Neurocomputing 133 (2014) 427–436

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Sparse tensor embedding based multispectral face recognition ☆

Haitao Zhao [a], Shaoyuan Sun [b,*]

[a] Automation Department, East China University of Science and Technology, Shanghai, PR China
[b] Automation Department, Donghua University, Shanghai, PR China

ARTICLE INFO

Article history:
Received 8 May 2013
Received in revised form
5 December 2013
Accepted 7 December 2013
Communicated by Ran He
Available online 9 January 2014

Keywords:
Multispectral face recognition
Multibiometrics
Sparse tensor embedding

ABSTRACT

Face recognition using different imaging modalities has become an area of growing interest. A large number of multispectral face recognition algorithms/systems have been proposed in last decade. How to fuse features of different spectrum has still been a crucial problem for face recognition. To address this problem, we propose a sparse tensor embedding (STE) algorithm which represents a multispectral image as a third-order tensor. STE constructs sparse neighborhoods and the corresponding weights of the tensor. One advantage of the proposed technique is that the difficulty in selecting the size of the local neighborhood can be avoided in the manifold learning based tensor feature extraction algorithms. STE iteratively obtains one spectral space transformation matrix through preserving the sparse neighborhoods. Due to sparse representation, STE can not only keep the underlying spatial structure of multispectral images but also enhance robustness. The experiments on multispectral face databases, Equinox and PolyU-HSFD face databases, show that the performance of the proposed method outperform that of the state-of-the-art algorithms.

1. Introduction

Multibiometrics refers to the use of a combination of two or more sensor modalities in a single identification system. The reason for combining different sensor modalities is to improve the recognition accuracy. Face recognition using different imaging modalities has become an area of growing interest [1–7]. Recent studies have found that multispectral face recognition offers a few distinct advantages, including invariance to illumination changes [8,9]. Multispectral image also reveals anatomical information of a subject that is useful in detecting disguised faces [1]. In [3,10] several different face recognition algorithms were tested and good performances were achieved. The effect of lighting, facial expression, and passage of time between the gallery and probe images was examined by Chen et al. [11]. Wang et al. [12] demonstrated that color space learning represents a viable approach for enhancing face recognition performance. The image-based fusion performed in the wavelet domain and feature-based fusion performed in the eigenspace domain were present in [13]. Heo et al. [14] proposed to fuse visual and thermal signatures with eyeglass removal for robust face recognition. Multi-sensory face biometric fusion methods were investigated for personal identification [15]. Pan et al. [4,5] analyzed the facial tissue spectral measurements for multispectral face recognition in the near-infrared spectral range (0.7–1.0 μm). Denes et al. [6] adopted three single visible bands (0.6 μm, 0.7 μm, and 0.8 μm) to test the spectral asymmetry. Chang et al. [16] fused the multispectral images in the visible spectrum (0.4–0.72 μm) into a single image and compared the result with the visible image to validate the improvement of face recognition due to the image fusion. Chou and Bajcsy [7] pre-processed the multispectral images (visible: 0.4–0.72 μm, near-infrared: 0.65–1.1 μm) by principal component analysis (PCA) to extract the first principle component information for face detection. Wong and Zhao [17] adopted kernel PCA for visual information aided thermal image reconstruction.

The above methods are mainly designed to preserve the global structure information of the data. Unsupervised algorithms [7,9,17] and supervised algorithms [33] have been widely used in multispectral face recognition. However, research results from manifold learning methods developed in the past decade show that the local geometric structure is more important than the global structure since the high-dimensional data lies on the low-dimensional manifold. Due to the low-dimensional manifold distributions of the face images, the manifold-learning-based linear feature extraction methods such as Laplacianfaces [18] become popular. Recent research shows that the high-order tensor based feature extraction methods such as tensor locality preserving projection (tensor LPP) [19,20] can obtain better performance than the classical feature extraction method. Unfortunately, the tensor data contain large quantities of information redundancy and thus not all the features/variables are important to feature extraction and classification [21–23]. It was shown that integrating sparse representation, manifold learning for feature extraction

☆This research is supported by the National Natural Science Foundation of China.
* Corresponding author. Tel.: +86 13817304566.
E-mail address: shysun@dhu.edu.cn (S. Sun).

0925-2312/$ - see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.12.019

may obtain better performance [24]. Many papers show that the sparse representation methods can obtain better performance than their corresponding non-sparse methods in the real data. And these sparse methods can give an intuitionistic or semantic interpretation for the transformed features [25].

Till now, the field in high order tensor data embedding with sparse manner has not been widely investigated and how to extend the manifold learning algorithms integrating sparseness and manifold structure for multispectral face recognition is unsolved. In this paper, motivated by tensor data embedding and sparse representation, we propose a novel method called sparse tensor embedding (STE) for multispectral image feature extraction. The multispectral image is considered as a third-order tensor. The aim of STE is to obtain transformation matrices through preserving the sparse information of the third-order tensors.

The rest of the paper is organized as follows. In Section 2, we give the related definitions to tensor. In Section 3, the introduction of tensor locality preserving projection is provided. In Section 4, a novel sparse tensor embedding method is presented. Experiments are carried out to evaluate the proposed tensor learning method in Section 5, and conclusions are given in Section 6.

2. Tensor fundamentals

A tensor is a multidimensional array [12,19]. It is the higher order generalization of scalar (zero-order tensor), vector (1st-order tensor), and matrix (2nd-order tensor). In this paper, lowercase letters (i.e. a, b, c) denote scalars, bold lowercase letter (i.e. $\boldsymbol{a}$, $\boldsymbol{b}$, $\boldsymbol{c}$) denote vectors, uppercased letters (i.e. A, B, C) denote matrices and bold uppercase letters (i.e. $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$) denote the tensors. It is assume that the training samples are represented as the nth order tensor $\{\boldsymbol{A}_i \in R^{m_1\times m_2\times\cdots\times m_n},\ i=1,\ 2,\ \ldots,N\}$, where N denotes the total number of training samples.

Definition 1. The inner product of two tensors $\boldsymbol{A}$, $\boldsymbol{B}\in R^{m_1\times m_2\times\cdots\times m_n}$ is defined as $\langle \boldsymbol{A},\boldsymbol{B}\rangle=\sum_{i_1}\sum_{i_2}\cdots\sum_{i_n}\boldsymbol{A}_{i_1,\ i_2,\ldots,\ i_n}\boldsymbol{B}_{i_1,i_2,\ldots,\ i_n}$. The Frobenius norm of a tensor $\boldsymbol{A}\in R^{m_1\times m_2\times\cdots\times m_n}$ is then defined as $\|A\|=\sqrt{\langle \boldsymbol{A},\boldsymbol{A}\rangle}$. And the distance between two tensors $\boldsymbol{A}$, $\boldsymbol{B}\in R^{m_1\times m_2\times\cdots\times m_n}$ is defined as $D(\boldsymbol{A},\ \boldsymbol{B})=\|\boldsymbol{A}-\boldsymbol{B}\|$.

Definition 2. The k-mode flattening of the nth-order tensor $\boldsymbol{A}_i\in R^{m_1\times m_2\times\cdots\times m_n}$ $(i=1,\ 2,\ \ldots,N)$ into matrix $A^{(k)}\in R^{m_k\times\prod_{i\neq k}m_i}$, i.e. $A^{(k)}\Leftarrow_k\boldsymbol{A}$, is defined as $A^{(k)}_{i_k,j}=\boldsymbol{A}_{i_1,i_2,\ldots,i_n}$, $j=1+\sum_{q=1,q\neq k}^{n}(i_q-1)\prod_{p=q+1,p\neq k}^{n}m_p$.

Definition 3. The k-mode product of a tensor $\boldsymbol{A}\in R^{m_1\times m_2\times\cdots\times m_n}$ by a matrix $U\in R^{m'_k\times m_k}$, denoted by $\boldsymbol{B}=\boldsymbol{A}\times{}_kU$, is an $(m_1\times m_2\times\cdots\times m_{k-1}\times m'_k\times m_{k+1}\times\cdots\times m_n)$ tensor of which the entries are given by: $\boldsymbol{B}_{i_1,\ldots,i_{k-1},i,i_{k+1},\ldots,i_n}=\sum_{j=1}^{m_k}\boldsymbol{A}_{i_1,\ldots,i_{k-1},j,i_{k+1},\ldots,i_n}U_{i,j}$ $(j=1,\ 2,\ \ldots,m_k)$.

The aim of tensor learning algorithm is to obtain a set of projection matrices $\{U_i\in R^{d_i\times m_i},\ d_i\leq m_i,\ i=1,\ 2,\ \ldots,\ n\}$ and map the original tensor into a new tensor:

$$\boldsymbol{B}_i=\boldsymbol{A}_i\times{}_1U_1\times{}_2U_2\cdots\times{}_nU_n. \tag{1}$$

3. Tensor locality preserving projection

Similar to Laplacian Eigenmap and LPP [18], tensor locality preserving projection (tensor LPP) [19,20] provides a way to linearly approximate the eigenfunctions of the Laplace Beltrami operator in a tensor space. Therefore, it can model the geometric and topological properties of an unknown manifold embedded in a tensor space with some data points sampled from the manifold.

Let $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, …, $\boldsymbol{A}_N$ be the multispectral face images in a high order tensor form, and $\boldsymbol{A}_i\in R^{m_1\times m_2\times\cdots\times m_n}$ $(i=1,\ 2,\ldots,N)$, N is the number of individual. We first construct a neighborhood graph to represent the local geometric structure. The corresponding similarity matrix $W=[w_{ij}]_{N\times N}$ is defined based on the heat kernel as follows:

$$w_{ij}=\begin{cases}\exp\left(-\|\boldsymbol{A}_i-\boldsymbol{A}_j{}^2\|/t\right) & \text{if } \boldsymbol{A}_j\in O(K,\boldsymbol{A}_i)\ \text{or}\ \boldsymbol{A}_i\in O(K,\boldsymbol{A}_j)\\ 0 & \text{otherwise}\end{cases} \tag{2}$$

where $O(K,\boldsymbol{A}_i)$ denotes the set of K nearest neighbors of $\boldsymbol{A}_i$ and t is a positive constant.

The aim of tensor LPP is to find transformation matrices $U_1,\ U_2,\ \ldots,\ U_n$, to project high dimensional $\boldsymbol{A}_i$ into low-dimensional representation $\boldsymbol{B}_i$, where $\boldsymbol{B}_i=\boldsymbol{A}_i\times{}_1U_1\times{}_2U_2\cdots\times{}_nU_n$ and $U_j\in R^{d_j\times m_j}$ $(d_j\leq m_j,\ j=1,\ 2,\ \ldots,\ n)$.

The optimization problem for tensor LPP can be expressed as follows:

$$\begin{aligned}&\operatorname{argmin} J(U_1,\ U_2,\ldots,\ U_n)=\sum_{i,j}\|\boldsymbol{B}_i-\boldsymbol{B}_j\|^2w_{ij}\\&=\sum_{i,j}\|\boldsymbol{A}_i\times_1U_1\cdots\times_nU_n-\boldsymbol{A}_j\times_1U_1\cdots\times_nU_n\|^2\ w_{ij},\\&\text{s.t.}\ \ \sum_i\|\boldsymbol{A}_i\times_1U_1\cdots\times_nU_n\|^2d_{ii}=1.\end{aligned} \tag{3}$$

In general, the larger the value of $d_{ii}=\sum_jw_{ij}$ is, the more important is the tensor $\boldsymbol{B}_i$ in the embedded tensor space for representing the original tensor $\boldsymbol{A}_i$. It is easy to see that the objective function will give a high penalty if neighboring tensors $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ are mapped far apart. Thus, if two tensors $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ are close to each other, then the corresponding tensors $\boldsymbol{B}_i$ and $\boldsymbol{B}_j$ in the embedded tensor space are also expected to be close to each other.

Similar to other tensor learning methods, this optimization problem can be solved by applying an iterative scheme. Assuming that $U_1,U_2,\ \ldots,\ U_{k-1},\ U_{k+1},\ldots,\ U_n$ are known, let $\boldsymbol{B}_i^k=\boldsymbol{A}_i\times{}_1U_1\cdots\times_{k-1}U_{k-1}\times{}_{k+1}U_{k+1}\cdots\times{}_nU_n$. In addition, since $B_i^{(k)}\Leftarrow_k\boldsymbol{B}_i^k$ and based on the properties of tensor and trace, we reformulate the optimization function in (3) as follows:

$$\begin{aligned}\operatorname{argmin} J_k(U_k)&=\sum_{i,j}\|\boldsymbol{B}_i^k\times_kU_k-\boldsymbol{B}_j^k\times_kU_k\|^2w_{ij}\\&=\sum_{i,j}\|U_kB_i^{(k)}-U_kB_j^{(k)}\|^2w_{ij}\\&=\sum_{i,j}\mathrm{tr}\{U_k((B_i^{(k)}-B_j^{(k)})(B_i^{(k)}-B_j^{(k)})^Tw_{ij})U_k^T\}\\&=\mathrm{tr}\left\{U_k\left(\sum_{i,j}(B_i^{(k)}-B_j^{(k)})(B_i^{(k)}-B_j^{(k)})^Tw_{ij}\right)U_k^T\right\},\\&\text{s.t.}\ \ \mathrm{tr}\left\{U_k\left(\sum_iB_i^{(k)}B_i^{(k)T}d_{ii}\right)U_k^T\right\}=1.\end{aligned} \tag{4}$$

The unknown transformation matrix U_k can be obtained by solving for the eigenvectors corresponding to the d_k smallest eigenvalues in the generalized eigenvalue equation

$$\left(\sum_{i,j}(B_i^{(k)}-B_j^{(k)})(B_i^{(k)}-B_j^{(k)})^Tw_{ij}\right)\boldsymbol{u}=\lambda\left(\sum_iB_i^{(k)}B_i^{(k)T}d_{ii}\right)\boldsymbol{u}. \tag{5}$$

The other transformation matrices can be obtained in a similar manner.

4. Sparse tensor embedding

Sparse representation algorithms have been widely studied in signal processing, computer vision and pattern recognition. Wright et al. [25] used sparse representation for robust face reconstruction and recognition, and Qiao et al. [26] proposed sparse preserving projections, Cheng et al. [27] used the L_1-graph for image clustering. He et al. [34,35] used nonparametric maximum entropy and L_{21} regularized correntropy for robust feature