Pattern Recognition 47 (2014) 509–524

Contents lists available at ScienceDirect

Pattern Recognition

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

An efficient 3D face recognition approach using local geometrical signatures

Yinjie Lei [a,b,*], Mohammed Bennamoun [b], Munawar Hayat [b], Yulan Guo [c,b]

[a] College of Electronics and Information Engineering, Sichuan University, Chengdu, Sichuan 610064, China
[b] School of Computer Science and Software Engineering, University of Western Australia, Crawley, WA 6009, Australia
[c] College of Electronic Science and Engineering, National University of Defense Technology, Changsha, Hunan 410073, China

ARTICLE INFO

Article history:
Received 26 March 2013
Received in revised form
19 June 2013
Accepted 26 July 2013
Available online 7 August 2013

Keywords:
3D biometrics
3D face recognition
3D representation
KPCA
SVM

ABSTRACT

This paper presents a computationally efficient 3D face recognition system based on a novel facial signature called Angular Radial Signature (ARS) which is extracted from the semi-rigid region of the face. Kernel Principal Component Analysis (KPCA) is then used to extract the mid-level features from the extracted ARSs to improve the discriminative power. The mid-level features are then concatenated into a single feature vector and fed into a Support Vector Machine (SVM) to perform face recognition. The proposed approach addresses the expression variation problem by using facial scans with various expressions of different individuals for training. We conducted a number of experiments on the Face Recognition Grand Challenge (FRGC v2.0) and the 3D track of Shape Retrieval Contest (SHREC 2008) datasets, and a superior recognition performance has been achieved. Our experimental results show that the proposed system achieves very high Verification Rates (VRs) of 97.8% and 88.5% at a 0.1% False Acceptance Rate (FAR) for the "neutral vs. nonneutral" experiments on the FRGC v2.0 and the SHREC 2008 datasets respectively, and 96.7% for the ROC III experiment of the FRGC v2.0 dataset. Our experiments also demonstrate the computational efficiency of the proposed approach.

1. Introduction

Face recognition has drawn considerable attention in the last few decades due to its non-intrusiveness and wide number of applications, e.g. surveillance and access control. Although considerable achievements have been attained with 2D face recognition, its accuracy is still challenged by pose and illumination variations [1,2]. With the rapid development of 3D imaging sensors, many researchers have now turned to 3D face recognition due to its potential capabilities to overcome the inherent limitations and drawbacks of 2D face recognition. Moreover, the geometrical information provided by 3D data has the potential to achieve a greater recognition accuracy compared to 2D data [3,4].

Based upon the type of strategy used to measure the similarity between two facial surfaces, the existing 3D face recognition approaches can be categorized into surface registration-based and feature matching-based approaches. The surface registration-based approaches commonly use the Iterative Closest Point (ICP) algorithm [5–7] or one of its variants [8,9]. The ICP error is used as a similarity measure between two matching facial surfaces for recognition. The ICP itself does not require any feature extraction or projection into another space. It is an iterative process and uses the whole surface, which makes it computationally expensive. The second type of approaches (i.e. feature matching-based) match two 3D facial surfaces defined in different coordinate systems based on object-centric shape features [10–13]. Some of them have achieved very promising performance in terms of recognition accuracy. However, despite achieving very high recognition accuracy, many of the existing approaches use complex shape features [14–16,13]. They require complicated mathematical transformations and are therefore not suitable for real-world applications due to the high computational costs associated with them. In this work, we develop an efficient Angular Radial Signature (ARS) to represent a 3D face. It encodes the complex 3D facial surface with a set of 1D feature vectors, which attain sufficient discriminating power while simultaneously achieving significant computational efficiency.

Similar to 2D face recognition, expression variation is also considered to be the main challenge of 3D face recognition since it introduces significant changes in the geometry of the facial surface. Two approaches are usually adopted in the literature of 3D face recognition to handle facial expressions. First, an expression deformable model is learned from facial scans under both neutral and nonneutral expressions. The learned model is then used to morph out the expression deformations from a nonneutral face

* Corresponding author. Postal Address: No. 24 South Section 1, 1st Ring Road, Chengdu, Sichuan 610064, P. R. China. Tel.: +86 028 85505099.
E-mail addresses: leolyj@gmail.com (Y. Lei), bennamou@csse.uwa.edu.au (M. Bennamoun), munawar@csse.uwa.edu.au (M. Hayat), yulan.guo@nudt.edu.cn (Y. Guo).

0031-3203/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.patcog.2013.07.018

[17–19]. The second approach is simply based on the fact that only few facial regions are affected by the distortions caused by expressions and most of the other regions remain invariant. It was proven that by only taking into account the least affected regions, a more accurate and robust face recognition can be achieved compared to the case of using the whole face [20,9,12]. Our expression invariant face recognition system combines the strengths of the above two approaches. First, when using the facial scans with various expressions of each individual for training, the learned model can accommodate expression variations across different individuals. Second, the proposed ARS features are only extracted from the semi-rigid facial region, which can remain relatively robust under the distortions caused by expressions. Third, unlike many existing 3D face recognition approaches which rely on surface registration or complex feature matching between a probe face and all gallery faces, our proposed ARS based approach is computationally efficient.

The rest of this paper is organized as follows. Section 2 describes the related work in the area of 3D face recognition and provides an overview of our proposed approach. Section 3 describes the 3D face normalization algorithm. Section 4 provides a detailed explanation of our proposed ARS facial representation. Section 5 presents the proposed two-stage mapping-based training and testing strategy for 3D face recognition. An experimental performance evaluation and a comparison with the state-of-the-arts are provided in Section 6. Finally, Section 7 comments on and concludes this work.

2. Related work and overview

2.1. Related work

The work in [1,4] provides a comprehensive survey of the 3D face recognition approaches. In the following, we will restrain our review to the approaches which are closely related to this work. More specifically, we cover the basic 3D face recognition approaches which were tested on the Face Recognition Grand Challenge (FRGC v2.0) dataset (Section 2.1.1), approaches developed to handle expression variations (Section 2.1.2) and the relevant 3D facial analysis approaches which were based on machine learning techniques (Section 2.1.3).

2.1.1. Three dimensional face recognition

3D face recognition approaches can be coarsely classified into two categories: surface registration-based and feature matching-based approaches. Chang et al. [6] segmented the 3D face into multiple regions. The regions around the nose of the gallery and probe faces were matched (using ICP) and their matching scores were combined to determine the final recognition results. In [21], Lu et al. constructed a database of 3D mesh models from several 2.5D images and proposed a recognition approach based on the ICP algorithm. To build their 3D meshes, they detected feature points in the 2.5D images, and extracted the maximum and minimum local curvatures. Next, the ICP was applied around these points to align all the 2.5D overlapping regions. Then, the local feature-based ICP registration error was used as a metric to perform matching between faces. In [7], Faltermier et al. presented a 3D face recognition approach based on the detection and registration of multiple small facial regions (some of them overlapping), which were extracted based on specified distances with respect to the location of the nosetip. These regions were then matched independently using the ICP algorithm. They analyzed their face recognition performance under different combinations of these facial regions. They concluded that one combination, which consists of 28 regions (around the nose) yielded the highest recognition accuracy. This indicates that the nose and the region around it are the least affected by expression deformations.

Al-Olsaimi et al. [10] proposed a 3D face recognition approach which integrated global and local geometric cues. They represented the 3D face with multiple rank-0 tensor fields. Then, the local and global fields were integrated into a 2D histogram. The Principal Component Analysis (PCA) coefficients were extracted from these histograms and combined into a single feature vector. The Euclidean distance between a pair of feature vectors was used as a similarity score to match two faces. Berretti et al. [15] proposed a 3D face recognition system that partitioned a 3D face into a set of isogeodesic stripes. Then, a descriptor named 3D Weighted Walkthroughs (3DWWs) was used to represent these stripes, and a graph-based matching algorithm was used to match a pair of faces. Gupta et al. [22] proposed a 3D face recognition approach that could automatically detect 10 anthropometric fiducial points on a 3D face. They used a stochastic pairwise method to calculate the Euclidean and geodesic distances between these points and combined them into a single feature vector. Then, a linear discriminant classifier was used for face recognition.

Most of the existing 3D face recognition approaches rely on a surface registration or on complex feature (surface descriptor) extraction and matching techniques. They are, therefore, computationally expensive and not suitable for practical applications. In this work, we effectively and simply represent a 3D face by a set of 1D feature vectors. Compared with the existing approaches, it can simultaneously achieve high recognition accuracy and computational efficiency.

2.1.2. Expression invariant approaches

The existing 3D face recognition approaches can be categorized, with respect to handling facial expressions, into two groups: (1) those which learn a model for the faces under different expression deformations and (2) those which only extract and process the rigid facial parts. In [17], Kakadiaris et al. proposed a 3D face recognition approach based on an annotated deformable model. They mapped the 3D geometric information onto a 2D regular grid to generate a so called geometric image. Two types of wavelet transforms were then applied on the geometric image for the extraction of features. In [18], Lu et al. developed a hierarchical geodesic-based resampling method to extract landmarks for the modeling of the deformations of a face under expression. Then, a deformable model was constructed for each individual in the gallery. The model is able to handle expressions and pose changes simultaneously. Their experimental results showed that their proposed model is robust under different facial expressions. Al-Olsaimi et al. [19] presented a so called expression model which was learned from pairs of neutral and nonneutral faces of each individual in the gallery. Next, a PCA subspace was constructed using the shape residues of the finely registered pairs of faces. The learned model was then used to morph out any expression deformation from a nonneutral face.

Amongst the approaches which belong to the second category, Zhong et al. [20] divided the 3D face into upper and lower (including the mouth) regions, but only used the upper region. Gabor filters were applied to extract Gabor features, and the centers of the filter response vectors were then learned by K-means clustering. Finally, the recognition results were obtained using a nearest neighbor classifier that was based on a learned visual codebook representation. Mian et al. [9] proposed an approach which automatically extracted two semi-rigid regions from a 3D face. Then, the extracted regions were matched individually using a modified ICP algorithm. The final matching score was obtained by combining the minimum Mean Squared Error (MSE) from each region. Lei et al. [12] developed a binary