

Shape Retrieval by Partially Supervised Fuzzy Clustering

G. Castellano¹ A.M. Fanelli¹ M.A. Torsello²

¹Department of Informatics, University of Bari "A. Moro"

²Department of Informatics, Systems, and Communication, University of Milano-Bicocca

Abstract

In this work we propose the use of partially supervised fuzzy clustering to create a two-level indexing structure useful for enabling efficient shape retrieval. Similar shapes are grouped by a fuzzy clustering algorithm that embeds a partial supervision mechanism exploiting domain knowledge expressed in terms of a set of labeled shapes. After clustering, a set of prototypes representative of shape clusters is derived and used as indexing mechanism for retrieval. A shape query is matched against prototypes, instead of the whole shape database, and then shapes belonging to clusters for which prototype similarity is higher are returned. Experimental results obtained on two different datasets are presented to show the effectiveness of the proposed approach.

Keywords: Shape retrieval, fuzzy clustering, partially-supervised clustering.

1. Introduction

Shape similarity and retrieval are very important topics in computer vision. The recent progress in this domain has been mostly driven by designing smart shape descriptors for providing good similarity measures between pairs of shapes. A large variety of shape descriptor methods and related matching criteria have been proposed in literature. Broadly speaking, these can be divided into region-based and contour-based methods. The first methods exploit only shape boundary information such as Fourier descriptors, moments analysis, scale space analysis. In region-based methods, shape representation is obtained by exploiting all pixels within a shape region. These methods use moment descriptors such as geometric moments, Legendre moments, Zernike moments and pseudo Zernike moments. Accordingly to the different shape descriptors, several matching algorithms have been developed [17], [2]. Nearly all of these approaches are based on computing pairwise shape similarity measure. However, with large databases, it is not practical to sequentially compare each object against the query. A large number of unnecessary comparisons is performed since only a small number of objects is likely to match the query.

To avoid expensive comparison, in this work we use a two-level index structure by considering the existing shapes as forming different clusters according to their similarity. Each cluster is represented by a prototype, hence the index structure includes a prototype layer and a shape layer. The prototype layer acts as a filter that reduces the search space quickly, discriminating the objects. In this way a database of shapes is organized so as to enable efficient searches.

The idea of exploiting clustering algorithms to group together similar shapes and to derive prototypes useful as a indexing mechanism has been investigated in different works [6], [18], [7], [14], [5].

In these works unsupervised clustering has been applied. However, this could generate not homogeneous clusters including shapes that are visually similar but belonging to different categories. This situation is not very surprising. Actually, as stated in [13] it is rather unrealistic to expect that unsupervised learning could produce a zero classification error of shapes. To improve the results of shape clustering it is often necessary to embed some domain knowledge about shape categories, thus considering a partially supervised clustering.

In this work we propose the use of partially supervised clustering to form clusters of shapes. A mechanism of partial supervision is applied to a fuzzy clustering algorithm in order to take advantage from domain knowledge expressed in terms of a number of labeled shapes. The clustering process is applied to shape boundaries represented by Fourier descriptors and for each cluster a prototype is identified. The derived prototypes are used as primary indexing mechanism to perform the retrieval process. Namely, a shape query is matched with all the prototypes and shapes belonging to clusters corresponding to the top-n similar prototypes are provided as a result.

The remainder of the paper is organized as follows: in Section 2 we describe the proposed approach. Section 3 gives the experimental results on two benchmark shape datasets to show the effectiveness of the proposed approach. Conclusions and discussions are given in Section 4.

2. The proposed approach

We assume that object shapes have already been extracted from the images and the shapes are available in form of contours. Of course, in many applications extraction of contours itself is a difficult problem but our focus here is on retrieving shapes once the contours are extracted. Therefore we consider object shapes that are described by boundary coordinates. To represent shape boundaries, we use Fourier descriptors that are well-recognized to provide robustness and invariance, obtaining good effectiveness in shape-based indexing and retrieval [3]. Based on this representation, each shape is described by means of M Fourier descriptors $\mathbf{x} = (x_1, x_2, \dots, x_M)$ ¹.

To formalize the shape retrieval problem we can consider the classical setting of retrieval that applies to many retrieval scenarios like keyword, document, image, and shape retrieval. Given is a set of objects $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ and a similarity function $sim : X \times X \rightarrow [0, 1]$ that assigns a similarity value to each pair of objects. We assume that \mathbf{x}_1 is a query object (e.g., a query shape), and $\{\mathbf{x}_2, \dots, \mathbf{x}_N\}$ is a set of known database objects (or a training set). Then, by sorting the values $sim(\mathbf{x}_1, \mathbf{x}_i)$ in decreasing order for $i = 2, \dots, N$, we obtain a ranking of database objects according to their similarity to the query, i.e., the most similar object has the highest value and is listed first. A distance measure could be also used. In this case the ranking should be obtained by sorting the objects in increasing order, i.e., the object with the smallest value is listed first. Usually, the first $n \ll N$ objects are returned as the most similar to the query \mathbf{x}_1 .

Similar shapes are grouped together by employing a fuzzy clustering algorithm equipped with a partial supervised mechanism originally proposed in [12]. This algorithm is a modified version of the Fuzzy C-Means (FCM) algorithm that takes into account domain knowledge expressed by a set of labeled shapes that provides a useful guidance during the clustering process thus improving the derivation of groups of similar shapes.

Let $X = \{\mathbf{x}_j\}_{j=1}^N$ be a set of N shapes represented by Fourier descriptors and K a given number of clusters, we denote by $X_1 \subset X$ a subset of labeled shapes and by $\mathbf{b} = [b_j]_{j=1}^N$ a boolean vector defined as follows:

$$b_j = \begin{cases} 1 & \text{if shape } \mathbf{x}_j \text{ is labeled} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Membership values of the labeled shapes to the clusters are arranged into a matrix $\mathbf{F} = [f_{jk}]_{j=1 \dots N}^{k=1 \dots K}$.

The partially supervised FCM algorithm works in the same manner of FCM, i.e. it iteratively de-

rives K clusters by minimizing an objective function. To embed partial supervision in the clustering process, the objective function of FCM is modified by adding a supervised learning component encapsulated in the form of \mathbf{b} and \mathbf{F} as follows:

$$O = \sum_{k=1}^K \sum_{j=1}^N u_{jk}^m d_{jk}^2 + \alpha \sum_{k=1}^K \sum_{j=1}^N (u_{jk} - b_j f_{jk})^m d_{jk}^2 \quad (2)$$

where d_{jk} represents the Euclidean distance between the shape \mathbf{x}_j and the center of the k -th cluster, m (the fuzzification coefficient) is any real number greater than 1 and α is a parameter that serves as a weight to balance the supervised and unsupervised components of the clustering process. The higher the value of α , the higher is the impact coming from the supervised component. It has been observed that a value of α proportional to the rate between the number of labeled shapes and the total number of shapes ensures that the impact of the labeled shapes is not ignored in the clustering process. The second term of the objective function captures the difference among the true membership of shapes (encapsulated in \mathbf{F}) and the membership computed by the algorithm. The aim to be reached is that, for the labeled shapes, these values have to coincide.

The objective function O is minimized by updating the partition matrix $\mathbf{U} = [u_{jk}]_{j=1 \dots N}^{k=1 \dots K}$ in the following way:

$$u_{jk} = \frac{1}{1 + \alpha} \left[\frac{1 + \alpha(1 - b_j \sum_{l=1}^K f_{lk})}{\sum_{l=1}^K d_{jk}^2 / d_{lk}^2} \right] + \alpha b_j f_{jk} \quad (3)$$

Also the centers of clusters are updated according to the following formula:

$$\mathbf{c}_k = \frac{\sum_{j=1}^N u_{jk}^m \mathbf{x}_j}{\sum_{j=1}^N u_{jk}^m} \quad (4)$$

The clustering process ends when the difference between the values of the objective function in two consecutive iterations drops below a prefixed threshold or when the established maximum iteration number is reached. Once the clustering process is completed, we consider the center \mathbf{c}_k as the prototype of the k -th cluster.

The choice of the subset of shapes to be labeled can be accomplished in different ways. For example, domain knowledge may be explicitly supplied by an expert that knows the categories of a certain number of shapes. Usually, the expert provides labels for the “most unclear” shapes, i.e. shapes that turn out to be visually similar to other shapes belonging to different categories. This approach, that was implemented in [4], gives good results but assumes the availability of some expert knowledge. In this work, we propose a novel way to acquire domain knowledge that consists in automatically selecting the shapes to be labeled among those shapes that are misclassified by FCM. In particular, among

¹As concerns the choice of the number M , an appropriate number has to trade off the accuracy in representing the original boundary with the compactness and simplicity of the representation. In this work, M has been empirically fixed to 32.

these, we choose to label a number of shapes that are randomly selected from each category according to the abundance of shapes misclassified in the corresponding category.

The derived prototypes are used as a first-level indexing mechanism for retrieval. Whenever a user submits a shape as a query, in the form of an object outline created by drawing an object on a sketch or on an image, we translate the outline into a set of Fourier descriptors \mathbf{x}_q . The query description \mathbf{x}_q is matched against descriptors of all prototypes \mathbf{c}_k to compute similarity values as follows:

$$\text{sim}(\mathbf{x}_q, \mathbf{c}_k) = 1 - \|\mathbf{x}_q - \mathbf{c}_k\| \quad (5)$$

where $\|\cdot\|$ is the Euclidean distance.

Then, prototypes are sorted in descending order on the basis of the similarity values and we return shapes indexed by the prototypes with high similarities.

3. Simulation results

To assess the suitability of the proposed approach, we performed a suite of experiments on two datasets, namely the Surrey Fish [1] and the MPEG-7 Core Experiment CE-Shape-1 [3] which are among the most widely used benchmarks to evaluate shape matching algorithms for retrieval tasks.

In our experiments we carried out several trials of the partially supervised FCM (hereafter indicated by psFCM) and the unsupervised FCM on the both considered datasets in order to evaluate the effectiveness of our approach in identifying shape categories and deriving significant prototypes. The retrieval performance is estimated using Precision and Recall.

In the following subsections, we provide results obtained by testing our shape retrieval approach on each considered dataset.

3.1. The Surrey Fish dataset

This dataset consists of 1,100 shape files where each file contains the coordinates of boundary points of the shape of a marine animal. In our experiments we considered a portion of the dataset composed of 265 shapes that have been manually classified into 10 different semantic categories, as follows: “Seamoths” (11), “Sharks” (58), “Soles” (52), “Tonguefishes” (19), “Crustaceans” (11), “Eels” (26), “U-Eels” (20), “Pipefishes” (16), “Seahorses” (11) and “Rays” (41). This manual classification was used as ground truth to assess the performance of the proposed shape retrieval approach. In Fig. 1 we show some sample shapes of the Surrey Fish dataset portion along with their respective semantic categories.

Each shape was represented by Fourier descriptors and the set of the obtained 265 shape descrip-

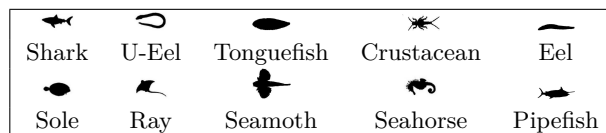


Figure 1: Sample images from the Surrey Fish dataset

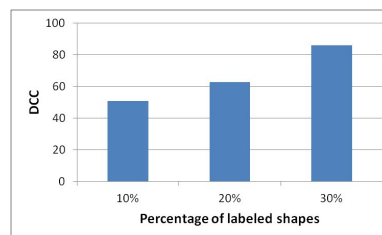


Figure 2: Dominant category cardinalities obtained by varying the percentage of labeled shapes

tors was divided into two parts: a training set, corresponding to the 90% of the whole set, that was employed for the creation of clusters and derivation of shape prototypes and a testing set, corresponding to the remaining 10%, that was employed for shape retrieval. To obtain more reliable results, a 10-fold cross-validation procedure was performed.

The psFCM was applied by varying the percentage of labeled shapes from 10% to 30%. We experimentally observed that better shape partitions are obtained by fixing in all runs the fuzzification coefficient $m = 2$ and the cluster number $K = 10$ (equal to the number of shape categories). The parameter α was modified so as to make it proportional to the percentage of labeled shapes. For each of the 10 runs, we evaluated the cardinality of the dominant category (DCC) in each derived cluster in order to verify the goodness of the obtained partitions. The obtained results are shown in Fig. 2. As expected, the higher the percentage of labeled shapes, the better the quality of the derived partitions in terms of compactness of clusters.

To assess the effectiveness of psFCM in creating good clusters of shapes, we performed a comparison with the unsupervised FCM that was run by varying the number of clusters K from 9 to 15 (being 10 the number of categories) and the value of the fuzzification coefficient m from 1.5 to 3. Figure 3 depicts the average results on the 10 runs of FCM varying the number of clusters and the values of the fuzzification coefficient. It can be observed

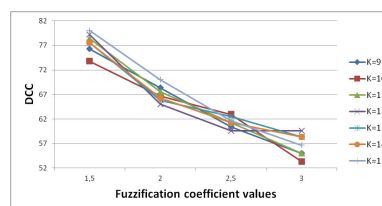


Figure 3: DCC values obtained by FCM with no supervision on the Surrey Fish dataset

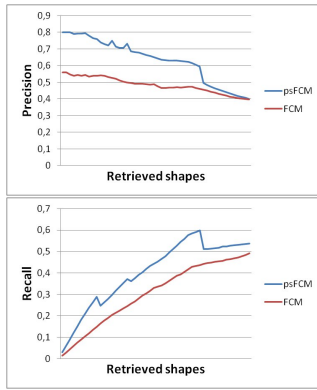


Figure 4: Precision and Recall obtained on the Surrey Fish dataset

that DCC values are crucially affected by such clustering parameters. Indeed, DCC values improve as the number of clusters increases. Moreover, regardless the number of clusters, highest DCC values are obtained in correspondence of the value 1.5 for the fuzzification coefficient.

Given the obtained clusters and corresponding prototypes, we performed a suite of experiments on the testing set. Precisely, each shape included in the testing set was considered as a query. Thus it was matched with the derived shape prototypes and the first 50 shapes indexed by the prototypes with higher similarity were considered as a result of the query. Then, Precision and Recall measures were computed on the testing set. Precision and Recall values were computed from each set of returned shapes as result to the query (from 1 to 50). Then average values of Precision and Recall were computed on the total number of queries included in the testing set. To better assess the retrieval accuracy of the proposed approach, we compared psFCM with 30% of labeled shapes with FCM with $m = 1.5$ and $K = 10$. Comparative results are shown in table 4. Here we present Precision and Recall values obtained in correspondence of each set of retrieved shapes (from 1 to 50). A method is considered to be better than another if it achieves better precision and better recall for larger sets of retrieved shapes. On the overall, the obtained results confirm that the supervision mechanism offers benefits in determining significant prototypes useful as indexing mechanism to retrieve shapes. Indeed, on the average psFCM overcomes FCM in terms of retrieval accuracy. These values could be considered as a good compromise between a satisfactory retrieval accuracy and the simplicity of the retrieval process. In effect, based on this method, the domain knowledge to be provided in input to the shape clustering algorithm is acquired in a completely automatic manner without asking the intervention of the expert. In this way, the process of shape retrieval will be faster and less prone to the subjectivity of the human expert.

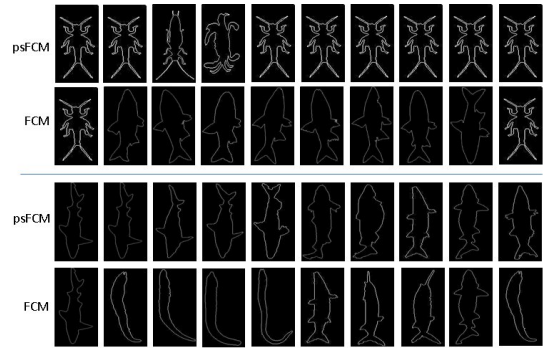


Figure 5: The first 9 shapes retrieved for two queries by applying psFCM and FCM

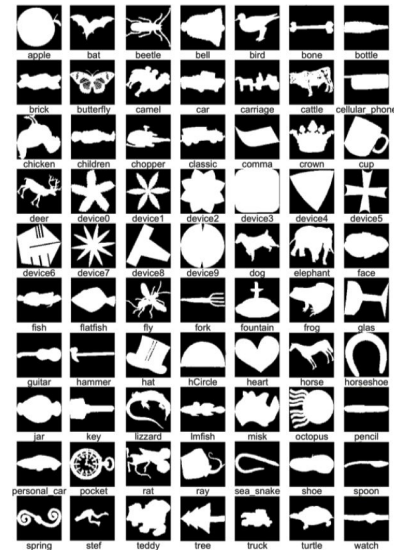


Figure 6: Sample shapes of the MPEG-7 dataset

To give an idea of the retrieval behavior of our approach, in Fig. 5 we show two sample results. For each query (the top-left shape), the first nine results are shown. As it can be observed, our approach permits to retrieve a higher number of relevant shapes in correspondence of the two considered queries.

3.2. The MPEG-7 dataset

The MPEG-7 Core Experiment CE-Shape-1 dataset contains 1400 binary images grouped into 70 different categories where each category contains 20 samples. Fig. 6 shows a sample image for each category of the dataset. We tested our approach on this dataset by following the same suite of experiments performed on the Surrey Fish dataset. The whole collection of 1400 shape descriptors was divided into a training set and a test set, and a 10-fold cross-validation procedure was performed. The psFCM algorithm was applied by setting the fuzzification coefficient $m = 2$ and the number of clusters $K = 70$ (corresponding to the 70 shape categories included in the dataset). Also in this case, we performed several trials of psFCM by varying the percentage of labeled shapes automatically selected

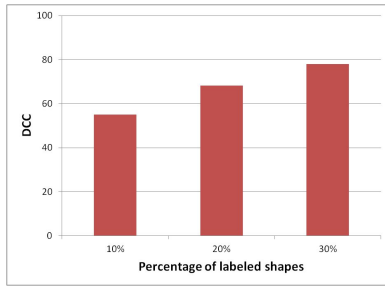


Figure 7: Dominant category cardinalities obtained by varying the percentage of labeled shapes

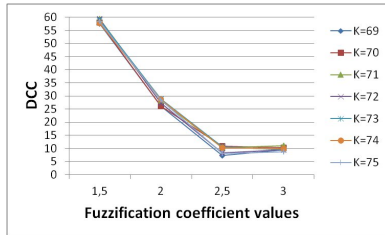


Figure 8: DCC values obtained by FCM with no supervision on the MPEG-7 dataset

from 10% to 30%. The parameter α was accordingly modified so as to make it proportional to the established percentage of labeled shapes. In Fig. 7 we show the average DCC values obtained in the 10 performed trials. As expected, the goodness of the derived partitions expressed in terms of compactness of clusters improves as the percentage of labeled shapes increases. As before, the psFCM was compared with the unsupervised FCM that was run by varying the number of clusters K from 69 to 75 (being 70 the number of categories) and the value of the fuzzification coefficient m from 1.5 to 3. Figure 8 shows the average results on the 10 performed trials obtained by FCM varying the number of clusters and the values of the fuzzification coefficient. Also in this case, we obtained highest DCC values in correspondence of $m = 1.5$.

The performance of our retrieval approach on this dataset was evaluated by computing Precision and Recall on the testing set by following the same procedure used for the Surrey Fish dataset. Then, we compared accuracy results obtained by psFCM with 30% of labeled shapes and FCM with $m = 1.5$ and $K = 70$. Fig. 9 shows the obtained comparative results. Also in this case, we may observe that the retrieval accuracy obtained by deriving shape categories and prototypes with psFCM is better than the accuracy obtained by applying FCM. This confirms the benefits of the partially supervised mechanism in deriving good shape partitions. As an example, in Fig. 10 we show the first nine returned shapes for two sample queries. We may observe that, as before, our approach allows to improve the retrieval results by returning a higher number of relevant shapes in correspondence of the two con-

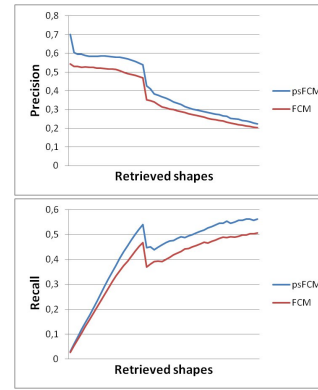


Figure 9: Precision and Recall obtained on the MPEG-7 dataset

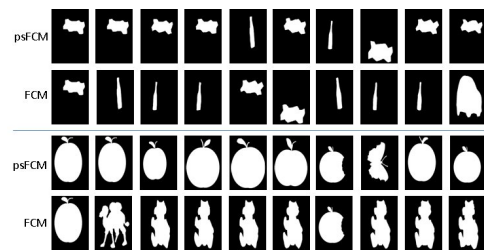


Figure 10: The first 9 shapes retrieved for two queries by applying psFCM and FCM

sidered queries.

Since MPEG-7 is among the most employed datasets in retrieval tasks, we also evaluated the retrieval performance of our approach on this dataset by using the Bulls Eye test that is the standard test used in the literature for MPEG-7. This is a leave-one-out test where 40 most similar shapes are determined for every query shape (being 40 the twice the number of relevant shapes, i.e. shapes belonging to the same category of the shape query). The final score, called Bulls Eye Percentage (BEP), is given by the ratio between the sum of the number of relevant shapes retrieved for each query and the highest number of relevant shapes that could be retrieved (in this case 20×1400). In our experiments, we obtained a BEP score equal to 77.40%. The performance of our approach was compared with other state-of-the-art shape matching algorithms applied on the same dataset. Table 1 summarizes the comparative results. Although not achieving the best result when compared to all those reported in the literature, our approach takes less time than other approaches. Indeed, the use of prototypes, that are an intermediate level of indexing, avoids comparison with all shapes in the dataset, thus allowing a faster retrieval process. In comparison most existing systems will probably take longer time to do the same task.

Algorithm	BEP
CSS [11]	75.44
SC + Dsc [10]	64.59
IDSC + Dsc [10]	68.83
IDSC + DP [10]	85.40
Visual Parts [8]	76.45
Our approach	77.40

Table 1: Performance comparison on MPEG-7 dataset

4. Conclusions

In this paper, a shape-based retrieval approach has been presented. The approach exploits a fuzzy clustering algorithm equipped with a partial supervision mechanism that is able to provide an helpful guide during the process of shape clustering in order to derive a set of prototypes representative of a number of shape categories. The derived prototypes are employed as an indexing mechanism to retrieve shapes that satisfy a query submitted by a user. The retrieval performance of our approach has been evaluated by computing standard measures on two different benchmarks, the Surrey Fish and the MPEG-7 datasets. The obtained accuracy results are encouraging. Beyond the stated goals, the proposed approach can contribute in developing robust algorithms for computer vision by incorporating shape-based recognition of objects. As a major conclusion we stand that our approach has demonstrated usefulness and effectiveness for both retrieval and recognition purpose, particularly if taken into account its simplicity. Of course, many issues could be addressed in order to improve the effectiveness of our approach. For example, the employment of more complex datasets and different descriptors for shape representation could be useful to definitely assess the suitability of our retrieval approach.

References

[1] Abbasi, S., Mokhtarian, F., Kittler, J.: SQUID Demo Dataset 1,500.

[2] A. Amanatiadis, V.G. Kaburlasos, A. Gasteratos, and S.E. Papadakis. Evaluation of Shape Descriptors for Shape-Based Image Retrieval. *IET Image Processing*, 5(5):493-499, 2011.

[3] I. Bartolini, P. Ciaccia, and M. Patella. WARP: Accurate retrieval of shapes using phase of Fourier descriptors and Time warping distance. *IEEE Trans. on Pattern Analysis and machine Intelligence*, 27(1), 142-147, 2005.

[4] G. Castellano, A.M. Fanelli, M.A. Torsello. Fuzzy image labeling by partially supervised shape clustering. In *Proc. of the 15th international conference on Knowledge-based and intelligent information and engineering systems*, pp. 84-93, 2011.

[5] I.E. Frigui, N. Boujemaa, L. Soon-Ann Lim. Unsupervised clustering and feature discrimination with application to image database categorization. In *Proc. of Joint 9th IFSA World Congress and 20th NAFIPS International Conference*, pp.401-406, vol.1, 2001.

[6] S. Joshi and A. Srivastava. A geometric approach to shape clustering and learning. In *Proc. of 2003 IEEE Workshop on Statistical Signal Processing*, pp. 302 - 305, 2003.

[7] P. Kotschieder, M. Donoser, and H. Bischof. Beyond Pairwise Shape Similarity Analysis. In *Proc. of the 9th Asian conference on Computer Vision*, pp. 655-666, 2010.

[8] L.J. Latecki, R. Lakamper, and U. Eckhardt. Shape descriptors for non-rigid shapes with a single closed contour. In *Proc. of Intl. Conf. on Computer Vision and Pattern Recognition*, pp 424-429, 2000.

[9] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain. Content-based multimedia information retrieval: state of the art and challenge. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2(1):1-19, 2006.

[10] H. Ling and D. W. Jacobs. Using the inner-distance for classification of articulated shapes. In *Proc. of Intl. Conf. on Computer Vision and Pattern Recognition*, pp. 719-726, 2005

[11] F. Mokhtarian, F. Abbasi, and J. Kittler. Efficient and robust retrieval by shape content through curvature scale space. *Image Databases and Multimedia Search*, 51-58, 1997

[12] W. Pedrycz and J. Waletzky. Fuzzy clustering with partial supervision. *IEEE Transaction System Man Cybernetics*, 27(5):787-795, 1997.

[13] W. Pedrycz, A. Amato, V. Di Lecce, V. Puri. Fuzzy Clustering with partial supervision in organization and classification of digital images. *IEEE Transactions on Fuzzy Systems*, 16(4):1008-1025, 2008.

[14] N. Rajpoot and M. Arif. Unsupervised shape clustering using diffusion maps. *Annals of the BMVA*, 2008(5):1-17, 2008.

[15] C. Shahabi and M. Safar. An experimental study of alternative shape-based image retrieval techniques. *Multimedia Tools and Applications*, 32(1):29-48, 2007.

[16] D. Zhang and G. Lu. A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures. *Journal of Visual Communication and Image Representation*, 14(1):41-60, 2003.

[17] D. Zhang. and G. Lu. Review of shape representation and description techniques. *Pattern Recognition*, 37:1-19, 2004.

[18] N. Xing and I.S. Ahmad. Fuzzy Clustering Paradigm and the Shape-Based Image Retrieval. In *Proc. of the 21st International FLAIRS Conference*, pp. 121-122, 2008.