

Gerald Schaefer
School of Computing and Technology
Nottingham Trent University
Nottingham, U.K.
Gerald.Schaefer@ntu.ac.uk

Abstract The images captured by a digital camera do not depend only on the characteristics of the objects in the scene but also on the colour of the scene illumination. To account for this confounding influence various colour invariants have been introduced whose aim is to provide descriptors that do not change with a change in light. In this paper we evaluate the usefulness of these invariants for general purpose content-based image retrieval. The results show that the application of invariants may lead to a severe degradation in retrieval performance.

Keywords: content-based image retrieval, colour, colour invariants, colour normalisation

1. INTRODUCTION

It is a well known fact that colour poses an important cue for object recognition and image retrieval (Swain and Ballard, 1991). Colour indexing - matching based on the colour distribution of images - is at the heart of many image database systems e.g. (Niblack et al., 1993). Yet, colour cues are not always as useful as one might expect. The reason for this is that the colours in an image are not only a function of the surface characteristics of the objects in the scene but also of the illumination incident on these objects. This becomes apparent when we look at the physical process of image formation. An image taken with a device such as a digital colour camera is composed (in case of Lambertian surfaces) of sensor responses that can be described by

$$\underline{p}^x = \underline{e}^x \cdot \underline{n}^x \int_\omega S^x(\lambda)E(\lambda)\underline{R}(\lambda)d\lambda \tag{1}$$

where λ is wavelength, $\underline{p}$ is a 3-vector of sensor responses (usually RGB pixel values), S^x is the surface reflectance at location x, E the spectral power dis-

K. Wojciechowski et al. (eds.), Computer Vision and Graphics, 381–386.
© 2006 *Springer. Printed in the Netherlands.*

tribution of the illumination, and $\underline{R}$ is the 3-vector of sensitivity functions of the device. Integration is performed over the visible spectrum ω. The light reflected at x is proportional to $S^x(\lambda)E(\lambda)$, its magnitude is determined by the dot product $\underline{e}^x \cdot \underline{n}^x$ where $\underline{e}^x$ is the unit vector in the direction of the light source, and $\underline{n}^x$ is the unit vector corresponding to the surface normal at x.

From Equation (1) we can indeed see that the sensor responses $\underline{p}$ are inherently dependent on the illumination $E(\lambda)$. Consequently, two objects captured under different illuminations $E_1(\lambda)$ and $E_2(\lambda)$, $E_1(\lambda) \neq E_2(\lambda)$, will produce different images $\mathcal{I}_1$ and $\mathcal{I}_2$ (the N response vectors $\underline{p}$ are stacked together to form an $N \times 3$ image $\mathcal{I}$), $\mathcal{I}_1 \neq \mathcal{I}_2$. Fortunately the differences in $\mathcal{I}_1$ and $\mathcal{I}_2$ are not arbitrary. It can be shown that under reasonable assumptions the relationship between $\mathcal{I}_1$ and $\mathcal{I}_2$ can be expressed as a diagonal transform (Worthey and Brill, 1986, Finlayson et al., 1994)

$$\mathcal{I}_2 = \mathcal{I}_1 \mathcal{D} \tag{2}$$

i.e. the channels of $\mathcal{I}_1$ are a scaled version of those of $\mathcal{I}_2$.

Based on this relationship, various approaches at deriving algebraic combinations of image RGBs that do not change with a change of light - so-called colour invariants - have been presented in the literature (Buchsbaum, 1980, Funt and Finlayson, 1991, Finlayson et al., 1995, Finlayson et al., 1998, Gevers and Smeulders, 1999). The experiments conducted in these papers, based on objects captured under a variety of differently coloured lights, seem to confirm that colour invariants provide a useful addition to colour-based image retrieval systems.

In this paper we are asking whether colour invariants are really as useful for general purpose image retrieval as one is made to believe. Through an extensive set of experiments based on the MPEG-7 Common colour dataset (Moving Picture Experts Group, 1999) we seek to provide an answer to this. The rationale behind this is that in many cases colour invariance is not necessary: many images are captured under similar lighting conditions or rather under a set of lighting conditions which the colour constancy software integrated in digital cameras can cope with well (provided a certain colour variety in the scene). On the other hand, in order to achieve invariance some information has to be discarded - information that could potentially help to distinguish and hence successfully retrieve images.

The rest of the paper is organised as follows. The next section describes the set of colour invariants that we explored. Section 3 describes the experiments we conducted while Section 4 concludes the paper.

2. COLOUR MODELS AND COLOUR INVARIANTS

Eight colour invariants/normalisations were investigates each of which is briefly described in the following.

2.1 RGB space

RGB space is the sensor space of the colour device itself, and thus follows directly from Equation (1). Clearly, the RGB space does not have any invariance properties. We consider it here to provide a reference to which other algorithms can be compared.

2.2 Chromaticity space

Colours defined in chromaticity space are independent of intensity, and hence independent of the imaging geometry (the position of the objects and the light source) and shading effects. There are many possible chromaticity representations, one method is to divide RGBs by their respective intensities:

$$(r\ ,\ g\ ,\ b) = (\frac{R}{R+G+B}\ ,\ \frac{G}{R+G+B}\ ,\ \frac{B}{R+G+B}) \tag{3}$$

Of course, while Equation (3) factors out light intensity, dependence on light colour still remains.

2.3 Greyworld normalisation

Greyworld normalisation (Buchsbaum, 1980) divides each pixel by the average of the image

$$(\frac{R}{\text{mean}(R)}\ ,\ \frac{G}{\text{mean}(G)}\ ,\ \frac{B}{\text{mean}(B)}) \tag{4}$$

As a result the scaling factors that describe illumination change according to the diagonal model in Equation (2) cancel out. Thus Greyworld normalised images are independent of the colour of the incident light.

2.4 Comprehensive normalisation

As can we have seen from above, illumination independence may be found by applying Greyworld normalisation. On the other hand, chromaticity transforms provide invariance to intensity and hence the lighting geometry. By recursively applying each of these normalisations in turn, a procedure called Comprehensive colour image normalisation (Finlayson et al., 1998), it is possible to arrive at a representation that encapsulates both types of invariance.

2.5 Colour constant colour indexing (CCCI)

Funt and Finlayson (Funt and Finlayson, 1991) suggested the use of colour ratios

$$(\frac{p_1^1}{p_1^2}\ ,\ \frac{p_2^1}{p_2^2}\ ,\ \frac{p_3^1}{p_3^2}) \tag{5}$$

for object recognition, where p_i^j denotes the pixel value of the ith channel at location j. From the diagonal model of illumination change in Equation (2) it is straightforward to see that colour ratios stay constant across a change of light.

2.6 Colour angles

If we denote

$$v_i = p_{i1}, p_{i2}, p_{i3}, ..., p_{iN} \ , \ i = 1, 2, 3 \tag{6}$$

as the vectors of responses for each channel, then the colour angles of the image are defined as (Finlayson et al., 1995)

$$\phi_{ij} = \cos^{-1}(\frac{v_i v_j}{|v_i||v_j|}) \tag{7}$$

The three angles between the R, G, and B channels are invariant to the colour of the light source.

2.7 c1c2c3 space

In (Gevers and Smeulders, 1999) Gevers and Smeulders discuss colour based object recognition in the context of changing lighting conditions. They also propose three new colour models each having specific invariance properties. The c1c2c3 space, defined as

$$\begin{aligned} &(c1, c2, c3) = \\ &(\arctan(\frac{R}{\max\{G, B\}}) \ , \ \arctan(\frac{G}{\max\{R, B\}}) \ , \ \arctan(\frac{B}{\max\{R, G\}})) \end{aligned} \tag{8}$$

is invariant to shading due to the lighting geometry. However, this invariance holds only under white illumination.

2.8 l1l2l3 space

The l1l2l3 space ((Gevers and Smeulders, 1999))

$$\begin{aligned} (l1, l2, l3) = (\frac{(R-G)^2}{M} \ , \ \frac{(R-B)^2}{M} \ , \ \frac{(G-B)^2}{M}) \\ M = (R-G)^2 + (R-B)^2 + (G-B)^2 \end{aligned} \tag{9}$$

is invariant not only to intensity but also to the effect of highlights (specularities). As c1c2c3 it is restricted to white illumination.

2.9 m1m2m3 space

Gevers and Smeulders (Gevers and Smeulders, 1999) extended the CCCI technique (Funt and Finlayson, 1991) to account not only for varying illumination colour but also for shading effects. Their m1m2m3 space is defined as

$$(m1, m2, m3) = (\frac{p_1^1 p_2^2}{p_1^2 p_2^1}, \frac{p_1^1 p_3^2}{p_1^2 p_3^1}, \frac{p_2^1 p_3^2}{p_2^2 p_3^1}) \tag{10}$$

(again, p_i^j denotes the pixel value of the ith channel at location j).

3. EXPERIMENTAL RESULTS

We conducted our experiments using the MPEG-7 Common colour dataset (Moving Picture Experts Group, 1999). This database consists of 5466 images and a set of 50 queries with predefined ground truth images and has been designed as a benchmark set for colour-based image retrieval.

We use the MPEG-7 Normalised Modified Retrieval Rank (NMRR) (Moving Picture Experts Group, 1999) as the standard performance measure for this data set. The NMRR is defined as

$$\text{NMRR} = \frac{\text{MRR}(q)}{K + 1/2 - N_G/2} \tag{11}$$

where $\text{MRR}(q) = \mu(q) - 1/2 - N_G(q)/2$ and $\mu(q) = \sum_{i=1}^{N_G(q)} r_i/N_G(q)$. $N_G(q)$ is the number of ground truth images for the qth query image and r_i denotes the retrieved rank. For K we use the MPEG-7 recommendation $K = \min(4N_G(q), 2\max_q(N_G(q)))$.

Table 1. Retrieval results expressed in mean NMRR over the MPEG-7 dataset

Colour model	mean NMRR
RGB	0.1084
chromaticities	0.2653
Greyworld	0.2300
comprehensive normalisation	0.3893
CCCI	0.3873
Colour angles	0.5470
c1c2c3	0.2189
l1l2l3	0.3458
m1m2m3	0.4856

We implemented all colour invariants described in Section 2 and used $8 \times 8 \times 8$ histograms as image features (except for the colour angle invariants which are described by a 3-vector). (Dis)similarity is established through histogram intersection (Swain and Ballard, 1991) (respectively by Euclidean distance between colour angles vectors). The results for all invariants plus those for RGB histograms (similar to the original colour indexing method proposed in (Swain and Ballard, 1991)) are given in terms of mean NMRR over all 50 queries and listed in Table 1.

Looking at Table 1 we can see that colour-based retrieval based on RGB histograms works fairly well which is not a surprise since the MPEG-7 dataset has been designed for colour-based evaluation. However, looking at the results that were achieved by indexing invariant colour features a huge difference in retrieval performance is apparent. All colour invariants perform much worse than the standard RGB descriptors; the 'best' invariant model, c1c2c3 normalisation, provides a mean NMRR of 0.2189 - significantly higher of the 0.1084 achieved by RGB histograms.

These results confirm our earlier suspicion that prompted this investigation, namely that colour invariants are not always useful for the purpose of image retrieval. Unless it is known that illumination conditions might play a role for a certain image dataset it seems to be advisable *not* to use invariants and rather rely on unmodified colour descriptors and integrated colour constancy algorithms.

4. CONCLUSIONS

An investigation into the use and usefulness of colour invariants for content-based image retrieval was conducted. Image retrieval based on various colour invariants was performed on the MPEG-7 Common colour dataset. The results show that through the application of invariant features retrieval performance drops significantly and hence suggest their utilisation only upon prior knowledge of the application domain.

REFERENCES

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310:1–26.

Finlayson, G., Chatterjee, S., and Funt, B. (1995). Color angle invariants for object recognition. In *IS&T and SID's 3rd Color Imaging Conference*, pages 44–47.

Finlayson, G., Drew, M., and Funt, B. (1994). Color constancy: Generalized diagonal transforms suffice. *Journal Optical Society of America A*, 11(11):3011–3020.

Finlayson, G., Schiele, B., and Crowley, J. (1998). Comprehensive colour image normalization. In *5th European Conference on Computer Vision*, pages 475–490.

Funt, B. and Finlayson, G. (1991). Color constant color indexing. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(5):522–529.

Gevers, T. and Smeulders, A. (1999). Color based object recognition. *Pattern Recognition*, 32:453–464.

Moving Picture Experts Group (1999). Description of core experiments for MPEG-7 color/texture descriptors. Technical Report ISO/IEC JTC1/SC29/WG11/ N2929.

Niblack, W., Barber, R., Equitz, W., Flickner, M., Glasman, D., Petkovic, D., and Yanker, P. (1993). The QBIC project: Querying images by content using color, texture and shape. In *Conf. on Storage and Retrieval for Image and Video Databases*, volume 1908 of *Proceedings of SPIE*, pages 173–187.

Swain, M. and Ballard, D. (1991). Color indexing. *Int. Journal Computer Vision*, 7(11):11–32.

Worthey, J. and Brill, M. (1986). Heuristic analysis of Von Kries colour constancy. *Journal Optical Society of America A*, 3(10):1708–1712.