3D Face Recognition Using Range Images

Literature Survey Report

Multidimensional DSP

Spring 05

Joonsoo Lee

Department of Electrical and Computer Engineering

The University of Texas, Austin, TX

Email: jslee509@mail.utexas.edu

Abstract

This literature survey will evaluate face recognition algorithms based on range images of human faces. Range images have several advantages over 2D intensity images and 3D meshes. Range images are robust to the change of color and illumination, which causes a significant problem in face recognition using 2D intensity images. Also, 3D information from range images is much easier to utilize than that from 3D meshes. Some previous approaches for face recognition using range images are focused on the data acquisition and preprocessing stage. This project will focus on the recognition stage itself. The previous feature extraction approaches can be divided into two types: geometrical and statistical. The geometrical approach utilizes the curvature information of face surface and requires additional calculation of specific descriptors. The statistical approach reduces the dimension of feature space largely by linear projection, but requires accurate pre-processing because it is sensitive to noise. The detailed algorithms of previous approaches will be compared and a new method which can improve the performance and reduce the computational complexity will be proposed.

I. Introduction

Human face recognition has received a great deal of attention in recent times and emerged as an active research area in response to numerous law enforcement and commercial applications [1]. The objective of face recognition is to develop an automatic system that can recognize the human face as humans do. The two important terms in the previous statement are *automatic* and *as humans do*. It means a good face recognition system should require as little manual intervention as possible but does not need to be able to distinguish faces which even people cannot distinguish.

With recent advances in image capture techniques and devices, various types of face-image data have been utilized and various algorithms have been developed for each type of an image [2]. Among various types of face images, a 2D intensity image has been the most popular and common image data used for face recognition because it is easy to acquire and utilize (Fig. 1). It, however, has the intrinsic problem that it is vulnerable to the change of illumination. Sometimes the change of illumination gives more difference than the change of people, which severely degrades the recognition performance. Therefore, illumination-controlled images are required to avoid such an undesirable situation when 2D intensity images are used. To overcome the limitation of 2D intensity images, 3D images are being used, such as 3D meshes and range images (Fig. 1). A 3D mesh image is the best 2D representation of 3D objects. It contains

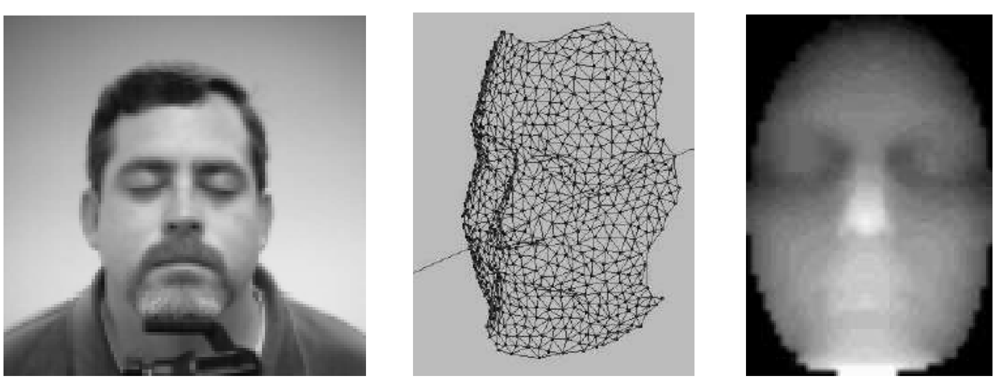

Fig. 1. Face images; 2D intensity image, 3D mesh image, and range image [3]

3D structural information of the surface as well as the intensity information of each point. By utilizing the 3D structural information, the problem of vulnerability to the change of illumination can be solved. A 3D mesh image is a suitable image data for face recognition, but the data is complex and difficult to handle. A range image can be a good alternative to a 3D mesh image. A range image contains the structural information of a face and also is simple to utilize for face recognition.

II. Background

A range image is simply an image with depth information as shown in Fig. 1. In other words, a range image is an array of numbers where the numbers quantify the distances from the focal plane of the sensor to the surfaces of objects within the field of view along rays emanating from a regularly spaced grid [4]. For example, a nose tip is the closest point to the camera on a face, so it has the highest numerical value. Range images have some advantages over 2D intensity images and 3D mesh images. First, range images are robust to the change of illumination and color because the value on each point represents the depth value which does not depend on illumination or color. Also, range images are simple representations of 3D information. The 3D information in 3D mesh images is useful in face recognition, but it is difficult to handle. Different from 3D mesh images, it is easy to utilize the 3D information of range images because the 3D information of each point is explicit on a regularly spaced grid. Due to these advantages, range images are very promising in face recognition.

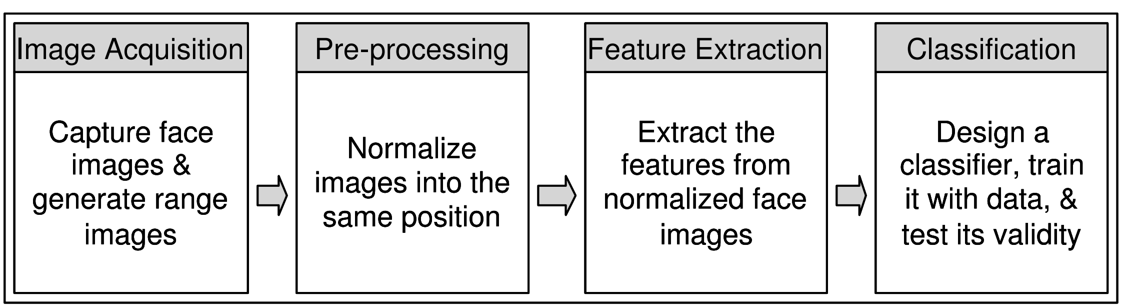

Fig. 2. Block diagram of a face recognition system

Usually, a procedure for face recognition using range images is composed of four steps (Fig. 2). First, face images are captured by appropriate image sensors. Normally, a 3D mesh image is captured by a 3D camera such as a Minolta Vivid 700 camera and a range image is generated from the 3D mesh image [3]. Then, range images are normalized to correct the difference of rotation, translation, and depth. Once all images are normalized, feature extraction algorithms are applied to range image data. The last step is to design a classifier by using the extracted features. The nearest neighborhood (NN) classifier and support vector machine (SVM) are commonly used for classification. This report will focus on the feature extraction step. Various algorithms for feature extraction will be analyzed and compared to each other. Basically, the previous approaches on feature extraction can be divided into two categories: geometrical and statistical.

III. Geometrical Approach

A. Principal Curvature

In this approach, a face recognition problem is considered as a 3D object recognition problem. Besl and Jain [4] studied the 3D object recognition using range images. They calculated Gaussian curvature and mean curvature and used the signs of these surface curvatures to classify range image regions. Based on this 3D object recognition problem, Gordon devised a solution for a face recognition problem using range images [5]. At each point P on the surface, a curve is formed by the intersection of the surface and the normal plane in a given tangent direction $\vec{t_i}$. The curvature of this planar curve is the normal curvature κ_n at P in the direction $\vec{t_i}$. The maximum and minimum normal curvatures at a point define the principal curvatures, κ_{max} and κ_{min}. The Gaussian curvature K, at a point, is defined as the product $\kappa_{max}\kappa_{min}$, and the mean curvature H is $(\kappa_{max} + \kappa_{min})/2$. From these curvature maps, a map of ridge lines or valley lines in a face is generated and several face-specific features, such as eyelids, eyeballs, and noses, are extracted from the line maps. Then, a simple brute force depth comparison recognition strategy is used for face recognition.

This approach shows an outline of the use of curvature information in the process of face recognition. It shows that a great deal of information about facial features that cannot be seen from intensity images is contained in the curvature maps, but it does not explicitly show how to utilize this information to extract specific features. Also, this approach can deal with faces different in size, but needs extension to cope with changes in facial expression.

B. Spherical Correlation

Tanaka and Ikeda [6] presented a better approach to using curvature information from range images. They also viewed a face recognition problem as a 3D shape recognition problem and improved previous approaches which only used the signs of curvatures for classification. First, they analyze face structure based on 3D principal curvatures and their directions from range images. Next, they extract convex-concave points with high curvature values as the discriminating features effective in curved-surface recognition. Then, they construct extended Gaussian images (EGI) of feature vector sets by mapping maximum and minimum principal directions on two unit spheres and use them as face representations. Finally, similarities among faces are measured using Fisher's spherical correlation on EGI's of faces. The EGI is a collection of impulses on the unit sphere [7]. Each face corresponds to an impulse with weight equal to its area, at a place on the sphere where the tangent is parallel to the face (Fig. 3). This representation is unique at least in the case of convex objects. A Fisher's spherical correlation coefficient ρ is a measure of similarity and defined as follows,

$$\rho = \frac{det\{E(X \cdot Y_t)\}}{\sqrt{det\{E(X \cdot X_t)\}det\{E(Y \cdot Y_t)\}}}, \tag{1}$$

where X and Y are two sets of n dimensional unit vectors. A pair of EGI's of ridge and valley line vectors on each face surface is used independently to evaluate the two types of similarity measures ρ_{ridge} and ρ_{valley}. Total similarity ρ is obtained as their product $\rho = \rho_{ridge} \times \rho_{valley}$.

This approach is the first work to investigate and evaluate free-formed curved-surface recognition.

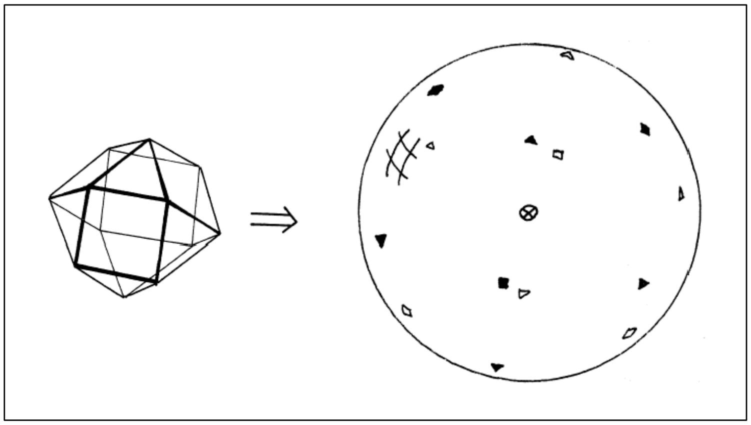

Fig. 3. Extended Guassian image [7]

Also, it is simple, efficient, and robust to distractions such as glasses and facial hair, but it hasn't been tested on faces in different sizes and facial expressions.

IV. Statistical Approach

A. Eigenface

A standard procedure for the analysis and recognition of 2D face images is the eigenface method described by Turk and Pentland [8]. They consider face images as vectors and apply principal component analysis (PCA) to get eigenfaces of the set of training images. Test images are first projected into the face space, and then it is determined which person's training image is the most similar in the face space. Achermann *et al*. [9] and Hesher *et al*. [3] applied this method to a set of range images. This method is optimal in the least mean square error sense and has been proved to perform well with 2D intensity images.

The advantage of this method is the large dimension reduction of feature space. But, it shows bad performance with a large database because it is badly affected by outliers, which appear more often when the size of the database gets larger.

B. Optimal Linear Components

There are several linear projection techniques which reduce the dimension of feature space, such as PCA, independent component analysis (ICA), Fisher's discriminant analysis (FDA), and so on. The performance of these techniques are compared in [10]. The performance of each technique varies according to the size of data and the variability in the data, such as facial expression change and pose variation. Therefore, it is hard to tell which technique is the best overall. Liu *et al*. [11] presented a technique for finding linear representations of images that are optimal for specific tasks and specific data sets, instead of choosing a standard projection technique. They also extended this work to range images [12].

The basic task in this approach is to find the optimal *k*-dimensional subspaces of R^n, where *n* is the size of an image and *k* is the desired dimension of feature space. Then, the problem of finding optimal linear subspaces for recognition becomes an optimization problem. [11] describes a numerical procedure for approximating the solution using a stochastic gradient algorithm. The basic idea is to construct a Markov chain that finds the points where the measure function has high values. It does so by using randomly-perturbed versions of the gradient directions to find candidates for updating the chain and these candidates are accepted and rejected according to a probability that depends upon the measure function.

By choosing the optimal projection, this approach shows much better performance than other standard projection techniques, but it requires a significant amount of computation because it includes an optimization problem and does not have a closed-form solution.

V. Comparison of previous approaches

The geometrical approaches utilize the curvature information of face surfaces and the statistical approaches use linear projection techniques. Therefore, the geometrical approaches don't need any training for recognition while the statistical approaches require some training of the classifier before the classification of test images. However, the geometrical approaches require several processing steps before the classification because they need to define specific face descriptors from the curvature information. The statistical approaches show great performance in the sense that they largely reduce the dimension of feature space, but they require accurate pre-processing because they are sensitive to noise and outliers. The comparison of previous approaches is presented in Table I. The recognition accuracy is not presented because they were tested on different databases, so it is meaningless to compare.

approach	training	complexity	different size	different expression
principal curvature	no	low	yes	no
spherical correlation	no	medium	no	no
eigenface	yes	medium	yes	no
optimal linear projection	yes	high	yes	yes

TABLE I

COMPARISON OF PREVIOUS APPROACHES

VI. CONCLUSION

Range images have several advantages over other image data for face recognition. Range images are invariant to the change of illumination and color and also represent the 3D information of face surface. There have been two types of approaches for face recognition using range images: geometrical and statistical. The geometrical approach utilizes the curvature information of face surface and requires additional calculation of specific descriptors. The statistical approach reduces the dimension of feature space largely by linear projection, but requires accurate pre-processing because it is sensitive to noise.

To improve the performance and decrease the complexity, I propose a new method combining advantages of both approaches. The statistical approach is good at reducing the dimension of feature space, but doesn't utilize any specific information of a face. I propose to use the curvature information which can be acquired by applying geometrical approaches in the process of linear projection. By doing this, I can reduce the complexity of linear projection, as well as improve the recognition accuracy.

References

[1] R. Chellapa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. of the IEEE*, vol. 83, no. 5, pp. 704–740, May 1995.

[2] K. Chang, P. Flynn, and K. Bower, "Multi-modal 2D and 3D biometrics for face recognition," in *Proc. IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures*, Oct. 2003, pp. 187–194.

[3] C. Hesher, A. Srivastava, and G. Erlebacher, "A novel technique for face recognition using range imaging," in *Proc. IEEE Int. Symposium on Signal Processing and Its Applications*, July 2003, pp. 201–204.

[4] P. Besl and R. Jain, "Invariant surface characteristics for 3D object recognition in range images," *Computer Vision, Graphics, and Image Processing*, vol. 33, pp. 33–80, 1986.

[5] G. Gordon, "Face recognition based on depth maps and surface curvature," *Geometric Methods in Computer Vision*, vol. SPIE 1570, pp. 1–12, July 1991.

[6] H. Tanaka and M. Ikeda, "Curvature-based face surface recognition using spherical correlation - principal directions for curved object recognition," in *Proc. IEEE Int. Conf. on Automated Face and Gesture Recognition*, Apr. 1998, pp. 372–377.

[7] B. Horn, *Robot Vision*. The MIT Press, 1986.

[8] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[9] B. Achermann, X. Jiang, and H. Bunke, "Face recognition using range images," in *Proc. IEEE Int. Conf. on Virtual Systems and Multimedia*, Sept. 1997, pp. 129–136.

[10] J. Li, S. Zhou, and C. Shekhar, "A comparison of subspace analysis for face recognition," in *Proc. IEEE Int. Conf. on Acoust., Speech, Signal Processing*, July 2003, pp. 121–124.

[11] X. Liu, A. Srivastava, and K. Gallivan, "Optimal linear representations of images for object recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, no. 5, pp. 662–666, 2004.

[12] A. Srivastava, X. Liu, and C. Hesher, "Face recognition using opimal linear components of range images," *Journal of Image and Vision Computing*, accepted in 2003, http://www.cavis.fsu.edu/.