Gestalt-based Feature Similarity Measure in Trademark Database [1]

Hui Jiang[*+] & Chong-Wah Ngo[*] & Hung-Khoon Tan[*]

[*]*Department of Computer Science, City University of Hong Kong*

[+]*Institute for Computational and Mathematical Engineering, Stanford University*

Abstract

Motivated by the studies in Gestalt principle, this paper describes a novel approach on the adaptive selection of visual features for trademark retrieval. We consider five kinds of visual saliencies: symmetry, continuity, proximity, parallelism and closure property. The first saliency is based on Zernike moments, while the others are modeled by geometric elements extracted illusively as a whole from a trademark. Given a query trademark, we adaptively determine the features appropriate for retrieval by investigating its visual saliencies. We show that in most cases, either geometric or symmetric features, can give us good enough accuracy. To measure the similarity of geometric elements, we propose a maximum weighted bipartite graph (WBG) matching algorithm under transformation sets which is found to be both effective and efficient for retrieval.

Key words: Trademark image retrieval, Gestalt Principle, Bipartite graph matching under transformation sets

[1] Chong-Wah Ngo is the corresponding author. Please contact him for any enquiry.
Email address: `cwngo@cs.cityu.edu.hk` (Chong-Wah Ngo[*]).

1 Introduction

To date, despite the numerous efforts in content-based image retrieval (CBIR), finding the best shape features and the best way of matching features for image retrieval remains challenging. One of core issues is in formulating a general-purpose shape similarity measurement that guarantees good retrieval performance, and with the baseline that the retrieved similar items should be consistent with human visual perception. Recently, Gestalt principle [1] is taken into account by researchers for the perceptual segmentation and grouping of shape features. Gestalt principle is one of the earliest studies conducted by a group of psychologists to model shape perception in the early 19th century. A number of principles have been experimentally studied and derived to govern the grouping of shape features.

Perception, in general, is viewed as an active process of organization, construction and analysis. Gestalt principle emphasizes the *wholistic* nature, where recognition is inferred more by the properties of an image as a whole, rather than parts, during visual perception. This is considered different from traditional pattern recognition where recognition is achieved by accounting image features of parts and their combinations. Take the image in Figure 3 as an example. Gestalt principle considers white regions (areas enclosed by five group of parallel lines) as a whole as the significant property rather than the shape of six independent black regions.

In this paper, we investigate the flexibility of applying Gestalt principle in trademark database since trademarks are images that usually contain rich *abstract* geometric features that are appropriate for the modeling of Gestalt principle. In particular, we focus on five *wholistic* properties: symmetry, continuity, proximity, parallelism and closure derived in Gestalt principles. The first property is described by Zernike moments, while the others are extracted and represented *illusively* [1] as a whole by our proposed geometrical features under the weighted bipartite graph (WBG) framework [2–6]. These five wholistic properties, in general, are not effective if they are jointly integrated in a linear weighted combination way for retrieval. To solve this problem, we propose a novel adaptive selection procedure of wholistic properties, which depends on the nature of a query image.

Gestalt principle has been investigated in [7–13] for trademark retrieval, however, only a subset of wholistic properties is utilized. No study has yet been carried out on how to systematically select and match these properties for trademark retrieval. In [7–13], clustering algorithms are employed to group semantically meaningful Gestalt components. After clustering, non-geometric features such as aspect ratio, circularity and right-angleness are extracted from each cluster for retrieval. Nevertheless, as pinpointed in [13], the incorrect clustering of elements is the major drawback that affects the retrieval accuracy. In this paper, instead of adopting clustering based approach, we encode directly the extracted geometric elements led by Gestalt principle in WBG for partial matching under a set of allowable transformations. Since our approach matches geometric elements as a whole directly, it leads to a more reliable framework for trademark similarity measurement.

[1] We use the word "illusively" to describe the nature of Gestalt principles and the motivation of our approach: Human always group low-level geometrical elements illusively as one or several "complete elements", even though a complete element is actually not connected and formed by several broken segments. For example, in Figure 5b, the inner circle is broken into three arcs, but our approach can detect them as a "whole" (a complete circle) which mimics the human perceptual organization.

2 Related Works

Numerous approaches have been proposed for trademark image retrieval. Representative works include [7–17]. As other CBIR problems, most approaches in trademark image retrieval consist of two major components: feature extraction and similarity measurement. The choice of features will normally affect the use of similarity measurement. In this section, we focus our attention on how they derive the shape features for similarity retrieval.

In the current literature, various visual features have been explored for trademark retrieval. The features adopted most frequently are: edge direction histogram [16, 18], moments [16–18] and shape descriptors [7–12, 19, 20]. Some of these features (e.g., edge direction histogram and moments) contain no geometric information. They are global and statistical in some sense. For convenience, we call them non-geometric features. Since no geometrical information is encoded in these features, two images with similar features can be very different (See Figure 1 for example). Similarly, two images with similar shape may have considerably different global and statistical features. One example is given in Figure 2, the four trademarks are perceptually similar to each other but their moments are very different. In [16], Jain and Vailaya proposed an image filling approach to solve this problem. However, this method can only handle the cases of Figure 2 (a) and (b). Because the moments do not fit human perception very well, recent approaches in [7–12, 20] consider only the edge points of regions for trademark retrieval.

Although the non-geometric features have the weakness introduced above, they have the advantages that they are easy to compute and compare. Most importantly, the weakness may be attenuated by integrating multiple features (a straightforward observation from figures 1 and 2). Nevertheless, the experimental results conducted

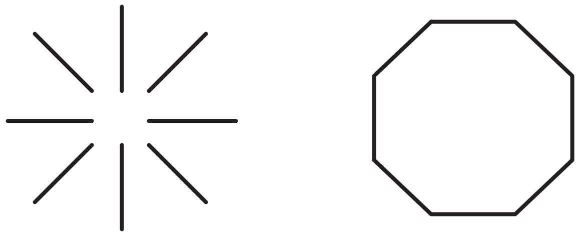

Fig. 1. An example showing the weakness of the non-geometric features: These two trademarks have similar edge direction histograms, but they are quite different.

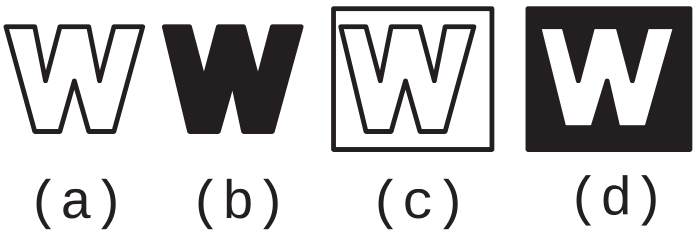

Fig. 2. An example showing the weakness of moments.

by Eakin et al in [12] indicated that it is not always true that the combination of multiple features can give better results than using them on their own. The key issue is how to effectively integrate multiple features, which is not a trivial problem. In [16], Jain and Vailaya employed a two-level hierarchical system: in the first stage, edge direction histograms and moments were used to rapidly filter the database; in the second stage, deformable template matching was used for final similarity ranking. The reason that they used such a framework is: edge direction histograms and moments are non-geometrical features, they are quick but coarse; deformable template matching takes into account the geometric information, it is accurate but slow. The experimental results in [16] showed that moments are not robust to trademarks with line drawings, and the deformable matching is not effective for the trademarks with many details in line drawings and holes. Their results are improved by filling-in the holes in the trademarks, but the major drawback is the non-utilization of information in holes. For example, the trademark in Figure 2 (d) becomes a square after image filling, the shape information “W” is missed after filling.

Instead of extracting the global features as a whole from the images as in [16, 17], there is a more general scheme in [7–12, 19, 20]: decompose the images into several components, and then use non-geometric features to encode each component. Decomposition of trademark images is a hard problem. In [19–22], trademarks are segmented into regions based on the pixel connectivity and the shape features are extracted from each region for retrieval. The segmentation by pixel connectivity, nevertheless, does not always reflect the segmentation by human. Consider the trademark shown in Figure 3, the shape of this trademark is inferred as a whole from the image, rather than from each individual region. To segment a trademark into perceptually meaningful components, Gestalt principle [1] is taken into account in [7–13]. In this principle, grouping is based on the proximity, similarity, symmetry and good continuation of edge points (rather than regions). For example, the approaches in [7–12] utilized the co-linearity, parallelism and good line continuation properties of edge points to segment the trademark in Figure 3 into five groups of parallel lines. Close figures (or a set of segments which lie on a closed loop) were further extracted from the so-called “Gestalt images” (i.e., images represented continuous lines, arcs, etc) by clustering algorithms [7–12]. Shape features are then extracted from each closed figure for retrieval.

Fig. 3. An example showing the weakness of image segmentation.

Indeed, twenty years ago Lowe [23] has made use of the Gestalt laws to perform perceptual organization and visual recognition for his SCERPO system. In his work, the significance of a grouping is determined by its non-accidentalness.

Certain image relations are carriers of statistical information indicating that they are non-accidental in origin. He proposed a probabilistic measure to quantify the degree of non-accidentalness, which forms the basis for assigning degrees of significance. Relations such as proximity, co-linearity and symmetry are of great significance since they remain invariant over positions and a wide range of viewpoints. Lowe also showed how view-point invariance can be used for interpretations of the three-space inference from the two-dimensional image groupings. Besides, using perceptual groupings helps to reduce the size of the search space over viewpoints and object parameters especially when the complexity of the relations between subparts of a grouping increases, thus improving the saliency of the grouping.

While the extraction of close figures based on Gestalt principle has shown advantages over the pixel connectivity approaches, there are limitations, for machine vision, on how to correctly perform perceptually meaningful clustering [7–13]. Consider the image in Figure 4, it is very difficult to tell whether (a) should be clustered into four triangles or just one polygon. If the close figures extracted from (a) are two squares, trademarks that consist of four triangles such as (b) will not be retrieved. On the other hand, if the close figures extracted from (a) are four triangles, trademarks consist of two squares such as (c) will not be retrieved. In [13], several experiments have been conducted to study the machine segmentation of trademark images by Gestalt principles. They compared the results with the segmentation of human subjects, and found that the agreements between machine and human segmentation are indeed limited. They also pinpointed that the major drawback of their system, ARTISAN [10–13] (which is regarded as one of the most comprehensive trademark retrieval system in the current literature), is the incorrect clustering of perceptually meaningful elements.

The remaining of this paper is organized as follows. In Section 3, we begin by

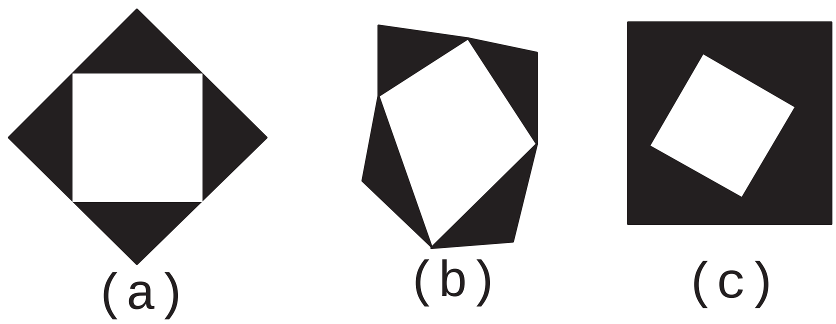

Fig. 4. An example showing the weakness of close figure.

describing the representation and extraction of the proposed geometric features. Their relationships with Gestalt principle are then outlined. To compare the geometric features, we also propose the novel maximum WBG matching algorithm under transformation sets for similarity measurement. In Section 4, we first introduce Zernike moments for incorporating symmetry property. Then, a procedure for the adaptive selection of geometrical and symmetric features is presented. Section 5 presents our major experimental results. Section 6 further discusses the empirical performance of our approach from both theoretical and practical aspects. Finally, Section 7 concludes our proposed works.

3 Retrieval with Geometric Features

Like most existing image retrieval systems, our approach consists of two major parts: feature exaction and similarity measurement. The features are composed of geometric features that can mimic Gestalt principle, while the similarity measurement is based on the maximum WBG matching. Because the geometric features we consider are not transform invariant, an iterative framework is proposed to simultaneously match features and estimate transformation. To speed up the retrieval, a hierarchical framework is also presented for the rapid filtering of irrelevant candidates.

3.1 Feature Extraction

We consider different kinds of Gestalt elements which include: lines, circles (arcs), parallel lines, concentric circles (arcs), and polygons. We employ Hough transform [24, 25] for primitive (line, circle and arc) detection in trademarks owing to its simplicity and robust, compared with other approaches [8, 10, 11]. Initially, traditional edge detection algorithm is performed and the Hough transform [24, 25] is employed to extract the Gestalt elements like lines, circles and arcs. Then we group lines which have almost the same directions and positions to form parallel lines. We further group circles and arcs which have almost the same centers to form concentric circles and arcs. Notice that in these processes, Gestalt principles such as continuity, proximity, and parallelism are utilized. To detect polygons, proximity principle and closure properties are used as in [7–12]. Line segments whose end-to-end distances is near are grouped together to form a closed polygon. However, only significant polygons like triangles, squares and rectangles are considered.

In our approach, Hough transform is implemented based upon [25], which requires the input of several empirical parameters [2] . These parameters perform satisfactorily for most trademarks regardless of some exceptional cases. Basically, edge detection in trademarks is less error-prone since trademarks are normally binary images with sharp and continuous edges. Thus, Hough transform is in general effective in detecting primitives of trademarks where geometric elements such as lines and circles are mostly distinctive. Hough transform, nevertheless, suffers from the mass requirement of computational cost. Fortunately, trademarks are in smaller size than general images. For instance, the trademarks in our database are in the resolution

[2] These parameters include the number of accumulator cells, and the distance between disconnected pixels identified during traversal of the set of pixels corresponding to a cell.

of 100×100 pixels. This feature indeed alleviates the need of heavy computational load.

Four trademark examples showing the Gestalt elements detected by our approach are given in Figure 5. In (a), three line segments and three arcs are extracted, and a triangle is further detected. In (b), six lines segments and two circles are initially extracted, and as in [7–12], the Gestalt elements with parallelism properties (e.g., parallel lines and concentric circles) are then detected. Nevertheless, unlike [7–12], no closed figure is extracted. We simply index the parameters of Gestalt elements (e.g., center position and radius of a circle) under WBG framework for similarity measurement. From these examples we can see that the major advantage of Hough transform is its robustness in handling occlusion and illusion, which is corresponding to the continuity and proximity properties in Gestalt principle. For instance, in (b), the inner circle is implicitly represented as a continuous circle, rather than three arc segments. Similarly, in (c), four line segments and an arc are extracted. In (d), a group of parallel lines and a group of concentric circles are detected. Figure 6 summarizes the geometric features and the corresponding feature extraction methods that we use in this paper. The relationships between the geometric features and Gestalt principles are also given.

3.2 Similarity Measurement by maximum WBG matching

The similarity of trademarks can be measured directly from the maximum WBG matching. Let $F_1 = \{f_{11}, f_{12}, ..., f_{1m}\}$ and $F_2 = \{f_{21}, f_{22}, ..., f_{2n}\}$, respectively, as the Gestalt features of two given trademarks T_1 and T_2. Each attribute f_{ij} represents a Gestalt element (e.g., a line segment, an arc or a circle) of trademark T_i. To compute the similarity between T_1 and T_2, we build a weighted complete bipartite graph $G =< V_1, V_2, E >$, where V_1 has m nodes $v_{11}, v_{12}, ...v_{1m}$ corresponding to

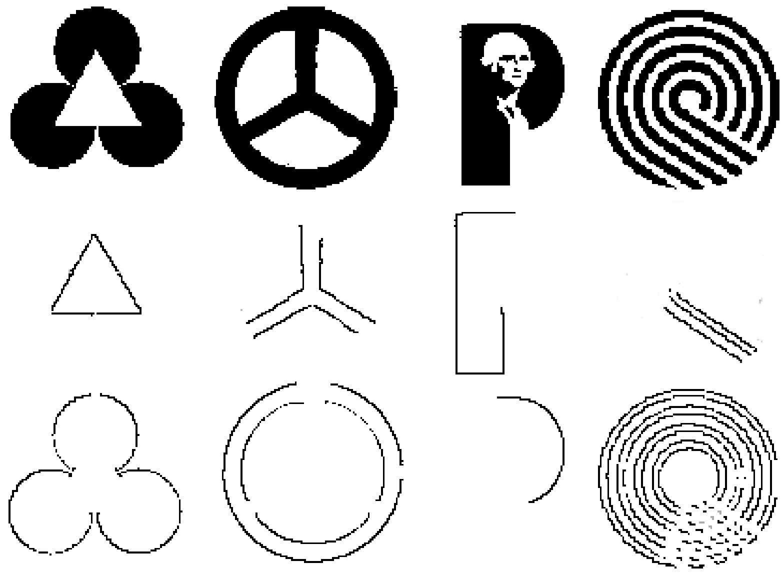

Fig. 5. Four examples showing the Gestalt elements detected from trademarks.

Geometric Elements	Method for Extraction	Gestalt Principle
Lines, Circles, Arcs	**Hough Transform (line segments or arcs which are co-linear and close enough are grouped)**	**Continuity, Proximity**
Parallel Lines, Concentric Circles (Arcs)	**Grouping (line segments or arcs which are parallel and close enough are grouped)**	**Parallelism, Proximity**
Polygons	**Grouping (line segments which are end-to-end close enough are grouped)**	**Proximity, Closure**

Fig. 6. Geometric features

f_{1i}, and similarly V_2 has n nodes corresponding to f_{2i}. For each node pair $< u, v >$, $u \in V_1$, $v \in V_2$, there is an edge between u and v. The weight on each edge represents the similarity between two Gestalt elements. The maximum weighted bipartite graph matching algorithm, or specifically Kuhn-Munkres algorithm [2–6], is employed to match the F_1 and F_2 of two trademarks.

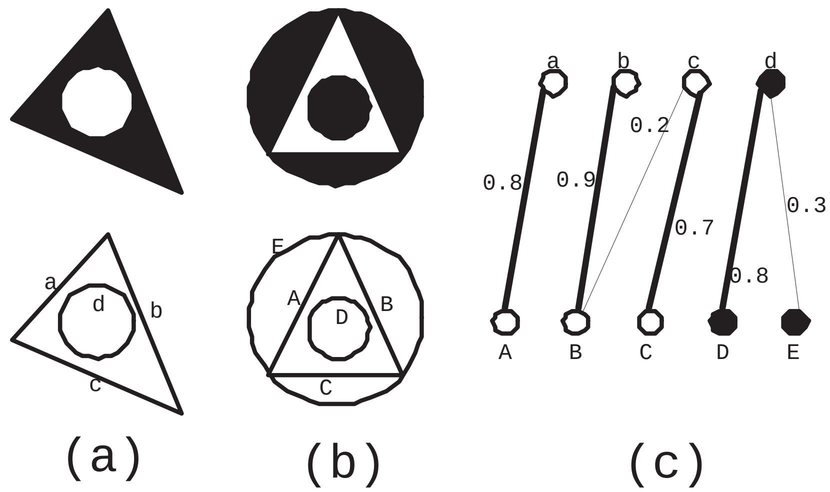

Fig. 7. An example showing the maximum WBG matching algorithm.

One simplified example (higher level Gestalt elements such as parallel lines, concentric circles and polygons are ignored) is illustrated in Figure 7. The trademark (a) has four elements (three line segments a, b, c and a circle d), while the trademark (b) has five elements (three line segments A, B and C, and two circles D and E). Their bipartite graph G is given in (c). It has nine nodes. Four nodes on the top represent the four Gestalt elements of trademark (a), while five nodes on the bottom represent the five elements of (b). The solid nodes represent the circles. For simplicity, we omit the edges whose weights are zero or almost zero (e.g., in the cases when a circle matches a line segment the weight are zero, and in the cases when two lines which are quite different matched to each other such as A and c, A and b, B and a, etc, the weights are very small) in this figure. Thick edges represent the maximum matching of two Gestalt elements.

One may argue that due to noise effects, there may be some extra line or curve segments in the image to be matched, and the maximum matching cannot always find the correct matching. Indeed, robustness is always related to some level of noises. When the noises become stronger and stronger, no system can be "robust" all the time. From the experiments we find that the maximum matching framework can tolerate a certain level of noises. One reason is that the outliers seldom overcome the true correspondence in a weighted bipartite matching even in the cases when the numbers of nodes in the two graphs are not equal.

The similarity of two trademarks T_1 and T_2 are then computed as

$$Sim_{BG}(T_1, T_2) = \frac{\sum_{u \in T_1} \sum_{v \in T_2} MW(u, v)}{\max(|T_1|, |T_2|)}$$

where $MW(u, v)$ is the edge weight between u and v after maximum matching, and $|T_i|$ is the number of Gestalt elements in T_i. $MW(u, v)$ in G is based on the similarity of u and v. For instance, given two line segments u and v, their similarity can be measured as

$$MW(u, v) = \sum_{i=1}^{3} w_i K_i(u, v)$$

where K_1 is a linear function of the distance d between the centers of u and v. $K_1(u, v) = 1$ when $d = 0$. $K_1(u, v) = 0$ when d is greater than a threshold, which depends upon the size of the trademark. The function $K_2 = \cos(\alpha)$, where α is the acute angle formed by u and v. K_3 is another linear function of length difference between two lines. The parameter w_i is a weighting factor for each function and $w_1 + w_2 + w_3 = 1$. When u and v are identical in terms of positions, lengths and directions, the value of $MW(u, v) = 1$. The similarity measurement between arcs or circles is similar, except that the geometrical parameters that are taken into account are: distance between two centers, difference of radiuses, difference of the arc lengths, overlapping of arc angles, and so on. When two compared elements belong to different geometric elements (e.g., a circle and a line), the similarity is set to zero.

In our current approach, higher level Gestalt elements such as parallel lines, concentric circles and polygons are implemented in the similar way. Each of these elements is represented by a node in the bipartite graph, and the weight between two elements of the same type is calculated with the geometric similarity between them.

3.3 Maximum WBG Matching under Transformation Sets

The Gestalt elements we consider are not rotational, translational and scale invariant. To handle this problem, we propose an iterative approach for maximum WBG matching under transformation sets as follows. The goal is to find a transformation that maximizes the weight of maximum WBG matching by given a set of allowable transformations.

Maximum WBG matching under transformation sets, intuitively, is an optimization problem that can be formulated through iterative maximization as follows:

$$Sim^{(k)} = \sum_{u \in T_1} \sum_{v \in T_2} MW(u, \Im^{(k)}(v)) \tag{1}$$

$$\Im^{(k+1)} = \arg\max_{\Im \in \Omega} \sum_{u \in T_1} \sum_{v \in T_2} MW^{(k)}(u, \Im(v)) \tag{2}$$

where $Sim^{(k)}$ and $\Im^{(k)}$ are the optimal matching and transformation respectively at step k, and Ω is the allowable transformation sets. The transformation begins with an initial transformation $\Im^{(0)}$. The Equation 1 and Equation 2 alternate between finding an optimal matching and an optimal transformation. This iterative algorithm is inspired from the iterative closest points (ICP) algorithm in [26] and the FT (flow transformation) algorithm in [27].

To find the optimal matching in Equation 1, we employ the maximum WBG matching algorithm described in Section 3.2. To compute the optimal transformation in Equation 2, we employ a gradient descent search. We run the iteration several times (5 times in our implementation) with multiple randomly generated initial transformations $\Im^{(0)}$ and keep the best result.

Apparently, the above iteration is always convergent. The reason is that, at each step of the iteration, the value of $Sim^{(k)}$ is non-decreasing and it has an obvious upper bound, *i.e.,* the number of edges of the bipartite graph. In our experiments, the iteration converges quickly. In most cases it converges within three steps. We compare our approach with brute force search. The results indicate that our approach has similar effectiveness as brute force algorithm, however, with significant improvement in speed efficiency.

3.4 Hierarchical Retrieval

To further speed up the retrieval, as in [16], we employ a two-stage hierarchical framework. In the first stage, edge direction histograms (EDH) is used to rapidly screen out potential candidates. In the second stage, we use both EDH and the proposed maximum WBG matching for similarity ranking. Unlike [16], we only choose EDH for filtering, rather than both EDH and moments. The reason is that EDH is efficient in filtering false matches even though some false positives (as shown in Figure 1) are included. Moments, on the other hand, is relatively unstable and can filter off correct matches especially for trademarks with holes and line drawing details.

In the second stage, the combined similarity measurement between a query Q and a trademark T is computed as

$$Sim(Q,T) = \frac{W_1 Sim_{EDH}(Q,T) + W_2 Sim_{BG}(Q,T)}{W_1 + W_2}$$

where W_1 and W_2 are the weighting factors, and Sim_{EDH} and Sim_{BG} are the similarity values based on EDH and maximum WBG matching. A simple way is to set $W_1 = W_2 = 0.5$. Here, we propose a novel approach based on the distribution of similarity values between Q and all the images in the database. Our intuition is that if one of the feature, EDH for instance, gives us many matches with similarity values close to 1, we can conclude that EDH is not a salient feature for this query in our database. As a result, the weight of EDH can be lower. The intuition is indeed similar to the "inverse document frequency" that frequently adopted in information retrieval literature [28]. In our approach, the weighting factors is a function Q as follows

$$W_1(Q) = \frac{1}{\frac{1}{|D|}\sum_{T \in D} Sim_{EDH}(Q,T)}$$
$$W_2(Q) = \frac{1}{\frac{1}{|D|}\sum_{T \in D} Sim_{BG}(Q,T)}$$

where D is the database and $|D|$ is its cardinality. In brief, the denominators of W_1 and W_2 are the average similarity values between Q and all the images in a database based on their features.

4 Retrieval and Fusion with Symmetry

Recent works of J. P. Eakins and et al in [11,12] experimented and analyzed a number of different shape measures, and drew out the conclusion that they are similar in retrieval effectiveness. They made the suggestion that integrating multiple features together and inventing novel methods for computing image similarity based on shape features may lead to better performance. Our maximum WBG matching approach for computing trademark similarity is indeed a novel method. This is because in [7–12], geometric shapes are extracted under the guidance of Gestalt principle, but discarded at the stage of feature generation. The geometric shapes are indeed converted to non-geometric information (e.g., aspect ratio, circularity). In other words, geometric elements are used intermediately but not being fully exploited for retrieval. Our work, in contrast, not only extracts and groups geometric primitives, but also fully utilizes the geometric parameters (e.g., distance, length, angle and etc) peculiar to geometric elements for maximum WBG matching.

Proper integration of multiple features is not trivial. In fact, it may deteriorate retrieval performance without careful integration. A property not considered by [7–12] is symmetry. However, [17, 29] show that Zernike moments do well with symmetry property of trademarks. We conduct experiments for trademark image retrieval with Zernike moments, and find that they do perform well with highly symmetric trademarks. However, for trademarks which are not highly symmetric, the performance drops greatly (see Section 5 for details). Although Zernike moments are thought to be the best invariant features among many moments including regular moments, Legendre moments, rotational moments and complex moments [30, 31], it is understandable that the retrieval performance using Zernike moments for asymmetric trademarks is not as good as that for symmetric trademarks. This is mainly because they belong to non-geometric features. Given the fact that large part of the trademarks are not highly symmetric, the best solution is to use Zernike moments only for query samples that are fairly symmetric.

4.1 Zernike Moments

Zernike moment of order (n, m) is computed as

$$A_{nm} = \frac{n+1}{\pi} \sum_{\rho} \sum_{\theta} [V_{nm}(\rho, \theta)]^* I(\rho, \theta), \text{ s.t. } \rho \leq 1$$

where $I(\rho, \theta)$ is the image pixel in polar coordinate, and $V_{nm}(\rho, \theta)$ is a Zernike basis polynomial defined as

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

where $R_{nm}(\rho)$ is defined as

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

where $n = 0, 1, 2, \infty$, $|m| \leq n$, and $n - |m|$ is even. In practice, the magnitudes of Zernike moments (ZMM) are used as the feature. The two parameters of Zernike moments, n and m, determine the properties of Zernike basis polynomials. The parameter m determines the symmetric property of Zernike basis polynomials (e.g., the ZMM of $m = 5$ is pentagonal-shaped), and $n - |m|$ determines the radial direction complexity of Zernike basis polynomials. Based on this characteristic of Zernike moments, we can know the best parameters of the Zernike moments for describing a trademark. As in [17, 29], the probabilistic distribution of ZMM for each (n, m) is modeled by a Gamma distribution whose parameters α and β can be estimated with ZMMs of the trademarks in our database. The degree of saliency, $DS(q, n, m)$ for a query trademark q, is defined [3] as

$$DS(q, n, m) = P(Z_{nm} \leq z_{nm}^{(q)})$$

where $z_{nm}^{(q)}$ is the ZMM of (n, m)-th order of the query trademark q. The larger the value of $z_{nm}^{(q)}$ is, the more the shape of q is affected by (n, m)-th order Zernike moment. Finally, the most salient feature (MSF) of the query trademark q is defined as the pair (N, M) that has the largest $DS(q, N, M)$, i.e.

$$MSF(q) = \operatorname*{argmax}_{N,M} DS(q, N, M)$$

We compute all the 100 moments up to order $n = 20$ for each trademarks in our database, and store the MSF of each trademark. For retrieval, the similarity between the query trademark Q and a trademark T in the database is computed as

$$Sim_{ZMM}(Q, T) = |Z_{NM}^{Q} - Z_{NM}^{T}|$$

[3] In [17, 29], $DS(q, n, m) = P(Z_{nm} \geq z_{nm}^{(q)})$ is used for the definition of saliency value, which is just one minus our saliency value.

where N and M is the MSF of Q. Some examples are illustrated in Figure 10. The column "N,M" and "DS" show the MSFs and the corresponding $DS(q, N, M)$ values for the query trademarks q, respectively.

4.2 Adaptive Selection of Geometric and Symmetric Features

Based on the observation above, we propose the approach for integrating Zernike moments and our geometric feature as follows. For query trademark q which satisfy $DS(q, N, M) > p$ with (N, M) as its MSF, we employ Zernike moments for retrieval. For others, we employ the proposed maximum WBG matching approach for the similarity measurement of geometric features. The threshold $p = 0.995$ is empirically determined. We can estimate the proportion of trademarks that use Zernike moments for retrieval. Suppose that the distribution of a trademark's ZMM is independent with the order (n, m), the probability P that the trademark's DS on its MSF is greater than p is

$$P = 1 - p^k = 1 - 0.995^{100} \approx 0.4$$

where $p = 0.995, k = 100$ in our implementation. But in fact the independent hypotheses is not satisfied, because a trademark q which has a large $DS(q, n, m)$ often has large $DS(q, n, km)$, where $k = 1, 2, ..., \infty$. So the actual probability should be less than 0.4. Experimental results justified this: for all the trademarks in our database, about 12 percent has a $DS(q, n, m) > 0.995$.

5 Experimental Results

5.1 Retrieval Accuracy

We use the benchmark trademark database in MPEG-7 dataset for performance evaluation. This database consists of about three thousand binary trademarks that are appropriate for testing. We select 50 trademarks from our database as query samples. The trademarks similar to these query samples are preselected manually. The numbers of manually preselected relevant trademarks for different query samples range from 10 to about 50. The evaluation is based on the normalized recall-precision [10–12] measures, where three measures of retrieval performance (normalized recall R_n, normalized precision P_n and last-place ranking L_n) are defined as follows:

$$R_n = 1 - \frac{\sum_{i=1}^{n} R_i - \sum_{i=1}^{n} i}{n(N-n)}$$

$$P_n = 1 - \frac{\sum_{i=1}^{n} (\log R_i) - \sum_{i=1}^{n} (\log i)}{\log(\frac{N!}{(N-n)!n!})}$$

$$L_n = 1 - \frac{R_l - n}{N - n}$$

where R_i is the rank at which the relevant trademark i is actually retrieved, R_l is the rank at which the last relevant trademark is found, n is the total number of relevant trademarks, and N is the size of the whole database. These measures can evaluate the retrieval performance from 0 (worst case) to 1 (perfect retrieval).

The mean and standard deviation of the performance scores from the 50 query samples are shown in Figure 8. The last row shows the results by the adaptive selection of proper features for retrieval. We can see that geometric features outperform Zernike moments, and by carefully integrating both features, better results are achieved.

Comparing to the retrieval performance in [10], which is $R_n = 0.90 \pm 0.12, P_n = 0.63 \pm 0.24, L_n = 0.56 \pm 0.31$, we can see that both approaches achieve similar performance, except that theirs did better in recall, while ours is better in precision and last-place ranking. Although we are using different databases and different queries, given the fact that [10] is one of the best works in trademark retrieval, we can see that our approach is very promising.

	Rn	Pn	Ln
Zernike Moments	0.76±0.18	0.55±0.21	0.42±0.35
Geometric Feature	0.84±0.12	0.61±0.18	0.56±0.27
Adaptive Selection between Zernike Moments and Geometric Feature	0.87±0.11	0.66±0.18	0.61±0.28

Fig. 8. Statistic (mean $\pm$ standard deviation) of experimental results on the trademark database in MPEG-7 dataset of about three thousand binary trademarks.

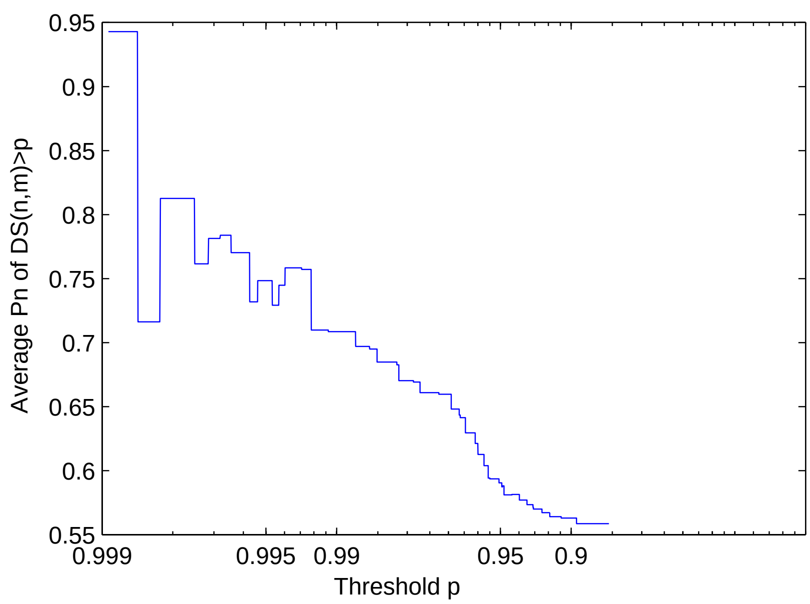

Fig. 9. The predefined threshold p versus the average P_n of the query trademarks whose $DS(N, M) > p$.

We take the threshold parameter $p = 0.995$ in all the experiments, i.e., a query trademark whose $DS(N, M) > 0.995$ is retrieved by Zernike moments, otherwise, the geometric features are used. In our experiments, 11 out of 50 queries are retrieved by Zernike moments. By investigating the retrieval performance of these 11 queries, we confirm that Zernike moments, on average, outperforms geometric features for trademark with high saliency value. From the results, we can guess that although Zernike moments are not as good as geometric features for retrieval, they are superior to geometric features for highly symmetric trademarks. Figure 9 shows the retrieval results by Zernike moments with different thresholds. We can see that by choosing 0.995 as the threshold, the average P_n for those query trademarks with $DS(N, M) > 0.995$ is about 0.75.

Figure 10 shows the results for ten trademarks. The column “N,M” and “DS” show the MSFs and the corresponding $DS(N, M)$ values for the query trademarks, respectively. During adaptive selection, Zernike moments are used for trademarks No.1 and No.4, while geometric features for the others. The features used are highlighted in Figure 10. Experimental results indicate that the proposed approach obtain good retrieval accuracy for some trademarks such as No.4 (a line of characters), No.3 (three circles) and No.10 (some characters in a box). It is because either they have large $DS(N, M)$ values (No.4), or they have describable geometric structures (No.3 and No.10). In our approach, the relevant trademarks in the first 20 of the retrieved images which are consistent with our manually labeled ground truth are shown in Figure 11. For some queries, such as No.9 and No.7, the retrieval performances are not so satisfactory. We present five relevant trademarks for each of them in Figure 12. We find that these trademarks are semantically similar rather than geometrically or symmetrically similar. For instance, the trademarks in (a) are all aircrafts and the trademarks in (b) are all letter ‘F’.

Query Image		Zernike Moment					Edge Histogram & Geometric Feature		
		Rn	Pn	Ln	N,M	DS	Rn	Pn	Ln
(1)		**0.87**	**0.66**	**0.51**	**18,8**	**0.997**	0.73	0.56	0.01
(2)		0.82	0.63	0.16	13,5	0.982	**0.83**	**0.67**	**0.28**
(3)		0.97	0.74	0.93	17,3	0.993	**0.99**	**0.92**	**0.95**
(4)		**0.98**	**0.89**	**0.92**	**6,2**	**0.996**	0.58	0.31	0.12
(5)		0.45	0.27	0.01	13,13	0.961	**0.86**	**0.71**	**0.28**
(6)		0.51	0.32	0.13	16,2	0.962	**0.93**	**0.73**	**0.74**
(7)		0.58	0.34	0.02	3,1	0.959	**0.71**	**0.48**	**0.15**
(8)		0.76	0.48	0.31	10,0	0.988	**0.93**	**0.79**	**0.62**
(9)		0.62	0.39	0.18	16,4	0.982	**0.58**	**0.41**	**0.16**
(10)		0.57	0.48	0.01	1,1	0.992	**0.96**	**0.85**	**0.81**

Fig. 10. Experimental results of ten queries on the trademark database in MPEG-7 dataset of about three thousand binary trademarks.

Through the experiments, we find that the DS value can basically represent the symmetry property of a query, and it roughly determines the retrieval accuracy of using Zernike moments. In general, the accuracy is high when the value of DS is large, and vice versa. Nevertheless, it is worth to notice that the value of DS does not always indicate the perception of human. For instance, queries No.2, No.3 and No.8 (in Figure 10) look quite symmetric, but their DS values are less than 0.995. Similarly, query No.10 has larger DS value than No. 2 and No. 8 although intuitively it is not as symmetric as them. In any case, the retrieval accuracy of these queries by Zernike moments is not better than by geometric features. As indicated in Figure 10, the effectiveness of retrieval by geometric features is clear. In most cases, the geometric features successfully encode the shape information of queries. For the six queries from No.5 to No.10, the retrieval accuracy by geometric features is significantly better than Zernike moments (except query No.9). For queries No.2 and No.3, although the retrieval results by Zernike moments are already good, geometric features can achieve even better accuracy.

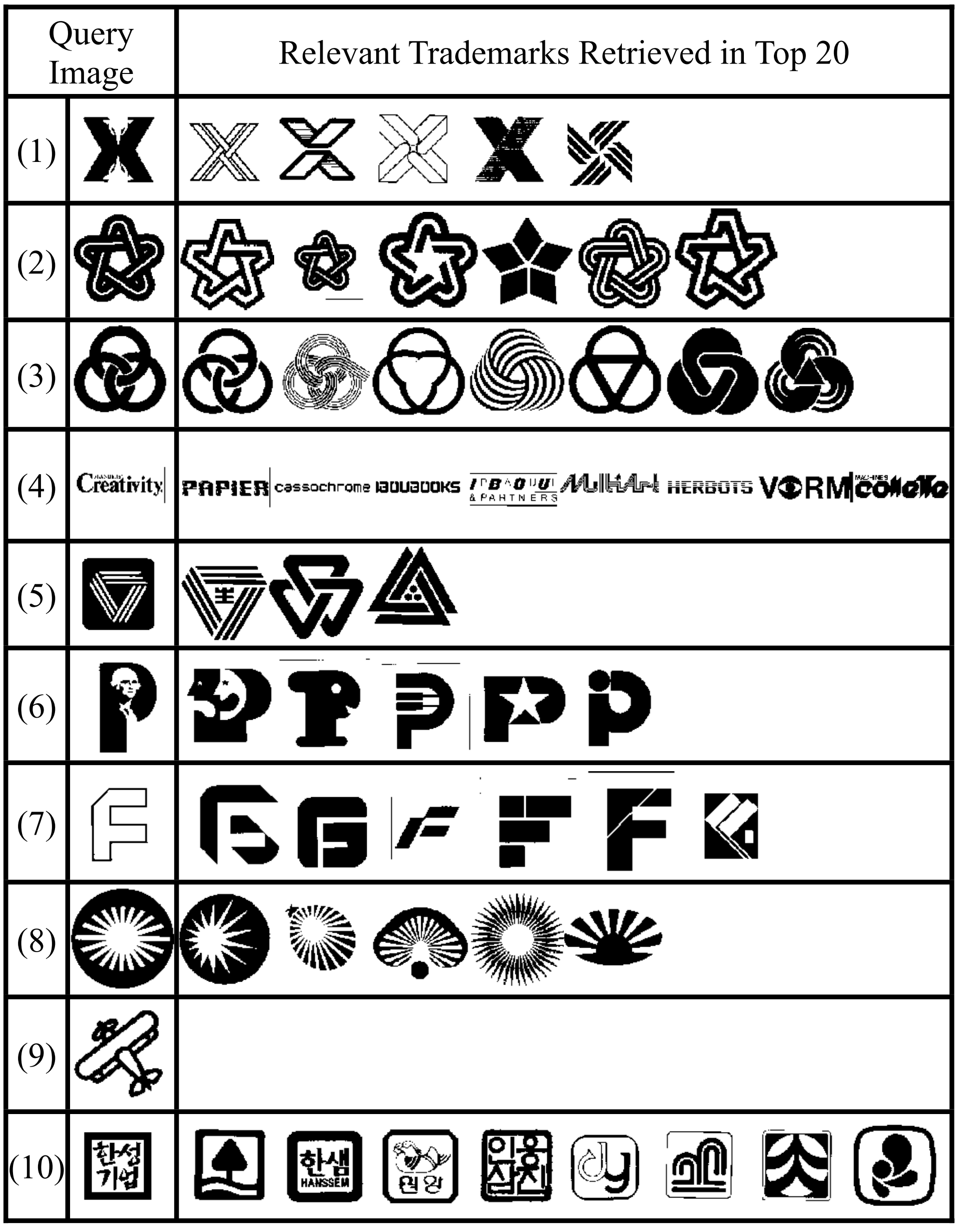

Fig. 11. Retrieval results for ten queries by adaptive selection.

Fig. 12. "Difficult" queries and the missed relevant trademarks.

For all 50 queries, the average retrieval time is approximately 5 seconds on a Pentium-III machine with 1G CPU and 256MB memory. For the hierarchical retrieval scheme with maximum WBG matching, the speed up is four times compared with pure maximum WBG matching. In particular, the average retrieval time using pure maximum WBG matching without hierarchical scheme is about 20 seconds on the same machine. In our current implementation, the top 600 retrieved trademarks from EDH are screened out for maximum WBG matching investigation. The retrieval accuracy of hierarchical retrieval is indeed close to that of using maximum WBG matching alone. In most cases, the weighting factor W_2 is greater than W_1 defined in Section 3.4.

5.2 Performance Comparison

Existing approaches that use bipartite graph matching for similarity measurement include [32,33]. We compare our method with Belongie's approach [33] since their shape matching algorithm performs quite well for trademark retrieval. In [33], bipartite graph matching is used to match two sets of edge points extracted from two images. Each point is attached with a descriptor named *shape context* which encodes the relative position of other points with respect to this point. The descriptor is represented as a log-polar histogram. The similarity between two shapes is based upon the upshot of point matching. An iterative framework, similar to ours in spirit, is used to improve both the correspondence and transformation. However, Belongie's approach employs regularized thin-plate splines while we use affine transformation to align two shapes.

To compare the performance of both approaches, we conduct experiments based on the same set of 50 queries on our database. Due to the fact that the efficiency of matching algorithm in [33] is directly impacted by the number of sample points, each trademark is represented by 100 feature points sampled from Canny edges. Figure 13 shows the retrieval performance of 10 query trademarks. The relevant images in the top 20 retrieved images and their corresponding R_n, L_n and P_n values are given in the table. The performance of our approach (in figures 10 and 11) is superior in term of the capability in recalling similar trademarks. Considering all the 50 queries, the average retrieval performance of Belongie's approach is:

$$R_n = 0.64 \pm 0.14, P_n = 0.39 \pm 0.19, L_n = 0.10 \pm 0.12$$

Compared with the results in Figure 8, our proposed approach is better in all performance measures. We repeat the same experiments for 200 and 300 sample points. No noticeable improvement is observed. Indeed, Belongie's approach performs well as long as critical points in shape are sampled for matching. In Belongie's approach, dummy nodes are added to increase robustness during bipartite graph matching. We experiment whether the same setting is useful for our approach. The retrieval performance, nevertheless, is not noticeably improved by using dummy nodes. We think the reason is that bipartite graph matching itself can handle partial matching, and outliers rarely overcome the true correspondence in a matching.

In term of speed efficiency, our method is favorable. The average computational time of Belongie's approach for a single image comparison increases exponentially with the increasing number of sample points. Even at a low number of sample

Query Image		Belongie's Approach [33]			
		Rn	Pn	Ln	Relevant TradeMarks Retrieved in Top 20
(1)		0.8185	0.4926	0.2447	
(2)		0.7470	0.4149	0.1376	
(3)		0.5052	0.2697	0.0536	
(4)		0.8334	0.6987	0.0572	
(5)		0.3276	0.1762	0.0122	
(6)		0.6492	0.2986	0.0104	
(7)		0.5536	0.2351	0.0478	
(8)		0.7266	0.2558	0.0190	
(9)		0.6770	0.2890	0.2674	
(10)		0.6624	0.3778	0.0483	

Fig. 13. Retrieval performance of Belongie's method.

points of 100, the typical retrieval time for a single query using their approach is around two hours [4] for a database of around three thousand images. Our approach completes the same task in 5 seconds, faster by three orders of magnitude. The typical number of feature points in Belongie's approach is about 10 times larger than the typical number of geometric elements extracted in our approach. Considering that the fastest version of weighted bipartite graph matching algorithm so far is $O(n^{2.5})$ [3], our approach is considerably efficient, in term of feature size, for

[4] We use the MATLAB code provided by [33] for the experiments, which might be speeded up if converted to C code.

trademark retrieval. In Belongie's approach, in addition to the complexity incurred in graph matching, the thin-plate spline coordinate transformation involves the inversion of a $p \times p$ matrix, where p is number of sample points. This impacts the speed of their algorithm as the number of sample points increases.

6 Discussions

In this section, we show the theoretical and practical arguments that support the use of particular techniques in our approach. The pro and con of these techniques, along with empirical evidence, are discussed.

6.1 Invariance to Transformation

As mentioned in Section 3.3, the geometric elements we adopt are actually not scale, rotational and translational invariant. Therefore, we embed the maximum WBG matching in an iterative framework to optimize the transformation that matches the two sets of geometric elements. Here, we give some analysis on how well this approach actually performs.

The iterative algorithm in our approach is essentially a local search algorithm. Although it cannot guarantee global optimum, it gives us an acceptable matching for most cases as justified by our experimental results. A global optimum may not be necessary since the similarity measurement is defined quite arbitrarily and our objective is to find a reasonable matching rather than a so-called global optimum of some arbitrary function.

An important factor that affects the performance of a local search algorithm is the smoothness of the neighborhood of a point in the state space for searching. The state space we are using is a subset of $\mathbb{R}^4$. In particular, it contains all the vectors that consist of 4 parameters for transformations (2 for translation, 1 for scaling and 1 for rotation). When the neighborhood of a point in the state space is smooth, the local search algorithm will perform well. Otherwise, there will be too many local optima and the search algorithm may easily get stuck into one of them which might be a poor matching. For example, the similarity of two lines is computed based on center positions, lengths and directions. The similarity values associated with the changes of each of these parameters should be smooth in the neighborhood of a particular point in the state space to achieve good local search results.

Here we use an example to show the effectiveness of our approach. In Figure 14, we use several transformed versions of a trademark as queries. In each of these retrievals, we manage to retrieve the original trademark as the most similar one in the whole database of about three thousand trademarks. This result is meaningful since we can achieve similar scale and rotation invariance by a local search framework, without using traditional normalization techniques such as shifting the trademark to its center of mass, scaling the trademark so that it has unit mass, rotating the trademark so that the longest axis is vertical. Most importantly, our approach can deal with more difficult cases, for instance the trademarks shown in Figure 2, where traditional invariant features such as moments cannot handle.

6.2 Adaptive Selection

Figure 8 indicates that our geometric features outperform Zernike moments, but adaptive selection between them performs even better. Here we give detailed analysis. Generally speaking, incorporating multiple features for retrieval does not always improve the retrieval performance. Since the performance of Zernike moments is worse than our geometric features for most trademarks, the retrieval accuracy may decrease if we simply incorporate them with simple method such as linear

Fig. 14. An example showing the effectiveness of our approach with the existence of transformations. The trademarks from left to right are respectively the original one, one that is scaled by 120%, one that is scaled by 150%, one that is rotated by 90 degree and one that is both rotated by 90 degree and scaled by 120%. In the experiments using the last four transformed trademarks as queries, the first one is always retrieved as the most similar one.

combination. Instead, we propose the employment of the saliency value $DS(N, M)$ inherent in the Zernike moments of a query image to guide the adaptive selection of features. Figure 15 shows the average normalized recall, precision and last-place ranking of the queries that have a saliency value greater than a particular value. We find that the average normalized recall, precision and last-place ranking decrease roughly monotonically as the saliency value decreases. For most cases, when the saliency value exceeds 0.995, the average retrieval performance using Zernike moments is better than geometric features, although geometric features outperform Zernike moments in retrieval in general. This is because Zernike moments work well for symmetric trademarks and the degree of symmetry can be measured by the saliency value.

However, there are counterexamples and cases that resemble counterexamples. For instance, although query No. 4 in Figure 10 is not symmetric from human perception, it is more effective to use Zernike moments compared to geometric features. This is because it is symmetric in terms of the distribution of pixels, which is indicated by the high saliency value with $M = 2$. Similarly but on the other end, query No. 5 and query No. 8 look symmetric from human perception but their saliency values are smaller than 0.995 and therefore geometric moments are the better feature descriptor. Query No. 5 does not have a large saliency value because it has both triangular and rectangular structures. Saliency value is large only if an image has a unique period in the polar angle axis. The saliency of query No. 8 is degraded because it has too many radial lines. Notice that its saliency value happens at $M = 0$. Normally, a zero value of M is not effective in describing symmetry.

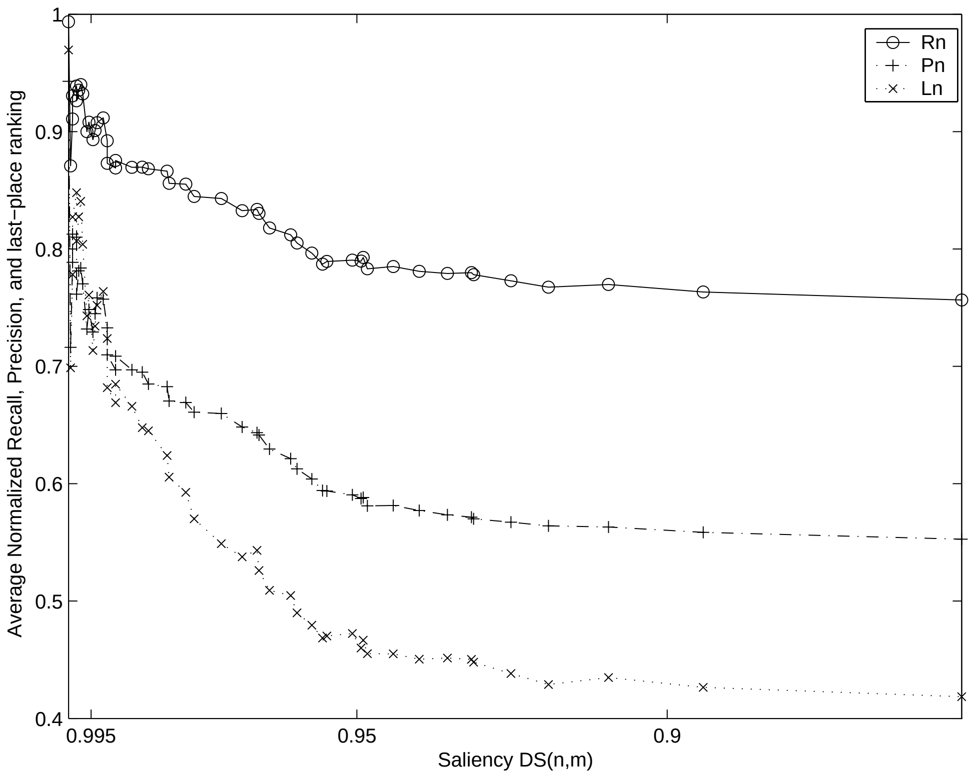

Fig. 15. The retrieval performance of various queries against their saliency values.

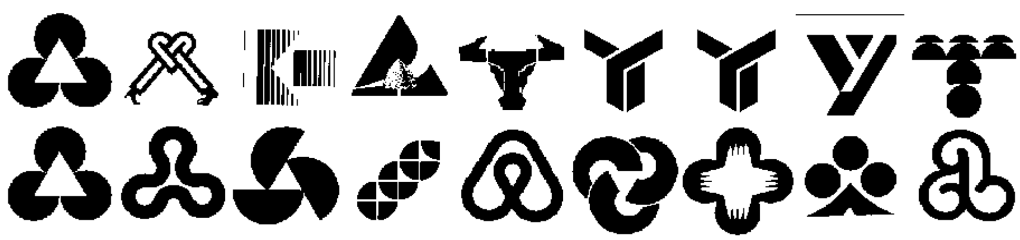

Fig. 16. An example showing the retrieval results of using Zernike moments (the first row) and using geometric features (the second row). The leftmost trademark is the query sample.

We also note some real counterexamples in our experiments. Figure 16 shows the retrieval results of using Zernike moments in the first row and the results using geometric features in the second row. The leftmost trademark is the query image. The saliency value of the query is 0.9982, achieved at $(N, M) = (7, 3)$. We can see that although the saliency value is larger than 0.995, the retrieval results using geometric features are better than using Zernike moments. Geometric moments successfully "catch" the characteristic of the trademark that it has three arcs but Zernike moments do not.

6.3 Retrieval with Zernike Moments and Its Variants

In [17, 29], Zernike moments are used for trademark retrieval. As shown in Section 5, this approach performs fairly well for symmetric trademarks. Deformed trademarks were also experimented in [17, 29]. Although the experiment is useful for general shape retrieval, it may not be critically important for trademark retrieval since trademarks are seldom deformed.

Besides [17, 29], other approaches that use Zernike moments for retrieval include [34,35]. In [34], 3D Zernike moments are utilized for retrieving 3D shapes. In [35], Zernike moments are employed for image reconstruction and recognition. The magnitudes of Zernike moments are used as feature values. However, the features are weighted by their contributions in the reconstruction procedures before comparison. This is different from our scheme, which uses the most salient feature (i.e., the saliency value $DS(N, M)$) for retrieval.

7 Conclusion

Based on the five wholistic properties of Gestalt principle, we have presented shape-based features that are appropriate for trademark retrieval. The effectiveness of our approach lies on the adaptive selection of features, and the maximum WBG representation for partial matching of geometrical elements inferred from Gestalt principle. Experimental results indicate that the adaptive selection scheme does improve the retrieval, in the sense that the retrieval performance using adaptive selection is better than that of using either of the two features for retrieval on their own. Experiments also show that Zernike moments work distinctly better for trademarks that have very high saliency values $DS(N, M)$, which usually refer to the highly symmetry of the trademarks, but not always the case. Also, experiments show that geometric features work reasonably well for trademarks that have describable geometric features. However, for the trademarks which are not symmetric and have no significant geometric characteristic (or simply because their geometric characteristics are difficult to be extracted by Hough transform), the retrieval performance of our approach is unsatisfactory. Future works will be concentrated on the incorporation of other feature extraction methods such as corner and texture detectors for more reliable interpretation of Gestalt principles by geometric features.

Acknowledgments

The work described in this paper was fully supported by a grant from City University of Hong Kong (Project No. 7001470).

We would like to thank the anonymous reviewers for the constructive comments. In addition, we thank Serge Belongie [33] for his MATLAB code for image matching.

References

[1] M. Wertheomer, Laws of Organization in Perceptual Forms, Humanities Press, 1950.

[2] J. E. Hopcroft, R. M. Karp, An $n^{2.5}$ algorithm for maximum matching in bipartite graphs, SIAM Journal on Computing 2 (1973) 225–231.

[3] H. N. Gabow, R. E. Tarjan, Faster scaling algorithms for network problems, SIAM Journal on Computing 18 (1989) 1013–1036.

[4] H. W. Kuhn, The hungarian method for the assignment problem, Naval Research Logistics Quarterly 2 (1955) 83–97.

[5] P. K. Agarwal, K. Varadarajan, Approximation algorithms for bipartite and non-bipartite matching in the plane, in: ACM-SIAM Symposium on Discrete Algorithms, 1999, pp. 805–814.

[6] J. A. McHugh, Algorithmic Graph Theory, Prentice Hall, 1990.

[7] S. Alwis, J. Austin, A novel architecture for trademark image retrieval systems, in: Proc. of the Challenge of Image Retrieval, 1998.

[8] S. Alwis, Content-based retrieval of trademark images, Ph.D. thesis, University of York (2000).

[9] S. Alwis, J. Austin, Trademark image retrieval using multiple features, in: Proc. of the Challenge of Image Retrieval, 1999.

[10] J. P. Eakins, J. M. Boardman, M. E. Graham, Similarity retrieval of trademark images, IEEE Multimedia 5 (2) (1998) 53–63.

[11] J. P. Eakins, J. D. Edwards, J. Riley, P. L. Rosin, A comparison of the effectiveness of alternative feature sets in shape retrieval of multi-components images, in: SPIE Storage and Retrieval for Media Database, Vol. 4315, 2001, pp. 532–540.

[12] J. P. Eakins, K. J. Riley, J. D. Edwards, Shape feature matching for trademark image retrieval, in: Intl. Conf. on Image and Video Retrieval, 2003, pp. 28–38.

[13] M. Ren, J. P. Eakins, P. Briggs, Human perception of trademark images: Implication for retrieval system design, Journal of Electronic Imaging 9 (4) (2000) 564–575.

[14] J.-K. Wu, C.-P. Lam, B. M. Mehtre, Y. J. Gao, A. D. Narasimhalu, Content-based retrieval for trademark registration, Multimedia Tools and Applications 3 (3) (1996) 245–267.

[15] M. Hussain, J. P. Eakins, Visual clustering of trademarks using a component-based matching framework, in: Intl. Conf. on Image and Video Retrieval, 2004, pp. 141–149.

[16] A. K. Jain, A. Vailaya, Shape-based retrieval: A case study with trademark image databases, Pattern Recognition 31 (5) (1998) 1369–1390.

[17] Y. S. Kim, W. Y. Kim, Content-based trademark retrieval system using visually salient feature, in: Proc. of the Intl. Conf. on Computer Vision and Pattern Recognition (CVPR), 1997, pp. 307–312.

[18] G. Ciocca, R. Schettini, Content-based similarity retrieval of trademarks using relevance feedback, Pattern Recognition 34 (8) (2001) 1639–1655.

[19] O. E. Badawy, M. Kamel, Shape-based image retrieval applied to tradmark images, Int. Journal of Image and Graphics 2 (3) (2002) 375–393.

[20] W. H. Leung, T. Chen, Trademark retrieval using contour-skeleton stroke classification, in: IEEE Intl. Conf. on Multimedia and Expo. Vol 2, Lausanne, Switzerland, 2002, pp. 517–520.

[21] K. T. Cheung, H. S. Ip, Semantic retrieval by spatial relationships, in: Int. Conf. of Pattern Recognition, 2002.

[22] I.-S. Hsieh, K.-C. Fan, Multiple classifier for color flag and trademark image retrieval, IEEE Trans. Image Processing 10 (6) (2001) 938–950.

[23] D. G. Lowe, Perceptual Organization and Visual Recognition, Kluwer Academic Publishers, 1985.

[24] R. C. Gonzalez, R. E. Woods, Digital image processing, Prentice Hall, 2002.

[25] R. O. Duda, P. E. Hart, Use of the hough transformation to detect lines and curves in pictures, Graphics and Image Processing 15 (1972) 11–15.

[26] P. J. Besl, N. D. McKay, A method for registration of 3-d shapes, IEEE Trans. on Pattern Analysis and Machine Intelligence 14 (2) (1992) 239–256.

[27] S. Cohen, L. Guibas, The earth mover's distance under transformation sets, in: Int. Conf. Comp. Vision, 1999.

[28] G. Salton, Automatic Text Processing, Addison-Wesley, 1989.

[29] Y.-S. Kim, W.-Y. Kim, M.-J. Kim, Development of content-based trademark retrieval system on the world wide web, ETRI Jounrnal 21 (1) (1999) 39–53.

[30] C. H. Teh, R. T. Chin, On image analysis by the methods of moments, IEEE Trans. on Pattern Analysis and Machine Intelligence 10 (4) (1988) 496–513.

[31] W. Y. Kim, P. Yuan, A practical pattern recognition system for translation, scale and rotation invariance, in: IEEE Conf. on Computer Vision and Pattern Recognition, 1994.

[32] S. Peleg, M. Werman, H. Rom, A unified approach to the change of resolution: Space and gray-level, IEEE Transaction on Pattern Analysis and Machine Intelligence 11 (7) (1989) 739–742.

[33] S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts, IEEE Transaction on Pattern Analysis and Machine Intelligence 24 (24) (2002) 509–522.

[34] M. Novotni, R. Klein, 3d zernike descriptors for content based shape retrieval, in: Proceedings of the eighth ACM symposium on Solid modeling and applications, 2003, pp. 216–225.

[35] Z. Khotanzad, Y. H. Hong, Invariant image recognition by zernike moments, IEEE Transaction on Pattern Analysis and Machine Intelligence 12 (5) (1990) 489–497.