Neurocomputing 149 (2015) 1451–1460

Contents lists available at ScienceDirect

Neurocomputing

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

Discriminant non-negative graph embedding for face recognition

Jinrong Cui [a,b,*], Jiajun Wen [a,b], Zhengming Li [c], Bin Li[a,b]

[a] Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China
[b] Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, China
[c] Guangdong Industrial training center, Guangdong Polytechnic Normal University, Guangzhou, China

ARTICLE INFO

Article history:
Received 23 January 2014
Received in revised form
23 August 2014
Accepted 23 August 2014
Communicated by X. Gao
Available online 2 September 2014

Keywords:
Non-Negative Matrix Factorization (NMF)
Graph preserve embedding
Feature extraction
Face recognition
Manifold learning
Sparse subspace

ABSTRACT

Non-negative Matrix Factorization (NMF) is an unsupervised algorithm for low-rank approximation of non-negative data and has been widely used in many fields, but its performance in feature extraction is not satisfactory. The main reason is that the model of NMF and its variants did not take into account the label information of the samples, which can add the discriminant ability of the methods. In this paper, we proposed a novel method, called discriminant non-negative graph embedding (DNGE) algorithm in which the label information of the samples and the local geometric structure are all integrated in the objective function. Furthermore, we incorporated the between-class graph and within-class graph into the objective functions to indicate that we not only used the local separability but also used the whole separability of the samples. To guarantee convergence, we use the KKT condition to calculate the non-negative solution of the DNGE. A convergent multiplicative non-negative updating rule is then derived to learn the transformation matrix. Experiments are conducted on the CMU PIE, ORL, Yale, FERET and AR database. The results show that the DNGE algorithm provides better facial representation and achieves higher recognition rates than naïve Non-Negative Matrix Factorization and its extension methods.

1. Introduction

IN recent years, Non-negative Matrix Factorization (NMF) [1] algorithm has become popular in data representation. NMF decomposes the data matrix as a product of two matrices which possess only non-negative elements. Compared with the other traditional matrix factorization algorithm, such as Singular Value Decomposition (SVD) [2], NMF can guarantee the factorization matrices' non-negativity, which makes NMF competitive in practice since most of the data in practice are non-negative. The non-negative constraints of NMF lead to a part-based representation because they only allow additive, not subtractive, combinations [4]. Lee and Seung [3] showed that NMF can learn a parts-based representation and the basis images of the face image consist of basis vectors representing eyes, mouths, nose and contour of face.

NMF have been successfully used for face recognition [5–8] and document clustering [9], where it is natural to consider the object as a combination of parts to form a whole. However, NMF did not take into account the geometrical structure of the data. In fact, the geometrical structure has played an important role in classification and clustering. In [4], Yuan et al. proposed a Projective Non-negative Matrix Factorization (PNMF) approach which was a new variant of the NMF method for learning spatially localized, sparse, part-based subspace representations of visual patterns.

In order to detect the underlying manifold structure, many manifold learning algorithms have been proposed, such as Locally Linear Embedding (LLE) [10], ISOMAP [27], Laplacian Eigenmap [28] and Locality Preserving Projection (LPP) [22]. Many experiments showed that the exploitation of the geometrical structure and the local invariance can improve the classification performance [11–13,34–35]. Motivated by the manifold learning method which use the graph modeling for the local geometric structure, Cai et al. [9] proposed a Graph regularized NMF (GNMF) approach to encode the geometrical information of the data space. In GNMF, a nearest neighbor graph is integrated into NMF to model the local manifold structure. In order to preserve the local coordinate structure in NMF, Chen et al. [6] proposed a Non-negative Local Coordinate Factorization (NLCF) method which adds a local coordinate constraint into the standard NMF objective function.

Based on the graph structure and sparse representation, many variants of NMF have been proposed. However, most of them are unsupervised and fail to discover the discriminant structure and non-negative representation in the data. The semi-supervised matrix decomposition method is also proposed and has the following rationale: many machine learning researchers have found that unlabeled data, when used in conjunction with a small

* Corresponding author.
E-mail address: tweety1028@163.com (J.-n. Cui).

http://dx.doi.org/10.1016/j.neucom.2014.08.047
0925-2312/

amount of labeled data, can produce considerable improvement in classification accuracy [14]. For semi-supervised manner, Liu et al. [15] proposed a Constrained Non-negative Matrix Factorization approach called CNMF. CNMF takes the label information as additional hard constraint, and the central idea of CNMF is that the data points from the same class should be merged together in the new representation space. Thus the obtained part-based representation of CNMF has the consistent label with the original data, and therefore can have more discriminant power.

In order to improve the discriminant ability of original NMF, some studies expand NMF to supervised matrix factorization approaches by introducing the label information. For example, the Discriminant Non-negative Matrix Factorization (DNMF) [16–18] imposes more constraint on the coefficients in order to take into account the class information. Liu et al. proposed the Projective Non-negative Graph Embedding (PNGE) [19] algorithm which is a general formulation for non-negative data factorization. In PNGE, the non-negative coefficient vector of each data is assumed to be projected from its original feature representation with a universal non-negative transformation matrix. An et al. proposed a manifold-respecting discriminant Non-negative Matrix Factorization in [33], which construct the intra-class neighborhoods graph and inter-class graph into the objective function. However, these methods did not incorporate the label information in the model and only used the label information in local with-class and inter-class scatter. In the first part of the objective function, the above mentioned NMF based methods use the form of $||X-UV||_F^2$, which cannot excavate any discriminant information in this term. Thus these methods did not simultaneously take advantage of the label information and the manifold structure.

Previous NMF-based methods such as GNMF, PNMF and PNGE did not sufficiently utilize the discriminant information since the part of $\|X-UV\|_F^2$ does not embedding any discriminant knowledge. Motivated by recent development in matrix factorization and manifold learning, in this paper, we develop a novel discriminant non-negative graph preserving (DNGE) algorithm for face recognition. In the first part of the objective function of DNGE, we use the form of $\|Y-XA\|_2^2$ and this term can obtain discriminant information since the label matrix Y is used. On the other side, for graph embedding, we construct the local between-class graph and local within-class graph for preserving the local information of the data. As shown in [29], the local within-class graph can assemble the data point with its neighborhood within the same class, and the between-class graph can exclude the data point without the same class. Our goal is to use the label information of the original data to find a parts-based representation space in which the data that come from the same class are near to each other and from different classes are far from each other. To this end, we propose a new NMF-based objective function which incorporates the discriminant graph structure simultaneously. In this new objective function, the local separability was introduced to approximate the global separability. We also develop an optimization scheme to solve the objective function based on KKT conditions. The convergence proof of our proposed optimization scheme is provided.

It is worthwhile to highlight several aspects of our proposed algorithm: (1) in our proposed DNGE algorithm we used two graphs to preserve the manifold structure of the original data. One local within-class graph is designed to characterize intra-class compactness, and the other local between-class graph is formulated for achieving interclass separability. In DNGE, the data points are mapped into a subspace in which the nearby points with the same label are close to each other while the nearby points with different labels are far apart. In this way, our proposed algorithm can obtain more discriminant information for classification. (2) We directly incorporate the label of the data for non-negative factorization, and we integrate the local separability and the global separability together. (3) A multiplicative non-negative updating rule is then derived to learn the non-negative transformation matrix and the updating rule is proved to be convergent.

The rest of the paper is organized as follows. In Section 2, we briefly review NMF, GNMF, and PNGE. The proposed DNGE algorithm and related analyses are described in Section 3. In Section 4, experiments are carried out to evaluate our DNGE algorithm. The conclusions are given in Section 5.

2. Non-negative Matrix Factorization and its variants

2.1. Non-negative Matrix Factorization (NMF) [20]

Given a data matrix $X=[x_1,x_2,\ldots,x_n]\in R^{m\times n}$, each column of X is a sample vector. NMF aims to find two non-negative matrices $U=[u_{ij}]\in R^{m\times k}$ and $V=[v_{ij}]\in R^{k\times n}$ by minimizing the following objective function:

$$E(U,V)=\|X-UV\|_F^2 s.t. U_{ij}\ge 0,\quad V_{ij}\ge 0 \tag{1}$$

where $\|\cdot\|_F^2$ denotes the matrix Frobenius norm.

Although the objective function $E(U,V)$ in Eq. (1) is convex only in U or V, it is not convex in both variables together. Therefore, it is unrealistic to expect an algorithm to find the global minimum of $E(U,V)$. Lee and Seung [3] presented an iterative updating algorithm to solve this problem. Two key iteration steps are shown in (2) and (3).

$$U_{ij}\leftarrow U_{ij}\frac{(XV^T)_{ij}}{(UVV^T)_{ij}} \tag{2}$$

$$V_{ij}\leftarrow V_{ij}\frac{(X^TU)_{ij}}{(V^TU^TU)_{ij}} \tag{3}$$

It is proved that the above update steps will find a local minimum of the objective function $E(U,V)$ [3]. The matrix U contains a basis that is optimized for the linear approximation of the data X, and the matrix V can be regarded as the low-dimensional representation. Thus the learned transformation matrix V can be directly used to obtain the encoding coefficient vectors of the new test samples.

2.2. Graph Regularized Non-negative Matrix Factorization (GNMF) [9]

The goal of GNMF [9] is to find a compact representation which uncovers the hidden semantics and simultaneously represents the intrinsic geometric structure. In GNMF, an affinity graph is constructed to encode the geometrical information, and thus it seeks a matrix factorization which represents the graph structure.

In GNMF, the geometrically based regularization is as follows:

$$R_1=\frac{1}{2}\sum_{i,j=1}^{N}\|z_i-z_j\|^2W_{ij}$$
$$=Tr(VDV^T)-Tr(VWV^T)=Tr(VLV^T) \tag{4}$$

where $Tr(\cdot)$ denotes the trace of a matrix and D is a diagonal matrix whose entries are column (or row, since W is symmetric) sums of W $D_{ii}=\sum_j W_{ij}$. $L=D-W$, which is called graph Laplacian [28].

Combining the geometrical based regularization with the original NMF leads to the objective function of GNMF, we have

$$E(U,V)=\|X-UV\|^2+\lambda Tr(VLV^T) \tag{5}$$