Locally Isometric and Conformal Parameterization of Image Manifold

Bernstein A.V.1,2, Kuleshov A.P.1,2, Yanovich Yu.A.1,2

1Institute for Information Transmission Problems RAS, Moscow
2National Research University Higher School of Economics, Moscow

ABSTRACT

Images can be represented as vectors in a high-dimensional Image space with components specifying light intensities at image pixels. To avoid the ‘curse of dimensionality’, the original high-dimensional image data are transformed into their lower-dimensional features preserving certain subject-driven data properties. These properties can include ‘information-preserving’ when using the constructed low-dimensional features instead of original high-dimensional vectors, as well preserving the distances and angles between the original high-dimensional image vectors. Under the commonly used Manifold assumption that the high-dimensional image data lie on or near a certain unknown low-dimensional Image manifold embedded in an ambient high-dimensional ‘observation’ space, a constructing of the lower-dimensional features consists in constructing an Embedding mapping from the Image manifold to Feature space, which, in turn, determines a low-dimensional parameterization of the Image manifold. We propose a new geometrically motivated Embedding method which constructs a low-dimensional parameterization of the Image manifold and provides the information-preserving property as well as the locally isometric and conformal properties.

Keywords: Image manifold, low-dimensional parameterization, isometry embedding, conformal embedding, manifold learning, tangent spaces, tangent vector fields.

1. INTRODUCTION AND PAPER MOTIVATION

Images can be represented as vectors in a high-dimensional Image space with components specifying light intensities at image pixels. Many methods used for solving of various Image analysis and Machine vision tasks, such as Pattern Recognition (for faces, hand-written characters, etc.), Image Classification, Clustering, and others [1-3], perform poorly in a high-dimensional space. Thus, various Representation learning (Feature extraction) algorithms are used as a first key step in solving of these tasks [4]. These algorithms transform the original high-dimensional image data into their lower dimensional features preserving certain subject-driven data properties such as local or global data geometry, proximity relations, geodesic distances, angles, etc.; in other words, an Embedding mapping from the high-dimensional ‘observation’ space to a low-dimensional Feature space is constructed. Then the low-dimensional features of the original data can be handled efficiently, avoiding the ‘curse of dimensionality’ phenomenon.

The possibility of constructing such lower dimensional features for image data is determined by the underlying low-dimensional structure in the data, and the features must retain most of this structure. As a rule, high-dimensional real-world data are ‘explained’ by a small number of factors; in appearance-based vision, for example, the observed image is often controlled by a small number of factors like the view angle and lighting direction. These properties of real-world data are defined formally as the Manifold assumption which was first proposed in [5] and is that the image data lie on or near certain unknown low-dimensional Image manifold (IM) embedded in an ambient high-dimensional ‘observation’ space. This model is now the most popular model of the imaging data [1,2,3,6]; intensive studies under this model have formed a new emerging field in Data analysis called Manifold Learning [7].

Further use of constructed features defines what particular properties should be preserved. In general, to prevent losses when using the low-dimensional features instead of original high-dimensional vectors, the Embedding mapping must preserve as much as possible available information contained in the high-dimensional data [8]; this means a possibility of constructing a reconstruction function from the Feature space to the high-dimensional observation space with small reconstruction error. Such possibility is required in many applied tasks, and Manifold Learning methods should ‘give natural tools to project and re-project new images onto the Image manifold’ [1]. For example, the frame rate conversion algorithms are widely used for video compression, format conversion and quality enhancement [9,10] to make movement of objects smoother and, therefore, more enjoyable to view. Many industrial companies use inter-frame interpolation to change flow from 50Gz to 100 or 200 Gz flow, but such interpolation of moving objects is quiet...
complicated task which can be considerably simplified by a transition to reduced interpolation problem in the Feature space. But a solution obtained in the reduced interpolation problem lies in the low-dimensional Feature space, so, it is necessary to accurately reconstruct the interpolated frames from their low-dimensional features.

Important property of a Representation method is how the low-dimensional features preserve distances and angles between the original high-dimensional image vectors. Desirability of these properties has motivated intensive research in imaging applications [11-15]. However an exact isometric Embedding mapping of the IM to a low-dimensional vector space is only possible in cases where the IM has zero sectional curvature [16] (coinciding with the Gaussian curvature for two-dimensional surfaces). That is why a strict isometric requirement is replaced with an appropriate local analogue.

In this paper we propose a new geometrically motivated Embedding method which has the ‘information-preserving’ as well the ‘locally isometric and conformal’ properties; strict definitions will be given in Section 2. Section 3 describes the proposed method; results of performed comparative numerical experiments are presented in Section 4.

2. INFORMATION-PRESERVING AND LOCALLY ISOMETRIC & CONFORMAL PROPERTIES

2.1 Information-preserving Property

Let $M$ be an unknown q-dimensional Image manifold embedded in an ambient p-dimensional ‘observation’ space $R^p$, $q < p$, and covered by a single chart. Let $X_n = \{X_1, X_2, \ldots, X_n\} \subset M$ be a dataset randomly sampled from the IM $M$. The inner manifold dimension $q$ is assumed to be known.

Denote by $h: M \subset R^p \rightarrow Y_h = h(M) \subset R^q$ the desired sample-based Embedding mapping from the IM $M$ to the Feature space (FS) $Y_h = h(M)$; this mapping determines a low-dimensional parameterization of the IM.

As was pointed out above, the ‘information-preserving’ property of the Embedding mapping $h$ means a possibility for reconstruction of the original vectors $X \in M$ from their features $y = h(X)$ with small reconstruction error. Denote by $g$ the desired reconstruction mapping from the FS $Y_h \subset R^q$ to the ambient ‘observation’ space $R^p$ which must ensure the approximate equality

$$r_{h,g}(X) \approx X \quad \text{for all } X \in M,$$

here $r_{h,g}(X) = g(h(X))$ is the reconstructed value of an original vector $X \in M$ as a result of successively applying the embedding and reconstruction mappings to the vector $X$. Reconstruction error $\delta_{h,g}(X) = |X - r_{h,g}(X)|$ is a valid evaluation measure (‘universal quality criterion’) of preserving the information contained in high-dimensional data [8].

The mappings $(h, g)$ determine a q-dimensional Image reconstructed manifold (IRM)

$$M_{h,g} = r_{h,g}(M) = \{X = g(y) \in R^p; \, y \in Y_h \subset R^q\}$$

embedded in $R^p$ and parameterized by a single chart $g$. The approximate equalities (1) can be considered as the manifold proximity $M_{h,g} \approx M$ meaning that the IRM $M_{h,g}$ (2) accurately reconstructs the IM $M$ from the sample.

The Reconstruction error $\delta_{h,g}(X)$ can be directly computed at sample points $X \in X_n$; for Out-of-Sample points $X \in M \setminus X_n$ it describes the generalization ability of the considered solution $(h, g)$ at a specific point $X$. As was shown in [17], the greater the distances between the tangent spaces $L(X)$ and $L_{h,g}(r_{h,g}(X))$ to the IM $M$ and IRM $M_{h,g}$ at the points $X \in M$ and $r_{h,g}(X) \in M_{h,g}$, respectively, the lower the local generalization ability of the solution $(h, g)$ and the less the mapping $r_{h,g}: M \rightarrow M_{h,g}$ preserves local structure of the IM $M$; here $L_{h,g}(r_{h,g}(X)) = \text{Span}(J_g(h(X)))$ is the linear space spanned by columns of the $p \times q$ Jacobian matrix $J_g(y)$ of the mapping $g(y)$ at the point $y = h(X) \in Y_h$. The tangent spaces are elements of the Grassmann manifold Grass$(p, q)$ consisting of all q-dimensional linear subspaces in $R^p$, and the distance in the Grass$(p, q)$ used in [17] is the projection 2-norm metric [18]. Thus, it is natural to require that the solution $(h, g)$ ensures not only Manifold proximity (1) but also Tangent proximity

$$L(X) \approx L_{h,g}(r_{h,g}(X)) \quad \text{for all } X \in M$$

between the tangent spaces in certain selected metric on the Grass$(p, q)$.
In topology, the set composed of points $X$ of the manifold $M$ equipped by tangent spaces $L(X)$ at these points is known as the Tangent bundle of the manifold $M$. Thus, the problem consisting in accurate reconstruction of both the IM and its differential structure (tangent spaces to the IM) from manifold-valued data can be referred to as the Tangent bundle manifold learning problem. Motivation of this problem is also that ‘natural IM are often very curved in ambient high-dimensional space and Manifold learning methods must be adapted to a differential structure of the IM’, whose curvature is characterized by a variety of the manifold tangent spaces [1].

Thus, the information-preserving property of the Embedding mapping $h$ will mean in what follows the possibility for constructing the mappings $h$ and $g$ in such a way that the pair $(h, g)$ provides the Manifold and Tangent proximities.

2.2 Locally Isometric and Conformal Properties

Locally isometric and conformal properties of the Embedding mapping $h$ will mean in what follows that the relations

$$(X' - X, X'' - X) \approx (h(X') - h(X), h(X'') - h(X)) \tag{4}$$

hold true for near points $X'$, $X''$ and $X$ from the IM $M$. Indeed, relation (4) with $X'' = X'$ means an isometric property

$$||h(X') - h(X)|| \approx |X' - X|; \tag{5}$$

relation (4) implies also local preservation of the angles:

$$\cos(\gamma_{h}(X', X'', X)) \approx \cos(\gamma_{p}(X', X'', X)), \tag{6}$$

which means a conformal property; here $\gamma_{p}(X', X'', X)$ and $\gamma_{h}(X', X'', X)$ are the angles between the original vectors $(X' - X)$ and $(X'' - X)$ and between their features’ vectors $(h(X') - h(X))$ and $(h(X'') - h(X))$, respectively.

Under the desired property (1), the Taylor series expansion of the mapping $g$ gives the approximate relation

$$X' - X \approx J_{g}(h(X)) \times (h(X') - h(X)), \tag{7}$$

from which, using the notation $D_{g}(h(X)) = J_{g}^{T}(h(X)) \times J_{g}(h(X))$, we obtain the relations

$$(X' - X, X'' - X) \approx (h(X') - h(X))^{T} \times D_{g}(h(X)) \times (h(X'') - h(X)) \tag{8}$$

for near points $X'$, $X''$ and $X$ from the IM $M$. Therefore, for satisfying the locally isometric (5) and conformal properties (6), the matrices $J_{g}(h(X))$ must be orthogonal.

3. THE PROPOSED EMBEDDING ALGORITHM

The proposed method is based on the Grassmann&Stiefel Eigenmaps (GSE) algorithm [17,19] which solves the Tangent bundle manifold learning problem; the GSE-Embedding algorithm being a part of the GSE provides the information-preserving property. However, the GSE-Embedding doesn’t provide the isometric and conformal properties, and, to achieve them, a new orthogonal version of the GSE (OGSE), whose Embedding part meets these properties, is proposed. The proposed OGSE is based on a solving of the Orthogonal Tangent manifold learning problem consisting in construction of smooth orthonormal tangent vector fields on the IM.

The required details of the GSE are contained in subsection 3.1; subsection 3.2 describes the proposed method.

3.1 Grassmann&Stiefel Eigenmaps Algorithm: Some Details

The GSE consists of three successively performed steps: Tangent manifold learning, Manifold embedding, and Manifold reconstruction.

In the Tangent manifold learning step, a sample-based family $H$ consisting of $p \times q$ matrices $H(X)$ smoothly depending on $X \in M$ is constructed to meet the relations $L_{q}(X) \approx L(X)$; here $L_{q}(X) = \text{Span}(H(X))$ are $q$-dimensional
linear spaces in $\mathbb{R}^p$ spanned by columns $H^{(1)}(X), H^{(2)}(X), \ldots, H^{(q)}(X)$ of the matrix $H(X)$. The linear space $L_d(X)$ will be the tangent space $L_{d,g}(h,g(X))$ to the further constructed IRM $M_{h,g}$ (2) at the point $r_{h,g}(X)$.

In the Manifold embedding step, given the family $H$ already constructed, the embedding mapping $y = h(X), X \in M$, is constructed to meet the relations $X' = X \approx h(X) \times (h(X') - h(X))$ for near points $X, X' \in M$. These relations are considered as regression equations for the features $\{h(X)\}$ and the standard least squares technique is used for solving this regression problem. The constructed mapping $h$ determines the Feature space $Y_h = h(M)$.

In the Manifold reconstruction step, given the already constructed family $H$ and mapping $h(X)$, the Reconstruction mapping $g$ is constructed to meet the desired proximities (1) and $J_d(h(X)) \approx H(X)$; the latter relation provides the tangent proximity (3).

The constructed matrices $H(X)$, which estimate the Jacobian matrices $J_d(h(X))$, are not orthogonal in general; thus, the GSE does not provide orthogonality of the matrix $D_d(h(X))$. In the proposed OGSE algorithm, the matrices $H(X)$ satisfy the additional orthogonal condition; other GSE steps are unchanged. Orthogonality of the matrices $H(X)$ means orthonormality of tangent vector fields $H^{(1)}(X), H^{(2)}(X), \ldots, H^{(q)}(X)$ on the IRM $M_{h,g}$.

### 3.2 Constructing the Orthonormal Tangent Vector Fields

The tangent space $L(X)$ at a point $X \in M$ is estimated by the $q$-dimensional linear space $L_{PCA}(X)$ spanned by columns of the $p \times q$ orthogonal matrix $Q_{PCA}(X)$ which is a result of the PCA applied to sample points from an $\varepsilon$-ball in $\mathbb{R}^p$ centered at $X$; the columns of $Q_{PCA}(X)$ are orthogonal principal vectors corresponding to the $q$ largest PCA eigenvalues.

The orthonormal vectors $H^{(j)}(X)$ are constructed step by step to meet the relations $H^{(j)}(X) \in L_{PCA}(X), j = 1, 2, \ldots, q$; at the $j$th step, the $(q-j+1)$-dimensional linear spaces $L_{PCA,j}(X)$, kernels $K(X, X')$, and vectors $H^{(j)}(X)$ are constructed.

At the first step ($j = 1$), we set $L_{PCA,1}(X) = L_{PCA}(X)$ and $K_1(X, X') = K(X, X')$; here $K(X, X') = K_0(X, X') \times K_0(X, X')$ is kernel in which $K_0(X, X')$ and $K_0(X, X')$ are Euclidean and Grassmann Binet-Cauchy kernels [17,18,19], respectively.

At the $j$th step, $j > 1$, given the constructed vector $H^{(j-1)}(X)$ and the space $L_{PCA,j-1}(X)$, introduce linear space $L_{PCA,j}(X)$ as $L_{PCA,j}(X) \cap L_{j+1}^{+}(X)$, where $L_{j+1}^{+}(X)$ is one-dimensional linear space spanned by the vector $H^{(j-1)}(X)$, and symbol $\perp$ denotes orthogonal complement to a linear space. Put $Q_j(X) = Q_{PCA}(X)$ and denote by $Q_j(X)$ the $p \times (q-j+1)$ orthogonal matrix whose columns form an orthogonal basis in the linear space $L_{PCA,j}(X)$; this matrix can be constructed by using standard projection and orthogonalization procedures applied to $Q_j(X)$. Let $K_j(X, X') = K_0(X, X') \times K_0(X, X')$ be the kernel in which $K_0(X, X')$ is the Binet-Cauchy kernel on the Grassmann manifold Grass($p, q-j+1$).

First, the vectors $H_{j,i} \in L_{PCA,i}(X), i = 1, 2, \ldots, n$, are constructed to minimize $\Delta_j = \sum_{i=1}^{n} K_j(X_i, X_j) \times \left( H_{j,i} - H_{j,i} \right)^2$ under the constraint $\sum_{i=1}^{n} K_j(X_i, X_j) \times \left( H_{j,i} \right)^2 - 1 = 0$ required to avoid a degenerate solution; here $K_j(X) = \sum_{i=1}^{n} K_j(X_i, X_j)$.

The condition $H_{j,i} \in L_{PCA,i}(X)$ implies a representation $H_{j,i} = Q_j(X_i) \times V_j$ in which $V_j$ is a $(q-j+1)$-dimensional vector. Under this representation, the quadratic form $\Delta_j$ and the constraint take a form $\Delta_j = 2 \left(1 - Tr(V_j^T \times F_j \times V_j)\right)$ and $V_j^T \times F_j \times V_j = 1$, respectively; here $V_j$ is an $(q-j+1)$-dimensional vector which consists of $n$ sequentially written vectors $v_{j1}, v_{j2}, \ldots, v_{jn}$, and $F_j$ and $\Phi_j$ are the $n \times (q-j+1)$ matrices written in explicit form. Therefore, the required vector $V_j$ is the eigenvector in the generalized eigenvector problem $\Phi_j \times V_j = \lambda \times F_j \times V_j$ corresponding to the largest eigenvalue; thus, the $(q-j+1)$-dimensional components $v_{j1}, v_{j2}, \ldots, v_{jn}$ of this vector $V_j$ determine the vectors $\{H_{j,i}\}$.

Consider the constructed vectors $\{H_{j,i}\}$ as preliminary values of the vector $H^{(j)}(X)$ at the sample points. The vector $H^{(j)}(X)$ for an arbitrary point $X \in M$ is chosen to minimize the quadratic form $\sum_{i=1}^{n} K_j(X_i, X_j) \times \left( H^{(j)}(X) - H_{j,i} \right)^2$ under the constraint $H^{(j)}(X) \in L_{PCA,j}(X)$. A solution of this problem in explicit form is given by the formula

$$H^{(j)}(X) = \pi(X) = \frac{1}{K_j(X)} \sum_{i=1}^{n} K_j(X_i, X_j) \times H_{j,i};$$

where $\pi(X) = Q_j(X) \times Q_j^T(X)$ is the projector onto the linear space $L_{PCA,j}(X)$.
Finally, we normalize the constructed vectors by putting $H(j)(X)$ to be equal to $H(j)(X) / |H(j)(X)|$. By the construction, the normalized tangent vector fields $\{H(1)(X), H(2)(X), \ldots, H(j)(X)\}$ are orthogonal to each other at each point $X \in M$.

4. RESULTS OF COMPARATIVE NUMERICAL EXPERIMENTS

Four two-dimensional manifolds (Fig. 1(a)) embedded into the three-dimensional space and having different Gaussian curvatures (zero, negative and positive curvature in the cylinder and cone, saddle, and ellipsoid, respectively) were used in numerical experiments performed to compare the proposed Embedding algorithm (as a part of the orthogonal version of the GSE, OGSE), standard GSE [17,19], and commonly used Embedding algorithms LLE [20], ISOMAP [21], Hessian eigenmaps (HLLE) [22], and LTSA [23]. All the compared algorithms were applied to the same random samples of different sizes (250, 500, 1000, 2000) uniformly distributed on these manifolds.

The algorithms were compared in the following characteristics: Reconstruction Errors (RE), Local Isometry Errors (LIE), and Local Conformal Errors (LCE). To estimate these characteristics, the corresponding Mean Squared Errors (MSE) were found using independent test samples of size $N = 1000$. Denote by $I$, $I_2$, and $I_3$ the sets consisting of the indices $\{1, 2, \ldots, N\}$ in test samples, pairs $(i, j)$, and triples $(i, j, k)$ of indices, $i, j, k \in I$, respectively.

The RE, LIE and LCE are estimated by the formulas:

$$ RE = \left( \frac{1}{N} \sum_{i=1}^{N} \left| g(i) - X_i \right|^2 \right)^{1/2}, $$

$$ LIE = \left( \frac{\sum_{(i,j) \in I_2} K_E(X_i, X_j) \left| h(X_i) - h(X_j) \right|^2}{\sum_{(i,j) \in I_2} K_E(X_i, X_j)} \right)^{1/2}, $$

and

$$ LCE = \left( \frac{\sum_{(i,j,k) \in I_3} K_E(X_i, X_j) \times K_E(X_i, X_k) \times \Delta_{i,j,k}^2}{\sum_{(i,j,k) \in I_3} K_E(X_i, X_j) \times K_E(X_i, X_k)} \right)^{1/2}, $$

where $\Delta_{i,j,k}^2 = (\cos(\gamma_{ij})(X', X'', X) - \cos(\gamma_{ij})(X', X'', X))^2$ and the LLE-reconstruction mapping $g$ [24] was used in (7) for the methods without their own reconstruction mapping.

Figure 1. Results of the performed numerical experiments.
The estimated errors RE, LIE and LCE for the compared methods are shown in Figs. 1(b), 1(c), and 1(d), respectively. These figures show that the GSE and OGSE outperform the other methods with respect to the RE; the OGSE has smaller both the LIE and LCE than the other methods. Note that, unlike the ISOMAP, an isometry property for the GSE was not expected.

5. CONCLUSION

Preserving the distances and angles between the original high-dimensional image vectors and their low-dimensional representations (features) is important and desired property of the Data Representation methods which are used in Image analysis and Machine vision tasks to avoid the ‘curse of dimensionality’ phenomenon. A new geometrically motivated embedding method which transforms the original high-dimensional image data into their lower-dimensional representations and has the ‘information-preserving’ property as well as the locally isometric and conformal properties is proposed. This method is based on construction of smooth orthonormal tangent vector fields on the Image manifold. In numerical experiments, the proposed method compares favourably with popular Manifold Embedding methods in terms of local isometric and conformal properties as well as the preserving information contained in the high-dimensional image data.

ACKNOWLEDGMENTS

The study was performed in the IITP RAS exclusively by the grant from the Russian Science Foundation (project № 14-50-00150).

REFERENCES


