MVA'90 IAPR Workshop on Machine Vision Applications Nov. 28-30,1990, Tokyo

A METHOD FOR THE SYNCHRONIZED ACQUISITION OF CYLINDRICAL RANGE AND COLOR DATA

Yasuhito SUENAGA and Yasuhiko WATANABE
NTT Human Interface Laboratories

1-2356, Take, Yokosuka Kanagawa, 238-03 Japan
Tel: +81-468-59-2864 Fax: +81-468-59-2898
E-mail: {suenaga, watanabe}%nttcvg.ntt.jp@relay.cs.net

ABSTRACT

This paper presents a method of 3D human face measurement using a newly developed device that acquires 3D range data and surface color data at the same time. The cylindrical range data is measured by a laser light source and a CCD sensor with a resolution of 512 vertical scan lines, 256 points per scan line. The color data is acquired as a cylindrical projection image having 512 by 256 pixels, 24 bits/pixel (8 bits each for red, green, and blue). The scanner has been successfully applied to the measurement of human faces and other 3D objects.

INTRODUCTION

The authors have been doing various research on computer vision and graphics for the recognition and synthesis of human images. These areas of research are important in realizing a better human interface and developing a model based CODEC in future [Mase'89] [Aki'86] [Aki'90] [Wat'89].

This paper presents a method of 3D sensing using a newly developed device that acquires 3D range data and surface color data at the same time. There have already been many reports of 3D range finders and scanners [Ino'90]. Almost all of them are designed for the detection of three dimensional distance or range data only. Although ordinary color data is commonly regarded as easy to acquire with ordinary color TV camera or color scanners, the acquisition is not so easy as explained below.

Roughly speaking, for such applications as recognition and synthesis of various objects, two kinds of data are used: 3D shape (range) data and texture (color) data. Usually, these two data types are measured separately, using different kinds of acquisition systems. Due to time differences, especially in the case of movable objects like human faces, it is sometimes difficult to match these two kinds of data. Moreover, camera angles and lighting conditions may differ for each type of acquisition. Though adjustment may be done to some extent, acquiring the consistent data for 3D objects having various shapes and colors is practically impossible.

Thus, it is desirable to acquire range and color data at the same time, since it becomes possible to acquire precise images of various objects from the synchronized 3D range and color data. The scanner is designed based on these considerations.

This paper presents a method of 3D sensing using the scanner for three kinds of applications: (1) facial expression generation, (2) 3D area extraction and (3) face image database preparation.

SCANNER

The Echo scanner (4020/PS Rapid 3D Digitizer), manufactured by Cyberware Laboratories, USA, is a laser scanner for measuring cylindrical range data of objects, as illustrated in Figure 1. According to our specifications, the manufacturer installed a TV-camera based color data acquisition subsystem in the digitizer unit of the scanner, while preserving its original functions. The object to be measured is placed on a center table, and an arm with a laser light source, a CCD sensor and a CCD color TV camera rotates 360 degrees horizontally around it, synchronously acquiring cylindrical range and color data. It takes the scanner fifteen seconds to acquire the data.

The cylindrical range data is measured by a laser light source and a CCD sensor with a resolution of 512 vertical scan lines, and 256 points per scan line. Measurement resolution is within 0.7 mm when measuring a cylinder 350 mm high and 350 mm in diameter. Actually, the cylindrical range data is converted to 512 x 256 sets of x, y, and z coordinate values. Surface color data is measured by a CCD color TV camera. The color data is acquired as a cylindrical projection image having 512 by 256 pixels, 24 bits/pixel (8 bits each for red, green, and blue). Since the cylindrical range and surface color are measured at the same time, all the data is acquired in a synchronized form.

Figures 2-1, 2-2, 2-3 and 2-4 show the measured results of a terrestrial globe, a doll, a pot and a ski boot, respectively. In each of the figures, (a) shows the wire-frame of the 3D range data, and (b) shows the cylindrical color data. Both range data and color data are easily processed to rebuild the complete 3D shape (with color) as shown in (c) in each figure. Figures 3 (a) and (b) show an example of acquired cylindrical range and color data for a human head. Figures 3 (c) and (d) show the human head images generated from the combination of range and color data.

APPLICATIONS

Facial Expression

There are many Computer Graphics (CG) applications producing facial expressions by controlling only a few points on a 3D face model. The texture mapping technique is used to produce a realistic appearance. Synthesized facial expressions were produced using the data acquired by the scanner. Since the 3D shape data includes both 3D range and color data, realistic facial expressions are easily produced without texture mapping as in conventional CG techniques. Figure 4 shows a smiling facial expression synthesized by controlling only four points.

3D Area Extraction

As described in the previous section, the acquired 3D shape data can be controlled using the color data. Since the color data is composed of a 2D (inverse panorama format) image, conventional image processing technologies can be applied to this data. Figure 5 (a) shows a face that was scanned. Figure 5 (b) shows the color data taken by the scanner. Figure 5 (c) shows the image clustered in RGB color space. The 3D shape data could be easily extracted using this clustered color image data. Figure 6 shows the extracted 3D area recognized as "skin area." This proves that the 2D color data is very useful in extracting the desired 3D shape from original data.

Face Image Database

In the development of face recognition systems, the most important thing is the construction of a face image database for matching which contains a big number of face images. Since the facial pattern matching must be done under various conditions, the preparation of all the template face images is not so easy.

When using the scanner, the 3D shape can be used for standardization of face position. Figure 7 shows the determination of standard axes in our experiment. Both of translation and rotations are done in 3D space. This process produces template face images for matching. Figures 8 shows the synthesized color data from the standardized 3D shape. This color data can be used as a face image in the database for matching.

CONCLUSION

The authors have been using the scanner mainly for acquiring 3D data of human faces, hands, legs, and trunks. The proposed method is very useful in making a complete 3D image of human body database needed for realizing better human interface and developing model based CODEC. The method proposed in this paper is a general purpose 3D data acquisition method which has opened a way for us to acquire highly accurate 3D data by directly measuring shape and color of objects in synchronization.

It is widely applicable to the preparation of 3D databases of various objects needed for many applications.

Acknowledgements

The authors wish to thank Mr. David Addleman, President, Cyberware Laboratory Inc. for his cooperation in realizing the scanner. The authors thank Dr. Takahiko Kamae, Director, NTT HI Labs, Dr. Yukio Kobayashi, Executive Manager, Visual Perception Lab, and the members of the VP Lab, for their encouragement and valuable discussions. The authors also thank Dr. Richard Wallace for improving the manuscript.

References

[Aki'86] T.Akimoto, "Expressive facial animation by 3D jaw model and automatic shape modification", NICOGRAPH'86, pp207-213, 1986 (in Japanese).

[Aki'90] T.Akimoto, R.S.Wallace, Y.Suenaga, "Automatic creation of facial model for generating facial images", IAPR Workshop on MVA'90 , Tokyo, November 1990.

[Ino'90] S.Inokuchi, "State of the art of 3D sensing technologies", Gazo Lab. , Vol.1, No.4, pp44-47, April 1990 (in Japanese).

[Mase'89] K.Mase, Y.Watanabe, Y.Suenaga, "A real-time head motion detection system", SPIE Workshop on Sensing and Reconstruction of Three-Dimensional Objects and Scenes, Vol.1290, pp262-269, Santa Clara, CA, February 1989.

[Wat'89] Y.Watanabe, Y. Suenaga, "Drawing human hair using wisp model", CG International'89, pp691-700, 1989.

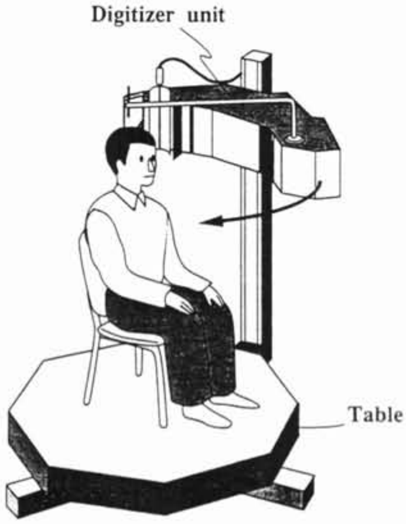


Fig. 1 3D face data acquisition by the synchronized cylindrical range and data scanner.

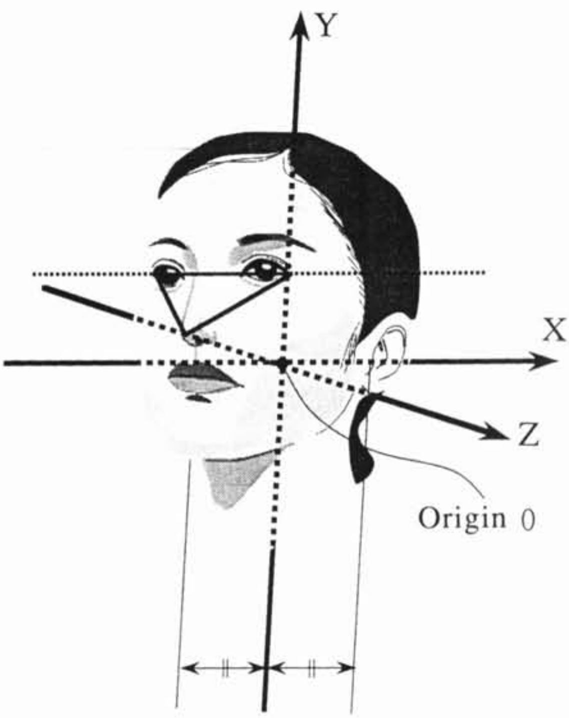


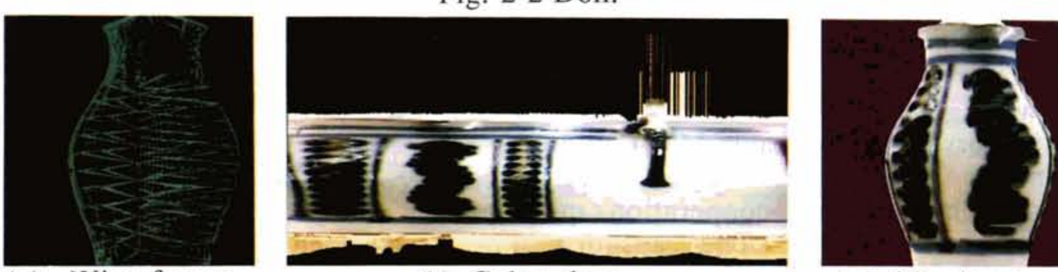
Fig. 7 Standard axes.

(a) Wire-frame (b) Color data (c) 3D shape

Fig. 2-1 Terrestrial globe.

(a) Wire-frame (b) Color data (c) 3D shape

Fig. 2-2 Doll.



(a) Wire-frame (b) Color data (c) 3D shape

Fig. 2-3 Pot.

(a) Wire-frame (b) Color data (c) 3D shape

Fig. 2-4 Ski boot.

Fig. 4 Facial expression (smile).

(a) Wire-frame

(b) Color data

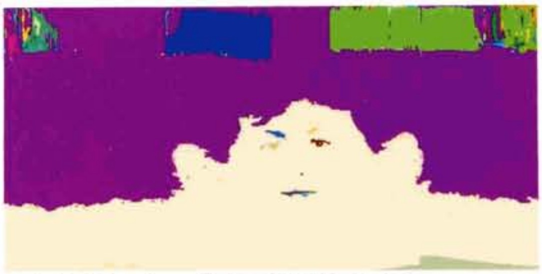
(c) (front view)

(d) (side view)

Fig. 3 Human head.

(a) Face image

(b) Color data



(c) Clustered in RGB space

Fig. 5 Area extraction.

Fig. 6 Extracted area.

Fig. 8 Synthesized color data