Image compression for indexing based on the encoding cost map

D. Simitopoulos, D. Tzovaras, M. G. Strintzis

Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki 54006, Greece

Informatics and Telematics Institute
1st Km Thermi-Panorama Rd,
Thessaloniki 57001, Greece
e-mail:dsim@iti.gr

ABSTRACT

The present paper proposes a hybrid content-based image search and retrieval technique based on the use of the DC coefficient information and the Encoding Cost Map (ECM), extracted from compressed JPEG data. The proposed scheme combines color-based indexing of the DC DCT coefficient data with shape-based indexing of the ECM. The use of ECM along with the information carried in the DC DCT coefficients was shown to result to very fast and accurate search and retrieval techniques. Additionally, an indexing oriented rate-distortion technique for the optimal thresholding of AC coefficients was introduced and tested. This technique selects the AC coefficients of each block that are going to be coded in order to achieve better indexing results, while producing JPEG compliant files.

1 Introduction

Due to the low cost of digital cameras, scanners, storage and transmission devices, digital images are now currently employed in an eclectic range of different areas such as entertainment, art galleries, advertising, medicine and geographic information systems among others. While significant advantages have been made in the development of efficient compression techniques [1, 2] which can be used to lessen storage and transmission requirements, efficient and effective techniques must still be investigated in order to improve retrieval of visual data from databases [3].

Visual elements such as color [4], texture, shape [5], structure and spatial relationships have been used as clues for retrieving database images with similar content. Such indexing and retrieval approaches try to improve indexing performance by exploiting existing image features [6]. A more sophisticated indexing optimization approach should also consider compressing images in an indexing-oriented way before applying the relevant indexing technique.

In the present paper, an indexing and retrieval scheme based on the exploitation of color and shape characteristics of the images is described. A query image is input to the system in order to retrieve similar, in color and shape characteristics, images. The indexing algorithm uses luminance and color information that exists in the DC coefficients of discrete cosine transformed blocks in the YUV color space. Specifically, the algorithm matches the global luminance and color characteristics of two images using histograms of their DC coefficients. It also uses the encoding cost map (ECM) [7] in order to perform a shape-based indexing of images. The ECM is a map related to the cost (in number of bits) of each DCT coefficient or of each 8×8 block of a JPEG image. Histograms of preprocessed ECM images are used for the indexing procedure.

In order to optimize the performance of the proposed indexing procedure, an indexing-oriented way of creating JPEG compliant files is introduced. Specifically a novel rate-distortion JPEG compliant image compression technique is proposed that is based on the appropriate thresholding of the AC DCT coefficients and is aiming at the simultaneous minimization of the distortion of the reconstructed image and the maximization of the performance of the proposed indexing algorithm.

2 Content-Based Search and Retrieval of Images Based on Color and Shape characteristics.

The images are stored in our database using the JPEG format. Color-related indices are extracted from the DC coefficients of the DCT blocks, while shape related indices are extracted from the AC coefficients of the DCT blocks as shown in Figure 1.

2.1 Colors Indexing of the DC Information

Color is one of the most recognizable elements of image content. For capturing the distribution of colors in images, color histogram, being invariant to image rotation, translation and viewing axis, is by far the most commonly used technique. Given an image query, an appropriate reference color histogram is formed. This histogram is then compared with all the color histograms of the stored images and those images whose color histograms are sufficiently close to the reference histogram are returned as answers to the query.

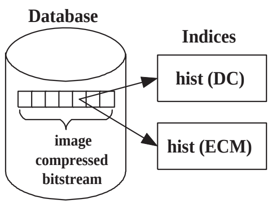

Figure 1: Indexing of compressed data.

DC DCT coefficient indexing based on color was used in the proposed scheme. The maximum possible DC coefficient value of an 8×8 discrete cosine transformed block (in case all pixels of the block take the value 255) is $255 \times 8 = 2040$. For the whole image, the DC coefficients of all three components are quantized by a uniform quantizer that has a step size of eight. Thus, the maximum value is 255. This is done in order to have a reduced size histogram of $K = 256$ bins. The histogram value for each bin is part of the feature vector which should not become too long. The histogram of the divided DC coefficients is evaluated for each one of the Y, U and V components, respectively.

In the following, the histogram values are scaled in order to produce a histogram that would correspond to the histogram taken from a resized copy of the original image to the default set dimensions. In our experiments, the *default height* and *default width* parameters were chosen to be 512. Finally, the histogram values evaluated for each image of the database are stored and form the color-based indices of the image.

When a query image is supplied, the same procedure is repeated for it and its histogram is evaluated. The histogram is then compared to all the histograms stored in the database. For each comparison, the sum of the absolute differences of all the histogram values is evaluated. This operation is performed independently for each one of the Y,U,V components resulting in $sum\{d_Y\}$, $sum\{d_U\}$, $sum\{d_Y\}$ which are given by the following equations.

$$sum\{d_Y\} = \sum_{i=1}^{K} \left| h_{Y_Q}(i) - h_{Y_{DB}}(i) \right|$$

$$sum\{d_U\} = \sum_{i=1}^{K} \left| h_{U_Q}(i) - h_{U_{DB}}(i) \right|$$

$$sum\{d_V\} = \sum_{i=1}^{K} \left| h_{V_Q}(i) - h_{V_{DB}}(i) \right|$$

where h is the histogram function, the indices Y, U, V denote the corresponding components and the indices Q, DB refer to the query image and the compared image of the database respectively.

Finally, each sum $sum\{d_Y\}$, $sum\{d_U\}$, $sum\{d_Y\}$ is divided by 256 (the number of histogram values) in order to obtain the mean absolute difference per histogram value.

2.2 The Encoding Cost Map (ECM)

The cost in an encoding process is defined as the number of bits required to encode a particular symbol or collection of symbols. An encoding cost map (ECM) can either be defined in a block-by-block basis or in a coefficient-by-coefficient basis. In the first case a ECM can be represented as an image with $(M/8) \times (N/8)$ dimensions, where M,N are the width and the height of the image respectively, and in the latter case ECM has the same dimensions with the image. The ECM is one key feature in processing compressed images and will be used in the next sections. The ECM has two basic properties.

First, the ECM provides a way to address the compressed data. After the ECM is derived, individual blocks can be addressed easily thus enabling operations such as cropping, segmentation, rotation, etc.

Second, the ECM conveys information pertaining the activity of a block. Because of JPEG's compression strategy, smooth areas generate low ECM entries, while edges generate high ECM entries. Using the ECM we take advantage of the JPEG computation for modeling edges and such by only measuring the degree of success or failure to compress a particular block. Thus, the ECM provides edge activity information.

The first of the aforementioned properties will not be exploited in the present paper. However, the second is the basis of the developed shape-based search algorithm.

2.3 Shape-Based Indexing of the ECM

The ECM visual properties are mainly related to the largest image components, and among their features the shape, the texture and the orientation play a major role. In many cases shapes can also be defined in terms of presence and distribution of oriented sub-components. A thick object has most of its lines arranged along its primary direction, while a thin elongated object has a peak in the line orientation distribution. Therefore, the orientation of objects within an image is a key attribute in the definition of the similarity with other images.

Following this assumption, we have defined a metric for image classification based on orientation in the two-dimensional space, that is quantified by signatures composed of histogram values of the ECM image components.

Shape-based indexing of the image is also integrated in the proposed scheme by extracting features from the histogram of the block-by-block ECM corresponding to the image. Prior to the evaluation of the histogram, appropriate preprocessing of the image has to be applied.

Firstly, every sum of the values of all pixels that belong to the same row is evaluated. Suppose that we evaluate these sums for two different ECM. Two problems may arise if we try to compare them. The images may be compressed with different quality level, or they may be of different size. This would make the sums incomparable. In order to solve the first problem, each sum of the compared images is normalized by a division with the mean value of the ECM image.

After this normalization, in order to be able to compare these sums with the ones computed from another image of different size, these sums must first be scaled. Each sum is multiplied by the scaling factor *default height / image height*. As mentioned earlier, *default height* and *default width* are both set to 512. This operation produces sums as if they were computed from an image with height equal to default height. But the number of sums equals the width of the image. In order to change the number of sums we use linear interpolation. As a result the number of sums becomes equal to the *default width*.

By following this procedure we obtain sums of rows of different sized images that are the same in number and have comparable values. Finally, a histogram of the sums values is evaluated. This histogram is evaluated and stored for every image of the database. In order to compare the query image with an image of the database, we evaluate its histogram and then, for each image of the database, compute the sum $sum\{d_{Ver}\}$ of the absolute differences of all the histogram values. Finally, the sum is divided by the number of histogram values in order to obtain the mean absolute difference per histogram value.

For every image four histograms are evaluated. First the one just described above. Then, using a similar approach, the histogram of the horizontal lines sums and two histograms for the two types of diagonal lines sums as shown in Figure 2. When two images are compared the values $sum\{d_{Ver}\}$, $sum\{d_{Hor}\}$, $sum\{d_{Diag1}\}$ and $sum\{d_{Diag2}\}$ that are calculated from the corresponding histograms define a similarity metric. The smaller their value is, the closer the images match.

2.4 Combining color and shape-based indexing

The two indexing methods described in the above sections can be combined in order to match images with respect to both their color and shape. Therefore, we define the metric D which is a measure of the difference between two images. D is a function of all the sums of absolute differences defined in the previous sections as shown in the following equation

$$D = w_Y sum\{d_Y\} + w_U sum\{d_U\} + w_V sum\{d_Y\} +$$
$$+ w_{Ver} sum\{d_{Ver}\} + w_{Hor} sum\{d_{Hor}\} +$$
$$+ w_{Diag1} sum\{d_{Diag1}\} + w_{Diag2} sum\{d_{Diag2}\} \quad (1)$$

where the weights w define the importance of the corresponding color component and the importance of each one of the four ECM based indices when the matching algorithm is applied.Experiments with different weighting factors are described in section 4.

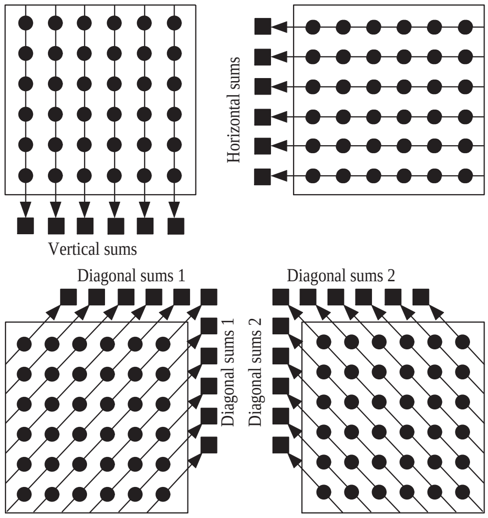

Figure 2: Histogram extraction of the vertical, horizontal and two diagonal lines.

3 Indexing oriented rate-distortion optimal thresholding of AC coefficients

In the previous section, indexing techniques on compressed images were described. Optimization of the indexing and matching results can only be achieved through an indexing-oriented optimization of the image compression. This leads us to the use of rate-distortion techniques as explained in [1, 8].

Rate-distortion optimal thresholding of the AC coefficients of each block is applied so that for a given bitrate R_{budget} of the image, a distortion metric D is minimized. In order to solve this constrained problem we follow the Lagrangian approach and solve the equivalent unconstrained problem, which requires the minimization of $J(\lambda)$

$$J(\lambda) = D + \lambda R$$

The distortion metric D takes into account the perceptual degradation of the image D_{percep}, as well as the deterioration in terms of indexing performance D_{index} that is introduced by thresholding AC coefficients of a DCT block. Thus, the distortion metric D is given by the following equation

$$D = D_{percep} + aD_{index} \quad (2)$$

where a is a constant that balances the contribution of indexing performance deterioration and perceptual quality degradation to the total distortion function. D_{percep} is assumed to be the sum of $D_{percep}(i)$ for all AC coefficients of an image block. $D_{percep}(i)$ is the perceptual distortion metric for each AC coefficient of an image block and is equal to the squared error between the original unquantized DCT coefficient C_i and the reconstructed coefficient $\hat{C}_i$ (after dequantization).

$$D_{percep}(i) = (C_i - \hat{C}_i)^2$$

D_{index} is the sum of the absolute differences of the corresponding values of the four ECM-based histograms h_x, extracted from the original unquantized DCT coefficients, and the histograms $\hat{h}_x$, extracted from the existing in each optimization step coefficients (x defines the orientation of the sums for which the histogram is calculated). The histograms are calculated using the procedure described in section 2.3.

$$D_{index} = \sum_{i=1}^{K} [(h_{Ver}(i) - \hat{h}_{Ver}(i)) + (h_{Hor}(i) - \hat{h}_{Hor}(i)) +$$
$$+ (h_{Diag1}(i) - \hat{h}_{Diag1}(i)) + (h_{Diag2}(i) - \hat{h}_{Diag2}(i))]$$

where K is the number of the histogram bins.

Therefore $J(\lambda)$ is given by

$$J(\lambda) = D_{percep} + aD_{index} + \lambda R$$

The minimization process for $J(\lambda)$ is performed for each DCT block of the image as in [1], using a dynamic programming algorithm.

4 Experimental Results

The proposed method was tested for the content-based search of images using the 502 photographs of the Corel Photo Gallery.

Content-based search was performed using the indices extracted from the database images with the techniques described in section 2. The database images were stored with the indexing optimized method described in section 3.

Figure 3 contains search results using only the color-based indices (the weighting factors concerning the shape-based indices of equation 1 are set to zero). Figure 4 depicts the search results in the case all weighting factors of equation 1 are equal. The results are improved when the combined color and shape-based search is performed.

(a) Input image (b) 1st match (c) 2nd match (d) 3rd match (e) 4th match (f) 5th match (g) 6th match

Figure 3: Image retrieval using color-based indices.

5 Conclusions

The present paper proposed a content-based image search and retrieval technique based on the use of the Encoding Cost Map, extracted from compressed JPEG data. The use of ECM along with the information carried in the DC DCT coefficients was shown to result to very fast and accurate search techniques for indexing-oriented compressed JPEG images.

(a) Input image (b) 1st match (c) 2nd match (d) 3rd match (e) 4th match (g) 5th match (f) 6th match

Figure 4: Image retrieval using combined color- and shape-based indices.

References

[1] K. Ramchadran and M. Vetterli, "Rate - Distortion Optimal Fast Thresholding with complete JPEG/MPEG Decoder Compatibility," *IEEE Trans. Image Processing*, vol. 3, no. 5, pp. 700–704, September 1994.

[2] D. Tzovaras and M. G. Strintzis, "Motion and Disparity Field Estimation Using Rate Distortion Optimization," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 2, pp. 171–181, April 1998.

[3] A. Del Bimbo, *Visual Information Retrieval*, Morgan Kaufmann, San Francisco, 1999.

[4] H. Yamamoto, H. Iwasa, N. Yokoya, and H. Takemura, "Content-Based Similarity Retrieval of Images Based on Spatial Color Distributions," in *Proc. International Conference on Image Analysis and Processing*, 1999, pp. 951–956.

[5] C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Region-based image querying," in *CVPR '97 Workshop on Content-Based Access of Image and Video Libraries*, 1997.

[6] Y. A. Aslandogan and C. T. Yu, "Techniques and Systems for Image and Video Retrieval," *IEEE Transactions on Knowledge and Data Engineering*, vol. 11, no. 1, January/February 1999.

[7] R. L. de Queiroz, "Processing JPEG-Compressed Images and Documents," *IEEE Trans. Image Processing*, vol. 7, no. 12, pp. 1661–1672, December 1998.

[8] K. Argyraki, D. Tzovaras, N. V. Boulgouris, and M. G. Strintzis, "Rate-Distortion Optimal Fast Thresholding for MPEG-2 Image Sequence Coding," in *Proc. Packet Video 2000*, Cagliari,Italy, May 2000.