Journal of Visual Communication and Image Representation 55 (2018) 106–114

Contents lists available at ScienceDirect

Journal of Visual Communication and Image Representation

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

Probabilistic collaborative representation based orthogonal discriminative projection for image set classification☆

Quan Zhang, Huaijiang Sun*

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, PR China

ARTICLE INFO

Keywords:
Image set
Face recognition
Probabilistic collaborative representation
Orthogonal discriminative projection

ABSTRACT

Image set based collaborative representation and classification(ISCRC) has been proposed and achieved state-of-the-art performance. Though ISCRC works well for Image set based face recognition(ISFR), the classification mechanism of ISCRC is still unclear. Besides, another challenge that ISCRC encountered is to deal with the high-dimensional data. In this paper, we first propose a novel Probabilistic Collaborative Representation based Classifier for Image Set (ProCRCIS), which is interpreted from a probabilistic viewpoint. Then, according to the reconstruction residual-based classification rule of ProCRCIS, we propose a novel dimensionality reduction method, called Probabilistic Collaborative Representation based Orthogonal Discriminative Projection for Image Set(ProCR-ODP-IS). The goal of ProCR-ODP-IS is to find a projection space such that the between-class reconstruction residual is maximized and the within-class reconstruction residual is minimized simultaneously. Hence, this projected space can fit ProCRCIS very well. Extensive experimental results on different datasets demonstrate the superiority of the proposed method compared to the state-of-the-arts.

1. Introduction

With the increase of the requirements in social security and the development of digital imaging techniques, image set based face recognition (ISFR) technology has attracted increasing interest in recent years. Collected from video sequences or personal albums [1] etc., the image set can cover more information about a subject than a single image. Compared with image based face recognition (IFR), the biggest difference lies in the fact that both query instance and training instance contain more than one image sample for ISFR [2].

Many ISFR models [2–9] have been proposed. Kim et al. [2] proposed a discriminative learning method for set classification. The image sets are firstly transformed by the discriminant function, then, the similarity between sets were compared by the canonical correlations. Wang et al. [4] modeled each image set as a manifold and defined manifold-manifold distance (MMD) to classify the query image set. In contrast to MMD which is a non-discriminative measure, the methods [7,8] were proposed. The method in [7] exploited statistics information as feature representations for image sets and proposed a localized multi-kernel multi-metric learning method to jointly learn multiple feature-specific distance metrics in the kernel spaces. In [8], a method called sparse discriminative multi-manifold Grassmannian analysis (SDMMGA) was proposed to seek multiple projection matrices, which can uncover the geometrical information of different manifolds. The affine hull and convex hull based methods were also applied to ISFR [3,5,9]. Hu et al.[3] defined a dissimilarity between two sets, which can be sparsely approximated from the image samples of their respective set. The method in [5] characterized each image set by convex or affine hull spanned by its feature points and used geometric distances between convex models to measure the set dissimilarity. However, the methods in [3,5] relied highly on the location of each individual image in the set [9]. In [9], a method called prototype discriminative learning (PDL) was proposed to discriminate each image set with its nearest neighbor prototype. Several LRC-associated approaches [10,6,11] have been proposed. In [6], linear regression models were used for ISFR. Shah et al. estimated regression models for each query image using the class specific gallery set. Based on the minimum reconstruction error between the restricted and original images, a weight voting strategy was used to classify the query set. Different from the distance-based classification criteria [6], Feng et al. [11] proposed superimposed sparse parameter (SSP) classifier, which used the sum of the linear regression parameters for classification.

However, in almost all the works mentioned above, the classification schemes do not consider the information between gallery sets. In recent years, sparse representation based classifier (SRC) has become a powerful tool for pattern recognition[12–17]. In SRC, the query image

☆ This paper has been recommended for acceptance by Yicong Zhou.

* Corresponding author.
E-mail address: sunhuaijiang@njust.edu.cn (H. Sun).

https://doi.org/10.1016/j.jvcir.2018.05.016
Received 23 July 2017; Received in revised form 6 March 2018; Accepted 19 May 2018
Available online 26 May 2018

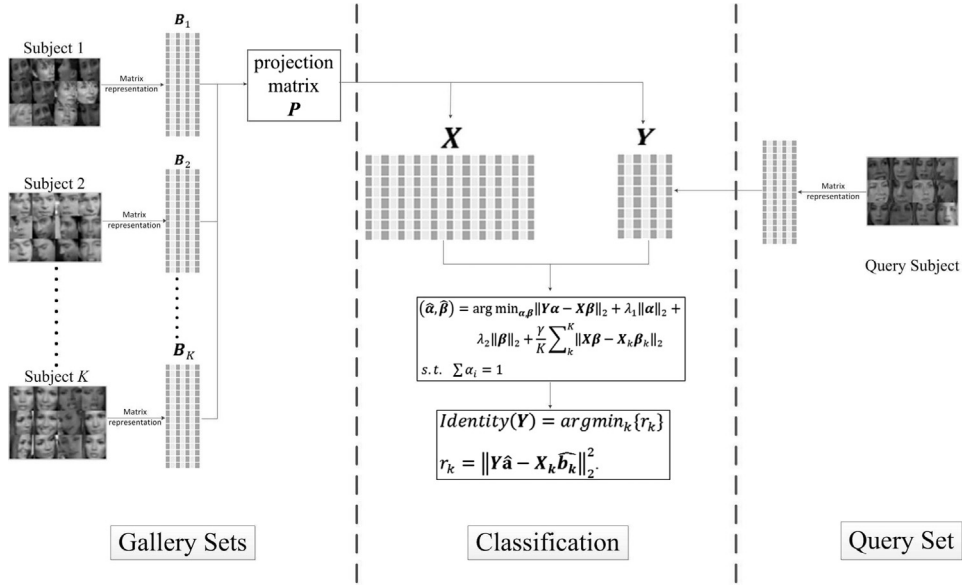

Fig. 1. A block diagram of the proposed method.

is represented by a sparse linear combination of all gallery images, and l_1-norm is used to regularize the representation coefficients [18–21]. Qrtiz et al. [12] presented an end-to-end video face recognition system, Mean Sequence SRC (MSSRC), which used a joint optimization leveraging all of the available video data. These methods based on image sets achieved better performance than ones based on individual images. However, Zhang et al. [22] argued that compared to the l_1-norm-based sparsity constraint, the collaborative representation plays a more important role in the success of SRC. They proposed a new method called collaborative representation based classifier (CRC) which replaces the sparsity constraint with l_2-norm constraint and has less complexity than SRC but is competitive in classification performance. Zhu et al. [20] applied collaborative representation to ISFR and proposed an image set based collaborative representation and classification (ISCRC) scheme. The work in [20] has a drawback while applying KSVD [23] to compress each image set into a dictionary [24]. Li et al. [24] use Laplacian sparse coding as a dictionary learning technique to compress each image set into a dictionary. Despite the fact that ISCRC achieved preferable performance, the classification mechanism of ISCRC still lacks a substantial understanding.

Another great challenge for ISFR is the high-dimensionality of face images. In many real-word applications of face recognition, the high dimensionality face image often brings a series of problems [25,26].On the one hand, it leads to an increase in cost for storage and computing along with curse of dimensionality problem; on the other hand, the solution may be unstable when the dimensionality of feature vectors is larger than the number of training samples for most classifiers. Hence, dimensionality reduction (feature extraction) becomes necessary in the preprocessing of face recognition. Popular dimensionality reduction methods can be divided into linear dimensionality reduction methods (such as principal component analysis (PCA) [27]28, linear discriminant analysis(LDA) [29], etc.) and nonlinear dimensionality reduction methods (such as manifold learning [30], which also includes Locally Linear Embedding(LLE) [31], Locality Preserving Projections (LPP) [32], Sparsity Preserving Projection (SPP) [33], Nonnegative Discriminant Matrix Factorization (NDMF) [34], etc.). Feng et al. [35] proposed a method called projection representation-based classification (PRC). PRC denoted that the 'ideal projection' of a sample point may be gained by iteratively computing the projection with the proper strategy. The pity is that these methods are not directed by classifiers. Yang et al. [36] claimed that the feature extraction method should be bound to classifier. Yin et al. [37] proposed optimized projection for collaborative representation based classification (OP-CRC) which was directed by CRC. OP-CRC maximized the ratio of between-class reconstruction residuals and within-class reconstruction residuals in the projected space and thus enabled CRC to achieve better performance. Up to now, OP-CRC is the most closely connected dimensionality reduction method with CRC. However, it does not consider the relationship between query images.

Motived by the above concerns, in this paper, we propose a probabilistic collaborative representation based classifier for image set (ProCRCIS). The goal of ProCRCIS is to interpret the classifier from a probabilistic viewpoint and define the probability that a query set belongs to each gallery set. According to the decision rule of ProCRCIS, in this paper, we propose a novel feature extraction method, called probabilistic collaborative representation based orthogonal discriminative projection for image set (ProCR-ODP-IS). ProCR-ODP-IS is applied to get a projection matrix by simultaneously maximizing the between-class reconstruction residual and minimizing the within-class reconstruction residual. Therefore, the features have better discriminative power and ProCRCIS can achieve better classification performance in the transformed low-dimensional space. Orthogonality constraint is applied to the discriminative projection vectors, considering the fact that orthogonal projections often demonstrates good performance empirically and has been widely used in dimensionality reduction methods [38,39]. The solution of the proposed method can be efficiently achieved via a trace ratio optimization method [40]. Fig. 1 shows the block diagram of the proposed method. Our experiments on CMU MOBO, YouTube Celebrities, and Extended Yale B datasets show that the proposed method is superior to the compared methods in terms of both accuracy and stability.

The main contributions of this paper are as follows:

(1) We analyzed the image-set based face recognition from a probabilistic viewpoint and proposed a new classifier, Probabilistic Collaborative Representation based Classifier for Image Set (ProCRCIS), which provides a probabilistic interpretation for ISCRC.
(2) Bound to ProCRCIS, we designed a new feature extraction method called Probabilistic Collaborative Representation based Orthogonal Discriminative Projection for Image Set(ProCR-ODP-IS) to solve the dimensionality reduction problem.
(3) An iterative strategy is applied to solving the proposed trace ratio optimization problem and improving the discrimination ability of the algorithm.