Effectiveness of 3-D Facial Information in Analyzing Facial Expressions

Mitsuru Iwata and Shun'ichi Tano

Graduate School of Information Systems, University of Electro-Communications
5-1, Chofugaoka 1-Chome, Chofu-shi, Tokyo 182-8585, Japan
E-mail: {miwata, tano}@is.uec.ac.jp

***Abstract*-- This paper mentions the use of range image sequence of face for analyzing six facial expressions, i.e., happiness, sadness, surprise, fear, anger, and disgust. In order to assess whether 3-D facial data is effective in analyzing facial expressions, we compare the amount of facial movements obtained from the 2-D facial image sequence and the amount of facial movements obtained from the range image sequence is performed by calculating optical flow. Although the qualitative difference between them is found in distinctive facial expressions such as happiness and surprise, there is no distinctive difference in ambiguous facial expressions such as fear and disgust. Nevertheless, range image can be used effectively to extract a face region from an image with background.**

I. INTRODUCTION

Human face-to-face communication is performed smoothly by using both verbal and nonverbal information such as facial expressions, gestures and so on. It is said that 65% of all the messages exchanged in human communication are nonverbal information [1]. If a recognition and expression system of nonverbal information, e.g., face and gesture, could be developed at the same level as human does, the human-computer interaction became smooth. Therefore many researches on anthropomorphic user interface have been done [2].

Facial expression is most powerful and important means to for human to communicate their emotions. Although it is difficult to recognize facial expressions absolute automatically by a computer, there have been many algorithms to recognize human facial expressions such as using the movements of facial feature points [3], using the 'Action Units' which Ekman proposed [4], and consideration of facial muscle movements [5]. These researches recognize facial expressions from a 2-D facial image. Although there are researches using a 3-D geometric facial model [6-8], these match with a 2-D input facial image and an internal 3-D facial model, and estimate the 3-D motion of facial muscles.

By the recent technological progress in semiconductor electronics, high-speed stereo vision system (stereo camera) which outputs a range image at the video rate can be available. In this paper, the amount of movements of facial components obtained from a 2-D image and a range image are analyzed in order to check whether the range image is effective in recognition of facial expressions.

II. PREVIOUS RECOGNITION MODEL OF FACIAL EXPRESSIONS FROM 2-D LINE DRAWING

Ekman had examined the relationship between facial expressions and his/her emotions. From the results, it has been said that humans have six basic emotions, i.e., happiness, sadness, surprise, fear, anger, and disgust, and facial expressions corresponding to these emotions are common in the world.

We have already developed a system which recognizes a degree of each emotion, i.e. facial expression, from 2-D facial line drawing using a neural network model. To construct a neural network model, questionnaire survey which examines the relation between facial attributes and facial expressions was carried out. 27 subjects were asked to read 216 line-drawn facial expressions and check the degree of each facial expression. The total number of valid data is 420. The half of these data was assigned to the training data set and the remainder was assigned to the test data set. We used three-layer neural network with one hidden layer. The input layer has six nodes which represent attributes of line-drawn face, i.e., 'size of eye', 'shape of eye' and so on, and the output layer has six nodes which give the degrees of six facial expressions, i.e., 'happiness', 'sadness', and so on. From experiments, we have found six hidden units are sufficient.

From experiment results of the system evaluation, it was

found that the facial expressions which the system recognizes were mostly agreed to. And it was also found that the recognition rates of fear and disgust are lower than those of others. This tendency was seen in some recognition systems using real facial images [9]. While an average of recognition rates within 4 facial expressions, i.e., happiness, sadness, anger, and surprise, is about 87%, an average of recognition rates within all 6 facial expressions falls to about 62%.

It is thought that the recognition rates of former systems were low, since 2-D facial feature points movement on which 3-D facial movement is projected was used.

III. Range Image of Facial Expression

In order to assess whether the range image is effective in analyzing facial expressions, 2-D facial images and its' range images are collected by a stereo camera.

A. Experimental Equipment

A high-speed stereo vision system connected to a standard PC is used to record the subjects' facial expressions. Figure 1 illustrates the stereo vision system equipped with 9 CCD lenses. The system can calculate depth of all pixels in a center RGB image and generate a range image sequence at 30 frames per second. Examples of the center RGB image recorded by the center CCD lens and its range image presented in false-color (or pseudo-color) are shown in figure 2. The image size is 280 x 200 pixels.

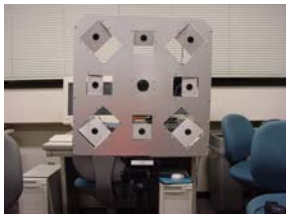
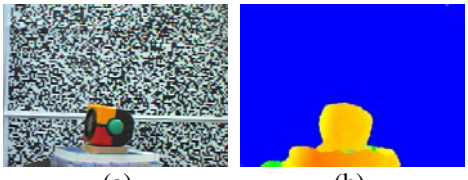


Fig. 1. Stereo vision system with 9 CCD lens.



(a) (b)

Fig. 2. Center image (a) and its range image (b).

B. Method

Ekman defined six basic emotions, i.e., happiness, sadness, surprise, fear, anger, and disgust, and described each basic emotion in terms of a facial expression that uniquely characterizes that emotion. For that reason image sequences by 13 Japanese subjects showing six distinct facial expressions have been recorded by a 3-D camera under constant illumination. Recording time was 20 seconds per one facial expression. Subjects were instructed to change their face into the specified facial expression from a neutral facial expression and then return to the neutral facial expression gradually within 20 seconds. They did not have practice in expressing facial expressions. However, they were required to express each facial expression naturally.

C. Calculation of Optical Flow

First of all the range image is used to extract a face region from the center RGB image with background. Since the subjects were sitting on about 1.4m distance from the camera, all the pixels of the center image further than the distance were deleted. The remained pixels constitute a region of the upper part of the body.

To calculate the magnitude and the direction of movement, i.e., optical flow, between an initial neutral and a specified facial expression image, we use the optical flow algorithm. A popular technique for optical flow is the block-matching (BM) algorithm as shown in figure 3. This method takes a template block P in the present frame and searches for a similar block Q in the next frame which minimizes absolute difference of pixels D in the matching blocks. The relative locations of blocks in present and the next frames $Q-P$ are then characterized as a motion vector.

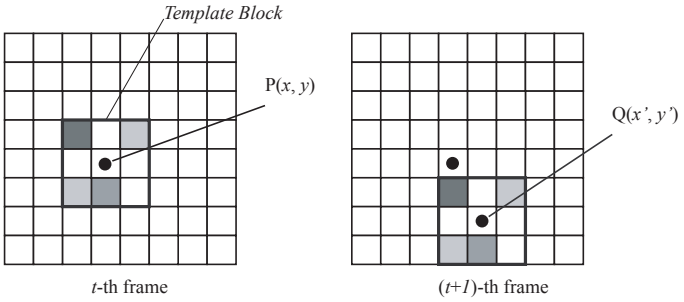


Fig. 3. Block matching algorithm.

The absolute difference of pixels D is defined as the following formula (1),

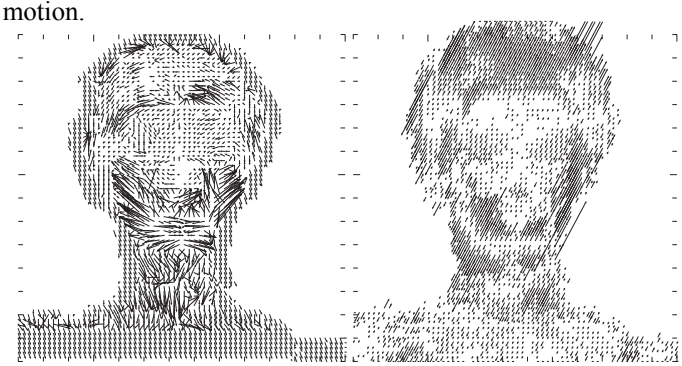
$$D = \sum_{m=-L}^{L} \sum_{n=-L}^{L} \left| I(x+m, y+n, t) - I(x'+m, y'+n, t+1) \right| \qquad (1)$$

where $I(x, y, t)$ is a brightness value of point (x, y) in the t-th frame of center image sequence, and $I(x', y', t+1)$ is a brightness value of point (x', y') in the (t+1)-th frame of center image sequence. The size of the template block is $(2L+1) \times (2L+1)$ *pixels*.

In this paper, the size of the template block is 5 x 5 pixels ($L = 2$). And the size of a search area in the (t+1)-th frame is 22 x 22 pixels.

D. Results

The magnitude and the direction of movement are calculated while expression changes from an initial neutral to the peak of a specified facial expression. An example of the 2-D optical flows calculated from a center RGB image sequence of happiness is shown in figure 4(a). Figure 4(b) illustrates optical flows in a depth direction calculated from the corresponding range image sequence. The arrow which points to the lower left in figure 4(b) represents a forward motion of the part, and the arrow which points to the upper right represents a backward motion.



(a) (b)

Fig. 4. Flow vectors in x-y direction (a) and depth direction (b).

From these results, the following tendencies are found about each facial expression.

HAPPINESS

- The area of a cheek and angle of mouth move outside above, and the upper lip and nasolabial sulcus also move upward.
- The area of a cheek, angle of mouth, and jaw become hollow.
- There is a difference between the amount of movements in *x-y* direction and that in depth direction around a mouth. There are few amounts of movements in depth direction on the lip. The amount of movements in depth direction increases especially in the part of a cheek and an angle of mouth.

SADNESS

- The part of a mouth moves downward and especially the part of eyes moves downward.
- The head droops and the lower part of a face moves forward.
- In the region of eyes, there are few amounts of movements in depth direction rather than that in *x-y* direction.

SURPRISE

- The whole face moves upwards. Although angles of mouth move upwards, the lower lip and a jaw move downward.
- There are many subjects whose eyes, eyebrows, and cheeks are dented.
- There is a difference between the amount of movements in *x-y* direction and that in depth direction around the cheeks and under the eyes.

ANGER

- The eyes and the eyebrows move in the *x-y* direction greatly.
- The surrounding area of eyes and eyebrows swells. The surrounding of a mouth dents. There are many subjects whose whole face moves forward.

FEAR

- Angle of mouth and eyes move downward, and eyebrows rise. A cheek moves.
- The surroundings of a mouth are dented (the lower part of a face goes back). Eyes, eyebrows, and a cheek are also dented.

DISGUST

- A cheek and a mouth move greatly. A mouth moves downward and eyes move downward and outside.
- The whole face move backward. The surrounding areas of a mouth and a cheek are dented greatly.

Though the qualitative differences between the amount of

movements in *x*-*y* direction and depth direction are found in facial expression of happiness or surprise, the remarkable qualitative differences have been found neither in the facial expression of fear nor the expression of disgust.

IV. Advantages of Range Image

We analyze advantages of using range information in facial expression analysis, gesture recognition and so on.

- In facial expression recognition, distinctive regions with many 3-D facial motions are easily extracted at the pixel level.
- It is easy to extract a face from an image with background.
- Since it is easy to pinpoint the position of a nose, it becomes easy to detect the locations of facial components.
- Even if the head moves somewhat, the frontal-view image can be reconstructed.
- Real-time 3D gesture recognition system will be developed more. Gestures have been one of the important interaction media in human-computer interaction.

V. Conclusion

For a preliminary analysis, we thought that the recognition rate of the former system was low because images by which the 3-dimensional facial shape was projected on the 2-dimensional plane were used. In this paper a comparison between the amount of movements of facial parts obtained from the 2-dimensional image and 3-dimensional image was performed in order to check whether the 3-dimensional facial data are effective in recognition of facial expression. Although the difference between them is found in the distinctive facial expressions such as happiness and surprise, there is no distinctive difference between them in the ambiguous facial expressions such as fear and disgust.

Nevertheless, it was thought that range information can be used effectively in extraction of distinctive regions with many 3-D facial motions, extraction of a face from an image with background, and detection of the locations of facial components since it is easy to pinpoint the position of a nose.

We will analyze the magnitude and the direction of facial movements obtained from the 2-D image sequence and from the range image sequence.

References

[1] T. Kurokawa, Nonverbal Interface, Ohmsha, Tokyo, 1994. (in Japanese)

[2] S. Hayamizu and T. Takezawa, "Trends in research on multimodal information integration system", Journal of Japanese Society for Artificial Intelligence, vol. 13, no. 2, pp. 206-211, 1998. (in Japanese)

[3] M. Wang, Y. Iwai, and M. Yachida, "Expression recognition from time-sequential facial images by use of expression change model", Proc. of the Third International Conference on Automatic Face and Gesture Recognition, pp. 324-329, Nara, Japan, 1998.

[4] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, "Classifying facial actions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 10, pp. 974-989, 1999.

[5] T. Ishikawa, S. Morishima, and D. Terzopoulos, "3D face expression estimation and generation from 2D image based on a physically constraint model", IEICE Trans. Inf. & Syst., vol. E83-D, no. 2, pp. 251-258, 2000.

[6] M. Kaneko, Y. Hatori, and A. Koike, "Coding of facial images based on 3-D model of head and analysis of shape changes in input image sequence", Transactions of The Institute of Electronics, Information and Communication Engineers B, vol. J71-B, no. 12, pp. 1554-1563, 1988. (in Japanese)

[7] C. S. Choi, H. Harashima, and T. Takabe, "Analysis of facial expressions using three-dimensional facial model", Transactions of The Institute of Electronics, Information and Communication Engineers D-II, vol. J74-D-II, no. 6, pp. 766-777, 1991. (in Japanese)

[8] H. Ohta, H. Saji, and H. Nakatani, "Recognition of facial expressions using muscle-based feature models", Transactions of The Institute of Electronics, Information and Communication Engineers D-II, vol. J82-D-II, no. 7, pp. 1129-1139, 1999. (in Japanese)

[9] H. Ohta, "Research on the analysis of facial movements and recognition of facial expressions using facial structure", doctoral dissertation, Graduate School of Science and Engineering, Shizuoka University, Japan, 1998.12. (in Japanese)