



A NEW APPROACH TO SHAPE-BASED IMAGE RETRIEVAL

Mehdi Chehel Amirani ¹⁾, Zahra Sadeghi Gol ²⁾, Ali Asghar Beheshti Shirazi ³⁾

¹⁾ Iran University of Science and Technology, Narmak, Tehran, Iran, amirani@ee.iust.ac.ir

²⁾ Iran University of Science and Technology, Narmak, Tehran, Iran, zsadeghigol@ee.iust.ac.ir

³⁾ Iran University of Science and Technology, Narmak, Tehran, Iran, abeheshti@iust.ac.ir

Abstract: Content-based image retrieval (CBIR) is very active research topic in recent years. This paper introduces a new approach to shape-based image retrieval. At first, feature points are determined at the boundary of the shape as the extremums of a new version of the curvature function and the initial features are calculated at these points. The proposed method utilizes a supervised system for nonlinear combination of initial features for extraction of efficient and low dimensional feature vector for each shape. The retrieval performance of the approach is illustrated using the MPEG-7 shape database. Our experiments show that the proposed method is well suited for object indexing and retrieval in large databases.

Keywords: Principle Component Analysis, Linear Discriminant Analysis, Neural Networks, Feature Extraction, Curvature Function, Image Retrieval.

1. INTRODUCTION

Recent years have seen a growing interest in developing effective methods for recognition of objects in multimedia content. Typically, visual features which provide complementary information, such as shape, color and texture are chosen for this purpose. Shape is clearly an important cue for recognition since humans can often recognize characteristic objects solely on the basis of their shapes.

Many shape analysis techniques have been proposed over the past three decades. An extensive survey of shape matching in computer vision can be found in [1] and [2]. Common criteria used for shape representation for reliable shape matching and retrieval include: uniqueness, invariance to translation, scale, rotation and symmetric transformations, scalability, compactness, extraction and matching efficiency, and robustness to different kinds of distortions [3].

A contour-based descriptor encapsulates the shape properties of the object's outline (silhouette). It should distinguish between shapes that have similar region-shape properties but different contour-shape properties. Such descriptors are usually very efficient in applications where high intraclass variability in the shape is expected, due to deformations in the object (rigid or nonrigid) or perspective deformations [4].

Since silhouettes do not have holes or internal

markings, the associated boundaries are conveniently represented by a single closed curve which can be parameterized by arc length. Early work used Fourier descriptors [5] which are easy to implement and are based on the well-developed theory of Fourier analysis. The disadvantage of this approach is that, after the Fourier transform, local shape information is distributed to all coefficients and not localized in the frequency domain [1]. A representation that has proven to be relevant in human vision is the medial axis transform (MAT) originally proposed by Blum [6], [7]. This approach led Sebastian *et al.* [8] to attempt to capture the structure of the shape in the graph structure of the skeleton. In this approach, the edit distance between shock graphs is used as a metric of similarity between shapes. Sebastian *et al.* claim that their approach is robust to various transformations including articulation, deformation of parts, and occlusion. The disadvantage of this description method is that matching two graphs is very computationally expensive.

Given the above observations, we conclude that a representation with enhanced discriminatory capabilities could be obtained by explicitly using the feature point. To evaluate this idea we introduced a new shape description method in [9] and expand it in this paper. Here we utilize a resistant version of curvature function against surface noise in compared to ordinary version. Furthermore, in the present approach, we use linear discriminant analysis (LDA)

that is a supervised method in the first stage of the hierarchical feature extraction stage instead of principal component analysis (PCA). Feature points are generated using the extremums of the new version of the curvature function. Then some information are calculated as initial features. The registration of shape for determining the starting point and so stabilizing of method against geometrical transforms such as translation, rotation and scaling is done by a PCA transform. After calculating of initial features from each feature counterpoint, a hierarchical feature extraction stage consist of LDA and a supervised neural network combines the most important of initial features to reduce their numbers and at the same time retain as much as possible of their class discriminatory. The retrieval performance of the approach is illustrated using the MPEG-7 shape database.

The paper is organized as follows. Section 2 explains the feature generation process. Section 3 discusses the concept of feature extraction using PCA, LDA and a supervised neural network. The proposed method is presented at Section 4. The experimental results are given in section 5. Finally, the conclusion is given in section 6.

2. FEATURE GENERATION

We assume that a size normalized contour C is represented by P contour points. Let the contour C be parameterized by arc length $t : C(t) = (x(t), y(t))$, where $t \in [0, P-1]$. At first, to remove boundary noise we smooth external boundary of original shape. For this purpose, the coordinate functions of C are convolved with a Gaussian kernel. This kernel is the result of n -times convolution of a uniform filter where n is proportional to the arc length.

Fig. 1 shows a Gaussian filter generated by the $n=s/30$ times convolution of a uniform filter where s is the arc length. In this shape s equal to 300.

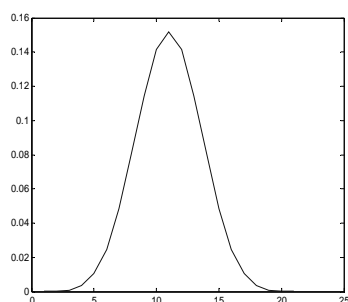


Fig. 1 – Gaussian filter generated by the $n=10$ time convolution of a uniform filter

Fig. 2 shows the effect of the Gaussian filter. It is obvious that the resulted contour after filtering has less surface noise (pixelization artefacts).

The curvature function $k(u, \delta)$ is calculated from the resulted shape as follows [4]:

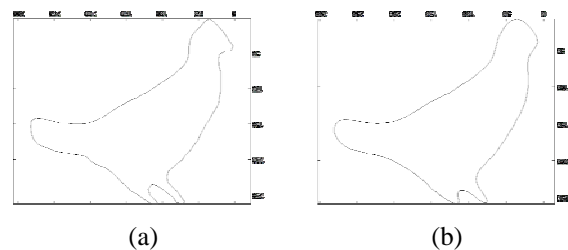


Fig. 2 –removing surface noise (a) and (b) shape before and after noise removing, respectively

$$k(u, \delta) = \frac{x_u(u, \delta)y_{uu}(u, \delta) - x_{uu}(u, \delta)y_u(u, \delta)}{(x_u(u, \delta)^2 + y_u(u, \delta)^2)^{3/2}} \quad (1)$$

$$x_u(u, \delta) = \frac{\partial}{\partial u}(x(u) * g(u, \delta)) = x(u) * g_u(u, \delta) \quad (2)$$

$$x_{uu}(u, \delta) = \frac{\partial^2}{\partial u^2}(x(u) * g(u, \delta)) = x(u) * g_{uu}(u, \delta) \quad (3)$$

$$y_u(u, \delta) = y(u) * g_u(u, \delta) \quad (4)$$

$$y_{uu}(u, \delta) = y(u) * g_{uu}(u, \delta) \quad (5)$$

As it can be seen in (1), the curvature function depends on the first and second derivation of x and y . Here, we have an explanation about curvature function and derivation of x and y and is a suitable explanation for shape description. Then a new equation for curvature function is created. that is a suitable explanation for shape description.

As we know, curvature is the relation of the measured angle to the measured distance $\kappa = d\phi / ds$. So curvature has a direct relation to the tangent at one contour point. As we know, tangent at one point $s(t)$ is defined as follow:

$$v(t) = \overset{\circ}{s}(t) = \lim_{h \rightarrow 0} \frac{s(t+h) - s(t-h)}{2h} \quad (6)$$

$$= (\overset{\circ}{x}(t), \overset{\circ}{y}(t))$$

$$\overset{\circ}{s}(t) = \lim_{h \rightarrow 0} \overset{\circ}{s}_h(t) \Rightarrow \quad (7)$$

$$v_h(t) = \overset{\circ}{s}_h(t) = \frac{s(t+h) - s(t-h)}{2h} = (\overset{\circ}{x}_h(t), \overset{\circ}{y}_h(t))$$

where

$$\overset{\circ}{x}_h(t) = \frac{x(t+h) - x(t-h)}{2h} \quad (8)$$

$$\overset{\circ}{y}_h(t) = \frac{y(t+h) - y(t-h)}{2h} \quad (9)$$

According to the above Equations, $\kappa_h(t), \theta_h(t)$ can be defined as below:

$$\kappa_h(t) = \frac{\overset{\circ}{x}_h(t)\overset{\circ\circ}{y}_h(t) - \overset{\circ\circ}{x}_h(t)\overset{\circ}{y}_h(t)}{(\overset{\circ}{x}_h(t) + \overset{\circ}{y}_h(t))^{3/2}} \quad (10)$$

$$\theta_h(t) = \arg(v_h(t)) = \arctan\left(\frac{\overset{\circ}{y}_h(t)}{\overset{\circ}{x}_h(t)}\right) \quad (11)$$

According to the new curvature Equation, for obtaining the first and the second derivation of x and y in the desired point, Instead of using its neighbourhood points, we can use farther points such as $h=s/10$. In this case, κ_h shows a more general behaviour toward κ .

Far distance from h causes more general behaviour of κ_h at that point and close distance from h to zero causes more local behaviour. For example, according to Equation (10), if $h=0 \rightarrow \kappa_h = \kappa$. Fig.3 shows a typical shape and its curvature functions. As it can be seen in the new function, the changes are not considered local but global. But the curvature function includes local changes.

If shape is stricken with the surface noise, the ordinary curvature function has more changes and experiences serious changes. But it has less effect on new curvature function.

The points corresponded to the extremum of the curvature function are considered as the feature points. Fig. 4 shows the feature points of a typical shape that are obtained according to extremums of the curvature function.

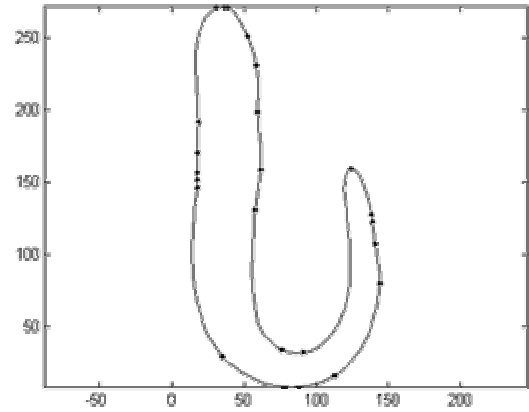
The obtained important points participate in several stages for extracting features. At first, principle axes are found for determining the starting point for feature generation. The computation of an object's principal axis will be showed at the next section.

The shape-based features included turn angle, central angle, distance between two feature points, distance between center of masse and feature point are generated at selected pints.

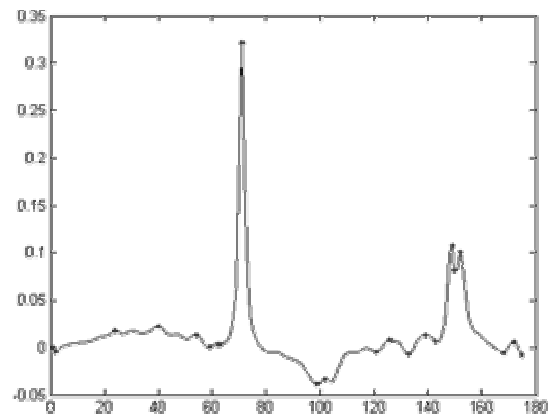
The curvature turning angles are defined as the angles made by two vectors: one that joins two contour's consecutive representative points and the other defined as the object's principal axis where the starting point in question is located (Fig.5). This feature is calculated as follows:

$$t(i) = \tan^{-1}\left(\frac{y'(i+1) - y'(i)}{x'(i+1) - x'(i)}\right) \quad (12)$$

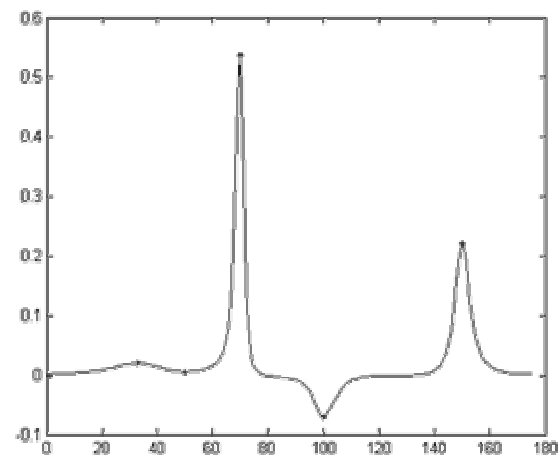
where $i=0,1,\dots,P'-1$ and $(x'(i), y'(i))$ is the contour's coordinates and P' is the number of selected points.



(a)



(b)



(c)

Fig. 3 – Curvature functions (a) Original shape (b) Its curvature function κ (c) Its new curvature function κ_h

Fig.6 shows a shape with the corresponded features, wherein g_1 is center of mass, g_2 and g_3 are feature points. 1 represents central angle, 2 and 3 show distance of center and feature points and 4 shows distances between two feature points.



Fig. 4 – Feature points of a shape image

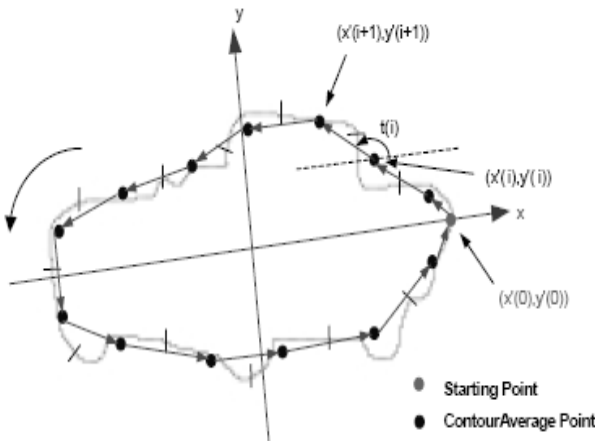


Fig. 5 – Extraction of curvature turning angles

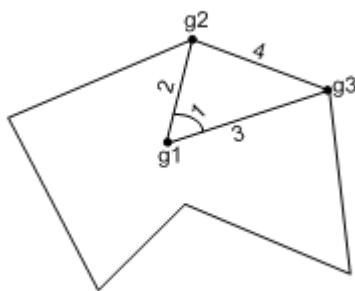


Fig. 6 – Representation of features

For calculation of these features, at first the object's centre of mass is determined:

$$m_x = \frac{1}{L} \sum_{i=0}^{L-1} x_i \quad (13)$$

$$m_y = \frac{1}{L} \sum_{i=0}^{L-1} y_i \quad (14)$$

where L represents the number of contour shapels (the shapels are the pixels belonging to the object support), (x_i, y_i) and (m_x, m_y) are the i -th contour shapel coordinates and the object's centre of mass, respectively.

Distance of center and feature points are obtained as follow:

$$D_1(i) = ((x'(i) - m_x)^2 + (y'(i) - m_y)^2)^{1/2} \quad (15)$$

$$D_2(i) = ((x'(i+1) - m_x)^2 + (y'(i+1) - m_y)^2)^{1/2} \quad (16)$$

$$D_3(i) = (x'(i) - x'(i+1))^2 + (y'(i) - y'(i+1))^2)^{1/2} \quad (17)$$

The fourth feature at each selected point is the central angle:

$$CA(i) = \cos^{-1} \left(\frac{D_1(i)^2 + D_2(i)^2 - D_3(i)^2}{2D_1(i)D_2(i)} \right) \quad (18)$$

Hence the generated features at each point are $t(i)$, $D_1(i)$, $D_3(i)$ and $CA(i)$.

3. FEATURE EXTRACTION

The problem of optimally selecting the statistical features is known as feature selection and feature extraction. While the feature extraction creates a smaller set of features from linear or nonlinear combinations of the original features, the feature selection chooses a subset of the original features. Feature extraction transforms a data space into a feature space that theoretically has the same dimension as the original data space. Such transformation is practically designed in a way that a reduced number of effective features represent the data set while retaining most of intrinsic information content of the data. A reduction in the dimension of the input space is generally accompanied by losing some of information. The goal of dimensionality reduction is to preserve the relevant information as much as possible. In this section, the concepts of PCA, LDA and supervised neural networks that are used to extract the efficient features are explained.

Principle component analysis: The PCA is a well-known feature extraction technique in the multivariate analysis. It is an orthogonal transformation of the coordinate in which the data is described. A small number of principle components are usually sufficient to account for the most of

structure of the data.

Let $\mathbf{X} = \{\mathbf{x}_n \in R^d / n = 1, \dots, N\}$ represents a d-dimensional dataset. The PCA tries to find a lower dimensional subspace to describe the original dataset while preserving the information as much as possible, so a new m-dimensional dataset $\mathbf{Z} = \{\mathbf{z}_n \in R^m / n = 1, \dots, N\}$ will be produced, where m is smaller than d. The orthogonal basis of the feature space is defined as the eigenvectors of the total class scattering matrix:

$$\mathbf{S}_t = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \bar{\mathbf{x}})(\mathbf{x}_n - \bar{\mathbf{x}})^T \quad (19)$$

where $\bar{\mathbf{x}}$ is the mean vector:

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n \quad (20)$$

Here we are dealing with solving the eigenproblem:

$$\mathbf{S}_t \mathbf{u} = \mathbf{u} \mathbf{\Lambda} \quad (21)$$

In which \mathbf{u} is a matrix where its columns are the eigenvectors of \mathbf{S}_t respectively, and $\mathbf{\Lambda}$ is a diagonal matrix consisting of all the eigenvalues $\{\lambda_n / n = 1, 2, \dots, N\}$ of \mathbf{S}_t along its principal diagonal, where its other elements are zero.

Practically, the algorithm proceeds by first computing the mean of \mathbf{x}_n vectors and then subtracting them from this mean value. The total class scattering matrix, or the covariance matrix, is then calculated and its eigenvectors and eigenvalues are founded. The eigenvectors corresponding to the m of largest eigenvalues are retained and the input vectors \mathbf{x}_n are projected onto the eigenvectors to give the components of the transformed vectors \mathbf{z}_n in the m-dimensional space. This projection results in a vector containing m coefficients a_1, \dots, a_m . Hereafter the vectors are represented by a linear combination of the eigenvectors with weights a_1, \dots, a_m .

It must be pointed out that if the number of selected features be equal to the number of initial features, the aim of PCA is a conversion of features distribution for a particular purpose or computation of the principle axes of a shape. We will use PCA for determining the initial point in feature generation.

Linear discriminant analysis: The goal of PCA on choosing a lower dimensional basis is to minimize the reconstruction error. This is not the major concern in pattern recognition applications, whose goal is to maximize the recognition rate.

The linear discriminant analysis provides a procedure to determine a set of axes whose projections of different groups have the maximum separation. This procedure can be described as follows.

The total class scattering matrix is calculated using Eq.(19). Then the between-class scatter matrix is constructed as:

$$\mathbf{S}_b = \frac{1}{N} \sum_{i=1}^c l_i (\bar{\mathbf{x}}_i - \bar{\mathbf{x}})(\bar{\mathbf{x}}_i - \bar{\mathbf{x}})^T \quad (22)$$

where l_i is the number of training samples in class i, $\bar{\mathbf{x}}_i$ is the mean vector of the training samples in class i and $\bar{\mathbf{x}}$ is the mean vector across all training samples. The $\mathbf{S}_t^{-1} \mathbf{S}_b$'s eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m$ corresponding to $m(m \leq d - 1)$ largest eigenvalues are calculated. Then LDA is carried out using

$$\mathbf{y} = \mathbf{G}^T \mathbf{x} \quad (23)$$

where

$$\mathbf{G} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m) \quad (24)$$

Supervised neural networks: Recently, several efforts have been done to use neural networks for feature generation and feature extraction. A possible solution for feature extraction is via multi layer perceptron (MLP). The employed network has d input and Q output nodes and a single hidden layer with m nodes and linear or nonlinear activations, where Q is the number of classes. During the training phase, the desired outputs are matched to the class number of labeled input, so the corresponded output is high and the others are low.

Such a network maps the input d-dimensional space onto the output m-dimensional subspace. This procedure is equivalent to the kernel Fisher discriminant (KFD) analysis [10]. Fig.7 illustrates the architecture of an MLP neural network. The outputs of hidden layer are the compressed data and are called the nonlinear discriminant components.

The first step of using a neural network is the training phase in which the back-propagation algorithm is used. In implementation of such algorithm, two distinct passes of computation may be distinguished. The former is referred to as the forward pass, and the latter as the backward pass. In

the forward pass, the synaptic weights remain unaltered throughout the network, and the function signals of the network are computed on a neuron-by-neuron basis. The Backward pass, on the other hand, starts at the output layer by passing the error signals leftward through the network, layer by layer. Weights and biases are updated iteratively until the MSE is minimized.

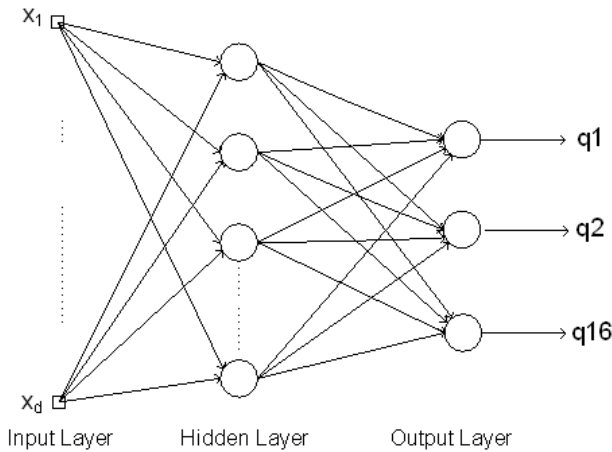


Fig. 7 – Architecture of an MLP neural network

4. PROPOSED METHOD

In this section we explain the proposed method based on generated features, which defined at section 2. Fig.8 shows the flowchart of the proposed method.

At first, the boundary of the shape is determined and after smoothing, the new explained curvature function is calculated. Its extremums are determined as feature points. Then, PCA maps their coordinates to a new space. The transform matrix is obtained by the feature points coordinates. In the new space the feature points are arranged by their angles from positive part of x axis. Since all of the feature points participate on the transform, missing or variation a few numbers of them does not serious effect on accuracy of the feature point’s arrangement.

For making the feature vector for each shape and for each kind of features that generated at previous stage, the new two dimensional plane after PCA is divided into N equal angle sectors. The vertex of these angles is the origin of the plane. The M features got from shape at each sector are localized in the feature vector regularly. M is the maximum number of features that is expected for considered database at each sector. If the number of feature points at each sector be smaller than M , the remained locations for features set to zero. Hence we have $M \times N$ features for each shape. This arrangement tends to meaning of the locations at feature vector. However, the multiplicity of features causes to fall down the speed and accuracy of the

recognition. Hence, a feature extraction method is needed to make a new dataset with the smaller dimensions. The LDA is a fast procedure, but it is poor when compared to the nonlinear mapping methods. On the other hand, using a MLP network with nonlinear transfer functions has a good performance. However, the computational load and the risk of being trapped in the local minima reduce its efficiency. To solve such problems, we use a hierarchical feature extraction method that serves LDA and a neural network in order to improve the speed and accuracy simultaneously.

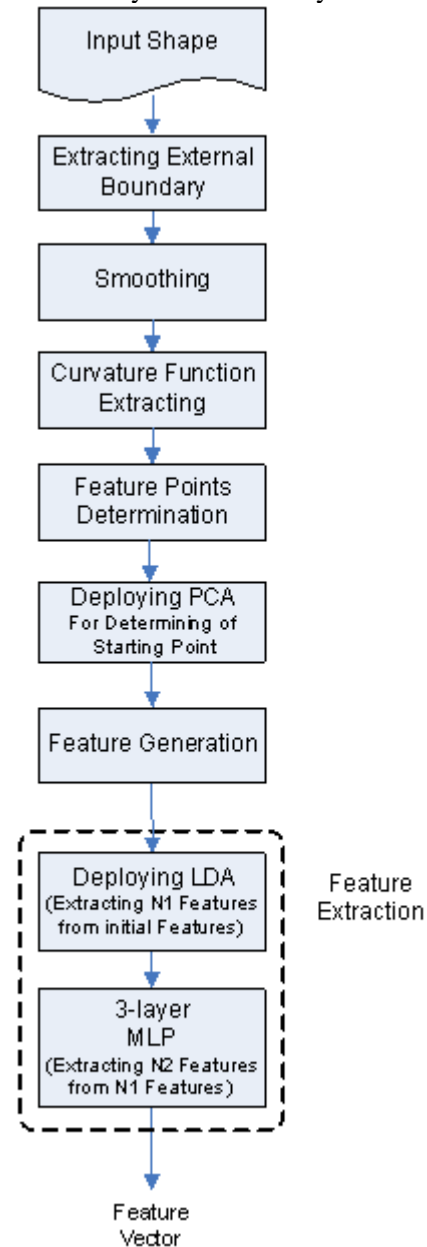


Fig. 8 – The proposed algorithm

At the training phase of the feature extraction stage, LDA projection matrix is calculated based on training dataset. The total class and the between-class scatter matrices are firstly calculated based on the training dataset, as explained in section 3. Then

the corresponding eigenvectors of the N_1 largest eigenvalues of the $S_1^{-1} S_b$ are selected to construct the projection matrix.

Output features from LDA are fed to a MLP neural network that its architecture was previously explained. The number of neurons at the first layer and the output layer are equal to N_1 and C, respectively. The number of neurons at the third layer is equal to N_2 (where $N_2 < N_1$). These N_2 features will be used at retrieval stage. This neural network is trained by using the back-propagation algorithm to minimize the MSE. The problem of trapping in the local minima at the training phase of neural network can be avoided by introduction of small disturbances on the weights after each epoch and evaluating and checking the results, thereby the step size or the learning rate can be controlled.

At the retrieval phase, the trained hierarchical feature extraction stage consists of the LDA and a MLP neural network provides the extracted features from the initial features got from the input shape. The similarity measure at retrieval stage is normalized Euclidean distance which is given by

$$NED(\bar{f}^x, \bar{f}^y) = \sum_m d_m(\bar{f}^x, \bar{f}^y) \quad (25)$$

where

$$d_m(\bar{f}^x, \bar{f}^y) = \frac{|f_m^x - f_m^y|}{\sigma(f_m)} \quad (26)$$

where $\sigma(f_m)$ is the standard deviation of the respective features over the entire database and are used to normalize the individual feature components. The numbers of ground truth images for each query image in the database are 9. The performance is measured in terms of the average retrieval rate which is defined as the average percentage number of patterns belonging to the same image as the query pattern in the top 9 matches.

5. EXPERIMENTAL RESULTS

We used 16 class of shapes obtained from the MPEG-7 database that each class consist of 12 shapes. Fig.9 shows one shape from each class. We chose 3 shapes from each class randomly for training of feature extractor. To making a rich dataset for training, at first each shape from training set was resized with different factors included 1/2, 1/4, 2 and 4. Then the premier shape and each of the yield shapes were rotated with 5 random angles from 0 to 360 degree. Therefore, training dataset has 75

shapes from each class and total number of training dataset is equal to 1200. The remained 9 shapes from each class were used for evaluation of retrieval performance based on extracted features.

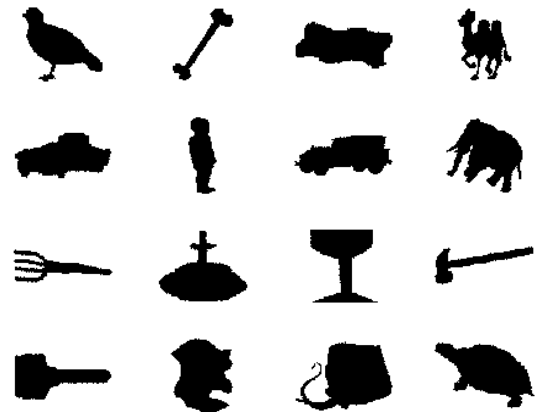


Fig. 9 – Sample shapes from database

At the feature extraction stage, at first 20 features from each kind of features are selected by LDA. Since 4 kinds of features are used for processing, the total number of features for each shape after PCA is equal to 80. Then MLP selects 12 features from 80 features for retrieval.

Fig.10 shows average retrieval accuracy rate for each shape, separately, based on proposed approach (Method II) as well as presented method in [9] (method I). It indicates that the proposed method is effective at all of classes. The average accuracy retrieval rate is equal to 91.7 % and is better than 87.5% given in[9].

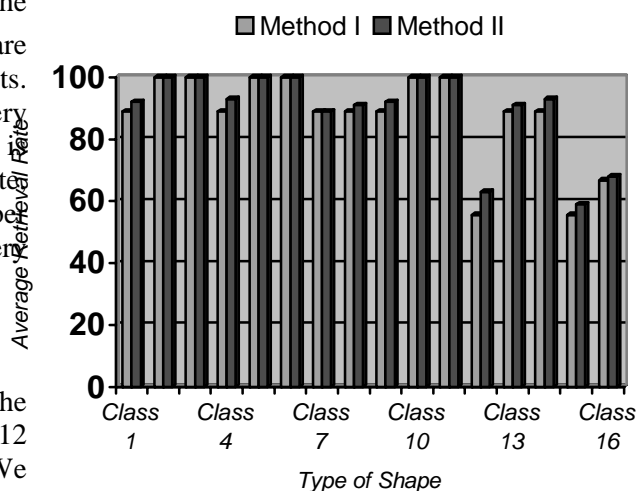


Fig. 10 – Average accuracy rate of retrieval for each shape

On the other hand, the correct retrieval rate without using neural network with same features is 84.7%. The other advantage of implementation of neural network is reducing of retrieval time due to

lower dimensional of feature vector at the similarity measurement stage.

6. CONCLUSION

We have introduced a new method for image retrieval based on shape features. Initial features are obtained at feature points, consist of turn angle, central angle, distance between two feature points, distance between center of masse and feature point. The registration of shape for determining the starting point and so stabilizing of method against geometrical transforms such as translation, rotation and scaling is done by a PCA transform. Then, for each shape, the smaller and more efficient set of features is obtained via a hierarchical supervised feature extractor consist of LDA transform and a supervised neural network (MLP). The proposed method showed high performance in the retrieval of the MPEG-7 shape database.

6. REFERENCES

- [1] S. Loncarin, A survey of shape analysis techniques, *Pattern Recognit.* 31 (5) (1998), pp. 983–1001.
- [2] R. Veltkamp and M. Hagedoorn, State of the art in shape matching, Utrecht, The Netherlands, Tech. Rep. UU-CS-1999-27, 1999.
- [3] E. G. M. Petrakis, A. Diplaros, and E. Millos, Matching and retrieval of distorted and occluded shapes using dynamic programing, *IEEE Trans. Pattern Anal. Machine Intell.* vol. 24 (11) (2002), pp. 1501–1516.
- [4] M. Bober, MPEG-7 visual shape descriptors, *IEEE Trans. Circuits Syst. Video Technol.* 11 (6) (2001), pp. 716–719.
- [5] H. Kauppinen, T. Seppanen, and M. Pietikainen, An experimental comparison of autoregressive and fourier-based descriptors in 2-D shape classification, *IEEE Trans. Pattern Anal. Machine Intell.* 17 (2) (1995), pp. 201–207.
- [6] H. Blum, A transformation for extracting new descriptors of shape, in *Whaten-Dunn, Editor, Models for Perception of Speech and Visual Forms.* Cambridge, MA: MIT Press, 1967, pp. 362–380.
- [7] -----, Biological shape and visual science (part1), *J. Theoretic. Biol.*, no. 38, 1973, pp. 205–287.
- [8] T. Sebastian, P. Klein, and B. Kimia, Recognition of shapes by editing shock graphs, in “*Proc. 8th Int. Conf. Computer Vision*”, Vancouver, BC, Canada, 2001, pp. 755–762.
- [9] M.Ch.Amirani, Z.Sadeghi and A.A.Beheshti,

Shape-based Image Retrieval Using a Neural Network, Proceedings of the “*Fifth International Conference on Neural Networks and Artificial Intelligence (ICNNAI’2008)*”, Minsk, Belarus, 27-30 May 2008.

- [10] Gaetan Kerschen1 and Jean-Claude Golinval, Feature extraction using auto-associative neural networks, *Smart Materials and Structures*, 13 (2004),pp. 211–219.



Mehdi Chehel Amirani, received the B.S. degree in Electronic Engineering from Urmia University, Iran, in 1993 and the M.S. degree in Communication Engineering from Iran University of Science and Technology (IUST) in 1998. He is currently working toward the Ph.D. degree in Department of Electrical Engineering, IUST. His research interests include image analysis, pattern recognition and artificial intelligence.



Zahra Sadeghi Gol, received the B.S. degree in Electronic Engineering from Iran University of Science and Technology (IUST) in 2004. She is currently working toward the M.S. degree in Department of Electrical Engineering, IUST. Her research interests include content based image retrieval.



Ali Asghar Beheshti Shirazi, received the B.S. and M.S. degrees in Communication Engineering from Iran University of Science and Technology (IUST) in 1984 and 1987, respectively and Ph.D. from Okayama University, Japan in 1995. In 1995, he joined the Department of Electrical Engineering, IUST, where he currently is an assistant Professor. His research interests include digital image processing, data communication networking and secure communication.