

Hybrid structure for robust dimensionality reduction

Xiaoqiang Lu, Yuan Yuan*

The Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, Shaanxi, PR China

ARTICLE INFO

Article history:

Received 22 November 2012
Received in revised form
1 July 2013
Accepted 26 July 2013
Communicated by Xiaofei He
Available online 31 August 2013

Keywords:

Dimensionality reduction
Pattern recognition
Manifold learning
Unsupervised learning
Face recognition

ABSTRACT

In recent years, dimensionality reduction has attracted a great deal of attention in the communities of machine learning and data mining. The basic goal of dimensionality reduction is to discover the low dimensional manifold embedded in a high dimensional space. Although some existing manifold learning algorithms (ISOMAP, LE, LLE, LTSA, etc.) can capture the local structure of data manifold, they have poor performance in some recognition tasks. This is mainly because that they cannot handle well with the “out of sample” problem. Moreover, these algorithms are sensitive to the choice of nearest neighbors, which is crucial in classification. To address these problems, this paper proposes a *Robust Dimensionality Reduction Algorithm With Local and Global Structure* (RLGS) based on a novel adaptive weighting mechanism. Hybrid structure of local and global structures is studied. By using the adaptive weight, RLGS has the capacity of adaptively exploiting non-linear structure of data manifold and is robust to parameters. Experiments demonstrate that RLGS performs better on public face databases compared with other reported algorithms.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Dimensionality reduction [23,24,17,22] has attracted considerable attention in a number of fields, including machine learning, pattern recognition, data compression and neural computation. In particular, matrix completion using low-rank matrix factorization has attracted considerable attention in the area of Dimensionality reduction [35]. In these fields, data analysis with a large number of features becomes prevalent. These features usually need to be represented as the points in high dimensional space. One may encounter the well-known problem of *curse of dimensionality* when dealing with the data in such high dimensional space. In order to obtain a compact and meaningful representation in low dimensional space, dimensionality reduction algorithms are applied to various data analysis. It aims to find the transformation between high dimensional data and its corresponding low dimensional data. In the ideal case, the dimensionality of reduced representation should be the same with the intrinsic dimensionality of the data.

Mathematically, dimensionality reduction can be described as follows. Given a $m \times n$ data matrix $\mathbf{X}$, which consists of n data vectors $\mathbf{x}_i$, with dimensionality m . Suppose the data matrix $\mathbf{X}$ has the intrinsic dimensionality d . The goal of DR is to find a mapping $\mathcal{L}: R^m \rightarrow R^d$ such that $\mathbf{y}_i = \mathcal{L}(\mathbf{x}_i)$, $i = 1, 2, \dots, n$, where $\mathbf{y}_i \in R^d$ corresponds to the data point $\mathbf{x}_i$ in high dimensional space.

In the past few years, numerous unsupervised dimensionality reduction algorithms have been developed [17,10,11,34]. In unsupervised dimensionality reduction, one hope that the low dimensional representation $\mathbf{Y}$ can capture as much information as possible in high dimensional data $\mathbf{X}$ [33]. There are some typical unsupervised methods such as *Principal Component Analysis* (PCA) [9,19,31,32], *Independent Component Analysis* (ICA) [27] and *Multi-dimensional Scaling* (MDS) [2] for dimensionality reduction. These methods are not complicated to implement, however, they cannot discover the non-linear structure of data manifold due to the linearity. Recently, manifold learning has attracted considerable attention. Some well known algorithms like *ISOMAP* [14], *Locally Linear Embedding* (LLE) [10], *Laplacian Eigenmaps* (LE) [7,17], *Local Tangent Space Alignment* (LTSA) [11], etc. have been proposed for exploiting the non-linear structure of data manifold. ISOMAP aims to preserve the local structure by preserving geodesic distance between any two points $\mathbf{x}_i$ and $\mathbf{x}_j$. LLE claims that the local structure of the manifold around a datapoint $\mathbf{x}_i$ can be represented by linear combination of its k nearest neighbors $\mathbf{x}_{i_j}$. LE attempts to keep pairwise distance between a datapoint and its k nearest neighbor by a weighted manner, which supposes that the distance in low dimensional representation $\mathbf{x}_i$ and its first nearest neighbor make a greater contribution than the distance between $\mathbf{x}_i$ and its second nearest neighbor [18]. LTSA tries to preserve local structure of data manifold through aligning a collection of local tangent subspaces onto a global coordinate system. Nevertheless, these methods cannot provide a explicit map to new testing data for recognition tasks. This is also known as the *out of sample* problem. To overcome the dilemma, He and Niyogi [16] proposed a novel

* Corresponding author. Tel.: +86 29 88889302.

E-mail addresses: luxq666666@opt.ac.cn (X. Lu), yuan369@hotmail.com, yuan@opt.ac.cn (Y. Yuan).

subspace method named *Locality Preserving Projections* (LPP), which can map the new datapoints into subspace directly. As a linear approximation to LE, LPP is the first linear algorithm that can discover the non-linear structure of data manifold. In addition, as the linear approximations to LLE and LTSA, *Neighborhood Preserving Embedding* (NPE) [5,12] and *Linear Local Tangent Space Alignment* (LLTSA) [13] are proposed to preserve the structure of data manifold from the points of locality. Recently, Zhu et al. presented a correlation analysis algorithm, called *Mixed Kernel CCA* (MKCCA), in the reproducing kernel Hilbert space (RKHS) with a mixture of kernels model to implement dimensionality reduction [29]. In [30], a novel *Self-Taught Dimensionality Reduction* (STDR) method was proposed for dealing with the high-dimensional and small-sized data.

Although manifold learning algorithm has been successfully applied in many fields, these local techniques still have some shortcomings. One of the problems is that these techniques explicitly fail to take into account the local structure in an adaptive manner. In other words, they are not robust to the choice of the nearest neighbors. In order to overcome this problem, this paper proposes a *Robust Dimensionality Reduction Algorithm With Local And Global Structure* (RLGS). In RLGS, a weight method pairwise similarity is introduced to adaptively capture local structure of data manifold. The introduced weight can preserve the structure of original data in the low-dimensional representation by adaptively learning the relation of different data points in original data space. The advantage of the proposed RLGS algorithm is demonstrated by comparing its performance with that of other dimensionality reduction algorithms over three widely used face databases.

The remainder of this paper is organized as follows. In Section 2, the proposed RLGS algorithm is described concretely. Section 3 demonstrates the robustness and performance of RLGS through recognition experiments. Finally, the conclusions are given in Section 4.

2. RLGS: robust dimensionality reduction with local and global structure

Some notations are reported first. The data set is given by $X = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{m \times n}$, correspondingly, $F \in \mathbb{R}^{n \times d}$ denotes the low-dimensional representation. W is the projection matrix, D is a diagonal matrix with the diagonal elements $D_{ii} = \sum_j P_{ij}$, $\mathbf{1} \in \mathbb{R}^{n \times 1}$ is a vector with all elements as 1, b is the vector with $d \times 1$, I denotes the identity matrix, and F_i denotes the i th row of F .

For a sample from the ‘outside’, traditional manifold learning algorithms like LLE, LE and LTSA, etc. cannot give an explicit map to the low dimensional subspace, which leads to the traditional manifold learning algorithms difficult to use in recognition tasks directly. However, LPP, NPE, LLTSA, as the linearized version to LE, LLE, and LTSA, respectively, they can effectively address the *out of sample* problem by explicit map, thus, they can yield a good performance in some recognition tasks. However, the choice of the nearest neighbors has a great influence on the performance of recognition. Motivated by the previous works [1,3,25]. This paper proposes an algorithm, named *Robust Dimensionality Reduction With Local And Global Structure* (RLGS). In RLGS, a weighting method is introduced to adaptively capture the local structure of the data manifold. The objective function is defined as follows:

$$\mathcal{Q} = \frac{1}{2} \sum_{i,j=1}^n \left\| \frac{F_i}{\sqrt{D_{ii}}} - \frac{F_j}{\sqrt{D_{jj}}} \right\|^2 P_{ij} + \mu \|X^T W + \mathbf{1}b^T - F\|^2 + \gamma \|W\|^2. \quad (1)$$

Our objective is optimizing the three variables together, then, the proposed objective function in Eq. (1) can be rewritten as

$$\mathcal{Q}(F^*, W^*, b^*) = \arg \min_{F, W, b, F^T L F = I} \text{Tr}(F^T P F) + \mu \|X^T W + \mathbf{1}b^T - F\|^2 + \gamma \|W\|^2, \quad (2)$$

where the weight P_{ij} is defined by

$$P_{ij} = \frac{\exp(-d_{ij}^2)}{\sum_{k \neq i} \exp(-d_{ik}^2)}, \quad (3)$$

where $d_{ij}^2 = \frac{1}{2} \|x_i - x_j\|^2$. Although the proposed objective function is almost the same as the related works in [20,28], there exists the difference in the construction of the weight P_{ij} in Eq. (1). In [20], the proposed objective function incorporates label fitness and manifold smoothness into a unified manifold learning framework for dimension reduction, while [28] performs co-clustering along both the sample and feature dimensions to capture the intersample and interfeature relationships by minimizing the fitting errors of structures of both sample and feature spaces. Compared to the works in [20,28], the introduced weight can preserve the structure of original data in the low-dimensional representation by adaptive learning the relation of different data points in original data space. It is easy to see that the weight matrix P is symmetric because $P_{ij} = P_{ji}$ for any i, j . When the high dimensional data point x_i and x_j are close in the intrinsic structure

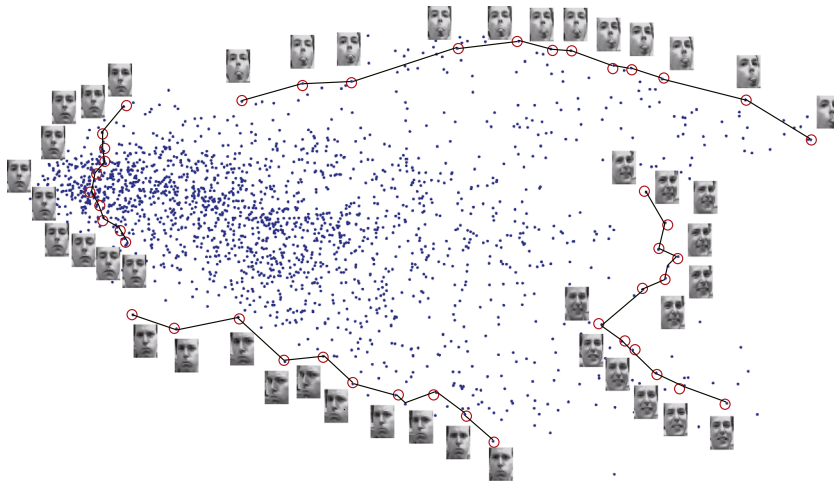


Fig. 1. 2-D visualization of Frey face images (20 × 28 pixels, 1965 examples). Images corresponding to the points on the bound lines (top, bottom, left, right).

Download English Version:

<https://daneshyari.com/en/article/6866844>

Download Persian Version:

<https://daneshyari.com/article/6866844>

[Daneshyari.com](https://daneshyari.com)