# TRADEMARK IMAGE RETRIEVAL USING HIERARCHICAL REGION FEATURE DESCRIPTION

*Feng Liu*[1], *Bin Wang*[1,*], *Fanqing Zeng*[2]

[1]School of Information Engineering, Nanjing University of Finance & Economics, Nanjing 210023, China
[2]Experimental Teaching Center of Economics & Management, Nanjing University of Finance & Economics, Nanjing 210023, China

## ABSTRACT

A novel trademark image retrieval method is proposed in this paper. In this method, the image region of the trademark is iteratively partitioned into progressively smaller one along various directions and four region measurements are then conducted for features extraction against the partitioned regions. The merit of the resulting descriptors is that the geometrical features of different directions are finely captured in a hierarchical manner. A shifting feature matching scheme effectively guarantees the matching invariant to the rotation of trademark images. The experiments results demonstrate that the proposed method outperforms the state-of-the-art approaches for the trademark image retrieval.

***Index Terms—*** Trademark image retrieval, region partitioning , region measurement, feature matching scheme

## 1. INTRODUCTION

With the rapid increase of the registered trademarks around the world, developing an effective trademark image retrieval (TIR) system to guarantee to exclude those similar trademarks from the existing registered trademark datasets is becoming increasingly urgent. Traditional classification of trademark images employ manually assigned codes based on their shape features and types of object depicted. Due to the different subjective perception of the trademark images, faults or slips may appear in the classification [1]. In this paper, we focus on developing an effective solution for TIR.

Trademarks can be categorised into a few different types: a word-only mark, a device-only mark, or a device-and-word mark. For a word-only mark, the design of the trademark consists purely of text words or phrases and the traditional character recognition techniques can be applied to it. While, the device mark contains graphical or figurative shapes. The composite mark is a combination of the former two types [2]. Some TIR systems are solely designed to accommodate one of the types. There are also several existing TIR systems which have been designed to handle all kinds of trademark images, however their performance are rather unfavourable when compared to those systems that are specifically designed to handle only one kind of trademark [3].

Two important issues are associated with the trademark image retrieval. One is how to extract appropriate features to effectively represent the visual content of the trademark images, and the other is how to measure the dissimilarity between any two given trademark images based on their descriptors. In this work, we treat the trademark as a general binary image and extract the regional features of the object for description. The contribution of our work is: A strategy which iteratively partitioning the region of the object into smaller parts along different directions is proposed for providing a hierarchical description and a shifting feature matching scheme is presented for a finely dissimilarity measure. The proposed method outperforms the state-of-the-art approaches which are used for trademark image retrieval in terms of the experimental results on the well-known MPEG-7 CE-1 and CE-2 tests.

## 2. RELATED WORK

Many methods have been proposed for trademark image retrieval (TIR). Zernike moments (ZM) [6] is a classical shape descriptor. It projects the shape image function on a set of orthogonal basis functions defined in the interior of a unit circle and the magnitudes of moments are used to form a numeric feature vector for description. To overcome the limitation of expensively calculation of the kernels of ZM, Yap et.al. proposed a new set of polar harmonic transforms (PHT) [7] for extracting rotation invariant features based on a set of orthogonal projection bases. The computation of its kernels is much simpler than that of ZM. Shape contexts [8] is a contour based descriptor in which each contour point (including the inner contour point) is taken as a reference point and the

*Corresponding author (Email: wangbin@njue.edu.cn). This work was partially supported by the National Natural Science Foundation of China under Grant 61372158, the Natural Science Foundation of Jiangsu Province under Grant BK20141487, the "333" Foundation for high level talents of Jiangsu Province under Grant BRA2015351, and the Policy guidance program (Cooperation of Industry, Education and Academy) of Jiangsu Province under Grant BY2016009-03.

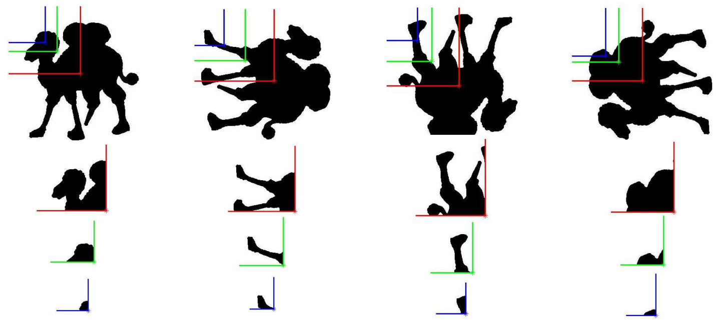

**Fig. 1**. A visual illustration of iteratively partitioning the image region in progressively smaller one along different directions.

distribution of contour points relative to it is captured for description.

There are also several remarkable TIR systems that have been developed in recent years. Wei et al. [3] proposed a two-component solution, in which Zernike moments, centroid distance, contour curvature were integrated for description, and a two-component strategy was applied in feature matching. Anuar et al. [9] proposed another TIR technique in which the Zernike moment was used as a global descriptor and the edge gradient co-occurrence matrix was used as the local descriptor, and their combination was taken as the description of the trademark image. Qi et al. [10] addressed the problem of TIR using a synthetic shape description of integrating contour-based feature and region-based feature, and a statistics-based feature matching strategy.

The closely relevant techniques to our work in this paper are the hierarchical decomposition techniques. Yang et al. [4] proposed to recursively decompose the binary image into progressively smaller regions, at each iteration step, the partitioning against the current region was determined by the estimation from the mean points. After a fixed number of iteration, the set of mean points, or centroids were taken as the description. Sidiropoulos et al. [5] proposed a similar methodology for splitting the image plane into regions. However they adopted a totally different way for the construction of the descriptor. Instead of using the set of mean points, this method used the estimation of point density histograms of image regions for description.

## 3. THE PROPOSED METHOD

In this section, we propose a novel description and matching method for trademark image retrieval.

### 3.1. Description Framework

A binary trademark image can be considered as a distribution of black pixels in a white two-dimensional space-background. Let $B_0$ be the set of the coordinate pair $(x, y)$ of all the black pixels. It represents the image region which the object pixels distribute over. We define its centroid coordinate pair $(\overline{x}_0, \overline{y}_0)$ as

$$\overline{x}_0 = \frac{\sum\limits_{(x,y)\in B_0} x}{N_0}, \quad \overline{y}_0 = \frac{\sum\limits_{(x,y)\in B_0} y}{N_0}, \tag{1}$$

where $N_0$ is the cardinality of the set $B_0$. Using $(\overline{x}_0, \overline{y}_0)$, we partition the set $B_0$ and obtain a subset $B_1$ of the set $B_0$ defined by $B_1 = \{(x, y)|x <= \overline{x}_0, y >= \overline{y}_0\}$. The subset $B_1$ denotes a subregion of the image region which the object pixels distributed over. We can calculate the centroid coordinate pair $(\overline{x}_1, \overline{y}_1)$ of the subset $B_1$ by Eq. 1 and use it to further partition the subset $B_1$ in the same way and obtain a subset $B_2$ of the set $B_1$. We continue the above partitions and obtain $l$ sets: $B_0 \supset B_1 \supset \ldots \supset B_l$, where $l$ is the pre-specified number of the iterations.

From the above process, we iteratively partitioned the image region into progressively smaller regions along a given direction. To capture the inner geometrical structure of the trademark image along various directions, we rotate the image region $B_0$ around its centroid point $(\overline{x}_0, \overline{y}_0)$ by an angle $\theta \in [0, 2\pi)$ and obtain a new set $B_0^\theta$. We repeat the same iterative partition against the set $B_0^\theta$ and obtain its $l$ subsets $B_0^\theta \supset B_1^\theta \supset \ldots \supset B_l^\theta$. We uniformly sample the angle range $[0, 2\pi)$ into $m$ angles and let $j = 0, 1, \ldots, m-1$ denotes the index of the sampled angle. Then along $m$ directions, we can iteratively partition the image region into progressively small regions and obtain $m * l$ sets $B_i^j : i = 1, 2, \cdots, l$ and $j = 0, 1, \cdots, m-1$. A visual illustration of the above iteratively partitioning is shown in Fig. 1.

It is worth noting that the proposed hierarchical description framework is completely different to the methods in [4, 5] as follows: (1) It decomposes the image region considering various directions while the methods in [4, 5] do it only in a single direction which make the proposed framework more insensitive to the rotation of the image than the methods in [4, 5], (2) the proposed framework splits the image region in an iterative manner, in each iteration, only one subregion is considered for the next partition, while the methods in [4, 5] split the image region in a recursive manner, and in each iteration, four subregions are considered for the next partition, so the proposed method has the lower computational cost than the methods in [4, 5].

### 3.2. Image Region Measure

In this section, we measure each image region $B_i^j, i = 1, 2, \cdots, l$ and $j = 0, 1, \ldots, m-1$ as follows:

**Density**: The density of the region $B_i^j$ is defined as

$$d_i^j = \frac{|B_i^j|}{|B_{i-1}^j|}, \tag{2}$$

where $|*|$ denotes the cardinality of the set.

**Compactness**: The compactness of region $B_i^j$ is defined as

$$c_i^j = \frac{|B_i^j|}{\pi h^2}, \tag{3}$$

where $h$ is the radius of the smallest circle centered at the point $(\overline{x}_i, \overline{y}_i)$ of enclosing the image region $B_i^j$ and can be calculated by

$$h = max\sqrt{(x-\overline{x}_i)^2 + (y-\overline{y}_i)^2}. \tag{4}$$

**Rectangularity**: Rectangularity is a measurement that represent how well a region fits its minimum enclosing rectangle(MER). The standard approach for estimating rectangularity of region $B_i^j$ is defined as

$$r_i^j = \frac{|B_i^j|}{(max(y) - min(y))(max(x) - min(x))}. \tag{5}$$

**Eccentricity**: Eccentricity has been widely used as a shape feature. It is a region based parameter and illustrates how the region points are scattered around the centre of the region [11].

The central moments of region $B_i^j$ is defined as

$$\mu_{p,q} = \sum_x \sum_y (x-\overline{x}_i)^p (y-\overline{y}_i)^q, \tag{6}$$

where $(x, y) \in B_i^j$ and $(\overline{x}_i, \overline{y}_i)$ is the centroid of $B_i^j$.

The principal moments of a region are eigenvalues of the matrix

$$\begin{bmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{bmatrix},$$

and the eccentricity of the region $B_i^j$ is

$$e_{cc} = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}, \tag{7}$$

where $\lambda_{max}$ and $\lambda_{min}$ are the eigenvalues of the matrix. This feature depends solely on the shape, not on size and orientation. Since eccentricity is greater than 1 , here we use the reciprocal value of eccentricity $e_i^j$ to measure the region, that is

$$e_i^j = \frac{1}{e_{cc}}. \tag{8}$$

### 3.3. Feature Matching

For an image region $B_i^j$, we obtain its four measurements $d_i^j$, $c_i^j, r_i^j, e_i^j$. By varying the index $j$ from 0 to $m-1$ and varying the index $i$ from 1 to $l$, we obtain four matrices $D = (d_i^j)_{m\times l}$, $C = (c_i^j)_{m\times l}$, $R = (r_i^j)_{m\times l}$ and $E = (e_i^j)_{m\times l}$. Consider that various types of features may have different contribution to the dissimilarity measure of trademark images, we combine the four matrices and assign each of them a different weighing factor, that is

$$F_0 = \begin{bmatrix} W_d \times D & W_c \times C & W_r \times R & W_e \times E \end{bmatrix}, \tag{9}$$

which is a $m * 4l$ feature matrix. We consider each row of the matrix $F_0$ as a vector $V_i, i = 0, 1, \ldots, m-1$, the matrix $F_0$ can be then denoted by a column vector as

$$F_0 = \begin{bmatrix} V_0 \\ V_1 \\ V_2 \\ \vdots \\ V_{m-1} \end{bmatrix}. \tag{10}$$

Since the rotation of the trademark image will result a circular shift of the column vector $F_0$, for a query image A, we should prepare the feature matrix $F_0$ and its $m-1$ shifting versions

$$F_1 = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_{m-1} \\ V_0 \end{bmatrix}, F_2 = \begin{bmatrix} V_2 \\ \vdots \\ V_{m-1} \\ V_0 \\ V_1 \end{bmatrix}, \cdots, F_{m-1} = \begin{bmatrix} V_{m-1} \\ V_0 \\ V_1 \\ \vdots \\ V_{m-2} \end{bmatrix}. \tag{11}$$

Then we measure the dissimilarity between the query trademark image A and a database trademark image B by

$$dis(A, B) = \min_{j=0,1,\cdots,m-1} \|F_j^A - F_0^B\|, \tag{12}$$

where $\|*\|$ represents $L$-1 distance.

## 4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed method, we conduct it on two well-known image databases, MPEG-7 CE-2 region shape database and MPEG-7 CE-1 contour shape database, and compare it with the five state-of-the-art approaches: Adaptive hierarchical density histogram (AHDH) [5], Zernike Moments [6], Polar harmonic transforms (PHT) [7], Shape contexts [8] and Zernike moment & edge gradient technique (ZMEG) [9]. In our experiments, the parameters for the image description are set to $l = 3$ and $m = 72$ and the weighting factors for the feature matching are set to $W_d = 0.4, W_c = 0.2, W_r = 0.2, W_e = 0.2$.

**Fig. 2**. Samples from MPEG-7 CE-2 database.

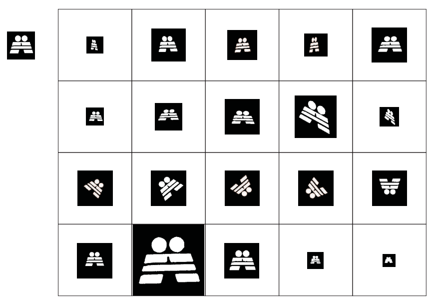

**Fig. 3**. All the samples in the group 8 of MPEG-7 CE-2 database.

The first group of experiments are performed on the MPEG-7 CE-2 shape database [12] which consists of 3621 trademark or logo images (some samples are shown in Fig. 2). Among them, there are 651 shapes which are organized into 31 groups with 21 samples in each one. Each one of the 651 shapes is taken as a query to retrieval similar shapes from the whole database. In each group, one member is regarded as a reference, ten members are its perspective transformed versions, five members are its rotated versions, and the other five members are its scaled versions. Fig. 3 shows all the members in the group 8.

We use the precision-recall curve [13] which is the s-tandard protocol for the retrieval performance evaluation in our experiments. The precision-recall curves of the proposed method and the other five state-of-the-art approaches for trademark image retrieval are plotted by in Fig. 4. It can be seen that the proposed method has the best precision-recall curve among all the competing methods. In each reference point (total 21 points) of the precision-recall curve, the proposed method achieves the higher precision score and the higher recall score over the compared methods which demonstrate the promising retrieval performance of the proposed method.

We also test the proposed method on the MPEG-7 CE-1 Part B database [14], it is widely used for evaluating those contour based methods which recognize objects only by their contours and ignoring their interior structures. Here, we take it as a test case of simple trademark images. This database

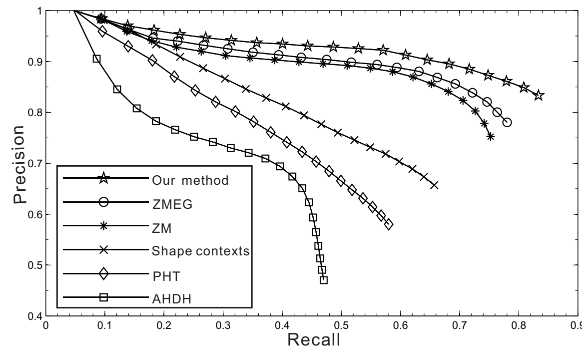


**Fig. 4**. The precision-recall curves of different methods on the MPEG-7 CE-2 database.

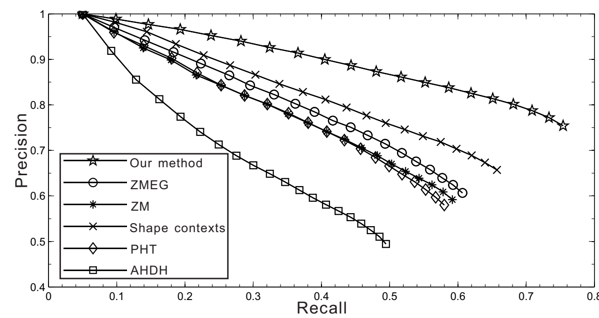


**Fig. 5**. The precision-recall curves of different methods on the MPEG-7 CE-1 database.

contains 1400 binary shapes organized into 70 groups with 20 similar shapes in each group.

We plot the precision-recall curves of the proposed method and the other five state-of-the-art approaches for trademark image retrieval in Fig. 5. It can be observed that the proposed method consistently achieves the best precision-recall curve among all the competing methods on the MPEG-7 CE-1 test.

## 5. CONCLUSION

We have presented a novel method for trademark image retrieval in this paper. A hierarchical description framework which iteratively partitions the image region into progressively smaller one along various directions is proposed and four region measurements are further conducted for features extraction against the partitioned regions. The resulting descriptors finely capture the geometrical features of different directions in a hierarchical manner. A shifting feature matching scheme guarantees the matching invariant to the rotation of trademark images. The experiments results on the well-known MPEG-7 CE-2 and MPEG-7 CE-1 image datasets demonstrate that the proposed method outperforms the state-of-the-art approaches for the trademark image retrieval.

## 6. REFERENCES

[1] S. Alwis, J. Austin, "A novel architecture for trademark image retrieval systems," in: *Proceedings of the Challenge of Image Retrieval*, British Computer Society, UK, 1998.

[2] J.P. Eakins, "Trademark Image Retrieval," *Principles of Visual Information Retrieval*, Springer, New York, pp. 319-350, 2001.

[3] C.H. Wei, Y. Li, W.Y. Chau, C.T. Li, "Trademark image retrieval using synthetic features for describing global shape and interior structure," *Pattern Recognition*, vol. 42, no. 3, pp. 386-394, 2009.

[4] M. Yang, G. Qiu, J. Huang, D. Elliman, "Near-Duplicate Image Recognition and Content-based Image Retrieval using Adaptive Hierarchical Geometric Centroids," in: *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong, China, pp. 958-961, 2006.

[5] P. Sidiropoulos, S. Vrochidis, I. Kompatsiaris, "Content-based binary image retrieval using the adaptive hierarchical density histogram," *Pattern Recognition*, vol. 44, no. 4, pp. 739-750, 2001.

[6] A. Khotanzad, Y.H. Hong, "Invariant Image Recognition in Zernike Moments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 5, pp. 250-261, May 1990.

[7] P.T. Yap, X. Jiang, A.C. Kot, "Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 7, pp. 1259-1270, July 2010.

[8] S. Belongie, J. Malik, J. Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 24, pp. 509-522, April 2002.

[9] F.M. Anuar, R. Setchi, Y.K. Lai, "Trademark image retrieval using an integrated shape descriptor," *Expert Systems with Applications*, vol. 40, no. 1, pp. 105-121, 2013.

[10] H. Qi, K. Li, Y. Shen, W. Qu, "An effective solution for trademark image retrieval by combining shape description and feature matching," *Pattern Recognition*, vol. 43, no. 43, pp. 2017-2027, 2010.

[11] F. Mokhtarian, S. Abbasi, "Matching Shapes with Self-Intersections: Application to Leaf Classification," *IEEE Transactions on Image Processing*, vol. 13, no. 5, pp. 653-61, May 2004.

[12] D. Zhang, G. Lu, "Shape-Based Image Retrieval Using Generic Fourier Descriptor," *Signal Processing: Image Communication*, vol. 17, no. 10, pp. 825-848, 2002.

[13] E.G.M. Petrakis, A. Diplaros, E. Milios, "Matching and Retrieval of Distorted and Occluded Shapes Using Dynamic Programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 11, pp. 1501-1516, November 2002.

[14] L.J. Latecki, R. Lakamper, U. Eckhardt, "Shape Descriptor for Non-rigid Shapes with a Single Closed Contour," in: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 424-429, 2000.