Multi-view Discriminative Manifold Embedding for Pattern Classification

X. Wang
Departmen of Information
Zhenghzou 450053, China

Y. Guo
Department of Digestive
Zhengzhou 450053, China

Z. Wang
Henan University of Technology
Zhengzhou 450001, China
13839527381@163.com

ABSTRACT: *While many dimensionality reduction algorithms have been proposed in recent years, most of them are designed for single view data and cannot cope with multi-view data directly. Dimensionality reduction algorithms in recent ten years, both in theory and application have great breakthrough. In the face of dozens, hundreds or even thousands of dimension by dimension reduction to the data from high dimensional space to a low dimensional space and extract the essential characteristics of low dimensional data. In many real-world pattern applications such as face recognition, multiple feature descriptors can provide complementary information in characterizing image from different viewpoints. Motivated by this concern, we propose a new multi-view discriminative manifold embedding (MDME) method for classification by making use of intra-class geometry and inter-class marginal information as well as complementary information of multiple feature representations. Experimental results on face recognition demonstrate the effectiveness of the proposed algorithm.*

Keywords: Discriminative manifold embedding, Multi-view learning, Dimensionality reduction, Pattern classification

Received: 2 February 2017, Revised 18 March 2017, Accepted 24 April 2017

1. Introduction

Dimensionality reduction is a fundamental problem in many machine learning and pattern recognition applications [1]. Recent researches show that higher dimensional data (such as face image) lie on a low-dimensional nonlinear manifold. To fully discover the manifold structure of high-dimensional data, the most representative manifold learning algorithm is a locality preserving projection (LPP) [2]. Although LPP can preserve the local geometry of samples with a weighted adjacency graph, it is unsupervised and ignores the interactions of samples from different classes. To utilize discriminant information for pattern classification, some supervised manifold learning algorithms have been developed. Despite the various motivations of different

manifold learning algorithms, they can be unified into a general graph embedding framework (GEF) [3] with different constraints.

Data dimension reduction of pattern recognition of face recognition has been widely used. Because usually face image data of high dimension will not only cause the problem of "dimension disaster", but also can make people found it difficult to structure of the underlying data set information. If the data dimension reduction preprocessing, on one hand, can overcome the "dimension disaster" problem; on the other hand can greatly reduce the computational complexity and noise. So data dimension reduction methods now attract the attention of the researchers.

2. Related Work and Background

Most existing dimensionality reduction algorithms assume that the data are represented in a single view. In many practical applications such as face recognition, multiple feature descriptors can provide complementary information in characterizing image information from different viewpoints. Therefore, existing single-view-based methods cannot directly cope with data described with multiple views. Recent studies have shown that leveraging information contained in multiple views potentially has a discriminating advantage over only a single view. Hence, it is more reasonable to incorporate multi-view learning into dimensionality reduction so that the complementary property of different views can be fully exploited for low-dimensional embedding. To achieve this goal, there have been some attempts which include distributed multi-view subspace learning (DMSL), multi-view stochastic neighbor embedding (MSNE), and multi-view spectral embedding (MSE) [4].

Although DMSL and MSNE can obtain low-dimensional embedding from multi-view high-dimensional data, they are unsupervised and classification abilities may be limited, since the class label information is not used in the learning process [5]. While MSE can solve the multi-view subspace learning problem by using the "patch alignment framework", it ignores the flexibility of allowing shared information between different views owing to the global coordinate alignment process [6]. More recently, multi-view subspace learning has been successfully applied to image annotation, cartoon retrieval and synthesis. By jointly considering the local manifold structure and the class label information, we propose a new method, called multi-view discriminative manifold embedding (DME) [7]. Experimental results on face recognition are presented to demonstrate the effectiveness of the proposed approach [8].

The human face is one of the most common mode of people's vision, face recognition have attracted much attention because of its non-contact, safe, natural, intuitive and convenient characteristics and has become one of the biometric identification technology currently. Face recognition is currently a hot research topic in the field of artificial intelligence and pattern recognition [9]. Extracting effective features (data dimension reduction) plays a key role in the pattern classification of high dimensional data. According to the biometric recognition of the human face, the dimension reduction algorithm directly affects the recognition rate of face recognition [10].

3. Proposed Approaches

In face recognition, generally use a high dimensional vector to represent an image, but because of the face recognition feature dimension is too high and not stable, dimension reduction is the most common method to solve this problem. The extraction of effective features is the key to reduce the impact of loss of information on the recognition rate [11,12]. Dimensionality reduction (feature extraction) simply by mapping function (including nonlinear mapping and linear mapping) mapping high-dimensional data into low dimensional space, reduce the data dimension, reduce the time complexity of the algorithm and extract the essential features of the data, for data classification and visualization analysis etc. So far, many linear and nonlinear dimensionality reduction methods have been proposed.

Recent studies show that although the different data dimensionality reduction algorithms based on different theories and methods, they also have some contact. So expect data dimensionality reduction methods into a unified framework to discuss and research [13, 14, 15], through this framework to a certain extent, can promote dimension reduction algorithm.

Given a multi-view data set consisting of n samples with m different features is denoted as a set of matrices $X = \{X^{(i)} \in R^{pi\times n}\}_{i=1}^{m}$, wherein $X^{(i)}$ is the feature matrix of the ith view representation, p_i is the dimensionality of the ith view $X^{(i)}$. The aim of MDME is to seek a low-dimensional and sufficient smooth embedding $Y^{(i)} \in R^{d\times n}$ of $X^{(i)}$ via $Y^{(i)} = V^T X^{(i)}$, wherein $d < p_i$ (i = 1, 2, ..., m).

In order to well explore the complementary information of multiple views for feature extraction, a set of nonnegative weights β

$= [\beta_1, \beta_2, ..., \beta_m]$ is imposed on the objective function of DME of each view. The larger β_i is, the more contribution of the view

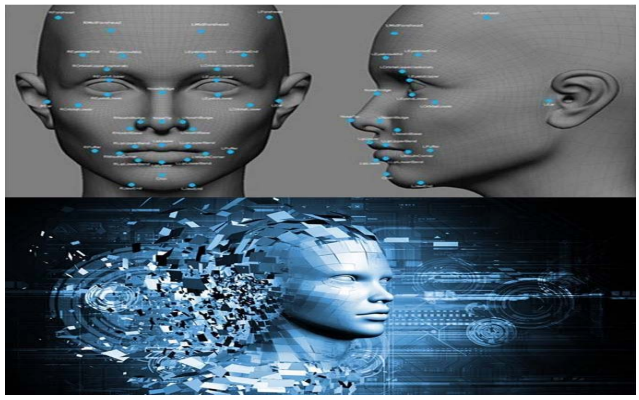

Figure 1. Face recognition

$X^{(i)}$ makes learning the low-dimensional embedding. Hence, by summing over all views, the optimal objective function of MDME can be formulated as follows.

$$\max_{V,\beta} J(V,\beta) = J_1(V,\beta) - \alpha J_3(V,\beta) = \sum_{i=1}^{m} \beta_i Tr\left(V^T H^{(i)} V\right) \tag{1}$$

With the constraint

$$V^T V = I\,,\ \sum_{i=1}^{m} \beta_i = 1,\ \beta_i \geq 0\,. \tag{2}$$

$$H^{(i)} = \left(X^{(i)} L^{(i)} X^{(i)^T} - \alpha X^{(i)} \left(I^{(i)} - R^{(i)}\right)\left(I^{(i)} - R^{(i)}\right)^T X^{(i)^T}\right) \tag{3}$$

The solution to (2) subject to (3) is $\beta_i = 1$ corresponding to the maximum $Tr\,(V^T H^{(i)} V)$ over different views, and $\beta_i = 0$ otherwise. This means that only the best view is finally selected by our method, thus resulting in failing to exploit the complementary information from different views for dimensionality reduction. To tackle this problem, following the trick adopted in, we modify β_i to be β_i^q with $q > 1$. Then, the new objective function is defined as follows.

$$\max_{V,\beta} \sum_{i=1}^{m} \beta_i^q Tr\left(V^T H^{(i)} V\right), \quad \text{s.t. } V^T V = I, \sum_{i=1}^{m} \beta_i = 1, \beta_i \geq 0. \tag{4}$$

Since (7) is a nonlinearly constrained nonconvex optimization problem which has no closed-form solution, we derive an alternating optimization-based iterative algorithm to obtain local optimal solution.

First, we fix V and update β. By utilizing the Lagrange multiplier λ to incorporate the constraint into objective function, we obtain the following Lagrange function:

$$L(\beta,\lambda) = \sum_{i=1}^{m} \beta_i^q Tr\left(V^T H^{(i)} V\right) - \lambda\left(\sum_{i=1}^{m} \beta_i - 1\right) \tag{5}$$

Let $\partial L(\beta,\lambda)/\partial \beta_i = 0$ and $\partial L(\beta,\lambda)/\partial \lambda = 0$, we have

$$\beta_i = \left(1/Tr\left(V^T H^{(i)} V\right)\right)^{1/q-1} \Big/ \sum_{i=1}^{m} \left(1/Tr\left(V^T H^{(i)} V\right)\right)^{1/q-1} \tag{6}$$

Then, we update V by using the obtained β. The optimal problem (7) is equivalent to

$$\max_V \sum_{i=1}^{m} Tr\left(V^T H^{(i)} V\right), \quad \text{s.t. } V^T V = I. \tag{7}$$

and V is given by the maximum eigenvalue solution to the eigenvalue problem.

$$\sum_{i=1}^{m} H^{(i)} V = \lambda V \tag{8}$$

4. Experimental Results

In this section, we evaluate the effectiveness of our proposed MDME approach for face recognition task. We also compare the proposed algorithm with some traditional single- view- based dimensionality reduction algorithms, such as PCA, LDA, LPP, and MFA, as well as the three latest multi-view dimensionality reduction algorithms, including DMSL, MSNE, and MSE. For a fair comparison, all the results reported here are based on the best tuned parameters of all the comparison algorithms.

We have used the ORL and AR face databases in our experiments. The ORL database contains 400 images of 40 people. Each person provides 11 different images with various facial expressions and lighting conditions. The AR database contains over 4000 face images of 126 people (70 men and 56 women). For all the images in the above two databases, the facial parts of each image was manually cropped, aligned, and resized to 32×32 pixels according to eye's positions, with 256 gray levels per pixel. In order to comprehensively describe face images, we extract three kinds of low-level visual features to represent three different views, i.e., SIFT, Gabor, and LBP features. Because the three views are generally complementary to each other, we empirically set q in MDME to 5. The dimensionality of the low-dimensional subspace d is set to 20 for ORL database and 30 for AR database. Table I tabulates the rank-1 recognition rates of different methods with different features on the ORL and AR databases.

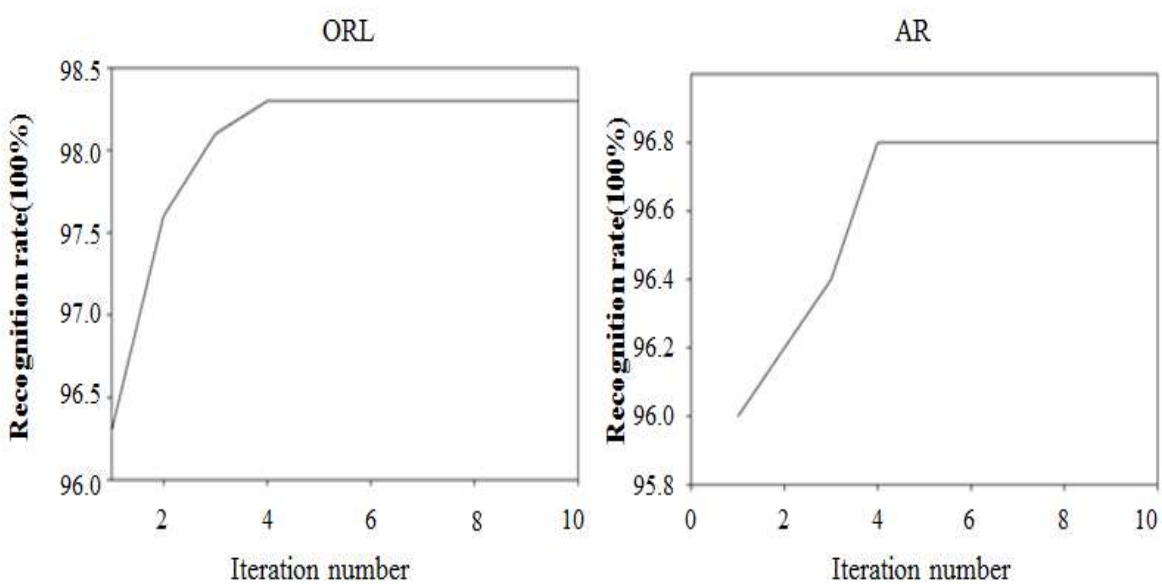

Figure 2. Recognition rate versus α in MDME on ORL and AR databases

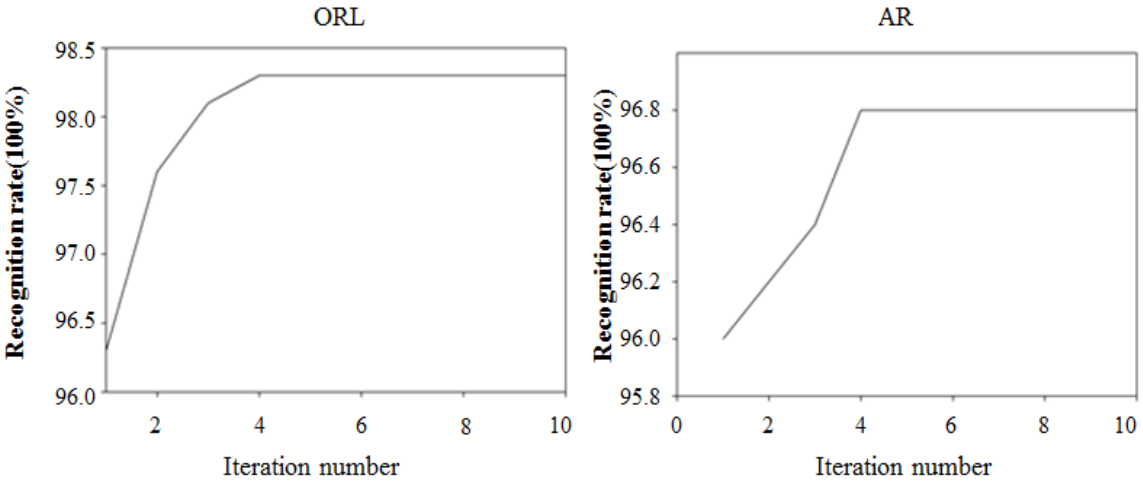

Figure 3. Recognition rate versus iteration number in MDME on ORL and AR databases

From the experimental results listed in Table 1, we can observe that the proposed MDME method significantly outperforms other multi-view and single view methods on the two databases, which implies that extracting the discriminative feature space by making use of intra-class geometry and inter-class marginal information and explicitly exploring the complementary characteristics of different visual features can achieve the best recognition performance. As shown in Figure 3.

In our proposed MDME algorithm, the parameter α is empirically set to 1 in the previous experiments. To investigate the influence of different choices of α, Fig.1 shows the recognition rate of MDME versus different values of α, we can observe that MDME demonstrates a stable recognition performance over a large range of α. Hence, the parameter selection is not a very crucial in MDME algorithm. In addition, since MDME is an iterative algorithm, we evaluate its recognition rate with different number of iterations in Fig.2. As can be seen, our proposed MDME algorithm can converge to a local optimal value in less than 5 iterations.

5. Conclusions

Data reduction is an important step towards the process of large data processing, and has become a hot research topic in the field of pattern recognition.

Feature extraction cannot only find the meaningful low dimensional structure information in high dimensional data, but also can reduce the impact of the curse of dimensionality to a certain extent, and promote the visualization and classification of data. Face recognition technology brings us to a lot of convenience, it has a wide range of application prospects. The development of face recognition technology not only improve the security of electronic technology products, but also can promote the development of artificial intelligence, image processing, computer graphics, cognitive science and psychology and other disciplines. Because the face structure model is too complicated and diverse, and easily influenced by illumination, facial expressions, gestures and other factors, how to effectively extract features to achieve the purpose of dimension reduction is the key in face recognition. Graph embedding framework is a better unified manifold learning linear dimension reduction algorithm and nonlinear dimensionality reduction algorithm, in order to understand the difference between the dimension reduction algorithm and the link is very helpful.

This letter proposes a new manifold learning algorithm, called multi-view discriminative manifold embedding (MDME) for feature extraction and classification. Compared with the conventional single-view-based subspace learning algorithms, MDME can consider local geometry and discriminative information as well as complementary information of multiple feature representations to obtain an effective low-dimensional embedding. Experimental results on two face databases have demonstrated the efficacy of the proposed approach.

Above is the main work done in this paper, although in the dimension of data reduction researchers has made some achievements,

but there still some deficiencies, so we need to continue to study the research direction, the complexity of the algorithm.

References

[1] Belhumenur, P. N., Hepanha, J. P. (1997). Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection. *IEEE Trans. Pattern Anal. Mach. Intell.* 19, 711-720

[2] He, X., Yan, S., Hu, Y., Niyogi, P., Zhang, H. -J. (2005). Face Recognition Using Laplacianfaces. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 328-340.

[3] Yan, S., Xu, D., Zhang, B. (2007). Graph Embedding and Extensions: A General Framework for Dimensionality Reduction. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 40-51

[4] Xie, B., Mu, Y., Tao, D. (2011). M-SNE: Multi-view Stochastic Neighbor Embedding. *IEEE Trans. Syst., Man, Cybern. B, Cybern.* 41 (6) 1088-1096.

[5] Belkin, M., Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation.* 15 (6) 1373-1396.

[6] Hardoon D R, Szedmak S, Shawe-Taylor J. (2004). Canonical correlation analysis: An overview with application to learning methods. *Neural Computation.* 16 (12) 2639-2664.

[7] Schölkopf B, Smola A, Müller K R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation.* 10 (5) 1299-1319.

[8] Zhang, L., Zhang, Q., Zhang, L., (2015). Ensemble manifold regularized sparse low-rank approximation for multi-view feature embedding. *Pattern Recognition*. 48 (10) 3102-3112

[9] Baudat, G., Anouar, F. (2000). Generalized discriminant analysis using a kernel approach. *Neural Computation.* 12 (10) 2385-2404.

[10] Zhang, J., Li S Z., Wang J. (2005). Manifold learning and applications in recognition//Intelligent multimedia processing with soft computing. Springer Berlin Heidelberg. 12 (10) 281-300.

[11] Phillips, P J., Wechsler, H., Huang, J., (1998). The FERET database and evaluation procedure for face-recognition algorithms. *Image and Vision Computing*. 16 (5) 295-306.

[12] Wiskott, L., Fellous, J M., Kuiger, N. (1997). Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 19 (7) 775-779.

[13] Yang, J., Zhang, D., Frangi, A F. (2004). Two-dimensional PCA: a new approach to appearance-based face representation and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 26 (1) 131-137.

[14] Parthiban, K., Wahi, Amitabh., Sundaramurthy, S, Palanisamy, C. (2014). Finger Vein Extraction and Authentication Based on Gradient Feature Selection Algorithm, *Journal of Information & Systems Management*, 4 (1) 1-10.

[15] Ibrahim, Muhammad Umar., Abbas, Mujahid . (2014). Biometric Authentication via Facial Recognition, *Journal of Information Security Research*, 5 (2) 61-68.