ARTICLE IN PRESS

Neurocomputing (xxxx) xxxx–xxxx

Contents lists available at ScienceDirect

Neurocomputing

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

Bi-weighted robust matrix regression for face recognition

Jianchun Xie[a,b], Jian Yang[a,*], Jianjun Qian[a], Lei Luo[a]

[a] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China
[b] School of Science, Nanjing University of Science and Technology, Nanjing, China

ARTICLE INFO

Communicated by Jiwen Lu

Keywords:
Face recognition
Low rank
ADMM
Matrix regression
Structural errors

ABSTRACT

Regression analysis based classification methods have attracted much interest in the face recognition area. However, dealing with partial occlusion or illumination is still one of the most challenging problems. In most of the current methods, the image needs to be stretched into a vector and each pixel is assumed to be generated independently, which ignores the dependence between pixels of the error image. That is, these methods do not consider the structure information of the image with continuous occlusion or disguise in modeling. In this paper, it is found that the non-convex function of the singular values can well describe the low rank structure of the image data. By virtue of this fact, we propose a bi-weighted robust matrix regression (BWMR) model for face recognition with structural noise, in which the non-convex function of the singular values is used as regularization. The alternating direction method of multipliers (ADMM) is applied to solving the proposed model. Experimental results demonstrate that the proposed method is more robust and effective than the state-of-the-art methods when handling the structural errors.

1. Introduction

Over the past years, face recognition has developed to be an important research area in image classification and computer vision. Like any other enduring discipline, face recognition also had its initial challenges. For example, the facial expression (laughter, anger, etc.), occlusion and illumination variations severely influence the face recognition performance [1]. Presently, there are various face recognition methods. For example, in order to handle the face with structural noise, Ou et al. [2] proposed to learn a clear dictionary and a noise dictionary simultaneously, then the clear dictionary was applied to classification task. You et al. [3,4] proposed a robust nonnegative patch alignment for dimensionality reduction, the experiments show that the learned representation is robust against occlusion, and extreme variations in illumination. Especially, regression analysis based methods have achieved promising results and become a popular tool for face recognition [5–11].

Linear regression classifier for face recognition aims to get the linear representation of a test sample, and classify it by checking the minimization of the representation error. In order to avoid over-fitting, the regression methods admit a tradeoff between the empirical loss and regularization as:

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) + \lambda g(\boldsymbol{x}) \tag{1}$$

where $f(\boldsymbol{x})$ describe the reconstruction residual, $g(\boldsymbol{x})$ is the regularization associated with coding coefficients $\boldsymbol{x}$, and λ is the regularized parameter.

Using the l_1-norm regularization on the coefficient vector $\boldsymbol{x}$, sparse representation based classifier (SRC) [5] has become one of the most famous regression based method. Some recent work began to investigate the role of sparsity in face recognition [12,13], Yang et al. [12] gave an insight into SRC and provided that it is l_1-norm constraint rather than l_0-norm that makes SRC effective. Meanwhile, Zhang et al. [13] found that it is not necessary to adopt the l_1-norm regularization, and proposed the collaborative representation classifier (CRC) based on the l_2-norm regularization, known as ridge regression. Compared with the SRC, the CRC has very competitive face recognition accuracy but lower time complexity. Both the SRC and CRC, the error loss is usually measured by the l_2-norm, which assumes the pixels of the error follow Gaussian distribution independently. However, the assumption might to be unreasonable in real situations, such as occlusion, disguise, etc.

Besides the representation coefficients, the reconstruction residual also influences the performance of object classification. As we known, the l_2-norm for the residual is sensitive to various types of outliers (e.g. occlusion, corruption, expression, etc.). Yang et al. [14] presented robust sparse coding (RSC) model and solved it by an effective iteratively reweighted sparse coding algorithm. RSC is essentially a sparsity-constrained robust regression process. Naseem et al. [15] presented robust linear regression classification (RLRC) algorithm by

* Corresponding author.
E-mail addresses: jcxie@njust.edu.cn (J. Xie), csjyang@njust.edu.cn (J. Yang), csjqian@njust.edu.cn (J. Qian), zzdxpyy3001@163.com (L. Luo).

http://dx.doi.org/10.1016/j.neucom.2017.01.028
Received 19 September 2016; Received in revised form 5 January 2017; Accepted 8 January 2017
0925-2312/

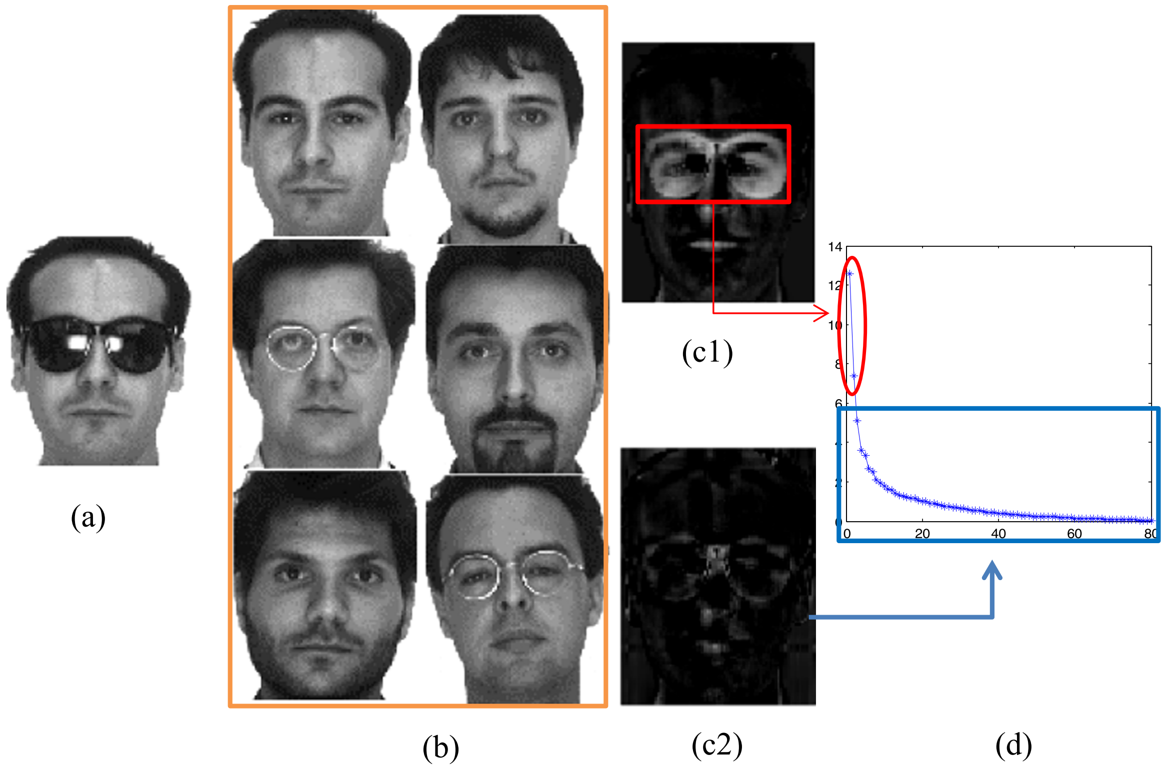

Fig. 1. The example shows the regression method with occlusion. (a) the test image with glasses, (b) the training images, (c1) the error image, (c2) the reconstructed image with some partial small singular values, (d) the singular values of the error image.

using the robust Huber estimation approach for dealing with the random pixel corruption and illumination variation. Based on the maximum correntropy criterion, He et al. [16] presented the correntropy based sparse representation (CESR) algorithm. All of these methods have achieved promising results in face recognition. In addition, some scholars considered the structural noise by dictionary learning and the.

However, in the linear regression for classification has been mentioned above, each image has to be stretched into a vector before classification and each pixel of error matrix is independently handled. In detail, $\boldsymbol{y} = [y_1; y_2; \ldots; y_n] \in \mathbf{R}^n$ is the testing image to be coded, n is the number of the pixels. $\mathbf{D} = [\boldsymbol{d}_1; \boldsymbol{d}_2; \ldots, \boldsymbol{d}_n] \in \mathbf{R}^{n\times m}$ denotes a dictionary with row vector $\boldsymbol{d}_i$. Let $\boldsymbol{e} = \boldsymbol{y} - \mathbf{D}\boldsymbol{\alpha} = [e_1; e_2; \ldots; e_n]$, where $e_i = y_i - \boldsymbol{d}_i\boldsymbol{\alpha}$. The RSC model can be expressed as $\min_{\boldsymbol{\alpha}} \sum_{i=1}^{n} w_i e_i^2 + \lambda \parallel \boldsymbol{\alpha} \parallel_1$, which imposes the different weights on error pixels. According to the RSC, the weights corresponding to noise pixels should be small. However, for many practical face variations, such as occlusion, disguise, or shadow caused by illumination change, the independent assumption does not hold. As shown in Fig. 1, the image (a) and (b) can be considered as the testing data and the training data, respectively. From the residual image (c1), we can find that the representation errors in the sunglasses part are correlated, because the pixel values of the occlusion part are zeros in the testing image (a). Therefore, assigning the different weights to error pixels for image classification with occlusions may be unsuitable.

Recently Yang et al. [8] found that the error image caused by occlusion or illumination changes has low rank structure and proposed the Nuclear norm Matrix Regression (NMR) model for face representation and classification. In modeling, NMR does not convert the matrix into a vector, that is, it directly deals with 2-D image. The initial idea of NMR is to achieve the best representation for the testing sample by minimizing the rank function of the error matrix. As we known, the rank minimization problem is NP hard. Yang et al. adopted the nuclear norm as the relaxation function of rank function and demonstrated the effectiveness of the nuclear norm for structural noise characterization by experiments on several public databases. Using the nuclear norm as matrix regularization and combining the idea of the RSC, Qian et al. [17,18] proposed robust nuclear norm regularized regression (RNR) method for face recognition, while the RNR model still assigns the weights to the each pixel of the residual image.

In this paper, we develop a robust matrix regression for face recognition problem. It is observed that the occlusion part can be considered as the set of outliers and the different singular values can describe the structure error, Thus, the weights is assigned to singular values rather than pixels of error image in our method. By a series of matrix completion experiments, we can find that the non-convex function of the singular value can characterize the essential low rank structure of the image matrix. Therefore, we adopt the non-convex function of the singular value as the regularization to depict the low rank structure information. The model can be solved via the ADMM method.

The remainder of the paper is organized as follows: Section 2 introduces the Bi-Weighted robust Matrix Regression model (BWMR) and presents the ADMM method to solve the model. In Section 3, we suggest the classifier criterion for robust classification. In Section 4, we conduct experiments and comparisons with the state-of-the-art methods. Finally, Section 5 concludes the paper.

Notations : Throughout this paper, $\mathbf{R}^n$ denotes the space of n-dimensional real column vectors, and $\mathbf{R}^{m_1\times m_2}$ denotes the space of $m_1 \times m_2$ dimensional real matrices. For a matrix $\mathbf{E} \in \mathbf{R}^{m_1\times m_2}$ (we assume that $m_1 \le m_2$ in this work), we write its singular value decomposition (SVD) as $\mathbf{E} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ with $\mathbf{U} \in \mathbf{R}^{m_1\times m_1}$, $\mathbf{V} \in \mathbf{R}^{m_2\times m_2}$ and $\boldsymbol{\Sigma} = [\mathrm{diag}\{\sigma_{i,i=1,2,\ldots,m_1}\}, 0]$, where σ_i is the i-th largest singular value of $\mathbf{E}$. $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \ldots, \sigma_{m_1})^T$ denotes the singular value vector. $rank(\mathbf{E})$ is the rank of $\mathbf{E}$, i.e. the l_0-norm of the singular value vector. $\parallel \mathbf{E} \parallel_* = \sum_{i=1}^{m_1} \sigma_i(\mathbf{E})$ is the nuclear norm of a matrix $\mathbf{E}$, and $\parallel \mathbf{E} \parallel_{w,*} = \sum_{i=1}^{m_2} w_i\sigma_i(\mathbf{E})$ denotes the weighted nuclear norm. The F-norm of matrix, $\|\mathbf{E}\|_F^2 = \Sigma_{ij} e_{ij}^2$, the l_1-norm of a vector is defined by $\parallel \boldsymbol{x} \parallel_1 = \sum |x_i|$. l_2-norm of a vector is defined by $\parallel \boldsymbol{x} \parallel_2 = (\sum x_i^2)^{1/2}$. $vec(\cdot)$ denotes an operator converting a matrix to a vector.