Neurocomputing 184 (2016) 176–187

Contents lists available at ScienceDirect

Neurocomputing

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

Face recognition using part-based dense sampling local features

Jiaqi Zhang [a], Yao Deng [a], Zhenhua Guo [a,b,c,*], Youbin Chen [d]

[a] Graduate School at Shenzhen, Tsinghua University, Shenzhen, China
[b] School of Computer Science and Software, Nanjing University of Information Science & Technology, Nanjing, China
[c] School of Computer Science & Software Engineering, Shenzhen University, Shenzhen, China
[d] School of Automation, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 2 February 2015
Received in revised form
15 May 2015
Accepted 14 July 2015
Available online 17 December 2015

Keywords:
Face recognition
Local Difference Feature
Dense sampling
Cosine similarity

ABSTRACT

For years, researchers have made great efforts to find an appropriate face representation for face recognition. A fusion strategy of Local Binary Pattern (LBP) and Gabor filters yields great achievements. LBP is good at coding fine details of facial appearance and texture, whereas Gabor features can encode facial shape and appearance over a range of coarser scales. Despite the great performance, this fusion representation suffers from low effectiveness and resolution variance. In this paper, we propose a novel representation strategy of face images which is fast and robust to resolution variance. We apply dense sampling around each detected feature point, extract Local Difference Feature (LDF) for face representation, then utilize Principal Component Analysis (PCA)+Linear Discriminant Analysis (LDA) to reduce feature dimension and finally use cosine similarity evaluation for recognition. We have utilized our proposed face representation strategy on two databases, namely self-collected Second Generation ID Card of China and Driver's License (SGIDCDL) database and public Facial Recognition Technology (FERET) database. Our experimental results show that the proposed strategy has good performance on face recognition with fast speed.

1. Introduction

Over the last few decades, face recognition has been an active research topic in computer vision and pattern recognition, and a lot of such systems emerged across various applications, such as video surveillance, access control, image retrieval and automatic log-on for personal computers or mobile devices. Face recognition is an easy task for human visual system, but it is quite challenging for automatic face recognition system due to the dramatic variations among the appearance of the same subject, which is caused by a lot of factors such as image resolution, illumination, expression, pose, occlusion, etc. These various visual complications deteriorate the performance of a face recognition system dramatically.

Automatic face recognition mainly involves three stages: detection, representation and classification. In the detection stage, the face is localized in an image. The representation stage refers to extracting specific features from a given face image. The classification stage deals with the final decisive process whether an unknown face image can be categorized into a target subject. Among the three main parts, face representation occupies a crucial position of a face recognition system. A critical point to the performance of a face recognition system is the face representation's discriminative power for various individuals and invariance to external environment changes, such as resolution and poses variation. A robust and effective face representation is indispensable to performance improvement.

Traditional face representations such as Local Binary Pattern (LBP) [1,2] and Gabor [3,4] achieved good performance on face images with good quality. LBP is good at coding fine details of facial appearance and texture, and Gabor features encode facial shape and appearance over a range of coarser scales. A fusion strategy of LBP and Gabor can obtain promising performance for face recognition [20,22–24]. But this strategy has the problem of complex computations. Meanwhile, the fusion of LBP and Gabor is sensitive to resolution variation.

In addition to efficient face representation, the performance of face recognition system relies on the metric techniques used to measure the similarity of two face representations. The recent frameworks mostly learn a dataset specific metric during the training stage, which can result in the improvement of the recognition performance. It is mainly because it learns the bias of the dataset and thus adapts to the specific distribution. In this paper, we show that our face representation can be used with a very simple face matching method that does not require complex

* Corresponding author at: Graduate School at Shenzhen, Tsinghua University, Shenzhen, China. Tel.: +86 755 26036434; fax: +86 755 26036443.
E-mail address: zhenhua.guo@sz.tsinghua.edu.cn (Z. Guo).

http://dx.doi.org/10.1016/j.neucom.2015.07.141
0925-2312/

metric learning algorithm. Once features are extracted from face images, we first use whitened Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) to project them into a low dimensional space, and then apply cosine similarity on the more compact features to measure the similarity between image pairs.

We experimentally evaluated our proposed face representation on two databases. The first one, which is composed of a subject's high resolution identification (ID) photo and relative low resolution driver's license photo, is collected by us. The second one is the public FERET database [5]. Experimental results show that our method has very good generalization capability and outperforms many traditional methods both in performance and efficiency.

The main contributions of this paper are summarized as follows:

1. We propose a novel face representation strategy for face recognition, which is effective and robust to resolution variation.
2. Experimental results show that the proposed face representation can significantly improve face recognition performance without time consuming Gabor convolution.

The rest of the paper is organized as follows: Section 2 reviews the related work of face recognition. Section 3 describes details of the whole framework and focuses on the proposed new face representation strategy. Section 4 shows experimental results and Section 5 gives conclusion of this paper.

2. Related work

Recent years, advances in the face representation have been a major source of progress in the field of face recognition. Many feature extraction methods have been introduced to characterize face images for various specific tasks, including holistic and local features.

The common goal of these holistic feature extraction methods is to learn a compact and low-dimensional feature subspace for face representation, so that the intrinsic characteristics of face images could be preserved. Among these methods, the most representative ones include principal component analysis (PCA) [6] (Eigenface), Fisher's linear discriminant (FLD) [7] (Fisherface), Locality Preserving Projections (LPP) [8], Marginal Fisher Analysis [9], their kernelized and tensorized variants [10–12]. In spite of different prior assumptions of these methods, they all attempt to find a linear projection matrix to project the image into a low dimensional subspace, so that the image can be unified under a general Graph Embedding (GE) framework [28]. The problem of small number of training samples is addressed in [13,14]. In [15], a sparse representation-based classification (SRC) method was proposed for face recognition which demonstrates striking performance with severe occlusion or corruption. Allen et al. [16] investigated a new solution based on a classical convex optimization framework, known as augmented Lagrangian methods, to recover human identities from high-dimensional facial images that may be corrupted by illumination, facial disguise, and pose variation. In [17], Gao et al. proposed Kernel Sparse Representation (KSR) to take advantage of the non-linear property of kernel trick. Compared with sparse coding, KSR can learn more discriminative sparse codes for face recognition.

The holistic feature extraction methods mentioned above did not make use of the significant spatial information of a face image. To capture the local spatial structure, local feature extraction methods have been developed by encoding different regions of a face image separately. Among the recent developed local features, local binary pattern (LBP) has gained great popularity for its computation simplicity and effectiveness. LBP was first proposed by Ojala et al. [1] for texture classification. The main idea of LBP is that face images can be viewed as a composition of various micro-patterns invariant to monotonic gray scale transformation and can be combined into a global face description. In [2], Ahonen et al. introduced LBP into face recognition for representing both shape and texture information by partitioning the face image into non-overlapping rectangular regions. In [18], Tan et al. extended the idea of LBP to Local Ternary Pattern (LTP), which takes the magnitude of derivatives into account along with its sign and generates a ternary series. Zhang et al. [19] proposed Local Derivative Pattern (LDP) to compute the higher order derivatives of a pixel with respect to its neighbors. Studies have shown that local features gained good performance for face recognition as an effective face representation because of their ability to robustly encode relevant facial traits as well as their computational simplicity.

After that, hybrid face representation is proposed. It considers both global and local features of the face image. In [20], Zhang et al. applied LBP on Gabor feature maps instead of original gray-scale face images. In [21], Zhang et al. developed a novel descriptor, named HGPP, which encodes the phase variation with the orientation changing of Gabor wavelet at a given scale, as well as the connection among local neighborhoods of the Gabor phase information. To eliminate the redundancy of the high dimensional feature vector generated by Gabor wavelet, in [22–24], the authors proposed to incorporate subspace learning algorithms like PCA or FLD, which result in the reduction of the dimensionality as well as the enhancement of the discriminative power.

Although these hybrid algorithms show great performance for face recognition, they have the limitations of high computation complexity and sensitivity to image resolution. To address these problems, we propose a new face representation scheme in the next section.

3. Proposed algorithm for face verification

In this section, we describe the algorithm proposed for face representation. We start from detecting facial feature points with Intraface system [25] on face images, extracting patches with fixed size centered on the detected feature points from face images over different scales, and applying dense sampling strategy to get a fixed number of blocks on each patch. Then on each block we extract histograms of a new proposed local feature, after that PCA+LDA are applied to generate compact descriptors. Finally, a simple cosine similarity based classifier is used to determine the recognition result. The overall framework of face recognition is depicted as Fig. 1.

In the following sub-sections, the paper is organized as follows: first, we present the face preprocessing procedure to eliminate variations of face images. Second, we introduce a LBP-like feature extraction method, namely Local Difference Feature (LDF), which captures the second order texture information from the image. Then, dense sample strategy is elaborated. Finally, a simple cosine similarity based classifier is depicted.

3.1. Face preprocessing

There are different types of variations in facial images, such as color and pose variations. To compensate color variation, a color face image is converted into a gray-scale image first. To reduce the effect of pose, a face image is geometric normalized based on 51 facial feature points detected by Intraface system [25]. Then, the normalized face image is used for the following feature extraction. Fig. 2 illustrates the preprocessing effect.