CONTENT-BASED IMAGE RETRIEVAL USING COLOUR N-GRAMS

Rick Rickman* & Paul Rosin+
*Dept. Electronic Engineering
+Dept.Computer Science
Brunel University
Uxbridge, Middx, UK

Colour histograms have been shown to be very effective in representing images for content-based retrieval systems. They are compact, robust and amenable to computer analysis. However, such histograms convey only global image properties and do not embody local colour information which is so important when comparing and contrasting images. We present a novel image coding scheme which captures some of this locally correlated colour information and improves the selectivity of the retrieval mechanism - an important issue for very large databases. The technique uses a histogram of features which represent local combinations of colour artifacts within an image, called colour n-grams. We outline the thrust of our approach and discuss the factors affecting the efficacy of the retrieval mechanism using a database of colour images.

Introduction

Content-based image retrieval systems utilise salient artifacts from within a target image as the basis of a query and matching mechanism to a large database of images. Database images which share a significant proportion of similar artifacts as the querying image. These feature must be capable of being extracted automatically as there is a general consensus that systems which require a high degree of user input are unlikely to provide an attractive or viable solution for most potential users. The index must be assigned *a priori* as each image is instantiated into the database to reduce the processing times during the matching.

Much of the work in content-based image retrieval suggests that, at the present time, the most effective features for practical near-market systems are comparatively simple ones such as texture[1][6][7] and/or colour[1][2][3][4][5]. More complex features, such as shape descriptors [10][11] and so-called 'eigen features'[6][12], have been employed but these have not proven to be quite so robust for practical, domain independent solutions. Colour features have been used extensively primarily because they can be readily extracted and represented.

Colour N-Grams

Colour is a powerful descriptor that often simplifies image understanding tasks. In most existing systems which use colour artifacts as the basis of the matching mechanism, the indices for each image are derived from the colour histogram representing the relative frequency of colour pixels across the image. Images are compared simply by matching their respective histograms. However, such histograms do not preserve local spatial colour information within an image and the matching mechanism is consequently impaired.

We propose a novel image coding technique which preserves some of the spatial colour correlates within an image to provide a more selective matching mechanism than can be achieved by global colour histograms. Here images are encoded with respect to a code book of features which describes every possible combination of a fixed number of coarsely quantised colour hues that might be encountered within local regions in an image. This enables images to be compared on the basis of their shared adjacent colour artifacts or boundaries. A histogram of the relative frequency of these features is formed for an image by incrementing the appropriate histogram 'bin' each time a particular combination of colours is encountered. This histogram 'bin' corresponds to the code book index for that particular combination of colours within a small local region.

This approach is analogous to a technique employed in text retrieval systems which uses character sub-strings as the basis of the indexing and matching mechanism [9]. These so-called n-grams are fixed length character sub-strings and in a typical n-gram retrieval system each combination of characters found in the database is assigned its own code. Documents are matched on the basis of the frequency of shared n-grams. The indexing advantage of this technique is afforded by the fact that the indexing domain for the n-grams sub-strings is significantly less than that for whole words. Whereas this technique uses small 1-D character strings, our approach employs a 2-D representation of local colour artifacts.

The Thrust Of The Colour N-Gram Coding Scheme

It is instructive to demonstrate the thrust of our approach with a simple example. The image in fig. 1 is comprised of a red car on a green and blue background and is to be encoded using colour n-grams. The histogram, which forms the basis of the image matching index, is derived from 3 n-grams F_0- F_2.

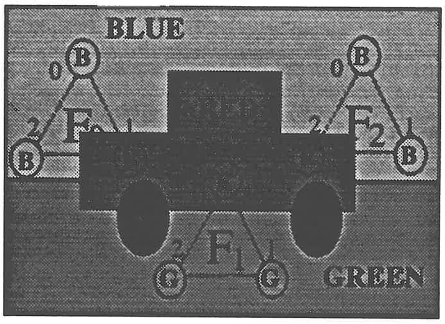

Fig. 1. Coding an image using colour n-grams.

The code book reference of each of these 3-gram features is determined by the combination of quantised colour hues at each of the 3 sampling sites located at the vertices 0,1 and 2 of the profiles of F_0- F_2. The hue is calculated by transforming the RGB values of the pixels at the 3 sampling sites into HSI (hue, saturation and intensity, respectively) co-ordinates. In the present system neither the saturation (S) or intensity (I) information is used. The hue is defined as an angle between 0° and 360° where red, green and blue are represented by hue angles of 0°, 180° and 270° respectively. This hue angle is quantised into q levels, say. In this example let q=8 so that the quantised hues for red, green and blue are 0, 3 and 5 respectively.

The code book index for the n-gram of each colour combination is given by:

$$I(F_x) = \sum_{i=0}^{n} H_{xi}\, q^{i}$$

Where:

F_x	is the x-th feature.
H_{xi}	is the quantised hue value at the i-th site of the x-th feature.
n	is the n-gram size (or number of sampling points per feature).
q	is the number of quantisation levels of the hue angle.

Then the index for each of the 3-gram features in fig.1 is given in Table 1.

Feature N°.	Hue at sampling points			n-gram Index
	0	1	2	
F_0	0 (BLUE)	0 (RED)	0 (BLUE)	$5x8^0+0x8^1+5x8^2 = 360$
F_1	3 (RED)	3 (GREEN)	0 (GREEN)	$0x8^0+3.8^1+3x8^2 = 216$
F_2	3 (BLUE)	3 (BLUE)	3 (RED)	$5x8^0+5.8^1+0.8^2 = 45$

Table 1. Calculating n-gram indices to construct the feature histogram for the image in Fig. 1.

The length of the colour n-gram code-book, C, is given by:

$$C = q^n$$

A scene typically contains many isomers of the same set of colours within a feature. For example, F_0 and F_2 contain different combinations of 2 blues and a red. The extent of the code book can be reduced significantly if each isomer is resolved into a single standard canonical form representing all combinations of a given set of colours. In this example, F_0 and F_2 are represented by a single canonical form blue, blue, red. The compacted canonical index greatly improves both the speed of match and memory requirements of the system. This is shown in table 2 below.

Quantisation Levels	n-gram size	Non-canonical code book size	Canonical code book size
8	3	512	120
8	4	4096	330
16	3	4096	816
16	4	65536	3876

Table 2. Canonical colour n-grams reduce the index length

Weighting Important N-Gram Features

The index for each image is derived from the frequency histogram of colour n-grams in the image. The image is sampled using a large number of randomly placed colour n-gram feature profiles. Images are matched by comparing the histograms of the query and database images. N-gram features which are very common throughout the database do not provide so much discriminatory information as features which occur less frequently. During the matching phase the comparison of respective n-gram 'bins' within the histogram is weighted by the probability of their occurrence across the database.

Image information is conveyed through local spatial variations of image attributes. These local variances tend to be rather smooth and, typically, those n-grams which contain a large number of different colours will provide more selective matching information than those of a single colour. Very localised n-gram feature profiles result in features which have the same colour at each sampling point and much of the spatially correlated colour information which provides improved selectivity in the matching phase is swamped by low-order features which contain recurrent colours. Indeed, as the extent of the features become more localised, the performance of the matching metric approaches that of a coarsely quantised straight colour histogram matching scheme.

Robust Features From Noisy Images

Problems associated with colour constancy are one of the most serious setbacks encountered when matching images using colour components. This is because the same scene captured under different lighting conditions across range of platforms may appear to be perceptually similar to a human observer but can have remarkably different spectral components. For most colour histogram matching techniques this shortcoming has not been successfully addressed. The finer the quantisation of the colour space, the more pronounced this effect becomes. Because the quantisation grid of the colour n-grams is very coarse it is less sensitive to small spectral differences and is not so prone to colour constancy problems. For straight histogram matching systems the decrease in selectivity brought about by a coarse quantisation grid would seriously impair the performance of the system. Whilst the quantisation of our scheme is comparatively coarse, the selectivity of the match is increased by the higher dimensionality afforded by the n-gram. Califano and Mohan [8] show that high dimensional coarsely quantised features produce improved image matching capability in terms of both accuracy and false positive retrieval suppression.

The effect of local images inconsistencies caused by noise or 'dithering', for example, can be reduced by taking the mean of the colour around a small local region centered at n-gram sampling point.

Dealing With Black and White Image Artifacts

The hue angle cannot be used to encode grey tones ranging from black to white. To alleviate this problem, any colour whose chrominance falls below a threshold is encoded using its quantised luminance value instead - the quantisation code being offset by the number of quantisation levels assigned to the hue.

Results

The test database contains 100 full colour (.gif) images of very general subject matter (facial images, flowers, animals, cars airplanes, vistas, abstract art and so on). The average resolution of each image was 284 x 235 pixels. The feature histogram was based on 3-gram features profiles with a side length of 20 pixels and 13 hue and 3 grey quantisation levels. The hue was taken to be the mean hue within a 3x3 pixel region centered on each sampling region. The number of n-gram samples was taken to be 25% of the number of pixels in the image - each sample was placed randomly within the image. Each n-gram isomer was resolved to its canonical form and the frequency of occurrence of each recorded and stored as the index. This was repeated for every image in the database and the probability of occurrence of each n-gram calculated for the whole database. This serves as the basis of the weighting function during the matching phase.

The querying image was encoded in the same fashion and the resulting histogram compared against those for the database images. The match score was derived from the number of filled histogram slots common to both the query and database image. The 10 highest scoring images were retrieved and displayed for final selection by the user.

We used the approach adopted by Faloutsos *et al* [10] to benchmark the efficacy of our retrieval system. Here the experimenter selects 5 images from within the database deemed to be the most similar to the target with respect to colour and subjective content. The number and rank of retrieved relevant images was used to assess the performance of the system. The following metrics are employed:

- AVRR is the average rank of all relevant images (the first position is the 0-th rank). Note that we include the rank of relevant images even if they are not in the 10 retrieved images.
- MT is the number of relevant images that were missed in the 10 retrieved images.

By way of comparison we have also applied the same retrieval to standard colour histogram matching using just the hue components. The histogram was resolved to 256 cells. These tests were repeated for 10 query sessions and the average results shown in Table 3.

Measure	Colour n-gram histograms	Colour histograms
AVRR	2.4	2.5
MT	1.9	2.1

Table 2 Comparing the efficacy of n-gram colour retrieval with straight colour histogram retrieval

Discussion

Colour information is useful for coding images for content-based retrieval systems because it often simplifies the classification and detection of features used to compare images. Colour histograms have been employed very effectively in current near-market content-based image retrieval systems but this approach tends to throw away much of the locally correlated colour information which is important in classifying images. This information could improve the selectivity of the retrieval mechanism and is especially relevant for very large databases. We have introduced a novel technique which captures some of this information and embodies it within a very compact index to facilitate rapid fuzzy matching of colour images in content-based retrieval systems.

The work is still at a formative stage and many issues still need to be addressed. At present the quantisation of the hues does not match the sensitivity of the human colour perception model. For example, humans are several orders of magnitude more sensitive to red and green that they are to yellow and blue and because we are attempting to match the users perception of visual similarity it would be useful to include this model in successive developments.

These results were obtained using a comparatively small image database. The authors feel that this technique would come to the fore for very large databases where enhanced selectivity is required. It is hoped that these issues will be addressed in future work.

Acknowledgments.

This research was funded by the British Library Research and Development Department research contract MO921

Bibliography

1. W.Niblack, R. Barber W. Equitz, M. Flickner, E. Glasman D. Petkovic, P. Yanker and C. *Faloutsos 'The Q-BIC Project: Querying Images Using Colour, Texture and Shape'* IBM Research Report RJ 9203 (81511) Feb. 1 1993

2. Gargi, S. Oswald, D. Kosiba, S. Devadiga and R. Kasturi.*'Evaluation of Video Sequence indexing and hierarchical video Indexing'* Proc. Conf. IS&T/SPIE -Storage and Retrieval for Image and Video Databases III. pp. 144-151. 9-10 Feb 1995 San Jose CA, USA.

3. Equitz *'Image Searching in a Shipping Product'* Proc. Conf. IS&T/SPIE -Storage and Retrieval for Image and Video Databases III. pp. 186-196. 9-10 Feb 1995 San Jose CA, USA.

4. Ogle and M. Stonebraker *'Chabot: Retrieval from a Relational Database of Images'* IEEE Computer Magazine September 1995 pp. 40-48.

5. Stricker and M. Orengo *'Similarity Of Color Images'* Proc. Conf. IS&T/SPIE -Storage and Retrieval for Image and Video Databases III. pp. 144-151. 9-10 Feb 1995 San Jose CA, USA.

6. Pentland, R. W. Picard and S. Sclaroff *' Photobook: Tools for content-based manipulation of Image Databases'*. IS&T/SPIE -Storage and Retrieval for Image and Video Databases II. pp. 34-47. 7-8 Feb 1994 San Jose CA, USA.

7. Jain, S. N. Jayaram Murthy and P Chen *'Similarity Measures for Image Databases'* Proc. Conf. IS&T/SPIE -Storage and Retrieval for Image and Video Databases III. pp. 58-65. 9-10 Feb 1995 San Jose CA, USA.

8. Califano and R. Mohan 'Multidimensional Indexing for Recognising Visual Shapes' IEEE Trans. PAMI Vol. 16 No.4 April 1994 pp 373 - 392.

9. Suen, Ching Y., "n-Gram Statistics for Natureal Language Understanding and Text processing," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. PAMI-, No.2 pp. 164-172, April, 1979.

10. Faloutsos, M. Flickner, W. Niblack, D. Petcovic, W. Equitz and R. Barber. 'Efficient and Effective Querying by Image Content' IBM Research Report RJ 4953 (83074) August 3, 1993.