Information Sciences 346–347 (2016) 198–215

Contents lists available at ScienceDirect

Information Sciences

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

Local structure based multi-phase collaborative representation for face recognition with single sample per person

Fan Liu [a,b,*], Jinhui Tang [c], Yan Song [c], Ye Bi [c], Sai Yang [d]

[a] College of Computer and Information, HoHai University, Nanjing, 210098, P. R. China
[b] Key Laboratory of Image and Video Understanding for Social Safety, Nanjing University of Science and Technology, Nanjing, 210094, China
[c] School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing, 210094, P. R. China
[d] School of Electrical Engineering, Nantong University, Nantong, 226019, China

ARTICLE INFO

Article history:
Received 21 February 2014
Revised 18 November 2015
Accepted 2 February 2016
Available online 10 February 2016

Keywords:
Face recognition
Single sample per person problem
Collaborative representation
Local structure
Subspace
Class selection
Entropy minimization

ABSTRACT

Many real-world face recognition applications can only provide single sample for each person, while most face recognition approaches require a large set of training samples, which leads to *single sample per person* (SSPP) problem. In this paper, we propose local structure based multi-phase collaborative representation classification (LS_MPCRC) to solve SSPP problem. By adopting the “divide-conquer-aggregate” strategy, we successfully alleviate the dilemma of high data dimensionality and small samples, where we first divide the face into local blocks, and classify each local block, and then integrate all the classification results by voting. For each local block, we further divide it into overlapped local patches and assume that these patches lie in a linear subspace. This subspace assumption reflects local structure relationship of the overlapped patches and makes CRC robust for SSPP problem. Motivated by the fact that the entropy of the class probability distribution is a measure about classification confidence, we further apply multi-phase technique to reduce entropy, where useless classes are eliminated after each phase classification. This strategy finally produces a sparse class probability distribution with higher classification confidence. Experimental results show that the proposed method generalizes well to SSPP problem and outperforms many state-of-the-art methods. It also shows strong robustness to the large variation of expression, illumination, little poses variation, occlusion and time variation.

1. Introduction

Face recognition has been an active research topic in computer vision and pattern recognition for many years [25] and it is still attracting much attention because of its scientific challenges and its potential applications. Within the last 2 decades, various face recognition algorithms have been devised [2,18,20,26,28] and face recognition systems are considered to be critically dependent on discriminative feature extraction, about which many approaches have been proposed, such as Eigenface [26], Fishfaces [1], ISOMAP [18], LLE [13], LPP [6] and Laplacian Eigen map [2]. Recently, the significance of feature extraction has been debated. Wright et al. indicate that once the test image can be approximated by a sparse linear combination of

* Corresponding author. Tel.: +86 15851819468.
E-mail addresses: fanliu.njust@gmail.com (F. Liu), jinhuitang@mail.njust.edu.cn (J. Tang), songyan@mail.njust.edu.cn (Y. Song), biyemail518@gmail.com (Y. Bi), yangsai166@126.com (S. Yang).

http://dx.doi.org/10.1016/j.ins.2016.02.001

the training images, the choice of feature space is no longer critical [20]. Thus, they propose a sparse representation based classification (SRC). Later, Zhang et al. [28] propose a collaborative representation based classification method with regularized least squares (CRC_RLS) to further reduce the complexity of SRC. However, no matter subspace learning methods (such as PCA, LDA and so on) or representation based methods (e.g. SRC and CRC) always require a rich set of training samples of each subject. Unfortunately, in some real-world applications, the number of samples of each subject is usually very small. For example, there is only one image available in the scenario of identity card or passport verification, law enforcement, surveillance or access control. This is the so called *Single Sample per Person* (SSPP) [19] problem which severely challenges existing face recognition algorithms.

In this paper, we propose a simple yet effective algorithm, namely local structure based multi-phase collaborative representation classification (LS_MPCRC) to solve SSPP problem. Motivated by the "divide-conquer" [16] strategy, we divide each face into many local blocks and ensemble the classification results on each block to make the final decision. To classify each local block, we divide each block into overlapped patches and assume that the local patches within the local block lie in a linear subspace. Intuitively, if two face images are from the same person, their subspaces corresponding to the blocks at the same position should also be the same. Then, the central patch of the test block can be approximately represented by a linear combination of the patches in the corresponding block from the same class. In other words, if we conduct collaborative representation based classification (CRC) for each local block, the subspace assumption makes the CRC classification under SSPP condition transfer to the classification problem with more training samples of each subject. The whole procedure can be summarized as "divide-conquer-aggregate" and defined as local structure based collaborative representation classification (LS_CRC). Due to the fact that the entropy of the class probability distribution is a measure about classification confidence, the performance of LS_CRC can be further improved by entropy reduction. Motivated by this idea, we further propose local structure based multi-phase collaborative representation classification (LS_MPCRC), which gradually reduces the entropy of the class probability distribution by multi-phase class selection scheme. As the voting technique of LS_CRC makes the true class always lie in the top of the voting rank, the classes in the top are selected for next phase LS_CRC. To make the result more robust, the voting result of current phase LS_CRC is aggregated with those of previous phases and the accumulated voting result is used for class selection and classification. Finally, this local structure based multi-phase (LS_MP) strategy produces a sparse probability distribution with higher classification confidence. Experimental results show that the proposed LS_MPCRC not only outperforms other common solutions to SSPP problem, but also has good generalization ability to expression, illumination, little poses variation, occlusion and time variation.

The rest of this paper is organized as follows. We start by introducing the related work in Section 2. Then in Section 3, we present the proposed local structure based multi-phase collaborative representation classification (LS_MPCRC). Section 4 demonstrates experiments and results. Finally, we conclude in Section 5 by highlighting key points of our work.

2. Related work

In the past, face recognition systems are considered to be critically dependent on discriminative feature extraction, about which many approaches have been proposed, such as Eigenface [26], Fishfaces [1], ISOMAP [18], LLE [13], LPP [6] and Laplacian Eigenmap [2]. Recently, Yan et al. [23] proposed a general framework called graph embedding, which generalizes the above-mentioned methods to a unified model within this common framework. Of late, the significance of feature extraction has been debated. Wright et al. [20] have demonstrated that, once the test image can be approximated by a sparse linear combination of the training images, the choice of feature space is no longer critical. In Wright et al.'s pioneer work, a testing sample is first coded as a sparse linear combination of all the training samples via l_1-norm minimization. Then the testing face image is classified to the class which yields the least representation error. The experimental results show that sparse representation based classification (SRC) with random projections-based features can outperform a number of conventional face recognition schemes, such as the nearest-neighbor classifier with Fisher face and Laplacian faces-based features. Later, Zhang et al. [28] propose a collaborative representation based classification method with regularized least squares (CRC_RLS) to further reduce the complexity of SRC. It was indicated that the costly l_1-norm sparse regularization on the representation vector in SRC is not necessary and l_2-norm regularization can lead to similar FR results but with much lower computational cost. Xu et al. [22] proposed the two-phase test sample representation (TPTSR) which can be viewed as a local or sparse version of CRC. However, all the aforementioned methods suffers serious performance drop or even fail to work when encountering *single sample per person* (SSPP) problem.

To overcome SSPP problem, some literatures proposed to use virtual samples or generic training set. Virtual sample generation methods aim to generate extra samples for each person in order to extract the discriminatory information embedded in the intra-personal variations. For example, Shan et al. [17] extended Fisherfaces by generating virtual face images via geometric transform and photometric changes. In [3], Chen et al. proposed to partition each face into a set of sub-images with the same dimensionality and use them as the training samples. Then the traditional FLD-based methods can be applied. In [5], Gao et al. utilized SVD to decompose each face image into two complementary parts: a smooth general appearance image and a difference image. The latter is used to approximate the intra-personal variations. However, these methods are highly limited to the prior information about human face. For generic learning methods, a generic training set, in which each person has more than one training sample, is adopted to extract the discriminatory information. For example, the works [8,11,21] are under this framework. All the generic learning methods are based on the assumption that both the intra-personal variations of different persons and the inter-personal variations of different populations are similar. However,