Estimating the Mean Manifold of a Deformable Object from Noisy Observations

Ziv Yavo*, Joseph M. Francos*, Ignacio Santamaria†, Louis L. Scharf‡
*Dept. of Electrical and Computer Engineering, Ben-Gurion University, Israel
†Dept. of Communications Engineering, University of Cantabria, Spain
‡Dept. of Mathematics, Colorado State University, Fort Collins, CO, USA

Abstract—Assume we have a set of noisy observations (for example, images) of different objects, each undergoing a different geometric deformation, yet all the deformations belong to the same family. As a result of the action of these deformations, the set of different observations on each object is generally a manifold in the ambient space of observations. It has been shown, [1], that in the absence of noise, in those cases where the set of deformations admits a finite-dimensional representation, the universal manifold embedding (UME) provides a mapping from the space of observations to a low dimensional linear space. The manifold corresponding to each object is mapped to a *distinct linear subspace* of Euclidean space, and the dimension of the subspace is the same as that of the manifold. In the presence of noise, different observations are mapped to different subspaces. In this paper we derive a method for "averaging" the different subspaces, obtained from different observations made on the same object, in order to estimate the mean representation of the object manifold. The mean manifold representation is then employed to minimize the effects of noise in matched manifold detectors and to improve the separability of data sets in the context of object detection and classification.

I. INTRODUCTION

Solutions to many problems in image and signal analysis have to cope with the effects of the multiplicity of appearances of objects. For example, in the problem of object recognition the "same" object may have a huge family of different appearances, and the first problem one needs to confront, is the understanding of the set of all possible appearances of that single object. One of the main reasons for the variability in the appearance of an object is a change in its underlying geometry.

In general, we are given a set of observations (for example, images) of different objects, each undergoing a different geometric deformation. As a result of the action of the deformations, the set of different realizations for each object is generally a manifold in the space of observations. The problem of finding and analyzing non-linear low-dimensional structures in high-dimensional data has been attracting considerable interest in recent years, see, *e.g.*, [7] for a recent collection of papers. The common underlying idea unifying the existing approaches is that although the data is sampled and presented in a high-dimensional space, for example because of the high resolution of the camera sensing the scene, in fact the intrinsic complexity and dimensionality of the observed physical phenomenon are very low. One of the dominant approaches among existing dimensionality reduction methods is to expand the principles of the linear spectral methods to more complex low-dimensional structures than a single subspace, by assuming the existence of a smooth and invertible isometric mapping from the original manifold to some other manifold which lies in a lower dimensional space, [2]-[4]. These dimensionality reduction methods make very modest assumptions on the reasons for the variability in the appearances of the object. The common general assumption is that the degrees of freedom act continuously on the objects and therefore the set of appearances of a single object is some continuous entity – the manifold. As a result of the very mild assumptions made, the only way to determine the structure of the manifold generated by a single object is to densely sample it such that any other appearance of the object can be approximated locally and linearly by the collected samples. In many cases this implies the collection of a very large number of samples. In the presence of noise, *i.e.*, when the observed data does not lie exactly on the manifold, the performance of these manifold learning methods degrades significantly. In [17] a method for manifold denoising using spectral graph wavelets is presented. The method employs the inherent smoothness of the manifold.

While there are many cases where no prior knowledge of the sources of the variability in the appearances of an object is available, there are many scenarios in which such information is inherently available, and hence can be efficiently exploited. We concentrate on the case where the geometric deformations, and the observation noise, are the major source for the variability in the appearances of the object. In this paper we present a method that exploits this type of a priori knowledge in order to enable efficient detection and recognition of *multiple and deformable* objects.

It has been shown, [1] that, in the absence of noise, in those cases where the set of deformations admits a finite-dimensional representation, the universal manifold embedding provides a mapping from the space of observations to a low dimensional linear space. The manifold corresponding to each object is mapped to a *distinct linear subspace* of Euclidean space, and the dimension of the subspace is the same as that of the manifold. In the presence of noise, different observations will be mapped to different subspaces. In this paper we derive a method for "averaging" the different subspaces, obtained from different observations made on the same object, in order to estimate the mean representation of the object manifold.

II. The UME of Affine Deformed Objects

The basic problem addressed in this paper is the following: Given a sequence of noisy observations of an object undergoing different geometric deformations, estimate the best representation of the noise free object manifold in a sense which we rigorously define in this paper. More specifically, assume we are given a sequence of observations such that any two of them, g and h are affine related, bounded, Lebesgue measurable functions from $\mathbb{R}^n$ to $\mathbb{R}$ with compact support, *i.e.*,

$$h\left(\mathbf{x}\right) = g\left(\mathbf{A}\mathbf{x} + \mathbf{c}\right) + n(\mathbf{x}) \quad \forall \mathbf{x} \in \mathbb{R}^n \tag{1}$$

where $\mathbf{A} \in GL_n(\mathbb{R})$ and $\mathbf{c} \in \mathbb{R}^n$

To simplify the presentation let us first define the problem in the noise-free setting. Let O be the space of observations, let Φ be the set of possible geometric deformations with N degrees of freedom, and let S be a set of known objects, where each object is modeled as a compactly-supported, bounded, and Lebesgue measurable (or more simply, integrable) function from $\mathbb{R}^n$ to $\mathbb{R}$. We assume that the observations are constructed by the following procedure: we first choose an object $g \in S$ and an arbitrary geometric deformation $\varphi \in \Phi$. Next, we define an operator $\psi : S \times \Phi \to O$ that acts on an object and a geometric deformation, producing an observation. The observation is $o = \psi(g, \varphi)$. For a specific object $g \in S$ we will denote by $\psi_g : \Phi \to O$ the restriction of the map to this object. We assume that the N parameters characterizing Φ are the coefficients of a linear combination of a priori known basis functions that completely specify the action of the group of geometric transformations the object may undergo. For example, if Φ is the set of functions describing invertible two-dimensional affine deformations then Φ is of dimension 6, as these 6 parameters define the geometric transformation along the two-axes. For any object (function) $g \in S$ the set of all possible observations on this particular function is denoted by S_g. We refer to this subset as the orbit of g under Φ. In general, ψ_g is not linear, and hence S_g is a nonlinear manifold, in the space of functions. We note here, that in the context of this paper the term "manifold" adopted from the machine learning and dimensionality reduction literature, refers to the orbit of g under Φ, *i.e.*, to the set of all possible observations on g due to the action of the group defined by Φ.

In general O has a very high dimension (*e.g.*, the number of pixels in an image). It is composed of the union of orbits, S_g, of the different objects g in S such that each orbit S_g is the result of the action of the group of coordinate transformations defined by Φ on the object g.

In [1], it is shown that under the above assumptions and for some specific choices of Φ (which include the group of affine transformations considered in this paper) there exists a linear space $H \subset \mathbb{R}^M$ and a map $T : O \to H$, such that the restriction of this map to S_g is such that the composed map $T^g = T \circ \psi_g$ is a linear and invertible map from the finite-dimensional representation of Φ to H. These properties hold for every object $g \in S$ and the map T is *independent* of the object. The map T, is called *universal manifold embedding* as
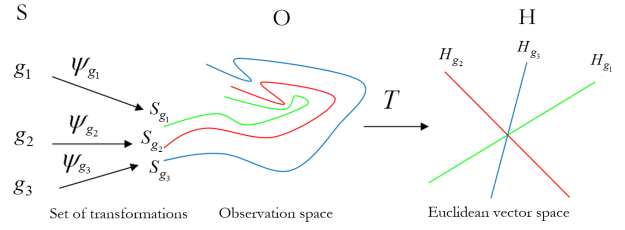

Fig. 1. The Universal Manifold Embedding framework (from left to right): The physical model that generates the observations - applying the set of possible deformations to some object g produces S_g which is the set of all possible observations on g. S_g is a subset of the space of observations O. The UME - all observations in S_g are nonlinearly mapped by T to a unique linear subspace $H_g = T(S_g)$.

it universally maps each of the different manifolds, each manifold corresponding to a single object, into a *different* linear subspace such that the overall map $T \circ \psi_g : \Phi \to H$ is linear in the parameterization of Φ. The map $\psi_g : \Phi \to O$ maps Φ nonlinearly and represents the physical relation between the object and the observations on it. The map $T : O \to H$ maps O nonlinearly such that the overall map $T \circ \psi_g : \Phi \to H$ is linear. This universal map allows us to represent the (mapped) observations in a space where the action of Φ is linear. Figure 1 schematically illustrates the concept of the UME in the noise free case.

We next describe the implementation of the concept of universal manifold embedding for the case where the geometric deformations are affine.

Consider the case where h is an observation of g undergoing an affine transformation. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, *i.e.*, $\mathbf{x} = [x_1, x_2, \ldots, x_n]^T$, $\mathbf{y} = [y_1, y_2, \ldots, y_n]^T$, such that $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{c}$, $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} + \mathbf{b}$. Let $\tilde{\mathbf{y}} = [1, y_1, \ldots, y_n]^T$. Thus, $\mathbf{x} = \mathbf{D}\tilde{\mathbf{y}}$ where $\mathbf{D}$ is an $n \times (n+1)$ matrix given by $\mathbf{D} = [\mathbf{b} \ \mathbf{A}^{-1}]$.

Let $M \in \mathbb{N}$ and let $w_l \ \ l = 1, \ldots, M$ be a set of bounded, Lebesgue measurable functions $w_l : \mathbb{R} \to \mathbb{R}$. Let $\mathbf{D}_k$ denote the kth row of the matrix $\mathbf{D}$. Then, [13],

$$\int_{\mathbb{R}^n} x_k w_\ell \circ h(\mathbf{x}) d\mathbf{x} = \left|\mathbf{A}^{-1}\right| \int_{\mathbb{R}^n} (\mathbf{D}_k \tilde{\mathbf{y}}) w_\ell \circ g(\tilde{\mathbf{y}}) d\tilde{\mathbf{y}} \tag{2}$$

Let f be some observation on a deformable object and let

$$\mathbf{T}(f) = \begin{bmatrix} \int_{\mathbb{R}^n} w_1 \circ f(\mathbf{y}) & \int_{\mathbb{R}^n} y_1 w_1 \circ f(\mathbf{y}) & \cdots & \int_{\mathbb{R}^n} y_n w_1 \circ f(\mathbf{y}) \\ \vdots & \ddots & & \vdots \\ \int_{\mathbb{R}^n} w_M \circ f(\mathbf{y}) & \int_{\mathbb{R}^n} y_1 w_M \circ f(\mathbf{y}) & \cdots & \int_{\mathbb{R}^n} y_n w_M \circ f(\mathbf{y}) \end{bmatrix}$$

Let $r = n + 1$. Hence $\mathbf{T}(f)$ is a rank r matrix. Denote $\tilde{\mathbf{D}} = [\mathbf{e}_1 \ \mathbf{D}^T]$ where $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$. Then, if h is an observation of g undergoing an affine deformation represented by the matrix $\mathbf{D}$, then from (2) we get:

$$\mathbf{T}(g)\left|\mathbf{A}^{-1}\right|\tilde{\mathbf{D}} = \mathbf{T}(h) \tag{3}$$

Since the deformations at hand are invertible, this implies that the column space of $\mathbf{T}(g)$ and the column space of $\mathbf{T}(h)$ are the same subspace. Thus, after applying the mapping $\mathbf{T}$ to the space of observations, the problem of object recognition becomes a problem of classifying subspaces.

III. CLASSIFICATION USING UME INDUCED SUBSPACE

A. Projection Matrices

Since projection matrices have a one-to-one correspondence to subspaces, a subspace can be classified using its respective projection matrix. That is, given two matrices $\mathbf{A}, \mathbf{B}$, we measure the distance between $range(\mathbf{A})$ and $range(\mathbf{B})$, via the orthogonal projections for both subspaces, $\mathbf{P_A}, \mathbf{P_B}$, respectively, *i.e.*,

$$\|\mathbf{P_A} - \mathbf{P_B}\|_F \tag{4}$$

As shown above, (see also, [15]) in the absence of noise, given two affine related observations on the same object, their distance should satisfy $d(h, g) = \|\mathbf{P}_{\mathbf{T}(h)} - \mathbf{P}_{\mathbf{T}(g)}\|_F = 0$. Since in practice we do not have noise-free observations on the objects, our goal here is to obtain a statistic representing the entire manifold, or equivalently, to estimate the *mean manifold* from the available observations. Since in the absence of noise, for any two observations on the same object $\|\mathbf{P}_{\mathbf{T}(h)} - \mathbf{P}_{\mathbf{T}(g)}\|_F = 0$, the choice of the function h or g as the representative of the equivalence class formed by the set of all possible observations on the object, which we denoted by S_g is arbitrary. Hence, in the following we present a method for "averaging" the noisy projection matrices and employ this average as a sufficient statistic of the entire object manifold.

B. The Subspace Mean

Intuitively, given a set of projection matrices, it is natural to estimate the orthogonal projection onto the "mean subspace" as the average of the realizations: $\mathbf{P}_{avg} = \frac{1}{L} \sum_{m=1}^{L} \mathbf{P}_m$. The problem with such an estimate is that a sum of projection matrices is not a projection matrix. Therefore, several other definitions of the mean have been used, the most commonly used among them is the Karcher-mean. In [16] a thorough analysis of different types of subspace means is presented. In the following we provide a derivation of the extrinsic mean adapted to the set-up of the above formulation.

Let $\mathcal{P}_r$ denote the set of orthogonal projections onto the *same* dimension-r subspaces in an M-dimensional ambient space. Given $\{\mathbf{P}_m\}_{m=1}^{L} \in \mathcal{P}_r$ find $\overline{\mathbf{P}} \in \mathcal{P}_r$- the "best representative" of the set in the following sense: Let $\mathbf{e}_i \in \mathbb{R}^M$ be the i-th vector of the standard basis. We wish to find a matrix $\overline{\mathbf{P}}$ such that

$$\overline{\mathbf{P}} = \arg\min_{\mathbf{P} \in \mathcal{P}_r} \frac{1}{L} \sum_{i=1}^{M} \sum_{m=1}^{L} \|(\mathbf{P} - \mathbf{P}_m)\mathbf{e}_i\|^2 \tag{5}$$

$$= \arg\min_{\mathbf{P} \in \mathcal{P}_r} \frac{1}{L} \sum_{i=1}^{M} \sum_{m=1}^{L} \mathbf{e}_i^T (\mathbf{P} - \mathbf{P}_m)^T (\mathbf{P} - \mathbf{P}_m)\mathbf{e}_i$$

Let $\mathbf{Q} = \frac{1}{L} \sum_{m=1}^{L} \mathbf{P}_m$. Then

$$\begin{aligned} \overline{\mathbf{P}} &= \arg\min_{\mathbf{P} \in \mathcal{P}_r} \sum_{i=1}^{M} \mathbf{e}_i^T (\mathbf{P} - \mathbf{Q}\mathbf{P} - \mathbf{P}\mathbf{Q} + \mathbf{Q})\mathbf{e}_i \\ &= \arg\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}(\mathbf{P} - \mathbf{Q}\mathbf{P} - \mathbf{P}\mathbf{Q} + \mathbf{Q}) \\ &= \arg\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}\left[(\mathbf{P} - \mathbf{Q})(\mathbf{P} - \mathbf{Q}) + (\mathbf{Q} - \mathbf{Q^2}) \right] \quad (6) \end{aligned}$$

Since the term $(\mathbf{Q} - \mathbf{Q^2})$ is not affected by the choice of $\mathbf{P}$, we obtain an equivalent problem.

$$\overline{\mathbf{P}} = \arg\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}\left[(\mathbf{P} - \mathbf{Q})(\mathbf{P} - \mathbf{Q}) \right] \tag{7}$$

Since $\mathbf{P}$ is a projection matrix, $\mathbf{P} = \mathbf{U}\mathbf{U}^T$. Since $\mathbf{Q}$ is the sum of projection (and hence normal) matrices, it is normal. Hence, its SVD decomposition is given by $\mathbf{Q} = \mathbf{F}\mathbf{K}\mathbf{F}^T$, $\mathbf{K} = diag(k_{i,i})$, $k_{1,1} \geq k_{2,2} \geq ... \geq k_{M,M}$, $\mathbf{F}$ being a unitary matrix. Then,

$$\begin{aligned} &\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}\left[(\mathbf{P} - \mathbf{Q})(\mathbf{P} - \mathbf{Q}) \right] \\ =\ &\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}\left[(\mathbf{U}\mathbf{U}^T - \mathbf{F}\mathbf{K}\mathbf{F}^T)(\mathbf{U}\mathbf{U}^T - \mathbf{F}\mathbf{K}\mathbf{F}^T) \right] \\ =\ &\min_{\mathbf{P} \in \mathcal{P}_r} \operatorname{tr}\left[\mathbf{U}\mathbf{U}^T \right] + \operatorname{tr}\left[\mathbf{F}\mathbf{K}^2\mathbf{F}^T \right] - 2\operatorname{tr}[\mathbf{U}^T\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{U}] \\ =\ &\min_{\mathbf{P} \in \mathcal{P}_r} \underbrace{\operatorname{tr}\left[\mathbf{U^T}\mathbf{U} \right]}_{r} + \sum_{i=1}^{M} k_{i,i}^2 - 2\operatorname{tr}[\mathbf{U}^T\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{U}] \\ =\ &\min_{\mathbf{P} \in \mathcal{P}_r} r + \sum_{i=1}^{M} k_{i,i}^2 - 2\operatorname{tr}[\mathbf{U}^T\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{U}] \quad (8) \end{aligned}$$

In order to minimize the error, we must minimize (8), which is the same as maximizing $\operatorname{tr}[\mathbf{U}^T\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{U}]$. Let $\mathbf{g}_i \triangleq \mathbf{F}^T\mathbf{u}_i$ and $\mathbf{G} \triangleq \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_M \end{bmatrix}$. Then,

$$\begin{aligned} &\max_{\mathbf{U}} \operatorname{tr}[\mathbf{U}^T\mathbf{F}\mathbf{K}\mathbf{F}^T\mathbf{U}] = \\ =\ &\max_{\mathbf{U}} \mathbf{G}^T \begin{bmatrix} k_{1,1} & 0 & \cdots & 0 \\ 0 & k_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k_{M,M} \end{bmatrix} \mathbf{G} = \\ =\ &\max_{\mathbf{U}} \sum_{i=1}^{M} k_{i,i} \|\mathbf{g}_i\|^2 = \max_{\mathbf{U}} \sum_{i=1}^{M} k_{i,i} \|\mathbf{F}^T\mathbf{u}_i\|^2 = \\ =\ &\max_{\mathbf{U}} \sum_{i=1}^{M} k_{i,i} \mathbf{u}_i^T \underbrace{\mathbf{F}\mathbf{F}^T}_{\mathbf{P_F}} \mathbf{u}_i = \sum_{i=1}^{M} k_{i,i} \mathbf{u}_i^T \mathbf{P_F} \mathbf{u}_i = \\ =\ &\max_{\mathbf{U}} \sum_{i=1}^{M} k_{i,i} \mathbf{u}_i^T \mathbf{P_F^T}\mathbf{P_F}\mathbf{u}_i = \max_{\mathbf{U}} \sum_{i=1}^{M} k_{i,i} \|\mathbf{P_F}\mathbf{u}_i\|^2 \end{aligned}$$

where we define $\mathbf{P_F} = \mathbf{F}\mathbf{F}^T$ since $\mathbf{F}$ is a unitary matrix. Since the space is r-dimensional, in order to minimize the error, we must choose $\mathbf{U} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_r \end{bmatrix}$, where $\mathbf{f}_i, i = 1, .., r$ are the r eigenvectors of $\mathbf{Q}$ that are associated with the r largest eigenvalues $k_{1,1}, ..., k_{r,r}$, and $\overline{\mathbf{P}} = \mathbf{U}\mathbf{U^T}$.

Following this derivation, each object C_j is represented by a single projection matrix $\overline{\mathbf{P}}_j$, and its corresponding *averaged UME* $\mathbf{U}_j$.

IV. CLASSIFICATION OF NOISY OBSERVATIONS ON DEFORMED OBJECTS

When using the UME it is necessary to first choose M, the dimension of the ambient space, *i.e.*, the number of w functions. We next show how to empirically choose the optimal value of M, such that the probability of correctly classifying the objects of interest is maximized.

To evaluate the performance of the new approach we compare the performances of two classifiers using the same training and test sets for both, as a function of both the noise level and the dimension M of the ambient space. Note, that since in this problem the objects are 2-D images undergoing affine transformations, and since in the definition of $\mathbf{T}(f)$ only moments of order zero and one of $w_\ell \circ f$ are employed, we have that $r = 3$. The first classifier employs the estimated mean manifold representation using the corresponding mean projection matrix $\mathbf{P}_j$ for each object j, evaluated from the observations in the training data set. The second classifier is based on the nearest neighbor method. Using the mean manifold based method, we decide that a noisy observation h belongs to class j if the distance of its corresponding projection matrix $\mathbf{P}_{\mathbf{T}(h)}$ to the mean projection matrix $\mathbf{P}_j$ is minimal. In the nearest neighbor method we decide that the observation belongs to class j if the distance of $\mathbf{P}_{\mathbf{T}(h)}$ to any of the projection matrices associated with observations that belong to object j is minimal. The presented results are based on 50 independent experiments. For each experiment, 10 images are chosen at random out of a fixed set of 240 different images. All the images are grey scale with values ranging between 0 and 1 and are 400×400 pixels. In each of the 50 tests, 40 new geometrically deformed images are synthesized using random affine deformations. To simulate different SNR's, WGN is added to each deformed instance. Figure 2 provides an example of two images from the set, an example of a deformed instance and a deformed and noisy instance. 20 of the generated images are used for training and the other 20 are used for performance evaluation. Since for each object there are 40 deformed observations, overall there are 200 images used for training and 200 images are used for performance evaluation in each test. The classifiers are evaluated at different SNR's where $0 \leq \sigma \leq 0.5$, and different choices for the dimensions of the ambient space. Thus, for each given σ and M, 50 tests are performed. The w functions used are the following indicators:

$$w_{i,M}(x) = \begin{cases} 1, & \frac{i-1}{M} < x \leq \frac{i}{M} \\ 0, & else \end{cases} \;, i = 1, ..., M \quad (9)$$

For a given SNR, and for a given dimension of the ambient space, M, the performance is measured as the ratio of the number of correct classifications to the number of attempted classifications, measured jointly for all the objects in the data set (and not tuned to detect any particular object). In Figure 3 the classification results are presented. We clearly see that the classification method based on the subspace means outperforms the nearest neighbor method in all cases, and

Fig. 2. First row: Examples of images from the data set. Second row from left to right: A deformed image, a deformed image in the presence of WGN with $\sigma = 0.3$
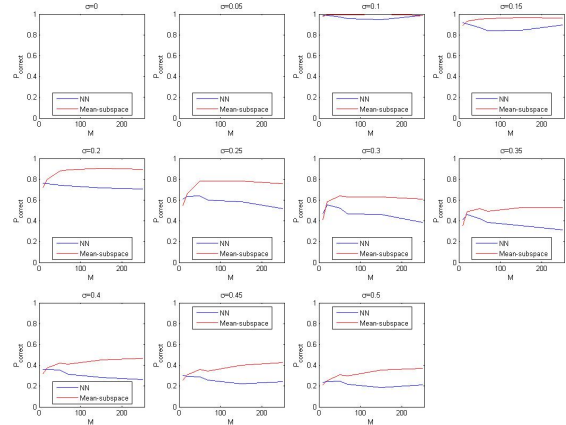

Fig. 3. The performance of the two classifiers for different SNRs as a function of the dimension of the ambient space, M. Red: performance of the classifier based on mean subspaces. Blue: performance of the nearest-neighbor approach.

that no significant gain performance is obtained by choosing ambient space dimension larger than 50 .

V. SUMMARY

We have presented a method for estimating the mean manifold, where the manifold is generated by the set of all possible affine transformations of the object, from noisy and geometrically transformed observations on the object. The universal manifold embedding provides a mapping from the space of observations to a low dimensional linear space. In the noise-free case, the manifold corresponding to each object is mapped to a *distinct linear subspace* of Euclidean space. In the presence of noise, different observations are mapped to different subspaces. We have presented a method for averaging the UME subspaces, obtained from different observations made on the same object, in order to estimate the mean representation of the object manifold. The mean manifold representation has been employed to estimate the optimal dimension of the ambient space such that the probability of correctly classifying noisy and affine deformed observations on a set of objects is maximized.

REFERENCES

[1] R. Hagege and J. M. Francos, "Universal Manifold Embedding for Geometrically Deformed Functions," *IEEE Trans. Info. Theory*, acceptted for publication, 2016.

[2] P. Dollar, V. Rabaud and S. Belongie, "Learning to Traverse Image Manifolds", *Proc. NIPS*, 2006.

[3] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, vol. 290, pp. 2319–2323, 2000.

[4] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science*, vol. 290, pp. 2323-2326, 2000.

[5] C. Walder and B. Schölkopf, "Diffeomorphic Dimensionality Reduction" *Proc. NIPS*, 2008.

[6] Z. Zhang and H. Zha, "Principal Manifolds and Nonlinear Dimension Reduction via Local Tangent Space Alignment," *SIAM Jou. Sci. Comp.* vol. 26 , pp. 313 - 338, 2005.

[7] "Special Issue on Dimensionality Reduction Methods," *Signal Process. Mag.*, Mar. 2011.

[8] K. Weinberger and L. Saul, "Unsupervised Learning of Image Manifolds by Semidefinite Programming," *Int. J. Comput. Vision*, pp. 77-90, 2006.

[9] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Info. Theory*, vol. 50, pp. 2231-2242, 2004.

[10] I. Tosic and P. Frossard, "Dictionary Learning," *Signal Process. Mag.*, pp. 27-38, Mar. 2011.

[11] R. Vidal, "Subspace Clustering," *Signal Process. Mag.*, pp. 52-67, Mar. 2011.

[12] R. Baraniuk and M. Wakin, "Random Projections of Smooth Manifolds," *Found. Comput. Math.*, pp. 51-77, 2009.

[13] R. Hagege and J. M. Francos, "Parametric Estimation of Affine Transformations: An Exact Linear Solution," *Jou. of Math. Imaging and Vision.*, pp. 1-16, Jan. 2010.

[14] R. Hagege and J. M. Francos, "Linear Estimation of Time Warped Signals," *IEEE Trans. Info. Theory*, pp. 4423 - 4439, July 2013.

[15] R. Sharon, R. Hagege and J. M. Francos, "Detection and recognition of deformable objects using structured dimensionality reduction", in *40th IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Brisbane, AU, pp. 3442-3446, Apr. 2015.

[16] T. Marrinan, J. Beveridge, B. Draper, M. Kirby and C. Peterson, "Finding the Subspace Mean or Median to Fit Your Need," in *IEEE conf. on Computer Vision and Pattern Recognition*, pp. 1082-1089, June 2014.

[17] S. Deutsch, A. Ortega and G. Medioni, "Manifold Denoising Based on Spectral Graph Wavelets," *Proc. ICASSP 2016,* Shanghai.

[18] I. Santamaria, L. L. Scharf, C. Peterson, M. Kirby and J. M. Francos, "An Order Fitting Rule for Optimal Subspace Averaging," submitted for publication, *SSP 2016*, Majorca.