Annals of the BMVA Vol. 2008, No. 5, pp 1–17 (2008)

Unsupervised Shape Clustering using Diffusion Maps

Nasir Rajpoot [a,⋆] and Muhammad Arif [b]

[a] Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK
[b] Department of Electrical Engineering, Pakistan Institute of Engineering & Applied Sciences (PIEAS), PO Nilore, Pakistan

⋆ Corresponding author's email address: ⟨nasir@dcs.warwick.ac.uk⟩

Abstract

The quotient space of all smooth and connected curves represented by a fixed number of boundary points is a finite-dimensional Riemannian manifold, also known as a *shape manifold*. This makes the preservation of locality a critically important issue when reducing the dimensionality of shapes on the manifold. We present a completely unsupervised clustering algorithm employing diffusion maps for locality-preserving embedding of shapes onto a much lower-dimensional space. The algorithm first obtains a non-linear low-dimensional embedding of shape context features of outer boundary contours of the shapes. Considering the embedded coordinates as a new minimalist representation of shapes, a clustering of shapes is obtained using a finite mixture model. The proposed clustering algorithm is computationally efficient, as it relies on clustering in a very low-dimensional space, and produces much improved results (88.6% for a 7-class dataset) as compared to clustering with conventional linear projections.

1 Introduction

With an ever increasing amount of image data at one's disposal, a *plausible* clustering of the data has become a well sought after goal in recent years. Such a clustering can be applied to data visualisation, data organisation, exploratory data analysis, only to name a few applications. Shape can be used as a vital clue for these applications when the data happens to contain images of objects. The 2D shape, also referred to as simply shape in the remainder of this paper, of a 3D object is a function of the object's 3D geometry and its reflectance function or texture (*internal* parameters) as well as illumination and camera location (*external* parameters). Shape analysis is one of the key problem areas in computer vision with a range of applications in areas such as object recognition, image retrieval, and object-based video coding. Most of the existing literature on shape analysis can be broadly categorised into two classes in terms of representation: (1) *landmark* or keypoint based methods, and (2)

© 2008. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

outer *contour* based methods. The former kind of techniques face the limitations that they require the keypoints to be marked, manually or automatically, and a correspondence between keypoints of different shapes should be found for analysis purposes. The latter type of methods overcome this limitation by first representing the object's 2D shape image as a silhouette. A *contour of the outer boundary* is obtained, often by edge detection, and is then regarded as shape of the object. A finite sampling, consisting of say n points $\mathbf{p}_i = \{x_i, y_i\}$ where (x_i, y_i) are coordinates of the ith boundary point, of the outer boundary can, therefore, be used to represent an instance (in $\Re^{2n}$) of 2D shape of the corresponding 3D object. A suitable representation of shapes should be invariant to affine transformations of shape Kendall [1984], Bookstein [1986], in turn requiring n to be large and potentially invoking the *curse of dimensionality*. To circumvent this problem, classical shape analysis methods such as Cootes et al. [1995], Leventon et al. [2000] employ principal component analysis (PCA) to reduce the dimensionality of the problem. One of the basic assumptions that these methods make is that the subspace corresponding to the major modes of variation for a particular class of shapes is linearised. The reality, however, is that such subspaces may be far from linear. In fact, the space of all 'smooth' connected closed curves is infinite-dimensional and is not a vector space, but locally linear (i.e., a *manifold*) Michor and Mumford [2003]. And the resulting quotient space is a finite-dimensional Riemannian manifold Michor and Mumford [2003], Klassen et al. [2004], Brun [2007], often termed as a *shape manifold*. For these manifolds, a locality preserving non-linear dimensionality reduction (NLDR) method (see van der Maaten et al. [2008] for a review of recently developed NLDR methods) is required. In this context, we define the problem of *shape clustering* as follows:

> **Problem Definition**: Given a set $\Omega = \{\Omega_1, \Omega_2, \ldots, \Omega_c\}$ containing shapes from c different classes, where $\Omega_i = \{C_j^i\}, j = 1, 2, \ldots, N_i = |\Omega_i|$, $\sum_i N_i = N$, the total number of shapes, and $C_j^i \in \mathcal{C}^1$ is a connected closed curve, represented by n boundary points of the form $\{\mathbf{p} = (x, y)\}$ in the Euclidean plane, find a non-linear mapping
>
> $$\Psi : \Re^{2n} \mapsto \Re^m,$$
>
> where $m << n$, such that clustering in $\Re^m$ yields a plausible grouping of similar shapes.

In a recent paper, Yankov & Keogh Yankov and Keogh [2006] have proposed an algorithm for clustering of shape manifolds using a variant of isomaps, termed by the authors as b-isomaps, in a semi-supervised manner. They showed that isomap projection results in a much better separation of clusters and estimated the intrinsic dimensionality of the shape manifolds using the 'elbow' of the curve of residual variance when plotted against the number of coordinates of the projection, resulting in 3 coordinates for a four-class shape dataset of diatoms. However, a limitation of the work in Yankov and Keogh [2006] is that it is not entirely unsupervised, in that the cluster quality is checked by comparing the resulting labels to ground truth labels and only the best clustering results are reported. Moreover, the authors found that the clustering accuracy using top 3 coordinates went down when a dataset with relatively large (four) number of classes was used, from 98% for two-class diatoms dataset to 80% for four classes from the Marine creatures dataset. In another very relevant paper Etyngier et al. [2007], Etyngier *et al.* proposed a deformable model framework integrating non-linear shape priors using diffusion maps. The work in Etyngier et al. [2007] did not,

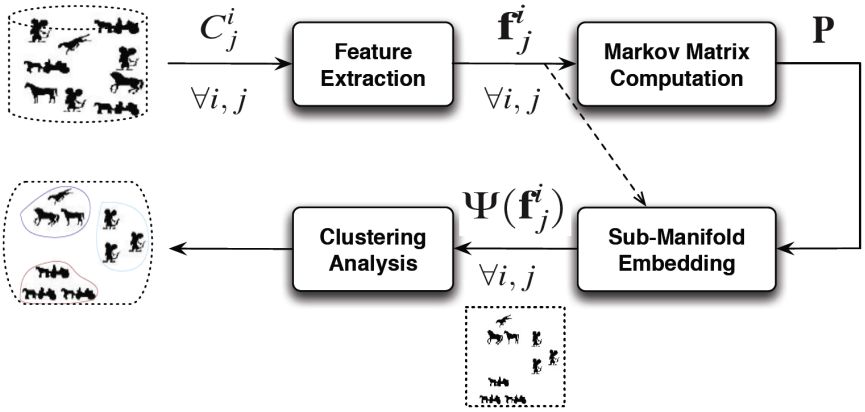


Figure 1: Block diagram of the proposed clustering algorithm.

however, address the clustering problem.

In our earlier work Rajpoot et al. [2007], we have shown a diffusion maps based framework for learning the shape manifolds to be significantly more effective than linear subspace projection. In this paper, we extend the framework proposed in Rajpoot et al. [2007] to unsupervised clustering of shape manifolds. We conjecture[1] that an isometry exists between Σ_2^n, the space of shapes with n points in 2D, and a shape manifold whose exact geometry is unknown to us but whose different sub-manifolds correspond to different shapes. We propose a shape clustering algorithm which employs diffusion maps for non-linear embedding of the feature vectors extracted from a given set of unknown and unseen 2D shapes into a very small dimensional space. It is shown in Section 3 that the low-dimensional embedding preserves the locality of shapes on a high-dimensional manifold. Clustering of low-dimensional embedding is obtained using a finite Gaussian mixture model, since the resulting clusters can be of elongated shapes as noted in Yankov and Keogh [2006], and Bayesian information criterion (BIC) Fraley and Raftery [1998] is employed to determine the number of modes in the mixture model. Results of clustering thus obtained can be used for exploratory analysis of a database of shapes and for visualisation purposes.

The remainder of this paper is organised as follows. A description of the proposed clustering algorithm is given in the next section, followed by a presentation of the experimental results and their discussion in Section 3. The paper concludes with a summary and some future directions.

2 The Clustering Algorithm

A block diagram of the proposed clustering algorithm is given in Figure 1. Input to the algorithm is a dataset consisting of shape images belonging to c different classes. We first extract feature vectors for each of the shape images using shape context, as described in Section 2.1. The feature vectors are used to form a Markov matrix, described below in Section 2.3, with

[1]This conjecture may be regarded as an extension of Kendall's theorem in Kendall [1984] which states that an isometry exists between Σ_2^3, the space of labelled triangles in 2D represented by 3 vertices in $\Re^2$, and $S^2(1/2)$, the sphere of radius $1/2$ in $\Re^3$.

the help of a shape similarity measure (Section 2.2). Low-dimensional embedding of all the shape images is computed using eigenvectors of the Markov matrix. The embedded coordinates are then considered as a minimalist representation of the shapes and a clustering is obtained using finite mixture models, as described in Section 2.5. Clustering by finite mixture models takes place by using the Expectation-Maximisation (EM) framework and can be achieved in both semi-supervised and unsupervised manners. In a semi-supervised setting, the algorithm requires the user to input the number of clusters (i.e., modes in the mixture model) to be found in the dataset. In an unsupervised setting, the algorithm determines the number of clusters by finding the number of modes that maximise the BIC Fraley and Raftery [1998].

2.1 Feature Extraction

As mentioned above, a desirable representation of shape should factor out any translation, scale, and rotation transformations. In order to learn the structure of shape manifolds of different classes in an unsupervised manner, it is crucial that such a representation employs features related to a shape that are associated with main shape characteristics. Several features for shape representation, also known as shape descriptors, have been proposed in the literature (see Zhang and Lu [2004] for a survey on the topic). For a given shape, we first extract its corresponding outer boundary $\mathcal{C}$ which consists of n boundary points $p_i = (x_i, y_i)$, for $i = 1, 2, \ldots, n$. The number of boundary points n is fixed by employing a cubic spline interpolation for re-sampling the boundary points. In our earlier work Rajpoot et al. [2007], we used Fourier descriptors (FDs) to represent $\mathcal{C}$. While being simple to compute, just requiring a fast Fourier transform in $O(n)$ time, the FDs are inherently rotation invariant and can be easily made scale invariant by dividing the magnitude of high frequency Fourier coefficients by the DC component.

In this paper, we employ shape context (SC), recently proposed by Belongie *et al.* Belongie et al. [2002] as a powerful descriptor for shapes. The main idea of SC is to describe a distribution of points in the log-polar neighbourhood of a given point p_i in the form of a coarse histogram h_i computed as follows Belongie et al. [2002],

$$h_i(k) = \#\{q \neq p_i : (q - p_i) \in \text{bin}(k)\} \tag{1}$$

$\forall k = 1, 2, \ldots, K$, where K denotes the total number of bins, which are uniform in the log-polar space. This idea is also illustrated in Figure 2, taken from Belongie et al. [2001]. The histogram h_i as computed by equation (1) defines the shape context for p_i. This rich description of contextual information about the neighbourhood of a point on the shape boundary lends itself to an efficient solution for the point correspondence problem and proves to be useful for our purpose of shape clustering as well.

A vector containing the histogram count values $h_i(k)$, $\forall i = 1, 2, \ldots, n$ and $\forall k$ for a shape $\mathcal{C}$ can then be regarded as a feature vector $\mathbf{f}$ for that shape. In all our experiments, $n = 300$ and the number of bins for histograms at each of these n points is 60, 5 bins for $\log r$ and 12 for orientation, yielding $18,000$-dimensional feature vector for every shape. It is worth noting here that since shape context rearranged as a feature vector is not inherently affine invariant, it poses an interesting challenge for our clustering algorithm. It will be shown that the nonlinear embedding of shapes by diffusion maps indeed preserves the locality of shapes on their manifold.

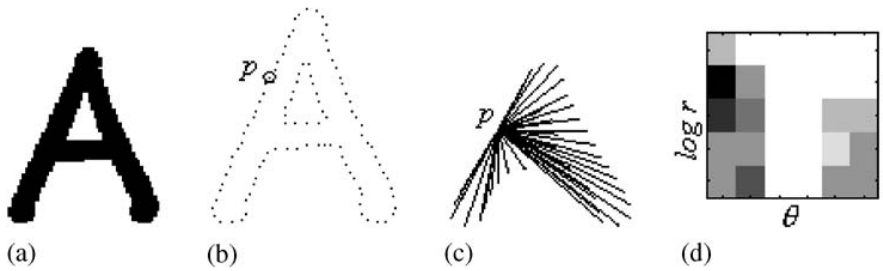


Figure 2: Shape context: (*a*) a character shape; (*b*) edge image of (*a*); (*c*) a point p on shape (*a*) and all the vectors originating from p to all other points on the edge image; (*d*) log-polar histogram of the vectors in (*c*), termed as the shape context of p (reprinted from Belongie et al. [2001]).

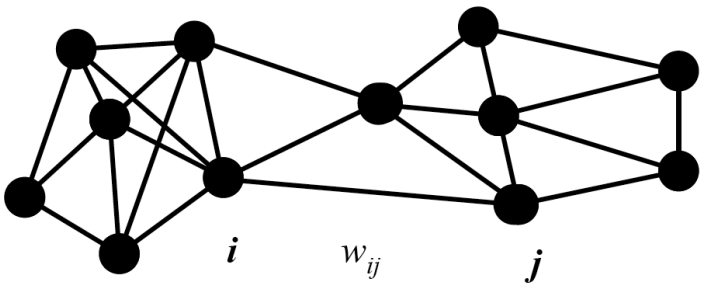


Figure 3: Adjacency graph depicting relationships between different shapes in the database; Nodes of the graph represent shapes and edges represent similarity between nodes with their lengths inversely proportional to w_{ij}.

2.2 The Shape Similarity Measure

In order to ensure that the low-dimensional embedding will preserve the locality of shapes of different internal and external parameters on the shape manifold, a suitable similarity measure is required. We define the similarity between two feature vectors[2], $\mathbf{f}_i$ and $\mathbf{f}_j$, $i, j = 1, 2, \ldots, N$, using a Gaussian kernel of width ϵ as follows,

$$w_{ij} = w(\mathbf{f}_i, \mathbf{f}_j) = \exp\left(-\frac{||\mathbf{f}_i - \mathbf{f}_j||^2}{2\epsilon}\right). \tag{2}$$

The above pairwise similarity measure is also used as the main property of an edge between nodes i and j in an adjacency graph, where node i of the graph represents feature vector for the ith shape image in the dataset, as shown in Figure 3.

[2]Since our goal in this work is *unsupervised* clustering of shapes, without loss of generality we will drop the superscript from the shape contour C_i and its corresponding feature vector $\mathbf{f}_i$, as in Figure 1, for the remainder of this paper.

2.3 Markov Matrix

A Markov random walk is defined on the graph by defining the state-transition matrix $\mathbf{P} = [p_{ij}]$, also known as the Markov matrix, with its (i,j)th element p_{ij} given by,

$$p_{ij} = \frac{w(\mathbf{f}_i, \mathbf{f}_j)}{d(\mathbf{f}_i)}, \tag{3}$$

where $d(\mathbf{f}_i)$ denotes the degree of node i in the graph and is calculated as follows,

$$d(\mathbf{f}_i) = \sum_{\mathbf{z} \in \Omega} w(\mathbf{f}_i, \mathbf{z}). \tag{4}$$

It is worth noting that unlike most similarity matrices used in other manifold learning methods, such as Isomaps Tenenbaum et al. [2000] and Laplacian eigenmaps Belkin and Niyogi [2003], this matrix is generally not symmetric.

2.4 Embedding in a Low-Dimensional Space

Let $\{\lambda_i\}$ be the sequence of eigenvalues of $\mathbf{P}$ such that $|\lambda_0| \geq |\lambda_1| \geq \cdots$, and $\{\psi_i\}$ be the corresponding eigenvectors. A mapping from the shape feature set Ω to a lower-dimensional Euclidean space $\Re^m$, where $m << n$ is the dimensionality of the lower-dimensional subspace, is given by (see Lafon and Lee [2006], Coifman and Lafon [2006] for details),

$$\Psi : \mathbf{f} \longmapsto (\lambda_1 \psi_1(\mathbf{f}), \lambda_2 \psi_2(\mathbf{f}), \ldots, \lambda_m \psi_m(\mathbf{f}))\,. \tag{5}$$

The above mapping gives an initial low-dimensional embedding at $t = 1$. Spectral fall-off is the main factor contributing to dimensionality reduction. The mapping Ψ evolves as time t increases. The above mapping at time $t > 1$ can be computed by scaling the diffusion coordinates $\psi_i(\mathbf{f})$, $i = 1, \ldots, m$, with λ_i^t as follows Coifman and Lafon [2006],

$$\Psi^{(t)} : \mathbf{f} \longmapsto \left(\lambda_1^t \psi_1(\mathbf{f}), \lambda_2^t \psi_2(\mathbf{f}), \ldots, \lambda_m^t \psi_m(\mathbf{f})\right)\,. \tag{6}$$

For large value of t, large-scale structures in the data can be captured with fewer diffusion coordinates Coifman and Lafon [2006].

2.5 Clustering in the Embedding Space

Spectral graph embedding of shapes residing on a very high-dimensional manifold to a much lower-dimensional space using diffusion maps results in a set of elongated clusters, representing shapes of similar kinds, provided an appropriate value is chosen for the Gaussian kernel parameter ϵ. We found that a mixture model representation of these clusters is more effective than the standard k-means (kM) or fuzzy C-means (FCM) clustering algorithms. Let $\mathbf{F}_i$, $i = 1, 2, \ldots, N$ denote embedding of the shape context feature vectors for all shapes in a given dataset. Using the GMM, the probability mass function of an arbitrary feature vector's embedding $\mathbf{F}$ is given by,

$$P(\mathbf{F}) = \sum_{j=1}^{c} a_j \, g_j(\mathbf{F}, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j), \tag{7}$$

where a_j denotes weight of the jth Gaussian $g_j(\mathbf{F})$, whose mean vector $\boldsymbol{\mu}_j$ and covariance matrix $\boldsymbol{\Sigma}_k$ determine its position and shape. The parameters a_j, $\boldsymbol{\mu}_j$, and $\boldsymbol{\Sigma}_j$ are estimated with the help of EM. Using priors a_j and likelihood $p(\mathbf{F}|a_j, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ of an embedding $\mathbf{F}$ for the jth Gaussian, its posterior probability $p(j|\mathbf{F})$ can be computed. A label k is assigned to the shape image corresponding to $\mathbf{F}$, where k maximises the posterior,

$$k = \underset{j=1,2,\ldots,c}{\operatorname{argmax}} \; p(j|\mathbf{F}). \tag{8}$$

In order to determine the number of clusters, c, we utilise the Bayesian information criterion (BIC), as suggested in Fraley and Raftery [1998]. Choosing a value of c which maximises the BIC value for a range of c ensures that the sum of squared errors between actual embeddings and the embeddings estimated by GMM is minimised while penalising the model complexity.

3 Experimental Results

Our experimental dataset consisted of a subset of the MPEG-7 core experiment (CE) Shape-1 Part-B dataset and comprised of 140 images of objects from seven classes (20 images per class): `apple`, `car`, `carriage`, `device7`, `flatfish`, `guitar`, `teddy`. Although the number of classes chosen for our experiments is not very high, it is in agreement with other shape clustering studies found in the literature, such as Yankov and Keogh [2006]. As mentioned earlier in Section 2.1, 18,000-dimensional shape context feature vectors were extracted per shape image in the dataset. These feature vectors were embedded in $\Re^{10}$, a Euclidean space with only 10 coordinates, using both conventional linear projection with principal component analysis (SC+PCA) and diffusion maps (SC+DM). Figure 4 shows top three coordinates of the embeddings thus obtained. It can be seen from Figure 4(*a*) that while two of the seven classes, namely `carriage` and `guitar`, are clearly separated using SC+PCA, the remaining classes are mixed up in the embedding. This is in sharp contrast with Figure 4(*b*), which shows that most of the classes have not only well separated but also relatively compact clusters. In order to illustrate this point clearly, we also show in Figure 5 shape images overlaid onto their embedding using only top two diffusion coordinates. An embedding of this kind can be used for exploratory data analysis, when no prior knowledge about the different kinds of shapes is available, and visualisation of relationships between shapes belonging to different classes. In this Figure, it can be seen that SC+DM effectively separates at least four of the seven classes using just two diffusion coordinates. The remaining three classes (`car`, `flatfish`, and `guitar`) that appear to be mixed up towards the bottom-right corner of Figure 5(*a*) can also be separated to some extent, provided an efficient clustering method is employed. It can also be noted from both Figures 4 and 5 that the resulting clusters in the embedded coordinate space are not necessarily spherical in nature. Therefore, we employ a general elliptical Gaussian mixture model (GMM) in an unsupervised EM framework to assign class labels to all shapes in our dataset.

Final clustering results obtained by using a fixed number of clusters ($c = 7$), or number of modes of the GMM, for both SC+PCA and SC+DM with five embedding coordinates are shown in Figures 6 and 7. These qualitative results show that SC+DM is more effective than SC+PCA in terms of grouping together shapes belonging to the same class. While SC+PCA divides `guitar` shapes into two clusters (1 and 3) and lumps together `car` and

`flatfish` shapes (cluster 4), SC+DM effectively groups most of the shape images with the exception of `guitar` (cluster 2). For an objective comparison, we use two clustering performance measures: adjusted Rand Index (aRI) Martinez and Martinez [2004], which is based on counting pairs of points found in same clusters in two clustering results and those found in different clusters in the two results, and variance of information (VI) Meilă [2007], which is an information theoretic metric for comparing two clustering results based on entropies and mutual information of the two associated random variables. A higher value of aRI means better grouping, whereas a lower value of VI implies better clustering. For both the performance measures, clustering results obtained by either of SC+PCA and SC+DM are compared to the ground truth labels. Performance curves with each of the measures plotted against the number of embedding coordinates m with a fixed number of clusters ($c = 7$) for both SC+PCA and SC+DM are shown in Figure 8. From this Figure, it can be seen that SC+DM outperforms SC+PCA in terms of both the measures, achieving steady values at $m = 5$ and suggesting that no further gains are made by adding more projected coordinates. The overall accuracy of SC+DM is 88.6% as compared to 72.9% for SC+PCA. When we utilise the BIC to determine the number of clusters for GMM, better clustering performance with a higher value of aRI and a lower value of VI than the steady-state values found in Figure 8 can be achieved, as shown in Figure 9. With $m = 6$, we get the minimum value for aRI and the maximum value for VI, suggesting perhaps that this may be the intrinsic dimensionality of the dataset. These results demonstrate the effectiveness of the proposed shape clustering method.

4 Conclusions

In this paper, we have extended the shape manifold learning framework presented in Rajpoot et al. [2007] to unsupervised clustering of shapes. The proposed clustering algorithm employs diffusion maps, a nonlinear dimensionality reduction method, to embed the feature vectors of shapes in a given dataset onto lower-dimensional spaces. An unsupervised clustering of the embedding employs the Bayesian information criterion (BIC) to determine the number of clusters and is shown to effectively reveal the prevalent groups of shapes in the dataset. The proposed method is also shown to be significantly more sensitive than PCA for grouping together shapes that may be considered relatively difficult. The method is yet to be tested for a relatively large dataset with large number of classes. Since solving the eigenvalue problem for a large matrix can be highly demanding in terms of computational power and storage requirements, Nyström's extension can be employed for embedding the shape context feature vectors Lafon et al. [2006]. To the best of our knowledge, selection of kernel width ϵ for diffusion maps is an open question. Finding a suitable answer to this question and investigating robustness of the proposed algorithm to noise are some of the future directions for this work.

References

M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

S. Belongie, J. Malik, and J. Puzicha. Shape matching and object recognition using shape

contexts. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(4):509–522, 2002. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp%3Farnumber%3D993558`.

Serge Belongie, J. Malik, and J. Puzicha. Matching shapes. In *Proceedings International Conference on Computer Vision*, volume 1, pages 454–461, July 2001.

F.L. Bookstein. Size and shape spaces for landmark data in two dimensions. *Statistical Science*, 1(2):181–242, 1986. URL `http://www.jstor.org/stable/2245441`.

Anders Brun. *Manifolds in Image Science and Visualization*. PhD thesis, Linköping University, Department of Biomedical Engineering, 2007.

R. Coifman and S. Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis, Special Issue on Diffusion Maps and Wavelets*, 21:5–30, July 2006.

T. Cootes, D. Cooper, C. Taylor, and J. Graham. Active shape models – their training and applications. *Computer Vision and Image Understanding*, 61(1):38–59, 1995.

P. Etyngier, F. Segonne, and R. Keriven. Shape priors using manifold learning techniques. In *International Conference on Computer Vision (ICCV)*, pages 1–8, 2007.

C. Fraley and AE Raftery. How many clusters? Which clustering method? Answers via model-based cluster analysis. *The Computer Journal*, 41(8):578–588, 1998.

D.G. Kendall. Shape manifolds, procrustean metrics,and complex projective spaces. *Bulletin of the London Mathematical Society*, 16(2):81–121, 1984. URL `http://intl-blms.oxfordjournals.org/cgi/content/abstract/16/2/81`.

E. Klassen, A. Srivastava, W. Mio, and S.H. Joshi. Analysis of planar shapes using geodesic paths on shape spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3): 372–383, March 2004. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1262333`.

S. Lafon, Y. Keller, and RR Coifman. Data fusion and multicue data matching by diffusion maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1784–1797, 2006.

S. Lafon and A.B. Lee. Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parmeterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9):1393–1403, September 2006.

M. Leventon, E. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proceedings International Conference onComputer Vision and Pattern Recognition (CVPR)*, pages 316–323, 2000.

W.L. Martinez and A.R. Martinez. *Exploratory data analysis with MATLAB*. Chapman & Hall/CRC, 2004.

M. Meilă. Comparing clusterings£an information based distance. *Journal of Multivariate Analysis*, 98(5):873–895, 2007.

P.W. Michor and D. Mumford. Riemannian geometries on spaces of plane curves. *Arxiv preprint math.DG/0312384*, 2003. URL `http://arxiv.org/abs/math.DG/0312384`.

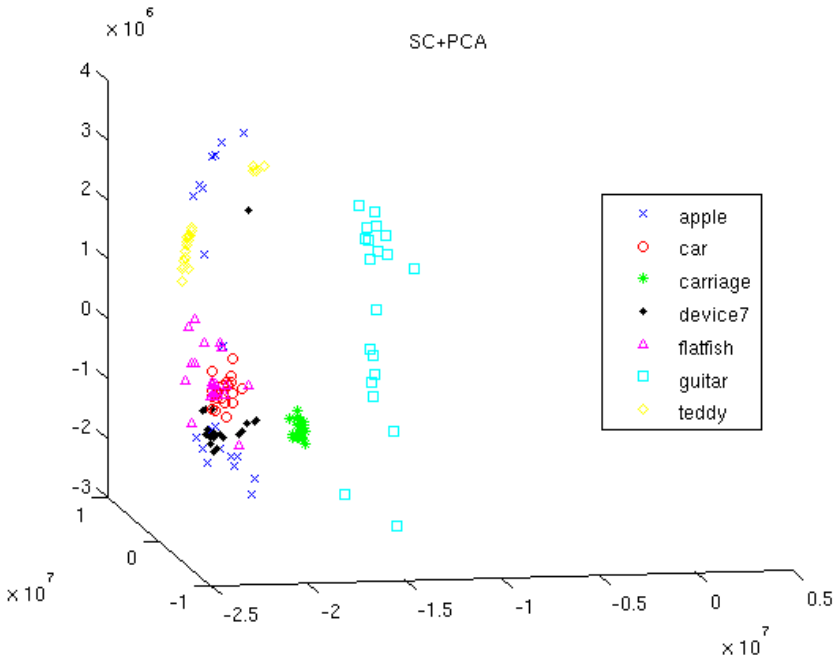
N. Rajpoot, M. Arif, and A. Bhalerao. Unsupervised learning of shape manifolds. In *Proceedings British Machine Vision Conference (BMVC'2007)*, pages 312–321, September 2007. URL `http://www.dcs.warwick.ac.uk/~nasir/papers/bmvc07a.pdf`.

J.B. Tenenbaum, V. Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000. URL `http://www.sciencemag.org/cgi/content/abstract/290/5500/2319`.

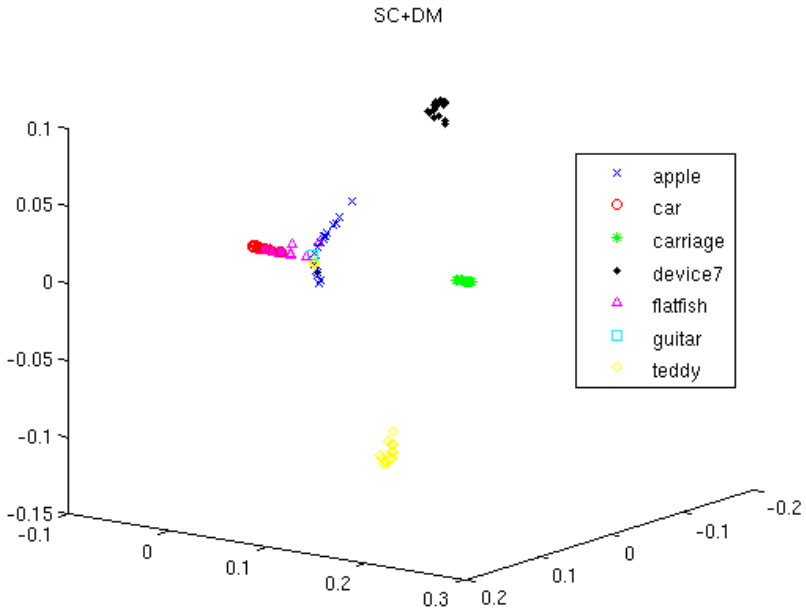
L.J.P. van der Maaten, E.O. Postma, and H.J. van den Herik. Dimensionality reduction: A comparative review. Submitted to Neurorecognition, 2008.

D. Yankov and E. Keogh. Manifold clustering of shapes. In *Proceedings International Conference on Data Mining (ICDM)*, pages 1167–1171, 2006. URL `http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4053173`.

D. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern recognition*, 37(1):1–19, 2004.

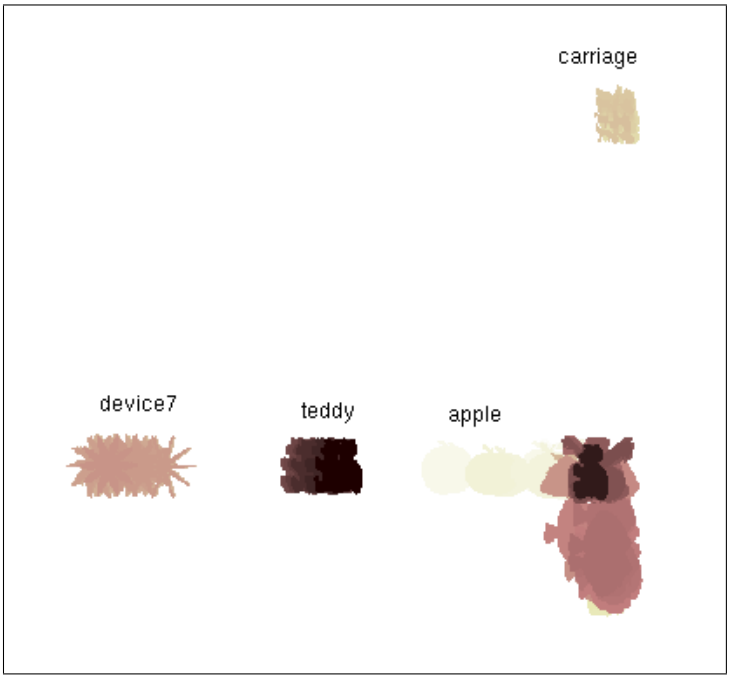


(*a*)

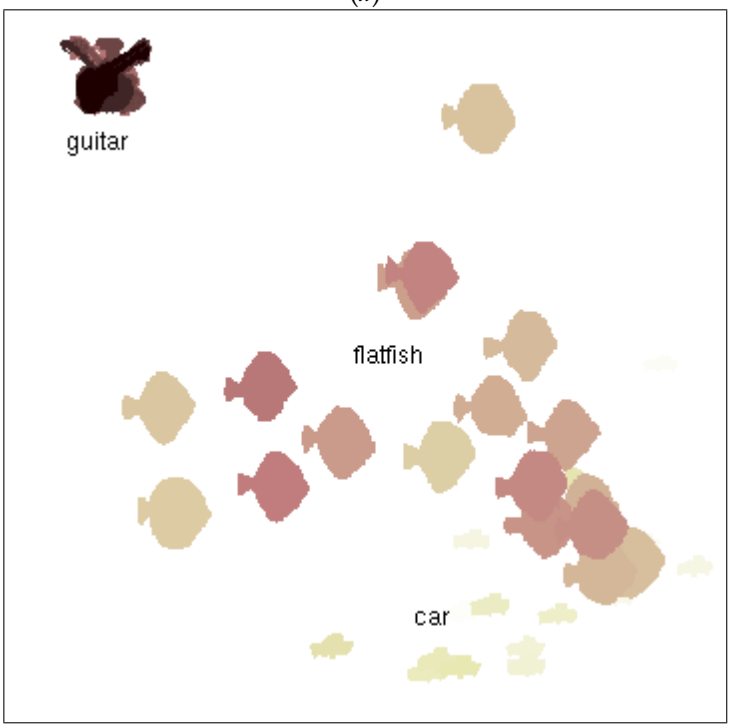


(*b*)

Figure 4: Embeddings of the SC feature vectors using top 3 coordinates of (*a*) PCA and (*b* Diffusion Maps.



(*a*)



(*b*)

Figure 5: Images of shapes overlaid on their respective positions in the embedded space using top two diffusion coordinates (SC+DM with $m = 2$), showing (*a*) all shapes with 4 relatively well-separated classes (`apple`, `carriage`, `device7`, and `teddy`) and (*b*) zoomed segment of the space containing remaining three classes (`guitar`, `flatfish`, and `car`) apparently mixed up in the embedding as seen towards the bottom right of (*a*).

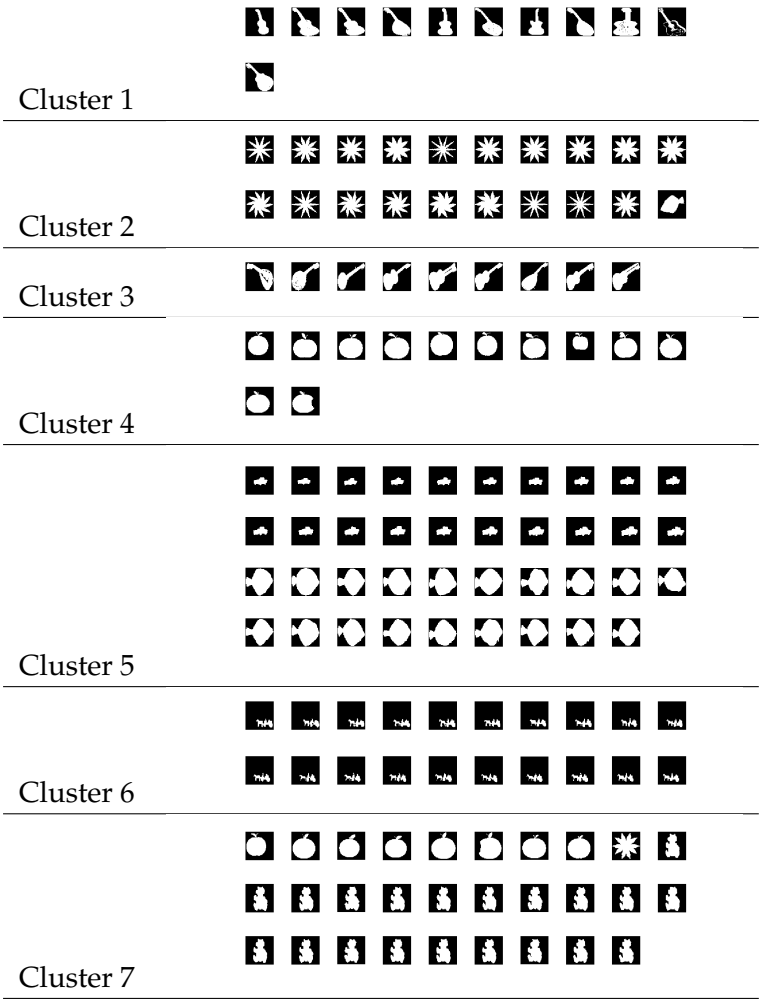


Figure 6: Clustering results for SC+PCA ($m = 5, c = 7$).

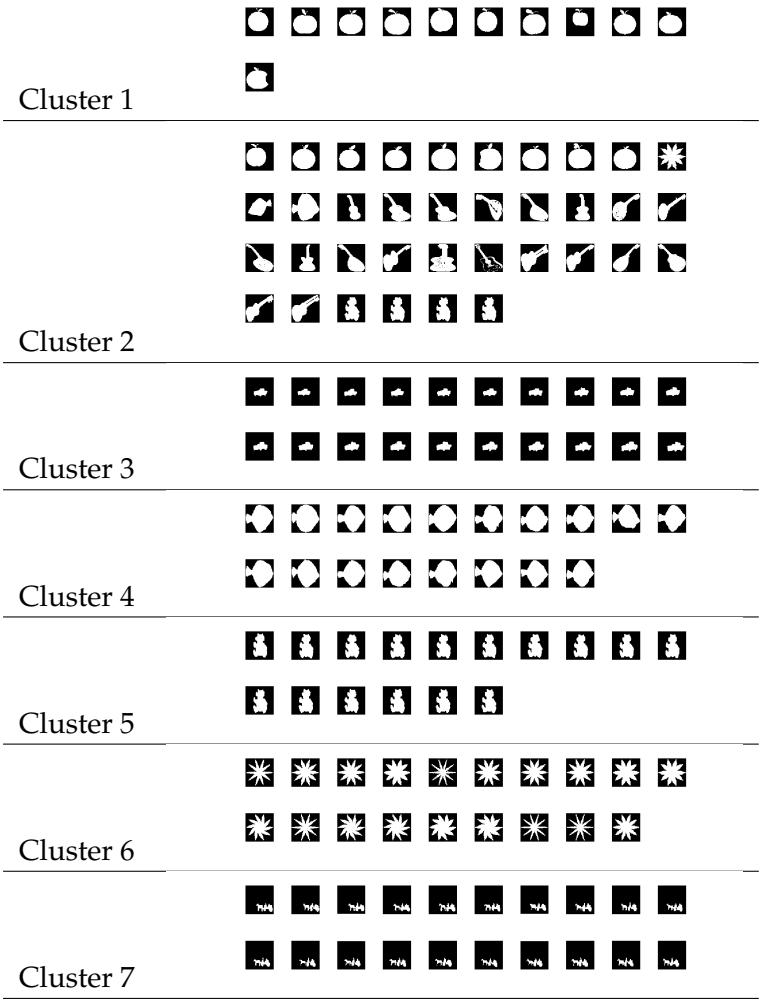
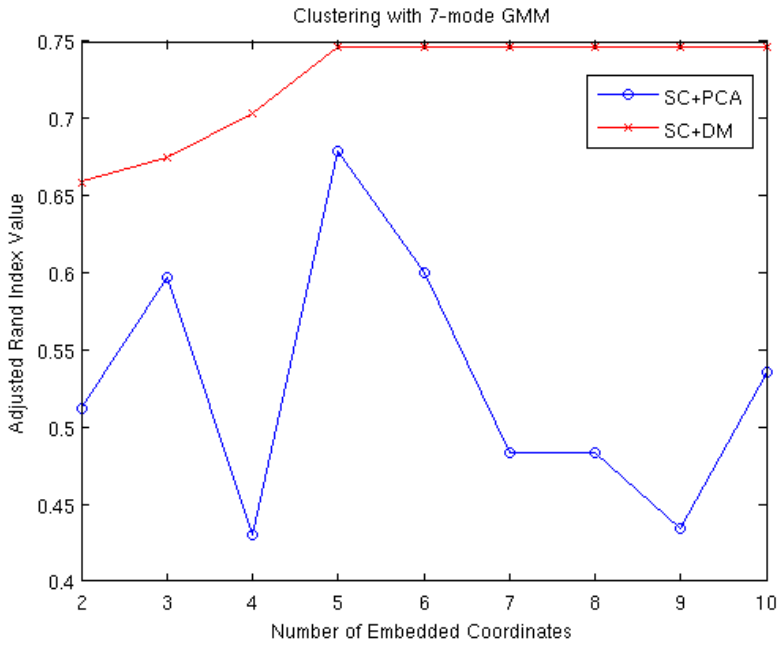
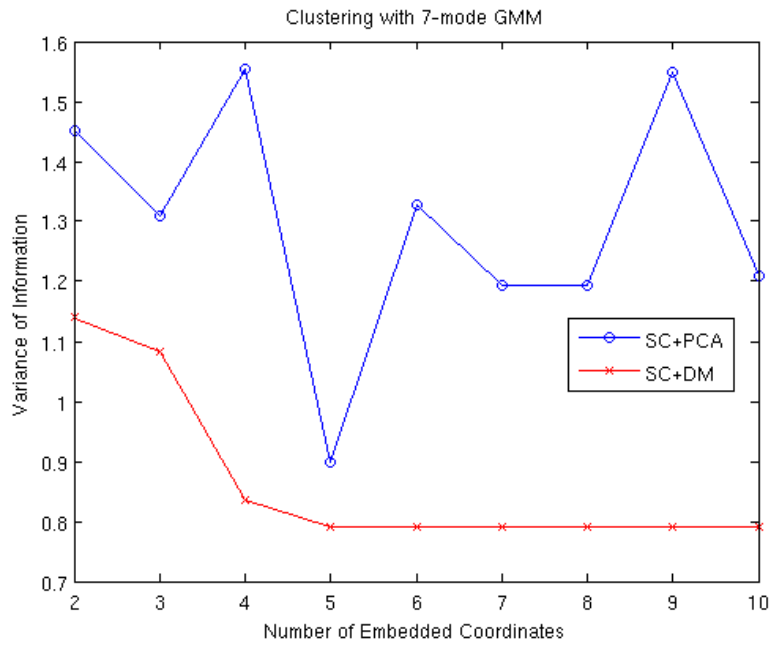


Figure 7: Clustering results for SC+DM ($m = 5, c = 7$).

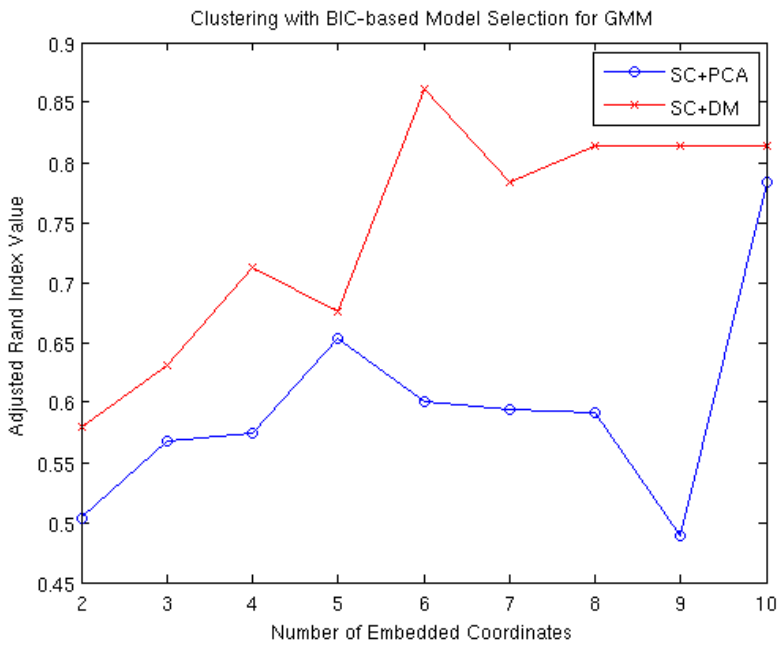


(*a*)

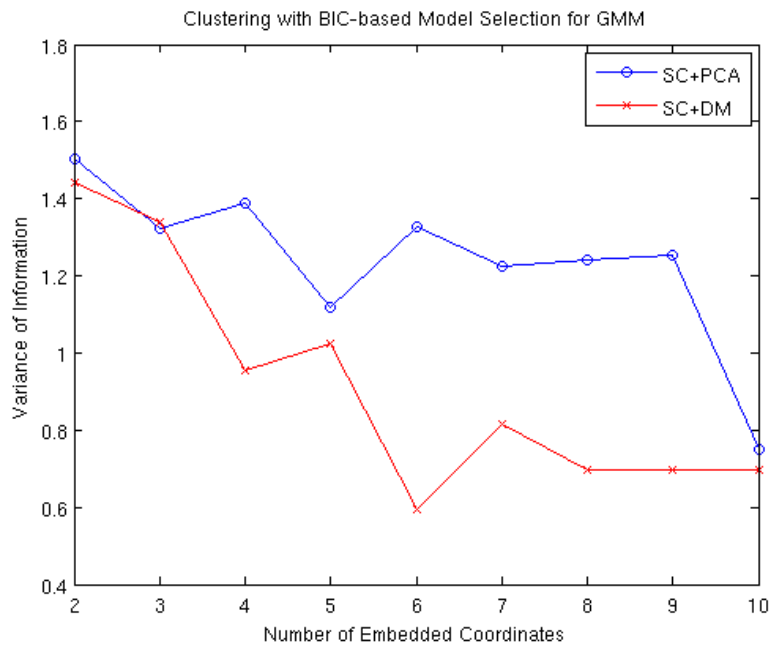


(*b*)

Figure 8: Plots of performance measures against number of embedded coordinates (m) for clustering results using SC+PCA and SC+DM for a fixed value for number of clusters ($c = 7$): (a) adjusted Rand Index (aRI) and (b) Variance of Information (VI).



(*a*)



(*b*)

Figure 9: Plots of performance measures against number of embedded coordinates (m) for clustering results using SC+PCA and SC+DM for BIC-based model selection: (*a*) adjusted Rand Index (aRI) and (*b*) Variance of Information (VI).