Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 52 (2018) 756–764

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Bidirectional feature selection with global and local structure preservation for small size samples

Yongjie Chu[a,*], Yong Zhao[b]

[a] School of Management, Nanjing University of Posts and Telecommunications, Nanjing, China
[b] Institute of Systems Engineering, Southeast University, Nanjing, China

Received 19 August 2018; received in revised form 29 August 2018; accepted 10 September 2018
Available online 20 September 2018

Abstract

Collecting samples is a challenging task for face recognition, especially for some real-world applications such as law enhancement and ID card identification, where there is usually single sample per person (SSPS) used to train a face recognition system. To extract discriminative features from the small size samples, in this paper we propose virtual samples via bidirectional feature selection with global and local structure preservation (VS-BFS-GL) to augment the number of training samples. In VS-BFS-GL, bidirectional feature selection is developed, which introduces $L_{2,1}$ norm to explore the face variations from both horizontal and vertical directions. Further, to include more variations in the virtual images, the global structure information and sample-specified local structure information of the SSPP training set are considered. By integrating bidirectional feature selection, global and local structure, the limited training samples are fully utilized and more knowledge are mined. To further improve the effectiveness of VS-BFS-GL, an auxiliary database containing different face variations can be used to explore the local structure information. We extensively evaluated the proposed approach on AR and FERET database. The promising recognition results demonstrate that VS-BFS-GL is robust to expression, pose and partial occlusion variations in the faces.

Keywords: Bidirectional feature selection; Local structure preservation; Small size samples

1. Introduction

As the wide potential applications in security, video surveillance, human-computer interface, face recognition is has been an important issue in both computer vision and pattern classification. In the situation where multiple face images are available for training a face recognition system, most of the present techniques (Ding, Choi, Tao, & Davis, 2016; Ghazi and Ekenel, 2016; Lu, Liong, Zhou, & Zhou, 2015; Xing and Wang, 2016; Yang et al., 2017; Zhao et al., 2016) are able to achieve satisfying recognition rate. Nevertheless, in some special scenarios, such as e-passport and ID card identification as well as law enforcement, in which only a single face image per person is available for training a face recognition system, those methods designed for addressing multiple training face images are far less capable of solving single sample per person (SSPP) problem. As a result, the SSPP face recognition problem has drawn considerable attention in the last few years and many efforts have been made to tackle it.

* Corresponding author.
E-mail address: yongjiechu@yeah.net (Y. Chu).

https://doi.org/10.1016/j.cogsys.2018.09.009
1389-0417/

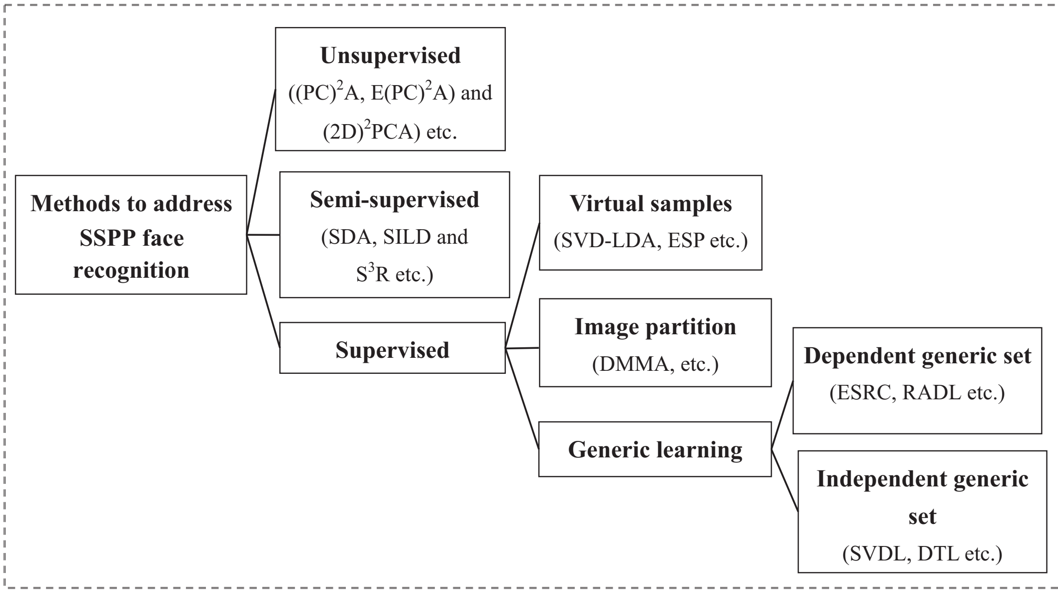

Fig. 1. The classification of methods for SSPP face recognition.

Methods used to settle the SSPP face recognition problem can be classified into three categories: the unsupervised methods, supervised methods, and semi-supervised methods. Fig. 1 illustrates our classification of these methods designed for SSPP face recognition. Though there is only one training image per person available in the training process, the unsupervised techniques, like principal component analysis (PCA) (Chen and Yang, 2007), two-dimensional PCA (2DPCA) (Wang and Wang, 2013), can still be applied to SSPP. These basic unsupervised methods can be heavily affected by variations presented in the probe images, such as expression, occlusion, illumination and pose. Therefore, based on the traditional unsupervised methods, some specifically designed unsupervised approaches are proposed for SSPP face recognition, such as projection-combined PCA ($(PC)^2A$) (Wu and Zhou, 2002), enhanced projection-combined PCA ($E(PC)^2A$) (Chen, Zhang, & Zhou, 2004). Semi-supervised approaches utilize part of label information to infer within-class variations. For example, the side-information (weak label information) is utilized to calculate the within-class and between-class scatter matrices when there is no full class label information (Kan, Shan, Xu, & Chen, 2011). Double linear regressions (DLR) seeks the best discriminating subspace and preserves the sparse representation structure by first propagating the label information to the unlabeled data and then extracting features with the help of the propagated labeled dataset (Yin, Jiao, Shang, Xiong, & Mao, 2014).

To utilize the supervised information and cover more intra-class variations in the training process, many supervised methods have been developed for SSPP problem, which can be further categorized into three subclasses: virtual samples based methods, generic set based methods, and patch-based methods. In the virtual samples based methods, multiple virtual training images per subject are generated from the gallery images, which have the same size as that of gallery images. Generating virtual images can be accomplished by a small SVD perturbation (Zhang, Chen, & Zhou, 2005), some transformations (Shan, Cao, Gao, & Zhao, 2002), or some decompositions (Gao, Zhang, & Zhang, 2008). In patch-based methods, each image in the gallery set is divided into small blocks or patches. Thus there are multiple patches for each subject and intra-class variations can be measured. After obtaining the block images, then liner discriminant analysis (LDA) and self-organizing map (SOM) are used for feature extraction (Chen, Liu, & Zhou, 2004) and (Tan, Chen, Zhou, & Zhang, 2005). Sparse representation is also used to address SSPP face recognition (Gao, Jia, Zhuang, & Ma, 2015), which imposes some sparsity constraints in learning the reconstruction coefficients and the intra-class variance dictionaries. Lu, Tan, and Wang (2013) formulate SSPP face recognition as a manifold-manifold matching problem and to extract features from multiple feature spaces to maximize the manifold margins of different persons, i.e., discriminant manifold-manifold analysis (DMMA). That model is further improved by employing multiple feature descriptors to learn manifolds (Yan, Lu, Zhou, & Shang, 2014). Based on collaborative representation (Zhang, Yang, & Feng, 2011), a patch-based collaborative representation method is proposed by operating collaborative representation on patches and combining the recognition outputs of all patches (Zhu, Zhang, Hu, & Shiu, , 2012). Unlike methods belonging to the first two categories, generic learning based methods adopt a generic training dataset to estimate the intra-personal and inter-personal variations of gallery set. The generic set contains multiple training samples per person and corresponding label information. The advantages of adopting a generic set