Optik 127 (2016) 883–889

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Improved the minimum squared error algorithm for face recognition by integrating original face images and the mirror images

Xiaojun Wen[a], Jie Wen[b,*]

[a] School of Computer Engineering, Shenzhen Polytechnic, Shenzhen, Guangdong 518055, China
[b] College of Automation, Harbin Engineering University, Harbin, Heilongjiang 150001, China

ARTICLE INFO

Article history:
Received 30 January 2015
Accepted 27 October 2015

Keywords:
Various poses and illuminations
Face recognition
Minimum squared error classification
Mirror face
Kernel minimum squared error

ABSTRACT

In order to improve the accuracy of face recognition and solve the problem of various poses and illuminations, we proposed an improved minimum squared error (IMSE) classification algorithm. Firstly, the mirror faces of the original training faces are generated through row preserved and columns flipped in the left/right direction. Secondly, the minimum squared error (MSE) algorithm is performed on both original faces and the mirror faces. Thirdly, the predicted errors of the test sample and standard class labels are obtained. In addition, the residual between the predicted labels of the test sample and each training sample can also be calculated. At last, the correct class can be determined by fusing the predicted errors and residuals. We also promoted the IMSE algorithm to the kernel MSE algorithm and proposed an improved kernel minimum squared error (IKMSE) algorithm for face recognition. The experimental results show our proposed IMSE and IKMSE algorithm are more robust than the conventional MSE and KMSE algorithm, respectively. In addition, our proposed algorithms improve the accuracy of face recognition effectively.

1. Introduction

Face recognition as a biological identification technology has attracted much attention and is widely applied in the field of public security system, banking system, electronic commerce and military security [1–4]. In the last decades, many face recognition algorithms have been proposed, such as Eigenface [5], Fisherface [6] and Laplacianface [7]. The conventional Eigenface and Fisherface classification algorithms need to convert the face matrix into vector firstly for face recognition; it will lose the structure information of face and is easily affected by the small-sample-size problem. In order to solve the problems above, the two-dimensional principal component analysis (2DPCA) [8,9] and two-dimensional linear discriminant analysis (2DLDA) [10] which can extract the features from matrix directly have been proposed.

The minimum square error (MSE) has good properties and is widely applied in the field of pattern recognition and linear regression [11,12]. The conventional MSE (CMSE) face classification algorithm focus on using the minimum square error method to obtain a mapping that can best transform all training samples into the corresponding class labels. For multi-class classification under conditions of high-dimensional and under-sampled data, reference [13] proved that the CMSE classification algorithm is equivalent to the linear discriminant analysis classification algorithm. In recent years, many improved MSE algorithms have been proposed. Xu et al. proposed an improved MSE algorithm by fusing the predicted class labels of the training sample and test sample for face recognition [14]. The kernel MSE (KMSE) classification algorithm is the well-known improved MSE algorithm, which can convert the nonlinearly separable problem to the linearly separable problem and thus obtain better performance than CMSE algorithm [15]. However, if a training set has a large number of samples, the KMSE algorithm will be inefficient for face recognition. In order to improve the efficiency of conventional kernel minimum square error (CKMSE) algorithm, certain fast KMSE algorithms based on selecting the important features which have much contribution have been proposed [16,17]. The representation-based classification (RBC) method can be viewed as another improved form of CMSE algorithm for face recognition [18–21]. Different with the CMSE algorithm, RBC method takes the training samples and test sample as the input and output of the representation equation, respectively. Similar to CMSE method, RBC method also uses the minimum squared error method to obtain a mapping that can best transform the training samples into the test sample. The sparse representation classification (SRC) method, collaborative representation classification (CRC) method and linear regression classification (LRC) method are the typical RBC method for face recognition. The conventional SRC algorithm and CRC algorithm obtain the mapping

* Corresponding author. Tel.: +86 15218720459.
E-mail address: wenjie@hrbeu.edu.cn (J. Wen).

http://dx.doi.org/10.1016/j.ijleo.2015.10.182

or representation coefficient by calculating the minimum representation error between all training samples and the test sample with l_1-norm constraint and l_2-norm constraint, respectively [22]. The LRC method can be viewed as the subspace form of the CRC method. If there are c classes, it needs to solve c representation equations using minimum squared error method [18]. Based on the CRC method, literature [23] proposed an improved CRC algorithm that is able to reduce the problem of noise for face recognition. Yang et.al proposed a Gabor-feature-based SRC algorithm that can reduce the influence of block occlusion and disguise [24].

The problem of face with various poses, expressions, illuminations and limited training samples are the main challenges for practical application of face recognition [25,26]. Some researchers used near-infrared image [27] and thermal infrared image [28] to eliminate the influence of various illuminations. In the real-world application, the training set usually does not contain too many training samples to sufficiently express various face poses, and sometimes it only has one sample per person, which lead to a bad performance by using the traditional classification methods [29–31]. Some researchers addressed on using the virtual face to improve the recognition accuracy with the problems above. Xu et al. proposed a virtual frontal-face-generated method based on the symmetry of face to enlarge the number of training sample that can improve the accuracy under various poses [32]. Chai et.al proposed a locally linear regression method that can convert non-frontal face images to the frontal view, pose-invariant face images [33]. Sharma et.al presented a method to obtain multiple virtual views of a person under different poses and illuminations from a single 2D face image for face recognition [34]. Literature [35] first used the mirror image of the original face image for face recognition. The mirror image holds good properties. For example, the mirror face and the original face have opposite illumination and pose information. Using the mirror face not only enlarges the number of training sample but also expresses possible face poses and illuminations. The mirror face is easy to be generated and can significantly improve the classification accuracy.

In this study, considering the good properties of the mirror face, we propose an improved MSE (IMSE) algorithm by using both original face images and the mirror images for face recognition. Moreover, our method is also promoted to the kernel MSE algorithm for face recognition. Both original training images and the mirror images are simultaneously used to learn the mapping matrix of MSE algorithm and KMSE algorithm. The experiments which are performed on FERET, Yale B and GT face databases show that our proposed algorithms can improve the face recognition rate effectively, especially under variable poses, illuminations and limited training samples.

2. The conventional MSE-based classification algorithm

2.1. The conventional MSE classification algorithm

The CMSE algorithm focuses on obtaining a mapping which can best transform the input into the output [14]. In the CMSE classification algorithm, the training sample and its corresponding class label are taken as the input and output, respectively.

Suppose there are c subjects and each subject contains n training images. Let a row vector $x_{i,j} \in \mathbb{R}^{1\times p}$ denotes the jth training sample of ith class. $x_{i,j}$ is formed by stacking the image columns. A row vector $g_i \in \mathbb{R}^{1\times c}$ is used to represent the class label of the ith class, and the ith element of class label g_i is set as 1 and rest elements of g_i are set as 0. It means, for the first class and last class, $g_1 = [1, 0, \ldots, 0] \in \mathbb{R}^{1\times c}$ and $g_c = [0, 0, \ldots, 1] \in \mathbb{R}^{1\times c}$, respectively. The CMSE algorithm uses a linear transformation to transform each training sample into its corresponding class label. That is

$$\mathbf{XA} = \mathbf{G} \tag{1}$$

where,

$$\mathbf{X} = \begin{bmatrix} x_{1,1} \\ x_{1,2} \\ \vdots \\ x_{c,n} \end{bmatrix}, \quad \mathbf{G} = \begin{bmatrix} g_1 \\ g_1 \\ \vdots \\ g_c \end{bmatrix}$$

$\mathbf{A} \in \mathbb{R}^{p\times c}$ denotes the linear transformation matrix. $\mathbf{G} \in \mathbb{R}^{N\times c}$ is the class label matrix. $N = cn$ is the total number of training samples. In order to obtain the transformation matrix $\mathbf{A}$, Eq. (1) can be converted to the following equation:

$$\mathbf{X}^T\mathbf{XA} = \mathbf{X}^T\mathbf{G} \tag{2}$$

Then we can obtain the transformation matrix $\mathbf{A}$.

$$\hat{\mathbf{A}} = (\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{G} \tag{3}$$

where, λ is a small positive constant. $\mathbf{I}$ is an identity matrix. The predicted class label of a test sample s can be calculated by the following equation:

$$g_s = s\hat{\mathbf{A}} \tag{4}$$

Then the predicted error between the predicted labels of test sample and each standard class label can be calculated

$$e_i = \left\| g_s - g_i \right\|_2, \quad i = 1, 2, \ldots, c \tag{5}$$

At last, the test sample s is assigned to the kth class which has the minimum predicted error with the test sample. It can be expressed as:

$$k = \underset{i}{argmin} e_i, \quad i = 1, 2, \ldots, c \tag{6}$$

2.2. The conventional kernel MSE classification algorithm

The kernel MSE (KMSE) algorithm, as an improvement algorithm of MSE, can deal with nonlinearly separable problems well and is widely applied to the field of pattern recognition [36]

The conventional KMSE (CKMSE) algorithm uses a nonlinear funtion φ to transform the original nonlinearly separable samples into a higher dimensional space, which can make the samples from different classes to be separated linearly.

In the new higher dimensional space, the training samples can also be mapped into the corresponding class labels by the following equation:

$$\boldsymbol{\phi}\mathbf{W} = \mathbf{G} \tag{7}$$

where,

$$\boldsymbol{\phi} = \begin{bmatrix} 1 & \varphi(x_{1,1}) \\ 1 & \varphi(x_{1,2}) \\ \vdots & \vdots \\ 1 & \varphi(x_{c,n}) \end{bmatrix}, \mathbf{W} = \begin{bmatrix} w_0 \\ w \end{bmatrix} = \begin{bmatrix} w_0 \\ w_{1,1} \\ \vdots \\ w_{c,n} \end{bmatrix}, \mathbf{G} = \begin{bmatrix} g_1 \\ g_1 \\ \vdots \\ g_c \end{bmatrix}.$$

It should be noted that the two training matrices of the CMSE algorithm and CKMSE algorithm are different. The purpose of adding the element of 1 to each training sample on the left side is to prevent the mapping result to pass the origin point.

According to the theory of reproducing kernel [37], w can be expressed as

$$w = \sum_{i=1}^{c}\sum_{j=1}^{n}\alpha_{i,j}\varphi^T(x_{i,j}) \tag{8}$$