Indexing chromatic and achromatic patterns for content-based colour image retrieval

G. Qiu*

School of Computer Science, The University of Nottingham, Jubilee Campus, Nottingham NG8 1BB, UK

Abstract

In this paper, we present a method to represent achromatic and chromatic image signals independently for content-based image indexing and retrieval for image database applications. Starting from an opponent colour representation, human colour vision theories and modern digital signal processing technologies are applied to develop a compact and computationally efficient visual appearance model for coloured image patterns. We use the model to compute the statistics of achromatic and chromatic spatial patterns of colour images for indexing and content-based retrieval. Two types of colour images databases, one colour texture database and another photography colour image database are used to evaluate the performance of the developed method in content-based image indexing and retrieval. Experimental results are presented to show that the new method is superior or competitive to state-of-the-art content-based image indexing and retrieval techniques. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Colour vision; Colour imaging; Chromatic patterns; Achromatic patterns; Content-based image indexing and retrieval; Image database; vector quantization

1. Introduction

Colour plays a very important role in current state-of-the-art content-based image retrieval from image databases. Since its first publication, Swain's colour histogram indexing technique [1] has played a very influential role in the research of content-based image indexing and retrieval. Since colour histogram is a global measure and contains no information about the spatial distribution of the colours, the success of the colour histogram method is quite limited. Recognising the importance of spatial information, researchers have proposed various methods to incorporate such information in the indexing and retrieval for image databases. Although it is not possible to be exhaustive, representative approaches to incorporating the spatial information into content-based image indexing and retrieval include

classical texture descriptors [2], image segmentation (blobworld) [3], joint histogram of colours and edges [4], and colour correlogram [5]. Recent work has also proposed the use of binary space partitioning tree as an efficient representation for capturing the colours and their spatial distribution information [6].

In this paper, we present an alternative approach to construct effective and efficient image features for content-based image indexing and retrieval. We have developed a method to capture statistically representative chromatic and achromatic spatial image patterns and use the distributions of these patterns within an image to characterise the image's visual content. In developing our new method, we have used some well-known and relatively new colour vision theories, and importantly, we have also based our development on modern digital signal processing technologies.

A well-known colour vision theory is the opponent colour theory [7], which suggests that there are three visual pathways in the human colour vision system. One

---

* Corresponding author. Fax: +44-115-951-4254.
  E-mail address: qiu@cs.nott.ac.uk (G. Qiu).

pathway is sensitive mainly to light–dark variations; this pathway has the best spatial resolution. The other two pathways are sensitive to red–green and blue–yellow variation (the opponent channels). The blue–yellow pathway has the worst spatial resolution. In opponent-colour representation, the spatial sharpness of a colour image depends mainly on the sharpness of the light–dark component of the images and very little on the structure of the opponent-colour image components. There is evidence to suggest that different visual pathways process colour and spatial pattern in the human visual system. The pattern-colour-separable model of human colour vision [8] is a relatively new colour vision theory derived from psychovisual experiments. This model suggests that the value of one neural image is the product of three terms. One term defines the pathway's sensitivity to the stimulus colour direction. A second term defines the pathway's sensitivity to the spatial patterns of the stimulus and the third term defines the pathway's sensitivity to the stimulus strength.

Colour vision theories could provide some guidance in the development of colour image processing algorithms.[1] According to the opponent colour vision model, chromatic and achromatic signals should be treated differently, achromatic signal should be given higher bandwidth and chromatic signal can be given lower bandwidth. In fact, this theory has long been successfully used in the broadcasting of colour television signals [9]. Based on the colour and pattern separable model [8], we can have separate channels to process spatial and spectral information. Based on above reasoning, we have developed a digital signal processing method to capture the statistics of colour images for content-based indexing and retrieval. The method uses an opponent colour representation, and derives an efficient representation and coding scheme to characterise the chromatic and achromatic spatial patterns of small neighbourhood of pixels. We use the statistics of these chromatic and achromatic spatial patterns as the images' indices for effective content-based image indexing and retrieval for image database applications.

The organisation of the paper is as follows. In Section 2, we present an efficient method for capturing the statistics of chromatic and achromatic spatial patterns of colour images. Section 3 presents the use of the method in content-based image indexing and retrieval. Section 4 presents extensive experimental results and Section 5 concludes the paper.

2. Modelling coloured image patterns

A coloured image pattern is defined as the spatial and spectral characteristics of a (small) block of pixels in a colour image. We aim to develop efficient and effective representation (coding) schemes to describe coloured image patterns. To fully exploit the colour vision properties of human visual system, we use an opponent colour space to represent colour images. Although many variations of colour space exist, we have chosen the YC_bC_r space [10] (similar colour space can also be used). The relation between this space and the better-known RGB colour space is as follows:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.169 & -0.331 & 0.500 \\ 0.500 & -0.419 & -0.081 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (1)$$

To make the computation manageable, we want to work on relatively small-size coloured image patterns.[2] For a coloured image pattern, we want to develop representation schemes which, (1) can faithfully capture the visual appearance of the pattern, (2) are computationally efficient and (3) are compact. We have developed the coloured pattern appearance model (CPAM) [12] which has two channels capturing the characteristics of the chromatic and achromatic spatial patterns. With reference to the schematic of CPAM in Fig. 1, the visual appearance of a small image block is modelled by three components:[3] the stimulus strength (**S**), the spatial pattern (**P**) and the colour pattern (**C**). For a small image area, the stimulus strength **S** is approximated by the local mean of the Y component. The pixels in Y normalised by **S** form the spatial pattern. Since C_b and C_r have lower bandwidth, they are sub-sampled by a factor of 2 in both dimensions. The sub-sampled pixels of C_b and C_r are normalised by **S**, to form the colour pattern (**C**)

---

[1] The author fully understands that our knowledge about our own visual system is still very limited, and theories concerning the visual system are widely and intensively debated in the community. It is not the intention of the author to go into the debate about the correctness of these theories. We shall simply use these theories and aim to draw useful guidance from them for building practical computational models for colour image processing.

[2] We will only work on a single spatial resolution in this paper. There are theories to suggest that human visual system represents visual signals in coarse and fine multi-resolutional decomposition [11]. We will extend the model to multi-resolution patterns in the future.

[3] Please note that the CPAM model is influenced by the pattern-colour separable model (PCSM) of human colour vision [8], which states that there are three (pattern, colour and strength) visual pathways within the human colour visual system. CPAM also has three channels resemble those of PCSM, however, it should be clear that CPAM is not a truthful interpretation of PCSM, rather, it is a pragmatic signal processing model.
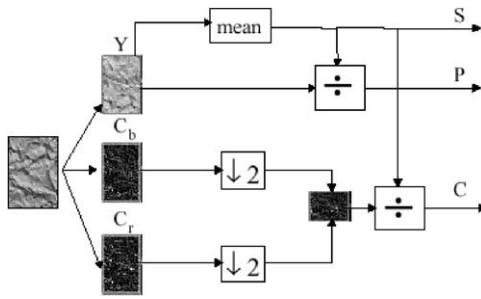
Fig. 1. Coloured Pattern Appearance Model (CPAM). The visual appearance of a small image block is modelled by three components: the stimulus strength (**S**), the achromatic spatial pattern (**P**) and the chromatic spatial pattern (**C**).

component of the appearance model. Normalising the pattern and colour channels by the strength has the effect of removing (to a certain extent) the effects of lighting conditions, making the visual appearance model somewhat "colour constant" [13]. Colour constancy is another important property of human colour vision system [7]. We shall demonstrate some possible benefits of this normalisation in the experimental results.

The **P** channel of CPAM captures the achromatic spatial pattern of the input coloured image pattern, and the **C** channel captures the chromatic spatial pattern. From now on, we shall call **P** and **C** the achromatic spatial pattern (ASP) and chromatic spatial pattern (CSP), respectively.

In constructing the ASP and CSP pattern representations, we have taken into account the argument that the appearance of a coloured image pattern is pattern and colour separable [8]. We interpret this as that we can model the spatial patterns in different chromatic signals independently, hence the chromatic and achromatic patterns are modelled separately. By using an opponent colour representation, we also have taken into account the fact that opponent chromatic signals have lower bandwidth, thus can be sub-sampled to form a lower dimensional vector (pattern). By separating achromatic and chromatic signals, we can work on two low-dimensional vectors rather than one very high dimensional vector. From a computational point of view, low dimensionality is advantageous because higher dimensional signals will suffer from the "curse of dimensionality" phenomenon.

Therefore, from both colour vision theories viewpoint and digital signal processing viewpoint, it is justifiable to represent the chromatic and achromatic spatial patterns in such a way. In order to use the representation scheme in content-based image indexing and retrieval, it is necessary to collect statistics of the distributions of these patterns (ASP and CSP) within the image. However, what shapes and forms of ASPs and CSPs will appear in the images are not known. It is therefore necessary to use some forms of statistical techniques to estimate the most likely

shapes and forms of ASPs and CSPs in natural colour images, i.e., we need a compact representation scheme for these patterns. In modern digital signal processing, vector quantization (VQ) [14] is a well-developed statistical technique which can estimate statistically most representative feature vectors in the feature space. We can therefore adopt VQ to estimate the models of ASPs and CSPs.

2.1. Vector quantization

VQ [14] is a mature method of lossy signal compression/coding in which statistical techniques are used to optimise distortion/rate trade-offs. A vector quantizer is described by an encoder Q, which maps the k-dimensional input vector X to an index $i \in I$ specifying which one of a small collection of reproduction vectors (codewords) in a codebook $C = \{C_i; \ i \in I\}$ is used for reconstruction, and there is also a decoder, Q^{-1}, which maps the indices into the reproduction vectors, i.e., $X' = Q^{-1}(Q(X))$. The key to the successful use of VQ technology is to obtain a well designed codebook.

There are many methods developed for designing the VQ codebook. The K-means types algorithms, such as the LGB algorithm [14], and neural network based algorithms, such as the Kohonen feature map [15] are popular tools. In this work, we used a specific neural network-training algorithm, the frequency sensitive competitive learning (FSCL) algorithm [16] to design our codebook. According to our own experience, we find FSCL is insensitive to the initial choice of codewords, and the codewords designed by FSCL are more efficiently utilised than those designed by methods such as the LGB algorithm. The FSCL method can be briefly described as follows:

2.1.1. FSCL VQ design algorithm

1. Initialise the codewords, $C_i(0)$, $i = 1, 2, \ldots, I$, to random numbers and set the counters associated with each codeword to 1, i.e., $n_i(0) = 1$,
2. Present the training sample, $X(t)$, where t is the sequence index, and calculate the distance between $X(t)$ and the codewords, $D_i(t) = D(X(t), C_i(t))$, and modify the distance according to $D_i'(t) = n_i(t)D_i(t)$
3. Find j, such that $D_j'(t) \leqslant D_i'(t)$ for all i, update the codeword and counter

$$C_j(t+1) = C_j(t) + a[X(t) - C_j(t)]$$

$$n_j(t+1) = n_j(t) + 1,$$

where $0 < a < 1$ is the training rate.
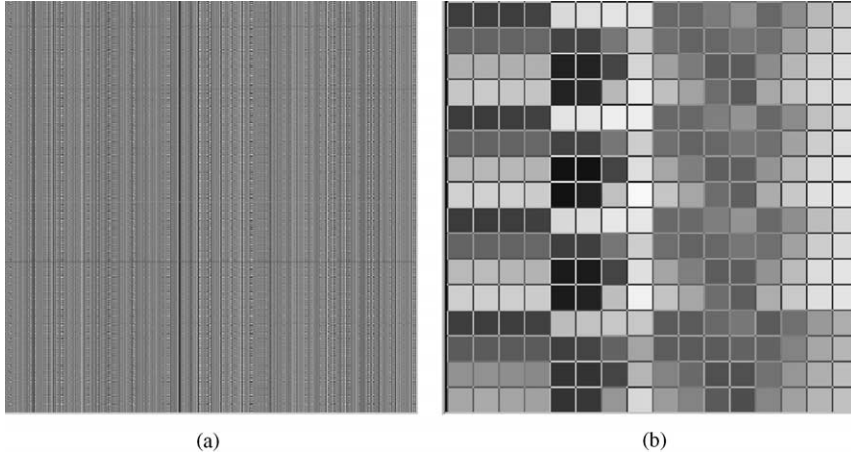4. Repeat by going to 2.

Fig. 2. The codebook image. (a) This is a 256×256 2-dimensional array of 4×4 pixel block patterns (the image is 1024×1024 pixels); each 4×4 pattern is constructed by one of the 256 ASP codewords and one of the 256 CSP codewords; the CSP codewords have been up-sampled by nearest-neighbour interpolation [9] to scale back to the original block size (4×4). For displaying purpose, all values of the codewords have been scaled by a constant factor of 128. (b) This is an enlarged portion of the image (a), as can been seen these 4×4 patterns exhibits various colours and spatial patterns.

2.2. VQ coding of ASP

Let $Y = \{y(i,j), \ i = 0, 1, 2, m, \ j = 0, 1, 2, n\}$ be the $m \times n$ Y image block. The stimulus strength of the block is calculated as

$$S = \frac{1}{m \times n} \sum_{i=0}^{m} \sum_{j=0}^{n} y(i,j). \tag{2}$$

Then the ASP pattern vector, $ASP = \{asp(i,j), \ i = 0, 1, 2, m, \ j = 0, 1, 2, n\}$ of the block, is formed as

$$asp(i,j) = \frac{y(i,j)}{S}. \tag{3}$$

An N-dimensional vector quantizer ($N = m \times n$), Q_p, with a codebook $C_P = \{C_P(i); \ i \in I\}$ of size I, where $C_P(i)$ is the ith N-dimensional codeword, is then designed for ASP using many training samples. In this work, two hundred and thirty-five 512×512 pixels true colour images, which form more than 15 million samples have been used to design C_P.

2.3. VQ coding of CSP

The CSP vector is formed by sub-sampling the two chromatic channels, C_b and C_r to form a single vector, which is also normalised by S. In the case where a $m \times n$ block size is used, the CSP vector is a $2M$-dimensional vector ($M = m \times n/4$). Let $C_b = \{c_b(i,j), \ i = 0, 1, 2, m, \ j = 0, 1, 2, n\}$ and $C_r = \{c_r(i,j), \ i = 0, 1, 2, m, \ j = 0, 1, 2, n\}$ be the two corresponding chromatic blocks of the C_b and C_r channels. Then the sub-sampled C_b signal, $SC_b = \{sc_b(k,l),$

$k = 0, 1, 2, m/2, \ l = 0, 1, 2, n/2\}$ is obtained as follows:

$$sc_b(k,l) = \frac{1}{4S} \sum_{i=0}^{1} \sum_{j=0}^{1} c_b(2k+i, 2l+j). \tag{4}$$

Similarly, the sub-sampled C_r signal, $SC_r = \{sc_r(k,l),$ $k = 0, 1, 2, m/2, \ l = 0, 1, 2, n/2\}$ is obtained as follows:

$$sc_r(k,l) = \frac{1}{4S} \sum_{i=0}^{1} \sum_{j=0}^{1} c_r(2k+i, 2l+j). \tag{5}$$

The CSP vector, $CSP = \{csp(k), \ k = 0, 1, \ldots, 2M\}$ is formed by concatenating SC_b and SC_r. A $2M$-dimensional vector quantizer, Q_c, with a codebook $C_c = \{C_c(j); \ j \in J\}$ of size J is then designed. Again over 15 million samples have been used to design C_c.

2.4. Designing ASP and CSP codebooks

The VQ codebooks should capture statistically representative feature vectors from the input space. It is therefore necessary to have as many training samples as possible to design the codebooks. In this work, we have used 235 true colour (24 bits/pixel) images from the MIT Media Labs VisTex image database[4] to obtain the training samples. The size of these images is 512×512 pixels. Although it is possible to choose various block sizes, we have chosen the size of the pattern to be 4×4 pixels for computational convenience. Therefore, the ASP vectors are 16-dimensional and the CSP vectors are 8-dimensional. We used the FSCL algorithm

_____

[4] http://vismod.www.media.mit.edu/vismod/imagery/Vision Texture /vistex.html

described above to train our codebooks and more than 15 million training samples were obtained from these 235 images. The training was stopped after each of the 15 million samples had been presented for training 20 times. We have trained a 256-codeword quantizer for the ASP vectors and a 256-codeword quantizer for the CSP vectors, respectively. Using a Pentium III PC, the total time taken to train the codebooks was about 9 h. In Fig. 2, we have plotted these codewords as an image. As can been seen, a combination of these two sets of codewords can represent a very diverse set of image patterns.

3. Indexing coloured image patterns for content-based image retrieval

The visual appearance of a coloured image pattern can be completely characterised by the 3 channels of the CPAM model. However, because the strength, $\mathbf{S}$, depends directly on the intensity of the illuminant [17], and can therefore be an unreliable variable in object recognition. Statistics (histograms) of local features, measured by local operators such as Gaussian derivatives or Gabor filters, can be used to characterise the global appearance of objects for recognition purposes [18]. By indexing the ASP and CSP patterns with the codebooks developed in the last section, we can construct image indices for content-based image indexing and retrieval. In the original colour histogram method, each pixel is quantized to one of a set of pre-determined colours. This can be viewed as a special case of indexing coloured pattern, whereby the size of the pattern is 1×1 pixel. For each image, we define an ASP histogram and a CSP histogram.[5] Let ASPH and CSPH be the achromatic spatial pattern histogram and chromatic spatial pattern histogram, respectively, they are defined as

$$\text{ASPH}(i) = \Pr(Q_P(ASP) = i),$$

$$\text{for } i = 1, 2, \ldots, I, \text{ where } I \text{ is size of } Q_p, \quad (6)$$

$$\text{CSPH}(j) = \Pr(Q_c(CSP) = j),$$

$$\text{for } j = 1, 2, \ldots, J, \text{ where } J \text{ is size of } Q_c. \quad (7)$$

Therefore, these histograms' dimensionality are the size (numbers of codewords) of the VQ codebooks, they capture the probability of ASPs and CSPs being quantized to a particular codeword. In collecting these statistics from an image, we can use either overlapping or non-overlapping patterns (blocks). Either way, the computational burden is light (they are mostly computed offline and needs to be done only once)

and the histograms' dimensionalities remain the same. We found that histograms collected using overlapping blocks and non-overlapping blocks have similar performance in content-based image retrieval. Since using non-overlapping blocks has a slight computational advantage, we have used non-overlapping block throughout this work for collecting the histograms.

To compare the similarity of two images, we can calculate the distance between their ASPHs and CSPHs. We have found the L_1 relative distance measure in Ref. [5] worked well. Let ASPH_m and CSPH_m be the achromatic spatial pattern histogram and chromatic spatial pattern histogram of image m, and ASPH_n and CSPH_n be the achromatic spatial pattern histogram and chromatic spatial pattern histogram of image n. The similarity of the two images, m and n can be measured as

$$d(m,n) = \lambda_a \sum_{\forall i} \frac{|\text{ASPH}_m(i) - \text{ASPH}_n(i)|}{1 + \text{ASPH}_m(i) + \text{ASPH}_n(i)}$$

$$+ \lambda_c \sum_{\forall j} \frac{|\text{CSPH}_m(j) - \text{CSPH}_n(j)|}{1 + \text{CSPH}_m(j) + \text{CSPH}_n(j)}, \quad (8)$$

where λ_a and λ_c are the weighting factors giving the chromatic and achromatic spatial patterns different weightings. A smaller value of $d(m,n)$ indicates image m and n are more similar.

In constructing image databases, we can compute the ASPHs and CSPHs for each image and store them as the image's metadata (image index). Content-based image retrieval can be done by comparing the query image's index with indices of the images in the database according to Eq. (8), and returning a set of images that are most similar to the query.

4. Experimental results

We have performed extensive experiments to evaluate the performance of indexing the chromatic and achromatic spatial patterns in content-based image indexing and retrieval. As a comparison, we have also implemented one of the state-of-the-art image indexing techniques, the colour correlogram method [5]. In all the results presented, the colour correlogram was implemented exactly as those described in Ref. [5], and the colour quantizer (64 colours) was trained using the same set of 235 true colour images used to train the ASP and CSP quantizers. Again, FSCL algorithm was used and each pixel was presented for training 20 times. The size of both ASP and CSP's codebooks are 256. Therefore, the colour (auto-) correlogram is 256-dimensional (64 colour $\times$ 4 distances $\{1, 3, 5, 7\}$ [5]), and the ASPH and CSPH are both 256-dimensional. Therefore, the storage and computational requirement of the current method

---

[5] A joint ASP and CSP histogram can also be constructed, but this will increase the histogram's dimensionality dramatically.

is 2 times that of colour correlogram's. However, with the rapid advancement in hardware technology, storing 512 numbers alongside each image is trivia, and computation is still simple enough. It is much simpler than the joint histogram of [14] and the blobworld method of Carson et al. [3]. Although our method requires significant amount of computation to design the codebooks, this can be done off-line and needs only to be done once. In terms of constructing the histograms, only table look-up is involved, so the computation required is almost trivia. And because we are using non-overlapping blocks, we need only to calculate the histograms for $(M \times N)/16$ blocks for an $M \times N$ image. An added advantage of our current strategy is that we can utilise the ASP and CSP quantizers to compress images, thus enable indexing and retrieval of image in the compressed domain without having to pre-compute the indexing meta-data. Preliminary results along this line will be presented in Ref. [12]. We will see in the following, with a slightly increase in complexity, our method performs better than the colour correlogram method in one set of image database and at least as well in another set of image database.

4.1. Colour texture database

In the first set of experiments, we used a texture database consisting of 120 difference texture classes, a subset of which is show in Fig. 3 as thumbnails. Each of the 300×200 pixels true colour (24 bits) images [6] was divided into six 100×100 pixels non-overlapping sub-images, thus creating a database of 720 texture images. In the experiment, each of the 720 images in the database was used in turn as the query image. For each query image q, the distances between q and images in the database, $p_i, d(q, p_i)$ where $i = 1, 2, \ldots, 720$, $p_i \neq q$ were calculated and sorted in increasing order and the closest set of images are then retrieved. In the ideal case, the top 5 retrievals should be from the same large image as the query image. The performance is measured in terms of the retrieval rate (RR) defined as follows: Let q be the query image from texture class T_k, i.e., $q \in T_k$ and p_j, $j = 1, 2, 3, 4, 5$, be the first 5 retrievals, the retrieval rate for query image q is

$$\text{RR}(q) = \frac{1}{5} \sum_{j=1}^{5} S(p_j) \quad \text{where} \quad \begin{cases} S(p_j) = 1, \text{ if } p_j \in T_k, \\ S(p_j) = 0, \text{ if } p_j \notin T_k. \end{cases}$$
(9)

A larger retrieval rate indicates better performance, and in the ideal case *RR* should equal to 100%. Table 1

[6] There are 120 images in this collection and each image was judged by human observers as having a single class of texture. This set of textures differs significantly from the MIT VisTex database both in terms of colours and surface textures.
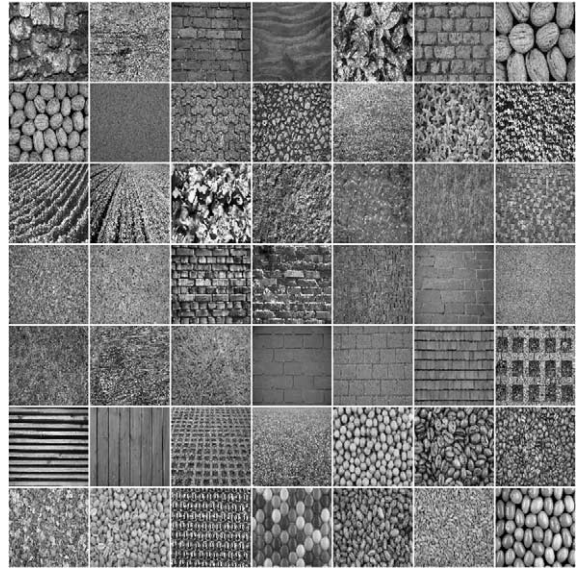

Fig. 3. Thumbnails images of a subset of the colour textures in the colour texture database.

Table 1
The number of queries and their retrieval rates for different content-based image indexing and retrieval methods performed on a texture database consisting of 720 images and 120 different texture classes. Each image was used in turn as the query image; therefore 720 queries were performed. This table should be interpreted as follows: for the colour correlogram method, 482 queries achieved a retrieval rate of 100% and 130 queries achieved a retrieval rate of 80% etc.

Methods	Retrieval rates				
	5/5	4/5	3/5	2/5	1/5
Colour correlogram	482	130	68	33	7
New ($\lambda_a = 1$, $\lambda_c = 0$)	394	166	78	54	28
New ($\lambda_a = 0$, $\lambda_c = 1$)	538	106	52	16	6
New ($\lambda_a = 1$, $\lambda_c = 1$)	626	58	24	8	4

shows the number of queries and their corresponding retrieval rates for different methods. The entries in the table are the number of queries. For the colour correlogram method of Huang et al. [5], 482 queries (out of 720) achieved a retrieval rate of 100%, 130 queries achieved a retrieval rate of 80% and so on. Using the achromatic spatial pattern indexing alone, 394 queries achieved a retrieval rate of 100%. Using the chromatic spatial pattern indexing alone, 538 queries achieved a retrieval rate of 100% and 106 queries achieved a retrieval rate of 80%. It is seen that chromatic spatial pattern indexing alone has outperformed colour correlogram in this context. Combining the chromatic and achromatic spatial pattern indexing, there are 626 queries which achieved a retrieval rate of 100%, which is significantly better than the colour
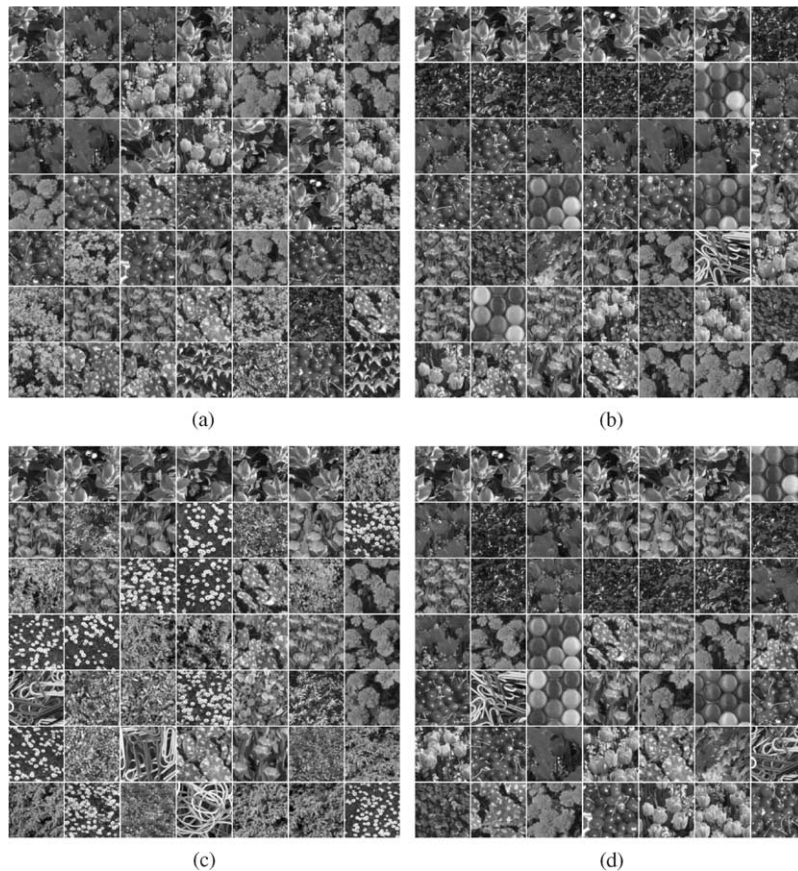
Fig. 4. Texture image database retrieval results of different methods. The image on the top-left hand corner is the query image, subsequent ones, from left to right, top to bottom, are retrieved images ordered according to their similarities to the query image. (a) Result of colour correlogram method. (b) Result of the new method with $\lambda_a = 0$, $\lambda_c = 1$. (c) Result of the new method with $\lambda_a = 1$, $\lambda_c = 0$. (d) Result of the new method with $\lambda_a = 1$, $\lambda_c = 1$.

correlogram technique. We have also experimented using a 4096-bin colour histogram indexing [1] as well, the performance was very poor, and less than half the queries achieved a retrieval rate of 100%. An example of the retrieval results for colour correlogram of Huang et al. [5] and the new method are shown in Fig. 4, where the top 48 returns of a particular query are displayed.

As we briefly mentioned in Section 2, ASP and CSP are scaled by the average intensity of the pattern. This has the effect of removing (to a certain extent) the influence of the lighting conditions on the performance of the indexing methods. Fig. 5 is one of the textures in the database. As can be seen, the lighting is uneven across the image. On the right-hand side of the image, the colour appears bright golden because light was reflected by the coins, whilst in the middle of the image, the coins appears darker and have different colour appearance from those on the right. In such a situation, recognition methods based on colour but without some

sort of colour constancy measures will very likely fail. Fig. 6 shows the results of a query using one of the sub-images shown in Fig. 5 by the colour correlogram and the new method ($\lambda_a = 1$, $\lambda_c = 1$). It is seen that while the new method retrieved all other five sub-images from the same image in the top five ranks, the colour correlogram has failed to do so. Observing the results of all other five queries by other five sub-images of the same image, the new method achieved a retrieval rate of 100% for all queries, whilst the colour correlogram has failed. This may suggest that the new method is somewhat "colour constant", i.e., the histograms of ASP and CSP may be invariant to illumination conditions. Of course, colour constancy is an important and challenging problem in colour imaging, more work is needed to make any general claim about the colour constancy properties of the new technique. Our future work will include evaluating the technique in this aspect and develop it to be invariant to illumination conditions.

Fig. 5. One of the texture images in the database. The lighting is unevenly distributed across the image.

(a)   (b)
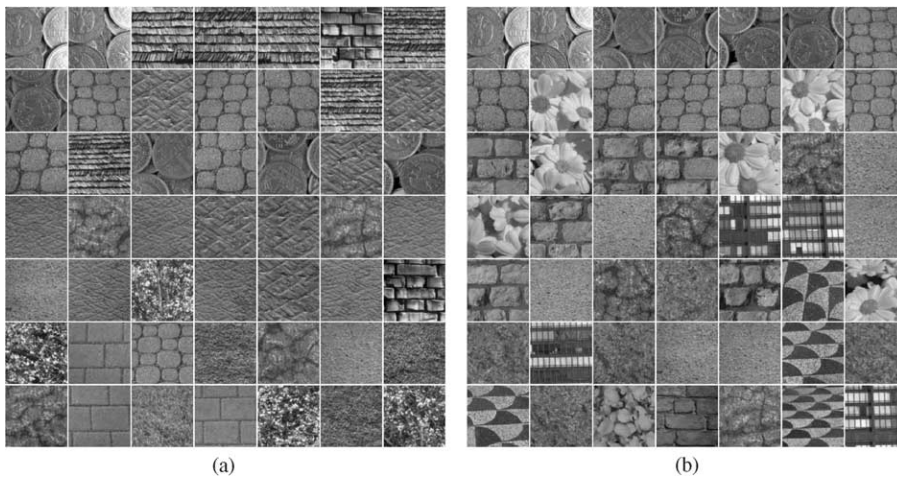
Fig. 6. Retrieval results of a query using one of the sub-images from the image showed in Fig. 5. (a) Result of colour correlogram method. (b) Result of the new method with $\lambda_a = 1$, $\lambda_c = 1$. Since the lighting is unevenly distributed across the image, the sub-images are in effect under different illuminant, therefore even though the whole image contains the same coloured objects (same type of coin), the colours of the coins are different at different locations because colours are illuminant dependant. Since the coins have specularly reflective surface, it makes the colour appearance of the scene even more complicated. It is seen the retrieval rate for colour correlogram is low whilst the retrieval rate of the new method is 100%.

4.2. Photography image database

To further evaluate the performance of our new method, we have tested it on two photography image databases. Both are a sub-set of commercially available Corel Photo collection. The first database consists of 4000 images and the second database consists of 20,000 images. We have hand picked 96 pairs of similar images from various sources and divided them into set A and set B. For each image in set A, there is a corresponding similar target image in set B, or vice versa. Fig. 7 shows a sub-set of these pairs of images. As can be seen, the corresponding images are similar, but differ in various ways. We used one set as query images, and embedded the other set into the database. The goal is to retrieve the corresponding target image from the database. As in the previous experiment, we have also implemented the colour correlogram method as a comparison. For each of the image in the database and in the query set, a colour correlogram, an achromatic and a chromatic

Fig. 7. A subset of 96 pairs of query images. These pairs are divided into set A and Set B, for each image in set A, there is a corresponding similar target image in set B, or vice versa.

Table 2
Two different methods' image retrieval performance on a photography image database consists of 4000 images. 96 queries were performed. The table shows the number of queries and their target images rank range. This table should be interpreted as follows: for the colour correlogram method, 76 queries found their corresponding targets in the first rank, 84 queries found their corresponding images within the first 10 returned images, etc.

Methods	Ranks				
	1	$\leqslant 10$	$\leqslant 30$	$\leqslant 100$	> 100
Colour correlogram	76	84	89	89	7
New ($\lambda_a = 1$, $\lambda_c = 1$)	72	86	89	90	6

Table 3
Two different methods' image retrieval performance on a photography image database consists of 20,000 images. Again 96 queries were performed. The table shows the number of queries and their target images rank range. This table should be interpreted as follows: for the colour correlogram method, 72 queries found their corresponding targets in the first rank, 81 queries found their corresponding images within the first 10 returned images, etc.

Methods	Ranks				
	1	$\leqslant 10$	$\leqslant 30$	$\leqslant 100$	> 100
Colour correlogram	72	81	85	87	9
New ($\lambda_a = 1$, $\lambda_c = 1$)	65	79	84	88	8

spatial pattern histograms were constructed. The distances between the query image and all images in the databases were calculated according to Eq. (8). For each query, the images in the database were ranked according to their distances to the query image, and a small distance has a higher rank. In the ideal case, the corresponding target image should be ranked in the No 1 position.

Table 2 shows the results of using set A as query and set B as target in a 4000 (4096 if the embedded target images were counted) image database. The entries in the table are the number of queries and the rank position range of their corresponding target images. It is seen for the colour correlogram method, 76 queries found their corresponding targets in the first rank (most similar), whilst the new method had 72 queries found their targets in the first rank. Both methods had 89 queries which found their corresponding targets within the first 30 retrieved images. On the other hand, the colour correlogram had the lowest rank in 1612 position, i.e., for this particular

query, you would have to look through the first 1611 returned images to find the target image, whilst the new method had the lowest rank of 282. However, in practice, about 10 images can be displayed on the screen at any one time and the user of the query will normally not go through more than 3 screens. We therefore, consider a method which does not find the target image within the first 30 returned images as failure. From Table 2, we can see that the success rate of colour correlogram was 89/96, i.e., 92.7%, whilst the success rate of the new method was also 89/96, i.e., 92.7%. The average rank for all the successful queries (i.e., those returned their targets with a rank $\leqslant 30$) for the new method was 2.0 and for the colour correlogram was 2.6. Overall, we see the new method is comparable to colour correlogram.

To investigate whether the method will scale to larger image databases, we repeated the experiments on a larger image database consists of 20,000 images. Table 3 shows the results. As can be seen, these results scale

Query	Target	Ranks
		Colour Correlogram: 20 New ($\lambda_a = 1, \lambda_c = 1$): 1
		Colour Correlogram: 1138 New ($\lambda_a = 1, \lambda_c = 1$): 1328
		Colour Correlogram: 9 New ($\lambda_a = 1, \lambda_c = 1$): 2
		Colour Correlogram: 285 New ($\lambda_a = 1, \lambda_c = 1$): 10
		Colour Correlogram: 950 New ($\lambda_a = 1, \lambda_c = 1$): 165
		Colour Correlogram: 1 New ($\lambda_a = 1, \lambda_c = 1$): 61
		Colour Correlogram: 13 New ($\lambda_a = 1, \lambda_c = 1$): 28
		Colour Correlogram: 20 New ($\lambda_a = 1, \lambda_c = 1$): 1
		Colour Correlogram: 3 New ($\lambda_a = 1, \lambda_c = 1$): 1
		Colour Correlogram: 12 New ($\lambda_a = 1, \lambda_c = 1$): 4
		Colour Correlogram: 288 New ($\lambda_a = 1, \lambda_c = 1$): 314
		Colour Correlogram: 1870 New ($\lambda_a = 1, \lambda_c = 1$): 86

Fig. 8. Some example query images and their target images' position when retrieved from a database of 20,000 images.

quite well. Again, on average, the new method performed better than the colour correlogram method. In this case, the colour correlogram had the lowest rank in 7299 position, i.e., for this particular query, you would have to look through the first 7298 returned images to find the target image, whilst the new method had the lowest rank of 5684. Again, consider a query returned its target at a rank higher than 30 as failure, the success rate of the new method was 84/96, i.e., 87.5%, and the success rate of the colour correlogram method was 85/96, i.e., 88.5%. The average rank for all the successful queries (i.e. those returned their targets with a rank $\leqslant 30$) for the new method was 1.9 and for the colour correlogram was 2.5. Again we see the new method is comparable to colour correlogram. Fig. 8 shows some examples of the query images and their corresponding target images' positions retrieved by both methods from this large database.

Visual assessment of the retrieved results showed that the new methods returned more meaningful images. Figs. 9 and 10 show some examples. In Fig. 10 for example, even though the colour correlogram method retrieved the target in the first rank position, browsing these first 48 retrieved images indicates that the new method returned visually more relevant images.

5. Conclusions

In this paper, we have developed a method to represent and indexing the achromatic and chromatic signals independently. We justify the scheme from both human colour vision theories' viewpoint and modern digital signal processing viewpoint. Through extensive experiments, we demonstrated that the scheme could be

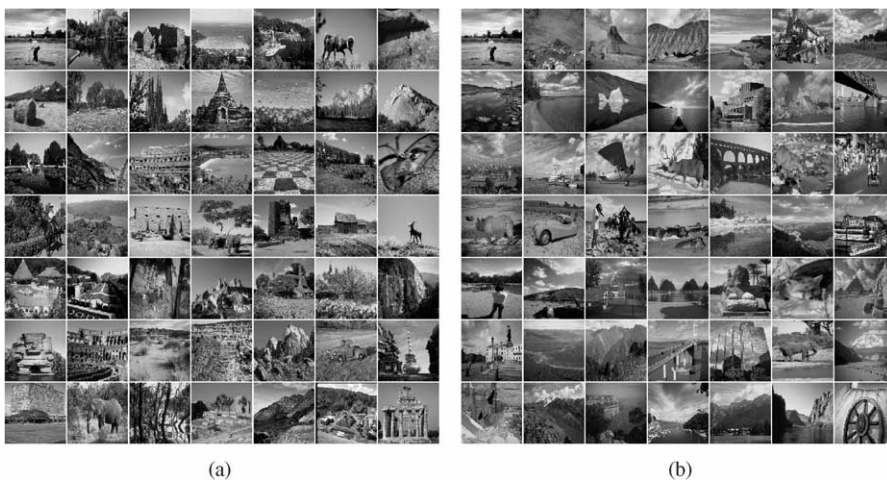(a)   (b)

Fig. 9. Retrieval results from the 20,000 image database. The image on the top-left hand corner is the query image, subsequent ones, from left to right, top to bottom, are retrieved image ordered according to their similarities to the query image. (a) Result of colour correlogram method. (b) Result of the new method with $\lambda_a = 1$, $\lambda_c = 1$. It is seen that the new method returned images which have more visual relevance.
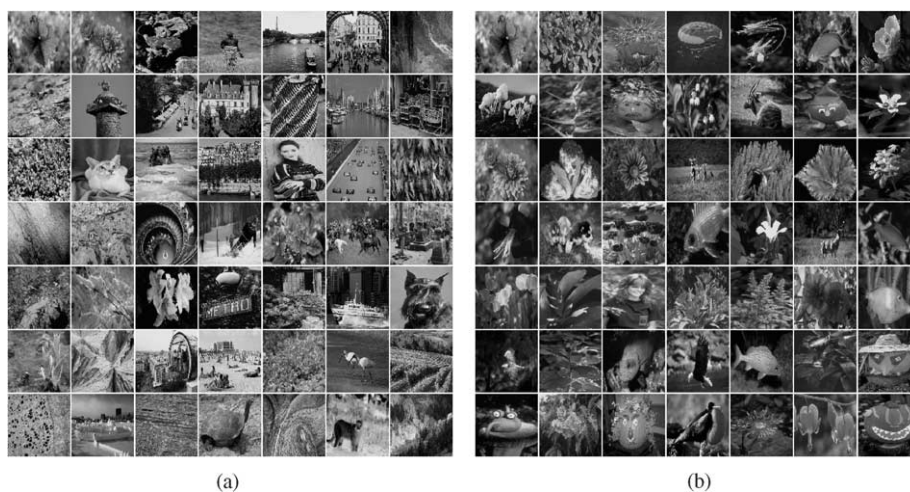
Fig. 10. Retrieval results from the 20,000 image database. The image on the top-left hand corner is the query image, subsequent ones, from left to right, top to bottom, are retrieved image ordered according to their similarity to the query image. (a) Result of colour correlogram method. (b) Result of the new method with $\lambda_a = 1$, $\lambda_c = 1$. In this example, even though the colour correlogram returned the target image in the first rank, the rest of the images returned by the new method are visually more relevant.

used for indexing images for effective content-based image retrieval from image databases. Through experiment, we also show the new method demonstrated some illumination invariant properties. We further demonstrated that the new method is better than or at least as good as state-of-the-art image indexing and retrieval techniques. The method can be extended to achieve image compression and perform compressed domain image indexing and retrieval [19]. Our future work will include studying the scheme's image coding performance and developing it into an illumination invariant indexing scheme.

References

[1] M. Swain, D. Ballard, Color Indexing, Int. J. Comput. Vision 7 (1991) 11–32.

[2] B.S. Manjunath, W.Y. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (1996) 837–842.

[3] C. Carson et al., Blobworld: a system for region-based image indexing and retrieval, Proceedings of the International Conferences on Visual Information Systems, 1999.

[4] G. Pass, R. Zabih, Comparing images using joint histograms, Multimedia Systems, Vol. 7, Springer, 1999, pp. 234–240.

[5] J. Huang et al., Image indexing using color correlogram, IEEE CVPR (1997) 762–768.

[6] G. Qiu, S. Sudirman, Representation and retrieval of color images using binary space partitioning tree, Eighth Color Imaging Conference, Scottsdale, Arizona, USA, November 2000, pp. 195–201.

[7] P.K. Kaiser, R.M. Boynton, Human Color Vision, Optical Society of America, Washington DC, 1996.

[8] A. Poirson, B. Wandell, Appearance of colored patterns: pattern-color separability, J. Opt. Soc. Am. A 10 (1993) 2458–2470.

[9] W. Pratt, Digital Image Processing, Wiley, New York, 1978.

[10] CCIR, Encoding parameters of digital television for studios, CCIR Recommendation 601-2, Int. adio Consult. Committee, Geneva, 1990.

[11] D. Marr, Vision: A Computational Investigation into Human Representation and Processing of Visual Information, Freeman and Co., New York, San Francisco, 1982.

[12] G. Qiu, Image coding using a coloured pattern appearance model, Visual Communication and Image Processing 2001, January 2001, San Jose, CA, USA.

[13] B.V. Funt, G.D. Finlayson, Color constant color indexing, IEEE Trans. Pattern Anal. Mach. Intell. 17 (1995) 522–529.

[14] A. Gersho, R.M. Gray, Vector quantization and signal compression, Kluwer Academic Publishers, Boston, 1992.

[15] T. Kohonen, Self-organization and associative memory, Springer, Berlin, 1989.

[16] S.C. Ahalt et al., Competitive learning algorithms for vector quantization, Neural Networks 3 (1990) 277–290.

[17] G.D. Finlayson, M.S. Drew, B.V. Funt, Color constancy: generalized diagonal transforms suffice, J. Opt. Soc. Am. A 11 (1994) 3011–3020.

[18] B. Schiele, J.L. Crowley, Recognition without correspondence using multiresolution receptive field histogram, Int. J. Comput. Vision 36 (1) (2000) 31–50.

[19] R.W. Picard, Content access for image/video coding: the fourth criterion, MIT Media Lab Technical Report No. 295, 1994.

About the Author—G. QIU is a Lecturer of Computer Science at the School of Computer Science and IT, University of Nottingham, UK. Before joining Nottingham in October 2000, he was a Lecturer at the School of Computing, University of Leeds, UK; and the School of Computing and Mathematics, University of Derby, UK. He received his Ph.D. in neural networks and image processing in 1993 from the University of Central Lancashire, Preston, UK, and his B.Sc. in Electronic Measurement and Instrumentation in 1984 from the University of Electronic Science and Technology of China. From 1987 to 1990 academic years, he was also a postgraduate student at the Department of Electronic Engineering, Beijing Institute of Technology, studied and researched Digital Signal Processing. In the summers of 1999 and 2000, he worked at British Telecommunications Laboratories as a Short-term Research Fellow. His research covers many aspects of visual information processing including colour imaging, image databases, colour image coding for indexing and retrieval and web-based informatics. He has published some 50 papers in these areas including recent papers in IEEE Transactions on IP and CASVT. His research is funded by European Union and Industry. He has been a consultant to UK and multinational companies.