# Shape-based Image Retrieval using VQ based Local Differential Invariants

Hyun-Sool Kim, Dae-Kyu Shin, Tae-Yun Chung and Sang-Hui Park

**Abstract** - In this study, for the shape-based image retrieval, a method using local differential invariants is proposed. This method calculates the differential invariant feature vector at every feature point extracted by Harris corner point detector. Then through vector quantization using LBG algorithm, all feature vectors are represented by a codebook index. All images are indexed by the histogram of codebook index, and by comparing the histograms the similarity between images is obtained. The proposed method is compared with the existing method by performing experiments for image database including various 1100 trademarks.

**Keywords** - content-based image retrieval, Harris corner point detector, differential invariants, vector quantization

## 1. Introduction

Recently many materials are being digitalized and made into multimedia data due to the rapid development of multimedia technology. Also through the spread of scanner and digital camera and the development of the large capacity of storage equipment, numerous multimedia contents are being constructed. The medium that plays an important role in multimedia is image data and many image databases are constructed and used in various fields. In these image databases, it is a very difficult work to search a desired image by manual work. So the need for the system of the automatic image retrieval from image database is urgent and the problems of image retrieval by content are becoming widely recognized and studied actively.

The early systems perform image retrieval by using keywords about an image as a feature. But these systems have many problems in indexing process and limitations in retrieval capacity. Problems with traditional methods of image indexing have led to the rise of interest in techniques for retrieving images on the basis of automatically-derived features such as color, texture and shape - a technology now generally referred to as Content-Based Image Retrieval(CBIR). In contrast to the text-based approach, CBIR operates on a totally different principle, retrieving stored images from a collection by comparing features automatically extracted from the images themselves. The commonest features used are mathematical measures of color, texture or shape[1-2].

Manuscript received: May. 29, 2001 accepted: Jan. 3, 2002.

Hyun-Sool Kim is with Mobile Communication Division, Samsung Electronics Co., Korea.
Dae-Kyu Shin and Sang-Hui Park is with Dept. of Electrical and Electronics Engineering, Yonsei Univ., Seoul, Korea.
Tae-Yun Chung is with Dept. of Control and Instrumentation Engineering, Kangnung National University, Korea.

In this study, we propose new retrieval methods using local shape information. In shape-based retrieval, local differential invariant feature vector obtained at every feature point is used as shape information. All feature vectors are represented with codebook index by vector quantization, and by comparing the codebook index histograms of images, the similarity between images is measured.

## 2. Image Retrieval Using Local Differential Invariant Vector

To extract the shape information, the methods such as chain code, Zernike moment, Fourier descriptor and invariant moments, etc. have been studied[1].

In this study, we use differential invariants as a shape feature vector at the automatically detected feature points and then apply vector quantization to all feature vectors for the efficient description and searching performance.

### 2.1 Extraction of Feature Points

Interest points are the points where the signal varies 2 dimensionally, and correspond to the corners, T-junctions, and also the points which texture changes rapidly at. Of the existing corner detectors, the Harris detector[3] has been known as the most efficient method considering repeatability and information content. Harris detector constructs the matrix related to the autocorrelation function in order to determine locations where the signal changes in two directions and then extracts the interest points by means of the eigenvalues of this matrix which become the principal curvatures of autocorrelation function.

The change, E, of the image intensity, I, for the small shifts $(x, y)$ can be written as

$$\begin{aligned} E(x,y) &= \sum_u\sum_v W_{u,v}(I_{u+x,v+y} - I_{u,v})^2 \\ &= \sum_u\sum_v W_{u,v}(xX + yY + O(x^2, y^2))^2 \\ &= Ax^2 + 2Cxy + By^2 = (x,y)\,M(x,y)^T \\ M &= \begin{bmatrix} A & C \\ C & B \end{bmatrix} \end{aligned}$$

$$A = X^2 * W, \quad B = Y^2 * W, C = (XY) * W \tag{1}$$

where, $W_{u,v} = \exp -(u^2 + v^2)/2\sigma^2$
$X = I * (-1, 0, 1) = \partial I/\partial x$
$Y = I * (-1, 0, 1)^T = \partial I/\partial y$
$*$ : *convolution*
$u, v$ : window area
$x, y$ : shift

Let $\alpha$ and $\beta$ be the eigenvalues of M. And $\alpha$ and $\beta$ will be proportional to the principal curvatures of the local autocorrelation function, so about the regions of an image, there are three cases to be considered. If both of these two eigenvalues are small, the windowed region is of approximately same pixel value, and if one is large and the other is small, the windowed region corresponds to the edge component. And if both of them are large, this indicates corner point.

Eigenvalues of this matrix have the following properties, so corner point region necessary for the experiment can be identified by using matrix component not by calculating eigenvalues directly.

$$\begin{aligned} Tr(M) &= \alpha + \beta = A + B \\ Det(M) &= \alpha\beta = AB - C^2 \\ R &= Det - k\,Tr^2 \end{aligned} \tag{2}$$

Where, $Det$ and $Tr$ indicate determinant and trace of matrix, respectively. Response function R is composed of A, B, C and the value of R is positive in the corner region, negative in the edge region, and small in the flat region.

## 2.2 Differential Invariants

Locally windowed image centered at feature points can be described by a differential invariant vector. Differential invariant vector is described by the set of differential values, named Local Jet[4]. If locally windowed image is $I$, and given scale is $\sigma$, then Nth-order local jet at arbitrary one point $x = (x_1, x_2)$ is

$$J^N[I](\boldsymbol{x}, \sigma) = \{ L_{i_1 i_2 \ldots i_n}(\boldsymbol{x}, \sigma) \mid (\boldsymbol{x}, \sigma) \in I \times R^+ ; n = 0, 1, 2, \ldots, N \} \tag{3}$$

$L$ is the component of convolution of locally windowed image $I$ and Gaussian function $G(x, \sigma)$. Subscript indicates the partial differential of that variable. As differentiation commutes with convolution, the following relation is formed.

$$L_i = \partial_i(I * G) = I * \partial_i G = \partial_i I * G \tag{4}$$

In equation (4), a simply way to stabilize the calculation of the derivatives $\partial_i(I * G)$ is therefore to perform convolution of a locally windowed image $I$ and the derivative of Gaussian function $\partial_i G$.

By using Einstein's addition form and rectangular coordinate, differential invariant vector $DI$ can be described like equation (5)[4].

$$DI[0,\ldots,4] = \begin{bmatrix} L \\ L_i L_i \\ L_i L_{ij} L_j \\ L_{ii} \\ L_{ij} L_{ji} \end{bmatrix} = \begin{bmatrix} L \\ L_x L_x + L_y L_y \\ L_{xx} L_x L_x + 2L_{xy} L_x L_y + L_{yy} L_y L_y \\ L_{xx} + L_{yy} \\ L_{xx} L_{xx} + 2L_{xy} L_{yx} + L_{yy} L_{yy} \end{bmatrix}$$

$$DI[5,\ldots,8] = \begin{bmatrix} \varepsilon_{ij}(L_{jkl} L_i L_k L_l - L_{jkk} L_i L_l L_l) \\ (L_{iij} L_j L_k L_k - L_{ijk} L_i L_j L_k) \\ -\varepsilon_{ij} L_{jkl} L_i L_k L_l \\ L_{ijk} L_i L_j L_k \end{bmatrix}$$

$$\text{Where, } \varepsilon_{12} = -\varepsilon_{21}, \quad \varepsilon_{11} = -\varepsilon_{22} = 0 \tag{5}$$

Subscript $i$ in the Einstein's addition form indicates the summation of the partial derivatives of the variables.

$$\begin{aligned} L_i &= \sum_i L_i = L_x + L_y \\ L_{ij} &= \sum_i\sum_j L_{ij} = L_{xx} + L_{xy} + L_{yx} + L_{yy} \end{aligned} \tag{6}$$

## 2.3 Retrieval Procedure

Vector Quantization is performed using LBG algorithm to the differential invariant vectors calculated on feature points of all the images in the database to describe the shape information of an image. After forming a codebook which efficiently describes all the feature vectors, feature vectors on all the feature points of an image are represented by the index of the codebook and forms the histogram about the codebook index. All images are represented by histogram describing shape information, and then shape-based image retrieval is performed by comparing the histograms between a query image and database images. To compare the histograms, we use histogram intersection method used by Swain in color-based image retrieval[5]. That is, if the feature vectors of two images $I_i, I_j$ are represented by the N-level codebook, shape information of two images can be described by the histogram $H_i[k], H_j[k]$, $k=0,1,\cdots,N-1$ and comparison of these two images can be performed by the equation (7).

$$HI(H_i, H_j) = \frac{\sum_{k=0}^{N-1} \min(H_i[k], H_j[k])}{\sum_{k=0}^{N-1} H_j[k]} \tag{7}$$

$HI(H_i, H_j)$ has a value ranged from 0 to 1 and the larger the similarity of two images is, the nearer value to 1 this has. Shape-based retrieval is performed using this measure.

## 3. Experiments and Results

### 3.1 Experiments

In this study, we performed shape-based image retrieval tests using 20 query images for 128 × 128 1280 trade-mark image database in various situations like rotation, scale change, brightness change and noise addition. Also experiment using color and shape information simultaneously was performed. In vector quantization procedure, codebook size was set to 50. And we compared the performance of the proposed method with the existing methods using Zernike moments(ZM), differential invariants(DI) and Jain's one using gradient[2]. As a criterion of evaluating performance, we used the average retrieval rank of the desired image like equation (8).

$$average\ retrieval\ rank = \frac{1}{M}\sum_{i=1}^{M} rank(i)$$

$$\%\ rank = \frac{average\ retrieval\ rank}{N} \times 100 \tag{8}$$

Where, $N$ and $M$ are the number of database images and the number of query images respectively.

### 3.2 Results

#### 3.2.1 Shape-based Retrieval

For a query image with elliptical shape, the retrieval results of shape-based retrieval methods are shown in Fig. 1.


**Fig. 1** Result of shape-based retrieval

#### 3.2.2 Retrieval for the queries with noise

We performed image retrieval for queries whose 5%, 10%, and 15% pixels are contaminated by white Gaussian noise. Results are shown in Table 1.

**Table 1** Retrieval results for noise added queries

| Noise | Jain | Zernike | DI | Proposed |
|---|---|---|---|---|
| 5% | 2.10(0.19%) | 1.25(0.12%) | 3.60(0.32%) | 3.80(0.34%) |
| 10% | 2.55(0.23%) | 10.55(0.95%) | 4.30(0.39%) | 3.63(0.33%) |
| 15% | 4.20(0.38%) | 41.35(3.73%) | 13.00(1.17%) | 7.11(0.37%) |

This result shows that the proposed method is more stable to the noise addition than other methods using Zernike moments and differential invariants. It is due to the robustness of VQ to noise effect.

#### 3.2.3 Retrieval for the queries with brightness change

We performed image retrieval for queries whose pixel values are changed by +10, -10. Results are shown in Table 2.

**Table 2** Retrieval results for brightness changed queries

| Brightness | Jain | Zernike | DI | Proposed |
|---|---|---|---|---|
| +10 | 9.80(0.88%) | 1.00(0.09%) | 7.11(0.64%) | 4.15(0.37%) |
| -10 | 1.00(0.09%) | 1.00(0.09%) | 4.65(0.36%) | 1.3(0.12%) |

The proposed method shows the good performance no less than other methods in brightness change.

#### 3.2.4 Retrieval for the queries with rotation

We performed image retrieval for queries rotated by 30, 45, and 90 degrees. In this experiment, Jain's method is not applicable. Results are shown in Table 3.

**Table 3** Retrieval result for rotated queries

| Rotation | Jain | Zernike | DI | Proposed |
|---|---|---|---|---|
| 30 ° | X | 13.63(1.22%) | 46.37(4.18%) | 12.42(1.11%) |
| 45 ° | X | 13.32(1.2%) | 41.79(3.77%) | 4.42(0.40%) |
| 90 ° | X | 1.00(0.09%) | 1.20(0.11%) | 1.05(0.09%) |

#### 3.2.5 Retrieval for the queries with scale change

We performed image retrieval for queries whose scale are changed by 0.9 and 1.1 times. Results are shown in Table 4. This result also shows the better performance of the proposed method than those of other methods. From this, we can see that VQ based feature is very effective for image retrieval under the shape change by scale.

**Table 4** Retrieval result for scale changed queries

| Scale | Jain | Zernike | DI | Proposed |
|---|---|---|---|---|
| 0.9 | 24.7(2.23%) | 22.80(2.01%) | 54.20(4.89%) | 5.85(0.53%) |
| 1.1 | 13.69(1.23%) | 20.40(1.84%) | 57.63(5.19%) | 7.19(0.65%) |

#### 3.2.6 Retrieval by a partial image

We performed the retrieval test for the parts of an image. A query image is translated to cut off some part of an image. The retrieval result is shown in Fig. 2.


(a) Query

(b) Jain (c) Zernike (d) DI (e)Proposed

**Fig. 2** Retrieval result by a partial image

The good retrieval performance of the proposed method indicates that the elements of histogram bin are reduced but the global shape of histogram is maintained. Also from this result we can infer that only a few points are necessary to recognize an image.

#### 3.2.7 Comparisons of required memory size and retrieval time

In retrieval system, two important factors - required memory size and retrieval time - are compared as an evaluating criterion for an efficient retrieval system. We used 5-order moments for Zernike method, and for differential invariants, the number of feature points for an image was restricted to maximum 100 and 5-dimensional vector at every feature point was used as a shape feature. The required memory size and retrieval time of each retrieval method are shown in Table 5 and Fig. 3. In the case of comparison of required memory size, we can see that the proposed method needs relatively small memory size to save features. In comparison of retrieval time, the proposed method shows the more efficient retrieval speed than DI method.

**Table 5** Comparison of memory size required for saving shape features

| Method | Jain | Zernike | DI | Proposed |
|---|---|---|---|---|
| Memory(Byte) | 720 | 60 | 1611 | 200 |

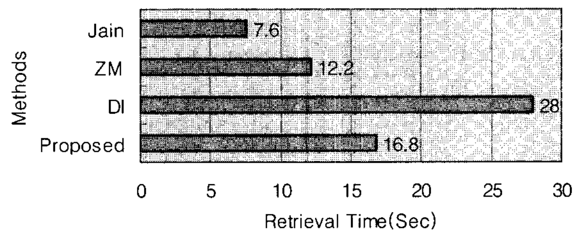



**Fig. 3** Comparison of shape-based retrieval time

## 4. Conclusion

In this study, we proposed new shape-based image retrieval method which uses the VQ-based local differential invariant feature vector. This method detects corner points by Harris detector, calculates differential invariant feature vector at every corner point and applies vector quantization to all feature vectors. It was shown that this method retrieves the desired image better and faster than other methods in the case of noise addition, scale change, rotation and brightness change in a query image. Also because the proposed method uses a local feature, it is possible to retrieve the wanted image using a partial image as a query.

## References

[1] J. Eakins and M. Graham, "Content-based Image Retrieval", JISC Technology Applications Report: 39, Oct., 1999

[2] A.K. Jain and A. Vailaya, "Image Retrieval using Color and Shape", Pattern Recognition, Vol. 29, No. 8, pp. 1233-1244, 1996

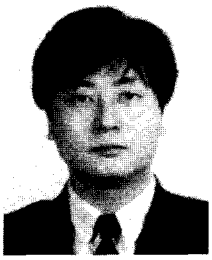
[3] C. Harris and M. Stephens, "A combined corner and edge detector", Alvey Vision Conference, pp. 147-151, 1988

[4] C. Schmid and R. Mohr. "Local Grayvalue Invariants for Image Retrieval," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 5, pp. 530-535, 1997

[5] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, Vol. 7, No. 1, pp. 11-32, 1991


**Hyun-Sool Kim** was born in Korea, in 1971. He received the B.S., M.S. and Ph.D degrees in Electrical Engineering from Yonsei University, Seoul, Korea, 1994, 1996, and 2001, respectively. Since 2001, he has been working in Samsung Electronics Co., where he is a senior engineer at mobile communication division. His research interests include video processing, wavelet analysis, signal processing, image processing and pattern recognition.

**Dae-Kyu Shin** was born in Korea, in 1972. He received the B. S. and M. S. degree in Electrical Engineering from Yonsei University, Seoul, Korea, in 1995 and 1997, respectively. He is presently working toward the Ph.D. degree at the Department of Electrical and Electronics Engineering, Yonsei University. His research interests include the digital signal processing, the image retrieval and the image recognition.



**Tae-Yun Chung** received the B.S., M.S., and Ph.D degrees in electrical engineering from Yonsei University, Seoul, Korea, in 1987, 1989, and 2000, respectively. Since 1989, he has been involved in research and development of video coding system as a senior engineer in Samsung Electronics, Korea. And since 2001, he has been with Kangnung National University, Korea, where he is currently a assistant professor in the department of control and instrumentation engineering. His research interests are video coding, multimedia signal processing and contents copyright protection.

**Sang-Hui Park** received the B.S., M.S., and Ph.D. degrees in Electrical Engineering from Yonsei University, Seoul, Korea, in 1962, 1964 and 1971, respectively. From 1964 to 1969, he was an Assistant Professor at Soodo Engineering College. Since 1970, he has been with Yonsei University, Korea, where he is currently a Professor at the Department of the Electrical and Electronics Engineering. His research interests are currently image and multimedia signal processing, neural network and biomedical engineering.