

Research Note

Improvement of Supervised Shape Retrieval by Learning the Manifold Space

Mohammad Ali Zare Chahooki
Electrical & Computer Engineering
Tarbiat Modares University
Tehran, Iran
chahooki@modares.ac.ir

Nasrollah Moghadam Charkari
Electrical & Computer Engineering
Tarbiat Modares University
Tehran, Iran
charkari@modares.ac.ir

Received: September 20, 2011- Accepted: October 28, 2011

Abstract—Manifold learning is the technique that aims for finding a constructive way to embed the data from a high-dimensional space into a low-dimensional one based on non-linear approaches. In this paper a supervised manifold learning method for shape recognition is proposed. The approach is based on learning the manifold space for training samples, and maps the test samples to the learned space by a Generalized Regression Neural Network (GRNN). The main goal in this paper is to propose a new feature vector to coincide semantic with Euclidean distances. To accomplish this, the desired topological manifold is learnt by a global distance driven non-linear feature extraction method. The experimental results indicated that the geometrical distances between the samples on the manifold space are more related to their semantic distance. To fuse the results of shape recognition based on contour and region based methods, the final result of shape recognition is based on committee decision in three manifold spaces. The experimental results confirmed the effectiveness and the validity of the proposed method.

Keywords—object recognition; shape retrieval; shape annotation; manifold learning; non-linear dimension extraction

I. INTRODUCTION

Today, due to advances in imaging devices and memory technologies, the storage of huge number of images has been possible in computers. Large volume of information is encoded in images. It is interesting to mention that the ancients were aware of the importance of images. Confucius- a Chinese philosopher said in 500 BC that "A picture is worth one thousand words". The study of automatically indexing the image databases has been one of important issues in information and communication technologies. Much research has been carried out on image retrieval (IR) in the last two decades. In general, IR research efforts are divided into two types of approaches. The first approach is based on annotation by textual information around the image [1].

Especially, this has been done in search engines like Google by file names, image captions, Alternate Text (ALT), HTML titles, and hyperlinks. The second approach is based on image labeling by low-level features extracted from image content. This approach has an important use in semantical web filtering, image mining and multidimensional indexing of image databases. Typically, the features used in content based image indexing are texture, colour, shape and spatial relation of objects. In many cases, shape in IR is more powerful for identification than colour and texture. On the other hand, it has been found that 71% of users are interested in retrieval by shape [2].

Topology describes the object property that is not changed by deformations. Only tearing cause changes in topology. For example, circle and ellipse have the same topology, and sphere is topologically equivalent

to ellipsoid. Topological space is defined by a set of topologically equivalent objects. A manifold M is a topological space that is locally Euclidean. It means that there is a neighborhood around every point of M that is topologically the same as the open unit ball in R^D . In general, any object that is nearly “flat” on small scales is a manifold [3]. An open line segment, a circle and a knotted circle are 1-manifold ($d=1$) that are mapped into one, two or three dimensional space, respectively. This means that, although the mapping spaces of these samples are different, they have similar intrinsic dimensions.

Intrinsic dimensions that are extracted from the samples of feature space. They are not structured in a linear manner. Therefore, linear feature extraction methods such as Principal Component Analysis (PCA) are not able to discover the latent structures [4]. In this regard, non-linear feature extraction algorithms, also known as manifold learning methods, would be suitable to discover the intrinsic dimensions by the use of graphs and new metrics like geodesic distance. The aim of this type of algorithms is to map a set of high-dimension data set to a lower dimension data set.

Non-metric Multidimensional Scaling (MDS) [5], Isomap [6], LLE [7], Hessian LLE [49], Laplacian eigenmap [8], Local Tangent Space Alignment (LTSA) [9], Maximum variance [10], Sammon’s nonlinear mapping (NLM) [11], Curvilinear Component Analysis (CCA) [12] and Curvilinear Distance Analysis (CDA) [13] are the basic manifold learning methods.

Using manifold learning approach in supervised shape retrieval is the main innovation in this paper. Isomap which is an unsupervised manifold learning method is used for supervised shape manifold learning in multiple feature spaces. We describe shapes in three feature spaces. For each one, manifold space is learnt for training samples. In manifold space related to each feature space, Euclidean distance between training samples is compatible with semantic distance. In order to have an extension to map out-of-samples (test samples) into manifold space, a Generalized Regression Neural Network (GRNN) is employed. Classification by kNN in each space by manifold space is more accurate than observation space. Final classification is based on committee of classifiers.

The experimental results demonstrated that the geometrical distance between the samples on the manifold space is closely related to the semantic distance between them.

The paper is organized as follows. An overview to supervised shape retrieval is introduced in section 2. Section 3 presents a brief review of the manifold learning in machine vision applications. Section 4 details our approach to shape classification. Section 5 presents experimental results. Section 6 concludes by summarizing the contribution as well as discussing potential future work.

II. SUPERVISED SHAPE RETRIEVAL

There exist many shape description methods, which are categorized into contour- and region-based ones. Contour-based methods use edge points, while

region-based methods employ all points of the shape. Some interesting surveys of various shape description methods have been introduced in [14][15][16]. In addition, there are some efforts in shape retrieval based on learning methods.

However, the majority of shape retrieval methods are based on unsupervised approaches. Bicego et al. [17] represented a supervised scheme for sequences based on Hidden Markov Models (HMMs). In their approach, shape was described by the vector of its similarities with respect to a predetermined set of other shapes, where the similarities were supported by HMMs. They also in [18] investigated HMMs for the purpose of classifying planar shapes represented by their curvature coefficients. Thakoor et al. [19] used HMM for identification of each shape class as well. Accordingly, a reference path for each class was built from the corresponding HMM, which was the optimal path followed by the most likely example shape. To classify a shape, its optimal path through HMM was calculated and warped to match the reference path using dynamic time warping (DTW). Correct class is identified as the one for which the warping cost is minimum. In [20], HMM is used for shape curvature as its 2-D shape descriptor. The method is different from the traditional maximum likelihood (ML) ones, in which classification is based on probabilities from independent individual class models, utilized information from all classes to minimize classification error.

Gorelick et al. [21] developed shape classification based on a decision trees framework. They assumed that some features are more prominent in some classes of shapes than the others. Therefore, they employed different combinations of features at different steps of the classification algorithm.

In the method developed by Daliri et al. [22], the dissimilarities between pairs of shapes were transformed into suitable kernels which were classified using support vector machines.

Wang et al. [23] used shape tree which consists of junction nodes connected with several main skeleton paths. They used Bayesian classifier in their supervised shape classification. Sun et al. [24] also employed Bayesian classification within a three-level framework which consists of models for contour segments, for classes, and for the entire database of training examples. Bayesian classification has been also used in Bai et al. [25] on graphs of shape skeleton.

Vassilis et al. in 2011 [26] used fuzzy lattice reasoning to learning the fusion of different shape descriptors.

III. MANIFOLD LEARNING IN MACHINE VISION

Manifold learning is used in different machine vision applications such as CBIR [27][28][29], shape analysis [30][31], face recognition [32][33][34][35], facial expression recognition [36][37][38], tracking [39][40], action recognition [41][42], and pose estimation [43][44]. A brief review of the manifold learning in machine vision applications are described below.



Shape Analysis: Xiao approach [30] describes each shape by an especial graph. Isomap was applied to the graph and an extracted set of points indicates each shape in the data set. So matching two shapes is performed by matching two point sets in different spaces and corresponding point-sets performed by semi-definite programming.

Face Recognition: Arandjelovic [34] considered face recognition from Face Motion Manifolds (FMMs). In addition, it is shown how geodesically local FMM structure is modelled. The model automatically leads to a stochastic algorithm for generalizing the unseen modes of data variation. Arandjelovic [35] recognized faces by using video sequences both for training and test samples, in a realistic, unconstrained setup in which lighting, posing, and user motion pattern have a wide variability and face images are of low resolution.

Facial Expression Recognition: Chang [38] used a modified embedding Lipschitz [4] to embed aligned facial features in a low-dimensional space while keeping the main structure of the manifold. In the embedded space, a complete expression sequence becomes a path on the expression manifold emanated from a centre that corresponds to the neutral expression.

Tracking: Lee [40] tightly couples the tracking and recognition modules within a single framework. The complex nonlinear appearance manifold of each registered person is partitioned into a collection of sub-manifolds where each models the face appearances of the person in nearby poses. The sub-manifold is approximated by a low-dimensional linear subspace computed by principal component analysis using images sampled from training video sequences. The connectivity between the sub-manifolds is modeled as transition probabilities between pairs of sub-manifolds and these are learned directly from training video sequences.

Action Recognition: Wang and Suter [42] assumed that a given sequence of moving silhouettes associated to an action video. By LPP project them into a low-dimensional space to characterize the spatiotemporal property of the action and to preserve much of the geometric structure as well.

Pose Estimation: Yan [43] considered the pose data space as a union of sub-manifolds which characterizes different subjects instead of a single continuous manifold as conventionally regarded. A manifold embedding algorithm, dually supervised by both identity and pose information, which is proposed for person-independent precise 3-D pose estimation means that the testing subject may not appear in the model training stage.

IV. LEARNING THE SHAPE MANIFOLD

In supervised feature extraction methods, dataset is divided into training and test data. The class labels are known in training data set. Therefore, the information provided by class labels can be used to execute the dimensionality reduction. This can be called supervised dimensionality reduction in comparing to the unsupervised scheme of the most dimensionality

reduction methods like PCA, MDS and Isomap. Classical manifold learning methods, like Isomap created by computational methods, are influenced on noise-less samples such as Swiss roll effectively.

In specific applications such as shape analysis and object recognition, multiple feature spaces are extracted. The method proposed in this paper is based on fusion of multiple shape feature spaces. Fusion is done in decision phase. Decision for shape classification is based on kNN for each space. Euclidean distance between the test and training samples is the basis for determining accuracy in kNN classification. In each space, to reduce the semantic gap, it is important the compatibility of Euclidean and semantic distance of feature vectors.

In this paper, samples are described in multiple feature spaces. The purpose of this approach is the reduction of semantic gap by analyzing multiple feature spaces as well as the classification methods.

Our proposed method for shape classification is done in four phases. (I) Learn the manifold space of trained samples in each feature space, (II) learn the map from observation space to manifold space, (III) map test samples to manifold space in each space and (IV) classifying the test sample based on committee of multi-classifiers in each space.

Phase I) Learning the manifold space

Learning the manifold space of trained samples in each space is explained as follow:

1. Start from a shape data set $X = \{x_1, x_2, \dots, x_N\}$ with N shape samples. Consider T and S are set of Train and Test shapes with N_T and N_S members where $X = T \cup S$ and $N = N_T + N_S$.
2. $S = \{s_1, s_2, \dots, s_K\}$ is the set of K shape feature spaces. x_i in each feature space is described as x_i^k , a D_k dimensional feature vector.
3. Construct a dissimilarity graph $G_k = (V, E_k)$ in each feature space $s_k \in \{s_1, s_2, \dots, s_K\}$ where K is the number of shape feature spaces. V is the set of N_T training samples and $E_k = V \times V$.
4. x_i and x_j are two training data samples which would be assigned to $(x_i, x_j) \in E_k$ and referred by $edge_{ij}^k$ that is equal to Euclidean distance between x_i and x_j in s_k ($\|x_i^k - x_j^k\|$). Because of dissymmetry, $edge_{ij}^k \neq edge_{ji}^k$.
5. For each $(x_i, x_j) \in E_k$, decrease the noise of dissimilarity by set $edge_{ij}^k = \varepsilon \times edge_{ij}^k$ where x_i and x_j are members of the same class, and set $edge_{ij}^k = \text{Inf} \times edge_{ij}^k$ where x_i and x_j are not member of the same class.
6. The neighbourhood graph, NG_k of G_k , is constructed. This is done by connecting nodes i and j if they are closer than e or j is one of the E -nearest neighbours of i .
7. The NG_k can also be defined as a geodesic distance matrix. The shortest path of pair wise nodes of G_k defines The geodesic distance matrix, GD^k .
8. The construction of the manifold space f_k , in each observation space s_k , is done by using MDS on geodesic distance matrix, GD_k , in each space.



Classical scaling is concerned with the converse problem: Given a matrix of Euclidean distances (GD_k), how we can determine the coordinates of a set of points in a dimension d_k . This is achieved via a decomposition of $N \times N$ matrix T^k between the individual sums of squares and the products matrix $T^k = Y^k Y^{kT}$. Here, $Y^k = \{y_1^k, y_2^k, \dots, y_N^k\}^T$ is the $N \times d_k$ matrix of coordinates.

$$T_{ij}^k = \frac{-1}{2} [GD_{ij}^{k2} - GD_{i.}^{k2} - GD_{.j}^{k2} + GD_{..}^{k2}] \quad (1)$$

Where, $GD_{i.}^{k2} = \frac{1}{N} \sum_{j=1}^N GD_{ij}^{k2}$, $GD_{.j}^{k2} = \frac{1}{N} \sum_{i=1}^N GD_{ij}^{k2}$ and $GD_{..}^{k2} = \frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N GD_{ij}^{k2}$. Equation 6 leads us to construct T^k from a given dissimilarity matrix GD^k . The matrix T^k is factorized to convert in the form of $T^k = Y^k Y^{kT}$. Since it is a real symmetrical matrix, it can be written in the form of $T^k = U \Lambda U^T$, where the columns of U are the eigenvectors of T^k and Λ is a diagonal matrix of eigenvalues. Therefore, we consider $V = U \Lambda^{\frac{1}{2}}$. If a representation in a reduced dimension is searched, then those eigenvectors that are associated with the largest eigenvalues will be used. d_k is chosen for some pre-specified threshold, $\alpha (0 < \alpha < 1)$.

$$\sum_{l=1}^{d_k-1} \lambda_l < \alpha \sum_{l=1}^N \lambda_l < \sum_{l=1}^{d_k} \lambda_l \quad (2)$$

Then, Y^k is found as follow;

$$Y^k = [u_1, \dots, u_{d_k}] \text{diag} \left(\lambda_1^{\frac{1}{2}}, \dots, \lambda_{d_k}^{\frac{1}{2}} \right) = U_{d_k} \Lambda_{d_k}^{\frac{1}{2}} \quad (3)$$

Phase II) Learn the map to manifold space

As mentioned, a manifold space, f_k , is learned for each observation space, s_k , by the proposed approach in previous phase. In the learned space, mapping the observations into a new space with fewer dimensions and more rapprochements between semantical and Euclidean dissimilarities are possible.

Phase I includes some technical difficulties that arise when one uses the above procedure to map out-of-samples without including them in the learning stage. The difficulty of how to insert the out-of-sample objects into the configuration of points that represents the original objects in Euclidean space is an out-of-sample embedding problem [45]. A proper approach to out-of-sample embedding depends on how the original sample was embedded. Classic linear feature extraction approaches, such as PCA, only make a transition matrix for mapping in to the new low-dimensional space. Out-of-sample is multiplied by the transition matrix and be mapped to new space. The use of transition matrix in non-linear feature extraction approaches is not correct in the presence of non-linear background.

Mapping out-of-samples into manifold space can be done by two analytical and soft computing approaches.

In analytical approaches, a function like $y=f(x)$ which maps new sample, x , to manifold space should be found.

On the other hand, in soft computing approaches, two neural networks, Radial Basis Function (RBF) and General Regression Neural Network (GRNN) can be used. The two networks have the property that after supervised feature extraction, mapping the training samples can be done with high precision [46].

The proposed method is non-analytical. The use of neural network, initially proposed in LLE [7], but it did not mention the type of neural network. In this study, we use GRNN to map out-of-samples to manifold space.

So, N^k , as a GRNN is trained to map samples from observation space, s_k , to manifold space, f_k . The input matrix to N^k is a $D_k \times N_T$, observation training samples and the output matrix become a $d_k \times N_T$, training samples in manifold space.

Phase III) Mapping test samples to manifold space

In each spaces s_k , map the $D_k \times N_S$, matrix of the test samples into the observation space Z^k , $d_k \times N_S$, matrix of the test samples in manifold space.

Phase IV) Committee of classifiers in manifold space

For each manifold space, use kNN^k as classifier. The final classification result is based on committee opinion of K classifiers.

The method is capable of learning separable manifold spaces for each observation space. The learned manifold spaces have lower dimensionality than the corresponding observation space. In the learned manifold space, samples with geometric distances are closer to their conceptual distance. Learning the new space is only based on the training (labeled) samples and the classification (labeling) of the test (unlabeled) samples is based on kNN, in each manifold space. The final classification is based on voting of classifiers. So that the vote of each classifier is based on the corresponding manifold space.

The label of training samples are used in classification and extracting a new space in which the dimensionality of samples are closer to their intrinsic dimensionality. Therefore, it improves the accuracy of the classification method.

V. EXPERIMENTAL RESULTS

This section describes the experiments of the proposed approach. Section 4.1 explains three shape indexing methods ($K=3$) and section 4.2 evaluates the proposed manifold learning methods on Kimia-216 [47] shape dataset.

A. Shape Feature Spaces

Three spaces ($K=3$) of shape descriptors are employed in the experiments. To obtain the description vector in the first space (s_1), the shape image should be normalized. So, the larger dimension of image is changed to 250 pixels without altering the aspect ratio of image size. Then, the contour pixels are smoothed by four neighborhood pixels and their distances from the centre gravity point are calculated, Fourier coefficients are computed and divided by the first Fourier coefficient. The absolute values of the coefficients result are computed and the description



vector is created with the first 60 absolute values. The first Fourier coefficients are the most important values. Therefore, other absolute values will be ignored. In the experiments, shapes are indexed with $D_1=60$ dimensional vectors.

In the second space (s_2), similar to s_1 , shape is normalized with 250 pixels and the contour points are smoothed. For each point on the contour, the distance to the farthest point on the contour is computed. Then, the Fourier transform of the distances and the absolute value of the coefficients will be found. The absolute values are divided by the first coefficient. The description vector is created by the first 30 absolute values for each shape image. Therefore, in this space, shapes are indexed with $D_2=30$ dimensional vectors.

Shape images are normalized to 151×151 pixels in the third space (s_3). The polar coordinates are evaluated by the normalized shapes. Zernike moments will be evaluated up to 8th level. Therefore, shapes are indexed with $D_3=25$ dimensional vectors in this space.

B. Kimia-216 dataset

In this section we test our manifold learning method for supervised shape retrieval on Kimia-216 [47] and compare the results with the state-of-the art approaches for the shape classification. Kimia-216 is a subset of dataset MPEG-7 Part B [48]. In particular, Kimia-216 includes 216 binary images classified in 18 classes with 12 images per class. Fig. 1 shows sample shapes for each class as well as elephant class shapes.

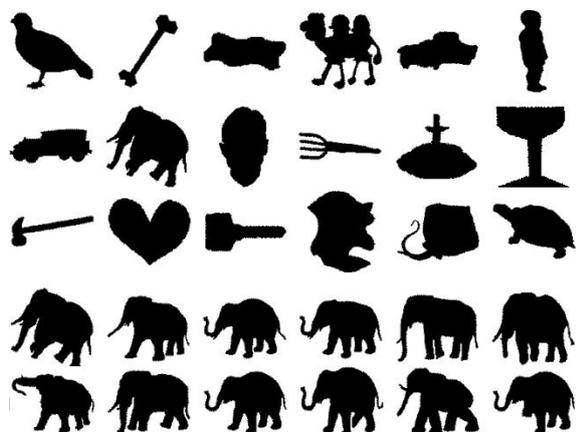


Figure 1. Classes of Kimia-216 dataset and samples of elephant class shapes.

Experiments are based on “standard” leave-one-out cross validation classification that each of the 216 samples, in turn, is left out for testing, whereas all the remaining 215 samples were employed for training that is, no validation data were employed at all.

First, we carried out the experiments by using feature vectors in observation spaces 1, 2 and 3 separately by a kNN ($k=3$) classifier. We recorded 6 misclassifications in 216 experiments based on a committee of three classifiers. The accuracy is 97.22% as shown in TABLE I.

Second, we carried out 216 Leave-1-Out classification experiments using learning the manifold space on three observation spaces by a kNN ($k=3$) classifier. Then, similar to classification in observation

space, we used a committee of three classifiers. We recorded 3 misclassifications in 216 experiments. The accuracy is 98.61% as shown in TABLE I. In learning the manifold space ϵ and Inf and E are set to $10^{-0.5}$, 3.5 and 23.

The results of Table I, shows the effectiveness of the proposed method in improving the accuracy for each of the three feature spaces. In each of observation space, initially the graph is constructed based on the labeled samples. Then, the conceptual mapping from semantical space to observation spaces is performed according to Phase 1 of the proposed method. Mapping to manifold space for new samples (Phase 3) is done by the learned GRNN as presented in Phase 2. The improvements are due to the closed distances between the geometrical and conceptual ones in manifold spaces. This is indebted to the first, the conceptual mapping from semantical spaces to the observation spaces, and second, removing the irrelevant neighborhoods which have had some negative effects in finding the geodesic distances of samples. GRNN has the property that mapping of training samples can be done with high precision. The same thing is true for the test samples. The experimental results are confirmed the proposed expectation. Fig. 2 illustrates the misclassification improvement in the proposed method. It shows the improvement in classification results. Therefore, it indicates that the geometrical distances from test samples to train ones on the manifold space are more related to their semantic distance.

TABLE I. COMPARISON OF ACCURACY IN OBSERVATION AND MANIFOLD SPACES. THE FINAL ACCURACY IS BASED ON COMMITTEE IN EACH SPACE.

Observation		Manifold		Improvement rate
Space	Accuracy	Space	Accuracy	
s_1	95.37	s_1	96.76	1.39%
s_2	93.52	s_2	95.83	2.31%
s_3	91.20	s_3	94.44	3.24%
Committee Accuracy		Committee Accuracy		
97.22		98.64		1.42%

TABLE II. CLASSIFICATION RESULTS OF LEAVE-1-OUT EXPERIMENTS ON THE KIMIA-216 DATASET

Method	Classification Accuracy
Committee in Manifold Space	98.64
Tree-Union [23]	97.7
Committee in Observation Space	97.22
Class-Segment-Sets [24]	97.2
Lattice-Computing [26]	95.4
Skeleton-Based [25]	94.1

Table II shows the comparison of the proposed method with others on the Kimia-216 Dataset. We obtained an accuracy of 98.64%, which shows the highest ever reported accuracy value in the literature on this data set. This table shows the highest ever reported accuracy value in the literature on this data set.



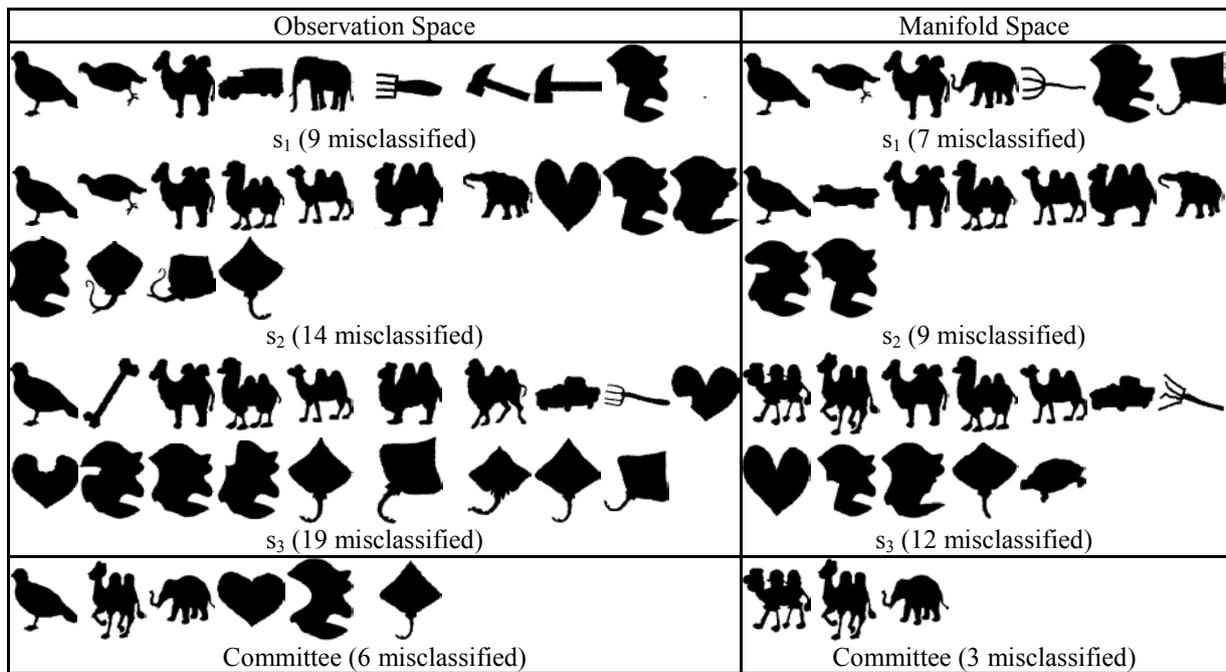


Figure 2. Illustration of misclassification improvement by manifold space.

VI. CONCLUSION AND DISCUSSION

This work has demonstrated novel supervised manifold learning for shape classification. In vision applications such as object recognition and image annotation, multiple array of sensors capture sample features in different feature spaces. It is important to map the samples in each space to enhance the compatibility of Euclidean distance of samples to semantic distance. The main innovation in the present research is the effective use of representing the shape samples in the multiple feature spaces to learn multiple shape manifold spaces. In each manifold space, Euclidean distances of shape feature vectors are more compatible with semantic distances. Final classification is done by committee of multiple classifier opinions.

Experiments on Kimia-216 dataset improve the results on manifold space than s_1 , s_2 and s_3 as 1.39%, 2.31% and 3.24%. The final classification result based on committee of classifiers improves the accuracy in manifold space compared with the observation space as 1.42%.

However, feature extraction is well done for available samples, but feature extraction of the out-of-samples is done only using GRNN. This also can be done by adding new samples to the existing graph by a semi-supervised approach. Also, in the proposed method, fusion of multiple feature spaces is performed in decision level. Feature spaces can be combined in feature level or in graph construction level. In the proposed method, non-linear feature extraction is based on a distance driven approach. In future, topological driven approaches [7][49][8] are used in feature extraction methodology.

The proposed method for shape classification can be used in other machine vision applications. In our research group, automatic image annotation (AIA) is the ongoing project based on the proposed method. AIA has an important role in information and

communication technologies like semantical web filtering, image mining and indexing of image databases. The main objective of this project is to reduce the semantic gap. So, the continuity between the instances of a semantic at the semantic space is also kept in feature space.

REFERENCES

- [1] Z. G. Hua, X. J. Wang, Q. S. Liu, and H. Q. Lu, "Semantic knowledge Extraction and Annotation for web images," In Proceeding of the 13th Annual ACM International Conference On Multimedia, 2005, pp. 467-470.
- [2] N. Alajlan, I. El Rube, M. S. Kamel, and G. Freeman, "Shape retrieval using triangle-area representation and dynamic space warping," *Pattern Recognition*, vol. 40(7), 2007, pp. 1911-1920.
- [3] M. Jordan, J. Kleinberg, and B. Scholkopf, "Nonlinear dimensionality reduction," Springer, pp. 11-14, 2007.
- [4] W. Johnson, and J. Lindenstrauss, "Extension of Lipschitz mapping into a hilbert space," *Contemporary Mathematics*, vol. 26, pp. 189-206, 1984.
- [5] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, 1964, pp. 1-27.
- [6] J. B. Tenenbaum, V. Silva, and J.C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, 2000, pp. 2319-2323.
- [7] S. T. Roweis, and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, 2000, pp.2323-2326.
- [8] M. Belkin, and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, 2003, pp. 1373-1396.
- [9] Z. Zhang, and H. Zha, "Principal manifolds and nonlinear dimensionality reduction via tangent space alignment," *Journal of Shanghai University (English Edition)*, vol. 8, 2004, pp. 406-424.
- [10] K. Q. Weinberger, and L. K. Saul, "Unsupervised learning of image manifolds by semidefinite programming," *International Journal of Computer Vision*, vol. 70, 2006, pp. 77-90.
- [11] J. W. Sammon Jr, "A nonlinear mapping for data structure analysis," *IEEE Transactions on computers*, vol. 100, 1969, pp. 401-409.
- [12] P. Demartines, and J. Herault, "Curvilinear Component Analysis: A Self-Organizing Neural Network for Nonlinear



- Mapping of Data Sets," IEEE Transactions on Neural Networks, vol. 8, 2002, pp. 148-154.
- [13] J. A. Lee, A. Lendasse, N. Donckers, and M. Verleysen, "A robust nonlinear projection method," In Proceeding of the European Symposium on Artificial Neural Networks, 2000, pp. 13-20.
- [14] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A survey of shape feature extraction techniques," Pattern Recognition, Peng-Yeng Yin (Ed.), 2008, pp. 43-90.
- [15] D. Zhang, and G. Lu, "Review of shape representation and description techniques," Pattern Recognition, vol. 37, 2004, pp. 1-19.
- [16] D. Zhang, and G. Lu, "Evaluation of MPEG-7 shape descriptors against other shape descriptors," Multimedia Systems, vol. 9, 2003, pp.15-30.
- [17] M. Bicego, V. Murino, and M. A.T. Figueiredo, "Similarity-based classification of sequences using hidden Markov models," Pattern Recognition, vol. 37, 2004, pp. 2281-2291.
- [18] M. Bicego, and V. Murino, "Investigating Hidden Markov Models' Capabilities in 2D Shape Classification," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 26(2), 2004, pp. 281-286.
- [19] N. Thakoor, and S. Jung, "Occlusion Resistant Shape Classifier based on Warped Optimal Path Matching," In Proceeding of the 18th International Conference on Pattern Recognition, 2006, pp. 60-63.
- [20] N. Thakoor, J. Gao, and S. Jung, "Hidden Markov Model-Based Weighted Likelihood Discriminant for 2-D Shape Classification," IEEE Transactions on Image Processing, vol. 16(11), 2007, pp. 2707-2719.
- [21] L. Gorelick, M. Galun, E. Sharon, R. Basri, and A. Brandt, "Shape Representation and Classification Using the Poisson Equation," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 28(12), 2006, pp. 1991-2005.
- [22] M. R. Daliri, and V. Torre, "Shape recognition based on Kernel-edit distance," Computer Vision and Image Understanding, vol. 114, 2010, pp. 1097-1103.
- [23] B. Wang, W. Shen, W. Y. Lin, X. G. You, and X. Bai, "Shape classification using tree-unions," In Proceeding of the IEEE 20th International Conference on Pattern Recognition (ICPR) Conference, 2010, pp. 983-986.
- [24] K. B. Sun, and B. J. Super, "Classification of contour shapes using class segment sets," In Proceeding of the IEEE Computer Vision and Pattern Recognition Conference, 2005, pp. 727-733.
- [25] X. Bai, X. W. Yang, D. G. Yu, and L. J. Latecki, "Skeleton-based shape classification using path similarity," International Journal of Pattern Recognition and Artificial Intelligence, vol. 22(4), 2008, pp. 733-746.
- [26] G. Vassilis, V. G. Kaburlasos, and A. Amanatiadis, "Binary Image 2D Shape Learning and Recognition Based on Lattice-Computing (LC) Techniques," Journal of Mathematical Imaging [online], June 2011.
- [27] X. He, W. Y. Ma, and H. J. Zhang, "Learning an image manifold for retrieval," In Proceeding of the ACM 12th annual international conference on Multimedia, 2004, pp. 17-23.
- [28] Y. Y. Lin, T. L. Liu, and H. T. Chen, "Semantic manifold learning for image retrieval," In Proceeding of the ACM 13th annual international conference on Multimedia, 2005, pp. 249-258.
- [29] X. He, D. Cai, and J. Han, "Learning a maximum margin subspace for image retrieval," IEEE Transactions on Knowledge and Data Engineering, vol. 20, 2007, pp. 189-201.
- [30] B. Xiao, E. Hancock, and H. Yu, "Manifold embedding for shape analysis," Neurocomputing, vol. 73, 2010, pp. 1606-1613.
- [31] M. A. Z. Chahooki, N. M. Charkari, "Learning the shape manifold to improve object recognition," Machine Vision and Applications [online], Jan 2012.
- [32] M. Cheng, and et al., "Incremental Embedding and Learning in the Local Discriminant Subspace With Application to Face Recognition," IEEE Transactions on Systems Man and Cybernetics, vol. 40 (5), 2010, pp. 580-891.
- [33] X. Gao, and C. Tian, "Multi-view face recognition based on tensor subspace analysis and view manifold modeling," Neurocomputing, vol. 72, 2009, pp. 3742-3750.
- [34] O. Arandjelovic, and R. Cipolla, "An information-theoretic approach to face recognition from face motion manifolds," Image and Vision Computing, vol. 24, 2005, pp. 639-647.
- [35] O. Arandjelovic, and R. Cipolla, "A pose-wise linear illumination manifold model for face recognition using video," Computer Vision and Image Understanding, vol. 113, 2009, pp. 113-125.
- [36] R. Xiao, and et al., "Facial expression recognition on multiple manifolds," Pattern Recognition, vol. 44, 2011, pp. 107-116.
- [37] Y. Cheon, and D. Kim, "Natural facial expression recognition using differential-AAM and manifold learning," Pattern Recognition, vol. 42, 2009, pp. 1340-1350.
- [38] Y. Chang, and et al., "Manifold based analysis of facial expression," Image and Vision Computing, vol. 24, 2006, pp. 605-614.
- [39] H. Qiao, and et al., "Learning an Intrinsic-Variable Preserving Manifold for Dynamic Visual Tracking," IEEE Transactions on Systems Man and Cybernetics, vol. 40 (3), 2010, pp. 868-880.
- [40] K. C. Lee, "Visual tracking and recognition using probabilistic appearance manifolds," Computer Vision and Image Understanding, vol. 99, 2005, pp. 303-331.
- [41] L. Wang, and D. Suter, "Visual learning and recognition of sequential data manifolds with applications to human movement analysis," Computer Vision and Image Understanding, vol. 110, 2008, pp. 153-172.
- [42] L. Wang, and D. Suter, "Learning and Matching of Dynamic Shape Manifolds for Human Action Recognition," IEEE Transactions on Image Processing, vol. 16 (6), 2007, pp. 1646-1661.
- [43] S. Yan, "Synchronized Submanifold Embedding for Person-Independent Pose Estimation and Beyond," IEEE Transactions on Image Processing, vol. 18 (1), 2009, pp. 202-210.
- [44] C. Shen, and et al., "Human pose estimation from corrupted silhouettes using a sub-manifold voting strategy in latent variable space," Pattern Recognition Letters, vol. 30, 2009, pp. 421-431.
- [45] M. W. Trosseta, and C. E. Priebe, "The out-of-sample problem for classical multidimensional scaling," Computational Statistics and Data Analysis, vol. 52, 2008, pp. 4635-4642.
- [46] S. Haykin, Neural networks: a comprehensive foundation, Prentice Hall, 2008.
- [47] T. B. Sebastian, P. N. Klein, and B. B. Kimia, "Recognition of shapes by editing their shock graphs," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 26 (5), 2004, pp. 550-571.
- [48] L. J. Latecki, and R. Lakamper, "Shape similarity measure based on correspondence of visual parts," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 22 (10), 2000, pp. 1185-1190.
- [49] D. L. Donoho, and C. Grimes, "Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data," In Proceeding of the National Academy of Sciences Conference, 2003, pp. 5591-5596.



Mohammad Ali Zare Chahooki received his B.Sc. degree in computer engineering from Shahid Beheshti University, Tehran, Iran in 2000 and his M.Sc. degree in software engineering from Tarbiat Modares University, Tehran, Iran in 2004. Currently he is a Ph.D. candidate of computer engineering in Tarbiat Modares University. His research interests include nonlinear feature extraction, image annotation, content based image retrieval, computer vision and AI.





Nasrollah Moghaddam Charkari received his B.Sc. degree in computer science from Shahid Beheshti University, Tehran, Iran in 1986 and his M.Sc. and Ph.D. degrees in computer engineering and information systems engineering from Yamanashi University, Japan in 1992 and 1995, respectively. Currently

he is an associate professor in faculty of electrical and computer engineering, Tarbiat Modares University, Tehran, Iran. His research interests include parallel processing, AI and robotics, computer vision, and image processing.

