Automatic 3D Facial Movement Detection from Mirror-reflected Multi-Image for Facial Expression Modeling

거울 투영 이미지를 이용한 3D 얼굴 표정 변화 자동 검출 및 모델링

*Kyu-Min, Kyung. **Mignon Park, ***Chang-ho, Hyun
*경규민, **박민용, ***현창호

Abstract - This thesis presents a method for 3D modeling of facial expression from frontal and mirror-reflected multi-image. Since the proposed system uses only one camera, two mirrors, and simple mirror's property, it is robust, accurate and inexpensive. In addition, we can avoid the problem of synchronization between data among different cameras. Mirrors located near one's cheeks can reflect the side views of markers on one's face. To optimize our system, we must select feature points of face intimately associated with human's emotions. Therefore, we refer to the FDP (Facial Definition Parameters) and FAP (Facial Animation Parameters) defined by MPEG-4 SNHC (Synthetic/Natural Hybrid Coding). We put colorful dot markers on selected feature points of face to detect movement of facial deformation when subject makes variety expressions. Before computing the 3D coordinates of extracted facial feature points, we properly grouped these points according to relative part. This makes our matching process automatically. We experiment on about twenty koreans the subject of our experiment in their late twenties and early thirties. Finally, we verify the performance of the proposed method by simulating an animation of 3D facial expression.

Key Words : 3D facial expression, mirror-reflected multi-image, marker, facial feature point, 3D face modeling

1. Introduction

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intentions. In the past decade, much progress has been made to build synthesize facial animation for new communication methods. This is similar to human-to-human communication with a face-to-face style and is sometimes called a Life-like Communication Agent. Most these systems attempt to recognized a prototypic emotional expressions, i.e., joy, surprise, anger, sadness, fear, and disgust. This practice may follow from the work of Darwin and more recently Ekman and Friesen , and Izard et al. who proposed that basic emotions have corresponding prototypic facial expressions. This paper proposes a method of automatic estimation of facial muscle parameters from marker movements in a face image to develop a system for realistic facial animation. For 3D position and motion estimation, many researchers often use more than two cameras. These systems are based on the epipolar constraint and the 8 points algorithm. On the contrary, we used mirrors to get new images with different view directions. We proposed a more robust and simpler algorithm to estimate accurate 3D position and motion from mirrored and front view images, since there are good properties of mirrored images that can be used. Besides, we conquest the cost problem of a system that uses several cameras and synchronization between data among different cameras using only one camera and two mirrors (a mirrored image of a scene is equivalent to a virtual camera behind it). With those, we can create, at a low cost, a robust system. Finally, we verify the performance of the proposed method by simulating an animation of 3D facial expression.

2. Camera Geometry and Calibration

Fig. 2.1 illustrates the basic geometry of the camera model. (X_W, Y_W, Z_W) is the 3D coordinate of the object point P_W in the 3D world coordinate system. (X, Y, Z) is the 3D coordinate of the object point P in the 3D

저자 소개
* 경 규 민 : 연세대학교 인지과학 전기전자전공 석사과정
** 박 민 용 : 연세대학교 전자전자공학과 교수
***현 창 호 : 연세대학교 전기전자공학 박사과정

This research was performed for the Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Ministry of Science and Technology of Korea.

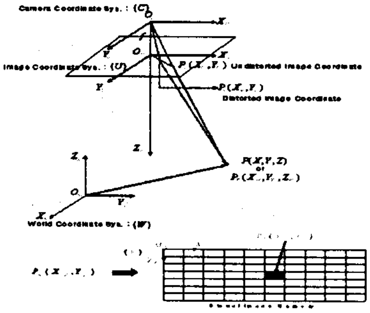

Fig. 2.1 Camera geometry with perspective projection and radial lens distortion.[11]

camera coordinate system, which is centered at point O_C the optical center, with the Z_C axis the same as the optical axis.

3. Facial Feature Points

We use colorful circle marker, red and green. Six red markers are located in the middle of face vertically to calculate the mirror's normal vector. We should observe these points all in three view, left mirror side , front side , and right mirror side. Green markers are attached on each feature point of the subject's face to measure and model facial expression. To optimize our system, we select feature points of face intimately associated with human's emotions. We refer to the FDP (Facial Definition Parameters) and FAP (Facial Animation Parameters) of MPEG-4 SNHC (Synthetic/Natural Hybrid Coding). Finally, we select 27 feature points of face for our experiment. (Fig. 3.1).

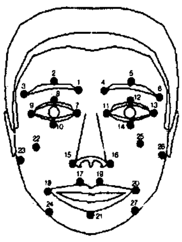

Fig. 3.1 Selected feature points for our experiment

4. 3D facial movement estimation

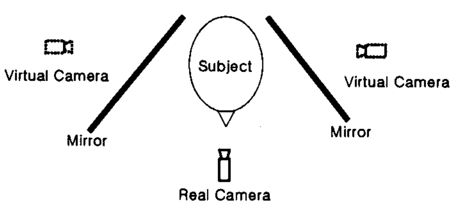

Fig. 4.1 The conceptual virtual camera diagram

With two mirrors next to a subject's face(Fig.4.1), we can acquire three different views of the face image data simultaneously and can also avoid the problem of synchronization between data among different cameras.

$f(x_i, y_j)$

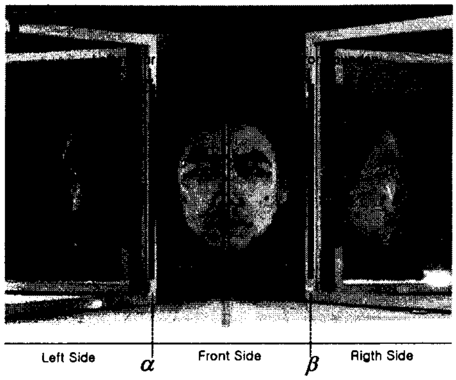

Fig. 4.2 Separating region in input image

An image is denoted by (x_i, y_j). We divide labeled feature point's pixel coordinate $L(x_i, y_j)$ into 4 parts by x_i pixel's coordinate value (Fig. 4.2) of that. Mirrors are fixed during experiment, we can know mirror position of pixel coordinate value α and β in image already.

$$L(x_i, y_j) = \begin{cases} Part\ 1-1, & x_i < \alpha \\ Part\ 2-2, & x_i > \beta \\ Part\ 1-2, 2-1, & else \end{cases} \quad (1)$$

Among the $Part\ 1-2, Part\ 2-1$,

$$y - y_1 = a(x - x_1), \quad (a = \frac{y_i - y_1}{x_i - x_1})$$

$$\frac{y - y_1 + ax_1}{a} = x, \quad (a = \frac{y_i - y_1}{x_i - x_1})$$

$$L(x_i, y_j) = \begin{cases} Part\ 1-2, & \alpha < x_i < x \\ Part\ 2-1, & x < x_i < \beta \end{cases} \quad (2)$$

The mirror's equation is $ax + by + cz = d$. m is the mirror reflected point of physical point M (Fig.4.3), then two points are same thing originally. t_1 , t_2 and t_3 are scalar value. $\|u_1\| = \|u_2\| = \|u_3\| = 1$. $P_i = (X_d, Y_d)$ is the projection of M on image plan, $P_i' = (X'_d, Y_d)$ is the projection of m on image plan.

$$\begin{bmatrix} X_f \\ Y_f \\ 1 \end{bmatrix} = I \begin{bmatrix} X_d \\ Y_d \\ 1 \end{bmatrix}$$

$$I^{-1}\begin{bmatrix} X_f \\ Y_f \\ 1 \end{bmatrix} = \begin{bmatrix} X_d \\ Y_d \\ 1 \end{bmatrix} \quad (3)$$

(I ; Camera Intrinsic Parameter Matrix)

$$m' = t_3\|n_3\|, \; M = t_2\|u_2\|, \quad (4)$$

$$((m+t_1u_1)-(O+t_2u_2)) \cdot u_1 = 0 \quad (5)$$

$$((m+t_1u_1)-(O+t_2u_2)) \cdot u_2 = 0$$

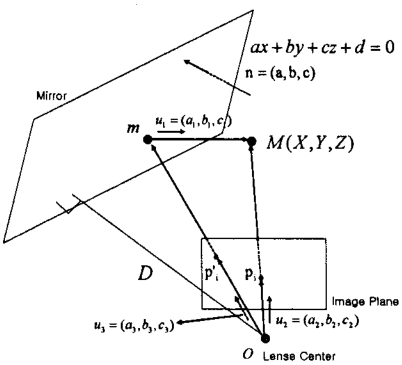

Fig. 4.3 Find shortest distance point between two vectors.

After calculation of t_1 and t_2, we find a 3D point $M(X, Y, Z)$

$$M(X, Y, Z) = \frac{(m+t_1u_1)+(O+t_2u_2)}{2} \quad (6)$$

5. Result and conclusion

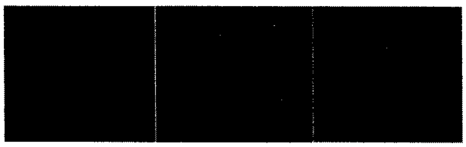

Fig 5.1.(a) Fig 5.1.(b) Fig 5.1.(c)

(a) : About 45degrees' right view of neutral emotion
(b) : Front view of neutral emotion
(c) : About 45degrees' left view of neutral emotion

We can acquire precise 3D coordinate of facial feature points and estimate those movements from front view and mirror-reflected images. We take pictures of each person's changing of face a frame rate more than 4 frames per second in each expression from neutral to 6 different kinds. Using the collection of facial motion dataset, we simulate human's six emotions of expressions successfully.

References

[1]. I-Chen Lin, Jeng-Sheng Yeh, and Ming Ouhyoung, "Realistic 3D Facial Animation Parameters from Mirror-reflected Multi-view Video", Computer Animation, 2001. The Fourteenth Conference on Computer Animation. Proceedings , 7-8 Nov. 2001, Pages:2 - 250

[2]. Tian, Y.-I., Kanade, T., and Cohn, J.F., "Recognizing action units for facial expression analysis", Pattern Analysis and Machine Intelligence, IEEE Transactions on , Volume: 23 , Issue: 2, Feb. 2001, Pages:97 - 115

[3]. J.-Y. Bouguet, Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/calib_doc

[4]. P.Ekman and W.V. Friesen, The Facial Action Coding System : A Technique for the Measurement of Facial Movement. San Francisco, Consulting Psychologists Press, 1976

[5] Abrantes, G.A. and Pereira, F., "MPEG-4 facial animation technology: survey, implementation, and results", Circuits and Systems for Video Technology, IEEE Transactions on Volume 9, Issue 2, March 1999 Page(s):290 - 305

[6] Rafael C. Gonzalez and Richard E. Woods., Digital Image Processing (2002), Second edition. Prentice Hall.

[7] David A. Forsyth and Jean Ponce, Computer Vision A modern approach (2003), Prentice Hall.

[8] Faugeras and Luong, The geometry of multiple images (2001), London, MIT press.

[9] Ramesh Jain, Rangachar Kasturi, and Brian G. Schunck, Machine vision (1995), McGraw-Hill

[10] Jonas Gomes and Luiz Velho, Image processing for computer graphics (1997), New York : Springer

[11] Tsai, R.A, "Versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses", Robotics and Automation, IEEE Journal of [legacy, pre - 1988] , Volume: 3 , Issue: 4 , Aug 1987, Pages:323 - 344