POSTER:
3D-Face model Reconstruction Utilizing Facial Shape Database from Multiple Uncalibrated Cameras

Toshio Takeuchi
Keio University
3-14-1 Hiyoshi Kohoku-ku
223-8522, Yokohama, Japan
takeuchi@ozawa.ics.keio.ac.jp

Hideo Saito
Keio University
3-14-1 Hiyoshi Kohoku-ku
223-8522, Yokohama, Japan
saito@ozawa.ics.keio.ac.jp

Masaaki Mochimaru
Digital Human RC, AIST
2-41-6, Aomi, Koto-ku
135-0064, Tokyo, Japan
m-mochimaru@aist.go.jp

ABSTRACT

In recent years, a surveillance camera has come to be attached in various places from a rise of the consciousness to security. Since the surveillance cameras are installed in variety of place, it is possible to take a picture of the same person from multiple uncalibrated cameras though it is asynchronous. In this article, we propose a method for reconstructing a face shape from multiple-view images taken with non - synchronous multiple cameras. In this method, we do not directly reconstruct the shape, but estimate a small number of parameters which represent the face shape. The parameter space is constructed with Principal Component Analysis of database of a large number of anatomical face shapes collected for different people. From the input multiple view images, the region of the face and three feature points on the face are manually extracted. Then the facial pose is estimated by optimizing the evaluation based on the silhouette shape, appearance, and the position of the feature points. According to the facial pose, the parameters representing the facial shape are also estimated by optimizing the same evaluation function. Those optimizing procedures are repeated for obtaining the facial shape for the object face captured with the non-synchronous multiple cameras. The experimental results demonstrate the effectiveness of the proposed method. Since the database used in this paper consists of anatomically aligned shape data, we can obtain anatomical shape of the face, which is suitable to represent the identity of each person.

Keywords

Asynchronous, Multi-view, 3D-Face model, Face shape database, PCA, Anatomical, pattern recognition

1. INTRODUCTION

In recent years, surveillance cameras have been attached in various places from a rise of the consciousness to security. By using surveillance cameras, we might be able to prevent unauthorized people entering into secured area. Surveillance cameras can be used not only as such a criminal deterrence purpose but also for the purpose of recording the action history of the people. For such a purpose, the identification of the person taken in images of the surveillance cameras is one of the important technologies. The surveillance cameras are installed in variety of places, so it is possible to take pictures of the same person from multi view cameras although it is asynchronous. Using the multiple viewpoint images, the reconstruction accuracy improves more than the reconstruction from one image. However, accurate position and poses of the capturing cameras should be estimated, which sometimes makes the computing time become huge. As a method for reconstructing the face shape from the image, there is a method named Active Shape Model that uses the data base with man's head shape information analyzed in PCA and that presumes shape from a peculiar space of the low level. There are researches of the Active Shape Model. Some are using the PCA to 3D shape and texture information and estimate the face shape [Vet99], [Vet02], [Vet03], others fit two dimensional vertex model to the image, and compare that from three dimensional shape model [Yan04], [Yan05]. There is a method using AR-toolkit for estimating face pose and position, and entering the characteristic points manually, then estimating three dimensional face shape [Saito06]. But using AR-toolkit, the scene used is limited. This research reconstructs three dimensional face model by using that the direction of the face is already-known.

In this paper, we will propose a novel method for recovering face shape from multiple viewpoint images that are taken with asynchronous uncalibrated multiple cameras. Our method estimates both face direction and three dimensional face shape at the same time. We use the anatomy face shape database[Saito06]. Adopting PCA to the database, our method express the face shape by small number of parameters of about 20 dimensions. By estimating the parameter, we can reconstruct face shape. To estimate the face shape, the estimation of face direction is necessary. Our method use random face shape that can be calculated by face database to estimate the face direction and then estimate the face shape. By repeating this face direction estimating and face shape estimating, we reconstruct the three dimensional face shape from asynchronous multi view cameras.

The data base used with this paper is made for an anatomy study, so that we can easily get the anatomy specific vertex after reconstructing the face shape. So, reconstructed three dimensional model can be used not only for the person identification but also for getting face information from pictures took by multi view cameras installed in a free position.

2. PROPOSED METHOD

In our method, we adopt the way to compare the model image described by a few principal ingredients with input image, optimize the face shape and get the answer. Thus, it is a final purpose of this method to estimate the face direction and three dimensional face shape that fit the input image best. To estimate that, we use three evaluated value, facial contour information, specific points information [Sai06] and appearance information. We describe the evaluated value of appearance. We choose the image that is the nearest the front direction among all images given as an input by using the result of face direction estimation. Afterwards, texture mapping is done to three dimensional model by the texture of an image that is the nearest the front at that time. Images in other three directions are synthesized by the three dimensional model which is texture mapped.

We show the outline of the proposed method in Fig.1. First, estimate the rough face direction by using the input image. Next, estimate the three dimensional face shape by using face direction. Then, repeating the face direction estimation and the three dimensional face shape estimation, the face direction and three dimensional face shape when evaluated value converge is assumed to be the face direction and three dimensional face shape of the input image.

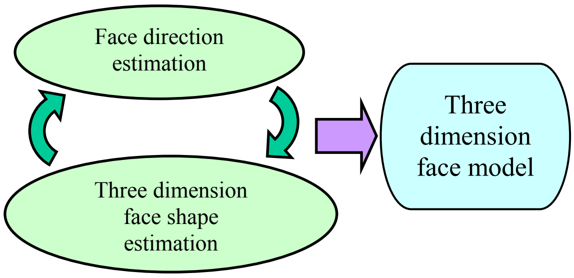


Figure 1. Outline of proposed method

3 . EXPERIMENT AND RESULT

3.1 Error evaluation using plastic model with known shape

To evaluate the accuracy of the three dimensional face model reconstructions, we use a plastic face model with known shape that is the average shape of the face database. We take the picture of the model by non-synchronous multiple cameras, and adopt this method to reconstruct the model.

The reconstruction accuracy can be evaluated by comparing the vertex of reconstructed shape with the correct answer shape, because correct answer shape is already-known. If difference between two become smaller after adopting our method, it can be said that near shape can be made by our method.

We use 4 input images, Fig.**2** shows the input image.

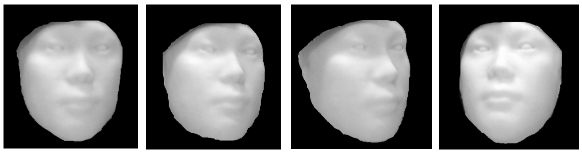


Figure 2. Input images for plastic model with known shape

Fig.**3** **(b)** shows the result of recovered shape form the input images shown in Fig.**2**. Fig.**3**(a) indicates the initial shape used for the recovery of the shape in the proposed method. Fig.**3** (c) is the correct shape of the object face. By comparing those images shown in Fig.**3**, it is obvious that the recovered shape shown in Fig.**3** (b) is getting close to the correct shape shown in Fig.**3** (c).

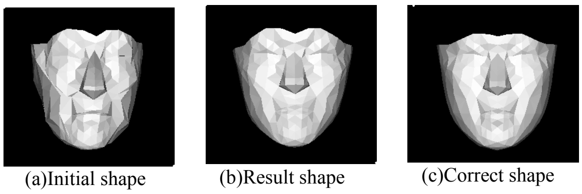


Figure 3. Result image

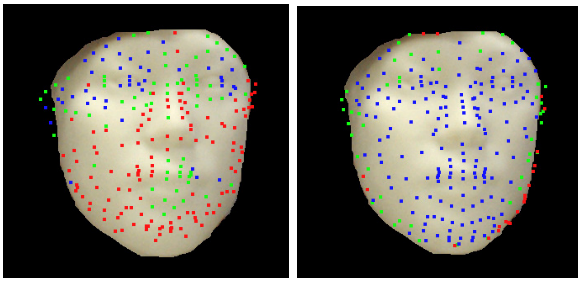
We also compare the error of the initial shape with the error of the recovered shape. Table.**1** shows the

 ISBN 978-80-86943-17-6

average error of the shape. Face area means the whole area of face and face front means the area forward from eyes. Table.**1** says that the error of the face front is more than the error of the face area. This is also indicated by Fig.**4**, which shows the distribution of error margin. Red shows the error margin equal or more than 8mm, and green shows the error margin less than 8mm and more than 5mm, blue shows the error margin equal and less than 5mm. Fig.**4**(a) shows initial shape error, and Fig.**4**(b) result shape error. Fig.**4**(a) shows that the top error margin has extended to the entire face, especially the outline of the shapely nose and the cheek. Fig.**4**(b) shows that the error margin decreases overall.

	Initial shape	Recovered shape
Face area	9.83	5.01
Face front	9.26	4.36

Table. 1. Average Error(mm)



(a) Initial shape (b) Recovered shape

Figure 4. Error margin

To examine the average and variance of 3D face model reconstruction error, we used 4 images from 10 input images. There are 210 kinds of combinations that take 4 images from 10 input images. We calculated the error of all 210 combinations and then, calculated the average and variance of 210 combinations. Table.2 shows the result of error. The average error of face area is 4.87mm, and face front is 4.76mm. The data shows face front is better accuracy than face area, but the extension of variance is larger than that of face area. It may be because of the delicacies of face parts. facial contour changes little by direction estimation error, but facial part changes large.

	average	variance
Face area	4.87	0.202
Face front	4.76	0.376

Table. 2. Average and variance

3.2 Performance evaluation using real face

We took a picture of five actual people, and estimate the direction and three dimensional model of the input image. We use the images that are assumed as an asynchronous multiple viewpoints as input images. Fig.**5** shows the example of the input images. Some expressions are different and the direction and the position of the face to the camera have changed into each image as shown here. We clipped the face area, and used it as gray scale.

Figure 5. Example of the input image

Fig.**6** shows input images and result shapes. Steps above are the input images, and below are result shapes. The input image was taken of a picture from various directions, and only an image near the front was taken out here for the comparison. It is understood that the shape that exists in the feature of each face such as nose, eye origin, mouths have been optimized.

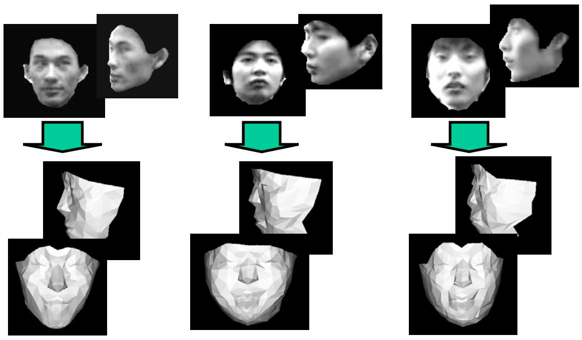
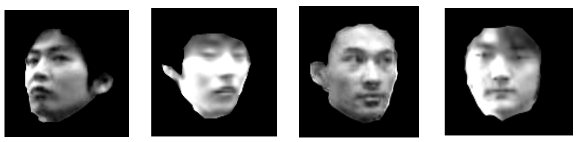


Figure 6. Comparison of image



(a)No.0 (b)No.1 (c)No.2 (d)No.3

Figure 7. Input image

Fig.**7** shows the input image of silhouette experiment. Fig.**8** shows comparison of silhouette of recovered face shape with silhouette of input image. The silhouette of the recovered face shape is closer to the input image silhouette than the initial shape. In Fig.**8**, the green part is a common of silhouette of the input and silhouette made from the model. A blue part is a part only of the silhouette of the input. A yellow part is a part only of the silhouette made from the model.

The protruding area is large in an initial model from the silhouette of the input image. By the direction and three dimensional model of the input image improving since the face direction estimation and the processing of three dimensional model estimation are repeated, it is understood that the area of the silhouette part of the protruding model decreased

 ISBN 978-80-86943-17-6

from the silhouette of the input image. As a result, it is understood that the direction and three dimensional model of the input image improved by repeating the processing of this method.

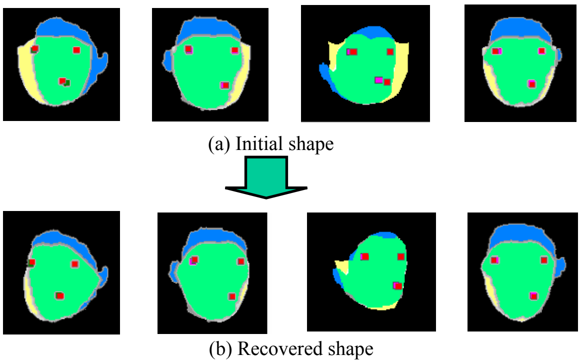


Figure 8. Comparison of the silhouette

Fig.**9** shows comparison of the texture mapping image of recovered shape with the texture mapping image of initial shape. Initial result means the result of the texture mapping on the initial three dimensional model. Reconstructed result means the result of texture mapping on the reconstructed model after repeating the direction estimation and three dimensions model estimation of the face. The texture mapping results for the initial shape are distorted more than the texture mapping results for the reconstructed shapes. This means that the accuracy of the direction estimation by the initial face is bad, the first three dimensional model estimation are thought to be similarly bad accuracy, and the texture mapping result, too. The distortion decreases in reconstructed result, and it is understood that it is near the correct answer image. As The accuracy of face direction estimation and face shape estimation become better by having repeated the direction estimation and the model estimation of three dimensions of the face, the texture mapping result improved as a result, too.

4. CONCLUSION

In this research, we propose the face reconstructing method by using the face image taken from an asynchronous multiple viewpoints. In the image, the face direction and face shape is unknown. And we use the anatomy face shape data base for reconstruction. We examine the process of optimization by this method.

In this method, we don't use that the direction of the face is already-known. We proposed the method for obtaining both by repeating the face direction estimation and the three dimensional face model estimation. In the experiment, we examined the appearance of optimization. It was shown for the distortion of the rotation by texture mapping to decrease, and to be able to make three dimensional models which looked like the object.

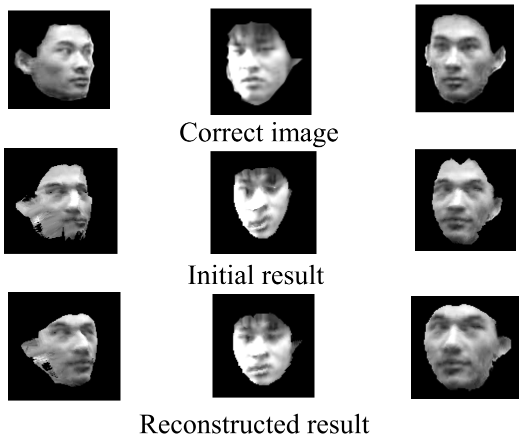


Figure 9. Optimization of the texture mapping

With the advantage that direct three dimensional shape information is obtained after it reconstructs it by using the database that contains a lot of anatomy shape information, it has the possibility to be able to use for the person attestation by the surveillance camera etc. in the future

5. ACKNOWLEDGMENT

This work was supported by Grant-in-Aid for Scientific Research (C)

(17500119).

6. REFERENCES

[Saito06]H. Saito, Y. Ito, M. Mochimaru, Face Shape Reconstruction from Image Sequence Taken with Monocular Camera using Shape Database, Proc. 14th International Conference on Image Analysis and Processing (ICIAP 2007), pp165-170, Sept. 10-14, 2007

[Vet99]T.Vetter, V.Blanz, A morphable model for the synthesis of 3D-faces, ACM SIGGRAPH, 1999

[Vet02]T.Vetter, S.Romdhani, V.Blanz, Face identification by fitting a 3D morphable model using linear shape and texture error functions, European Conf. on Computer Vision, vol.V4, pp.3-19

[Vet03]T.Vetter, V.Blanz, Face recognition based on fitting a 3D morphable model, IEEE Pattern Analysis and Machine Intelligence, vol.25, 2004

[Yan04]S.Yan, L.Zhang, H.Zhang, Y.Hu, Automatic 3D reconstruction for face recognition, IEEE International Conference on Automatic Face and Gesture Recognition, 2004

[Yan05]S.Yan, L.Zhang, H.Zhang, W.Gao, D.Jiang, Y.Hu, Efficient 3D reconstruction for face recognition, IEEE Pattern Recognition, vol38, pp.787-798, 2005