Pattern Recognition 56 (2016) 77–87

Contents lists available at ScienceDirect

Pattern Recognition

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

Corrupted and occluded face recognition via cooperative sparse representation

Zhong-Qiu Zhao [a,b,*], Yiu-ming Cheung [b], Haibo Hu [d], Xindong Wu [c]

[a] College of Computer Science and Information Engineering, Hefei University of Technology, China
[b] Department of Computer Science, Hong Kong Baptist University, Hong Kong, China
[c] Department of Computer Science, University of Vermont, USA
[d] Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong, China

ARTICLE INFO

Article history:
Received 11 June 2014
Received in revised form
14 December 2015
Accepted 19 February 2016
Available online 27 February 2016

Keywords:
Face recognition
Cooperative sparse representation
Recognition rate
Time consumption

ABSTRACT

In image classification, can sparse representation (SR) associate one test image with all training ones from the correct class, but not associate with any training ones from the incorrect classes? The backward sparse representation (bSR) which contains complementary information in an opposite direction can remedy the imperfect associations discovered by the general forward sparse representation (fSR). Unfortunately, this complementarity between the fSR and the bSR has not been studied in face recognition. There are two key problems to be solved. One is how to produce additional bases for the bSR. In face recognition, there is no other bases than the single test face image itself for the bSR, which results in large reconstruction residual and weak classification capability of the bSR. The other problem is how to deal with the robustness of the bSR to image corruption. In this paper, we introduce a CoSR model, which combines the fSR and the bSR together, into robust face recognition, by proposing two alternative methods to these two key problems: *learning bases* and *unknown faces help* to enrich the bases set of the bSR. Thereby, we also propose two improved algorithms of the CoSR for robust face recognition. Our study shows that our CoSR algorithms obtain inspiring and competitive recognition rates, compared with other state-of-the-art algorithms. The bSR with the proposed methods enriching the bases set contributes the most to the robustness of our CoSR algorithm, and *unknown faces* works better than *learned bases*. Moreover, since our CoSR model is performed in a subspace with a very low dimensionality, it gains an overwhelming advantage on time consumption over the traditional RSR algorithm in image pixel space. In addition, our study also reveals that the sparsity plays an important role in our CoSR algorithm for face recognition.

1. Introduction

Recently, sparse representation (SR) based image classification has undergone a great development [1–7]. Given some training images with class labels, sparse representation selects several training images which give the most accurate and compact representation of one test image according to visual content similarity, and then classifies the test image by the reconstruction residual associated with each object class. The sparse representation classifier can be considered as a generalization of the nearest neighbor (NN) [8,9] or the nearest subspace (NS) [10]. The NN classifier approximates the test image with a single training image which is the most similar, while the NS classifier is based on *the strategy* of the best linear representation of the test image with all of the training images in each class. However, the NN can be easily affected by noise, especially for real applications, while the NS may not work well when classes are highly correlated to each other [3]. As a tradeoff between the NN and the NS, the sparse representation (SR) [3,11] can automatically select a small number of training images and approximate the test image with a weighted sum of the selected ones. The motivation of finding a sparse solution rather than a dense one lies in that a dense solution usually results in a large number of training images selected to approximately represent the test image, while these training images may come from various classes. Therefore, this dense solution is not especially informative for classification [3].

Robustness is an important and open problem in real face recognition systems since test face images are usually partially corrupted or occluded [12,13]. One of the most successful applications of sparse representation based image classification is to deal with corruptions and occlusions in face recognition [3,14]. Face recognition can be performed in the subspaces such as principal component analysis (PCA) [15,16], linear discriminant analysis (LDA) [17], and

* Corresponding author.
E-mail addresses: z.zhao@hfut.edu.cn (Z.Q. Zhao), ymc@comp.hkbu.edu.hk (Y.M. Cheung), haibo.hu@polyu.edu.hk (H. Hu), xwu@cs.uvm.edu (X. Wu).

http://dx.doi.org/10.1016/j.patcog.2016.02.016
0031-3203/

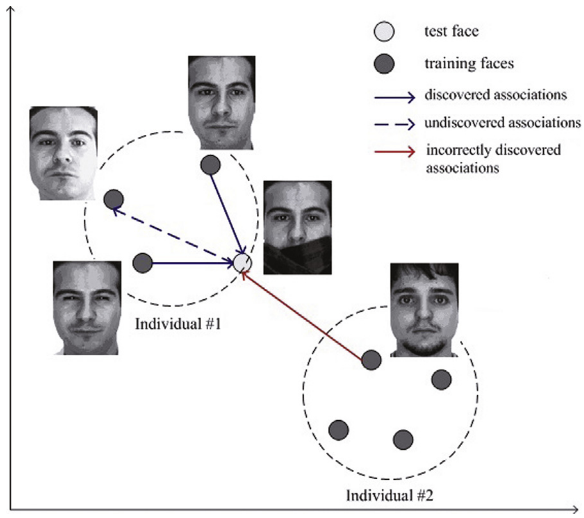

Fig. 1. A sketch of the SR bridging between test face and training faces.

locality preserving projection (LPP) [18] where the dimensionality is reduced below the number of training faces so that the approximation problem is underdetermined and simplified. However, Wright et al. [3] believed that noise on the original pixels cannot be eliminated by the subspace projections, and no bases or features are more spatially localized than the original image pixels themselves. So they proposed to perform face recognition with sparse representation in the original pixel space. In order to do that, besides training images, the bases set of robust sparse representation (RSR) is enlarged to reduce the reconstruction residual with the identity matrix which can linearly reconstruct any noise. We call this algorithm 'Pixel+RSR' in the following sections. In addition, Wang et al. proposed a sparse residue method for occluded face image reconstruction, which considers an occluded face image as the summation of a non-occluded face and a sparse occlusion [14]. Both of the above two methods can achieve satisfactory recognition accuracy for robust face recognition.

For face recognition, the SR bridges between the test face and the training faces from the same individual, and avoids associations between the test face and the training faces from the incorrect individuals. However, it is *NOT* definite that the training faces from the correct individuals can be all *selected*, while the ones from the incorrect individuals can be all excluded by the sparsity constraint, especially when the test image may contain noise. On the contrary, the case in Fig. 1 usually appears. Comparing to the general SR which represents each test image with training images and we call the forward sparse representation (fSR), the backward sparse representation (bSR), in an opposite direction, representing each training image with test images. These two SRs may contain some complementary information for classification. That is to say, the bSR can remedy the correct associations undiscovered by the fSR, and comparatively weaken the incorrect associations by strengthening the correct associations which are discovered by the fSR. The reason is that two opposite SRs are conditionally independent due to the sparsity of the SR and the diversity of images [4]. In [4], we validated the complementarity between the fSR and bSR, and proposed a cooperative sparse representation (CoSR) model for scene image annotation. Recently, Xu et al. [19] cast this idea into linear regression classification and ℓ_2-norm based representation methods.

In this paper, we will introduce the CoSR into robust face recognition and address two key problems. First, in the CoSR application in image annotation, all test images to be annotated can be utilized so that there are enough bases for the bSR to represent each training image with small residual. However, in face recognition, there is only one ready-made basis for the bSR, viz., the test face image itself. So how to construct and expand the bases set, viz. dictionary, of the bSR to reduce the reconstruction residual is one problem to be addressed in this paper. Another key problem is how to deal with the corruption in test images. If the corruption or noise in any test image is large, the bSR tends to utilize newly added bases rather than the test image to represent the training image, even if the training image and the test image come from the same individual, which will result in smaller reconstruction error. This will lead to the failure of recognition of the test image. Imposing the noise similar to that in the test face on the learned bases of the bSR or using additional unknown faces with the similar noise as the expanded bases of the bSR can reduce the correlation between the additional bases and the training images, so that the test image is more probably selected to represent the training images from the correct individual. Thereby the robustness of the bSR is improved. The primary motivation of this paper is to extend the CoSR model to solve the robust face recognition problem, namely corrupted and occluded face recognition, and to improve the robustness and performance of the CoSR for robust face recognition by utilizing this robustness of the bSR to noise in face images.

The main contributions of this paper are two-fold. (1) There have been no previous work on the complementarity between the fSR and the bSR for face recognition, and no previous work to explore to compensate the general SR for face recognition. This paper introduces the CoSR model into the field of face recognition, obtaining very competitive recognition rates and showing high efficiency for robust face recognition. (2) This paper proposes two methods, namely '*learning bases*' and '*unknown faces help*', to expand the bases set of the bSR, which are very robust to noise or corruption in face images, especially when the noise or corruption is very large. (3) This paper also reveals that the sparse constraint is very important for the success of our CoSR model on face recognition.

2. Related work

For robust face recognition, the 'Pixel+RSR' algorithm which solves sparse representation in the original pixel space inevitably brings the problem of high time complexity due to large feature dimensionality and large number of bases. Traditional ℓ_1 minimization solution algorithms by linear programming (LP) [20] have shown their high computational complexity for real-world applications with large size of face images. Heuristic greedy algorithms such as orthogonal matching pursuit (OMP) [21] and least angle regression (LARS) [22], which are faster than using LP methods, can often fall into local optimum [23], and these sub-optimal solutions result in poor face recognition performance, even in the original pixel space. Newly developed algorithms such as gradient projection (GP) [24], augmented lagrange multiplier (ALM) [25], homotopy [26], and feature-sign search sign (FSS) [27] methods, which attempt to improve the solution efficiency of ℓ_1 minimization, however, cannot avoid the problem of high time complexity when working in the original pixel space. Though the computational complexities of GP and ALM methods are difficult to estimate exactly, the solutions all involve the matrix operations w.r.t. the dictionary of sparse representation which is a huge matrix when in pixel space. And the computational complexity of the homotopy algorithm is $O(d^2+dn)$ for each iteration [28], where d and n denote the feature dimensionality and the number of bases, respectively. Comparatively, our proposed CoSR algorithm works well in subspaces such as the PCA, LDA, and LPP, which will largely reduce time consumption.