Applied Mathematics and Computation 219 (2013) 6410–6419

Contents lists available at SciVerse ScienceDirect

Applied Mathematics and Computation

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

Median null(S_w)-based method for face feature recognition

Jian-qiang Gao [a,*], Li-ya Fan [b], Li-zhong Xu [a]

[a] College of Computer and Information Engineering, Hohai University, Nanjing 210098, PR China
[b] School of Mathematics Sciences, Liaocheng University, Liaocheng, Shandong 252059, PR China

ARTICLE INFO

Keywords:
M-N(S_w)
Linear discriminant analysis (LDA)
Within-class median
Null space
Face recognition

ABSTRACT

With the progress of science and technology artificial intelligence is being paid more and more attention. People want to use computers to deal with complex practical problems. So, linear discriminant analysis (LDA) is widely used as a dimensionality reduction technique in image and text recognition classification tasks. However, a weakness of LDA model is that the class average vector in the formula completely depends on class sample average. Under special circumstances such as noise, bright light, some outliers will appear in the practical input databases. Therefore, by employing several given practical samples, the class sample average is not enough to estimate the class average accurately. So, the recognition performance of LDA model will decline. Compared to human intelligence, computers are far short of necessary fundamental knowledge of judgment which people normally acquire during the formative years of their lives. In order to solve the problem and also to render LDA model more robust, we propose a within-class scatter matrix null space median method (M-N(S_w)), which first transforms the original space by employing a basis of within-class scatter matrix null space, and then in the transformed space the maximum of between-class scatter matrix is pursued. In the second stage, within-class median vector is used in the traditional LDA model. Experiments on ORL, FERET and Yale face data sets are performed to test and evaluate the effectiveness of the proposed method.

1. Introduction

Face recognition has been researched in many areas such as pattern recognition and computer vision. Because it is a natural and direct biometric approach.

Under controlled or uncontrolled conditions, many developments have been made towards recognizing faces as described in [1–7]. Linear discriminant analysis (LDA) [8] and principal component analysis (PCA) [8] methods introduced by Turk and Pentland are two popular methods used in face recognition tasks. As we all know, PCA aims to generate a set of orthonormal projections by maximizing the covariance over all the sample. Therefore, it is an effective approach to represent each face image. However, from classification point of view, this method is not the best due to it does not make full use of the classification information.

LDA is a well-known linear learning method, whose goal is seeking optimal linear projections vectors so that the follow fisher criterion of the between-class scatter versus the within-class scatter is maximized. So, it has got better recognition performance than PCA.

* Corresponding author.
E-mail address: jianqianggaoHH@126.com (J.-q. Gao).

0096-3003/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.amc.2013.01.005

$$J(W) = \frac{W^T S_b W}{W^T S_w W}, \tag{1}$$

However, we often meet small sample size and high dimensional data problems in face classification and recognition tasks. Therefore, the traditional LDA cannot be used directly due to the fact that the within-class scatter matrix is always singular. In order to overcome the problem some corresponding methods were proposed [9–17]. LDA/QR [14] method introduced by Ye and Li is a popular approach used in face recognition. LDA/QR is a two-stage linear discriminant analysis learning approach, which aims to overcome the singularity problems of traditional linear discriminant analysis. It can achieve efficiency and stability simultaneously. A more effective null subspace discriminant method by Chen et al. [15] proposed to handle small sample size problem. A direct linear discriminant analysis method by Yu and Yang [10] proposed to deal with high dimensional image data. Song et al. [18] suggested a maximum scatter difference method, which adopts the difference of both within-class scatter and between-class scatter as discriminant criterion. Because the inverse matrix need not be calculated, the small sample size problem is avoided in nature.

There is a common weakness in using these approaches should be mentioned, the inverse matrix of between-class, within-class and total scatter matrices must be calculated. Therefore, this is a quite complex procedure.

An important problem in using the existing linear discriminant analysis models should be mentioned. In traditional linear discriminant analysis models, the class average vector is always estimated by the class sample average. However, by employing several given practical samples, the class sample average is not enough to estimate the class average accurately in particular when there are outliers in the images with noise and occlusion sample sets [19]. Li et al. [20] proposed a median-based maximum scatter difference method to deal with the complex procedure of within-class scatter matrix problem. We propose a two-stage linear discriminant analysis method called median-based null space of S_w. We call this method M-N(S_w) for short. Thus, the M-N(S_w) method should be more robust than the class sample average based traditional linear discriminant analysis models. Experiments on ORL, FERET and Yale face databases are performed to test and evaluate the robustness of the proposed method.

The rest of this paper is organized as follows. The traditional linear discriminant analysis model is briefly introduced in Section 2. The current methods PCA and LDA/QR are produced in Section 3. Section 4 introduces the concept of median. Our proposed method is introduced in Section 5. Experimental results and conclusion are summarized in Sections 6 and 7, respectively.

2. Traditional linear discriminant analysis (TLDA)

In this section, we first introduce some important notations used in this paper. For matrix $A \in R^{n\times N}$, we consider seeking a linear transformation $G \in R^{n\times l}$ that maps each a_i of A to l-dimensional space $y_i \in R^l$ with $y_i = G^T x_i$. Assume that the original data in A is partitioned into c classes as $A = [A_1, \ldots, A_c]$, where $A_i \in R^{n\times N_i}$ contains sample data points from the ith class and $N = \sum_{i=1}^{c} N_i$. Finding optimal transformation matrix G is a core problem of traditional linear discriminant analysis model, and then the class structure of the original high-dimensional space is preserved in the reduced-dimensional space.

In discriminant analysis, the between-class, within-class and total scatter matrices are defined as follows (see [21]):

$$S_b = \frac{1}{N}\sum_{i=1}^{c} N_i (m_i - m)(m_i - m)^T = H_b H_b^T, \tag{2}$$

$$S_w = \frac{1}{N}\sum_{i=1}^{c}\sum_{j=1}^{N_i} (x_i^j - m_i)(x_i^j - m_i)^T = H_w H_w^T, \tag{3}$$

$$S_t = S_b + S_w, \tag{4}$$

where the precursors H_b and H_w of the between-class and within-class scatter matrices in (2) and (3) are

$$H_b = \frac{1}{\sqrt{N}}\left[\sqrt{N_1}(m_1 - m), \ldots, \sqrt{N_c}(m_c - m)\right], \tag{5}$$

$$H_w = \frac{1}{\sqrt{N}}\left[A_1 - m_1 e_1^T, \ldots, A_c - m_c e_c^T\right], \tag{6}$$

$e_i = (1, \ldots, 1)^T \in R^{N_i}$. A_i is the data matrix of the ith class, m_i is the center of the ith class and m is the total centroid of the training sample set. It is worthwhile to note that the total scatter matrix S_t can be called covariance matrix in statistics.

We can obtain the between-class, within-class and total scatter matrices $S_b^l = G^T S_b G$, $S_w^l = G^T S_w G$ and $S_t^l = G^T S_t G$ by using linear transformation matrix G, in the reduced-dimensional space, respectively. An optimal transformation G would