Proceedings of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)

Image Indexing and Retrieval in Compressed Domain Using Color Clusters

Meekal Bajaj
Electronics and Communication Engineering Department
Netaji Subhas Institute of Technology
New Delhi 110045, India
meekalbajaj@ieee.org

Jose A. lay
School of Electrical and Information Engineering
University of Sydney
NSW 2006, Australia
jlay@ee.usyd.edu

Abstract– **A compressed domain color-based image indexing method that avoids overheads associated with full decompression and color space transformation by operating in the YCbCr space has been presented. The proposed method performs a hard segmentation of the color space into predefined clusters based on perceptual similarity. A three dimensional color descriptor is generated by extracting the average intensity value of the Y, Cb and Cr components for each 8X8 DCT block. Each blocks descriptor is subsequently allocated to its corresponding cluster, whose centroid is updated using a weighted mean. The similarity measure is evaluated using a compact four dimensional feature vector from each cluster, which consists of the quantized centroid co-ordinates and the percentage contribution to the image composition by the cluster. Retrieval is performed by comparing the feature attributes of the clusters having significant membership in the query image, to those with the database images; the obtained matches are then integrated to evaluate the final ranking. The experimental results of the proposed model on a database of 7380 images are reported.**

Index Terms– **Color indexing, compressed domain, content based retrieval, DCT.**

I. Introduction

Contemporary search engines such as Google and Yahoo rely on metadata for image indexing; thereby creating an inherent limitation in their retrieval efficacy. Obtaining meaningful results while searching through the large amount of visual data available online necessitates some form of feature extraction from images for classifying relevant results. Color is an intuitive cue for image retrieval, and is a very effective low-level feature for indexing. Color histograms for image indexing were proposed by [1], and remain a popular feature representation for searching image databases. Other proposed approaches for indexing include, the use of color moments by [7] and color sets as an approximation to color histograms by [8]. Most approaches, including [4], towards color based retrieval convert the color space to the more tractable HSV for similarity matching. This adds computation time, specifically when multiple transformations are applied. Popular compression standards like JPEG, almost unexceptionally, store images using the YCbCr color space, because of the higher levels of compression that can be achieved through subsampling of the CbCr components. We hold that color as a stand alone query in the compressed domain indexing of images using the original color space has not been adequately concentrated on. We propose a model to address this that performs retrieval within the compressed domain YCbCr color space, thus eliminating the need for repeated transformations and the associated overhead.

Images as a norm are now being stored in the compressed domain, so as to cater to the constraints imposed by storage and transmission costs in managing large amounts of image data. However, most content-based retrieval systems assume full images to be available for feature extraction and analysis; this translates into decompression being a prerequisite for implementation of the system. Our approach models a compact, global color descriptor in the compressed domain that can be used for efficient image retrieval. The color space is coarsely divided into visually similar regions, which define the cluster membership boundaries. Faster retrieval is achieved by selecting from each cluster a representative color corresponding to its centroid. Consequently, the feature space is defined by the Y, CB, Cr tuple along with the clusters contribution to the image composition, is indexed in the feature space. Ranking of the results is performed by employing a modified Euclidean distance measure. Retrieval time is further reduced by quantization of the YCbCr color space into 343 colors, and by only evaluating the members of the query color's cluster.

II. Background

A. Existing image indexing techniques in the compressed domain

Jiang et al. [10] proposed an image indexing algorithm analyzing the relationship between DCT coefficients of one block of 8X8 pixels with its four sub-blocks of 4X4 pixels to extract an approximated smaller sized image. Ngo et al. [2] apply Mandala transformations in the DCT domain to extract ten blocks. One of these blocks is used to compute the color description after being transformed into the HSV color space. Color space transformation is also performed in [12]. While [13] uses the YCbCr space, color as an individual descriptor is

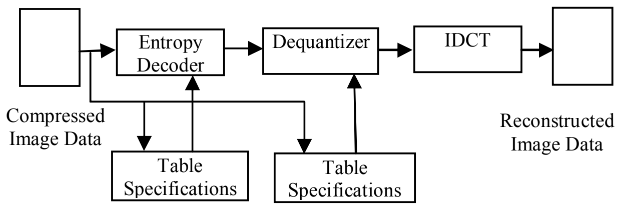

Fig. 1. DCT based decoder processing steps [3]

not available, since extraction is performed on an integrated 12-D feature vector. Additionally, no effort has been made to address the issues regarding the implementation of YCbCr color space.

B. JPEG compression scheme

In the JPEG standard, the source image samples are grouped into 8 X 8 blocks that are transformed independently by DCT using (1)

$$F_{u,v} = \frac{1}{4} C(u)C(v) \sum_{i=0}^{7} \sum_{j=0}^{7} \frac{\cos(2i+1)u\pi}{16} \frac{\cos(2j+1)v\pi}{16} f_{i,j} \tag{1}$$

Where $C(u),C(v) = 1/\sqrt{2}$ if $u,v = 0$, otherwise $C(u),C(v) = 1$
$F_{u,v}$ is the 2-D DCT coefficient and $f_{i,j}$ is the image spatial value. $F_{0,0}$ is normally referred to as the DC coefficient, whose value is obtained from (1), so that

$$F_{0,0} = \frac{1}{8} \sum_{i=0}^{7} \sum_{j=0}^{7} f_{i,j} \tag{2}$$

The zero frequency coefficient of the JPEG image in the DCT domain is a measure of the average value of the 8 X 8 block of image pixels, and is used for the extraction of color data.

The key steps in DCT based decoding are illustrated in Fig.1. We avoid the computational expense incurred during IDCT by using the value after the dequantizer stage in the partial reconstruction of the compressed image data.

III. Feature Extraction and Image Indexing

The DC coefficient maps linearly to the uncompressed domain; therefore, its value is indicative of the position of the component in the fully reconstructed image's YCbCr color space. One of the major drawbacks of the YCbCr space in retrieval is its perceptually non-uniformity, so that moving identical distances along different directions, does not result in similar changes in perception of color. The color space has been represented in Cartesian co-ordinates, with luma varying along the z-axis and Cb and Cr varying along the y and x-axis respectively. The range of values along each axis is normalized so as to extend from zero to unity. The achromatic components are located at the centre of the normalized Cb-Cr planes, and vary in intensity from black to white as we move along the z-axis. Each tuple of co-ordinates in YCC space identifies a color. To augment computational efficiency while performing retrieval, we quantize the color space into 343 colors, such that there are seven divisions along each axis. Each quantized color can be visualized as a cube in space. These cubes are used as the unit for measurement of spatial distance in the enumerated method

We empirically determine perceptual planes which demarcate visually similar regions of color. The perceptual planes are found to intersect the YCC space at an incline, so that the perceptually similar regions for different colors occupy unequal volumes in space. The color space is then coarsely divided into a predetermined number of clusters, whose dimensions are decided by the regions marked off by the planes.

A three dimensional color descriptor is extracted for each 8X8 block, whose coefficients are determined by the DC values of the Y, Cb and Cr channels. The physical location of the point, specified by the 3-D vector in the color space, is viewed as the criterion for membership to the appropriate cluster. Each cluster nominates a representative color that is determined by its centroid's final position. The cluster's centroid is updated with the inclusion of a new member using (3)

$$C_{i\,(new)} = C_{i\,(old)} + \alpha\,(f_i - C_{i\,(old)}) \qquad i \in [1,n] \tag{3}$$

Where α = 1 / (No. of members belonging to the cluster)
C_i and f_i are the values of the cluster centroid and the color descriptor vector along the i th dimension; n is the number of significant clusters.

The quantized co-ordinates of the centroid and the number of members for each cluster are then indexed as the image's feature vector and stored in a database.

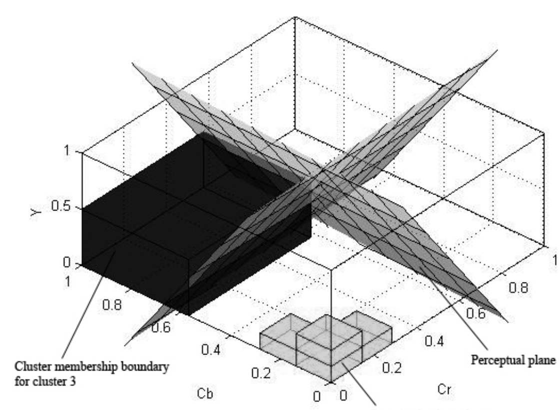

Fig. 2. Representation of constituent elements of the color space.

IV. Image Retrieval

The offline database for the given set of images is used for retrieval, as shown in Fig. 3, such that for every new query image, the quantized co-ordinates of each clusters centroid, I_x, I_y and I_z and the contribution of the cluster to the overall composition of the image is extracted. The extracted feature vectors from the significant clusters are then assessed using the distance measure with the local feature attribute database.

The evaluation of the measure of similarity takes into account similar colors by including a weighted contribution of the physically adjacent cubes in the color space, and thus it is performed for color cubes in the query cluster satisfying the co-ordinates:

$$(I_{x+u} , I_{y+v} , I_{z+w}) \qquad u,v,w \in -1, 0, 1$$

An inherent limitation is imposed on the division of the YCbCr color space due to its non uniform nature; so that perceptually similar colors may occasionally also be quantized separately. A wider scope serves to mitigate the effect of segregation of similar colors in the same cluster into different cubes. While this may lead to an increase in false retrievals, we emphasize the need for retrieving all relevant images.

A. Distance Measure

A modified Euclidean distance measure has been formulated so as to be a function of two factors, first, the overall contribution, η of the given cluster to the image, i.e. the percentage of the image that is composed of members from a given cluster. Second, the city-block distance, d between the evaluated and the query cube. For the block of cubes being examined, *d* is also a measure of the number of dimensions being traversed. A Boolean operator m, tests whether the evaluated and the query cube are members of the same cluster. The similarity measure is assessed using (4).

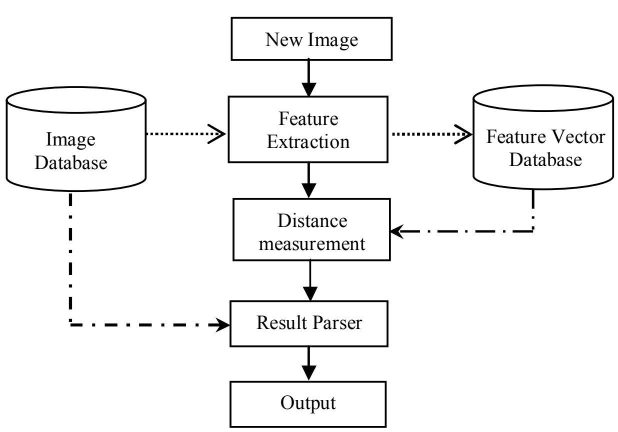

Fig. 3. Image retrieval model.

$$\text{Similarity Measure } (SM_c) = \sum_{e=1}^{27} \frac{(1 - |\eta_q - \eta_e|) \cdot m}{\exp(|d_x| + |d_y| + |d_z|)} \qquad (4)$$

Where, *e* is the cube number currently being evaluated. $(|d_x| + |d_y| + |d_z|) \in 0,1,2,3,\ \eta \in 0\text{-}1$ and $m \in 0,1$

The perceptual similarity between colors decreases sharply as the number of dimensions along which the color space is traversed increases. In order to model this behavior, our distance measure is made inversely proportional to the exponential of the city block distance.

$$\text{Rank}_i = \sum_{c=1}^{n} SM_c \cdot wt_c \qquad (5)$$

The final ranks of the returned images are determined by a weighted match value that incorporates the similarity measure of all significant clusters as calculated in (5). The clusters are initially equally weighted, however by making use of relevance feedback [14], the user is allowed the flexibility in determining the color features that they found to be more congruous, by marking them as *highly relevant, relevant, no opinion, irrelevant and highly irrelevant*. The system returns a more apt set of results by updating the weights according to the user's feedback. The contribution to the distance measure of the color features that have been marked relevant is increased, while the existing weights are retained in case of no opinion.

V. Experimental Results

We implemented the proposed system on a single uncategorized database of 7380 JPEG photographs having a visually dominant hue, depicting a wide range of real-life activities The collection consists of images drawn from [12], [16] and [18].

Subjective testing was performed before the evaluation to determine the relevant matches in the database for the query image. The images selected as being perceptually similar in color content were marked, and considered as ground truth to evaluate the retrieval accuracy. Precision and recall are used as measures of retrieval accuracy, defined as

$$\text{Precision}_k = N_k / k \text{ and } \text{Recall}_k = N_k/m$$

Where, k is the number of retrievals, N_k is the number of relevant retrievals and m is the total number of relevant images in the database.

The average Precision-Recall (P-R) curves for a set of 112 query images from the database have been shown in Fig. 4. The query images were pre-selected manually to be the representatives of different classification groups. A cubic regression on the data points was used in determining the curve. The refined P-R curve was generated after performing a

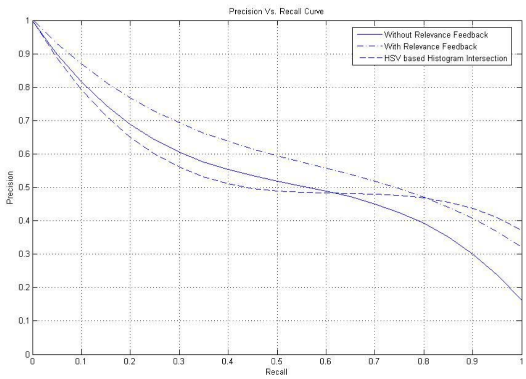

Fig. 4 Average Precision-Recall Curves

single iteration of relevance feedback. The P-R curves generated from the results of the proposed method have been found to be comparable with those obtained by evaluating the popularly used Histogram Intersection distance in the HSV color space, as shown in Fig. 4. Since each image was classified into only one group on the basis of its dominant color while establishing the ground truth, the actual performance in general was found to be better than that shown in Fig. 4. Additionally, multiple iterations of relevance feedback can be employed to obtain better results.

VI. CONCLUSION

In this work a color-based image indexing method in compressed domain has been proposed that operates in the YCbCr color space. The presented method performs a hard segmentation of the color space into predefined clusters based on perceptual similarity. A feature vector is extracted from each cluster, consisting of the representative color and the cluster's percentage contribution to the image composition. Retrieval is performed by the application of a modified Euclidean distance similarity measure on the feature vector database. The results are ranked by a weighted sum of the similarity measure, where the user is given the flexibility of modifying the weight through relevance feedback. Experimental results demonstrate the effectiveness of the proposed method.

Further research could look into modification of cluster shapes to follow the perceptual curves more accurately. Additionally, adaptive binning of the color space may be examined. The inability of the current system to assimilate spatial relationships between colors in the image is another area that can be explored.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of Dr. David Levy and Dr. Jonathan Randall to this work, and thank them for their support and advice.

REFERENCES

[1] M.J. Swain and D.H. Ballard, "Color Indexing," *Int. J. Comput. Vis.*, 1991, vol. 7, no. 1, pp. 11-32.

[2] C.W. Ngo, T.C. Pong and R.T. Chin, "Exploiting image indexing techniques in DCT domain", *Pattern Recognition*, 2001, vol. 34, pp. 1841-1851.

[3] G.K. Wallace, "The JPEG Still Picture Compression Standard", *Communications of the ACM*, 1991, vol. 34, no. 4, pp. 30-44.

[4] B.S. Majnunath, J.-R. Ohm, V.V. Vasudevan and A. Yamada, "Color and Texture Descriptors", *IEEE Trans. on Circuits and Systems for Video Technology*, 2001, vol. 11, no. 6.

[5] Y. Deng et al., "An Efficient Color Representation for Image Retrieval", *IEEE Trans. on Image Processing,* 2001, vol. 10, no. 1, pp. 140-147.

[6] T.S. Huang, Y. Rui and S.-F. Chang, "Image Retrieval: Past, Present and Future",*Proc. of Int. Symposium on Multimedia Information Processing*, Taiwan, Dec. 1997

[7] M. Stricker and M. Orengo, "Similarity of Color Images", *Proc. SPIE Storage and Retrieval for Image and Video Databases*, 1995.

[8] J.R. Smith and S.-F. Chang, "Single Color Extraction and Image Query", *Proc. Int. Conf. on Image Proc.*, 1995.

[9] J. Jiang, A. Armstrong, G.C. Feng, "Web-based image indexing and retrieval in JPEG compressed domain", *Multimedia Systems*, 2004, vol. 9, no. 5, pp. 424 – 432.

[10] J. Jiang, A.J. Armstrong, G.C. Feng, " Direct content access and extraction from JPEG compressed images", *Patt Recog*, 2002, vol. 35, pp. 2511-2519.

[12] James Z. Wang, Jia Li, Gio Wiederhold, ``SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol 23, no.9, pp. 947-963, 2001.

13] H. Nezamabadi-pour and S. Saryazdi, "Object based Image Indexing and Retrieval in DCT Domain using Clustering Techniques", *Trans. on Engineering, Computing and Technology*, 2004, vol. 3, pp. 207-210.

[14] Y. Rui, T.S. Huang, M. Ortega and S. Mehrotra, "Relevance Feedback: A Power Tool for Content Based Image Retrieval", *IEEE Trans. on Circuits and Systems for Video Technology*, 1998, vol. 8, no. 5.

[15] J.A. Lay and L. Guan, "Image retrieval based on Energy Histograms of the Low Frequency DCT Coefficients", *Proc. Int. Conf. Acoustics, Speech and Signal Processing*, 1999, vol. 6, pp. 3009-3012.

[16] Media Graphics International, "*Photo Gallery* 5,000 *Vol. 1", CD-ROM,* http://www.media-graphics.net.

[17] P. Sallee , "Matlab JPEG Toolbox", 2003, http://redwood.ucdavis.edu/phil/demos/jpegtbx/jpegtbx.htm

[18] M. Bajaj, Personal Image Collection