A Novel Architecture for Trademark Image Retrieval Systems

Sujeewa Alwis and Jim Austin

Department of Computer Science

University of York

York, YO1 5DD, UK

email: sujeewa@cs.york.ac.uk and austin@cs.york.ac.uk

Abstract

This paper describes the first phase of an ongoing research project aimed at implementing a trademark retrieval system using an associative memory neural network. The novel aspect of the work described in this paper is the presentation of a new integrated methodology for employing multiple interpretations from different analytical levels of images for image retrieval. In achieving this objective, we extract local features as well as features of the closed figures of images. In deriving alternative interpretations of the images, a segment level gestalt grouping method based on a modification of Sarkar and Boyer's method is used. In designing the search engine of the system we have adopted a novel similarity assessment criteria based on local features as well as features of the closed figures, which may feasibly be implemented using an associative memory neural network to achieve high performance in retrieval.

1 Introduction

Retrieval of images by shape feature has proved a great challenge, though there has been considerable research into this topic [1, 2, 3]. Defining a shape similarity criteria which correspond to human visual perception is an unsolved problem [4]. Trademark image retrieval provides a good avenue of investigation in this regard since an effective trademark retrieval system should necessarily be able to retrieve images which humans perceive as similar. This is a difficult problem since results from psychological experiments show that most of the existing shape similarity assessment methods fail to identify some of the images perceived as similar by humans [5]. In addition, individual differences in perception of images makes the problem even more complicated. This evidence has motivated us to use concepts from visual cognitive psychology for investigating new shape retrieval systems. Because of diversity and complexity of the trademark images, it is difficult to capture their visual properties such as shape, structure and complexity. Unfortunately, there has been relatively little work done on trademark image retrieval systems [4, 6, 7]. In building up trademark retrieval systems, the issue of how humans judge these images should always be addressed.

Traditionally, classification of trademarks is based on limited vocabulary descriptions. Most of the trademark offices use manually assigned codes to represent these descriptions such as human beings, animals, geometrical figures. The internationally accepted Vienna classification for trademarks has 29 major categories such as human beings, animals, plants, landscapes, foodstuffs and each of these major categories have sub-categories and subsub-categories. However, it has been shown that all these methods suffer due to a number of problems including: the assignment of classes to trademarks is subjective, the classes become either too specific or too broad depending on how users use the classes, there is no mechanism to handle the generation of new classes [7], and there is a large fraction of images with little or no representational meaning which makes such a classification extremely difficult [4].

In our study, we attempt to find a solution to these problems using a content based image retrieval approach. The rest of the paper is organised as follows. In section 2, we summarize the previous work done in this field and in section 3, we explain the architecture of the proposed system. Finally we present some of the results obtained in the preliminary experiments.

2 Previous work

Several research groups have been working on finding a solution to this problem using automatic shape retrieval systems. Kato *et al* [6] described a system known as TRADE MARKS. Their approach is based on mapping normalised trademarks to an 8X8 pixel grid and calculating a GF-vector for each image from various pixel distributions. The query phase consists of matching GF-vectors between the query image and the stored images. The STAR system [7] offers a collection of modules for trademarks image retrieval based on shape, the meaning of the trademark and the words in the trade mark. They use Fourier descriptors, gray level projection and moment invariants to capture the shape features. The ARTISAN system [4] presents an approach which incorporates principles derived from Gestalt psychology. It uses a novel boundary family extraction method based on gestalt principles in an attempt to capture human perceptual similarity judgements about images. In this approach, gestalt principles of proximity and similarity are used to group boundaries into perceptually significant families. Feature vectors obtained from these families and boundaries alone are used for database indexing. Matching is based on calculating the Euclidean distance between feature vectors using several different strategies.

The retrieval of images using only the features of closed boundaries would not be always successful for a number of reasons: for example, objects under occlusion would not give the expected feature vector of the closed figure, the objects that do not have a closed contour would not have such a feature vector. On the other hand, local feature-based matching alone would also not be always robust enough due to the fact that it would be too sensitive to small changes in the object and there is enough psychological evidence to argue that human image understanding is not always driven by detailed segment by segment analysis [8]. These facts show the necessity in image retrieval of performing an analysis based on the features of closed figures as well as local features. In our study, we attempt to make use of the concepts put forward by several researchers in the visual cognitive psychology community including Biederman [9], Witkin and Tenenbaum [10], the Gestalt psychologists [11], for image retrieval.

3 Architecture of the system

Our system performs two major functions: feature extraction and retrieval. In the feature extraction phase, the system extracts information required for indexing. In the retrieval phase, the information acquired during the feature extraction phase is utilised to assess similarity with an query image.

The feature extraction phase goes through the following steps as shown in figure 1:

- extraction of local features and perceptual relationships from the first level image (*raw image*)
- perceptual organization to obtain the *Gestalt image*
- extraction of local features and perceptual relationships from the *Gestalt image*
- extraction of features from the closed figures in *raw* and *Gestalt images*

Figure 1: The overview of the feature extraction phase.

The retrieval phase goes through the following steps, using both the *raw image* and the *Gestalt image*:

- similarity assessment of the *raw images* based on local features and perceptual relationships
- similarity assessment of the *raw images* based on features of closed figures
- similarity assessment of the *Gestalt images* based on local features and perceptual relationships
- similarity assessment of the *Gestalt images* based on features of closed figures

The following sections describe each of these steps in detail.

3.1 Feature extraction phase

3.1.1 Extraction of local features and perceptual relationships from the raw image

In the local feature extraction from the *raw image* stage, constant curvature edge tokens are extracted and segmented using the method proposed by Wuescher and Boyer [13]. These two steps give straight line segments having the following properties: starting point, end point, orientation, and pixel points on the line; and arc segments with following properties: starting point, end point, center, curvature, gradients to the curve at the end points and the pixel points on the arc. We use this local feature information to obtain the perceptual relationships between the segments.

In the ARTISAN system, principles of perceptual organization are used for content based image retrieval. ARTISAN groups boundaries into different proximal and shape families. However, the use of

boundary based grouping alone suffers some limitations such as an inability to group parts of different boundaries and to group image components with cluttered boundaries. In our system, we perform segment based perceptual grouping in an attempt to overcome some of the limitations of the boundary based grouping.

In our approach to perceptual relationship extraction, seven graphs are created representing the following pairwise relationships: end point proximity, orientation similarity, curvature similarity, parallel line segments, parallel curve segments, co-linear segments, and co-curvilinear segments. Each graph has nodes representing the segments and their attributes, and edges representing the corresponding perceptual compatibility relationships between segments.

To avoid combinatorial search in the creation of these graphs, bin spaces are prepared for the compatibility relationships. In preparing the bin spaces and the end-point proximity graph, we followed the method proposed by Sarkar and Boyer [14]. To obtain parallel segments, they created a proximity graph, taking the position of each pixel of every segment into consideration. This method is expensive in terms of memory and speed, due to pixel voting. Alternatively, we make use of orientation and curvature similarity graphs to reduce the search space in proximity analysis, as detailed below.

We first prepare one dimensional bin spaces of orientation and curvature of segments, recording their segment identification numbers as tags in the corresponding bins. During the process, pairs of segments with the same orientation or curvature share the same bin positions. To detect parallel segments, these pairs of segments are further examined to assess the degree of proximity between them. This is performed by constructing a two dimensional bin space of (x, y) image co-ordinates. A tag is stored representing each pixel point of the pair of segments under consideration. Then, the number of bins shared by both segments is counted and normalized by the length or perimeter of the longer segment. This value gives an indication on the degree of proximity between the pair of segments. If this value exceeds a pre-determined threshold, they are considered to be parallel segments. This step is needed to filter out the segments which show the same orientation or curvature but are far away from each other.

Sarkar and Boyer use the earlier prepared end-point proximity graph to obtain the co-linear and co-curvilinear segments. This limits the detection of such grouping to a relatively small subregion. This limitation can be avoided by preparing a new end-point proximity graph using a higher proximity threshold for co-linear and co-curvilinear grouping.

In constructing the co-linearity graph that shows the lines which are co-linear, the end point proximity graph and orientation similarity graph are ‘AND’ed. Then, the non-colinear associations in the new graph are removed. When a line segment has two or more co-linear segments on the same side of it, the closest segment is chosen. In constructing the co-curvilinearity graph, the same procedure for the co-linearity graph is followed after replacing the orientation similarity graph with the curvature similarity graph.

Figure 2: Perceptual features captured using Gestalt principles.

Figure 2(b) shows co-linear and co-curvilinear segments extracted from the the image in figure 2(a), and figure 2(c) shows the parallel segment pairs extracted from the image in figure 2(a).

The algorithm proposed by Reingold *et al.* [15] to find the fundamental cycles in graphs was used to obtain the closed figures or components from these connected nodes. This could be easily done by detecting closed loops in the end point proximity graph. Closed figure detection using this method gives the ability to detect the figures that have cluttered boundaries in poor quality images. Figure 3 shows some of the extracted closed figures with cluttered boundaries, extracted using this method.

Figure 3: Some of the closed figures extracted using Gestalt principles.

In some images, a boundary segment may be contained in more than one closed figure (e.g., the edges of the triangle in figure 4(b)). It is difficult to obtain some of the expected closed figures using pixel based linking methods which use boundary pixels to obtain only the closest boundary approximation.

Figure 4(b) shows the extracted edges of the image in figure 4(a) while figure 4(c) and figure 4(d) show the expected triangle and circle, obtained using the end point proximity graph based closed figure extraction method.

Figure 4: Some of the closed figures extracted using Gestalt principles.

The fundamental cycles detection algorithm fails to detect all the different closed figure interpretations of an image. But, it was noticed that an algorithm which extracts all the cycles, gives a large number of closed figures which could not be handled by a practical image retrieval system.

In preparing the *Gestalt image*, the co-linear segments are replaced by new continuous line segments, while co-curvilinear segments are replaced by new continuous curve segments. This step gives rise to a new image structure. Figure 5 shows some of the *raw images* and their *Gestalt images* (note that the grouped curves cannot be seen as continuous curves due to pixel by pixel display of curves in our interface).

Some structural patterns which include parts of the co-linear or co-curvilinear segments can not be seen in the new image after grouping (e.g., figure 5(a)). A different set of closed figures may be created at this

stage. This makes the similarity matching using both the *raw image* and the *Gestalt image* useful.

Table 1 illustrates the performance of the grouping process (obtained under the same constraint thresholds using a Silicon Graphics Indy workstation).

Image	Co-linear grouping			Co-curvature grouping		
	false pos	false neg	expected	false pos	false neg	expected
5(a)	0	1	10	0	0	5
5(c)	0	0	34	0	0	0
5(e)	0	0	18	0	2	9
5(g)	0	0	5	0	0	2
5(i)	0	2	18	0	0	0

Image	Parallel grouping			Closure			Time
	false pos	false neg	expected	false pos	false neg	expected	
5(a)	0	5	10	0	0	8	50s
5(c)	1	6	19	1	0	14	170s
5(e)	0	6	17	5	0	15	70s
5(g)	0	0	6	3	1	4	25s
5(i)	0	6	15	0	0	1	165s

Table 1: Performance of the system in obtaining co-linear, co-curvilinear and parallel segment pairs as well as closed figures, using the images in figure 5.

3.1.2 Extraction of local features and perceptual relationships from the Gestalt image

The *Gestalt image* may have fewer segments as a result of the replacement of the co-linear and co-curvilinear segments by continuous line and arc segments, respectively. The attributes of the segments of the new image structure can be easily derived from the attributes of the *raw image*.

The *Gestalt image* structure goes through the above mentioned procedure for perceptual relationship extraction to obtain the end-point proximity graph and the parallel lines and curves graph. The end point proximity graph for the new image is used to detect the closed figures in the new image, using the same procedure for the *raw image*.

3.1.3 Feature extraction of the closed figures

In this step, features of the closed figures (components) are extracted from the *raw image* and the *Gestalt image*. The feature measurements set includes a subset of features proposed by Umetani and Taguchi [16]. In many systems, including QBIC [17], ARTISAN [4] and SAFARI [18], some or most of these features have been used for image retrieval. The difference in our system is the method we adopt to extract the closed figures and the similarity assessment criteria using these features. Our feature parameter set includes:

- Circularity = $4\pi A/P^2$
- Right-angleness = r/n
- Sharpness = $\sum \max(0, 1 - (2|\theta_i - \pi|/\pi)^2)/n$
- Complexity = $10^{-(7/n)}$

- Directness = M/P
- Straightness = S/P
- Aspect Ratio = p_1/p_2
- Stuffedness = A/R

where:

A, area of the figure enclosed by the segment boundary;

P, perimeter of the figure enclosed by the segment boundary;

r, number of discontinuity angles equal to a right angle within a specified tolerance;

M, total length of straight line segments parallel to mode direction of straight line segments within a specified tolerance;

S, total length of straight line segments in segment boundary;

n, number of sides of polygon enclosed by segment boundary;

R, area of the circumscribed rectangle of minimum area;

p_1 , length of the closed figure;

p_2 , width of the closed figure;

θ_i , discontinuity angle between $(i - 1)^{th}$ and i^{th} boundary segment.

The relative positions of the closed figures are also extracted and currently we extract five interrelationships, namely above, below, left, right and enclosed. All the closed figures or components are stored as a graph in which nodes represent the feature vector and arcs represent different types of relative spatial relationships.

3.2 Retrieval phase

In the retrieval phase, we use local features as well as the features of closed figures of both *raw* and *Gestalt images*, in separate modules.

3.2.1 Local features based retrieval

Using the raw images

In this stage we represent each *raw image* using five graphs obtained as explained in section 3.1.1, which represent the non-accidental properties identified by Witkin and Tennenbaum [10]. Biederman postulates that the detection of these relationships is one important step in object recognition [9]. The graphs are the endpoint proximity graph, parallel lines graph, parallel curves graph, co-linearity graph and co-curvature graph; the nodes represent a segment identification number and the edges represent the corresponding associations. In the next step all these graphs are combined so that the new graph represents all the above mentioned relationships. As a result, the new graph will have five different types of edges representing different associations.

The matching process consists of matching the query graph with model graphs in the database, using the relaxation matching by elimination method [19]. This process consists of a starting phase which finds out candidate node matches, a constraint propagation process to filter out false matches and a final aggregation phase.

As the first step in performing matching, all the nodes y in the model graph x , which have similar attributes to each node j in the query graph i , are identified to assign an initial possibility for a match between node ij and node xy , expressed as:

if $\|att_{ij} - att_{xy}\| < \epsilon$

then $S^{(0)}(\theta_{ij} = \bar{\theta}_{xy}) = 1/Z$

else

$S^{(0)}(\theta_{ij} = \bar{\theta}_{xy}) = 0$

where $S^{(n)}(\theta_{ij} = \bar{\theta}_{xy})$ - possibility function for $(\theta_{ij} = \bar{\theta}_{xy})$ after n^{th} iteration.

θ - the query graph

$\bar{\theta}$ - the model graph

θ_{ij} - the j^{th} node of the i^{th} graph which represents the j^{th} segment of the i^{th} image

Z - the normalization constant (during the experiments described in this paper, Z is made 1)

To reduce the time taken for comparison of attributes, one dimensional bin spaces are created for each feature parameter. The current attributes include orientation (with bin step of 20°) and length (with bin step of 100 pixel units) for straight line segments and curvature for curve segments. Since we use images of size less than 500×500 pixels, it can be seen that the system can tolerate some amount of scale variation. However, the local feature matching stage is not rotation invariant since it takes the orientation as an attribute of straight line segments.

Optimisation of evidence obtained from initial hypotheses is performed using the relaxation matching framework recently proposed by Turner and Austin [19]. It has been shown that this performs better than the conventional probabilistic relaxation method in the light of experimental results in the domain of chemical graph matching. Moreover, this methodology gives the advantage of implementation using a Correlation Matrix Memory (CMM) neural network [20] which would give very high speed processing rate.

In optimising the evidence obtained from initial hypotheses, the evidence from perceptual neighbours for each possibility is taken into account, as expressed below:

$$S_k^{(n+1)}(\theta_{ij} = \bar{\theta}_{xy}) = S_k^{(n)}(\theta_{ij} = \bar{\theta}_{xy}) \cdot U_k^{(n)}(\theta_{ij} = \bar{\theta}_{xy})$$

where $U_k^{(n)}(\theta_{ij} = \bar{\theta}_{xy})$ is the contextual support for $(\theta_{ij} = \bar{\theta}_{xy})$ from the neighbourhood nodes connected by perceptual relationship k .

Then only the possibilities which have enough evidence, above a certain threshold ϵ_{ij} , are considered for the next iteration as follows:

if $\sum_k S_k^{(n+1)}(\theta_{ij} = \bar{\theta}_{xy}) > \epsilon_{ij}$

$$S^{(n)}(\theta_{ij} = \bar{\theta}_{xy}) = 1/Z$$

else

$$S^{(n)}(\theta_{ij} = \bar{\theta}_{xy}) = 0$$

The threshold ϵ_{ij} has to be incremented to re-start the relaxation process once the update process of labels becomes stable. It is done only at the most dependable positions. The selection of such nodes is done by choosing the nodes that show the lowest entropy (see [19] for more details).

The final similarity assessment is performed by calculating the contribution from each model graph x for possible matches at all nodes j of the query graph i , as explained below.

The possibility of a node to node match between the query graph and the each model graph is calculated first:

$$C(\theta_{ij} = \bar{\theta}_{xy}) = \sum_k S_k^{(t)}(\theta_{ij} = \bar{\theta}_{xy})$$

In the next step each such possibility at each node in the query graph is normalized by the sum of all the possibilities at each node in the query graph, as follows:

$$C(\theta_{ij} = \bar{\theta}_{xy}) = C(\theta_{ij} = \bar{\theta}_{xy}) / \sum_x \sum_y C(\theta_{ij} = \bar{\theta}_{xy})$$

Then, all the possibility values for a match with each model graph x , at all the nodes in the query graph i are summed.

$$C(\theta_i = \bar{\theta}_x) = \sum_j \sum_y C(\theta_{ij} = \bar{\theta}_{xy})$$

Finally, the similarity measure score between image i and x is expressed as $C(\theta_i = \bar{\theta}_x)$.

Use of relaxation matching improves the robustness to noise and errors that occur due to use of arbitrary parameters and due to anomalies of bin voting at the extraction of perceptual relationships phase. A relaxation matching framework helps in retrieving images that have partial similarities in the arrangement of segments.

Using the Gestalt images

The same procedure for the *raw image* is followed except the *Gestalt image* is represented using three graphs: the endpoint proximity graph, the parallel lines graph, and the parallel curves graph instead of five graphs as in the earlier case. This is obvious since in this image, all the co-linear and co-curvilinear segments are replaced by continuous segments.

3.2.2 Retrieval based on features of the closed figures

At this stage, retrieval is performed using feature vectors of closed figures. The most common method in matching shapes using feature vectors is by calculating the distance between the vectors. However, this method suffers some drawbacks with multi-component images for a number of reasons: a single pair of highly different components would distort the overall distance measure of partially similar pair of images, a relatively large difference in one element of the feature vector will have a large effect on the final result, all the feature parameters have the same effect on the final similarity measure which makes the representation of sensitivity of each parameter in human judgements difficult and integration of the knowledge about the relative positions of objects may be difficult using such a criteria.

The similarity measuring criteria used in our system adopts an evidence counting method which could feasibly be implemented using a CMM neural network. The evidence counting method is based on finding the candidate matches and calculating the matching possibility for each such image. First, we find the figures of the stored images which have at least a certain number of sufficiently closer feature elements to each figure in the query image. To reduce the time taken to compare each feature parameter, a separate one dimensional bin space is created for each feature parameter so that each figure stores its identification number as a tag in the corresponding bins. The number of common bins is calculated for each pair of figures from the query image and one of the stored images. If it exceeds a certain threshold, evidence count ($S(C_{ij} = \bar{C}_{xy})$) for a match between query figure ij and stored figure xy is set to unity or otherwise zero. After this step is performed, the evidence counts for each query figure are normalized by the total evidence count for that particular query figure as follows:

$$S(C_{ij} = \bar{C}_{xy}) = S(C_{ij} = \bar{C}_{xy}) / \sum_x \sum_y S(C_{ij} = \bar{C}_{xy}) \text{ (step 2)}$$

In the final step, the total contribution of evidence counts for a match between query image i and stored image x is calculated as follows.

$$S(C_i = \bar{C}_x) = \sum_j \sum_y S(C_{ij} = \bar{C}_{xy}) \text{ (step 3)}$$

The same procedure is followed for both *raw* and *Gestalt* images.

The system has the flexibility in considering only pairs of figures which do not have large differences in enclosed area simply by using another bin space for area. It avoids making similarity judgements between figures with highly different sizes while imposing some restrictions on scale invariance. We hope to evaluate the advantages and disadvantages of this step in future experiments using large and more diversified trademark images.

We wish to integrate the spatial inter-relationships of the figures with the evidence counting procedure and observe whether it can improve the overall similarity assessment measure. Unfortunately, the use of inter-relationships such as left, right, above and below will make this module rotationally sensitive. Alternatively, we could represent all the relationships with a single connectivity type relationship in the graph based representation of the multi-component image. In doing so, we require an additional step before step 2, in the earlier procedure. The initial evidence count $S(C_{ij} = \bar{C}_{xy})$ will be multiplied by the total contextual evidence (denoted by $U_k S(C_{xy} = \bar{C}_{ij})$ where k is the type of the relationship) from the neighbour components for $S(C_{ij} = \bar{C}_{xy})$, as follows:

$$S(C_{xy} = \bar{C}_{ij}) = S(C_{xy} = \bar{C}_{ij}) \cdot \sum_k U_k S(C_{xy} = \bar{C}_{ij})$$

Table 2 illustrates the similarity measures obtained from the new evidence counting method and the Euclidean distance measuring method, with the same query images. (R_{norm} , P_{norm} and L_n and the dataset used are described in the next section). In calculating the Euclidean distance, an *asymmetric simple matching* method, which was identified as the most effective Euclidean distance measure criteria in ARTISAN experiments [4], was used.

Image	Evidence counting criteria			Euclidean distance criteria		
	R_{norm}	P_{norm}	L_n	R_{norm}	P_{norm}	L_n
fig 6(c)	0.99	0.91	0.95	0.95	0.88	0.80
fig 6(d)	0.99	0.97	0.99	0.97	0.88	0.80
fig 6(f)	0.99	0.97	0.92	0.97	0.82	0.92
fig 7(a)	0.99	0.99	0.96	0.98	0.71	0.80
fig 8(a)	1	1	1	1	1	1
fig 8(b)	0.99	0.96	0.97	0.98	0.91	0.95

Table 2: Comparison of retrieval performance between the usage of evidence count criteria and Euclidean distance measure criteria.

4 Preliminary Results

We have conducted preliminary experiments on the performance of the system using a smaller image database of 210 trademark images which includes nine groups of perceptually similar images (in total 61, some examples are shown in figures 6-11) which have been pointed out by trademark examiners during evaluation experiments of the ARTISAN system [4], and 149 arbitrary selected images. We have summarised some of the results obtained in these experiments below. For the experiments presented in this paper, we count the ranks starting from 0 and retrieval failures are marked by a dash.

To measure the performance, we have used the following measures proposed by Salton [21] which has been regarded as the most widely accepted measures in assessing the precision in information retrieval [4].

normalized recall : $R_{norm} = 1 - \frac{\sum_{i=1}^n R_i - \sum_{i=1}^n i}{n(N-n)}$

$$\text{normalized precision : } P_{norm} = 1 - \frac{\sum_{i=1}^n \log(R_i) - \sum_{i=1}^n \log(i)}{\log \frac{N!}{n!(N-n)!}}$$

In addition we have used last place ranking which has been defined as: $L_n = 1 - \frac{R_n - N}{N - n}$ where R_i is the rank at which relevant image i is actually retrieved, n is the total number of relevant image and N is the size of the whole image collection.

In this paper, we use the following equation to obtain the *combined score* for each image, though we are currently investigating a better method. In the calculation, we give the lowest possible ranking (i.e., 209) for the images which are not retrieved by the system:

$$\text{combined - score} = \sum_i \frac{1}{rank_i + 1}$$

where $rank_i$ is the rank of the retrieved images obtained using the retrieval module i . The following tables illustrate the results obtained using the above mentioned four matching modules and the new rank obtained by combining them.

It can be seen that during the tests the query image is not always retrieved as the best candidate. This is due to the effect of one to many matching possibilities between the parts of the the query and model images. However, it can be seen that the query image has been retrieved within a reasonable distance.

Image	Obtained rank				
	local (raw)	local (gest)	closed fig. (raw)	closed fig. (gest)	combined
6(a)	6	10	9	-	13
6(b)	2	8	3	10	3
6(c)	1	9	5	6	2
6(d)	0	4	0	0	0
6(e)	3	6	2	-	5
6(f)	8	16	10	11	12
7(a)	5	15	6	8	6
7(b)	14	19	1	-	4
7(c)	12	14	7	9	11
R_{norm}	0.99	0.97	0.99	-	0.98
P_{norm}	0.95	0.71	0.97	-	0.90
L_n	0.97	0.94	0.99	-	0.97

Table 3: Results obtained with query image figure 6(d).

Image	Obtained rank				
	local (raw)	local (gest.)	closed fig. (raw)	closed fig. (gest.)	combined
8(a)	9	11	9	2	6
8(b)	0	4	2	0	1
8(c)	3	3	5	6	3
8(d)	1	0	0	1	0
8(e)	12	6	1	5	2
8(f)	2	5	8	12	5
9(a)	7	8	6	7	9
9(b)	10	9	3	3	4
9(c)	5	7	7	8	8
9(d)	11	10	4	4	7
R_{norm}	0.99	0.99	1	1	1
P_{norm}	0.95	0.90	1	0.99	1
L_n	0.98	0.99	1	0.99	1

Table 4: Results obtained with query image figure 8(a).

Image	Obtained rank				
	local (raw)	local(gest.)	closed fig. (raw)	closed fig. (gest.)	combined
10(a)	0	16	1	0	2
10(b)	-	17	-	8	22
10(c)	77	7	4	-	8
10(d)	15	12	3	2	3
10(e)	103	1	11	-	4
10(f)	10	5	12	-	9
11(a)	98	43	-	7	23
11(b)	9	16	2	-	7
11(c)	2	0	0	1	0
R_{norm}	-	0.96	-	-	0.98
P_{norm}	-	0.78	-	-	0.86
L_n	-	0.83	-	-	0.92

Table 5: Results obtained with query image figure 10(a).

5 Future Work

We hope to conduct further evaluation experiments on the performance of the system, with a larger and more diversified image database including more noisy and gray-level images. We are currently implementing the local feature based matching module using Correlation Matrix Memories. Our final objective is to implement all the matching modules using the PRESENCE hardware platform [22] which is a binary neural network, to achieve rapid retrieval.

Acknowledgements

We would like to thank Dr. John Eakins at the University of Nothumbria at Newcastle for the substantial benefits obtained from the discussions with him and providing the trademark image collection for our ex-

periments.

References

- [1] S. Sclaroff and A. Pentland. Object recognition and categorization using model matching. In *Proceedings of the 2nd CAD Based Vision Workshop*, pages 258-265, 1994.
- [2] W. I. Grosky and Z. W. Jiang. A hierarchical approach to feature indexing. *Image and Vision Computing*, vol. 5, pages 275-283, 1994.
- [3] R. Mehrotra and J. E. Gray. Similar-shape retrieval in shape data management. *IEEE Computer* vol 28, no 9, pages 57-62, 1995.
- [4] J. P. Eakins, M. E. Graham, J. M. Boardman et al. Retrieval of trademark images by shape feature. *British Library Research and Innovation Report* 26, 1996.
- [5] D. Mumford. Mathematical theories of shape: Do they model perception. *Geometric methods in computer vision-SPIE*, vol 2185 pages 2-10, 1992.
- [6] T. Kato. Database Architecture for content based image retrieval. *Image storage and Retrieval Systems-SPIE*, vol 2185, pages 112-123, 1992.
- [7] J. K. Wu, C. P. Lam, B. E. Mehtre et al. Content-based retrieval for trade mark registration. *Multimedia Tools and Applications*, vol 3, pages 245-267, 1996.
- [8] M. Eysencj and M. Keane. *Cognitive Psychology*. Lawrence Erlbaum Associates Publishers, 1995.
- [9] I. Biederman. Recognition by components - A theory of human image understanding. *Psychological Review*, vol 94, no 2, pages 115-147, 1987.
- [10] A. Witkin and J. Tennenbaum. *On the role of structure in human and machine vision*. Academic Press, pages 481-543, 1982.
- [11] D. G. Lowe. *Perceptual Organization and Visual Recognition*. Kluwer Academic Publishers, Boston, 1985.
- [12] J. Canny. A computational approach to edge detection. *IEEE PAMI*, vol 8, pages 679-714, 1986.
- [13] D. Wuescher and K. L. Boyer. Robust contour decomposition using a constant curvature criterion. *IEEE PAMI*, vol 13, pages 41-51, 1991.
- [14] S. Sarkar and K. Boyer. *Computing perceptual organization in computer vision*. World Scientific Publishers, 1994.
- [15] E. M. Reingold. *Combinatorial Algorithms*, Prentice Hall Publishers, pages 318-392, 1977.
- [16] Y. Umetani and K. Taguchi. Discrimination of general shapes by psychological properties. *Digital Systems for Industrial Automation*, vol 1, pages 179-198, 1982.
- [17] M. Flickner. Query by image and video content. *IEEE Computer*, vol 28, no 9, pages 23-32, Sep 1995.
- [18] J. P. Eakins. Design criteria for a shape retrieval system. *Computers in Industry*, vol 21, pages 167-184, 1993.
- [19] M. Turner and J. Austin. A neural relaxation technique for chemical graph matching. In *Proceedings of the Fifth International Conference on Artificial Neural Networks, Cambridge, UK*, Editor. M Niranjana, IEE Publishers, July 1997.
- [20] M. Turner and J. Austin. Matching Performance of Binary Correlation Matrix Memories. *Neural Networks*, 1997.
- [21] G. Salton. *The SMART Retrieval System*. Prentice Hall Publishers, 1971.
- [22] J. V. Kennedy and J. Austin. A Parallel Architecture for Binary Neural Networks. In *Proceedings of MicroNeuro '97, 6th International Conference on Micro Electronics for Neural Networks, Evolutionary Systems and Fuzzy Systems*, 1997.

Figure 5: Figures 5(a) - 5(h) show the *raw image* (left) and *Gestalt image* (right) pairs. figure 5(j) is the *raw image* obtained from figure 5(i) (after rejecting small segments) and figure 5(k) is the corresponding *Gestalt image*.

Figure 6: Some of the cited images similar to figure 6(d).

Figure 7: Some of the cited images similar to figure 6(d).

Figure 8: Some of the cited images similar to figure 8(a).

Figure 9: Some of the cited images similar to figure 8(a).

Figure 10: Some of the cited images similar to figure 10(a).

Figure 11: Some of the cited images similar to figure 10(a).