A High Speed Face Measurement System

Kazuhide HASEGAWA, Kazuyuki HATTORI and Yukio SATO
Department of Electrical and Computer Engineering,
Nagoya Institute of Technology
Gokiso, Showa, Nagoya, Japan, 466-8555
kazuhide@hilbert.elcom.nitech.ac.jp, sato@elcom.nitech.ac.jp

Abstract

In this paper, A rangefinder system which measures both 3-D shape and color texture of a human face is described. Measurement speed is an important factor for human body sensing, to reduce a burden on the subject. In our system, high-speed face measurement is achieved by using a space-encoding method with laser scanning. This system equips a color CCD camera as an image input device which grabs both range and color texture images. There is the possibility that laser radiation is harmful for human eyes, and our system meets the Japanese radiation safety standards for laser products. In our experiment, this system obtained 3-D shape and color texture information in about 1 second. The measurement error on shape was less than 1 mm.

1 Introduction

The research focused on processing face information has been extensively studied in many years. Most of the research has done with intensity images of the face. However, with the progress of shape measurement, virtual reality and computer graphics technology, there has been growing interest in new fields, such as man-machine interfaces, security and teleconferencing. In these fields, obtaining 3-D face model is highly demanded [4][5]. A face measurement system described in this paper can be used in the above application fields.

In recent years, rangefinder systems obtaining both range and intensity images, especially the color texture on the object surface, have been developed and used for research. For example, Suenaga used a system based on triangulation using slit-ray[3]. However, this measurement system takes about 15 seconds to measure the upper part of a body.

One problem in face sensing, not only the face but also the body, is measurement time. Obtaining a range image takes time compared to an intensity image. Thus, a person must remain stationary during the measurement. This places a burden on the subject, but less attention has been paid. Even if the actual measurement time is about 10 seconds, the person unconsciously moves slightly. Therefore, it is difficult to obtain stable measurement results.

In this paper, a high speed measurement system for 3-D face model at a speed high enough to reduce the burden on the subject is described. In our system, high speed measurement is achieved by using a space-encoding method with laser scanning. Through this method, both high-resolution range and intensity images are obtained in 1 second. We also describe that our system meets Japanese Radiation Safety Standards for laser products.

2 High Speed face Measurement System

2.1 Concepts for the System

We considered the following properties for the face measurement system.

(a) Obtaining both 3-D shape and color texture.

(b) Occlusion-less measurement.

(c) High-speed measurement.

Property (a) is demanded in some application fields, such as computer graphics. Property (b) means that a whole facial surface must be measured from one view point. Based on the active stereo method, there are some unmeasured areas that correspond to no projection or unobserved areas caused by occlusion. Property (c) is an important factor, as mentioned before, if the object, such as a human being, cannot keep a stationary pose for a long time. Movement of the object during the measurement causes measurement errors. One solution for this problem is to use a fast measuring method to obtain 3-D shape information.

Considering above properties, we developed a prototype of the face measurement system. We describe this system as follows.

2.2 Measurement Principle and Optical Section

The measurement time depends on a measurement principle. The space-encoding method is one of the fastest principles when using an ordinary video camera as an image input device. Our system is based on this space-encoding method with laser scanning [1][2].

The developed face measurement system is shown in Fig. 1. The optical section in this system has a color CCD camera at the center, two laser scanners at both sides of the camera and a white light source located above the camera. The color camera (CN-411, ELMO) plays two roles. One role is color image obtainment, and the other is 3-D shape measurement. By using one camera to obtain both images, the range and color images correspond to each other by pixel order. Thus, property (a) is satisfied. With two laser scanners, patterns are projected to the face from both sides of the camera to complement each other. As a result, the lack of range data caused by occlusion is filled. This function satisfies property (b). This laser scanner equips a semiconductor laser (680 nm, 40 mW), a cylindrical lens and a polygonal mirror (12-surface, 3600 rpm). By scanning the laser slit with the mirror and switching the laser at the proper timing, the laser scanner generates and projects encoded stripe patterns to an object. Thus, spatial light patterns are generated by temporal switching patterns. We can obtain a range image in 17/60 seconds with each scanner and a color image in 1/30 seconds. Therefore, both range and color images of a human face are measured in 36/60 seconds at the highest speed. This speed is sufficient to satisfy property (c). 3-D shape measurement must be executed in a dark room, because visible lasers and a color CCD camera are used. Thus, the white light source (fluorescent lamp) is used for obtaining the color image.

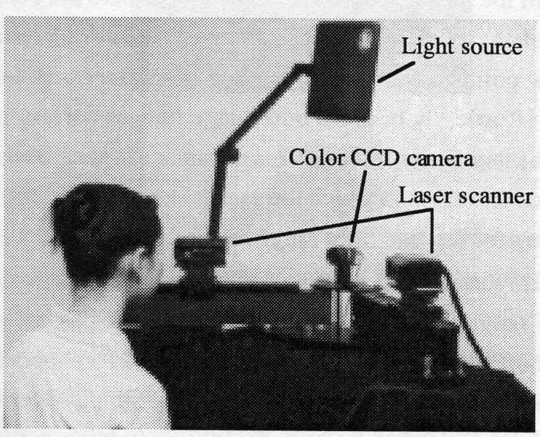

Fig.1 Face measuring system

2.3 Image Processing and Control Section

The image processing and control sections of our system are built into a personal computer as extension boards. These sections are shown in Fig. 2.

The image processing section has a decoding part and a memory part. The decoding part converts stripe pattern images obtained by the camera to a decoded image at a real time. The memory part stores the decoded image and the color texture image. In our system, laser scanners project complementary patterns for thresholding binary patterns.

The control section has a synchronous signal generator, a laser scanner controller and a light source controller. The synchronous signal generator produces various signals for the entire system and supplies them. The laser scanner controller manages the speed of the polygonal mirrors, and switches the semiconductor lasers for projecting the encoded stripe patterns. The light source controller turns the light source on when obtaining a color image and turns it off when obtaining range images.

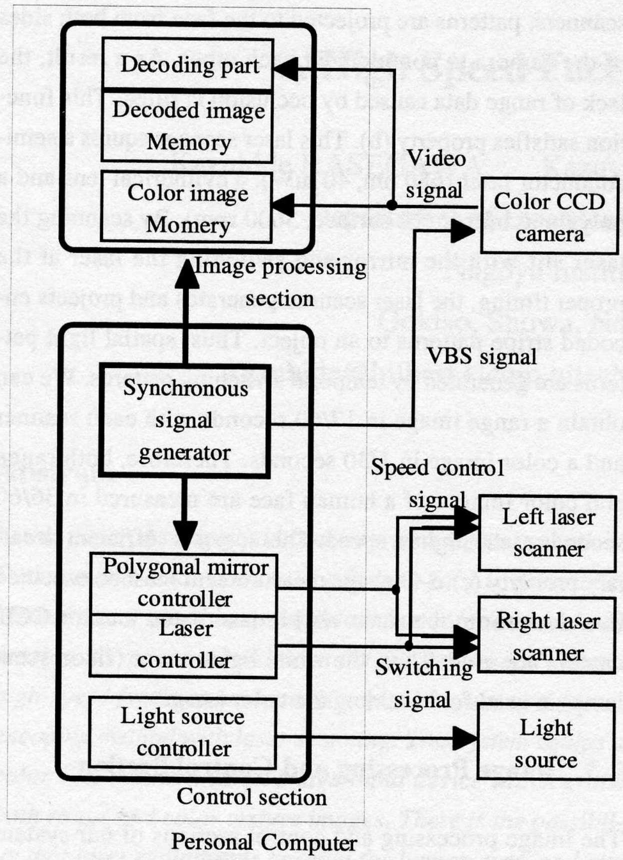

Fig.2 Configuration of image processing and control section

2.4 Measurement Procedure and Measurement Time

Our system uses the following procedures to execute a measurement. Numerical values in parentheses are the processing time for each step.

Step 1: Turn the light source on beforehand.

Step 2: Obtain a color image, and then turn the light off. (1/30 seconds)

Step 3: Measure a range image with the right side laser scanner. (17/60 seconds)

Step 4: Measure a range image with the left side laser scanner. (17/60 seconds)

In the first step, the white light source is turned on beforehand. This step is to prevent the subject from moving. Because, a person might be moved by the reflex movement if suddenly projected laser patterns in a dark room. Another reason is to adjust the position of the face. Moreover, to shorten the measurement time, because the light source in prototype system takes time to turn on. In the next step, a color image is obtained and the light is turned off. Then, we let the laser scanners at both sides work successively to measure the range images.

The total of each steps' processing time and the lag time between steps is the measurement time. If there is no lag time in switching steps, both range and color images are obtained in 36/60 seconds. In our actual system, entire measurement time is about 1 second.

3 3-D Face Model Processing

3.1 Range Image Integration

We need to calibrate the optical system beforehand. The calibration process was carried out as follows. A standard plane printed prearranged dots was mounted on a slide stage. Then, intensity images of the dots and the projected stripe patterns were obtained in 15 steps of 20 mm interval. With these images, internal and external camera [6][7], distortion correction [8] and laser projection parameters were calculated. As a result of calibration, the average depth errors are estimated between 0.93 mm and 0.96 mm.

Images obtained by the system and the processing procedure are shown in Fig. 3. The obtained range images have some unmeasured areas that correspond to no projection or unobserved areas caused by occlusion. Thus, the two range images are integrated into one face range image to fill the lack of range data. Generally, there are complex problems to integrate range images obtained from different view points. In our system, integration process is carried out effortlessly because both range images correspond to each other by pixel order. This advantage is generated from the construction in which the same camera is used for measuring the left and right range images.

However, there are small differing values between the two range images. This difference is approximately 0.6 mm on average when a white polystyrene face model was measured. Therefore, if two images are simply integrated, jump edges might be occurred on the surface. Thus, the two

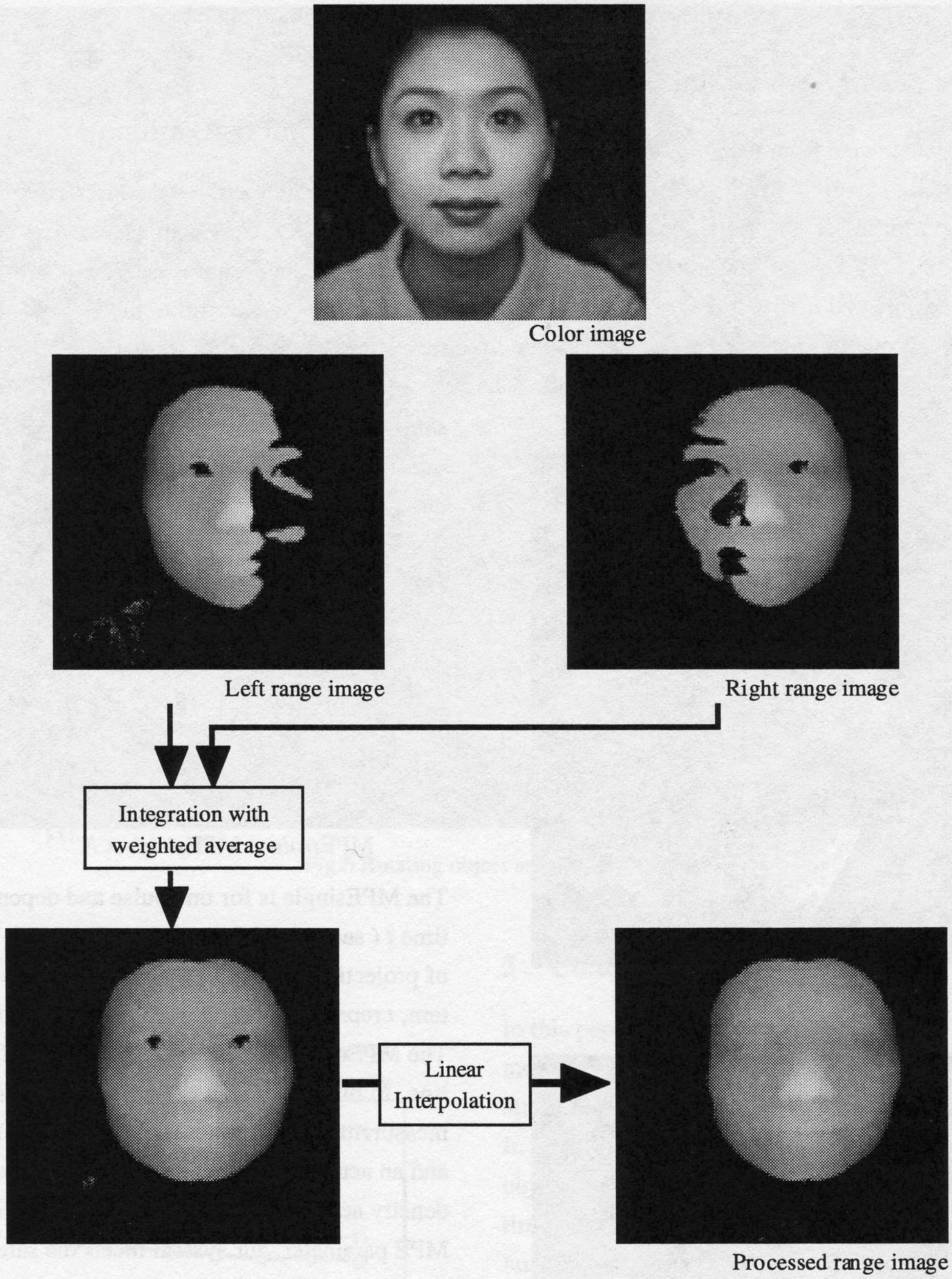

FIg.3 Obatained images and processing procedure

images are integrated smoothly with a weighted average. After integration, the average error of the overlapped areas was 0.70 mm when the standard plane used in calibration process was measured. In case of a white polystyrene cylinder that is more similar to a face, the average error was 0.63 mm. The measurement error on shape was estimated less than 1 mm, since the areas measured by laser scanner at one side are included.

Despite of weighted averaging, some unmeasured areas still exist. Because, the dark color on the surface, such as pupils or eyebrows, does not reflect the projected laser patterns to the camera enough to measure. Thus, these areas are filled with a linear interpolation based on the neighborhood pixel value. Finally, an occlusion less and smooth face range image (512 by 242 pixels) is obtained.

3.2 3-D Face Model Reconstruction

3-D display reconstructed from the integrated range image is shown in Fig. 4. The mesh model in Fig. 4 is generated with an adaptive mesh algorithm [9]; mesh sizes are decomposed adaptively to the surface shape. Mapping the obtained color texture to the mesh model is shown in Fig. 5. And the results of rotating in nine directions are shown in Fig. 6. As can be seen, the 3-D face model retains the 3-D structure and color texture information of the face. Thus, it is very simple to generate 3-D images even when the view point is shifted or the light source angle changed.

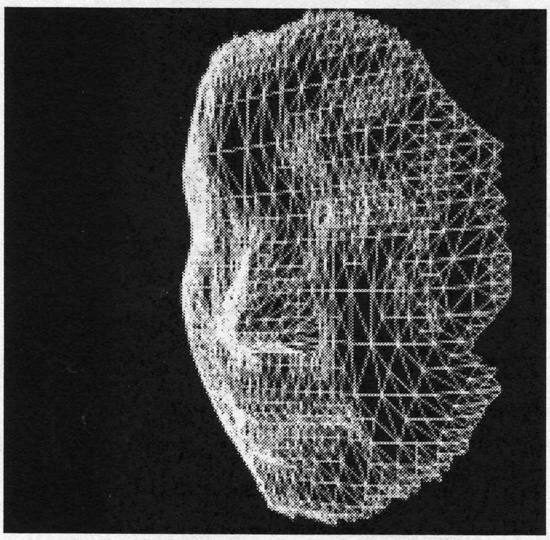

Fig.4 3D display

Fig.5 3D display with texture mapping

4 Laser Safety Examination

We now need to examine the laser safety for the measurement to human beings. Our system uses semiconductor lasers for shape measurement. There is the possibility that laser radiation is harmful to the human body, especially to the retina. The Japanese Radiation Safety Standards for laser products are regulated in JIS C 6802. In JIS C 6802, the safety standard is regulated with the MPE (Maximum Permission Exposure) parameters (J/m^2). The MPE parameters are classified by the wave length of laser, exposure time, radiation form, observation state, and defined by following three equations.

$$\mathrm{MPE}single = \begin{cases} 5\times 10^{-3} : 10^{-9} \le t \le 1.8\times 10^{-5} \\ 18\times t^{0.75} : 1.8\times 10^{-5} \le t \le 10 \end{cases}$$

$$\mathrm{MPE\text{-}T} = 18\times t^{0.75} : 1.8\times 10^{-5} \le t \le 10$$

$$\mathrm{MPE}train = \mathrm{MPE}single \times N^{1/4}$$

The MPEsingle is for one pulse and depends on projecting time t (seconds) of this pulse. The MPE-T is in condition of projecting same time as an entire pulse line. In our system, t represents the entire measurement time (1 second). The MPEtrain is considered the number of the pulses in the line. In our system, the number of pulses N is 204 in the measurement procedure. Fig. 7 shows the MPE parameters and an actual projecting power of the system. If the energy density actually projected by our system is less than each MPE parameter, our system meets the safety standards. As can be seen from Fig. 7, the energy density is much less than the MPE parameters on each distance from the laser scanners to the retinas. Therefore, our system is definitely safe to the human eyes.

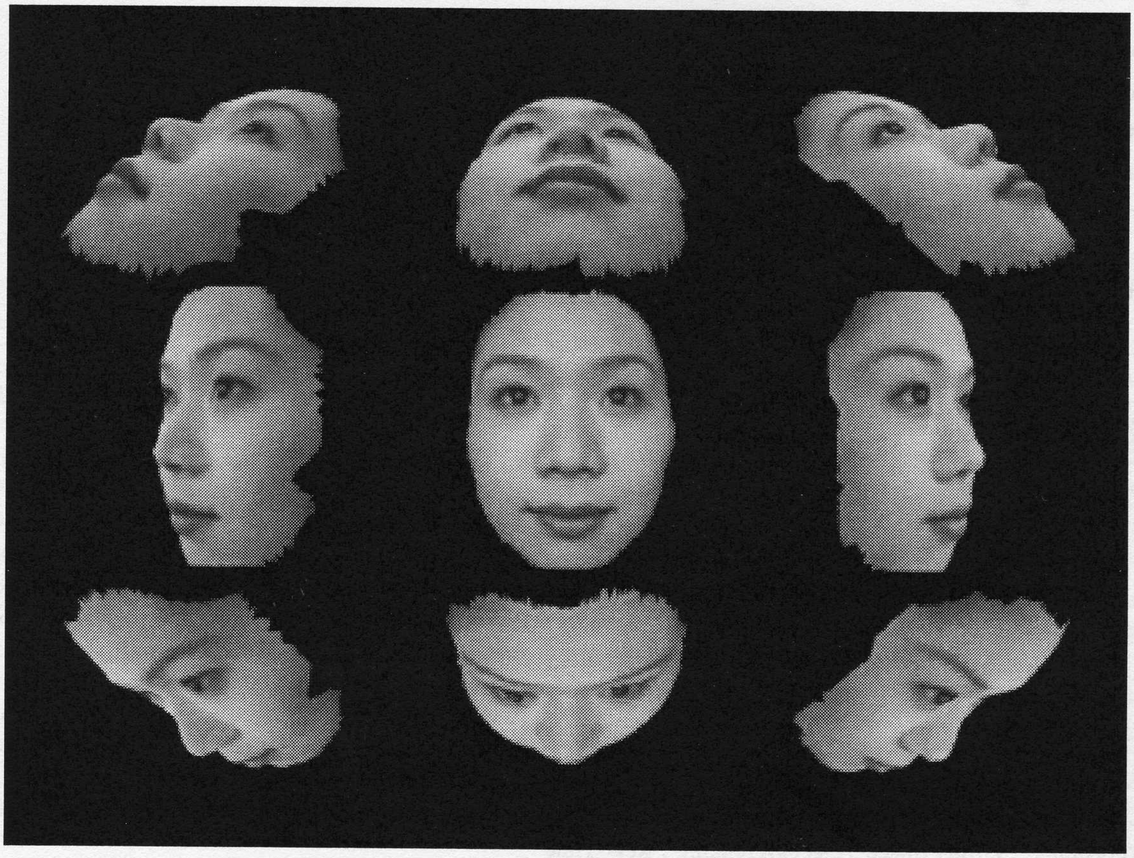

Fig.6 Rotating object and 3D display

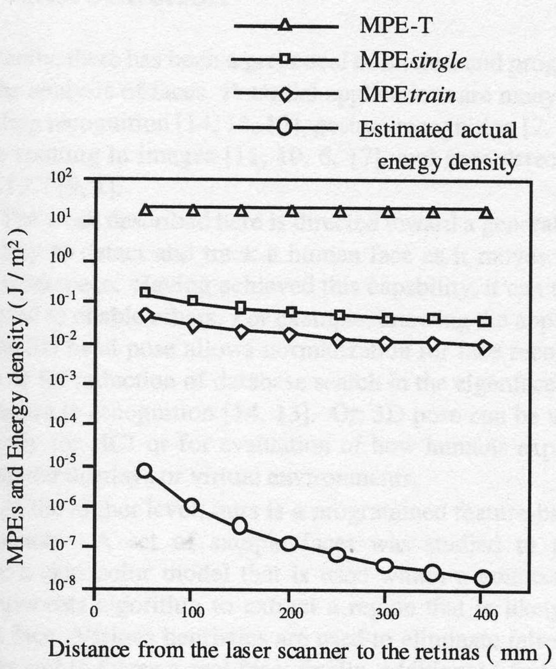

Fig7. MPE parameters and energy density of our system

5 Conclusion

In this paper, we presented a prototype of high-speed face measurement system. This measurement system can ideally obtain both range and color images in 36/60 seconds. In our experiment, the prototype system measured a 3-D object shape and texture in 1 second. Despite of some lag time, this measurement speed is high enough to obtain a human face without forcing a person in a stationary pose for a long time.

Our system must be practical enough for measuring 3-D human faces. Then, it is very simple to input a 3-D face shape together with color texture into a computer and reconstruct the 3-D face model in it. Eventually, 3-D face models of individuals will be demanded in various kinds of new application fields.

Acknowledgement

We would like to thanks CKD Corporation and Ohshima Prototype Engineering for their help on system development. This research is supported by the Grant-in-Aid for Scientific Research (B) of Ministry of Education, Science, Sports and Culture, Japan.

References

[1] K. Hattori and Y. Sato, "Handy Rangefinder for Active Robot Vision", In Proc. of Int. Conf. on Robotics and Automation, pages 1423 - 1428, May, 1995

[2] K. Hattori and Y. Sato, "Accurate Rangefinder with Laser Patter Shifting", Proc. of ICPR, volume c, pages 849 - 853, 1996

[3] Y. Suenaga and Y. Watanabe, "A Method for Synchronized Acquisition of Cylindrical Range and Color Data", Trans. of IEICE, 25(12) : pages 3407 - 3416, 1991

[4] J. Ohya and F. Kishino, "Detecting Facial Expressions from Face Images Using a Genetic Algorithm", In Proc. of Int. Conf. on Pattern Recognition, volume c, pages 649 - 653, Aug 1996

[5] J. Ohya and Y. Kitamura and H. Ishii and F. Kishino and N. Terashima, "Virtual Space Teleconferencing : Real-Time Reproduction of 3-D Human Image", Journal of Visual Communication and Image Representation, 6(1) : 1 - 25, Mar 1991

[6] R. Y. Tsai, "A Versatile Camera Clibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses", Computer Science, No.10, 1985

[7] R. K. Lenz, R. Y. Tsai, "Techniques for Calibrationof the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology", Proc.IEEE ICRA, pages 68 - 75, 1987

[8] J. Weng, P. Cohen, M. Herniou, "Camera Calibration with Distortion Models and Accuracy Evaluation", IEEE Trans. PAMI, 14, 10, pages 965 - 980, 1992

[9] H. Tanaka and F. Kishino, "Adaptive Mesh Generation for Surface Reconstruction", In Proc. of CVPR, pages 88 - 94, Aug 1993