3D FACE RECOGNITION BASED ON 3D RIDGE LINES IN RANGE DATA

Mohammad H. Mahoor and Mohamed Abdel-Mottaleb *

Department of Electrical and Computer Engineering,University of Miami
1251 Memorial Drive, Coral Gables, FL 33146

Emails: mmahoor@umsis.miami.edu and mottaleb@miami.edu

ABSTRACT

In this paper we present an approach for 3D face recognition from range data based on the principal curvature, k_{max}, and Hausdorff distance. We use the principal curvature, k_{max}, to represent the face image as a 3D binary image called ridge image. The ridge image shows the locations of the ridge lines around the important facial regions on the face (i.e. the eyes, the nose, and the mouth). We utilize Hausdorff distance to match the ridge image of a given probe to the created ridge images of the subjects in the gallery. For pose alignment, we extract the locations of three feature points, the inner corners of the two eyes and the tip of the nose using Gaussian curvature. These three feature points plus an auxiliary point in the center of the triangle, made by averaging the coordinates of the three feature points, are used for initial 3D face alignment. In the face recognition stage, we find the optimum pose alignment between the probe image and the gallery, which gives the minimum Hasusdorff distance between the two sets of features. This approach is used for identification of both neutral faces and faces with smile expression. Experiments on a public face database of 61 subjects resulted in 93.5% ranked one recognition rate for neutral expression and 82.0% for the faces with smile expression.

***Index Terms*—** 3D Face Recognition, Hausdorff Distance, Range Data, Ridge Image.

1. INTRODUCTION

Face recognition has become one of the most important applications of image analysis and computer vision in recent years. As mentioned in [1], there are three main contenders for improving face recognition algorithms: high resolution images, three-dimensional (3D) face recognition, and new preprocessing techniques. In current two-dimensional (2D) face recognition systems, changes in lighting (illumination), and pose of the face always have been challenging problems [2]. The 3D face recognition algorithms identify faces from the 3D shape of a person's face. Because the 3D shape of a face is less affected by changes in lighting or pose, 3D face recognition has the potential to improve performance under these conditions [2]. As Bowyer *et al.* mentioned in [2], the literature appears to be split on whether using a single 3D face data outperforms using a single 2D face data. We believe that 3D data has more potential for face recognition and the results in [3, 4, 5, 6, 7] support this opinion.

In this paper, we present a method for 3D face recognition from range images. Figure 1 shows the block diagram of our method. In the first step, because of noise and artifacts in the range images, we use median filtering and low-pass filtering to remove sharp spikes and smooth the images and then we use interpolation to fill the gaps. In the next step, we roughly find the tip of the nose which is the closest point to the scanner. Then, we apply template matching to localize the face region in the filtered range data. Afterwards, we use Gaussian curvature to label three feature points, i.e., the inner corners of the two eyes and the accurate position of the nose tip. We represent the range images by the points on the 3D surface of the face which have maximum principal curvature, k_{max}, greater than a threshold. We build an image that represents ridge-ravine lines on the 3D surface of the face. In fact, for each facial range data, a 3D binary image called ridge image is created and then used for face recognition. For recognition, we use robust Hausdorff distance to find the best match of a given probe image to a gallery image. In the recognition process, we apply similarity transformation (scale, rotation and translation) to find the best pose that matches the probe image with the gallery image. The three labeled feature points plus an auxiliary point in the middle of the triangle formed by the three labeled points are used to find the initial transformation that aligns the probe ridge image to the test ridge image. After initial alignment, we apply an iterative algorithm to find the optimum pose that results in the minimum Hausdorff distance. Compared to other approaches for 3D face recognition, such as [8, 9, 7, 5], our approach requires less computations. Rather than matching the entire surfaces of two faces (the probe image and a gallery image), we only match the detected ridge lines on the face based on the principal curvature. This reduces the computations during the matching process. In this work, the number of points in the ridge images that represent the lines around the facial feature regions are $14\% \pm 2\%$ of the total number of points that cover the face (the size of the face area in the images that we used is at most 120×120).

The rest of this paper is organized as follows. Section 2 introduces the preprocessing module, the face localization and the method for extracting three salient feature points from the range data. Section 3 presents the matching algorithm using robust Hausdorff distance. Section 4 illustrates the experimental results using a public database to validate our algorithm and test the identification performance. Conclusions and future work are discussed in Section 5.

2. PREPROCESSING AND LOCALIZING SALIENT POINTS IN FACIAL RANGE IMAGES

The range images captured by laser scanners have some artifacts, noise and gaps. In the preprocessing step, we apply median filtering to remove sharp spikes that occur during scanning the face. Afterwards, we use interpolation to fill the gaps on the face region and finally apply a low pass filter to smooth the surface of the face that suffers from rapid changes due to facial hair or other artifacts. Next, the goal is to localize the area of the face in the range data and discard the neck, hair and the background areas of the range image.

*Corresponding Author.

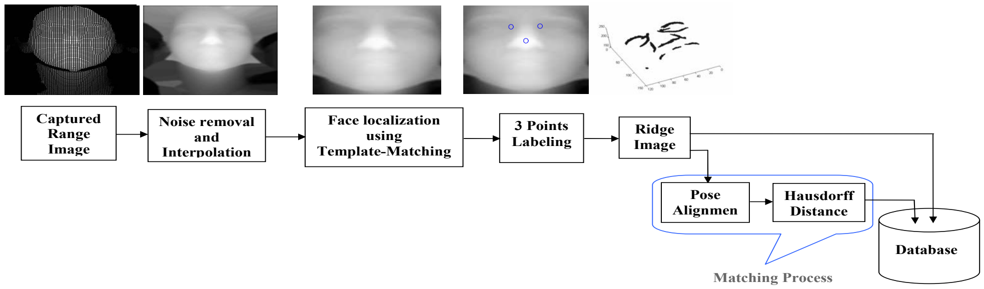

Fig. 1. Block diagram of our system for 3D face recognition from range data.

	$K>0$	$K=0$	$K<0$
$H<0$	peak	ridge	saddle ridge
$H=0$	none	flat	minimal
$H>0$	pit	valley	saddle valley

Fig. 2. The relations between surface types and their mean (H) and Gaussian (K) curvatures.

Therefore, we developed a method for localizing faces in range data using template matching. We apply normalized cross correlation to match a facial template range image to the range images. For template matching, at first we roughly detect the location of the nose tip. Then, we translate the template face such that the detected tip of the nose is placed on the location of the nose tip of the range image under test. Afterwards, we iteratively apply a similarity transformation (scale, and rotation) to the template image and calculate the normalized cross correlation to find the best pose. The area underneath the template with the best calculated pose is taken as the facial region. In order to extract the three feature points that utilized in face alignment during the recognition process, we use Gaussian curvature. The extracted feature points are the two inner corners of the eyes and the tip of the nose. Since the calculation of Gaussian curvature involves the second derivative of the surface function, the noise and the artifacts affect the final result and applying a low-pass filter to smooth the data is required. Figure 2 shows the representation of different shapes and their corresponding mean curvatures and Gaussian curvatures. As the Figure shows, the surface that either has a peak or a pit shape has a positive Gaussian curvature value ($K > 0$). Each of the two inner corners of the eyes has a pit surface type and the tip of the nose has a peak surface type that are detectable based on the Gaussian curvature. These points have the highest positive Gaussian curvature values among the points on the face surface.

Figure 3 shows the result of calculating Gaussian curvature for one of the sample range images in the gallery. The highest points in Figure 3.(a) correspond to the points with pit/peak shape. We threshold the Gaussian curvature to find the areas that have positive Gaussian curvature values greater than a threshold, producing a binary image (Fig. 3.b). This threshold is calculated based on a small training data set (15 facial range images) different from the images used in the recognition experiments. The three regions with the largest average value of Gaussian curvature are the candidate regions that include the feature points. The locations of the points with maximum Gaussian curvature in these regions are labeled as feature points. Figure 3.(c) shows a final result of feature points labeling.

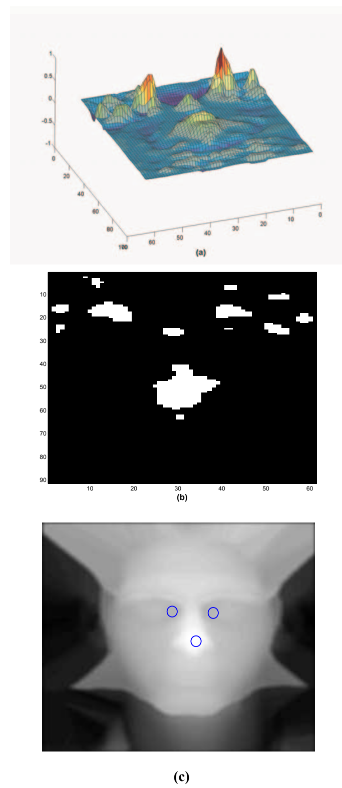

Fig. 3. Three feature points labeling: (a) Gaussian curvature on a patch around the nose and eyes (b) Result of thresholding the Gaussian curvature image(c) Final result of feature points labeling.

3. 3D FACE IDENTIFICATION USING RIDGE IMAGES AND HAUSDORFF DISTANCE

For face identification, we find the location of the ridge lines in the range image. These are the lines around the eyes, the nose, and the mouth. We call a 3D binary image that contains only these lines a ridge image. For recognition, we use Hausdorff distance to find the best match for a given probe image from the facial range images in the gallery. In the following subsections, we explain in detail the process of face identification based on ridge images and Hausdorff

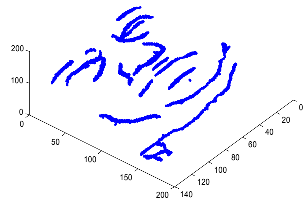

Fig. 4. An example of ridge image extracted for one subject in the gallery.

distance.

3.1. Ridge Image

Our goal is to extract and use the points lying on ridge-ravine lines as the feature points. These points correspond to the extreme ridge points on the considered surface. An extreme ridge point is a point where the principal curvature k_{max}, has large positive value. There are different approaches to locate the ridges, here we threshold the k_{max} values to find these points. The suitable threshold is obtained based on a small training set that is different from the images in the gallery. Figure 4 shows an example of the ridge images obtained by thresholding the k_{max} values. These are 3D binary images that show the location of the ridge lines on the surface of the face. The lines on the boundary of the face are filtered out and are not considered as feature points for recognition. To filter out the points on the boundary of the face, we ignore the points on the boundary of the matched template within a margin. In other words, after localizing the face by template matching, the points that are within a certain distance (for example 15 pixels) from the boundary of the matched face template are excluded from the process of ridge creation.

3.2. Hausdorff Distance

Huttenlocher *et al.* originally proposed Hausdorff distance (HD) [10] as a measure for object matching in computer vision. Unlike other shape matching methods, HD can be calculated without knowing the exact correspondences of the points. Given two sets of points $A = \{a_1, a_2, ..., a_{N_A}\}$ and $B = \{b_1, b_2, ..., b_{N_B}\}$ of size N_A and N_B respectively, the undirected partial Hausdorff distance between the two sets of points A and B is defined as:

$$H(A, B) = max(h_K(A, B), h_K(B, A)) \quad (1)$$

where $h_K(A, B)$ and $h_K(B, A)$ represent the directed distance between the two sets A and B. The directed distances of the partial HD are defined as:

$$h_K(A, B) = K^{th}_{a \epsilon A} d_B(a), \;\; h_K(B, A) = K^{th}_{b \epsilon B} d_A(a) \quad (2)$$

where $d_B(a)$ represent the minimum distance (e.g. Euclidean distance) value at point a to the point set B, $d_A(b)$ represent the minimum distance (e.g. Euclidean distance) value at point b to the point set A, $K^{th}_{a \epsilon A}$ denotes the K^{th} ranked value of $d_B(a)$, and $K^{th}_{b \epsilon B}$ denotes the K^{th} ranked value of $d_A(b)$. Sim *et al.* [11] applied the robust statistic techniques of regression analysis to the computation of the HD measures for object matching, resulting in two robust HD measures: M-HD based on M-estimation and least trimmed square-HD (LTS-HD) based on LTS. In the LTS-HD [11], $h_{LTS}(A, B)$, the directed distance is defined by a linear combination of order statistics:

$$h_{LTS} = \frac{1}{H} \sum_{i=1}^{H} d_B(a)_{(i)} \quad (3)$$

where H denotes $h \times N_A$ ($0 \leq h \leq 1$) as in the partial HD case, and $d_B(x)_{(i)}$ represents the ith distance value in the sorted sequence $d_B(x)_{(1)} \leq d_B(x)_{(2)} \leq \cdots \leq d_B(x)_{(N_A)}$. The measure $h_{LTS}(A, B)$ is calculated by eliminating the large distance values and only keeping the h fraction of the minimum distances. So, even if the object is partially occluded or degraded by noise, this matching scheme yields good results. In this work, the value of h that results in the best recognition rate is 0.8. In our case, the calculation of LTS-HD is between the two point sets of two binary images, one is the ridge image of the probe and the other is the ridge image of a subject in the gallery. The procedure for calculating the directed LTS-HD between a probe and a gallery ridge image is as follows:

1. Set $\hat{h}_{LTS} := +\infty$, and $t := 0$
2. Initially align the 3D ridge image of the test image P (i.e. translate, rotate and scale), to the gallery image P', by using the three labeled feature points and the auxiliary point. This similarity transformation is calculated by procrustes analysis [12].
3. Set $Success := 0$
4. Place the aligned probe ridge image, $T(P)$, over the gallery ridge image. For all the points in the aligned probe image, find the distance to the closest point in the gallery image, P', using:
$$D_{P'}(x) = min_{y \epsilon P'} ||x - y|| \quad (4)$$
5. Sort the minimum calculated distances and then calculate the robust Hausdorff distance, h_{LTS}, using Eq. 3.
6. If ($h_{LTS} <= \hat{h}_{LTS}$), set the following items:
$\hat{h}_{LTS} := h_{LTS}$
$t := t + 1$
$Success := 1$
7. Change the parameters of the similarity transformation, (i.e., translation, rotation, and scale).
8. If $Success = 1$ AND $(t < Max\_Iterations)$ goto 3.
9. Return h_{LTS}.

This optimization problem is solved using Matlab optimization toolbox (i.e. *fminsearch* Matlab function). The *fminsearch* uses the simplex search method of [13]. The maximum number of iteration is set to 20 in the simplex search method. This procedure is repeated to find the matching score between a probe image and all the images in the gallery. The gallery face image that results in the minimum distance, is considered as the best match.

4. EXPERIMENTS AND RESULTS

We use the GavabDB database [14], a database of 3D face images, for our experiments. GavabDB contains 549 3D facial surface images corresponding to 61 individuals. For each person, there are nine different images, two neutral frontal images, four neutral images with pose (looking left, right, down and up), and three frontal images in which the subject presents different and accentuated facial expressions. In our experiments, we used the two neutral frontal images and the frontal image with smile expression. One neutral frontal image is used as gallery, and the other two images, the second frontal

Table 1. First ranked recognition rate for the GavabDB.

Facial Expression	1^{st} Rank Recognition Rate(%)	Database	Number of Subjecs
Neutral (3D)	93.5	GavabDB	61
Neutral (2D)	82.0		
Smile (3D)	82.0		

neutral and the frontal with smile expression are used as probe images. The process of labeling three feature points is successful but fails for few subjects (15% of the images in the database), where the noise and the disturbance in the image around the eyes (i.e., eyelash) are high. Also, for few cases (10% of the images in the database), where the face has severe pose (looking upward or downward), the initial detection of the nose is difficult and the nose is mistaken with the chin or forehead. For all of the these difficult cases, we manually label the 3 feature points. In the recognition process of neutral frontal images, four subjects out of 61 subjects were not recognized by our algorithm (93.5% recognition rate). In another experiment, we projected the binary images to 2D (ignoring the 3^{rd} dimension) and used Hausdorff distance for recognition. By ignoring the 3^{rd} dimension, we obtained a recognition rate of 82%. We implemented our algorithm in Matlab 2006. Using our algorithm, it takes 1.275 seconds to match one probe ridge image to a gallery ridge image on a PC with Intel Core Duo 1.86Ghz processor.

For faces with smile expression, we considered only the upper part of the face (i.e., 3D ridge lines around the eyes and the nose) for recognition and excluded the lower part of the face, i.e., the mouth, which is affected by the smile expression. 50 subjects out 61 subjects with smile expression were recognized by our approach (recognition rate is 82%). Table 1 summaries the face recognition results based on our algorithm using the GavabDB database. We compared our method with other methods that used the same data set. Two different approaches for 3D face recognition were presented by Moreno *et al.* in [8, 9] and were evaluated using GavabDB. In [8], they segmented the range images into isolated subregions using the mean and the Gaussian curvatures. Then, they extracted 86 descriptors such as the areas, the distances, the angles, and the average curvatures of the subregions. They selected 35 best features and utilized them for face recognition based on the minimum Euclidean distance classifier. They achieved a first ranked recognition rate of 78.0% for neutral frontal images and 62% for smile expression (only 60 subjects out of the 61 from the database were utilized). In [9], they selected a set of 30 features out of the 86 features and obtained recognition rates of 82.0% and 90.16% when the images are frontal views with neutral expression using Principal Component Analysis (PCA) and Support Vector Machines (SVM), respectively. The recognition rates decreased to 76.2% and 77.9% using PCA and SVM matching schemes, respectively, under expressions and slight face rotation. As the results show, our method has better performance for recognition of both images with neutral expression and with smile expression.

5. CONCLUSIONS AND FEATURE WORKS

This paper presented a novel algorithm for 3D face recognition from range data based on 3D binary images, created using principal maximum curvature, using LTS-Hausdorff distance. The 3D binary image shows the locations of the ridge lines in the range facial image. The GavabDB database was utilized to test our algorithm. Our algorithm outperforms other approaches presented by Moreno *et al.* on the GavabDB database. In the future, we will apply our approach to cases with other types of expressions and use other databases such as Face Recognition Grand Challenge (FRGC).

6. REFERENCES

[1] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[2] Kevin W. Bowyer, K. Chang, and Patrick Flynn, "A survey of approaches and challenges in 3d and multi-modal 3d+2d face recognition," *Computer Vision and Image Understanding*, vol. 101, no. 1, pp. 1–15, 2006.

[3] A. N. Ansari and M. Abdel-Mottaleb, "Automatic facial feature extraction and 3d face modeling using two orthogonal views with application to 3d face recognition," *Pattern Recognition*, vol. 38, no. 12, pp. 2549–2563, 2005.

[4] T. Maurer, D. Guigonis, I. Maslov, B. Pesenti, A. Tsaregorodtsev, D. West, and G. Medioni, "Performance of geometrix activeidtm 3d face recognition engine on the frgc data," in *IEEE Workshop on Face Recognition Grand Challenge Experiments*, 2005.

[5] T. D. Russ, K. W. Koch, and C.Q. Little, "A 2d range hausdorff approach for 3d face recognition," in *IEEE Workshop on Face Recognition Grand Challenge Experiments*, 2005.

[6] G. Passalis, I. Kakadiaris, T. Theoharis, G. Toderici, and N. Murtuza, "Evalutaion of 3d face recognition in the presence of facial expressions: an annotated deformable model approach," in *IEEE Workshop on Face Recognition Grand Challenge Experiments*, 2005.

[7] Xiaoguang Lu, Anil K. Jain, and Dirk Colbry, "Matching 2.5d face scans to 3d models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 31–43, 2006.

[8] Ana Belen Moreno, Angel Sanchez, Jose Fco. Velez, and Javier Dkaz, "Face recognition using 3d surface-extracted descriptors," in *Irish Machine Vision and Image Processing Conference 2003 (IMVIP'03)*, 2003.

[9] Ana Belen Moreno, Angel Sanchez, Jose Fco. Velez, and Fco. Javier Dkaz, "Face recognition using 3d local geometrical features: Pca vs. svm," in *Fourth International Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, 2005.

[10] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[11] Dong-Gyu Sim, Oh-Kyu Kwon, and Rae-Hong Park, "Object matching algorithms using robust hausdorff distance measures," *IEEE Transactions on Image Processing*, vol. 8, no. 3, pp. 425–429, 1999.

[12] J. R. Hurley and R. B. Cattell, "The procrustes program: Producing direct rotation to test a hypothesized factor structure," *Behavioral Science*, pp. 258–262, 1962.

[13] J.C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder-mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[14] A. B. Moreno and A. Sanchez, "Gavab: A 3d face database," in *2nd COST Workshop on Biometrics on the Internet: Fundamentals, Advances and Applications*, 2004.