A Hybrid Trademark Retrieval System Using Four-Gray-Level Zernike Moments and Image Compactness Indices

Innchyn Her her@mail.nsysu.edu.tw
Professor, Dept. of Electro-mechanical Engineering
National Sun Yat-sen University
Kaohsiung 80424, Taiwan

Kazi Mostafa kazimostafa@gmail.com
Student, Dept. of Electro-mechanical Engineering
National Sun Yat-sen University
Kaohsiung 80424, Taiwan

Huan-Kai Hung hhung23@gmail.com
Student, Dept. of Electro-mechanical Engineering
National Sun Yat-sen University
Kaohsiung 80424, Taiwan

Abstract

There are just too many trademarks out there so that a good automated retrieval system is required to help to protect them from possible infringements. However, it is from people, i.e., the general consumers' viewpoint how similar or confusing two trademarks can be. Thus, in this paper we propose a hybrid system where we patently incorporate human inputs into a computerized trademark retrieval scheme. Various surveys involving general consumers' cognition and responses are conducted, and the results are used as benchmarks in developing the automated part. The core mathematical features used in the scheme are four-gray-level Zernike moments and two new image compactness indices. Experimental results show that this hybrid system, when compared with human-generated results, is able to achieve an average accuracy rate of 95% while that of the closest competing existing method is 65%.

Keywords: Image matching, Information retrieval, Trademarks.

1. INTRODUCTION

Owing to the constant development of commercial activities the number of trademarks used by companies increases dramatically by the year. Consequently, to design a new trademark without infringing others has become a critical and complex issue. With the advent of computers, traditional ways of archiving many documents have been replaced by computerized methods, which feature automated processes and fast and accurate information retrieval. When it comes to trademarks, unfortunately, there is no universally accepted sorting and retrieval scheme so far. For instance, through keywords or *search codes* that are manually assigned to the trademark images, an official retrieval system has been used for some time [1][22]. This procedure, though feasible, contains major drawbacks. It is very difficult for an operator to describe all images evenly objectively and consistently, especially for abstract or complex ones and over a long period of time. On the other hand, a different approach called Content-Based Image Retrieval (CBIR) proposes an image-based strategy instead of a tag-based one. Each trademark is processed beforehand, to leave and save its pertinent features in the computer. Then for a new trademark application, the features of the new image are compared with the stored data to see if there have been similar entries in the database. Such CBIR trademark retrieval systems are intrinsically less susceptible to human errors. A good CBIR scheme tends to exhibit robustness in terms of image size, position, and orientation variations. In general, these methods are often classified into contour-based schemes and region-based schemes.

A very early effort to digitally define the features of an image was the work by Freeman [2]. He invented the chain code, a typical contour-based approach. Many researchers later refined his method to improve the applicability. For example, Peng and Chen [3] more recently proposed to use the chain code to describe simple and properly segmented trademarks, while encoding them according to the angles. Contour-based methods are usually for simple, binary images. An advantage is that the coded contour is usually invariant with respect to translation and rotation [4]. Yet the newer region-based methods have a broader application range and a larger toolbox to use. For example, there are moments of various kinds, often used as descriptive features of an image. A benefit of using the moments is their insensitivity to noise. As a result, researchers have begun to resort more to region-based techniques in these years. For instance, Yin and Yeh [5] have proposed a method to employ a fuzzy approach to expedite the image classifying and retrieval processes. The image features used were the area, the number of closed objects, the location of the centroid, and symmetry of the image, etc.

The invariant moments, devised by Hu [6] based on geometric moments, have long been valued as useful image descriptive features. But they are comparatively sensitive to noise and the accompanying image reconstruction procedure was difficult. Shih and Chen [7] later suggested to use invariant moments with Fourier transforms and the histogram of image boundary orientations as features in a trademark retrieval scheme. Ciocca and Schettini [8] also demonstrated a way for smartly using the invariant moments. They combined moments with boundary directions and the results from a multi-resolution wavelet transform. Through a relevance feedback scheme they could compute the similarity between trademarks. On the other hand, there were researchers who had chosen to use Zernike moments for trademark retrieval schemes. For example, Kim and Kim [9] constructed a Gamma distribution model with Zernike moments to describe visually salient features of monochrome trademarks. Subsequently, Kim et al. [10] also presented a modified Zernike-moment shape descriptor, by first partitioning a trademark image into an inner and an outer region. More recently, Kamila et al. [11] normalized the regular Zernike moments via geometric moments and used them with binary images.

Regardless of these many research efforts on automated trademark retrieval methodology, the governing institution that supervises the examination and approval of trademarks still, for the time being, relies largely on manual and tag-based methods [1], [12], [13], [14]. From scientific point of view, a trademark is invariably an image, a discrete two-dimensional mathematical function. Using *search codes* to describe and to retrieve an image risks the danger of inaccuracy, inconsistency, and inefficiency. But why do people still stick to them? One blunt answer: Undeniably, trademark is not an engineering or scientific term. It is purely commercial; trademarks are meant to interact with people. They contain subtle feelings difficult to perceive or describe by the computer. For the task of trademark comparison and retrieval, human inputs can never be dismissed.

In light of this we now propose a hybrid retrieval system for trademarks. In the core of our numerical scheme, we use Zernike moments for their resistance to noise, their invariance to rotation, and multi-resolution capabilities. Besides, two new image features are proposed as additional sorting criteria for the computer. However, the term "hybrid" implies that human perception works hand in hand with computer's speed. In this system humans set up a benchmark while the computer emulates humans' cognition and feeds back a list of candidates for human to make final judgments. We will show that the hybrid system is robust and able to provide an output very usable for a professional trademark examiner.

2. PRELIMINARIES

2.1 Trademarks

A trademark is a mark that identifies one person's goods [13]. In practice, the word "trademark" is used to refer to any class of mark, or a word, symbol, or phrase, used to distinguish a particular manufacturer's product from those of others. The enactment of the trademark law aims to protect the legal use of the trademarks, so that legitimate transactions and fair business morals are guaranteed. It is obvious that there is little room for confusion in the use of trademarks. However, if there is any, in particular deliberate confusion, an act of infringement is very likely evolving somewhere.

Thanks to the collaborated efforts of industrialized nations and the WTO, trademark definition and laws around the world are quite similar now. According to a typical trademark regulation [14], there are four different kinds of trademarks, as depicted in Fig. 1. In this paper, we will deal mostly with graph and sign trademarks for they are more difficult to describe in words.

There are a few rules in determining the similarity of trademarks. The most important aspect is the impression the whole trademark imposes on the consumers. It is by this crude total impression that people decide whether two trademarks are confusingly similar or not. As confirmed by a senior officer of local Patent and Trademark Office [15], people are more impressed by the rough shape of a trademark, rather than by its details or color. Moreover, people tend to neglect the frame or circle outside the trademark. For computers, a framed trademark can be a far cry from an unframed one. For people, strangely, they produce almost identical impressions.

FIGURE 1: Different kinds of trademarks: (a) Word trademarks, (b) Graph trademarks, (c) Sign trademarks, and (d) Combined trademarks.

To verify the above statements, we have conducted a questionnaire survey that involved 100 general consumers [16]. When presented with a single trademark design, 59 consumers said that the color was the most impressing feature, versus 34 for the shape and 5 for its pronunciation. However, when presented with sets of similar trademarks for comparison, a total of 83 consumers confirmed that the rough shapes of the trademarks, rather than their colors or details, were the most discerning feature for trademark identification. In view of this double testimony, we decided to devise two new features solely for the rough shape of the trademark image. They will be discussed in Section 3.2.

2.2 Moments

Moments are handy properties that are repeatedly used in region-based image retrieval systems, while moments in various orders are regarded as descriptive features of an image. In this paper we will use and compare two types of moments, and their formulations are briefly discussed as follows.

The invariant moments of Hu were proposed in 1962 [6]. In this paper author anticipated a theory of two-dimensional moment invariants for planar geometric figures. In order to make the moments invariant to scaling, translation, and rotation, they need to be normalized. Based on the normalized moments, Hu proposed the seven invariants. Hu's invariants can be used as image descriptive features. The computation of image reconstruction from these invariants, however, is no piece of cake. Besides, they can still be affected by the image noise [8], [11], [17].

On the other hand, the Zernike moments, proposed by Teaque in 1980 [18], are constructed on orthogonal Zernike polynomials. These complex number Zernike functions were invented by the Nobel laureate in 1961, defined within a unit circle. A typical Zernike polynomial is expressed as,

$$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)\, e^{jm\theta} \tag{1}$$

(x, y) represents an arbitrary point in the unit circle, and within the circle, we have $x^2+y^2 \le 1$. The integer n, starting from zero, is the order of the polynomial. m is also an integer that satisfies $|m| \le n$, and $n-|m|$ must be an even number. ρ is the magnitude of vector (x, y), and θ is the orientation, measuring counterclockwise from the x-axis. j is the unit imaginary number $\sqrt{-1}$. Now the radial component R_{nm} (ρ) of Zernike polynomial is given by:

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

According to the above expression, we know $R_{n,-m}(\rho) = R_{nm}(\rho)$. Moreover, the complete orthogonal property of Zernike polynomial V(x,y) within the unit circle can be demonstrated by this equation:

$$\iint_{x^2+y^2\le 1} V_{nm}^{*}(x,y)\, V_{pq}(x,y)\, dxdy = \frac{\pi}{n+1}\delta_{np}\delta_{mq}$$
$$\text{and } \delta_{ab} = \begin{cases} 1 & (a=b) \\ 0 & (\text{otherwise}) \end{cases} \tag{3}$$

Where * denotes the conjugate of a complex number.

Zernike moments are the projection of the image function f(x, y) onto the above Zernike orthogonal basis functions [11]. We denote Zernike moment of order n for a continuous function as A_{nm}, and

$$A_{nm} = \frac{n+1}{\pi}\iint_{x^2+y^2\le 1} f(x,y)\, V_{nm}^{*}(\rho,\theta)\, dxdy \tag{4}$$

For a digital image, this equation is simplified as:

$$A_{nm} = \frac{n+1}{\pi}\sum_x\sum_y f(x,y)\; V_{nm}^{*}(\rho,\theta), \quad x^2+y^2 \le 1 \tag{5}$$

Note that the conjugate of A_{nm} is identical to $A_{n,-m}$, and the magnitude (denoted by Z_{nm}) of Zernike moment also has the following property:

$$Z_{nm} = \|A_{nm}\| = \|A_{n,-m}\| \tag{6}$$

In a typical image analysis, an image is decomposed via above equations into a series of Zernike moments A_{nm} of various orders. Conversely, by properly assembling these moments with corresponding Zernike polynomials V_{nm}, the original image function f(x, y) can also be reconstructed [16].

$$f(x,y)=f(\rho,\theta)\cong\sum_{n=0}^{N}\sum_{m}A_{nm}(\rho,\theta)\;V_{nm}(\rho,\theta) \tag{7}$$

Now let us assume that an image f(ρ, θ), the polar form of f(x, y), has undergone a rotation about the origin by an angle α. The rotated image is then $f^{r}(\rho,\theta)=f(\rho,\theta-\alpha)$. Rewrite the previous Zernike moment equation in polar form, we have

$$A_{nm}=\frac{n+1}{\pi}\int_{0}^{2\pi}\int_{0}^{1}f(\rho,\theta)\;R_{nm}(\rho)e^{-jm\theta}\rho\,d\rho\,d\theta \tag{8}$$

But the Zernike moment of the rotated image $f^{r}(\rho,\theta)$ is

$$A_{nm}^{r}=\frac{n+1}{\pi}\int_{0}^{2\pi}\int_{0}^{1}f(\rho,\theta-\alpha)\;R_{nm}(\rho)e^{-jm\theta}\rho\,d\rho\,d\theta \tag{9}$$

Substituting $\theta=\tilde{\theta}+\alpha$ into above equation yields

$$A_{nm}^{r}=\left[\frac{n+1}{\pi}\int_{0}^{2\pi}\int_{0}^{1}f(\rho,\tilde{\theta})R_{nm}(\rho)e^{-jm\tilde{\theta}}\;\rho\,d\rho\,d\tilde{\theta}\right]e^{-jm\alpha}=A_{nm}\;e^{-jm\alpha} \tag{10}$$

This expression leads to the equivalence of the magnitude:

$$Z_{nm}^{r}=\left\|A_{nm}^{r}\right\|=\left\|A_{nm}\right\|=Z_{nm} \tag{11}$$

Which means invariance of rotation within a unit circle.

3. METHODOLOGY

3.1 Smallest Enclosing Circle

Zernike moments are to be used as a kernel feature in our scheme, and it has an intrinsic property of rotational invariance. However, in order to achieve scaling and translational invariance, additional arrangements are needed. Since Zernike polynomials are defined within the unit circle, a method for conveniently finding the smallest enclosing circle for a trademark is necessary. This will take care of the translation concern. Then, the whole thing will be scaled to a fixed size for use with Zernike moments and other retrieval computations.

There exist many schemes for finding the smallest enclosing circle for different applications. For instance, Berg et al. [19] included in their book a recursive formula as such: For a set P= $\{p_1, p_2, \ldots p_i, \ldots, p_n\}$ of n points in a plane, if D_i is the smallest circle enclosing, $P_i = \{p_1, p_2, \ldots p_i\}$, the following rules hold.

(a) If $p_i \in D_{i-1}$, then $D_i = D_{i-1}$

(b) If $p_i \notin D_{i-1}$, then p_i is on the boundary of D_i. So starting from D_1, eventually we get D_n.

The point-by-point method is simple in formulation (as shown in Figure 2), however computationally intensive in practice, especially for digital images that normally contain hundreds of points (even on the boundary). In this paper, we use a simplified version, based upon the above scheme, for locating the smallest enclosing circle for a trademark image. Firstly, for a given trademark, extract digitally its contour, which should contain all boundary points of the image. Then, search among the boundary points and find the two points that are most distantly apart, say, points a and b in Fig. 2(a).

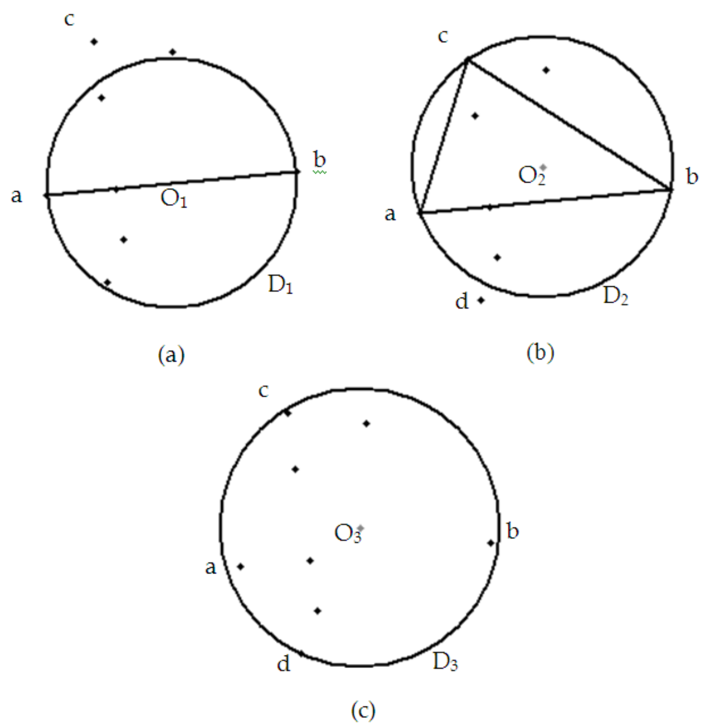


FIGURE 2: A simplified scheme for locating the smallest enclosing circle.

Construct circle D_1 so that line $\overline{ab}$ is a diameter of the circle. Next, if there are some points outside circle D_1, locate among them the point c, which is farthest from the center O_1 of D_1. Construct circle D_2 by points a, b, and c, as shown in Fig. 2(b). Lastly, if there still are some remaining points outside D_2, there is a final modification to make. Find point d that is most distant from O_2 (the center of D_2). Let $\overline{O_2d} + \overline{O_2c}$ be the length of the new diameter, and the new center be adjusted to O_3, which is on line $\overline{O_2d}$. Construct D_3 as illustrated in Fig. 2(c), which is the smallest enclosing circle used in this paper. Within four steps, we swiftly determine a usable smallest enclosing circle for a trademark image. Although the above method was heuristic, the enclosing circle it finds is well defined and suffices for our purpose. In our experience, we have not yet encountered a trademark that has a problem with this simplified scheme. When the enclosing circle is found, a transformation matrix can easily be formulated [20] to translate the trademark to the origin and to scale the image to the prescribed dimension. This dimension was 64×64 in most of our experiments.

3.2 Wrap Contour and Compactness Indices

Besides Zernike moments, which are a region-based method, we also require in this paper some contour-based features that can represent the gross shape of the trademark. We now propose a wrap contour concept, which we think is more describing than the common convex hull concept for two-dimensional objects (Fig. 3). A convex hull, having to maintain a convex shape, often has only loose contact with the object, and therefore poorly represents the contour of the object (Fig. 3b). The wrap contour (Fig. 3c), on the contrary, is allowed to shrink snugly inward the object. Thus, in the end it better represents the approximate shape of the object.

Technically, the smallest enclosing circle of the trademark is first to be located. Then, from the center of the enclosing circle, a radial line is drawn for every possible angle θ (according to image resolution) to intersect the outer contour of the trademark (Fig. 3a). The intersections are recorded as a function $r(\theta)$, where r is the distance from the center of the circle. If there is no intersection for some θ, then assign r=0 for that θ. Finally, let the wrap contour $r(\theta)$ become a closed and continuous curve in the polar plane, as illustrated in Fig. 3c.

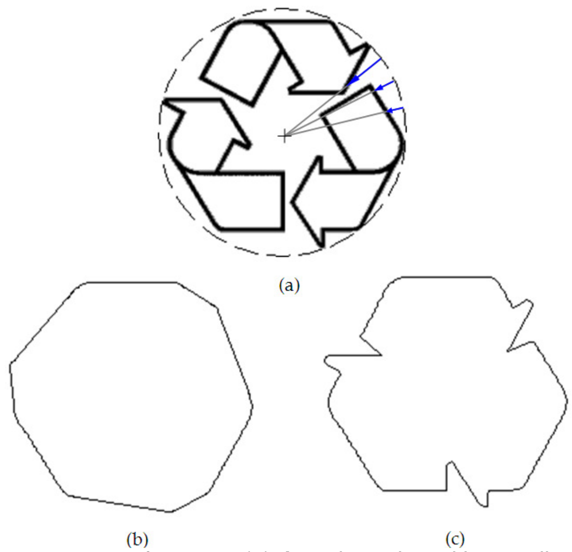


FIGURE 3: Using contours as features. (a) A trademark and its smallest enclosing circle, (b) The convex hull for that trademark, (c) The wrap contour for that trademark.

The wrap contour of a trademark is used in our scheme to help us define two new features called image compactness indices, CI_1 and CI_2. They can be regarded as mixed contour-based and region-based features, and are defined as:

$$\text{Compactness index 1 } (CI_1) = \frac{\text{Area of Wrap Contour}}{\text{Area of Smallest Enclosing Circle}} \tag{11}$$

$$\text{Compactness index 2 } (CI_2) = \frac{\text{Image Area}}{\text{Area of Wrap Contour}} \tag{12}$$

Note that in the above definition of CI_2, the "Image Area" is the foreground area of the trademark. However this area might change when we alter the color depth of the image. We will discuss this in the next section.

3.3 Four-Level Images

Generally, trademark designs are not limited to binary images. In fact, most modern trademarks use vivid colors to be visually conspicuous and attractive to consumers. On the other hand, a trademark normally uses only a scant palette of colors. This is because a trademark is not like a photograph. For trademarks, the simpler the design is, the quicker people tend to remember.

To compute the moments, a color image has to be transformed into a gray-scale image function, and traditionally a scale of 255 gray levels is used for picture-like images. For simple images as trademarks, some researchers suggested using only a two-level (binary) scale [11]. We have found, however, transforming trademarks into binary images sometimes incurs significant losses in the features. Fig. 4 shows examples of such losses in two trademark images, even though the thresholds used were computed by a sophisticated automated scheme [16]. As a compromise, we propose using four gray levels that correspond to levels 0, 85, 170, and 255, in a 0 to 255 grayscale. A trademark image is then transformed, with evenly spaced threshold values, to an image that contains only these four gray levels. By thus using more than two gray values, the loss of image features is effectively reduced. Yet we also take into account the fact that trademarks are less chromatically complex. Our experiments showed that by this scheme the total impression of a trademark image is satisfactorily preserved. Detailed examples of producing Zernike moments and about reconstructing such four-level trademark images from moments can be found in [19].

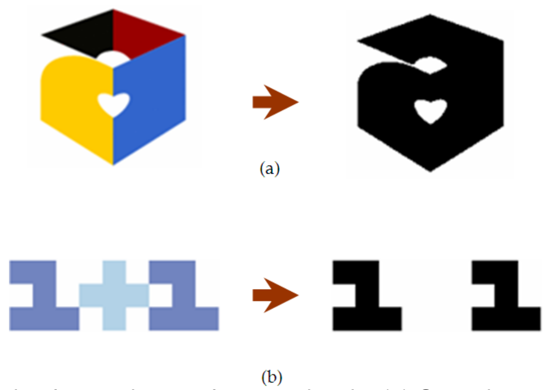


FIGURE 4: Losing boundaries from using too few gray levels. (a) Some internal boundaries are missing, (b) The plus sign is gone.

3.4 Exclusive Trademark Features

The feature values of an image are collectively used to identify the image. Thus, these features must be appropriately extracted. For trademarks, there are additional yet unique concerns. First, a trademark can be registered with a principal pattern and that pattern plus a variety of frames, e.g., circular, square, or polygonal shapes. Framed versions are regarded as the same as the original in trademark evaluation. Second, a chromatically altered image of a trademark pattern is also considered equivalent to the original one. For example, a dark-colored trademark on a light background is the same as the inversed, light trademark pattern on a dark background.

In this paper we highlight the quest of emulating the real trademark evaluation process. Hence, in extracting the features of one trademark we may need to consider not only the original design, but also its inversed image, and its core image. The core image refers to the central portion of the original trademark that lies within a smaller concentric circle whose radius is 2/3 of the enclosing circle. This core image is used to exclude the trademark frame and thus reduce the chances of misinterpretation. It is also noted that in real trademark retrieval the most important task is to detect confusingly similar designs, rather than to recover exactly like ones. Care has been taken in using features, so as to make our retrieval system neither under-restraining nor over-restraining. In this paper, we use Zernike moments of orders 0 to 19 as main region-based features for trademarks. The lower-order (0 to 12) moments represent the rough content of the trademark, whereas the higher-order moments delineate the minutia in the trademark image.

3.5 Weighting and Normalization

According to the trademark law, the degree of similarity between trademarks is the seriousness of confusion they cause for consumers. Obviously, this is a subjective judgment, and which may vary among different designs. Thus, we employ for our hybrid system a human-machine interface that allows users to feedback and to adjust the weighting of any feature type. Weighting means to prescribe significance for each feature group, which is among, say, m distinct image feature categories.

Combining the values from different feature categories requires a work called normalization. This process is to ensure that all image features possess influences of similar order. Now let P be an input trademark image, to be compared with the trademark database, which comprises n image entries Q_i, i=1 to n. Regarding the j^{th} feature category, which contains k values, the similarity between images P and Q_i is then defined as a Euclidean distance on the coordinate hyperspace [21], given by

$$\mathrm{dist}\left(P^j, Q_i^j\right) = \left[\sum_{s=1}^{k}\left({}_s p^j - {}_s q_i^j\right)^2\right]^{1/2} \tag{13}$$

where ${}_{s}p^{j}$ and ${}_{s}q_{i}^{j}$ are corresponding feature values of P and Q_i. Now for the j^{th} feature category, we scan through the entire database for the two extreme values:

$$\mathrm{Min}^{j}=\min_{i\le n}\left\{\mathrm{dist}\left(P^{j},Q_{i}^{j}\right)\right\}$$
$$\mathrm{Max}^{j}=\max_{i\le n}\left\{\mathrm{dist}\left(P^{j},Q_{i}^{j}\right)\right\} \tag{14}$$

The normalized Euclidean distance, meaning the normalized similarity between images P and Q_i regarding the j^{th} feature category, is then

$$\mathrm{Dist}\left(P^{j},Q_{i}^{j}\right)=\frac{\mathrm{dist}\left(P^{j},Q_{i}^{j}\right)-\mathrm{Min}^{j}}{\mathrm{Max}^{j}-\mathrm{Min}^{j}} \tag{15}$$

Adding the consideration of the weighting factor w^{j} for the j^{th} of a total of m feature categories, the gross similarity between images P and Q_i is then:

$$S_{G}\left(P,Q_{i}\right)=\sum_{j=1}^{m}w^{j}\cdot\mathrm{Dist}\left(P^{j},Q_{i}^{j}\right) \tag{16}$$

4. EXPERIMENTS & RESULTS

In this paper, 2020 trademark images were collected from various sources and compiled to form the experimental database. These images had been of different sizes, colors, and complexity. To begin with, since we wanted to emphasize human perspectives, we recruited ten adult volunteers, aged from 20 to 30, three among them female, and asked them to pick from the sample database some groups of similar images visible to their naked eyes. Listed in Table 1 are the five groups of similar trademarks hand-picked by the volunteers. These human culled images were to be used as benchmarks as we tested our hybrid scheme.

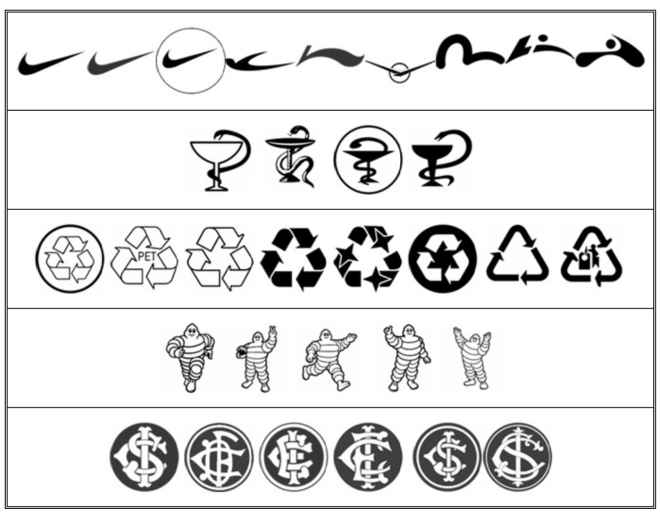


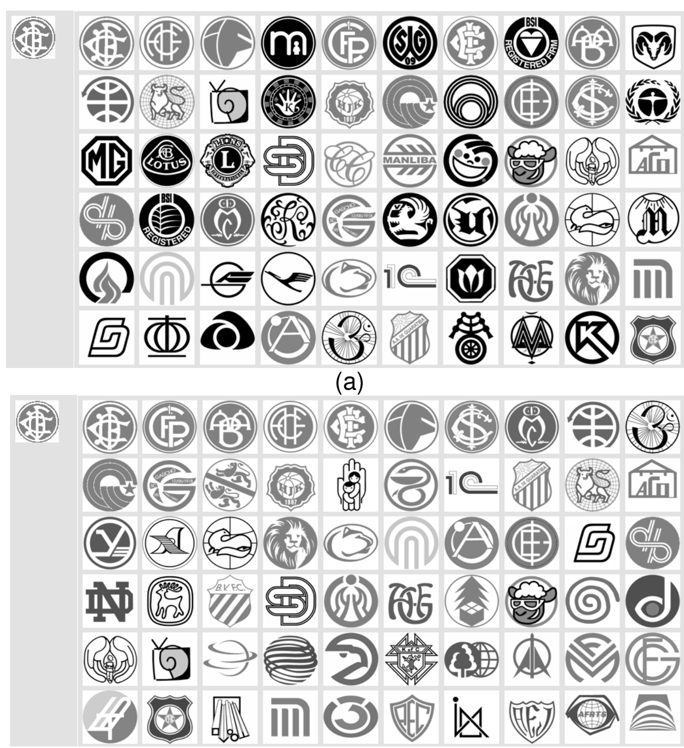
Table 1: Human-Picked Groups of Similar Trademark Patterns

Note that each group in Table 1 actually stands for a certain aspect of trademark identification. The images in the first group, "Nike," all contain an abstract notion of a checkmark. In the second group, "Goblet," either the snake or the goblet can deform its shape or change its color. The third group, "Recycle," is an example of inversed and framed trademark images. The fourth group involves various postures of the "Michelin" Man. These figures are line drawings containing no solid areas. The last group is "Monogram." Monograms are commonly used as trademarks in western countries and are also common victims of infringements.

As for the specific image features used, in each of our experiments we chose either to use only lower-order Zernike moments or to incorporate higher-order moments as well, according to the character of the input image. A collection of eight weighting factors $w^j = \{w^1, w^2, \ldots w^8\}$ was used in the experiments. Their meanings are described as follows. w^1: compare the whole input image with whole images from the database. w^2: compare core images only. w^3: compare the core of the input image with the whole images from the database. This is to remove the frame of the input image. w^4: compare the inversed core of the input with the whole database images. w^5: compare CI_1's of whole images. w^6: compare CI_2's of whole images. Lastly, w^7 and w^8: compare CI_1's and CI_2's of the core images, respectively.

We first took a "Monogram" trademark from Table 1 as the input image, and inspected exclusively the effects of using four-gray-level Zernike moments. Shown in Fig. 5 are the results from the conventional binary scheme (Fig. 5a) and from our alternative four-level scheme (Fig. 5b). The input image is shown on the left. Here, all Zernike moments of orders 0 to 19 for the whole images are used as features (w^j = {1,0,0,0,0,0,0,0}). For each scheme we list, out of 2,020 database images, 60 images closest to the input. The order is from left to right and from top to down. Since our trademark retrieval system was to provide a list of ranked similar entries for human examiners to inspect, a better scheme was the one that produced a better candidate list. We see that Fig. 5b gathers more monogram-like images in its top rows, and it contains fewer samples that are too different to human's eye.

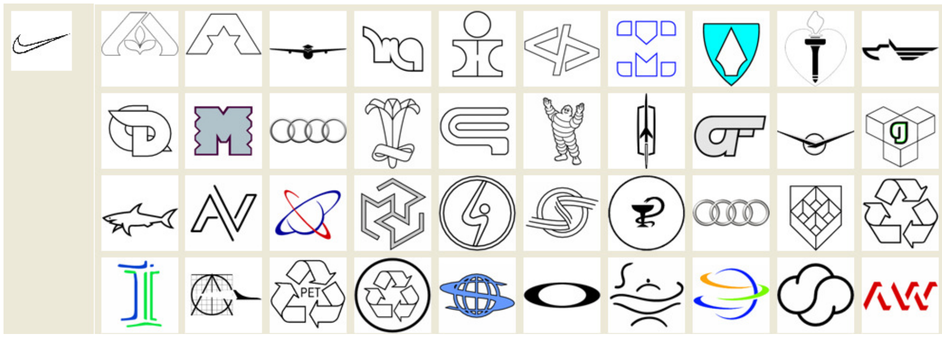
We also tested the effects of the proposed wrap contour concept and the related compactness index feature. We used an outlined "Nike" image as the input. Although to human's eye an outlined trademark gives similar rough impression as the solid one, we found that it literally confounded all the region-based methods. Compared in Fig. 6 are results from Hu's moment invariants (Fig. 6a), conventional binary Zernike moments (Fig. 6b), Kim's improved Zernike moments (Fig. 6c), and our first compactness index CI_1 (Fig. 6d, and w^j = {0,0,0,0,1,0,0,0}). In this experiment, only the compactness index scheme is capable of yielding meaningful outputs.



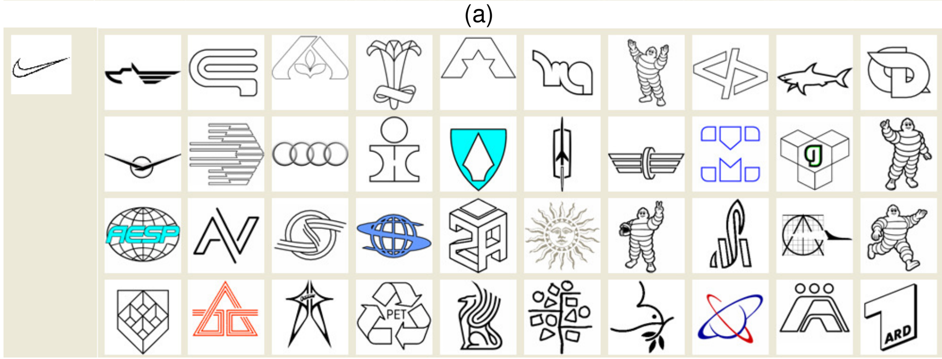
(a)

(b)

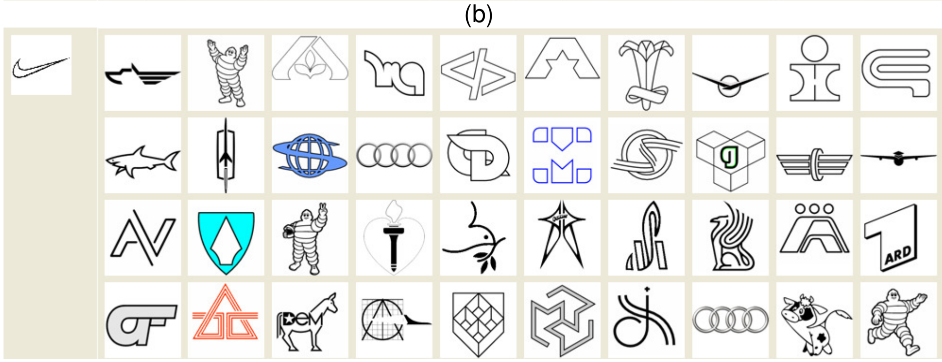
FIGURE 5: Using different gray-levels with Zernike moments. (a) The top output list from the two-level scheme, (b) The top output list from the four-level scheme.



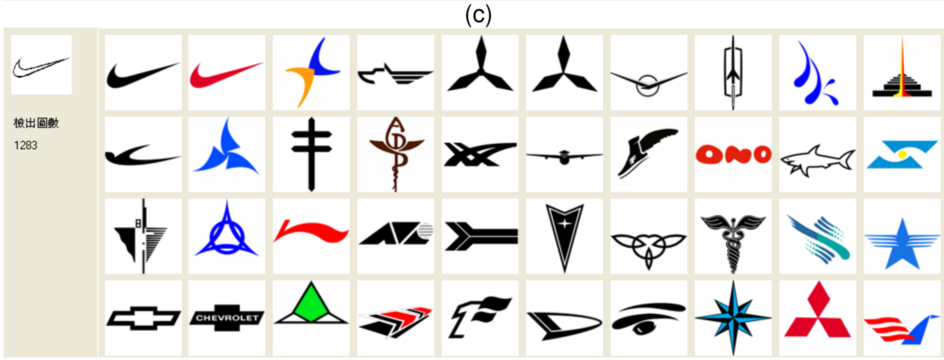
(a)



(b)



(c)



(d)

FIGURE 6: Search results for the outlined Nike. (a) Hu's moment invariants, (b) Traditional Zernike moments, (c) Kim's modified Zernike, (d) Our compactness index CI_1.

Then we tried using more image features to improve the search efficiency. Again in Fig. 7 we put Hu's method (Fig. 7a), the original Zernike (Fig. 7b), Kim's modified Zernike (Fig. 7c) and our scheme in comparison. A regular "Nike" logo is used as the input. Due to its simple construction, only lower-order moments are required. With the above-mentioned w^j notation, Fig. 7d shows results of the vector {1,4,0,0,0,0,0,0}, denoting a double four-level Zernike search stressing the importance of the core image. Fig. 7e corresponds to the weighting vector {0,0,0,0,1,0,4,1}, showing the effects of combining the two compactness indices, also emphasizing the core. The weightings of Fig. 7f are {1,4,0,0,1,0,4,1}, equivalent to the sum of the weightings of Fig. 7d and Fig. 7e. From these figures we see that while Fig. 7d already produces a better listing than Figs. 7a to 7c, Fig. 7f produces the best result of all. All but one trademark in the "Nike" group in Table 1 are listed in the top 20 (top 1%, out of 2,020) candidates of Fig. 7f.

Similar satisfactory results of the proposed scheme were obtained for a monogram trademark. Again, Figs. 8a to 8c are results from the three existing schemes, and Figs. 8d to 8f are results from three variations of the proposed scheme. The weighting vectors for Figs. 8d to 8f are {5,5,0,0,0,0,0,0}, {0,0,0,0,10,1,0,0}, and {5,5,0,0,10,1,0,0}, respectively. For monogram images we do not need to emphasize the core. Note that different image type may require different weighting factor assortments. In the previous two examples we have not used w^3 and w^4. They are by no means less useful. Many additional examples involving framed or inversed trademarks or even various kinds of noise can be found in [16].

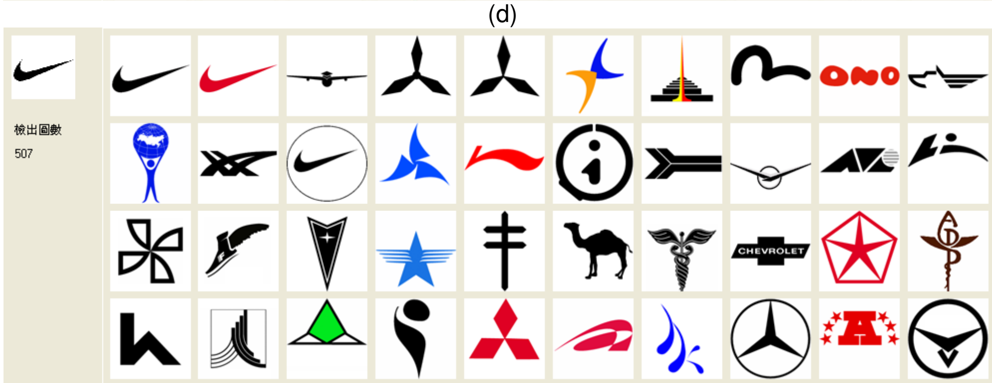
(a)

(b)

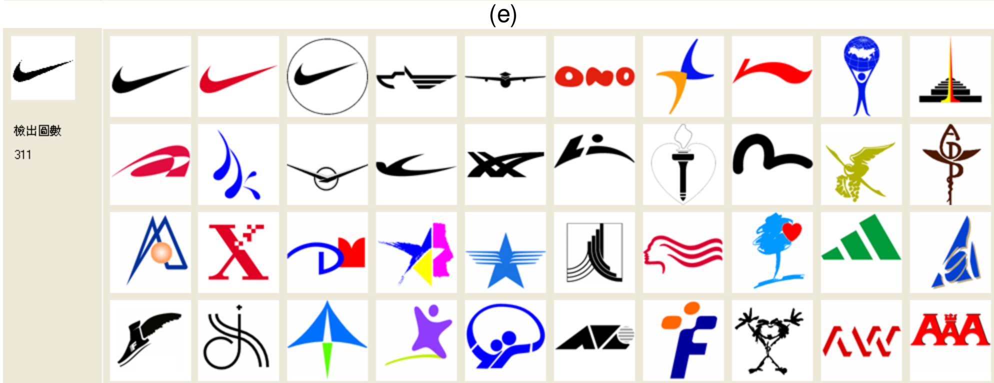
(c)

FIGURE 7: Comparison between schemes. (a) Hu's moment invariants, (b) The original Zernike, (c) Kim's modified Zernike, (d) Our double four-level Zernike, (e) Our Compactness indices, (f) Combining double Zernike and Compactness indices.

(d)



(e)



(f)

FIGURE 7: Comparison between schemes. (*Continued*)

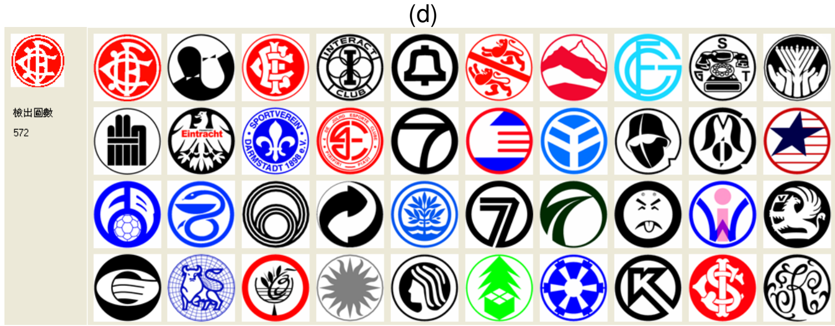
(a)

FIGURE 8: Another comparison between schemes. (a) Hu's moment invariants, (b) Original Zernike, (c) Kim's modified Zernike, (d) Our double four-level Zernike, (e) Our Compactness indices, (f) Combining double Zernike and Compactness indices.

(b)

(c)

(d)



(e)

(f)

FIGURE 8: Another comparison between schemes. (a) Hu's moment invariants, (b) Original Zernike, (c) Kim's modified Zernike, (d) Our double four-level Zernike, (e) Our Compactness indices, (f) Combining double Zernike and Compactness indices. (*Continued*)

On the whole, for the five groups of trademark images in Table 1, the performance of our method is summarized as follows. In average, both our double Zernike schemes (as Figs. 7d and 8d) and CI-only schemes (as Fig. 7e and 8e) got roughly 60% of human-picked images in their top 40 (top 2%) image listings. This was no better than Kim's modified Zernike (65% in average) and more or less equivalent to traditional Zernike (58% in average). But it is noted that since our double-Zernike and CI's are very different image features, the combination of the two (as Figs. 7f and 8f) yielded very good (95% in average) results. But with what kind of trademark did we show greatest advantage? We will say in dealing with the "Goblet" images, our method outperformed hugely other methods (100% vs. 25%). The next were the "Monogram" and the "Nike" trademarks. For the "Recycle" and the "Michelin" groups, only small improvement was achieved. This was because both traditional Zernike or Kim's modified Zernike already treated them quite well. Hybrid system may not be able to retrieve distributed shape or image like shape. Because we were concentrated only trademark retrieval and this kind of shape is not familiar as a trademark. However, our method was still a competitive choice because we could deal with outlined, framed, and inversed variation of the trademarks.

In all above experiments, a 64×64 image size was used for the input and for all the database images. An input image must go through a series of preprocessing processes such as conversion to four gray-level, obtaining the smallest enclosing circle, and proper geometric transformation. All image features are then extracted, paired with their weighting factors and compared with stored data in the database. The whole process for a 200×200 input image takes about 8.5 seconds on an Intel Celeron 2.66 GHz computer with 512 Mb RAM. Lastly, we think it is interesting to mention that when we doubled the resolution of all images (input and database) to 128×128, only slight change was observed in the results.

5. CONCLUSIONS

We have presented in this paper the framework of a hybrid trademark retrieval system. Surveys of human responses to different trademark aspects were conducted and the results were used as benchmarks for the system. The task of the computer was only to provide, according to the benchmarks, a good ranked list of similar candidates. Both region-based image features (i.e., four-gray-level Zernike moments) and contour-base features (the proposed compactness indices) were used. A simplified method for finding the smallest enclosing circle was also presented. Features of different categories were conjoined via a weighting scheme particularly useful for dealing with framed or inversed trademark variations. Experiments have verified that, when human viewpoints obtained from consumer surveys were used as standards, this hybrid scheme performed considerably better than some existing retrieval methods.

6. ACKNOWLEDGEMENTS

The authors are greatly indebted to Director Jiou-Ru Chang of Taiwan Intellectual Property Office (Kaohsiung Branch) for his kind assistance and precious comments offered to this research.

7. REFERENCES

[1] United States Patent and Trademark Office, "2007 Design Code Updates", *USPTO Design Search Code Manual*, http://tess2.uspto.gov/tmdb/dscm/dsc_ap.htm, accessed 2010

[2] H. Freeman, "*On the Encoding of Arbitrary Geometric Configurations*". IRE Trans. Electronic Computing, 10: 260-268, 1961

[3] H. L. Peng and S. Y. Chen, "*Trademark Shape Recognition Using Closed Contours*". Pattern Recognition Letters, 18: 791-803, 1997

[4] C. C. Huang and I. Her, "*Homomorphic Graph Matching of Articulated Objects by An Integrated Recognition Scheme*". Expert Systems with Applications, 31: 116-129, 2006

[5] P. Y. Yin and C. C. Yeh, "*Content-based Retrieval from Trademark Databases*". Pattern Recognition Letters, 23: 113-126, 2002

[6] M. K. Hu, "*Visual Pattern Recognition by Moment Invariants*". IRE Trans. Information Theory, pp. 179-187, 1962

[7] J. L. Shih and L. H. Chen, "*A New System for Trademark Segmentation and Retrieval*". Image and Vision Computing, 19: 1011-1018, 2001.

[8] G. Ciocca and R. Schettini, "*Content-based Similarity Retrieval of Trademarks Using Relevance Feedback*". Pattern Recognition, 34: 1639-1655, 2001

[9] Y. S. Kim and W. Y. Kim, "*Content-based Trademark Retrieval System Using a Visually Salient Feature*". Image and Vision Computing, 16: 931-939, 1998

[10] H. K. Kim, J. D. Kim, D. G. Sim and D. I. Oh, "*A Modified Zernike Moment Shape Descriptor Invariant to Translation, Rotation and Scale for Similarity-based Image Retrieval*". IEEE International Conference on Multimedia and Expo, 1: 307-310, 2000

[11] N. K. Kamila, S. Mahapatra and S. Nanda, "*Invariance Image Analysis Using Modified Zernike Moments*". Pattern Recognition Letters, 26: 747-753, 2005

[12] United States Patent and Trademark Office, *Trademark Manual of Examining Procedure (TMEP)*, 5th ed., USPTO, 2007

[13] United States Patent and Trademark Office, *U.S. Trademark Law*, USPTO, 2010

[14] Taiwan Intellectual Property Office, *Taiwan Trademark Law*, TIPO, 2003

[15] J. R. Chang, Director, Taiwan Intellectual Property Office, Kaohsiung Branch, Interview by author, 2006

[16] H. K. Hung, "*A Computer-assisted Trademark Retrieval System with Zernike Moment and Image Compactness Indices*". Master thesis, Electro-Mechanical Engineering Department, National Sun Yat-sen University, Taiwan, 2006

[17] S. D. Lin, S. C. Shie, W. S. Chen, B. Y. Shu, X. L. Yang and Y. L. Su, "*Trademark Image Retrieval by Distance-Angle Pair-wise Histogram*". International J. Imaging Systems and Technology, 15: 103-113, 2005

[18] M. Teague, "*Image Analysis via the General Theory of Moments*". J. Optical Society of America, 70: 920-930, 1980

[19] M. de Berg, M. van Kreveld, M. Overmars and O. Schwarzkopf, *Computational Geometry: Algorithms and Applications*, New York : Springer, (2000)

[20] I. Her, "*Geometric Transformations on the Hexagonal Grid*". IEEE Trans. Image Processing, 4: 1213-1222, 1995

[21] I. Her, "*A Symmetrical Coordinate Frame on the Hexagonal Grid for Computer Graphics and Vision*". ASME J. Mechanical Design, 115: 447-449, 1993.

[22] J. Schietse, J. Eakins & R. Veltkamp, "*Practice and Challenges in Trademark Image Retrieval*". ACM International Conference on Image and Video Retrieval, Amsterdam, The Netherlands, 2007