Neurocomputing 129 (2014) 445–457

Contents lists available at ScienceDirect

Neurocomputing

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

Rigid-area orthogonal spectral regression for efficient 3D face recognition ☆

Yue Ming *

Beijing Key Laboratory of Work Safety Intelligent Monitoring, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, PR China

ARTICLE INFO

Article history:
Received 1 November 2012
Received in revised form
9 September 2013
Accepted 29 September 2013
Communicated by Jinhui Tang
Available online 7 November 2013

Keywords:
3D face recognition
Facial rigid area
Orthogonal Spectral Regression
Curvature
Appearance-based methods

ABSTRACT

A new framework is proposed for 3D face recognition, called Rigid-area Orthogonal Spectral Regression (ROSR). We utilize the depth images of 3D facial rigid area for efficiently discriminant feature extraction. The framework can effectively estimate the regression matrix to describe intrinsic facial surface features. Large expressions, treated as non-rigid transformations, along with data noise, are the major obstacles that significantly deteriorate the facial linear structure. In our framework, we first utilize the curvature information to remove the non-rigid areas in the 3D face images. Orthogonality can minimize the reconstruction errors and Spectral Regression can accurately describe the manifold structure of the samples. We take these advantages into consideration and propose the ROSR framework, employed for 3D face recognition. Additionally, regression analysis is much faster than the traditional methods. CASIA, Bosphorus and FRGC 3D face databases are introduced for experimental evaluation. Compared with the other commonly used algorithms, our framework has a consistently better performance in terms of efficiency and robustness.

1. Introduction

Face recognition is a continuously developing field receiving lots of attention [1]. So far, most researchers have focused on recognizing 2D images and only a few of them have utilized the depth images converted from the 3D scans [2–4]. Although 2D face recognition has solved some difficulties in dealing with illumination variations to some extent, 2D projection of 3D non-rigid objects is quite sensitive to expression and pose variations, especially in uncontrolled environments [5]. The experiments established by [6] indicate that the performance on the 3D shape channel is better than on 2D texture alone. Extensive research efforts [7–9] have been dedicated to 3D information retrieval and recognition. With the development of 3D digital scanners and sensors [10], utilizing 3D facial information is possible for effective face recognition and discriminant feature extraction [11–13]. 3D facial surface data provides a promising way to overcome the influence of large poses and expressions. The technology of annotation and signature based on the range images can also precisely describe facial surface shape and geometric properties [14–16]. The internal anatomical structure provided by the facial shape can improve face recognition performance significantly instead of 2D facial appearance with external intensity [17,18].

3D face recognition has been successfully employed in numerous recent studies. Lee et al. [19] utilized curvature information to divide the original 3D face models into multiple convex areas. Then a similarity metric, defined on the Extended Gaussian Images (EGI) between two regions, was used to match the pairwise images. Gordon [20] developed facial surface representation for recognition with the depth and curvature information and evaluated the performance on 24 faces. Curvature, as a feature descriptor, also converted face recognition into the rigid shape recognition problem [21]. Chua et al. [22] extended a facial representation, called point signature, as a free-form surface to recognize 3D face images with the different expressions. For 3D face verification, Beumier et al. [23] extracted the facial surface profile as the discriminant feature of the depth images. A fusion scheme of 3D shape and 2D texture information is another active field for face recognition [11]. Chang et al. [24] simultaneously applied PCA to 3D and its corresponding 2D face images and demonstrated the improvements. Based on these studies, curvature, as a surface shape representation, is quite helpful in reflecting the facial rigid areas. However, all of the above algorithms are based on a small sample size from the 3D face databases.

☆This work was done when Yue Ming visited Carnegie Mellon University. The work presented in this paper was supported by the National Natural Science Foundation of China (Grants no. NSFC-61170176), Fund for the Doctoral Program of Higher Education of China (Grants no. 20120005110002), Beijing Municipal Commission of Education Build Together Project, CCF-Tencent Open Research Fund and Principal Fund Project. Portions of the research in this paper use the CASIA-3D FaceV1 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA).

* Tel.: +86 13269066919.
E-mail address: myname35875235@126.com

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.09.014

When the data size is significantly increased, the performance will be significantly lowered. Thus, finding an effectively compact and discriminant feature has become an essential problem to be solved in this area.

Recently, a lot of databases, that include many challenging images, have been available for 3D face recognition evaluation, particularly the CASIA [25], Bosphorus [26] and FRGC [6] 3D face databases. Faltemier et al. [27] introduced 3D face recognition by segmenting a face into a group of regions. By independently matching the different regions, the fusion of results led to a better performance of 3D face recognition. An annotated face model (AFM) used by Kakadiaris et al. [28] was proposed to align the facial shape changes and wavelet features based on the deformed images and it demonstrated a robust recognition performance. Mian et al. [29] dealt with expression variations with a multimodal fusion framework and showed the potential of solving the problem of expressions by combining appearance and shape information. Passalis et al. [30] focused on 3D face recognition in real-world scenarios. They employed an automatic landmark detector to overcome the influence of occluded areas and an annotated face model for fitting the scan. The experimental results show that facial symmetry has a significant advantage for handling pose variations. Ocegueda et al. [31] presented a Markov Random Field model for the analysis of 3D face meshes to obtain the discriminative information. Drira et al. [32] proposed a novel geometric framework for comparing, matching, and averaging 3D face shapes. The representation based on radial curves and elastic shape analysis can analyze shapes of full facial surfaces, which is robust to pose variations and missing. But it was sensitive to expressions with open mouths. Berretti et al. [33] introduced isogeodesic strips and 3DWWs for representing facial local morphology, which exhibits the smooth variations for facial expressions. Mohammadzade and Hatzinokos [34] evaluated the performance of their proposed iterative closest normal point algorithm, which can enable effective application of discriminant analysis. Although these algorithms have achieved good results for the large challenging databases, all of them require complex computation and are difficult to implementation.

Though a huge number of schemes have focused on 3D face recognition, there still exist some unsolved problems that would improve their performance, such as hair occlusions, expressions and spikes. In the past few decades, most researchers have been concerned with 2D face recognition [1]. Several methods have become quite accurate. From the perspective of the imaging principle, in the front viewpoint, 2D gray values are sensitive to illumination and require complex algorithms to solve. However, 3D depth images can describe the actual depth changes and are quick to recognize lighting variations. The literature supports introducing some mature 2D appearance-based recognition methods combined with distance information for 3D images [29]. Although the shape information is more robust to the changes caused by pose and lighting, it is still sensitive to expressions and occlusions [6]. To address the above challenges, we utilize the rigid area of the face to overcome the facial expressions. We mainly focused on solving the influence of expression variations based on facial rigid-area depth information.

Currently, numerous geometrically inspired methods have been explored for visual analysis. Considerable work has focused on seeking a compact low-dimensional feature to describe the intrinsic sample spaces, such as two traditional methods, PCA [35] and LDA [36]. The facial shape and geometrical characteristics on the sub-manifold spaces are essential for discriminant feature extraction. Many subspace learning approaches have been developed for this field, including LPP [37], NPE [38] and OLPP [39].

PCA [35] minimized the reconstruction errors along the direction of maximum variance. LPP [37] modeled the data as the local manifold structure. OLPP [39] shared the same locality preserving characteristics as LPP, but simultaneously it required the basic functions to be orthogonal. Moreover, an empirical study [39] showed that orthogonality was essential for finding optimal feature vectors in representation and discrimination. However, the OLPP algorithm resulted in a huge computational cost. Additionally, these algorithms run in an unsupervised setting, making it difficult to reflect the label information.

In statistical machine learning, a regression analysis has been introduced to construct the bridge between response and explanatory variables. Spectral Regression (SR) [40] showed the connection between regression and graph embedding [37], which efficiently combined the label information into the response variable. SR only required a linear complexity for the regression matrix estimation. Moreover, SR can extract the discriminant feature vectors when the data contains additive Gaussian noise or corruptions as long as the magnitude is not quite large. However, as discussed above, SR without orthogonal constraints cannot obtain the optimal regression matrix in the sense of the minimal reconstruction errors and thus is not suitable for 3D depth data.

To efficiently describe the intrinsic surface of the 3D data, we impose an orthogonal constraint on the SR framework and estimate the optimal regression matrix of an orthonormal discriminant basis, called Orthogonal Spectral Regression. Based on the subspace embedding, a group of compact feature vectors can be obtained, which can reflect the internal anatomical structure among the different individuals. Fig. 1 provides the proposed framework for 3D face recognition. To better extract the discriminant features and decrease the influence of non-rigid regions for a regression estimation, our framework can be divided into two major parts: 3D facial rigid area segmentation and discriminant feature extraction. Since the noises and facial expressions are two major challenges [41], we extract the facial rigid regions from the original 3D scatter point clouds based on curvature information [42]. Facial rigid areas have the distinctive characteristics of linear structure and are insensitive to expression variations. Based on the rigid area, a human face can be referred to by a limited length of its feature vector. To better reflect the geometric information of 3D facial manifold structure, OSR is used to determine the effective regression matrix. Computational complexity analysis shows that the estimation process allows a huge savings in terms of both computational cost and storage space. Finally, different metrics can be incorporated into nearest neighbor classifiers for evaluation in the recognition and verification scenarios.

The main contribution of this work is two-fold. First, facial rigid-area extraction combined curvature information is introduced to overcome the influence caused by large expression variations and registration errors. Second, Orthogonal Spectral Regression describes the manifold structure of the samples with orthogonal constraint. The proposed algorithm can effectively reflect facial geometric relationship and is much faster than the traditional methods. Then, we take the two above advantages into a whole ROSR framework for 3D face recognition, which reveals the efficiency, generality and robustness based on the different challenging 3D face databases.

The rest of this paper is organized as follows: Section 2 introduces the 3D facial rigid area segmentation using curvature

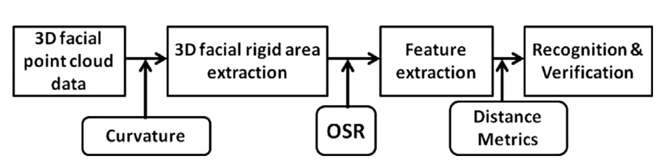

Fig. 1. The flow chart of our proposed 3D face recognition framework.