Using Local Binary Pattern Operators for Colour Constant Image Indexing

David Connah and Graham D. Finlayson; Department of Computer Science, University of East Anglia, UK

Abstract

The Local Binary Pattern (LBP) operator computes a local texture measure that is invariant to monotonic transformations of the image grey-scale. As a result of this property, calculating the LBP value in each channel of a colour image results in a triplet of values that are invariant to changes in illumination colour. Previous research has shown that histograms of grey-scale LBP values, and histograms of LBP values calculated in the R, G and B channels independently, form useful feature vectors for image retrieval. In this paper we generate 3D histograms of LBP values for colour images. Our experiments demonstrate that image retrieval performance, on a database of objects viewed under different illuminants, is greatly improved when using the 3D histogram compared to when the histogram is calculated for each colour channel independently. Furthermore, we find that this method can give better performance than using colour-histogram based features.

Introduction

Indexing images in large databases, that can be searched quickly, requires that the information in each image can be represented by a small number of features. It is therefore important to discover image features that facilitate good discrimination between different images and at the same time group similar images together.

Swain and Ballard [1] used the colour histogram of an image as a feature. Colour histograms have the advantage that they are robust to changes in the composition of scenes, and to partial occlusion of objects by other objects. A further property of colour histograms is that their elements, i.e. histogram bins, have a meaningful order; that is, bins which are close to each other in the histogram represent colours which look similar to one another. A problem with colour histograms however, as outlined in the original work of Swain and Ballard, is that they are not robust to changes in the colour of scene illumination; two scenes with identical composition but with differing illuminant colour would have different histograms, and hence be classed as different images. One solution to this problem is to apply a colour-constancy algorithm, or normalisation scheme, to the image prior to the calculation of the histogram [2, 3, 4]. A second approach is to derive properties from the image that are invariant to the colour of the illumination without explicitly transforming the image (e.g. [5]). In this work we follow the second approach and extract local texture features from an image that are invariant to changes in illumination colour.

Texture has also previously been employed as a feature for image retrieval [6, 7]. Textural features are generally extracted from a grey-scale image, which can be constructed from a colour image using a weighted average of the three colour channels. While this approach may be effective in some situations, it suffers from the same problem as the colour histogram; if the colour of the illumination in the scene changes, so does the grey-scale image and consequently so do the textural features. More recently, texture analysis methods have been proposed that consider colour and texture information together [8, 9, 10]. One such method, the Local Binary Pattern (LBP) operator [11] has been shown to be useful in several texture classification tasks and has been implemented, although as a grey-scale texture measure, in existing image retrieval systems [12].

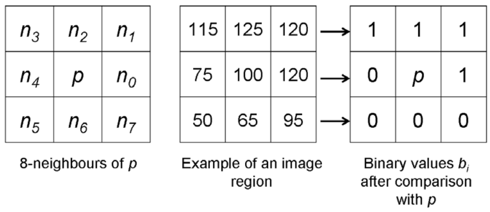

Figure 1. *An illustrative example of the LBP calculation*

The LBP operator characterises the local variation in grey-scale around a pixel, and is invariant to any monotonic transformation of the grey-scale. It follows that the LBP value calculated at each pixel in each colour channel will be unchanged by a positive scaling of the channel output. This is a very useful property, since a change in illumination colour generally results in an independent scaling of each colour channel [13], thus the LBP value in each colour channel is approximately illumination independent. Moreover, even when the scalar model of image change does not hold, typically the rank order of intensities in each channel is preserved. Surprisingly, this property is even true of images taken with different cameras [14]. As long as the rank ordering of pixel intensities is maintained, the LBP value will also be unchanged. As a result of this, LBP-based features should be robust to changes in both the colour of the illumination and the camera used to capture the image.

In this work we calculate the LBP value in the *R*, *G* and *B* colour channels, thereby creating a new colour texture image. We then characterise the images by their 3D LBP histogram. This is an extension of the work of Mäenpää and Pietikäinen [9], who computed LBP histograms in three colour channels independently. However, computing a 3D histogram retains more information about the underlying image; information which allows us to improve the performance of the LBP operator in a colour image retrieval task.

Background

When the colour or geometry of the light-source illuminating a scene is changed, there is a change in the image recorded by a digital camera. A result of this is that the distribution of camera responses elicited by different objects can no longer be used as a cue to identify them [1]. One way to tackle this problem is to transform the images in such a way as to discount the effect of the illumination (i.e. apply a colour-constancy algorithm) and to use colour distributions in these transformed images as features.

An alternative approach is to derive features directly from the image that are invariant to a change in illumination. The success of both these approaches is reliant upon the nature of the image transformation induced by an illumination change; this topic has therefore been well studied [13, 15].

If we represent the camera response at a single pixel as a vector $\mathbf{r} = [R \quad G \quad B]^T$, where R, G, and B are the responses of the red, green and blue channels respectively, then the change of illumination can be written as a transformation of responses $\mathbf{r}$ to responses under a second illuminant $\mathbf{r}'$:

$$\mathbf{r}' = T(\mathbf{r}). \tag{1}$$

For most changes in illumination that will be encountered in the natural world, it has been found that $T(\cdot)$ is well approximated by a 3×3 linear transform $\mathbf{A}$:

$$\mathbf{r}' = \mathbf{A}\mathbf{r}. \tag{2}$$

The structure of the matrix $\mathbf{A}$ has also been observed to be tightly constrained; in many cases $\mathbf{A}$ can take the form of a positive diagonal matrix [13], i.e. with positive values along the leading diagonal and zeros elsewhere. Even when $\mathbf{A}$ is not diagonal, a fixed transformation can be applied to the camera responses such that that a diagonal transformation does serve as a good model [15]. More recently it has been found that if the the camera responses in each channel are ranked in order of their magnitude, a change in illumination will not alter this ranking [14]. While this is clearly true in the diagonal case (multiplying a set of numbers by a positive scalar cannot change their ordering), it is an interesting empirical result that this holds for non-diagonal illumination changes. Surprisingly, the rank-ordering of responses in each channel is also preserved across images taken with different cameras. It is precisely this property that we aim to exploit in this paper by computing an image feature that is invariant to changes that preserve the rank ordering of pixel responses.

The feature that we employ is the Local Binary Pattern (LBP operator), which was developed by researchers at Oulu University [11, 9, 12]. The goal of the operator is to characterise the pattern around each pixel in an image. To understand the concept of LBP we start by considering a grey-scale image whose intensity I at pixel location (x,y) can be written $I(x,y)$. For a given pixel $p = I(x_0,y_0)$, the 8 neighbours of p can be written as $n_i, i = 0\ldots,7$, where $n_0 = I(x_0+1,y_0), n_1 = I(x_0+1,y_0+1),\ldots,n_7 = I(x_0+1,y_0-1)$ (see Figure 1). To compute an LBP value, the grey-level value of each neighbour n_i is compared to the value at the central pixel p to determine whether it is greater than or less than p. This amounts to a function which maps each n_i onto a value b_i as follows:

$$b_i = \begin{cases} 1 & \text{if } n_i \geq p \\ 0 & \text{if } n_i < p. \end{cases} \tag{3}$$

The LBP value for pixel (x_0,y_0) is derived by choosing an arbitrary starting point, say b_0 and concatenating the 8 binary values b_i into a an 8-bit number; this operation can be written as follows:

$$\mathrm{LBP}(x_0,y_0) = \sum_{i=0}^{7} b_i 2^i. \tag{4}$$

For a given pixel the LBP value is, by definition, invariant to any transformation of the image grey-scale that preserves the rank-ordering of pixel intensitites (i.e. a monotonic transformation). This property is a consequence of the binary operation outlined in Equation 3.

Mäenpää and Pietikäinen [9] extended the use of LBP to colour images. They investigated two different methods for doing this. One method is to compute the LBP value in each of the R, G, and B colour channels separately. Thus at each pixel one obtains a triplet of LBP values. Given that the LBP value in each channel is invariant to image transformations that preserve pixel rank-ordering, and that most illumination changes preserve rank-ordering [14], we expect this triplet of values to be invariant to illumination changes. Their second approach is to compute LBP with "cross-terms". In this method the central pixel p and neighbourhood pixels n_i are taken from different colour channels. This approach is interesting as it takes into account interactions between different channels. However, this approach is unlikely to yield features that are invariant to illumination change. Indeed, Mäenpää and Pietikäinen employ an image normalisation to discount illumination effects prior to computing LBP values in this way.

After the computation of the LBP value at each pixel, the whole image can be characterised by the histogram of LBP values [11]. In previous work, LBP histograms have been derived for both grey-scale [11] and colour images [9]. For colour images the histogram has been computed for each colour channel separately, giving three histograms which can then be grouped into a single feature vector [9]. However, whether computing LBP histograms for a grey-scale image, or for the individual channels of a colour image, some information is invariably lost. Clearly the grey-scale approach loses colour information, but it is also not invariant to illumination change. The colour approach is both invariant to illumination change and preserves some colour information. Unfortunately, by computing three separate histograms, one for R, G and B, information about the interaction between the channels is lost. In this work we aim to keep some of this information by, instead of computing a histogram in each colour channel, computing a single joint histogram for the whole image. We will evaluate the efficacy of this approach by using it as a feature for an image retrieval task.

Implementation

In the following experiments we test two approaches to computing LBP histograms for colour images; computing the histograms separately for each colour channel, and computing a joint histogram. The first stage in the histogram calculation is to generate an LBP image for a each colour channel. For a given pixel (x,y) and colour channel k, $k \in \{R,G,B\}$ the LBP value can be denoted $\mathrm{LBP}_k(x,y)$ and can take a discrete integer value from 1 to Q, where Q is the total number of possible LBP values.

The number of possible LBP values Q is a variable that is dependent upon several factors. In the treatment in Section the image is sampled at 8 points around the central pixel p. This produces an 8-bit number at each pixel and hence $Q = 256$. However, we are not restricted to this neigbourhood. Ojala *et al.* [11] use a circularly symmetric neighbourhood, sampling at points that are equidistant from p and from one another. When a sample point does not fall exactly on a pixel, its value can be calculated using bilinear interpolation. Employing this method facilitates the use of any number of sample points. Similarly the radius of the sampling circle can be varied freely in order to capture features at different spatial scales. Here we employ this circular neighbourhood, and incorporate the radius size as a variable to be optimised.

The LBP operator can be implemented in a number of different forms. In [11] Ojala *et al.* derive a form, LBP^{ri}, which is invariant to rotations in the underlying pattern. The computation of this rotation invariant form can be written as follows:

$$\mathrm{LBP}^{ri} = \min\{\mathrm{ROR}\,(\mathrm{LBP}, i)\,|\, i = 0, 1, 2, \ldots, N-1\}, \tag{5}$$

where the ROR function shifts the N-bit binary value LBP, i bits to the right with wrap-around. This step not only makes the operator invariant to rotations in the image, it reduces the number of possible patterns Q significantly.

A further reduction in the number of patterns can be achieved by only considering patterns of a specific "uniformity" [11]. Uniformity is defined as the number of transitions that occur in the binary string $b_0 b_1 \ldots b_{N-1}$, where a transition is a change from a 1 to a 0 or vice versa. The uniformity calculation can be written as follows:

$$\text{uniformity} = \|b_0 - b_{N-1}\| + \sum_{i=0}^{N-2} \|b_i - b_{i+1}\|. \tag{6}$$

It has been observed [11] that in many natural images, patterns with uniformity ≤ 2 account for the majority of patterns in the image. As a result, little information is lost by only considering those patterns. These patterns are also important as they correspond approximately to image features such as edges, line endings and corners. Patterns with uniformity ≤ 2 each consist of a chain of ones (or zeros if you prefer) with a variable length. The length of the chain determines which histogram bin the pattern corresponds to. Each chain can be of length zero up to length N, where N is the number of sample points. Following [11] we can collect patterns of a higher uniformity than two in a single histogram bin, which gives $N+2$ different bins in each channel. A good property of this approach is that the histogram bins have a meaningful order; elements that are close to one another have similar chain lengths and thus represent similar patterns.

The next step in the procedure is to build the histograms. If we consider the colour channels independently, then we need to create three "marginal" histograms, H_k:

$$\begin{aligned} H_k(u) &= \sum_x \sum_y f(x,y) \qquad (7) \\ \text{where} \quad f(x,y) &= \begin{cases} 1 & \text{if } \mathrm{LBP}_k(x,y){=}u \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

where $H_k(u)$ denotes the u-th histogram bin. The final colour histogram H_M is formed by concatenating these three histograms:

$$H_M = [H_R, H_G, H_B]. \tag{8}$$

The joint histogram H_J is calculated as follows:

$$\begin{aligned} H_J(u,v,w) &= \sum_x \sum_y g(x,y) \qquad (9) \\ g(x,y) &= \begin{cases} 1 & \text{if } \mathrm{LBP}_R(x,y){=}u \\ & \text{and } \mathrm{LBP}_G(x,y){=}v \\ & \text{and } \mathrm{LBP}_B(x,y){=}w \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

As a result of the refinement of the LBP calculation outlined above, the number of possible LBP patterns is reduced from 256 to 10 for each pixel in each colour channel. Thus when the LBP histogram is computed independently for each channel, there are a total of 30 bins over all three colour channels. A joint histogram requires $10 \times 10 \times 10 = 1000$ bins. In preliminary expermiments we found that the bins which correspond to patterns of "uniformity $= 0$" and "uniformity > 2" were not helpful in discriminating between the different objects, thus we discarded them. This leaves 7 unique LBP values, which in turn gives 21 bins for the independent histograms and 343 bins for the joint histograms.

Experimental

We test the efficacy of the different features using an image retrieval task. We use a database of images created by researchers at Simon Fraser University [1], which contains 20 objects, each under 11 different conditions. In each condition both the object's pose and the colour of the illumination are changed. For each object we choose a single condition to act as a model image; in this case we use the object captured under the "solux 3500" light source so that we can compare our results with those from an earlier study by Finlayson and Xu [4]. In total this gives us 20 model images which comprise the "image database" and 200 test images.

In the study by Finlayson and Xu the authors used colour-based image features. These included standard colour histograms and colour histograms computed after image "normalisation". The normalisation schemes removed both the effect of illuminant colour and the effect of lighting geometry. The results from their work act as a benchmark for our experimental results.

The distance between two images is measured using the L_1 norm of the difference between their LBP histograms. For a pair of Q-bin histograms H^1 and H^2 this distance is given by:

$$L_1(H^1, H^2) = \sum_{i=1}^{Q} |H^1(i) - H^2(i)| \tag{10}$$

For each test image we compute the distance to each of 20 model images and then rank these distances from 1 for the smallest L_1 norm to 20 for the largest. We then follow Swain and Ballard and compute the match percentile for each test image, which is given by $(20-R)/19$ where R is the rank of the correct model-image. This value is 1 when the correct model-image is ranked first and 0 when it is ranked last (20^{th}).

In addition to the Simon Fraser Data we also compare the methods on the images used by Swain and Ballard [1]. In this data set there are 66 model images and 30 test images. Each test image contains an object from the model database that is either rotated or partially occluded. The illumination colour is the same for both model and test images. As before, for each test-image we compute the distance to each model image and rank these resulting distances. The accuracy of the classification is summarised by the match percentile.

Results

Tables 1 and 2 show the retrieval results for the different methods. From both Tables it is clear that the performance for the joint histogram (LBP 3D) is significantly better than that when the histograms are calculated independently from one another (LBP 1D). In Table 1 the performance of the 3D LBP histogram exceeds that of the best colour normalisation scheme. Unsurprisingly, the colour histogram is not a good feature for colour constant image retrieval. In Table 2 we can see that, when the illumination colour is not varied, the colour histogram performs well. For this data the 3D LBP histogram also gives good results, although not as good as the best normalisation routine. A contributory factor to this result may be that the image-normalisation

[1] This database can be downloaded from http://www.cs.sfu.ca/ colour/data/

Figure 2. *A ball viewed under syl-wwf illumination*

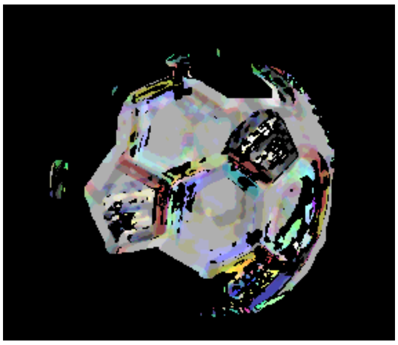

Figure 3. *An "LBP image" for the ball under syl-wwf, black areas correspond to regions that were removed from the histogram*

takes account of lighting geometry as well as illumination colour, whereas the LBP measure is invariant only to the colour of the illumination.

Figure 2 shows an object from the database under the "syl-wwf" lighting condition. In Figure 3 is a visual representation of the LBP image, where the display colour at each point is the triplet $[\mathrm{LBP}_R, \mathrm{LBP}_G, \mathrm{LBP}_B]$ re-scaled onto the range 0 to 255. It is this image that is invariant to changes in the colour of the illumination.

Discussion and Conclusions

We investigated the application of the LBP texture operator to the retrieval of coloured objects under changes in illumination colour and object pose. We compared two LBP-based approaches with two colour-based approaches. Using the joint histogram of LBP values gave significantly better performance than calculating histograms for the colour channels separately, although we concede that the joint histogram also uses significantly more parameters. The joint LBP histogram also compared favourably with the best colour-based approach, especially when model and test images were captured under different illuminants. This experimental result provides support for the proposition that the LBP operator is robust to real changes in illumination colour. However, the LBP histogram computed independently in each channel should be just as invariant to illumination change. This indicates that the performance of the joint LBP histograms is a function of both their illumination invariance, and their ability to encode additional information about the interaction between the colour channels.

Acknowledgments

The authors would like to thank Steve Hordley for useful comments on the manuscript. Graham Finlayson also gratefully acknowledges the support of the Leverhulme Trust.

References

[1] M. J. Swain and D. H. Ballard. Color indexing. *International Journal of Computer Vision*, **7**(1), 11–32 (1991).

[2] B. Funt and K. Barnard. Is machine color constancy good enough? In *Proc. 5th European Conference on Computer Vision*, pp. 455–459 (1998).

[3] G. D. Finlayson, S. Hordley, and P. Hubel. Illuminant estimation for object recognition. *COLOR research and application*, **27**(4), 260–270 (2002).

[4] G. D. Finlayson and R. Xu. Illuminant and gamma comprehensive normalisation in log rgb space. *Pattern Recognition Letters*, **24**, 2679–2690 (2003).

[5] D. A. Slater and G. Healey. The illumination-invariant recognition of 3d objects using local color invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**(2), 206–210 (1996).

[6] A. Berman and L. Shapiro. A flexible image database system for content based retrieval. *Computer Vision and Image Understanding*, **75**, 175–195 (1999).

[7] M. Lew. Next-generation web searches for visual content. *IEEE Computer*, **33**, 46–53 (2000).

[8] D. K. Panjwani and G. Healey. Markov random field models for unsupervised segmentation of textured color images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**(10), 939–954 (1995).

[9] T. Mäenpää and M. Pietikäinen. Classification with color and texture: jointly or separately? *Pattern Recognition*, **37**, 1629–1640 (2004).

[10] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih. Image indexing using color correlograms. In *IEEE Computer Vision and Pattern Recognition*, pp. 762–768 (1997).

[11] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(7), 971–987 (2002).

[12] T. Mäenpää and M. Pietikäinen. Texture analysis with local binary patterns. In C. H. Chen and P. S. P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision, 3rd Edition*, pp. 197–216. World Scientific (2005).

[13] J. A. Worthey and M. H. Brill. Heuristic analysis of von kries color constancy. *Journal of the Optical Society of America - A*, **3**(10), 1708–1712 (1986).

[14] G.D. Finlayson, S. Hordley, G. Schaefer, and G.Y. Tian. Illuminant and device invariant colour using histogram equalisation. *Pattern Recognition*, **38**(2), 179–190 (2005).

[15] G.D. Finlayson, M.S. Drew, and B.V. Funt. Color constancy: Generalized diagonal transforms suffice. *Journal of the Optical Society of America - A*, **11**(11), 3011–3019 (1994).

Author Biography

David Connah obtained a BSc in Biology in 1997 and an MSc in Machine Perception and Neurocomputing in 1998, both from the University of Keele, England. In 2004 he completed his PhD in Colour Science at the Colour and Imaging Institute, University of Derby. After his PhD he worked as a researcher in

Results of different matching regimes for objects under different lights (Simon Fraser Data)

Method	Average percentile	Rank 1	Rank 2	Rank > 2	Worst Rank
Colour histogram	73.50	28	12	60	20 out of 20
Best normalisation	99.13	91	3.5	5.5	5 out of 20
LBP 1D	90.58	45	21	34	16 out of 20
LBP 3D	99.50	94.5	3.5	2	6 out of 20

Results of different matching regimes for objects under the same light (Swain and Ballard data)

Method	Average percentile	Rank 1	Rank 2	Rank > 2	Worst Rank
Colour histogram	98.54	93.33	3.33	3.34	9 out of 66
Best normalisation	99.59	83.33	13.33	3.34	5 out of 66
LBP 1D	94.05	46.67	16.67	36.66	19 out of 66
LBP 3D	98.72	80	16.67	3.33	21 out of 66

the Norwegian Color Research Lab at Gjøvik University College, and is currently working with the Colour Group at the University of East Anglia.