Information Sciences 320 (2015) 406–417

Contents lists available at ScienceDirect

Information Sciences

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

Bayesian multi-distribution-based discriminative feature extraction for 3D face recognition

Ronghua Liang *, Wenjia Shen, Xiao-Xin Li, Haixia Wang

College of Information Engineering, Zhejiang University of Technology, Hangzhou, China

ARTICLE INFO

Article history:
Received 18 June 2014
Received in revised form 21 March 2015
Accepted 27 March 2015
Available online 4 April 2015

Keywords:
3D face recognition
Single training sample per person
Depth image
Bayesian learning

ABSTRACT

Due to the difficulties associated with the collection of 3D samples, 3D face recognition technologies often have to work with smaller than desirable sample sizes. With the aim of enlarging the training number for each subject, we divide each training image into several patches. However, this immediately introduces two further problems for 3D models: high computational cost and dispersive features caused by the divided 3D image patches. We therefore first map 3D face images into 2D depth images, which greatly reduces the dimension of the samples. Though the depth images retain most of the robust features of 3D images, such as pose and illumination invariance, they lose many discriminative features of the original 3D samples. In this study, we propose a Bayesian learning framework to extract the discriminative features from the depth images. Specifically, we concentrate the features of the intra-class patches to a mean feature by maximizing the multivariate Gaussian likelihood function, and, simultaneously, enlarge the distances between the inter-class mean features by maximizing the exponential priori distribution of the mean features. For classification, we use the nearest neighbor classifier combined with the Mahalanobis distance to calculate the distance between the features of the test image and items in the training set. Experiments on two widely-used 3D face databases demonstrate the efficiency and accuracy of our proposed method compared to relevant state-of-the-art methods.

Published by Elsevier Inc.

1. Introduction

Face recognition has received considerable attention from the scientific and industrial communities over recent decades. Although 2D face recognition technologies have been extensively studied [46] and are effective under constrained conditions, the variations in illumination, pose and expression still present challenges to existing recognition approaches [1,46]. In this respect, 3D face recognition can be advantageous, since a 3D face model represents the geometry underlying the face image and is thus independent of the ambient illumination and viewpoint [30].

3D face recognition methods can be roughly classified into two categories: spatial matching methods and feature-based matching methods [30]. The former match 3D faces directly by comparing the surface similarity; they include Hausdorff distance [2,17,18,36], iterative closest point (ICP) and its extension [9,32,47]. The main problem of spatial matching methods lies in their high computational cost since 3D face models are usually very large, frequently containing millions of points.

* Corresponding author. Tel./fax: +86 571 85290565.
E-mail address: rhliang@zjut.edu.cn (R. Liang).

http://dx.doi.org/10.1016/j.ins.2015.03.063
0020-0255/Published by Elsevier Inc.

As we are seeking computational efficiency, we shall concentrate on feature-based matching methods which are more efficient as they generally describe the 3D models as multiple 2D features [14–16,19,27,44] – commonly utilized 2D features include raw depth images and surface properties such as gradient and curvature [30].

The surface properties, especially curvature, are widely exploited from different views [5,11,35,39]. However, surface properties are difficult to exploit as the discrete surface property values are calculated from 3D face models, and are thus affected by different levels of noise in the acquisition settings. To enhance the robustness, pre-processing methods such as key point extraction and region segmentation have to be used to deal with the surface properties [29,31], which increases the complexity of the algorithms.

Depth images are usually less subject to noise [30], are usually robust to changes of illumination and viewpoint, and have good computational efficiency [21]. As a result, depth images have been widely used in 3D face recognition, frequently being integrated with various 2D recognition methods [13,20–23,33,34]. Achermann et al. [3] conducted early work on PCA-based depth images, while Lin et al. [26] extracted semi-local summation invariant features from a rectangular region surrounding the nose of a 3D facial depth image. Tsalakanidou et al. [40] and Cook et al. [12] both used depth images combined with intensity information – Tsalakanidou et al. [40] constructed embedded hidden Markov model techniques for face recognition, whereas Cook et al. [12] presented a method based on Log-Gabor templates which were insensitive to variations in expression. Further, Queirolo et al. [33] used a simulated annealing-based approach for depth image registration, with surface interpenetration as a similarity measure, in order to match two face images.

However, a disadvantage of methods based on depth images is that they may lose important 3D face features, such as information on the geometric structure; depth images from different subjects can be similar, which increases the difficulty of recognition. Moreover, for 3D face models, the number of training samples from the same subject is usually very small, as the collection and storage costs of 3D samples are high [38]. In fact, the popular 3D-face databases 3D_RMA [6] and CASIA HFB [25] each contain only a single sample per person (SSPP), which provides insufficient discriminative features from that individual [37]. These issues increase the difficulty of performing accurate recognition.

To address the two problems mentioned above, this paper proposes a novel discriminant analysis method for robust 3D face recognition. To solve the SSPP problem, we divide each depth image into several patches to enlarge the sample size per person. For the problem of similarity of depth image from different subjects, we build a Bayesian multi-distribution model which imposes different distributions on intra-classes and inter-classes, respectively, and hence extract the discriminative features from depth image patches. Experiments on two widely-used 3D face databases have validated the efficiency of the approach proposed.

The main contributions of our proposed method are as follows.

(1) An approach based on a depth image obtained from the 3D face model is adopted for the sake of computational efficiency. By dividing the depth image into local patches, the size of the training sample is enlarged; as a result, we can generate enhanced intra-class information for feature extraction.
(2) Our feature extraction method can identify discriminative features from 3D depth image patches by use of discriminative Bayesian multi-distribution analysis. By this means, we can detect increasingly discriminative inter- and intra-class information related to an individual and thus enhance the recognition rate.

This paper is organized as follows. In Section 2, we describe the motivation of our method, Section 3 introduces the method in detail, and Section 4 presents the experimental results. We summarize our conclusions and indicate our directions for future work in Section 5.

2. Motivation

The SSPP problem mentioned in Section 1 is associated with two main challenges: how to represent the 3D face model with a reduced-dimension method, and how to uncover more discriminative features from the small-sized sample.

2.1. Depth images

Feature-based methods for 3D face recognition typically use depth images to represent the 3D face models. These images are extracted from projections of the 3D models and can provide depth information (i.e., distance from the clipping plane) as well as silhouette information of the 3D models [30]. Being based on 2D, this approach provides greater computational efficiency than 3D spatial matching methods [8]. Furthermore, a depth image is independent of illumination and pose so is advantageous over other 2D representations of 3D faces.

However, 3D face recognition based on depth images has limitations [30] as some 3D face information is lost during the projection that creates the depth image. Fig. 1 presents a 3D model and the corresponding depth image from two different subjects. It can be seen that the depth images of different subjects are similar, which illustrates the difficulties associated with precise recognition from a depth image.

In overcoming these limitations, our method addresses two issues: acquiring a more reliable depth image from the original 3D face model databases is described below; investigating discriminative features between different subjects is discussed in Section 2.2.