Neurocomputing 143 (2014) 134–143

Contents lists available at ScienceDirect

Neurocomputing

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

Multi-feature multi-manifold learning for single-sample face recognition

Haibin Yan [a], Jiwen Lu [b], Xiuzhuang Zhou [c,*], Yuanyuan Shang [c]

[a] Department of Mechanical Engineering, National University of Singapore, Singapore 117576, Singapore
[b] Advanced Digital Sciences Center, Singapore 138632, Singapore
[c] School of Information Engineering, Capital Normal University, Beijing 100048, China

ARTICLE INFO

Article history:
Received 6 November 2013
Received in revised form
15 April 2014
Accepted 2 June 2014
Communicated by Xu Zhao
Available online 17 June 2014

Keywords:
Single-sample face recognition
Multi-feature learning
Multi-manifold learning

ABSTRACT

This paper presents a Multi-feature Multi-Manifold Learning (M^3L) method for single-sample face recognition (SSFR). While numerous face recognition methods have been proposed over the past two decades, most of them suffer a heavy performance drop or even fail to work for the SSFR problem because there are not enough training samples for discriminative feature extraction. In this paper, we propose a M^3L method to extract multiple discriminative features from face image patches. First, each registered face image is partitioned into several non-overlapping patches and multiple local features are extracted within each patch. Then, we formulate SSFR as a multi-feature multi-manifold matching problem and multiple discriminative feature subspaces are jointly learned to maximize the manifold margins of different persons, so that person-specific discriminative information is exploited for recognition. Lastly, we present a multi-feature manifold–manifold distance measure to recognize the probe subjects. Experimental results on the widely used AR, FERET and LFW datasets demonstrate the efficacy of our proposed approach.

1. Introduction

Face recognition has received increasing attention in both academic and industrial communities in recent years, and a large number of face recognition approaches have been proposed in the literature [1–15]. Generally, these approaches can be mainly classified into two categories: geometry-based and appearance-based. Since it is challenging to precisely localize and extract geometrical features from many real facial images especially when they were captured in unconstrained environments, appearance-based methods are more popular in face recognition and many such algorithms have been proposed over the past two decades [1,2,5,8].

The performance of appearance-based face recognition methods is heavily affected by the number of training samples per person [16]. Specifically, if the number of training samples per person is much smaller than facial feature dimension, it is usually inaccurate to estimate the intra-class and inter-class variances for existing appearance-based methods [2,6,8]. In many practical face recognition applications, such as law enhancement, e-passport and ID card identification, there is only a single sample per person registered in these systems because it is generally difficult to capture more additional samples. Therefore, many existing supervised appearance-based methods cannot be directly applied to solve the single-sample face recognition (SSFR) problem due to the non-existence of intra-class samples to estimate the within-class variance. While there have been some attempts to address this problem in recent years [17–27], there is still some room to obtain a further improvement on the recognition performance.

Existing SSFR methods can be classified into four categories: virtual sample generation [18,19], local feature representation [28,20,29–31], generic learning [21,22,32,33], and image partitioning [23,25–27]. For the first category, virtual samples of each training image are generated so that multiple samples per person are obtained for discriminative feature extraction. However, there is high correlation among the virtually generated samples, which make the extracted features are highly redundant. For the second category, each face image is represented by a discriminative feature descriptor so that different persons are expected to be separated as much as possible. These methods require high computational complexity, which may limit their effectiveness in practical applications. For the third category, an additional generic training set with multiple samples per person is employed to extract discriminative features. Even if these methods work around the SSFR problem, the performance of this type of methods is heavily affected by the selected generic training set, which is difficult to construct in practical applications. For the last category, each face image is first partitioned into several local patches and

* Corresponding author. Tel.: +86 10 68902367; fax: +86 10 68416830.
E-mail addresses: eyanhaibin@gmail.com (H. Yan), jiwen.lu@adsc.com.sg (J. Lu), xiuzhuang.zhou@ie.cnu.edu.cn (X. Zhou), syy@bao.ac.cn (Y. Shang).

http://dx.doi.org/10.1016/j.neucom.2014.06.012
0925-2312/

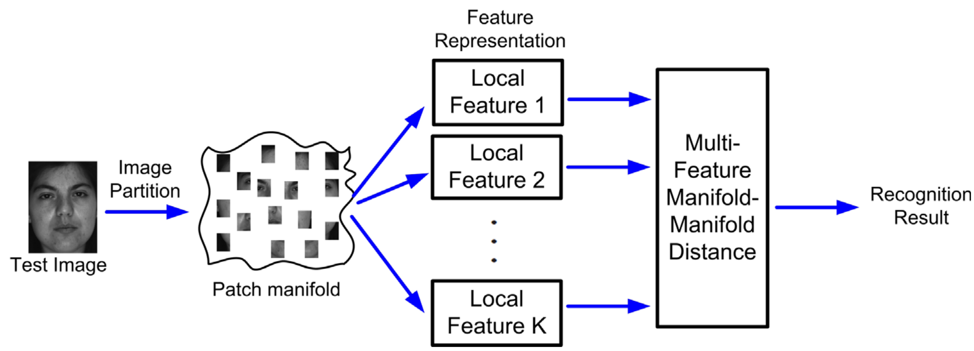

Fig. 1. Scheme illustration of the basic idea of our proposed M3L-based SSFR approach. Given a testing face image, we first partition it into several non-overlapping patches and extract K different local features within each patch. Then, SSFR is converted into a multi-feature manifold–manifold matching problem. Lastly, the minimal multi-feature manifold–manifold distance between the gallery sample and probe sample is used to recognize the testing subject.

then discriminative learning techniques are used for feature extraction [24]. However, methods in this category ignore the geometrical information of local patches because when a face image is partitioned into several local patches, different parts of the original face image cannot be modeled accurately by a simple distribution. It is more likely that these patches reside in a manifold and each patch corresponds to a point in the manifold. Motivated by this intuition, we proposed a discriminative multi-manifold analysis (DMMA) [27] method recently by modeling local patches of each face image as a manifold and performing discriminative manifold matching for SSFR. However, only the raw intensity feature was utilized for facial patch representation, which is not discriminative enough for robust face matching.

Since local features are more robust to raw pixels in face recognition, it is desirable to exploit local features for SSFR under the multi-manifold matching framework. In this paper, we propose a Multi-feature Multi-Manifold Learning (M^3L) method for SSFR. Fig. 1 illustrates the basic idea of our proposed approach. Given each face image, we first extract multiple features for each patch so that they are robust to variations and more complementary information can be exploited. Since different features correspond to different manifolds with different intrinsic dimensions, the SSFR problem is formulated as a multi-feature multi-manifold matching problem. To exploit discriminative information among these manifolds, we maximize the manifold margins by learning multiple discriminative subspaces. Experimental results on three datasets show the efficacy of the proposed approach.

The rest of this paper is organized as follows. Section 2 reviews the related work. Section 3 details our proposed approach. Section 4 provides the experimental results and Section 5 concludes the paper.

2. Related work

A number of manifold learning algorithms have been proposed to discover the intrinsic low-dimensional embedding of the original data in recent years [34–36], and most of them have been successfully applied to face recognition. The basic assumption of these methods is that high-dimensional data can be considered as a set of geometrically related points lying on or nearby a smooth low-dimensional manifold. While encouraging performance can be obtained, these methods simply assume that samples from different classes define a single manifold in the feature space, which may not usually hold in many practical applications because samples from different classes could lie on different sub-manifolds [37]. Inspired by this observation, several multi-manifold learning algorithms have been proposed in recent years [38,39,27], which model samples from the same class as a manifold and the recognition task can be converted as a manifold–manifold matching problem. Existing multi-manifold learning methods assume that data are drawn from a vector space and thus cannot handle multi-feature representations directly. To address this problem, we propose a new multi-feature multi-manifold learning method to learn multiple discriminative feature spaces to maximize the multi-feature manifold margins of different persons. To our best knowledge, this is the first attempt on multi-feature learning in the context of multi-manifold learning. Specifically, our recently proposed DMMA method [27] is the special case of the proposed M^3L when there is only one single feature used in our approach.

3. Proposed approach

In this section, we first formulate the SSFR problem with multi-feature multi-manifold matching, and then present the proposed M^3L method.

3.1. Problem formulation

Let $X=[x_1,x_2,\dots,x_N]$ be the training set, x_i is the training example of the ith person with a size of $p\times q$, $1\le i\le N$, N is the number of subjects in the training set. Assume each face image x_i is divided into r non-overlapping local patches with an equal size of $s\times t$, where $r=p\times q/s\times t$. Let $\mathcal{M}_i^k=[x_{i1}^k,x_{i2}^k,\dots,x_{ir}^k]$ be the image patch set of the ith person extracted by the kth feature representation method, which consists of a manifold $\mathcal{M}_i^k$, and $\mathcal{M}_i=[\mathcal{M}_i^1,\mathcal{M}_i^2,\dots,\mathcal{M}_i^K]$ contains K manifolds of the ith person extracted by K different feature representation methods. Let $\mathcal{M}=[\mathcal{M}_1,\mathcal{M}_2,\dots,\mathcal{M}_N]$ be the training set of N persons and $\mathcal{M}_i^k=[x_{i1}^k,x_{i2}^k,\dots,x_{ir}^k]$ be the kth manifold of the ith person, where $x_{ij}^k\in R^{d_k}$, $1\le i\le N$, $1\le j\le r$, and $1\le k\le K$. Similarly, each test image T is also partitioned into r non-overlapping local patches and the multi-feature manifold $\mathcal{M}_T=[\mathcal{M}_T^1,\mathcal{M}_T^2,\dots,\mathcal{M}_T^K]$ containing K manifolds can be also constructed. Now, SSFR can be formulated as the following multi-feature manifold–manifold matching problem:

$$c=\arg\min_i d_i(\mathcal{M}_T,\mathcal{M}_i) \tag{1}$$

where

$$d_i(\mathcal{M}_T,\mathcal{M}_i)=\sum_{k=1}^{K}\alpha_k d_i(\mathcal{M}_T^k,\mathcal{M}_i^k), \tag{2}$$

$d_i(\mathcal{M}_T^k,\mathcal{M}_i^k)$ is the manifold–manifold distance between $\mathcal{M}_T^k$ and $\mathcal{M}_i^k$, which are the kth manifold of the testing sample and the ith training sample, respectively, $\alpha_k>0$ is the weighting coefficient, and $\sum_k\alpha_k=1$. We see that the multi-feature manifold–manifold