Biometric Driver Authentication Based on 3D Face Recognition for Telematics Applications

Hyeonjoon Moon[1] and Kisung Lee[2]

[1] School of Computer Engineering, Sejong University
hmoon@sejong.ac.kr
[2] College of Health Sciences, Korea University, Seoul, Korea
kisung@korea.ac.kr

Abstract. In this paper we developed driver authentication system based on face recognition. Since 2D based face recognition has been shown its structural limitation, 3D model based approach for face recognition has been spotlighted as a robust solution under variant conditions of pose and illumination. Since a generative 3D face model consists of a large number of vertices, a 3D model based face recognition system is generally inefficient in computation time. In this paper, we propose a novel 3D face representation algorithm to reduce the number of vertices and optimize its computation time. Finally, we evaluate the performance of proposed algorithm with the Korean face database collected using a stereo-camera based 3D face capturing device. In addition, various decision making similarity measures were explored for final results. Our experimental results indicated that the algorithm is robust for driver authentication inside the vehicle and is also reasonably fast for real-time processing.

Keywords: biometrics, 3D face recognition, 3D Model, driver authentication, driver identification, telematics application.

1 Introduction

Biometrics is a rapidly evolving technology which has been widely used in variety of applications such as criminal identification and car theft, etc. In addition, it can make an important role in telematics applications. Especially biometric authentication methods based on face recognition can be a valuable tool to prevent unauthorized access to the vehicles, to detect drowsiness of the driver, and to do automatic adjustment of driver's settings for in-vehicle devices such as back mirrors, driver seat, and other telematics devices etc.

Traditional face recognition systems have primarily relied on 2D images. However, they tend to give a higher degree of recognition performance only when images are of good quality and the acquisition process can be tightly controlled. Recently, many literatures on 3D based face recognition have been published with various methods and experiments. It has several advantages over traditional 2D face recognition: First, 3D data provides absolute geometrical shape and size information of a face.

C. Stephanidis (Ed.): Universal Access in HCI, Part I, HCII 2007, LNCS 4554, pp. 473–480, 2007.

Additionally, face recognition using 3D data is more robust to pose and posture changes since the model can be rotated to any arbitrary position. Also, 3D face recognition can be less sensitive to illumination since it does not solely depend on pixel intensity for calculating facial similarity. Finally, it provides face segmentation information since the background is typically not synthesized in the reconstruction process [1].

Image captured by in-vehicle camera is very sensitive to pose and illumination variation which makes this application differentiated from the other 2D based approaches. In this paper we propose a novel 3D face representation. Proposed algorithm has flexibility to reduce the vertex number of each 3D face remarkably and all the 3D faces can be aligned with correspondence information based on the vertex number of a reference face simultaneously.

This paper is organized as follows. The next section presents the proposed driver authentication process for face recognition. Section 3 describes the procedure of the 3D face recognition process. Experimental results are presented in Section 4 based on a Korean database. Finally, conclusions and future work are discussed in Section 5.

2 Driver Authentication Process

We assume several constraints for this study.

- Car is used by limited number of drivers
- Medium size and quality of video/digital camera is utilized for realistic experi ments.
- Camera is fixed at upfront of driver seat and centered to the forehead of driver.
- Maximum 100 drivers are to be identified per vehicle.

Figure 1 shows the block diagram to explain the driver authentication process.

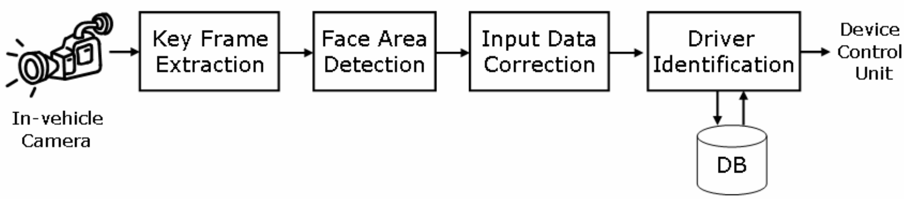


Fig. 1. Block diagram of the driver authentication system

When driver sits in the vehicle, video frames are acquired by in-vehicle camera. Since face recognition algorithm is sensitive to shooting conditions in the vehicle such as amount of lights, angle of camera to the face, etc, each frame is investigated to choose multiple key frames which have good image quality and good position of the face(i.e. fore front of the face). Then normalization techniques are applied to improve the quality of the images for further processes.

We assumed that the face detection algorithm is applied to produce normalized face images. Then the 3D based face recognition system process the 2D face image and look up the driver database to identity or verify the specific person. Authenticated driver is informed to the device control unit for user-dependent customization of telematics services.

2.1 Key Frame Selection

Key frame extraction technique has been widely used for video abstraction which can summarize a video clip with the limited number of major frames that contains significantly important information of the clip. Most of the algorithms observe variation of information in the video frame among adjacent frames, such as motion and object properties, etc [2].

In this study, however, change in adjacent frames is not important. Because of the variation of environment inside car, we observed that many different conditions of face images had acquired even within the same video clip. Since the face recognition algorithms is very sensitive to light conditions such as intensity, contrast, etc, we developed key frame extraction method which can adequately reflect those properties.

In addition, the algorithm focused on the quality of the foreground object (i.e. driver) by eliminating background scene to improve the reliability of the algorithm. We used object segmentation algorithm in [3] to extract object area. The algorithm is suitable for real time process and performs well especially for the object with stationary background. We present the pseudo-codes for key frame extraction method as below.

```
  for (each_frame) {
   Do_object_extration();
   if (extraction_is_failed()) continue;
   get_histogram_luminance_contrast();
   score_frame();
}
Select_best_scored_frames();
```

2.2 Face Area Detection and Input Data Correction

The extracted object contains driver's face and upper body. In order to segment the face we applied filtering followed by masking procedure. If the segmented face is tilted, geometrical transform is applied to correct the position of face. Histogram equalization and histogram specification [4] was also performed for improve the contrast of the face.

2.3 Driver Identification

Driver identification module is divided into feature extraction and similarity measure. Principal component analysis (PCA) a based algorithm [5] has been widely used for many face recognition applications. The algorithm, however, it has drawback for pose and illumination variation. Since in-vehicle environment in this study is highly affected by the driver's pose and light from outside, we used 3D based face recognition algorithm which is more robust to the pose and illumination variation. For similarity measure, we explored L1 + Mahalanobis distance, L2 + Angle was performed for the test [6].

3 3D Based Face Recognition

In general, a 3D face scan consists of texture intensity in 2D frame and geometrical shape in 3D. Especially, the shape information is represented by close connections of many vertices and polygons. Since the vertex number of each of 3D face scans is different from each other, it is necessary to manipulate them to have the same number of vertices for consistent mathematical expression. We propose a novel 3D face representation algorithm as shown in Figure 2.

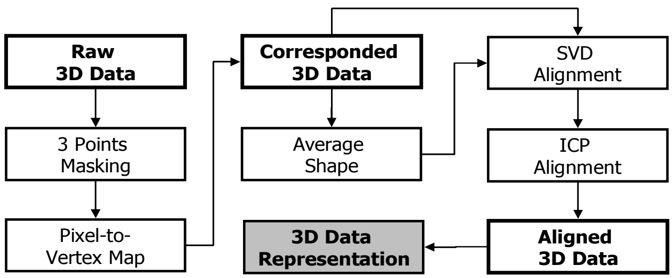


Fig. 2. 3D face representation process

3.1 Face Alignment and Model Generation

It is possible that all 3D face scans in our database are expressed with the same number of vertex points. To construct a more accurate model, it is necessary to utilize some techniques for face alignment, which is transforming the geometrical factors (scale, rotation angles and displacements) of a target face based on a reference face. Face alignment in our research is achieved by adopting singular value decomposition (SVD) [7] and Iterative Closest Points (ICP) [8][9] sequentially.

We constructed separate models from shapes and textures of 100 Korean people by applying PCA [10] independently. The separate models are generated by linear combination of the shapes and textures as given by equation below.

$$\mathbf{S} = \mathbf{S}_0 + \sum_{j=1}^{N_S} \alpha_j \mathbf{S}_j, \qquad \mathbf{T} = \mathbf{T}_0 + \sum_{j=1}^{N_T} \beta_j \mathbf{T}_j. \tag{1}$$

where $\boldsymbol{\alpha} = [\alpha_1\, \alpha_2 \cdots \alpha_{N_S}]$ and $\boldsymbol{\beta} = [\beta_1\, \beta_2 \cdots \beta_{N_T}]$ are the shape and texture coefficient vectors (should be estimated by a fitting procedure). Also, $\mathbf{S}_0$ and $\mathbf{S}_j$ are the shape average model and the eigenvector associated with the jth largest eigenvalue of the shape covariance matrix, $\mathbf{T}_0$ and $\mathbf{T}_j$ in textures likewise.

Fig. 3. A generative 3D model. These are rotated versions of the 3D model in Y-axis direction. Each version from the first is ranged from -45 degrees to 45 degrees at 15 degrees interval.

3.2 Fitting the 3D Model to a 2D Image

Shape and texture coefficients of the generative 3D model are estimated by fitting it to a given 2D input face image. This is performed iteratively as close as possible to the input face. Fitting algorithms, called stochastic Newton optimization (SNO) and inverse compositional image alignment (ICIA) were utilized in [11] and [12], respectively. It is generally accepted that SNO is more accurate but computationally expensive and ICIA is less accurate but more efficient in computation time [12].

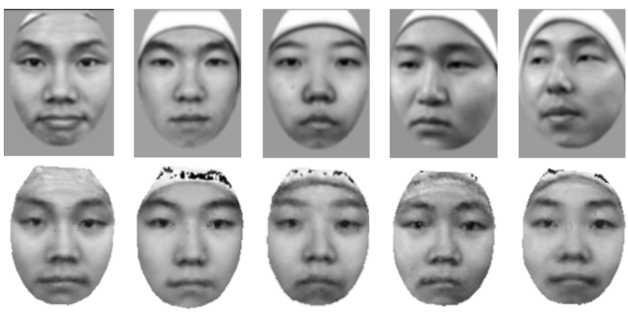
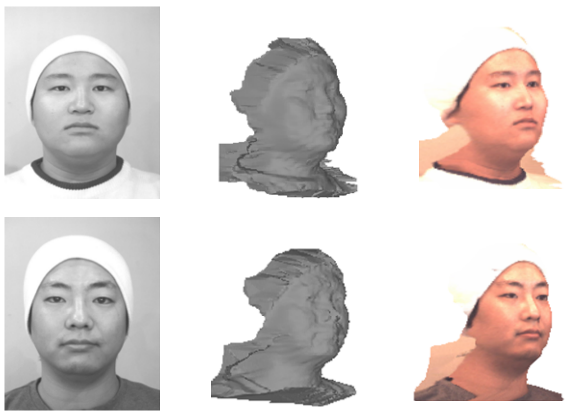


Fig. 4. Fitting results. The images in top row are input images and those in bottom row are the fitted versions of our 3D model. Especially, the inputs to the third column are frontal and the others are rotated 30 degrees approximately.

We also explore the ICIA algorithm as a fitting method to guarantee the computational efficiency. Given an input image, initial coefficients of shape and texture and projection parameters for the model are selected appropriately. Initial coefficients of shape and texture usually have zero values but projection parameters are manually decided by the registration of some important features. Then, fitting steps are iterated until convergence to a given threshold value, minimizing the texture difference between the projected model image and the input image. During the fitting process, texture coefficients are updated without an additive algorithm at each iteration. But in case of shape coefficients, their updated values are not acquired with ease because of the nonlinear problem of structure from motion (SFM) [13]. To solve it, we recover the shape coefficients using SVD based global approach [14] after the convergence.

4 Experimental Results

We evaluate our face recognition system based on a 3D model generated using proposed representation algorithm. As mentioned in previous sections, 3D Korean faces are collected using a Geometrix Facevision 200 [15], which is a stereo-camera based capturing device offering a 3D face scan including approximately 30,000 ~ 40,000 vertices and corresponding 2D texture image. There are 218 3D face scans collected from 110 people during 3 sessions, which are limited with frontal views and neutral expressions, in our database. We used 100 scans in session 1 for 3D model generation and 52 scans in other sessions for the performance evaluation. Also, 7 sided face images with range from 15 to 40 degrees are acquired separately using the same device for variant pose test. Some example scans are showed in Figure 5.



(a) (b) (c)

Fig. 5. 3D face examples : (a) texture images (b) shaded shapes (c) rendered 3D views

The experimental results to frontal and sided faces are shown in Table 1. In both experiments, we utilized the L2 norm and angle combined with Mahalanobis distance as a distance metric, denoted by L2+ Mahalanobis and Angle+ Mahalanobis respectively [16]. Also, we performed additional experiments on two cases, one is to use only texture coefficients and the other is to combine texture and shape coefficients. Recognition accuracy with rank 1 in both tests was 90.4% (47 out of 52 subjects) and 85.7% (6 out of 7 subjects) respectively. The average fitting time taken without the shape recovery was 3.7s on 1.73GHz Pentium-M and 1GB RAM, but when the recovery process is included, it required 11.2s on the same machine. A remarkable result is that utilizing the shape coefficients doesn't improve the performance meaningfully in spite of increased computation time.

Table 1. Recognition accuracy with rank 1 to frontal faces and pose variant faces

	Only Texture		Texture + Shape	
	L2+Mah	L2+Angle	L2+Mah	L2+Angle
Frontal faces	90.4%	86.5%	90.4%	88.5%
Pose variant faces	71.4%	85.7%	71.4%	85.7%
Computation time	3.7s		11.2s	

5 Conclusion

In this paper we developed driver identification system based on face recognition. Since the algorithm is video-based, key frame extraction was performed followed by face area detection and input data correction. For pose and illumination-invariant properties, we presented a novel 3D face representation algorithm for 3D model based face recognition system. On the basis of the presented method, an original 3D face scan including 30,000 ~ 40,000 vertices could be represented with about 5,000 vertices. We have generated 3D model using 100 3D face images (each 3D face image composed of 4822 vertices). Then, shape and texture coefficients of the model were estimated by fitting into an input face using the ICIA algorithm. For 3D model generation and performance evaluation, we have made the Korean 3D face database from a stereo-camera based device. Experimental results show that face recognition system using the proposed representation method is more efficient in computation time. Romdhani et. al. [12] presented that the fitting time takes about 30 seconds with almost same condition as our proposed algorithm, though it is less accurate in terms of recognition performance. Our preliminary results indicated that the algorithm is robust for in-vehicle face video and is also reasonably fast for real-time processing. Our future works will focus on automatic initialization in the fitting procedure and real-time processing.

Acknowledgement. This research was performed for the Promotion and Support Program of Strategic Industry Innovation Cluster, one of the Collaboration Programs for Industry, Academy and Research Institute funded by City of Seoul, Korea.

References

1. Papatheodorou, T., Rueckert, D.: Evaluation of 3D Face Recognition Using Registration and PCA. In: Kanade, T., Jain, A., Ratha, N.K. (eds.) AVBPA 2005. LNCS, vol. 3546, pp. 997–1009. Springer, Heidelberg (2005)
2. Winnemoller, H., Olsen, S.C., Gooch, B.: Real-time Video Abstraction. In: ACM Transactions on Graphics, vol. 25, pp. 1221–1226. ACM Press, New York (2006)
3. Kim, C., Hwang, J.-N.: Fast and Automatic Video Object Segmentation and Tracking for Content-Based Applications. In: IEEE Transactions on Circuits and Systems for Video Technology (CSVT), vol. 12, pp. 122–129 IEEE, (2002)
4. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 2nd edn. Prentice-Hall, Englewood Cliffs (2002)

5. Turk, M., Pentland, A.: Eigenfaces for Recognition. J. Cognitive Neuroscience 3, 71–86 (1991)
6. Development of the Performance Evaluation Tool for the Face Recognition System, Technical Report, Korea Information Security Agency (2004)
7. Horn, B.K.P., Hilden, H.M., Negahdaripour, S.: Closed-Form Solution of Absolute Orientation Using Orthonormal Matrices. Journal of the Optical Society of America A 5, 1127–1135 (1988)
8. Besl, P.J., Mckay, N.D.: A Method for Registration of 3D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2), 239–256 (1992)
9. Lu, X., Jain, A., Colbry, D.: Matching 2.5D Face Scans to 3D Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(1), 31–43 (2006)
10. Vetter, T., Poggio, T.: Linear Object Classes and Image Synthesis from a Single Example Image. IEEE Transactions on Pattern Analysis and machine Intelligence 19(7), 733–742 (1997)
11. Blanz, V., Vetter, T.: Face Recognition Based on Fitting a 3D Morphable Model. IEEE Transactions on Pattern Analysis and Machine Intelligence 25(9), 1063–1074 (2003)
12. Romdhani, S., Vetter, T.: Efficient, Robust and Accurate Fitting of a 3D Morphable Model. In: IEEE International Conference on Computer Vision (2003)
13. Bascle, B., Blake, A.: Separability of pose and expression in facial tracking and animation. In: Sixth International Conference on Computer Vision. pp. 323–328 (1998)
14. Romdhani, S., Canterakis, N., Vetter, T.: Selective vs. Global Recovery of Rigid and Non-rigid Motion. Technical report, CS Dept. Univ. of Basel (2003)
15. http://www.geometrix.com
16. Moon, H., Phillips, P.: Computational and Performance Aspects of PCA-Based Face-Recognition Algorithms. Perception 30, 303–321 (2001)