Trademark Image Retrieval Using Synthetic Features for Describing Global Shape and Interior Structure

Chia-Hung Wei[1], Yue Li[2], Wing Yin Chau[3], and Chang-Tsun Li[2]

[1]*Department of Information Management, Ching Yun University, Taiwan*
rogerwei@dcs.warwick.ac.uk

[2,3] *Department of Computer Science, University of Warwick, UK*
[2]*{yxl,ctli}@dcs.warwick.ac.uk,* [3]*awychau@yahoo.co.uk*

Abstract

A trademark image retrieval (TIR) system is proposed in this work to deal with the vast number of trademark images in the trademark registration system. The proposed approach commences with the extraction of edges using the Canny edge detector, performs a shape normalization procedure, and then extracts the global and local features. The global features capture the gross essence of the shapes while the local features describe the interior details of the trademarks. A two-component feature matching strategy is used to measure the similarity between the query and database images. The performance of the proposed algorithm is compared against four other algorithms.

Keywords: Content-based image retrieval, trademark image retrieval, feature extraction, Zernike moments, multimedia database

1. Introduction

With the rapid increase in the amount of registered trademark images around the world,

trademark image retrieval (TIR) has emerged to ensure that new trademarks do not repeat any of the vast number of trademark images stored in the trademark registration system. As the traditional classification of trademark images is based on their shape features and types of object depicted by employing manually-assigned codes, faults or slips may appear because of different subjective perception of the trademark images. Evidence has been provided that the traditional classification is not feasible in dealing with a large fraction of trademark images with little or no representational meanings [1].

Trademarks can be categorised into a few different types. A trademark can be 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. However, for a device-only mark, the trademark only contains symbols, icons or images. If a trademark comprises both words and any iconic symbols or images, it can be regarded as a device-and-word mark [2]. Since different algorithms have to be used in describing different kinds of trademark images, a trademark image retrieval system can only be designed to accommodate one of the types. Although several trademark image retrieval systems have been designed to handle all kinds of trademark images, the performance of these systems are rather unfavourable when compared to those systems that are specifically designed to handle only one kind of trademark. Another challenge in trademark image retrieval is the difficulty in modelling human perception about similarity between trademarks. As human perception of an image involves collaboration between different sensoria, it is in fact difficult to integrate such human perception mechanisms into a trademark image retrieval system.

The contributions of this paper are summarized as follows. 1) Novel algorithms are proposed to describe the shape of device-only marks and device-and-word marks; 2) A two-component feature matching strategy is applied to compare global and local features; 3) This study not only evaluates the proposed method, but also investigates the retrieval performance

of another four algorithms.

The rest of this paper is organized as follows. Section 2 reviews the related studies regarding the existing trademark image retrieval systems and techniques. Section 3 provides an overview of the proposed system architecture. Section 4 presents the algorithms proposed for extracting global and local features of trademarks. Section 5 describes a two-component matching strategy for measuring similarity between trademarks. Section 6 evaluates the performance and analyzes the results. Finally, conclusions are drawn in Section 7.

2. Existing Trademark Retrieval Systems

There are several remarkable trademark image retrieval systems that have been developed in recent years. TRADEMARK [3], STAR [4] and ARTISAN [5] are three of the most prominent trademark image retrieval systems. Different methodologies have been employed in these trademark systems. The TRADEMARK system uses graphical feature vectors (GF-vector) to interpret the image content automatically and calculate the similarity based on human perception [3]. The STAR system adopts mainstream content-based image retrieval (CBIR) techniques. The techniques adopted in this system include the Fourier descriptors, grey level projection and moment invariants. The STAR system works by considering both shape components and the spatial layout of an image [4]. However, the recognition of some perceptually significant components has been considered to be too difficult to be done by an automated process [1]. Therefore, to a certain extent, manual operation is needed for the segmentation of some abstract trademark images. The ARTISAN system, however, introduces an innovative approach that incorporates principles derived from Gestalt psychology to cope with device-only marks which consist of some abstract geometric designs [5].

Apart from the TRADEMARK, STAR and ARTISAN trademark image retrieval systems, a significant amount of research work has also concentrated on trademark or logo image

retrieval. Hussain and Eakins [6] employed the topological properties of the self-organising map for similarity retrieval from a trademark image database. Cerri et al. [7] utilised geometrical-topological tools for describing trademark shapes and matching their similarity based on size functions of the trademarks. Jiang et al. [8] presented a new approach by using the adaptive selection of visual features with different kinds of visual saliencies, including symmetry, continuity, proximity, parallelism and closure property of the trademark images. Hung et al. [9] exploited the contour and interior region for retrieving similar trademark images. Petrakis et al [10] utilised relevance feedback for logo and trademark image retrieval on the web. Shen et al. [11] used block feature index to describe trademark shape. An enhanced normalization technique for the wavelet shape description of trademarks was developed by Li and Edwards [12]. Kim and Kim [13] used Zernike moments as shape descriptor and conducted experiments based on the MPEG-7 core experiment procedure. Improved performance was reported when compared to other methods based on various shape descriptors; however Zernike moments can only describe global shape. This leaves room for further improvement. Jain and Vailaya proposed another trademark retrieval system [14], using the histograms of edge directions as shape descriptor. Although the descriptor is invariant to translation, as the authors have indicated, it is not invariant to scaling and rotation. The performance in terms of accuracy and time complexity is also sensitive to the size of histogram bins. Moreover, like [13], the descriptor cannot represent local features. In addition to the aforementioned methods, the reader is referred to [15-24] for earlier studies on trademark image retrieval.

3. Overview of the Proposed System

Our trademark image retrieval system consists of an *offline database construction* part and an *online image retrieval* part as shown in Figure 1. The offline database construction part is

intended to ensure high retrieval efficiency by extracting a feature set for each of the images in the database in an offline manner and storing the feature set along with its corresponding image in the database so that when a query image is presented to the system, the system does not have to perform online feature extraction on each database image.

To access the database, the user initiates the online image retrieval process by providing a query image as input, and then the system starts with extracting the features from the query image. Afterwards, the system measures the similarity between the feature set of the query image and those of the images stored in the database. Finally, the system ranks the relevance based on the similarity and returns the results to the user.

4. Feature Extraction

As shown in Figure 1, feature extraction has to be done in both offline database construction and online image retrieval processes. Feature extraction is about extracting from an image a set of attributes / features that can feasibly describe / represent the image in order to facilitate the ensuing feature matching and ranking processes. The feature extraction stage in the proposed algorithm involves image pre-processing and feature representation as shown in Figure 2. The purpose of the image pre-processing module is to detect the edges in order to pave the way for the feature representation module. The task of the feature representation module is to extract a *set of feature descriptors* (also called *feature vector*) of the images from their corresponding edge maps generated by the image pre-processing module. In this work, we divide the feature descriptors into a subset of *local* features descriptors and a subset of *global* features descriptors. As tabulated in Table 1, the global feature vector is intended to capture the gross essence of the shape of the trademark while the local feature vector is meant for capturing the interior details of the trademark in the edge map. In this work, *curvature* and *distance to centroid* are used for describing the local features whereas *Zernike moments* are

employed to extract the global features of the images. Detailed descriptions regarding the image pre-processing procedure, global and local feature extraction are successively presented in the following subsections.

4.1 Image Pre-processing

Trademarks come in different sizes and shapes and their key components appear in different locations. An effective retrieval system should be invariant to the scaling, rotation and translation of trademarks. To achieve these objectives, the proposed CBIR system performs the following step to normalise each trademark image.

- Apply median filtering to remove noise: This step is only necessary if noise is a concern.
- Apply Canny edge detector [25] to find the intensity discontinuities in an image: The Canny operator produces an *binary* edge map, which preserves the crucial structure properties and significantly reduces the amount of data in an image. The resulting edge map represents the contour of the underlying trademark image.
- Find the geometrical centroid, (x_0, y_0), of the edge points
- Create a circular image by superposing a circle centred at (x_0, y_0), with the radius equal to the distance between (x_0, y_0) and the most distant edge point, on the trademark image.
- Scale the circular image to a fixed size and then transform all the pixels within the circular image from the Cartesian coordinates system into the polar coordinates system (ρ, θ), where $\rho = ((x - x_0) + (y - y_0))^{1/2}$ and $\theta = \tan^{-1}((y - y_0)/(x - x_0))$.

4.2 Global Feature Extraction

The Zernike basis functions satisfy the orthogonal property [26, 27], implying that the contribution of each moment coefficient to the underlying image is unique and independent, i.e. no redundant information overlapped between the moments. Motivated by this property, in this study, we use the Zernike moments to describe the essential features of the trademark shapes. The Zernike moments are derived from a set of complex polynomials orthogonal over the interior of a unit circle $U : x^2 + y^2 \le 1$ and defined in the polar coordinates. The form of 2-dimensional Zernike polynomial V is expressed as

$$V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta), \tag{1}$$

where n and m are called order and repetition, respectively. The order n is a non-negative integer, and the repetition m is an integer satisfying $n - |m| =$ an *even* number and $|m| \le n$. j is an imaginary unit $\sqrt{-1}$. $R_{nm}(\rho)$ is the 1-dimensional radial polynomial, which is defined as

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{(n-|m|)}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \tag{2}$$

As the Zernike moments are the projection of image $I(x, y)$ onto these orthogonal basis functions, the image can be decomposed into a weighted sum of the Zernike polynomials

$$I(x, y) = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} A_{nm} V_{nm} \tag{3}$$

where A_{nm} are the Zernike moments, i.e., the coefficients of the Zernike polynomials. The Zernike moments of an image $I(x, y)$ with continuous intensity are calculated according to the following equation

$$A_{nm} = \frac{n+1}{\pi} \iint_U I(x, y) V_{nm}(\rho, \theta) dx dy \tag{4}$$

For a digital image of $N \times N$ pixels, the discrete form of the Zernike moments for an image is expressed as

$$A_{nm} = \frac{n+1}{\lambda} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x,y) V_{nm}(\rho, \theta) \tag{5}$$

where $\lambda = \delta A/\pi$ iis a normalizing constant. δA is the elemental area of the square image when projected onto the unit circle of the Zernike polynomials. As the Zernike basis functions take the unit disk as their domains, the disk must be specified before the moments are calculated. In this study, all the trademark images are extracted and projected onto a unit circle of fixed radius of 100 pixels. This step makes the resulting moments invariant to translation and scale, in addition to the rotation-invariant nature possessed in Zernike polynomials.

Since the purpose of using the Zernike moments is to expand an image into a series of orthogonal bases, the precision of shape representation depends on the number of moments used from the expansion. To human eyes, the coefficients of the orders higher than the fifth or sixth order are too small to be measurable reliably so high-order coefficients are always ignored [28]. For this consideration, this work computes the first 4-order Zernike moments so as to yield the most effective and reliable measurement of the global shape. The first 4-order Zernike moments are listed Table 2 [28]. The reader is referred to [26, 28] for more details about the expression of these 15 Zernike polynomials.

4.3 Local Feature Extraction

The local feature descriptors are obtained by calculating the curvature at each boundary / edge point and the distance between the centroid and each boundary / edge point in the edge map. The two features are called *curvature* and *distance to centroid*, and will be described in this section, respectively.

The curvature of each boundary point in the edge map is calculated by the reciprocal of

the radius of an osculating circle on the Cartesian plane. For instance, when the magnitude of the curvature is approaching zero, a straighter line can be identified. According to the illustration and formula in Figure 3(a), a straight line can be identified when all the points on the curve have zero curvature, while on the contrary, a circle can also be identified when all the points on the curve have 1/*r* curvature. Given a plane curve $y = f(x)$, the curvature κ can be expressed as [29]

$$\kappa = \frac{y''}{[1+(y')^2]^{3/2}} \tag{6}$$

where y', y'' are the 1st and 2nd order derivatives with respective to *x*. After obtaining all the curvature value of each boundary point of the image, the curvature feature is expressed as $\sigma(\kappa_1, \kappa_2, ..., \kappa_n)$, where σ is the standard deviation. The curvature feature not only captures interior minutiae, but also overcomes the problem caused by anisotropic stretch, an action of uneven stretch from different directions [30].

As for the *distance to centroid* feature, the edge map projected onto the polar coordinates will be employed to express the interior structure of the corresponding trademark image. The feature is obtained by calculating the distance from each boundary point to the centroid (x_0, y_0). As the edge map is projected onto a fixed unit circle, a histogram of all the distances to the centroid can be generated.

Given an image space containing *b* distance bins in a fixed unit circle, the distance histogram of the edge map of image $I(x, y)$ is represented as $H(I) = [h_1, h_2, \cdots, h_b]$, where h_i is the total number of boundary points in the i^{th} distance bin. The total distance can be expressed as

$$D = \sum_{i=1}^{b} h_i d_i \tag{7}$$

where d_i is the average of the values in the i^{th} distance bin. As a result, we can obtain the two

features μ_D and σ_D, which represent the mean and standard deviation of the distance from the boundary points to the centroid, respectively.

Figure 4 shows that although the four bat trademarks are differently oriented, their distributions of the features of *distance to centroid* extracted by our method remain very similar, indicating that *distance to centroid* is indeed invariant to rotation. Figure 5 indicates that the two features based on *distance to centroid* enable the proposed algorithm to classify different classes. The images on the left-hand side of Figure 5(a) and 5(b) are the original trademarks while the ones on the right hand side are the corresponding edge maps. These two significantly different histograms show strong evidence that *distance to centroid* is capable of capturing the unique local details of the trademarks and differentiating their difference even their global shape contours are the same. Figure 5(c) depicts the histograms of the distances to the centroid extracted from two similar trademark images of the same class. Although Figure 5(a) and 5(b) suggest that *distance to centroid* is capable of differentiating trademarks with significantly different local details, the similarity between the two histograms of Figure 5(c) demonstrates that *distance to centroid* is not sensitive to relatively insignificant differences between local details. This is an important characteristic in content-based image retrieval. Figure 5(d) shows the four histograms in the same plot.

5 Two-Component Feature Matching

Feature matching is about measuring the similarity between the feature vectors of the query image and the database images. An appropriate choice of feature matching techniques can enhance the system performance while an inappropriate choice may lead to unexpected results from the system even though an effective approach has been employed for feature representation. As trademark images can be structurally and conceptually alike with different

interior details, during the feature matching stage, global features and local features indicated in Table 1 are compared separately in order to enhance the system performance. The reader is reminded that the global feature component includes the 15 Zernike moments of orders 0 to 4 while the local feature component includes the standard deviation of *curvature*, the mean and the standard deviation of *distance to centroid.*

By utilising the Euclidean distance to compute the similarity between the query image features and the features stored in the database, the similarity between two images can be obtained. In order to discriminate the relevant images from the irrelevant images through the computation of Euclidean distance, the Euclidean distance is normalized into the range [0,1] and a threshold of 0.3 is set up for both feature components. In other words, if the Euclidean distance between a database image and the query image is greater than 0.3 for only *one* of the two feature components, a penalty value 1 will be added to its current Euclidean distance value, whereas a penalty value 2 will be added to its current distance value if the values obtained are greater than 0.3 for *both* components. The strategy ensures that the relevance is determined by both feature components, and precludes extreme cases. The final value of Euclidean distance obtained for each image can be used to rank images in ascending order. The candidate images with a smaller Euclidean distance are placed on the top while the candidate images with a larger Euclidean distance are placed at the bottom.

6 Performance Evaluation

In this section, the results obtained using the proposed algorithm under various conditions are presented. Precision and recall were computed for performance evaluation. The performance of the proposed algorithm is compared against two region-based descriptors: *moment invariants* [31] and *Zernike moments only* and two contour-based descriptors: *Fourier descriptors* [32, 33] and *curvature scaled space* (*CSS*) *descriptors* [32]. Some prior studies

[34, 35] have reported that the combined descriptors have better performance than individual algorithms in shape description. Since Zernike moments are considered to perform well in shape description, another two descriptors which combine Zernike moments and conventional algorithms are also investigated for their performances. The one combined descriptor *Zer.+Fou.* includes Zernike moments and Fourier descriptors; the other *Zer.+Mom.* combines Zernike moments and moment invariants. The proposed algorithm, these four conventional algorithms, and these two combined descriptors are examined based upon the graphs generated from the average precision and recall.

6.1 Trademark Dataset

One of the significant issues regarding the performance evaluation of trademark retrieval systems is the database. Although MPEG-7 provides a set of trademarks for evaluation, the images are not deliberately designed for performance evaluation of CBIR systems. For instance, the trademarks from MPEG-7 are usually semantically and perceptually alike but not structurally and conceptually alike. However, the greatest value of a trademark image retrieval system is to locate the trademark images that are highly-homogeneous, rotated and scaled to the query example. Therefore in this study, we create a trademark dataset with 1,003 trademark images for testing the algorithms against structural changes of trademark images. The trademark dataset contains 14 different classes and the classification is shown in Table 3. We follow several rules during the creation of the trademark image database: 1) all the trademark images within the same class should be conceptually similar except in the miscellaneous class (Class 14); 2) all the trademark images should be classified into different classes according to their visual properties; 3) all the trademark images should be binary images of 200 × 200 pixels;

6.2 Performance Comparison

The precision-recall graph is commonly used for evaluating the performance of CBIR systems. The *x*-axis and *y*-axis represent recall and precision rates, respectively. *Precision* is the ratio of the number of *relevant* images retrieved to the total number of images retrieved while *recall* is the number of *relevant* images retrieved to the total number of *relevant* images stored in the database. The evaluation procedure of this study is described as follows. For each query image presented to the system, the system returns 11 pages of hits with descending similarity rankings (i.e., ascending Euclidean distance ranking), each page containing 9 trademark images. This allows the performance of the system to be evaluated on a *page-wise* manner. However, in order to obtain a more objective picture of the system's performance, instead of plotting the precision-recall graph for each individual query image, we plot the average performance of the system after 20 query images of the *same* class have been presented to the system. As can be readily seen in Table 3, there are only 5 classes (i.e., Class 1, 3, 9, 10, and 12) in our database with more than 99 similar / relevant images, the query images will only be taken from those 5 classes. The performance of the proposed system and the other four algorithms with respect to the query images from the five classes are plotted in Figure 6(a) to (e), respectively. Each curve represents the performance of one algorithm and consists of 11 data point, with the i^{th} point from the left corresponding to the performance when the first i pages of hits are taken into account. A precision-recall line stretching longer horizontally and staying high in the graph indicates that the corresponding algorithm performs relatively better.

As the Example Images column of Table 3 shows, Class 1 (Apples) consists of trademarks with simple shape and fewer disjoint or non-connected components. This class was designed

to test the robustness against rotation, scaling, translation and non-rigid deformation of each algorithm. In Figure 6(a), the graph shows that the proposed algorithm outperforms other algorithms in retrieving the Class 1 images. The two combined descriptors, like the proposed algorithm, perform two-component feature matching strategy to measure the similarity, which makes the combined descriptors perform better than those conventional algorithms. The algorithm using *moment invariants* gives the poorest results among all the algorithms. We observed that the algorithm using *moment invariants* is most sensitive to the change of size (scale) of the trademarks. This suggests that normalising the size of the trademarks is a possible way of boosting the performance of the algorithm using *moment invariants*.

In Figure 6(b), the graph shows that the Fourier and the CSS algorithms perform relatively poorer than others in retrieving relevant images from Class 3. Trademarks in Class 3 are mainly circular-shaped, formed by non-connected components. We observed that the Fourier descriptors perform poorly in describing trademarks with non-connected components while the CSS algorithm has difficulty dealing with circular objects. Again the proposed algorithm outperforms others although the performance curve is not as good as its counterpart in Figure 6(a).

Figure 6(c) shows that the proposed algorithm performs well when describing encircled image objects such as the bat trademarks in Class 9. This can be attributed to the use of the mean and standard deviation of *distance to centroid*. These two features based on *distance to centroid* tend to produce more distinctive results for encircled image objects. The images from Class 9, which are relevant to the query image, tend to have a *distance to centroid* closed to zero, and the irrelevant images tend to have a greater distance. As a result, the precision and recall rates of the proposed algorithm for this class are increased.

In Figure 6(d), the Fourier and the CSS descriptors again perform relatively poorer than others when Class 10 images, which are of the same shape with different interior details, are presented as queries. This can be attributed to the inherent properties of contour-based shape descriptors such as the Fourier and the CSS descriptors. Since the contour-based shape descriptors mainly depict the global shape of trademarks, they are unable to distinguish the images with similar shapes, even though their interior structures are different. Although the algorithm using only the Zernike moments performs reasonably well in most image classes, it does not give better performance in Class 10 than the algorithm using moment invariants does. This can be attributed to the fact that the Zernike moments tend to capture the essential features of the shapes, but not the interior details. Furthermore, it can also be seen in Table 3 that Class 2 and 8 both contain images with contours / shapes similar to those of the images in Class 10, so during the search, the Zernike moments tend to retrieve the trademark images which belong to Class 2, 8 and 10. As a consequence, a comparatively low precision and recall rates were obtained. An example that illustrates the retrieval results obtained by the algorithm using only Zernike moments is shown in Table 4.

Figure 6(e) shows a sharp fall of the performance curve of the proposed algorithm after its recall rate reaches 13%. This is due to the degree of variation present in Class 12, which includes some hollow shape images as well as some deformed images. Half of the class is made up with these types of images. This shows that the proposed algorithm, although still outperforms others, is not sufficiently capable of describing hollow-shaped or significantly deformed images. Some example images from Class 12 are shown in Figure 7. The last column of Figure 7 shows two images with identical shape, but if the filled image is taken as the query image, the proposed algorithm is unable to retrieve the hollow image. The inter-class

variation widens the semantic gap between the low level features that are calculated by the computer and the subjective meaning perceived by the user, and entails the needs for the involvement of user's relevance feedbacks. Another related issue is that many different trademarks are enclosed in a rectangle, which is usually not perceived as relevant by humans while it is likely to appear as part of the trademarks by the system. Again, this issue entails the need for relevance feedback and we will investigate into methods for addressing it in the future.

In Figure 6(f), the data points shown on the graph are the average of all the precisions and recalls displayed in Figure 6(a) to (e), i.e., the average performance based on the 100 rounds of search. We can see that the proposed algorithm performs significantly better than others. It is also interesting to see that the performance of the algorithm using only the Zernike moments is only second to the proposed algorithm and significantly better than the other three conventional algorithms. Since the difference between the proposed algorithm and the algorithm using only the Zernike moments is that in addition to these 15 Zernike coefficients we also add the three descriptors for depicting the local / interior details of trademarks. The performance difference between these two algorithms illustrated in Figure 6(f) manifests the effectiveness of the three local features. In addition, that fact that those two combined descriptors perform better than other individual algorithms indicates that trademarks need to be interpreted using different algorithms for global and local structure, respectively.

7 Conclusions

In this work, we have proposed a novel content-based trademark retrieval system with a feasible set of feature descriptors, which is capable of depicting global shapes and interior /

local features of the trademarks. We have also proposed an effective two-component feature matching strategy for measure the similarity between feature sets. By utilising the curvature feature and the distance to centroid, the proposed algorithm is robust against rotation, translation, scaling and stretching. As for the image retrieval stage, a two-component matching strategy was used in feature matching. With this strategy, the images can be compared with the query image with their local and global features taken into account separately, and therefore enabling the system to be insensitive to noise or small regional changes.

The performance of the proposed algorithm was evaluated in terms of the precision and recall rates. The precision-recall graphs show that the proposed algorithm outperforms other conventional algorithms, including moment invariants, the Fourier descriptors, the Zernike moments only and the CSS. Nevertheless, the proposed scheme is not sufficiently capable of relating the trademarks with similar semantic meanings but significantly different low level features. We are currently investigating ways of incorporating relevance feedback into the proposed system in order to tackle this challenge.

8 References

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

[2] J.P. Eakins, "Trademark image retrieval", In S.W. Lew (Ed.), *Principles of Visual Information Retrieval* (pp. 319-354), New York: Springer-Verlag, 2001.

[3] T. Kato, "Database architecture for content-based image retrieval", *Proceedings of SPIE Image Storage and Retrieval Systems*, vol. 1662, pp. 112-123, 1992.

[4] J.K. Wu, C.P. Lam, B.M. Mehtre, Y.J. Gao, and A. Narasimhalu, "Content-based retrieval for

trademark registration", *Multimedia Tools Application*, vol. 3, no. 3, pp. 245-267, 1996.

[5] J. P. Eakins, M.E. Graham, and J.M. Boardman, "Evaluation of a trademark image retrieval system", Information Retrieval Research, *the 19th Annual BCS-IRSG Colloquium on IR Research*, 1997.

[6] M. Hussain and J.P. Eakins, "Component-based visual clustering using the self-organizing map," *Neural Networks*, vol. 20, no. 2, pp. 260-273, 2007.

[7] A. Cerri, M. Ferri and D. Giorgi, "Retrieval of trademark images by means of size functions," *Graphical Models*, vol. 68, no. 5-6, pp. 451-471, 2006.

[8] H. Jiang, C.-W. Ngo and H.-K. Tan, "Gestalt-based feature similarity measure in trademark database," *Pattern Recognition*, vol. 39, no. 5, pp. 988-1001, 2006.

[9] M.H. Hung, C.H. Hsieh, and C.-M. Kuo, "Similarity retrieval of shape images based on database classification," *Journal of Visual Communication & Image Representation*, vol. 17, no. 5, 970-985, 2006.

[10] E.G.M. Petrakis, K. Kontis, and E. Voutsakis, "Relevance feedback methods for logo and trademark image retrieval on the Web," *Proceedings of the 2006 ACM Symposium on Applied Computing*, pp. 1084-1088, 2006.

[11] D.-F. Shen, J. Li, H.T. Chang, and H.H.P. Wu, "Trademark retrieval based on block feature index code," *Proceedings of IEEE International Conference on Image Processing 2005*, vol. III, pp. 177-180, 2005.

[12] Q. Li and J. Edwards, "An enhanced normalisation technique for wavelet shape descriptors," *The Fourth International Conference on Computer and Information Technology*, pp. 722-729, 2004.

[13] W.-Y. Kim and Y.-S. Kim, "A region-based shape descriptor using Zernike moments," *Signal Processing: Image Communication*, vol. 16, no. 1-2, pp. 95-102, 2000.

[14] A.K. Jain, and A. Vailaya, "Shape-based retrieval: a case study with trademark image databases,"

Pattern Recognition, vol. 31, no. 5, pp. 1369–1390, 1998.

[15] J.-K. Wu, C.-P. Lam, B.M. Mehtre, Y.J. Gao, and A.D. Narasimhalu, "Content-based retrieval for trademark registration," *Multimedia Tools Applications*, vol. 3, no. 3, pp. 245–267, 1996.

[16] P.N. Suganthan, "Shape indexing using self-organizing maps," *IEEE Transactions on Neural Networks*, vol. 13, no. 4, pp. 835 – 840, 2002.

[17] J.P. Eakins, J.M. Boardman, and M.E. Graham, "Similarity retrieval of trademark images," *IEEE Multimedia*, vol. 5, no. 2, pp. 53–63, 1998.

[18] I. King and Z. Jin, "Integrated probability function and its application to content-based image retrieval by relevance feedback," *Pattern Recognition*, vol. 36, no. 9, pp. 2177-2186, 2003.

[19] G. Ciocca and R. Schettini, "Content-based similarity retrieval of trademarks using relevance feedback," *Pattern Recognition*, vol. 34, no. 8, pp. 1639-1655, 2001.

[20] O.E. Badawy, and M. Kamel, "Shape-based image retrieval applied to trademark images," *International Journal of Image Graphics*, vol. 2, no. 3, pp. 375-393, 2002.

[21] M. Ren, J.P. Eakins, and P. Briggs, "Human perception of trademark images: implication for retrieval system design," *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 564–575, 2000.

[22] M.-T. Chang and S.-Y. Chen, "Deformed trademark retrieval based on 2D pseudo-hidden Markov model," *Pattern Recognition*, vol. 34, no. 5, pp. 953-967, 2001.

[23] G. Cortelazzo, G. A. Mian, G. Vezzi and P. Zamperoni, "Trademark shapes description by string-matching techniques," *Pattern Recognition*, vol. 27, no. 8, pp. 1005-1018, 1994.

[24] P.-Y. Yin and C.-C. Yeh, "Content-based retrieval from trademark databases," *Pattern Recognition Letters*, vol. 23, no. 1-3, pp. 113-126, 2002.

[25] J. Canny, "A computational approach for edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, 1986.

[26] C.-Y. Wee and R. Paramesran, "On the computational aspects of Zernike moments," *Image and Vision Computing*, vol. 25, no. 6, pp. 967-980, 2007.

[27] G. A. Papakostas, Y.S. Boutalis, D.A. Karras, and B.G. Mertzios, "A new class of Zernike moments for computer vision applications," *Information Sciences*, vol. 177, no. 13, pp. 2802-2819, 2007.

[28] W.N. Charman, "Wavefront technology: Past, present, and future," *Contact Lens and Anterior Eye*, vol. 28, no. 2, pp. 75-92, 2005.

[29] G. James, *Modern Engineering mathematics*, 3ed. Prentice Hall, 2001.

[30] Z. Hu, D. Metaxas and L. Axel, "In vivo strain and stress estimation of the heart left and right ventricles from MRI images," *Medical Image Analysis*, vol. 7, no. 4, pp. 435-444, 2003.

[31] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, *Digital Image Processing Using MATLAB*, Prentice Hall, 2004.

[32] D. Zhang and G. Lu, "A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval," *Journal of Visual Communication and Image Representation*, vol. 14, no. 1, pp. 39-57, 2002.

[33] R.B. Yadav, N.K. Nishchal, A.K. Gupta, and V.K. Rastogi, "Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study," *Optics and Lasers in Engineering*, vol. 45, no. 6, pp. 695-708, 2007.

[34] J.P. Eakins, K.J. Riley, and J.D. Edwards, "Shape feature matching for trademark image retrieval, *Proceedings of the Second International Conference on Image and Video Retrieval*, pp. 28-38, 2003.

[35] B. Mehtre, M. Kankanhalli, and W. Lee," Shape measures for content based image retrieval: A comparison, *Information Processing & Management*, vol. 33, no. 3, pp. 319-337, 1997.

Table 1. The features extracted to describe the trademark images.

	Features	**Description**
Global Feature	The first 4–order Zernike moments	Gross essence of the trademark shapes

Local Feature	Mean of curvature	Edge direction of the trademark contour
	Mean and standard deviation of distance to centroid	Histogram of distances to centroids of the trademark contour

Table 2. Zernike polynomials up to 4th order and their physical meanings relative to primary aberrations.

n	m	Zernike Polynomial	Physical Meaning
0	0	1	Piston: constant term
1	-1	$2\rho\sin\theta$	Distortion: tilt in x direction
1	1	$2\rho\cos\theta$	Distortion: tilt in y direction
2	-2	$\sqrt{6}\rho^2\sin 2\theta$	Astigmatism with axis at $\pm 45^\circ$
2	0	$\sqrt{3}(2\rho^2-1)$	Spherical defocus: field curvature
2	2	$\sqrt{6}\rho^2\cos 2\theta$	Astigmatism with axis at 0° or 90°
3	-3	$\sqrt{8}\rho^3\sin 3\theta$	Triangular astigmatism, based on *x*-axis (Trefoil)
3	-1	$\sqrt{8}(3\rho^3-2\rho)\sin\theta$	Primary coma along x axis
3	1	$\sqrt{8}(3\rho^3-2\rho)\cos\theta$	Primary coma along y axis
3	3	$\sqrt{8}\rho^3\cos 3\theta$	Triangular astigmatism, based on *y*-axis (Trefoil)
4	-4	$\sqrt{10}\rho^4\sin 4\theta$	Quatrefoil
4	-2	$\sqrt{10}(4\rho^4-3\rho^2)\sin 2\theta$	5th order astigmatism
4	0	$\sqrt{5}(6\rho^4-6\rho^2+1)$	Spherical
4	2	$\sqrt{10}(4\rho^4-3\rho^2)\cos 2\theta$	5th order astigmatism
4	4	$\sqrt{10}\rho^4\cos 4\theta$	Quatrefoil

Table 3. Our dataset of the trademark images.

Class	Name	Quantity	**Example Images**
1	Apple	120	
2	Fan	43	
3	Abstract Circle 1	110	
4	Abstract Circle 2	34	
5	Encircled Cross	67	
6	Abstract Saddle	20	
7	Abstract Sign	35	
8	Triangle	52	
9	Bat	136	
10	Peacock	113	
11	Abstract Flower	26	
12	Rabbit	119	
13	Snow Flake	47	
14	**Miscellaneous**	**81**	

Table 4. Some retrieval results obtained by the algorithm using only Zernike moments.

Query Image	**Rank 1**	**Rank 2**	**Rank 3**	**Rank 4**	**Rank 5**
					
					

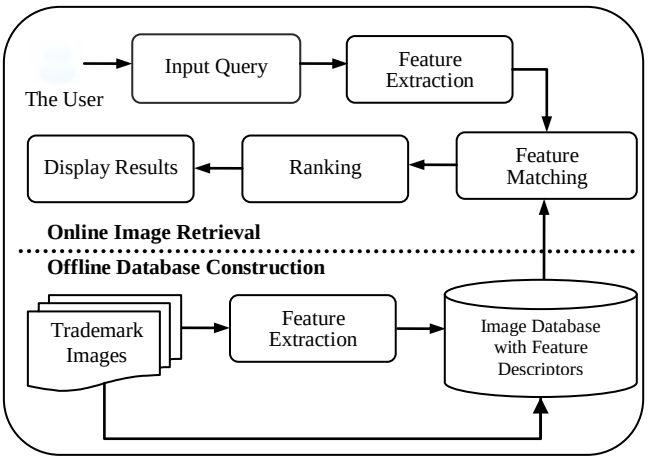


Figure 1. The architecture of the proposed trademark image retrieval system.

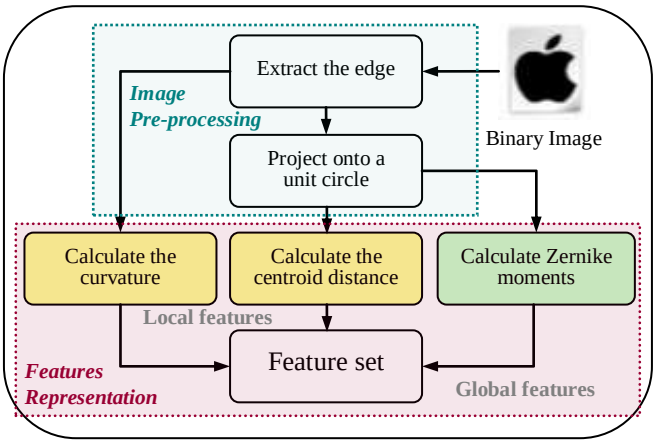
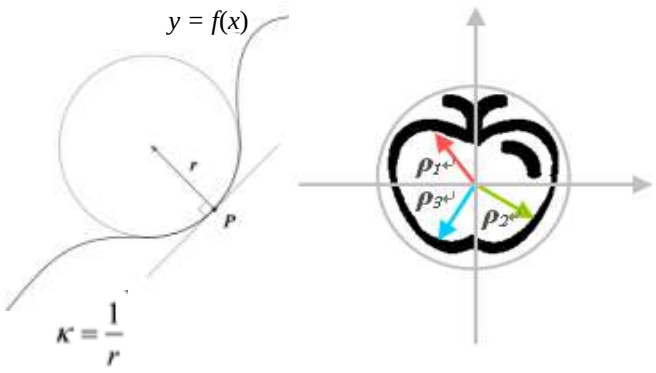


Figure 2. The feature extraction stage of the proposed algorithm.



(a) **(b)**

Figure 3. (a) The curvature at point *P*; **(b)** The distances between the centroid (i.e., the centre of mass) and the boundary points of the image.

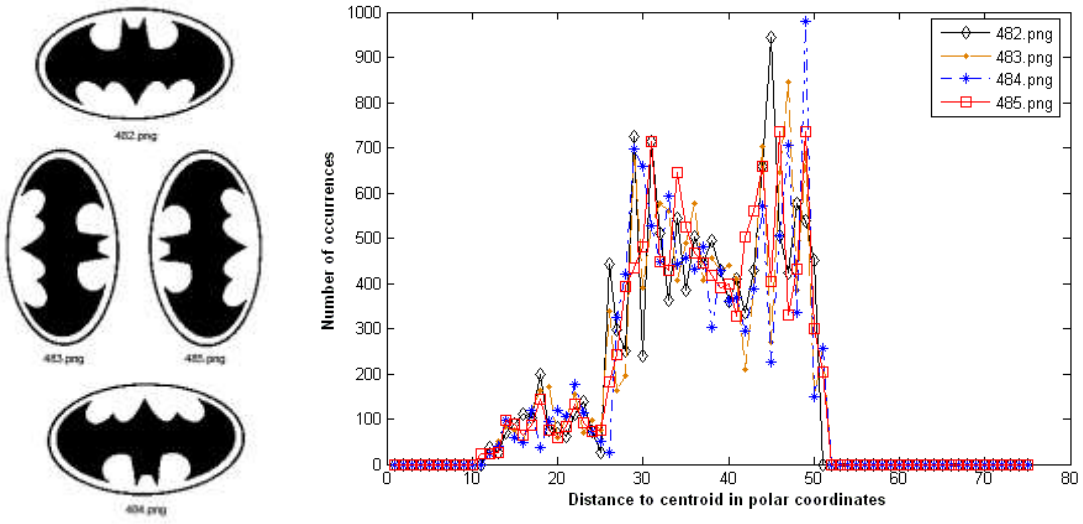


Figure 4. Histograms of *distance to centroid* of four differently oriented bat trademarks. The high similarity of these four histograms illustrates that the *distance to centroid* feature is invariant to rotation.

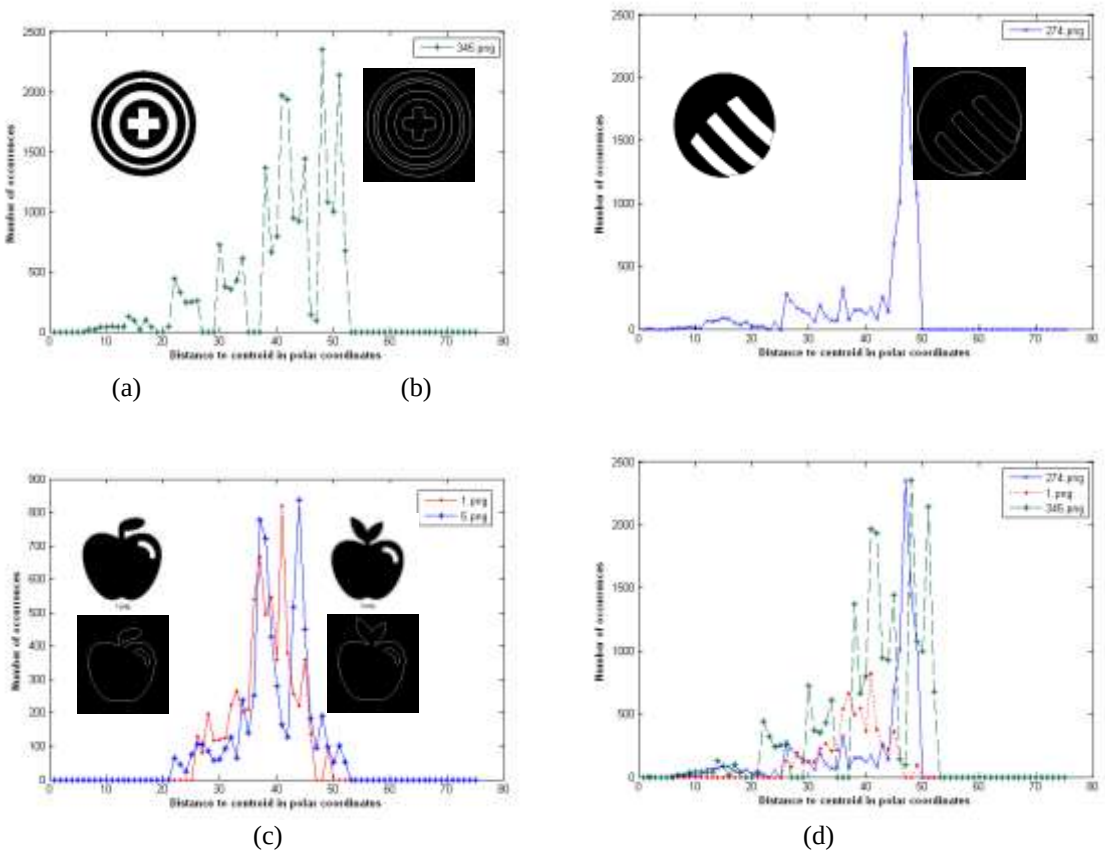


Figure 5. Histograms of distances to centroid extracted from different edge maps of trademark images: **(a)** encircled cross; **(b)** abstract circle 2; **(c)** histograms obtained from two trademarks of the same class; **(d)** histograms of (a), (b) and (c) superposed in the same plot.

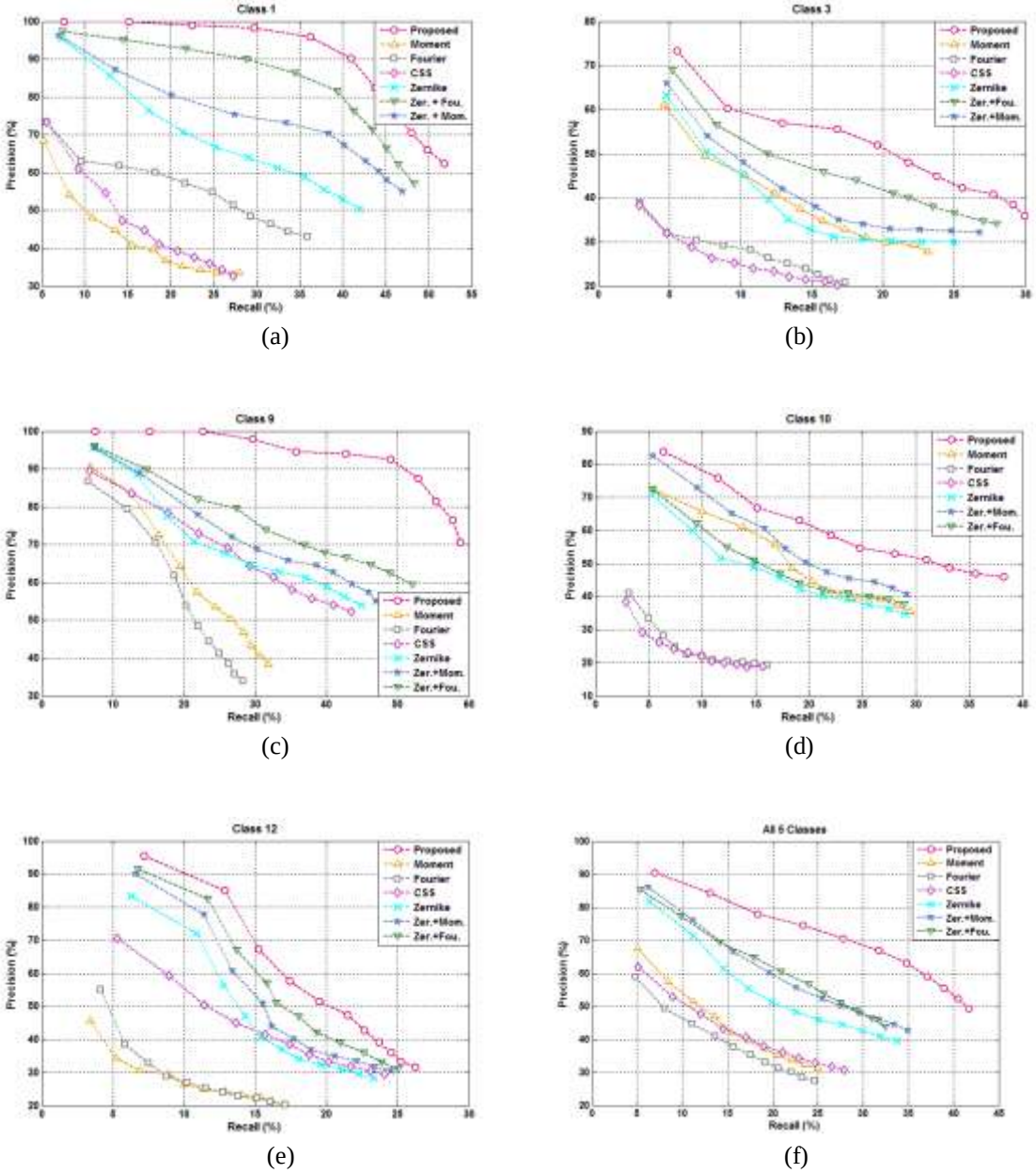


Figure 6. Precision-recall graphs of five algorithms when query images are taken from: **(a)** Class 1, **(b)** Class 3, **(c)** Class 9, **(d)** Class 10 and **(e)** Class 12 of our trademark dataset. **(f)** Overall performance.

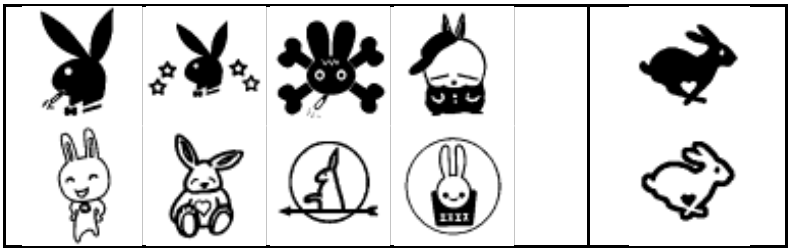


Figure 7. Variety of images within Class 12 of our trademark dataset**.**