Image Coding for Classification

G. Qiu
School of Computer Studies, The University of Leeds, United Kingdom

G. D. Finlayson
School of Information Systems, The University of East Anglia, United Kingdom

Abstract

Query by image contents is a successful approach to image database indexing and retrieval.[1,2] Because image data take up huge amount of storage space and network bandwidth, compressing image is a standard practice. Given that images are compressed/coded in a database, it would be more efficient if the compressed data could also be directly used as indices for content-based query. The compressed data of traditional image coding/compression techniques such as the popular JPEG[3] cannot be readily used to classify image contents. In this work, we set out to develop image coding schemes that will enable classification of image contents to be performed in the compressed domain without the need to perform image decoding and any extra processing. One of the novel features of this work is that we exploit some of the well known properties of human colour visual system for image representation. We outline a general scheme of colour image coding for classification and describe a specific implementation.

Introduction

Image compression aims at reducing the amount of data needed to represent a digital image whilst trying to maintain high perceptual image quality. In order to achieve reasonable compression ratios, the psychovisual redundancies, which exist in natural images, have to be exploited; i.e. visually meaningful information is preserved while visually insignificant data is discarded. If images are stored in a compressed format that conveys explicitly image contents, such as colour distribution, edge and texture patterns, image classification can be done on the compressed image data, this will facilitate image data management on the Web and ordinary image databases. We introduced such a scheme in this paper.

There is evidence to suggest that different visual pathways process colour and pattern in the human visual system. In [4], experiments were carried out using square wave patterns with a range of different spatial frequencies, colours and stimulus strengths to measure how colour appearance depends on spatial pattern. Results suggest that the value of one neural image is the product of three terms. One term defines the pathway's sensitivity to the square wave's colour direction. A second term defines the pathway's sensitivity to the spatial pattern and the third term defines the pathway's sensitivity to the square wave's stimulus strength.[4] There is also physiological evidence to suggest the existence of opponent colour signals in the visual pathway.[5] The opponent colour theory suggests that there are three visual pathways in the human colour vision system. One 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 blue-yellow pathway has the worst spatial resolution. In opponent-colours representations, 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.

Colour image coding systems, e.g. JPEG, MPEG and others[3] developed over the years have taken advantage of the low spatial acuity of the opponent colours representations where more bandwidth is allocated to the light-dark (luminance) channel and less bandwidth is given to the chromatic channels.

Query by image contents is a successful approach to image database indexing and retrieval. The contents used by different systems include colours, shapes and textures.[1,2] Because image data take up huge amount of storage space and network bandwidth, it is necessary to implement some sorts of compression/coding. Images in a database are normally in some sorts of compressed formats. Recent work on image classification on the World Wide Web[6] indicates that the vast majority of the images on the Web are Jpeg and Gif formats and from these image formats it is impossible to deduce the image contents such as colour resolution without decompressing the images. What is more, even though most of the images on the Web have far less than 24-bit colours, researchers have to assume 24-bit colour resolution for all the images. In this work, we set out to develop image coding schemes that will enable classification of image contents to be performed in the compressed domain without the need to perform image decoding.

Colour Image Coding for Classification

The schematic diagram of the image coding technique is shown in Fig. 1. An image is divided into small image blocks and coded one block at a time (at this stage, we have not considered multi-resolution hierarchical structure which could be easily extended to and maybe necessary in some applications). The method first estimates the luminance of the image block, and then the image pixels are normalized with respect to this block luminance value. The purpose of doing so is to eliminate the influence of the illuminant (as will become clearer later). After normalization, opponent colour mechanism is then invoked. The colour components are then sub-sampled (this is consistent with opponent colour theory of the human visual system[5]). The full (spatial) resolution black and white component, the sub-sampled colour components and the luminance of the block are then coded by their respective encoders (the system is somewhat reminiscent of the pattern colour separable model of the human visual system[4]). To make the compressed data directly usable for classification the encoders have to be designed in an appropriate manner.

Vector Quantization

Vector quantization is a well-known technique for image coding (see e.g. [10]). Typically, an image is divided into non-overlapping small blocks, and the pixels within each block form a vector. A schematic of a vector quantization system is shown in Fig. 2, where the code-book size is N and the code-words are C_1, C_2, …, C_N. The input vector X_n is quantized by being encoded into a scalar (integer) index I(n), which then serves as the input to the decoder, and the reproduction of the input is approximated by $C_{I(n)}$.

The encoding is based on minimum distance classification and therefore the input X_n is encoded to I(n) if

$$\left\| X_n - C_{I(n)} \right\| \le \left\| X_n - C_j \right\| \qquad \text{for } \forall \text{j}$$

Many approaches have been developed for designing the code-book. Clustering based algorithms, such as the generalized Lloyd algorithm (GLA) which is also sometimes referred to as the LBG algorithm[7] is a popular method for designing the code-book. The algorithm starts by choosing an initial set of code-words, then it iteratively updates the code-word using a minimum distance mapping.

Another technique, which has become popular in recent years is the neural network approach.[11] The frequency sensitive learning algorithm, which we used in this paper for code-book design, is described as follows:

Step 1, initialise the code-words C_1, C_2, C_N, and set the counters associated with each code-word, $count_n$, to 1.

Step 2, present the training sample, X_n, and calculate the distance between the training sample and the code-words as:

$$D_j = \left\| X_n - C_j \right\| \text{ and } D_j^* = count_j * D_j \text{, for } \forall j$$

Step 3, update the code-words according to

$$y_j = \begin{cases} 1 & \text{if } D_j^* \le D_k^* \quad \text{for } \forall k \\ 0 & \text{otherwise} \end{cases} \qquad \text{and}$$

$$C_j^{new} = C_j^{old} + \eta y_j \left(X_n - C_j^{old} \right) \qquad \forall j$$

where $0 < \eta < 1$ is a learning rate

Step 4, fetch the next training sample and repeat by going to step 2.

We have implemented the scheme in Figure 1 using vector quantization. The image is divided into 4 x 4 blocks. A 16-dimenssional vector quantizer is designed to code the black and white component and a 2-dimensional vector quantizer is designed to code the sub-sampled colour components. The luminance component is coded using a scalar quantizer. An example image is shown in Fig.3, here an 8-bit 16-D VQ encoder was used to code the black and white component, also an 8-bit 2-D VQ encoder was used to code the colour components. An 8-bit scalar quantizer codes the luminance. Note that this image was not included in the training image set used to train the code book. The compression ratio achieved is 16:1 (higher compression can be achieved if entropy coding is employed). At this level of compression, a large set of images was judged to have very good image quality. It should be noted that our objective in this work is not to strive for high compression ratio (JPEG will achieve higher compression than the present method), the goal instead is to achieve sufficient compression and at the same time facilitate image content classification (management). The merits of the present scheme should therefore be regarded as an elegant compromise.

Content Classification

Colours are excellent cue for image classification. In a recent work,[8] it was shown that using a chromaticity image of the opponent red-green and yellow-blue components (without using the luminance information), excellent image classification/recognition results could be achieved. In a 2-d chromaticity image, each location of the image defines the chromaticity of a particular colour and the pixel brightness encodes the frequency of occurrence of a particular chromaticity in an image). Image recognition is achieved by matching the chromaticity distribution of the images. In the above image-coding scheme, the output of the colour quantizer is the chromaticity index of the opponent colour components of a small block. The distribution of these chromaticity indices can then be directly used to form the opponent chromaticity distribution image, which can then be used to classify images based on their colours. What is important here is that the colour indices are illuminant intensity *independent*, which means recognition performance will not be affected by variance in illuminant intensity.

Low-level statistics can also be used for image classification. It was found that low-level statistics such as the orientation patterns of small image blocks could be used to classify city/suburb photographs.[9] The code-words of the black and white component contain such low level information that can be easily exploited.

Summary and Future Work

A colour image coding system suitable for simultaneous image data compression and classification has been developed. The compression performance of the system was shown to be very good. Importantly, the compressed data of the system can be directly used for image classification without extra processing.

Currently, we are working on combining the two cues, colour and pattern (code-word) for image indexing and retrieval. We are working on multi-resolution vector quantization on the luminance channel to capture the spatial structural properties of images. A unified image compression, indexing and retrieval scheme is currently under investigation and we will report results in the future.

References

1. De Marsicoi, M. Cinque, L. Levialdi, S. (1997) “Indexing pictorial documents by their content: a survey of current techniques”, *Image and Vision Computing*, Vol. **15**, pp. 119-141
2. Niblack, W. and Barber, R. (1993), “The QBIC project: Querying images by content using color, texture and shape”, *Proc. of SPIE, Storage and Retrieval for Image and Video Databases I*, Vol. **1908**
3. Netravali, A. N and Haskell, B. G. (1995), Digital Pictures, Representation, Compression and Standards, 2nd edition, Plenum Press, New York
4. Poirson, A. B. and Wandell, B. A (1993), “Appearance of coloured patterns: pattern-color separability”, *J. Opt. Soc. Am. A*, Vol. **10**, pp. 2459 –2470
5. Kaiser, P. K. and Boynton, R. M. (1996), *Human Color Vision*, Optical Society of America, Washington DC
6. M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection", Compaq Cambridge Research Laboratory, Technical Report Series, CRL 98/11, 1998
7. A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Boston:Kluwer, 1992
8. Berens, J., G. D. Finlayson and Qiu, G., (1999), “A statistical image of colour space”, IEE International Conference on Image Processing and Its applications, IPA’99, Manchester, UK
9. M. Gorkani and R. W. Picard, "Texture orientation for sorting photos", *MIT Media Lab Technical Report* No. 292, 1994
10. *IEEE Trans on Image Processing*, vol. **5**, No. 2, 1996, Special Issue on Vector Quantization
11. S. C. Ahalt, A. K. Krishnamurthy, P. Chen, and D. E. Meton, 'Competitive learning algorithms for vector quantization', *Neural Networks*, Vol. **3**, pp. 277-290, 1990

Acknowledgement

This work was done while the authors were at the School of Maths and Computing (GQ) and the Colour and Imaging Institute (GDF) at the University of Derby, UK

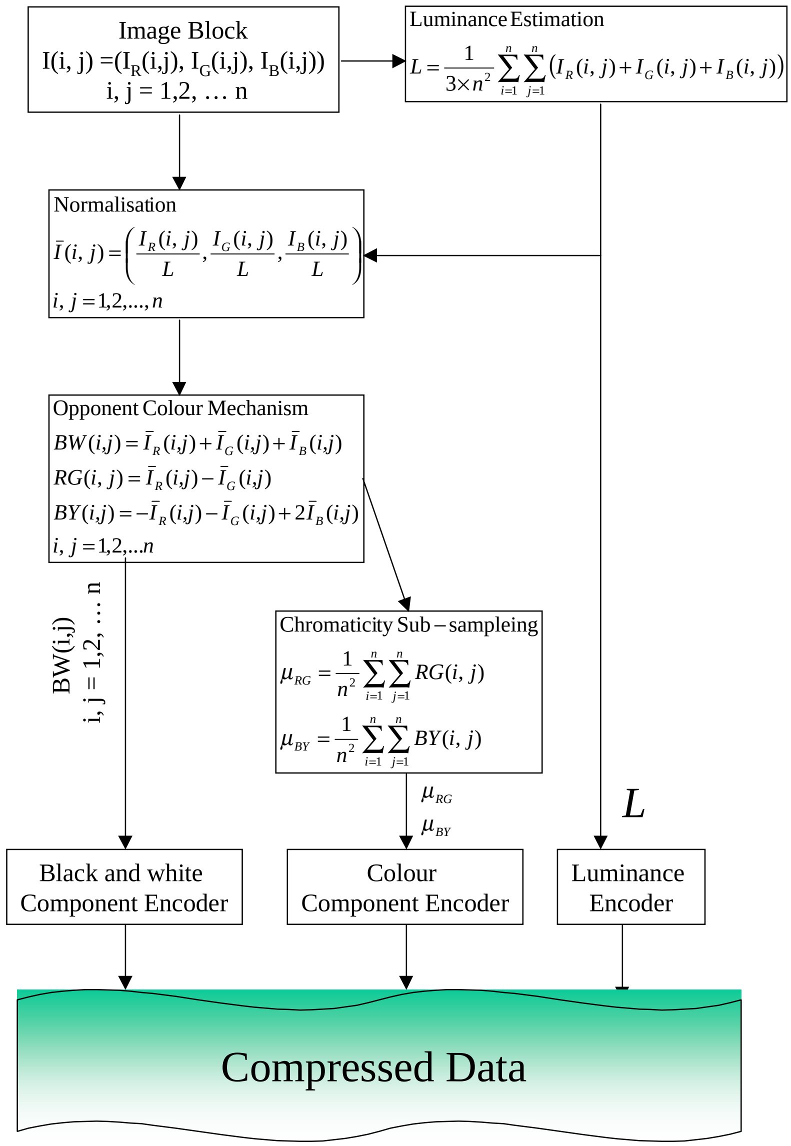

Figure 1. Colour Image Coding for Classification

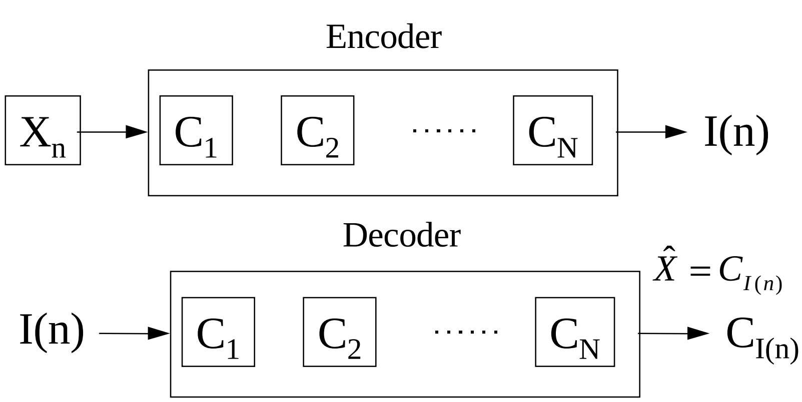

Figure 2. Vector quantizer

Figure 3. A Coded image at a compression ratio of 16 : 1