International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 6 418 - 421

Study of a imaging indexing technique in JPEG Compressed domain

Pooja .D. Phumatiya
Department of Master of Computer Application (MCA)
University Of Mumbai
Dr.G.D.Pol Foundation YMT College Of Management
Sector-4 Kharghar, Navi Mumbai-410210, India
phumatiyapooja7791@gmail.com

Abstract- In our computers all stored images are in JPEG compressed format even when we download an image from the internet that is also in JPEG compressed format, so it is very essential that we should have content based image indexing its retrieval conducted directly in the compressed domain.
In this paper we used a partial decoding algorithm for all the JPEG compressed images to index the images directly in the JPEG compressed domain. We also compare the performance of the approaches in DCT domain and the original images in the pixel domain. This technology will prove preciously in those applications where fast image key generation is required. Image and audio techniques are very important in the multimedia applications. In this paper, we comprise an analytical review of the compressed domain indexing techniques, in which we used transform domain techniques such as Fourier transform, karhunen-loeve transform, Cosine transform, subbands and spatial domain techniques, which are using vector quantization and fractrals. So after comparing other research papers we come on the conclusion that when we have to compress the original image then we should convert the image by using the 8X8 pixels of image blocks and after that convert into DCT form and so on. So after doing research on the same concept we can divide image pixels blocks into 4X4X4 blocks of pixels. So by doing the same we can compress the original image by using the steps further.

Keywords: *JPEG,MPEG,DCT,ISDN,KLT,VQ*

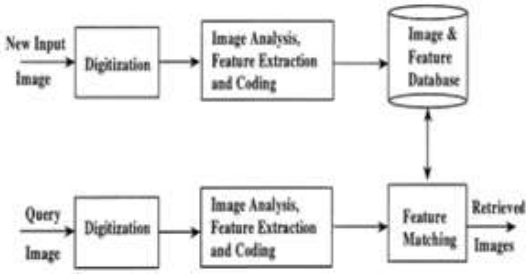
*****

I. INTRODUCTION

As the larger part of the image energy is compressd in the DC coefficient and some part in the AC coefficients along with the frequency bands. All the coefficients can be rated according to their position inside the block. The majority of the coefficients are not required, and it depends upon the application. When we used content descriptin then we can used partial decoding, which can be performed to approximate the pixel information and in the construction of the indexing keys. Its simplest method of doing is that keeping the first few coefficients and ignoring the rest of them. It is reffered as partial IDCT (Inverse Direct Cosine Transform) or partial decoding of the JPEG images. According to this condition the processor time can be saved by partial IDCT. In the JPEG compressed domain the texture indexing method is not new. In the development of computer hardware texture based keys is used to provide as interim solution to the retrieval problem. So our approaches is based only on the improving of the approximated image quality for the key generation and accuracy of its content description relates to the original JPEG image.

Digital image and video indexing techniques are becoming increasingly important with an recent advances in very large scale integration technology (VLSI, broad-band networks (ISDN, ATM) and the image/video compression standards (JPEG/MPEG). The goal of an image indexing is to develop techniques that provide the ability is to store and retrieve images based on their content Some of the potential applications of image and video indexing are multimedia information systems and digital libraries , remote sensing and natural resources management movie industry and video on demand.. Traditional databases use keywords as labels to quickly access large quantities of text data. However the representation of a visual data using text labels needs a large amount of manual processing and entails the extra storage. A more serious problem is that the retrieval of results might not be satisfactory since the query was based on features that may not react the visual content. Hence, there is a need for novel techniques for content based indexing of visual data.

Fig. 1. Schematic of an image archival and retrieval system



II. Image indexing in pixel domain

The visual data's pixel domain indexing is based only on some features such as sketch, histogram, texture, shape, color, moments etc. By the retrieves images which are based on texture, color, sketches and shape a system is developed which is called "query by image content" (QBIC). The content-based retrieval engine (CORE) is used for the Multimedia Information Systems, which is proposed by employing colour and word similarity, which is measures to retrieve images based on content and text annotation. We describe here the art of approaches in image indexing.

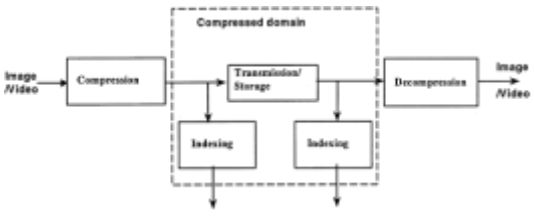


Fig. 2. Block diagram of a compressed domain indexing system.

III. Image indexing in the compressed domain

When volume of visual data is in very large then the use of a compression techniques is very important. So the visual data in the future multimedia databases is required to be stored in the compressed form. For the decompression of the image data and for applying the pixel domain indexing techniques, it is essential to index image/video in the compressed form. These techniques are inexpensive for the computing and storing any data. Compressed domain indexing (CDI) techniques are mainly classified into the two categories: transform domain techniques and the spatial domain techniques.

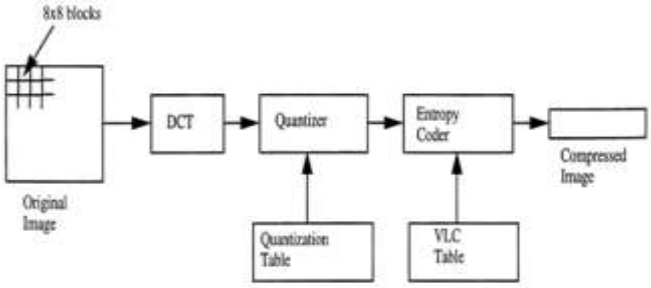
The transform domain techniques is mainly based on the DFT (discrete Fourier transform) and KLT (Karhunen-Loeve transform). In the Spatial domain techniques, vector quantization (VQ) and fractals are used.

3.1. Discrete Fourier transform

Fourier transform is an very important in image and signal processing. DFT employs complex exponential basis functions and provides an good coding performance since it has been a good energy compaction property. DFT has several properties that are useful in the indexing or pattern matching. Firstly the magnitude of the DFT coeffcients are translation invariant. Secondly the spatial domain are correlated and can be effcient to compute using DFT coeffcients. We now present selected Fourier-domain indexing techniques.

Stone and Li have proposed and evaluated an image retrieval algorithm in Fourier domain. The algorithm has a two thresholds that allow the user to independently adjust the closeness of a match. One threshold controls the intensity match while the other controls a texture match. The thresholds have correlated values that is computed effciently using an Fourier coeffcients and are particularly effcient when the Fourier coeffcients are mostly zero.

3.2 Discrete cosine transform



DCT a derivative of DFT employs a real sinusoidal basis functions and has energy compaction inefficiency of close to the optimal KL transform for most of natural images. As a result, all international image and video compression standards such as an JPEG, MPEG 1 and 2, H.261/H.263 employ DCT. We now provide a brief description of a DCT based JPEG baseline algorithm hence all the above mentioned standards are employ a similar algorithm for coding. In JPEG, compression is performed in three steps DCT computation, quantization and variable-length coding. The original image is First partitioned into non-overlapping blocks of 8 X 8 pixels as shown in Fig. 4. The 2D DCT of the block is computed and quantized using the visually adopted quantization table suggested by JPEG. Each 8 X 8 block generates one DC coeffcient and 63 AC or an high-frequency coeffcients. The quantized coeffcients are reordered using zigzag scan and pattern to form an 1D sequence of quantized coeffcients. The DC coeffcients from each block is DPCM coded and all other coeffcients, i.e., the AC coeffcients, are compressed using a combination of Huffman and run-length coding. We have note that the above given DCT based standards do not address the aspect of indexing. We have discuss some DCT based indexing techniques that has been appeared in the recent literature.

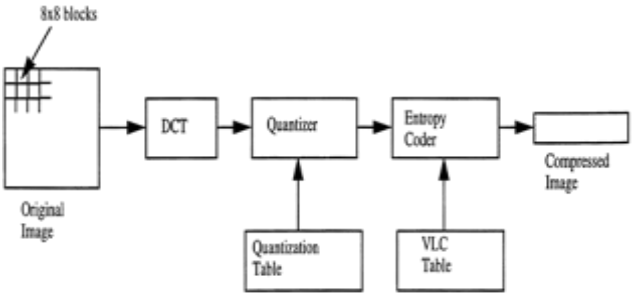


Fig. 3. Baseline JPEG encoder.

IV. Video indexing in compressed domain

In this we present a review of video segmentation techniques in DCT, DWT and VQ domains. We saw that motion vectors are not available for image indexing is an important feature for video segmentation. Hence a review of motion vector-based video segmentation will also be presented.

4.1 DCT coeffcients

We saw that the international standards for image and video compression (JPEG, MPEG, H.261 and H.263) are based on DCT. The transform coeffcient are in the frequency domain is related to the pixel domain. Therefore the DCT coeffcients can be used for see change detection in compressed video sequences.

Before discussing the indexing techniques we provide anoverall description of the MPEG algorithm. In MPEG a block-based motion compensation scheme is employed to remove the temporal redundancy. Because of the connecting requirements of random access and high compression ratio, the MPEG standard suggests that frames canbe divided in three categories I, P and B frames. The organization of the three frame types in a sequence is very flexible. It illustrates the relationship among the three different frame types in the group of pictures (GOP). Intra-coded frames (I frames) is coded without reference to other frames and employ the coding scheme similar to a JPEG baseline scheme. Predictive coded frames (P frames) is coded more effciently using motion-compensated predition from the past I or P frames, and are generally used as an reference for further prediction. Bi-directionally predictive coded frames (B frames) provide an the highest degree of compression but require both the past and future reference frames for motion compensation. We note that is the B frames are never used as a reference for prediction

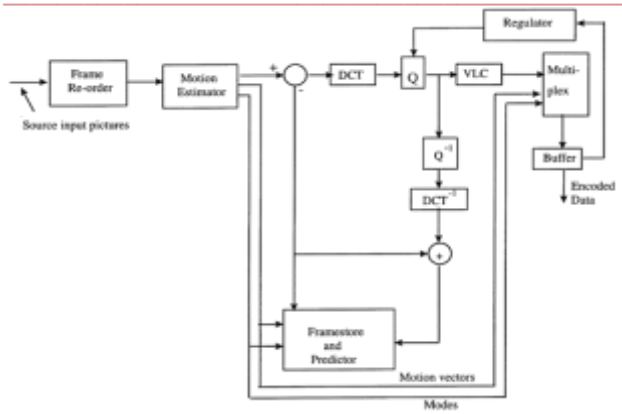


Fig.4 Block diagram of MPEG video encoder

Zhang et al. [80] have presented a pair-wise comparison technique for the intracoded (I frame) where the corresponding DCT coeffcients in the two frames f_m and f_n are matched. This is similar to the pixel intensity matching technique (see Section 4) in the uncompressed domain. Here, the pair-wise normalized absolute difference $D(f_m, f_n, l)$ of the I block in two frames f_m and f_n is deter-mined using

$$D(f_m,f_n,l) = \frac{1}{64} \sum_{k=1}^{64} \frac{|c(f_m,l,k) - c(f_n,l,k)|}{\max(c(f_m,l,k), c(f_n,l,k))} \quad (1)$$

$$w = 1 - \frac{Vf_m \cdot Vf_n}{|Vf_m||Vf_n|} \quad (2)$$

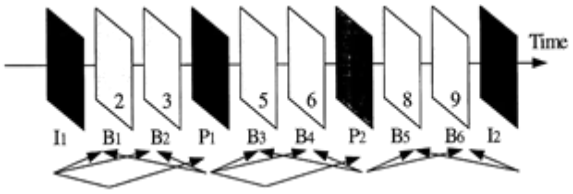
where $c(f_m,l,k)$ is the kth coeffcient of block l in f_m. If the difference $D(f_m, f_n, l)$ is larger than a threshold, the block l is considered to be changed. If the number of changed blocks exceeds a certain threshold, a scene change is declared in the video sequence from frame f_m to frame f_n.

Arman et al. [81] have proposed a technique based on the correlation of corresponding DCT coeffcients of two neighboring frames. For each compressed frame f_m, B blocks are ®rst chosen a priori from R connected regions in f_m. A set of randomly distributed coeffcients $\{c_x, c_y, c_z, \ldots\}$ is selected from each block where c_x is the xth coeffcient. A vector $Vf_m = \{c_1, c_2, c_3, \ldots\}$ is formed by concatenating the sets ofcoeffcients selected from the individual blocks in R. The vector Vf_m represents f_m in the transform domain. The normalized inner product is used as a metric to judge the similarity of frame f_m to frame f_n

$$w = 1 - \frac{Vf_m \cdot Vf_n}{|Vf_m||Vf_n|} \quad (4)$$

A scene transition is detected if W is greater than a threshold. In the case of false positives, which result from camera and object motion, f_m and f_n are decompressed and their color histograms are compared to detect camera breaks. We note that the technique of Zhang et al. [80] is computationally less intensive compared to that of Arman et al., although the former is more sensitive to gradual changes.

We note that the previous two algorithms are applied on video sequences compressed using motion JPEG. In the case of MPEG video, only I frames are compressed with DCT coeffcients and hence the previous two techniques cannot be directly applied to the B and P frames. In addition, the techniques based on I frames may result in false positives. To overcome these problems, Yeo and Liu [82] have proposed a united approach for scene change detection in motion JPEG and MPEG. This algorithm is based on the use of only the DC coeffcients which is explained in the following.



To start with, a DC frame f_m^{DC} is constructed for every frame in the sequence. The DC coeffcients in JPEG and I frames in MPEG are obtained directly from each block. For P and B frames in MPEG video, the DC coeffcients are estimated. The sum of the difference magnitude of the DC

V. JPEG IMAGECOMPRESSION

The JPEG standard defines a lossy method for compressing and decompressing.

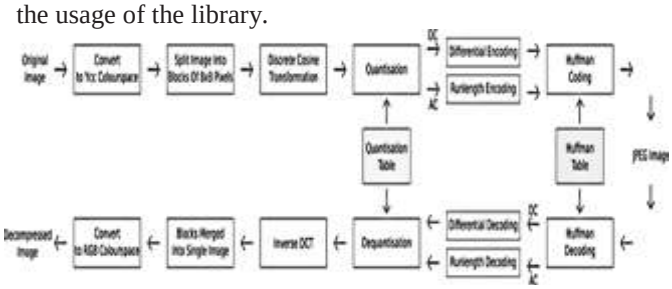
In typical CBIR systems, image features for a dataset are extracted in an offline stage, and are then loaded once and cached in a database. During retrieval, corresponding features for a query image are calculated and then compared to the features stored in the database. In contrast, in a non-line retrieval scenario features for database images are not available and hence have to be extracted during the retrieval process, requiring highly efficient methods to allow for interactive retrieval times.

The vast majority of images are stored in compressed form, typically in JPEG format[1]. While CBIR feature extraction needs to be performed quickly, especially for large image datasets, this is hindered by relatively slow image decompression that needs to be employed before features can be extracted from the pixel domain. A faster approach is to perform this directly in the compressed domain, that is to perform compressed domain image retrieval and support so called mid-stream content access.JPEG is a lossy image compression technique that splits the image into blocks of 8×8 pixels and applies the discrete cosine transformation to each block. This separates high and low frequency information, making it easier to discard visually less important information in a quantization process. After quantization, the DC and AC coefficients are differentially respectively run-length coded, while entropy coding is applied in a final step. By working in the compressed domain of JPEG, image retrieval can be performed on the raw DCT coefficients, avoiding the computationally expensive inverse-DCT and thus yielding a significant speed-up.

In this paper, we present JIRL, an open source C++ library for performing compressed domain CBIR of JPEG images, originally introduced in JIRL contains

implementations of twelve methods, representing the state of the art of image retrieval in the compressed JPEG domain, and provides functionality for both feature extraction and feature comparison.

Included with the library are tools which can be used for benchmarking existing and new compressed domain techniques. The tools allow for evaluation of both retrieval performance (based on a ground truth) and an accurate speed calculation. Also included is an application that performs query-by-example (QBE)retrieval and also serves as a demonstration of the usage of the library.



The JPEG compression scheme

VI. Conclusion

The demanding approach of the multimedia data services constrain the development of storing, navigating and retrieving techniques of visual data. Existing text indexing techniques for the image & video indexing is not efficient and it is very complex. This approach is not comprehensive and it is not very useful in the wide variety of applications. Content-based indexing techniques should be engaged to search for desired images and video in the database. In this paper reviews and summarizes compressed domain indexing techniques proposed in recent literature. The main focus of the review of image feature vectors are generated by using the transform coeffcients. In video indexing the use of motion vectors was also reviewed. Additionaly we also discussed of using pixel domain techniques, which were imported in the compressed domain. The techniques used in this paper is associated with coding techniques that is developed to provide a high compression ratio. For obtaining superior overall performance integrated coding and indexing techniques should be developed. The second generation image and video coding technique which are emerging, are expected to provide a better joint coding and indexing performance. These techniques are generally based on the segmentation or model-based schemes.

Shortly we conclude that in JPEG compressed images and video we can use another method by using pixel blocks 4X4X4, which makes the same thing as 8X8 so we can compressed the original images and make it in the compression form.

VII. Acknowledgement

We thank our colleagues from YMT who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper.

We thank Miss. Swapnali Mahadik for assistance us by providing proper IEEE formats and all faculties for guiding us.

References

[1] B. Furht, S.W. Smoliar, H. Zhang, Video and Image Processing in Multimedia Systems, Kluwer Academic Publishers, Dordrecht, 1995.

[2] M. O'Docherty, C. Daskalakis, Multimedia information systems: the management and semantic retrieval of all electronic datatypes, Computer J. 34 (3) (1991) 225±238.

[3] Digital Libraries, Commun. ACM, 38 (4, Special Issue) (April 1995).

[4] M. Ehlers, G. Edwards, Y. Bedard, Integration of remote sensing with geographic information systems: a necessary evolution, Photo-grammetric Engineering and Remote Sensing 55 (11) (1989) 1619± 1627.

[5] T.D.C. Little, G. Ahanger, R.J. Folz, J.F. Gibbon, F.W. Reeve, D.H. Schelleng, D. Venkatesh, A digital video-on-demand service support-ing content-based queries, Proc. of International Conference on Multimedia, August 1993, pp. 427±436.

[6] V.N. Gudivada, V.V. Raghavan, Content based image retrieval systems, IEEE Computer 28 (9) (1995) 18±22.

[7] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, P. Yanker, Query by image and video content: the QBIC system, IEEE Computer September (1995) 23±32.

[8] J.R. Bach, C. Fuller, A. Gupta, A. Hamapur, B. Horowitz, R. Humphrey, R. Jain, C.F. Shu, The Virage image search engine: an open framework for image management, Proc. SPIE 2670 (1996) 76±87.

[9] J.K. Wu, A.D. Narasimhalu, B.M. Mehtre, C.P. Lam, Y.J. Gao, CORE: a content-based retrieval engine for multimedia information systems, Multimedia Systems 3 (1995) 25±41.

[10] F. Arman, A. Hsu, M.-Y. Chiu, Image processing on compressed data for large video databases, Proc. of ACM International Conference on Multimedia, Anaheim, CA, USA, 1993, pp. 267±272.

[11] R.C. Jain, S.N.J. Murthy, P.L.J. Chen, Similarity measures for image databases, Proc. SPIE: Storage and Retrieval for Image and Video Databases III 2420 (February) (1995) 58±65.

[12] G.K. Wallace, The JPEG still picture compression standard, Commun. ACM 34 (4) (1991) 31±45.

[13] D.L. Gall, MPEG: A video compression standard for multimedia applications, Commun. ACM 34 (4) (1991) 59±63.

[14] M. Liou, Overview of the p 3 64 kbits/s video coding standard, Commun. ACM 34 (4) (1991) 46±58.

[15] P. Aigrain, H. Zhang, D. Petkovic, Content-based representation and retrieval of visual media: a state of the art review, Multimedia Tools and Applications 3 (1996) 179±202.