Fast Extraction of Edge Histogram in DCT Domain based on MPEG7

Minyoung Eom, and Yoonsik Choe

***Abstract*—**In these days, multimedia data is transmitted and processed in compressed format. Due to the decoding procedure and filtering for edge detection, the feature extraction process of MPEG-7 Edge Histogram Descriptor is time-consuming as well as computationally expensive. To improve efficiency of compressed image retrieval, we propose a new edge histogram generation algorithm in DCT domain in this paper. Using the edge information provided by only two AC coefficients of DCT coefficients, we can get edge directions and strengths directly in DCT domain. The experimental results demonstrate that our system has good performance in terms of retrieval efficiency and effectiveness.

***Keywords*—**DCT, Descriptor, EHD, MPEG7.

I. Introduction

ALMOST all digital images are stored in compressed formats, now among them, the formats standardized by joint picture expert group (JPEG) are widely used on internet or image databases [1]. To do the feature extraction using descriptors provided by MPEG7, firstly, the conventional approaches need to decode the compressed image to the pixel domain. In case of edge histogram descriptor (EHD), due to the decoding procedure and filtering for edge detection, the feature extraction process is time-consuming as well as computationally expensive. To improve efficiency of compressed image retrieval, we introduce a fast algorithm, which is utilized for the edge histogram generation, to get the direction and strength of the edge directly in DCT domain what is called EHDiD (Edge Histogram Descriptor in DCT Domain). The experimental results demonstrate that our retrieval system using EHDiD has good performance in terms of retrieval efficiency and effectiveness.

Manuscript received October 14, 2005. This work was supported by grand No. R01-2003-000-10412-0 form the Basic Research Program of the Korea Science and Engineering Foundation.

Minyoung Eom is with the Department of Electrical and Electronics Engineering, Yonsei University, 134 Shichon-dong, Sodaemoongu, Seoul 120-749, Korea (phone: 82-2-2123-2774; fax: 82-2-313-0960 ; e-mail: emy0606@ yonsei.ac.kr).

Yoonsik Choe is with the Department of Electrical and Electronics Engineering, Yonsei University, 134 Shichon-dong, Sodaemoongu, Seoul 120-749, Korea (phone: 82-2-2123-2774; fax: 82-2-313-0960 ; e-mail: yschoe@ yonsei.ac.kr).

II. MPEG-7 Edge Histogram Descriptor

A. *Definition and Semantics*

Spatial distribution of edges in an image is another useful texture descriptor for similarity search and retrieval. The EHD in MPEG-7 represents local edge distribution in the image. Specifically, dividing the image space into 4×4 subimages as shown in Figure 1, the local-edge distribution for each subimage can be represented by a histogram. To generate the histogram, edges in the subimages are categorized into five types; vertical, horizontal, 45-degree diagonal, 135-degree diagonal and nondirectional edges. Since there are 16 subimages, a total of 5×16 =80 histogram bins are required. Table I summarizes the semantics of the 80-bin EHD [2][3].

TABLE I
Semantics of The Histogram Bins of The EHD

Histogram bins	Semantics
BinCounts[0]	Vertical edge of sub-image at (0,0)
BinCounts[1]	Horizontal edge of sub-image at (0,0)
BinCounts[2]	45-degree edge of sub-image at (0,0)
BinCounts[3]	135-degree edge of sub-image at (0,0)
BinCounts[4]	non-directional edge of sub-image at (0,0)
BinCounts[5]	Vertical edge of sub-image at (0,1)
.	.
.	.
BinCounts[78]	135-degree edge of sub-image at (3,3)
BinCounts[79]	non-directional edge of sub-image at (3,3)

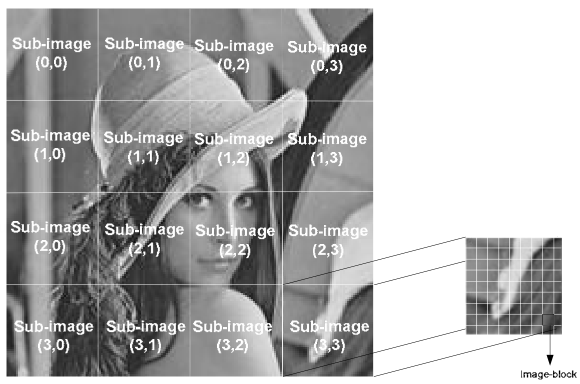

Fig. 1 Definition of subimage and image-blocks

B. *Extraction of Edge Histogram*

For the EHD, it is required to detect nondirectional edges as well as four directional edges. The nondirectional edges include the edges with no particular directionality. The five types of edges can be extracted by a block-based edge extraction scheme. To do that, as shown in Figure 1, each subimage is further divided into nonoverapping square image-bocks. The size of the image-block is dependent on the image resolution. Regardless of the size of the given image, the image space is divided into a predetermined number of image-blocks. Commonly the number of image-blocks is around 1100.

Each of the image-blocks is then classified into one of the five edge categories mentioned above or as a non-edge block. A simple method to do this classification is to treat each image-block as a 2×2 super-pixel image-block and apply appropriate oriented edge detectors (Figure 2) to compute the corresponding edge strengths. The edge detector with maximum edge strength is then identified. If this edge strength is above a given threshold, then the corresponding edge orientation is associated with the image-block. If the maximum of the edge strengths is below the given threshold, then that block is not classified as an edge block.

1	-1	1	1	$\sqrt{2}$	0	0	$\sqrt{2}$	2	-2
1	-1	-1	-1	0	$-\sqrt{2}$	$-\sqrt{2}$	0	-2	2

Fig. 2 Filter coefficients for edge detection

II. Edge Histogram Descriptor in DCT Domain

A. The Meaning of Two AC Coefficients

Before describing how to extract low level feature information, which is edge histogram, directly from DCT compressed images, we briefly describe the meaning of AC coefficients. The heart of the JPEG compression scheme is two-dimensional discrete cosine transform (DCT) which is defined as

$$F(u,v) = K_{u,v} \sum_{i=0}^{7} \sum_{j=0}^{7} \cos\frac{(2i+1)u\pi}{16} \cos\frac{(2j+1)v\pi}{16} f(i,j) \qquad (1)$$

where $K_{u,v} = C_u C_v / 4$ and $C_u, C_v = \begin{cases} \sqrt{1/2} & u = v = 0 \\ 1 & otherwise \end{cases}$

One simple observation is that each DCT coefficient $F(u,v)$ is a linear combination of all pixel values within the block. Our approach for direct extraction of edge histogram is based on the relationship between the pixels' values in a block and its DCT coefficients. For example, the upper left coefficient in DCT block is the DC coefficient and it means average luminance of the block. Other coefficients are all called AC coefficients and they reflect variations in gray level values in certain direction at certain rate. To see this relationship, consider the coefficient AC_{10}. From definition of DCT

$$AC_{10} = K_{10} \sum_{i=0}^{7} \sum_{j=0}^{7} \cos\frac{(2i+1)\pi}{16} f(i,j) \qquad (2)$$

$$= K_{10} \sum_{i=0}^{7} \cos\frac{(2i+1)\pi}{16} \sum_{j=0}^{7} f(i,j)$$

Using the fact that $\cos(\pi - \theta) = -\cos\theta$, the equation (2) can be expanded as

$$\begin{aligned} AC_{10} = K_{10}[&\cos\frac{\pi}{16}(\sum_j f(0,j) - \sum_j f(7,j)) \\ &+ \cos\frac{3\pi}{16}(\sum_j f(1,j) - \sum_j f(6,j)) \\ &+ \cos\frac{5\pi}{16}(\sum_j f(2,j) - \sum_j f(5,j)) \\ &+ \cos\frac{7\pi}{16}(\sum_j f(3,j) - \sum_j f(4,j)) \end{aligned} \qquad (3)$$

The above equation (3) shows that the value AC_{10} essentially depends on intensity differences in the vertical direction between upper and lower parts of the given block. In other words, it provides information of edge strength in the horizontal direction. Similarly, the value AC_{01} provides information of edge strength in the vertical direction for the given block. Figure 3 shows the meaning of two AC coefficients.

AC01 AC10

Fig. 3 Physical meaning of two AC coefficients

B. Decision of Edge Orientation using two AC Coefficients

In order to get the EHDiD, we use just two AC coefficients AC_{01} and AC_{10}. As mentioned above, the two AC coefficients denote the edge strength in the vertical and horizontal orientation of the given block, respectively. So we can obtain vertical edge strength and horizontal edge strength with two AC coefficients. From the ratio of these coefficients, therefore, we may roughly know the edge orientation of the block.

Currently, many researchers have been study on the extraction of edge orientation directly in DCT domain [4][5]. To provide high accuracy, these methods use various kinds of AC coefficients. Thereby they have a bit complexity. However, because we need just three types of orientation which are horizontal, vertical and diagonal, although we use only two coefficients, we may obtain satisfied results.

Before classifying the edge orientation of the given block, we determine whether the block contain an edge or not with its variance. By Parseval theorem, the variance of the block in the spatial domain can be expressed in DCT domain identically with the squared average of AC coefficients as below,

$$\sigma^2 = \frac{1}{N^2} \sum_{u=0}^{7} \sum_{v=0}^{7} X(u,v)^2 \qquad (u,v) \neq (0,0) \qquad (4)$$

Instead of using the squared summation, it is possible to use the following relationship with less computation.

$$\sigma^2 \cong \sum_{u=0}^{7}\sum_{v=0}^{7} | X(u,v) | \qquad (u,v) \neq (0,0) \qquad (5)$$

If the variance of the given block is greater than certain threshold value, it is possible to think it as high activity region. It means that the block contains edges. In that case classification of the edge orientation is performed as following way.

Let R1 and R2 be the ratio of AC_{01} and AC_{10} as below.

$$R1 = \frac{| AC_{01} |}{| AC_{10} |} \qquad R2 = \frac{| AC_{10} |}{| AC_{01} |} \qquad (6)$$

Then, as shown in Figure 4, by the value of R1 and R2, the edge orientation of the given block is classified. In case of vertical dominant edge, R1 is used and the other case uses R2.

As mentioned previous, because the coefficient AC_{01} denotes the edge strength in vertical direction and AC_{10} does it in horizontal direction, if $| AC_{01} | > | AC_{10} |$, the block has the vertical dominant edge, otherwise it has the horizontal dominant edge. In case of the vertical dominant edge, the value of R1 belongs to the region of 1, 3-A and 4-A in Figure 4. If it satisfies Condition 1, it is the block having 90-degrees orientation edge and belongs to the region 1. Condition 2 is the case of 45-degrees orientation and 135-degrees orientation has to satisfy Condition 3.

Condition 1: R1 >Threshold_1

Condition 2: $AC_{01} \times AC_{10} > 0$, R1 <Threshold_1

Condition 3: $AC_{01} \times AC_{10} < 0$, R1 <Threshold_1

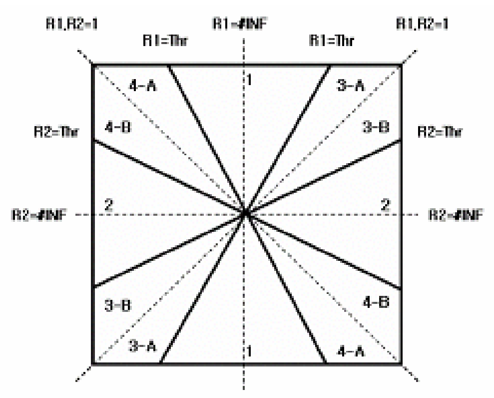

Fig. 4 Direction of several edges

Similarly, in case of the horizontal dominant edge, that is $| AC_{01} | < | AC_{10} |$, the value of R2 belongs to the region of 2, 3-B and 4-B. Therefore, according to following conditions, 0-degree, 45-degrees and 135-degrees are classified, respectively.

Condition 1': R2 >Threshold_1

Condition 2': $AC_{01} \times AC_{10} > 0$, R2 <Threshold_1

Condition 3': $AC_{01} \times AC_{10} < 0$, R2 <Threshold_1

We determine non-directional edge among 5 edge histograms by edge strength. In case of vertical dominant, satisfying the Condition 4, the block is classified as non-directional edge. And if the horizontal dominant edge block satisfies the Condition 5, it is classified as that.

Condition 4: $| AC_{10} |$ >Threshold_2

Condition 5: $| AC_{01} |$ >Threshold_2

Here we applied empirical threshold values. Figure 5 shows that the block diagram of the edge orientation decision algorithm used in this paper.

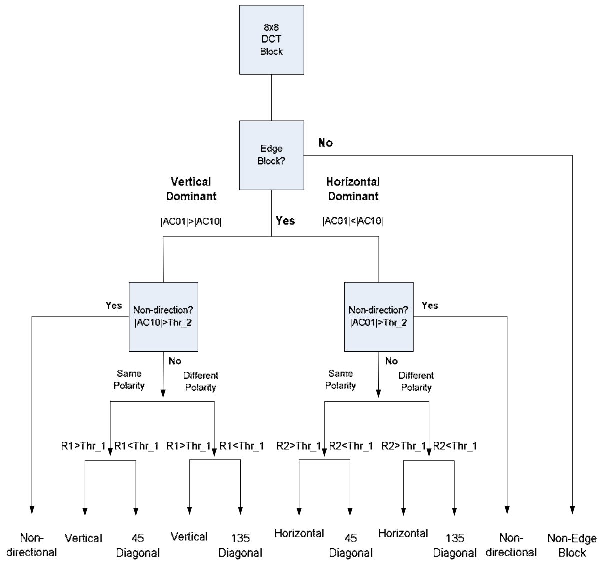

Fig. 5 Block Diagram of Edge Orientation Detection

C. EHDiD and Similarity Matching

To generate the histogram, edges in the sub-images are categorized into five types; vertical, horizontal, 45 degrees diagonal, 135 degrees diagonal and non-directional edges as a basic unit of 8×8 image-block. Since there are 16 sub-images, a total of 5×16 =80 histogram bins are generated. For similarity matching, we use the intact matching method of EHD as equation (7).

$$D(A,B) = \sum_{i=0}^{79} | h_A(i) - h_B(i) | + 5\sum_{i=0}^{4} | h_A^g(i) - h_B^g(i) | + \sum_{i=0}^{64} | h_A^S(i) - h_B^S(i) | \qquad (7)$$

where $h_A(i)$ and $h_B(i)$ represent the normalized histogram bin values of image A and B, respectively. $h_A^g(i)$ and $h_B^g(i)$ represent the normalized bin values for the global edge histograms of image A and B respectively, which are obtained from the corresponding local histograms $h_A(i)$ and $h_B(i)$. Similarly, $h_A^S(i)$ and $h_B^S(i)$ represent the histogram bin values for the semiglobal-edge histograms of image A and image B, respectively.

III. Experimental Results

Experiments were conducted using an image database containing 3022 JPEG natural images and the Threshold_1 is set by 3 and the Threshold_2 is 100. Six images selected in the database were used to make query images and ground-truth images are categorized by image classes based on the query images which were building, car, cell, eagle, elephant and penguin as shown in Figure 6.

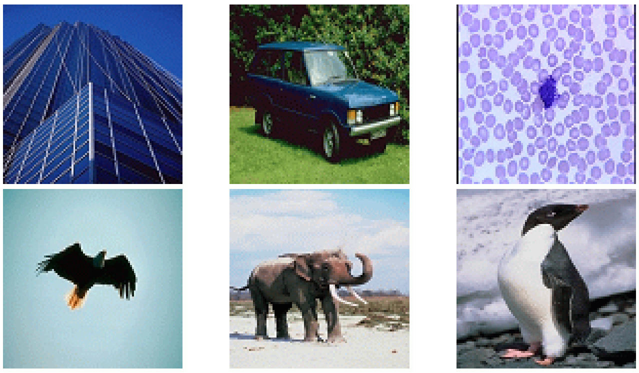

Fig. 6 Query images

As a criterion of evaluating performance, we used NMRR (Normalized Modified Retrieval Rank) and ANMRR (Average NMRR) based on the ground-truth size [3], which were the special metrics for calculating a quantitative evaluation of descriptors in terms of the retrieval rank and compared the performance of the proposed method (EHDiD) with EHD defined by the MPEG7 standard. The results are shown in Table II and Figure 7.

TABLE II
COMPARISON OF NMRR AND ANMRR BETWEEN EHD AND EHDiD

		EHD	EHDiD
NMRR	query1	0.7324561404	0.6315789474
	query2	0.8135451505	0.5911371237
	query3	0.2668495298	0.4868077325
	query4	0.1331360947	0.1528599606
	query5	0.7835497835	0.6594516595
	query6	0.670212766	0.6773049645
ANMRR		0.5466544966	0.5331900647

The experimental results demonstrate the efficiency of the proposed edge histogram generation algorithm in DCT domain. Table 2 shows performance comparisons between EHD and EHDiD via the NMRR and ANMRR and Figure 7 shows that the complexity comparisons of the two methods. In terms of NMRR and ANMRR, the retrieval performance of EHDiD is slightly superior to EHD or both of a sort. For processing time, however, the method using EHDiD is about 581 times faster than the method using EHD in terms of multiplication and about 8 times faster for addition.

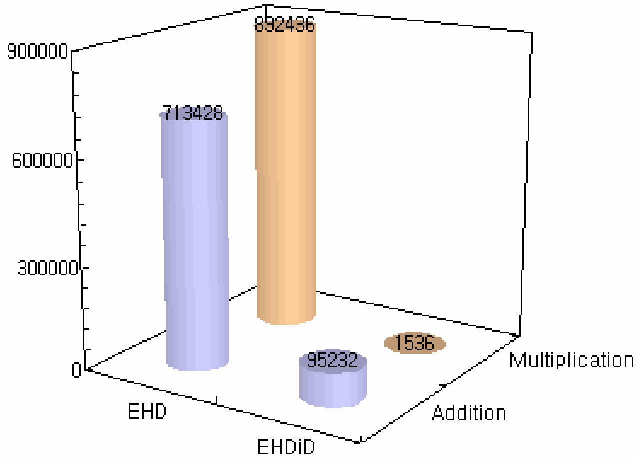

Fig. 7 Comparison of complexity between EHD and EHDiD

IV. CONCLUSION

In this paper, we proposed a fast edge histogram generation method in DCT domain using the properties of AC coefficients. Firstly, we verified the meaning of two AC coefficients. Secondly, we measured edge orientation using the ratio of the two AC coefficients. Finally, the edge histogram for retrieval is generated similarly to the EHD defined by MPEG7 standard. The proposed method is comparable to the EHD in complexity, and is found to be about 581 times faster in terms of multiplication with the retrieval performance slightly superior.

ACKNOWLEDGMENT

This work was supported by grant No. R01-2003-000-10412-0 from the Basic Research Program of the Korea Science & Engineering Foundation

REFERENCES

[1] K. R. Rao, J. J. Hwang, " Techniques and Standards for image, Video, and Audio Coding, " Prentice Hall PTR, 1996.
[2] Overview of the MPEG-7 standard, December 2001. ISO/IEC/TC/SC29/WG11 N3914
[3] B.S. Manjunath, Philippe Salembier, Thomas Sikora, "Introduction to MPEG-7", JOHN WILLEY & SONS, LTD , pp. 183-184, 2002
[4] Bo Shen, Ishwar K. Sethi, 'Direct feature extraction from compressed images', SPIE vol. 2670, Storage & Retrieval for Image and Video Databases , 1996.
[5] S. W. Lee, Y.M. Kim and S. W. Choi, 'Fast scene change detection using direct feature extraction form MPEG compressed videos', IEEE Trans. On Multimedia, Vol. 2, No. 4, December 2000