Gen. Math. Notes, Vol. 22, No. 2, June 2014, pp. 133-142
ISSN 2219-7184;
www.i-csrs.org
Available free online at http://www.geman.in

Dimensionality Reduction Via Proximal Manifolds*

Randima Hettiarachchi[a] **and James Peters**[a,b,†]

[a]Computational Intelligence Laboratory
Department of Electrical & Computer Engineering
University of Manitoba
Winnipeg, Manitoba R3T 51-V6, Canada
Email: hettiarr@cc.umanitoba.ca,james.peters3@ad.umanitoba.ca
[b]Department of Mathematics, Faculty of Arts and Sciences
Adıyaman University, Adıyaman, Turkey

Abstract

The focus of this article is on studying the descriptive proximity of manifolds in images useful in digital image pattern recognition. Extraction of low-dimensional manifolds underlying high-dimensional image data spaces leads to efficient digital image analysis controlled by fewer parameters. The end result of this approach is dimensionality reduction, important for automatic learning in pattern recognition.

Keywords *Manifolds, Descriptive Proximity, Dimensionality Reduction.*
2010 MSC No: 57M50, 54E05, 54E40, 00A69

1 Introduction

This article introduces an application of descriptive isometry [11] to achieve dimensionality reduction via proximal manifolds [12], which simplifies the analysis of images via low-dimensional spaces that preserve the relevant structure in high-dimensional spaces. According to [17], a manifold is supposed to be "locally" like the metric space $\mathbb{R}^n$, so the simplest example of a manifold is just $\mathbb{R}^n$ itself. L. Nicolaescu [8] introduces a manifold as a Hausdorff space. Every manifold has a specific and well-defined dimension, the number of linearly

*This research has been supported by the The Scientific and Technological Research Council of Turkey (TÜBİTAK) Scientific Human Resources Development (BIDEB) under grant no: 2221-1059B211301223 and Natural Sciences & Engineering Research Council of Canada (NSERC) discovery grant 185986. †Corresponding Author.

independent vectors in the basis needed to specify a point [5].

A digital image is a typical example of a subset of a high-dimensional vector space. This high dimensionality becomes an issue in efficient processing and analysis of images. Manifold learning is used to address the problem of dimensionality reduction by mapping the high dimensional data into a low-dimensional space, while preserving the intrinsic structure of the data.

Two well-known manifold learning algorithms are Isometric Feature Mapping (ISOMAP) [18] and locally Linear Embedding (LLE) [15, 14]. LLE attempts to compute a low-dimensional embedding so that nearby points in the high-dimensional space remain nearby and similarly co-located with respect to one another in the low-dimensional space [15]. This approach to achieving dimensionality reduction can be extended to preserving the nearness of the description of points in the high-dimensional space mapped into a low-dimensional manifold.

To achieve this, we extend the conventional isometry defined for the mapping between high-dimensional data to low-dimensional manifold with the descriptive form of isometry introduced in [11] and elaborated in [13].

In pattern recognition, dimensionality reduction serves as an automatic learning approach to feature extraction, by combining all important cues (e.g., shape, pose, and lighting for image data) into a unified framework [6]. Thus, by learning the underlying low-dimensional manifold for a given image structure, the optimal feature vector for pattern recognition can be determined.

2 Preliminaries

Let X be a metric topological space endowed with one or more proximity relations, 2^X the collection of all subsets of X, $A, B \in 2^X$, $\mathrm{cl}(A)$ the Kuratowski closure of A, and δ the Efremovič proximity [2, 1] defined by

$$\delta = \left\{(A, B) \in 2^X \times 2^X : \mathrm{cl}(A) \ \cap \ \mathrm{cl}(B) \neq \emptyset\right\}.$$

The mapping $\Phi : X \longrightarrow \mathbb{R}^n$ is defined by $\Phi(x) = (\phi_1(x), \ldots, \phi_n(x))$, where $\mathbb{R}^n$ is an n-dimensional real Euclidean vector space. Let $\Phi = \{\phi_1, \ldots, \phi_n\}$ be a set of probe functions that represent features of each $x \in X$. Let $\mathcal{Q} : 2^X \longrightarrow 2^{R^n}$ be defined by $\mathcal{Q}(A) = \{\Phi(a) : a \in A\}$ for $A \in 2^X$. The *descriptive intersection* $\underset{\Phi}{\cap}$ of A and B is defined by

$$A \underset{\Phi}{\cap} B = \{x \in A \cup B : \mathcal{Q}(A) \ \delta \ \mathcal{Q}(B)\}.$$

The descriptive proximity relation δ_Φ is defined by

$$\delta_\Phi = \left\{(A, B) \in 2^X \times 2^X : \mathrm{cl}(A) \underset{\Phi}{\cap} \mathrm{cl}(B) \neq \emptyset\right\}.$$

The expression $A\ \delta_\Phi\ B$ reads A *is descriptively near* B. The pair (X, δ_Φ) is called a *descriptive proximity space*. For more details about such spaces, see [10].

A descriptive isometry is defined in the context of descriptive proximity spaces [9]. Before going into details about descriptively near manifolds, consider first the definition of a manifold [17].

Definition 1. Manifold
A *manifold* is a metric space M with the following property: If $x \in M$, then there is some neighbourhood U of x and some integer $n \geq 0$ such that U is homeomorphic to $\mathbb{R}^n$.

Continuous maps and homeomorphisms provide the backbone for metric spaces called manifolds. The definitions of continuous map and homeomorphism [19] are given below:

Definition 2. Continuous map
If X and Y are topological spaces, a map $f : X \to Y$ is said to be continuous if for every open set $U \subset Y, f^{-1}(U)$ is open in X.

Definition 3. Homeomorphism
If X and Y are topological spaces, a homeomorphism from X onto Y is defined to be a continuous bijective map $f : X \to Y$ with continuous inverse.

The following lemma [5] extends the definition of manifolds given in [17].

Lemma 1. *A topological space M is locally Euclidean space of dimension n, if and only if, either of the following properties holds:*
1^o Every point of M has a neighbourhood homeomorphic to an open ball in $\mathbb{R}^n$.
2^o Every point of M has a neighbourhood homeomorphic to $\mathbb{R}^n$.

Proof. Given in [5] □

According to [5], every manifold has a specific and a well-defined dimension. A k-dimensional manifold can be defined as follows:

Definition 4. A k-dimensional manifold in $\mathbb{R}^n$ is a subset $\mathbb{S}^k \subset \mathbb{R}^n$ such that each point has a neighbourhood in $\mathbb{S}^k$ that is homeomorphic to $\mathbb{R}^k$.

3 Descriptively Near Manifolds

Let $\mathbb{X} = [x_1, x_2, ..., x_N]^T$ be a set of points sampled from the high dimensional data space. The description of each point x_i is given by a D dimensional

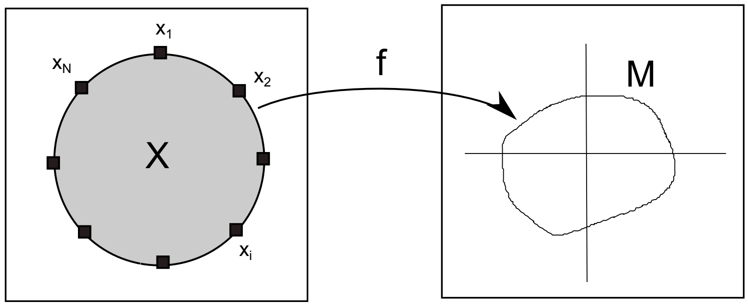


Figure 1: Mapping from high dimensional data to low dimensional Manifold

feature vector $\Phi(x_i) = (\phi_1(x), \phi_2(x), ..., \phi_D(x))$. Let M be the d-dimensional manifold extracted from the D-dimensional dataset $\mathbb{X}$ such that $d \leq D$.

In the conventional manifold learning techniques like LLE, the mapping f from $\mathbb{X}$ to M is a local isometry defined by definition 5 [4].

Definition 5. Local Isometry
The mapping f is locally isometric: For any $\varepsilon > 0$ and x in the domain of f, let $N_\varepsilon(x) = \{y : \|y - x\|_2 < \varepsilon\}$ denote an ε-neighbourhood of x using Euclidean distance. We have

$$\|f(x) - f(x_0)\|_2 = \|x - x_0\|_2 + o(\|x - x_0\|)$$

for any x_0 in the domain of f and $x \in N_\varepsilon(x_0)$.

The above definition indicates that in a local sense, f preserves the Euclidean distance. Thus, the nearby points in the high dimensional space remain nearby and similarly co-located with respect to one another in the low dimensional manifold.

In addition to the conventional mapping from $\mathbb{X}$ to M, which is a spatial isometry, we focus on a descriptive form of isometry introduced in [11] and elaborated in [13]. The definition of descriptive isometry is given below.

Definition 6. Descriptive Isometry
Let (X, δ_Φ, d_X), (Y, δ_Φ, d_Y) be metric descriptive proximity spaces and $A \subseteq X$, $B \subseteq Y$. A mapping $f : A \longrightarrow B$ is a *descriptive isometry*, provided $d_Y(\Phi(f(x_i)), \Phi(f(x_j))) = d_X(\Phi(x_i), \Phi(x_j))$, $x_i, x_j \in A$.

Lemma 2. *Let X, M be metric descriptive proximity spaces and $f : X \to M$ is a descriptive isometry on X to M. Descriptively near sets in X remain descriptively near in M.*

Proof. Let $A, B \subset X$ be descriptively near sets. Consequently, $\Phi(a) = \Phi(b)$ for at least one pair of points $a \in A, b \in B$. From definition 6, $\Phi(f(a)) = \Phi(f(b))$. Hence, $f(A), f(B) \subset M$ are descriptively near sets in manifold M. □

Lemma 3. *Let X, M be metric descriptive proximity spaces and $f : X \to M$ is a descriptive isometry on X to M. Descriptively remote sets in X remain descriptively remote in M.*

Proof. Let $A, B \subset X$ be descriptively remote sets. Consequently, $\Phi(a) \neq \Phi(b)$ for all points $a \in A, b \in B$. From definition 6, $\Phi(f(a)) \neq \Phi(f(b))$. Hence, $f(A), f(B) \subset M$ are descriptively remote sets in manifold M. □

From lemma 2, we can obtain the following results.

Theorem 1. *Descriptively near objects in an image have descriptively near manifolds.*

Proof. Let X and Y be descriptively near objects in an image endowed with a proximity relation δ_Φ. Then $\Phi(x) = \Phi(y)$ for at least one pair of points $x \in X, y \in Y$. Assume X maps to a manifold M_x and Y maps to a manifold M_y. Then, from lemma 2, $\Phi(x) = \Phi(y) \Rightarrow \Phi(m_x) = \Phi(m_y)$, where x maps to m_x and y maps to m_y. Hence, M_x is descriptively near M_y. □

Theorem 2. *Descriptively near digital images have descriptively near manifolds.*

Proof. Let A and B be descriptively near images endowed with a proximity relation δ_Φ. Consequently, $X\delta_\Phi Y$ for at least one pair of objects $X \in A, Y \in B$. Assume X maps to a manifold M_x and Y maps to a manifold M_y. Hence, M_x is descriptively near M_y. □

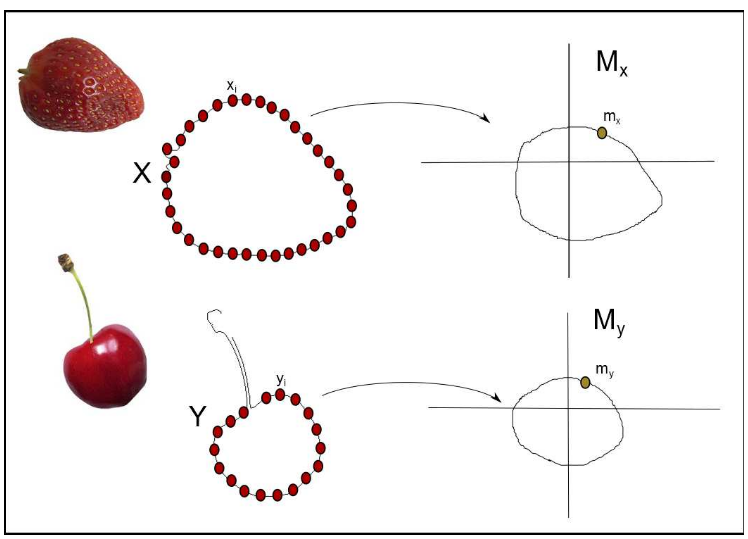
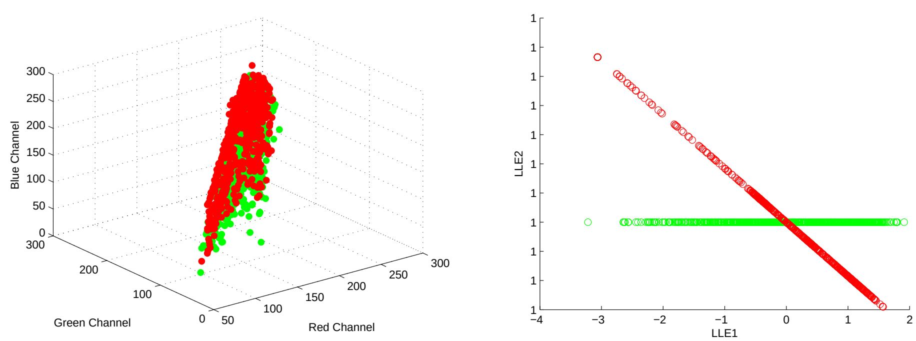


Figure 2: Descriptively Near Manifolds

Example 1. In figure 2, X and Y are two data sets obtained through uniform sampling of edge maps of two objects (a strawberry and a cherry) in an image. Each sample in X and Y is described by a feature vector $\Phi = \{R, G, B\}$, where R, G, B refers to red, green and blue colour channel values of the point respectively. Assuming X and Y are mapped to two dimensional manifolds M_x and M_y respectively. Given that point x_i maps to point $m_x \in M_x$ and point y_i maps to point $m_y \in M_y$. For any pair of points $x_i \in X$ and $y_i \in Y$, if the description of x_i is similar to description of y_i, then m_x and m_y also have similar descriptions.

In this example, it is essential that the same feature vector is selected to describe points in X and Y. However, the feature vector of x_i is not necessarily be the same as the feature vector of m_x.

Figure 3 illustrates the results of LLE algorithm for the two images. The 8 nearest neighbours were considered in obtaining the 2-dimensional manifold using LLE.



3.1: High Dimensional Data Space

3.2: 2-D Manifolds

Figure 3: Descriptively near objects. Data and manifolds representing Strawberry and Cherry images are given in green and red respectively.

From lemma 3, we can obtain the following results.

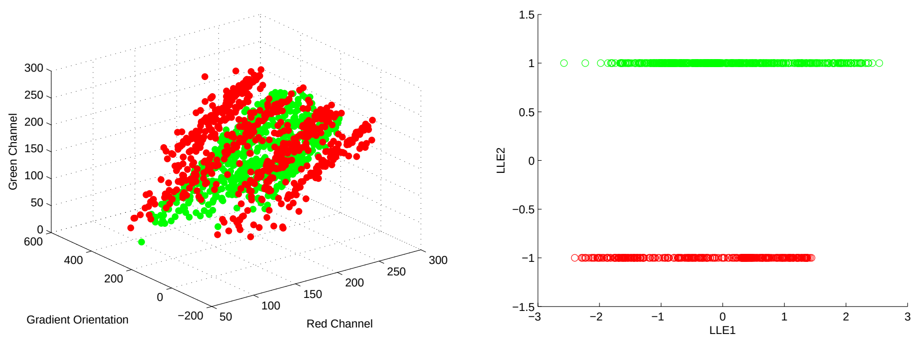
Theorem 3. *Descriptively remote objects in an image have descriptively remote manifolds.*

Proof. Let X and Y be descriptively remote objects in an image endowed with a proximity relation $\underline{\delta}_\Phi$. Then $\mathcal{Q}(clX) \cap \mathcal{Q}(clY) = \emptyset$. Assume X maps to a manifold M_x and Y maps to a manifold M_y. Then, from lemma 3, $\mathcal{Q}(clM_x) \cap \mathcal{Q}(clM_y) = \emptyset$. Hence, M_x is descriptively remote M_y. $\square$

Theorem 4. *Descriptively remote digital images have descriptively remote manifolds.*

Proof. Let A and B be descriptively remote images endowed with a proximity relation $\underline{\delta}_{\Phi}$. Consequently, $X\underline{\delta}_{\Phi}Y$ for all objects $X \in A, Y \in B$. Assume X maps to a manifold M_x and Y maps to a manifold M_y. Hence, M_x is descriptively remote M_y. □

Example 2. We take the same objects X and Y introduced in example 1. In this example, we change the feature set used to describe points in X and Y to $\Phi = \{R, G, B, Gdir\}$, where R, G and B refer to red, green and blue colour channel values and Gdir refers to gradient orientation of the point respectively. Thus, this example represents 4-dimensional data space. By incorporating gradient orientation along the edges, in this example we take the shapes of objects X and Y in to consideration. Thus, the objects X and Y are no longer descriptively near each other. The high dimensional data space and 2-dimensional manifolds extracted through LLE algorithm is illustrated in figure 4. The 8 nearest neighbours were considered in obtaining the 2-dimensional manifold using LLE. By observing figure 4, it is evident that the resulting manifolds are no longer descriptively near each other.

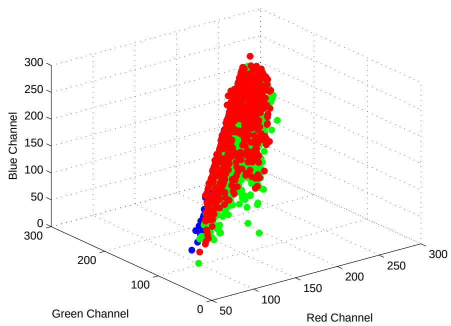


4.1: High Dimensional Data Space

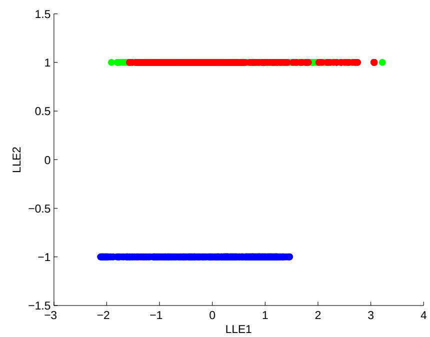
4.2: 2-D Manifolds

Figure 4: Descriptively remote objects. Data and manifolds representing Strawberry and Cherry images are given in green and red respectively.

Example 3. In this example we consider three objects, Strawberry, Cherry and Pear. We select the feature vector $\Phi = \{R, G, B\}$, where R, G and B refer to red, green and blue colour channel values. When considering R, G, B only, Strawberry and Cherry can be taken as descriptively near each other, while Pear can be considered descriptively remote to both Strawberry and Cherry. The 3-dimensional data space and 2-dimensional manifolds extracted through LLE algorithm is illustrated in figure 5. The 8 nearest neighbours were considered in obtaining the 2-dimensional manifold using LLE. By observing figure 4, it is evident that the resulting manifolds of Strawberry and Cherry are descriptively near each other, while manifold of Pear is descriptively remote to both of them.



5.1: High Dimensional Data Space



5.2: 2-D Manifolds

Figure 5: Descriptively near and descriptively remote objects. Data and manifolds representing Strawberry, Cherry and Pear images are given in green, red and blue, respectively.

4 Comparing Manifolds

As given in [7], when comparing manifolds mapped with the conventional spatial isometry, manifold distances are derived by matching spanning sets. For examples, see the tangent distance [16] and the joint manifold distance [3]. The assumption made here is that comparable manifolds have equal dimensions.

The concept of descriptively near manifolds can be used to solve classification problems. In such applications, it is important to define a metric for the comparison of two manifolds. We use a descriptive form of distance as the metric for manifold comparison.

Based on the definition of descriptive distance given in [11], the descriptive distance D_Φ between manifolds M_1 and M_2 can be defined as:

Definition 7. Descriptive Distance

$$D_\Phi(M_1, M_2) = inf\{d(\Phi(m_1), \Phi(m_2)) : m_1 \in M_1, m_2 \in M_2\}.$$

Theorem 5. *Let A and B be a pair of images and let A be mapped to a d-dimensional manifold M_1 and let B be mapped to a d-dimensional manifold M_2. M_1 is descriptively near M_2 if the descriptive distance $D_\Phi(M_1, M_2) < \varepsilon$, where ε is a threshold.*

Proof. If M_1 is descriptively near M_2, then $\Phi(m_1) = \Phi(m_2)$ for at least on pair of points $m_1 \in M_1$ and $m_2 \in M_2$. Thus, it satisfies the condition $D_\Phi(M_1, M_2) < \varepsilon$. □

5 Conclusion

The descriptively near manifolds introduced in this paper can be used in solving image pattern recognition and image classification problems. The descriptive nearness of low-dimensional manifolds arising from high-dimensional image data reduces the complexity of such problems and becomes an efficient and effective tool in image analysis.

References

[1] A. D. Concilio, Proximity: A powerful tool in extension theory, function spaces, hyperspaces, boolean algebras and point-free geometry, Contemporary Math. 486 (2009) 89–114.

[2] V. Efremovič, The geometry of proximity I (in Russian), Mat. Sb. (N.S.) 31(73) (1) (1952) 189–200.

[3] A. Fitzgibbon, A. Zisserman, Joint manifold distance: A new approach to appearance based clustering, in: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, vol. 1, IEEE, 2003.

[4] X. Huo, A.K.Smith, Performance analysis of a manifold learning algorithm in dimension reduction, Technical Paper, Statistics in Georgia Tech, Georgia Institute of Technology.

[5] J. Lee, Introduction to Topological Manifolds, vol. 202, Springer, 2011, dOI: 10.1007/978-1-4419-7940-7.

[6] T. Lin, H. Zha, Riemannian manifold learning, Pattern Analysis and Machine Intelligence, IEEE Transactions on 30 (5) (2008) 796–809.

[7] Y. Lui, J. Beveridge, Grassmann registration manifolds for face recognition, in: Computer Vision–ECCV 2008, Springer, 2008, pp. 44–57.

[8] L. Nicolaescu, Lectures on the Geometry of Manifolds, World Scientific, 2007.

[9] J. Peters, Near sets: An introduction, Math. in Comp. Sci. 7 (1) (2013) 3–9, dOI 10.1007/s11786-013-0149-6.

[10] J. Peters, Proximal relator spaces, FILOMAT (2014) 1–5*, in press*.

[11] J. Peters, Topology of Digital Images. Visual Pattern Discovery in Proximity Spaces, Springer, Berlin, 2014.

[12] J. Peters, R. Hettiarachchi, Proximal manifold learning via descriptive neighbourhood selection, Applied Mathematical Sciences 8 (71) (2014) 3513–3517.

[13] J. Peters, E. İnan, M. Öztürk, Spatial and descriptive isometries in proximity spaces, General Mathematics Notes 31 (73) (2014) 543–574.

[14] S. Roweis, K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.

[15] K. Saul, S. Roweis, Think globally, fit locally: unsupervised learning of low dimensional manifolds, The Journal of Machine Learning Research 4 (2003) 119–155.

[16] P. Simard, Y. LeCun, J. Denker, Efficient pattern recognition using a new transformation distance, in: Advances in Neural Information Processing Systems 5,[NIPS Conference], Morgan Kaufmann Publishers Inc., 1992.

[17] M. Spivak, A comprehensive Introduction to Differential Geometry, Volume I, vol. 1, 3rd ed., Publish or perish, Berkeley, 1979.

[18] J. Tenenbaum, V. D. Silva, J. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.

[19] S. You, H. Ma, Manifold topological multi-resolution analysis method, Pattern Recognition 44 (8) (2011) 1629–1648.