MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

A Support System for the Visually Disabled to Recognize 3D Objects

Yoshihiro Kawai Fumiaki Tomita *
Intelligent Systems Division
Electrotechnical Laboratory

Abstract

We have been developing a support system for the visually disabled to actively recognize three-dimensional objects or environments, an application system of 3D computer vision. This is a total system which has input, processing, and output functions. Using a stereo camera system as an input device allows 3D visual information to be obtained. The visual information is converted into tactile and auditory information which can be understood easily by the visually disabled. As one of the output devices, we have developed an interactive tactile display, which presents visual patterns by tactile pins arranged in two-dimensional format. The pin height can be set to several levels to increase the touch information and to display a 3D surface shape. Also, each pin has a tact switch at the bottom for the user to make the system know the position by pushing it. This paper describes the hardware and software of the system.

1 Introduction

There has been a lot of research on supporting systems for the visually disabled in the world. However, most of them are for character recognition. There are only a few which touch upon recognition of 2D figures and 3D objects [1]-[8]. Our group has been developing 3D computer vision technologies based on a stereo vision method, and has come to apply these technologies to some fields. Also, the support systems should be developed from a 3D vision field.

We are developing a computer aided system for the visually disabled people to obtain 3D visual information by incorporating high-level image processing technology. It consists of three functions: input, processing, and output of visual information (Fig. 1). The input function consists of a stereo vision system from which a variety of 3D geometric data are obtained without touching objects. In the processing function, the data is transformed into tactile or auditory data. The output function consists of a tactile display and a voice synthesizer. The 3D data obtained is not just raw image data, it is analyzed structurally and only the required visual information is selected according to the user's demand.

There are two senses, tactile sense and auditory sense, for the visually disabled instead of sight. The former is superior to understanding the relative position in 2D and 3D space and the shape of objects. The latter is suitable for knowing the concepts. On these busies, we developed a 3D tactile display with the function of a digitizer. Since the tactile display offers familiarity with image data and our technologies of 3D tactile display, we selected it and added the function of the digitizer so users can not only be given tactile and auditory information but they can also select the required information from the system to make it easier to understand the 3D world. There is no other total system which includes everything input to output and has interactive interface. In this paper we introduce the hardware and software of this interactive tactile display system.

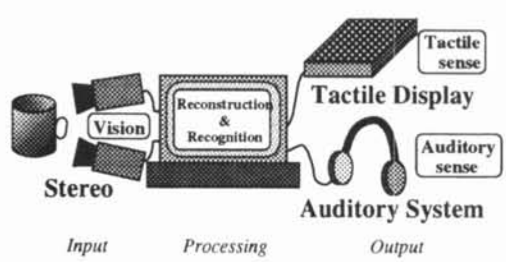

Figure 1: Conceptual diagram of the support system.

2 Tactile Display System

The overview of the system is shown in Fig. 2 and the composition of the system in Fig. 3.

2.1 Stereo Camera

The stereo camera system [9] has two cameras, which are controlled by a personal computer. Each camera is panned and tilted by pulse motors with an angular speed of 50 degrees/sec. The speed is fast enough to track moving objects.

*Address: 1-1-4 Umezono, Tsukuba, Ibaraki 305 Japan. E-mail: kawai@etl.go.jp

2.2 Tactile Display

The tactile display (Fig. 4) represents visual patterns by tactile pins arranged in a two-dimensional lattice. The pin height can be set to several levels to increase the touch information and to represent a three-dimensional surface shape. The major difference is the function of the digitizer. Also, the display has three pushbutton keys for selecting the display mode and hearing the voice message.

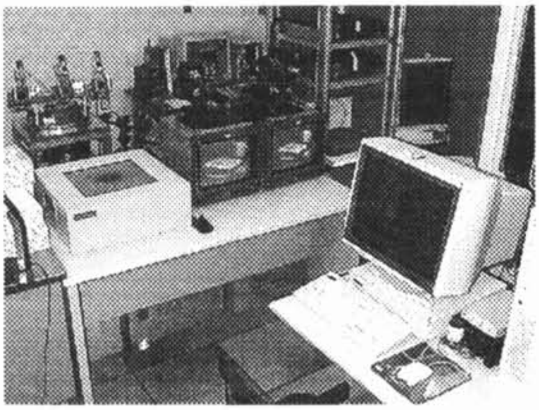

Figure 2: System.

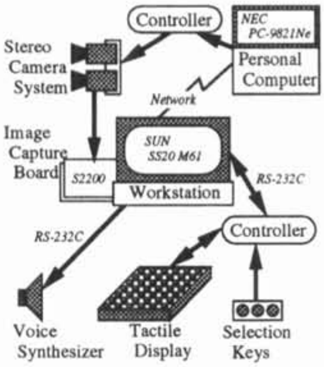

Figure 3: Composition of the interactive tactile display system.

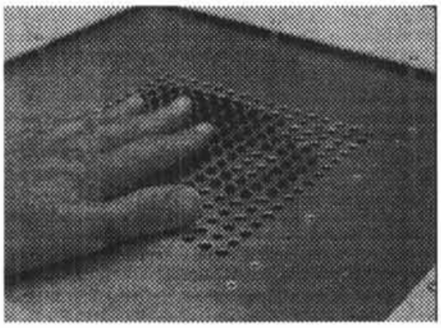

Figure 4: Tactile display.

The specifications of the tactile display are as follows. Pin arrangement: 16 x 16 pins; Pin area: 175 x 175 mm; Diameter of pin: 5 mm; Spacing between pins: 10 mm; Height of pins: 0-6 mm; Drive: stepping motor; Sensor: tact switch; Size: 550(W) x 530(L) x 195(H) mm; Weight: 28 kg.

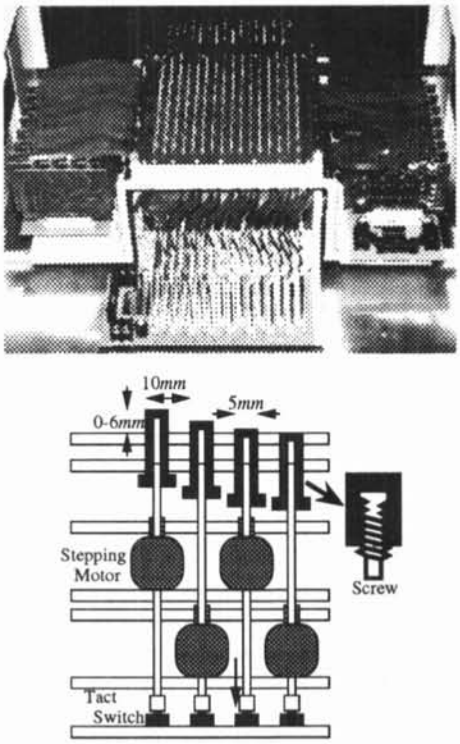

Figure 5: Tactile display mechanism.

The pins are arranged both horizontally and vertically. The mechanism of this display is shown in Fig. 5. The upper part of the figure is a picture of the interior. There are 32 driver boards for stepping motors, with each board controlling two pins. The diameter of the motor used is 10 mm. The motors are sufficiently powerful to maintain the same height under finger pressure. As depicted in the lower part of the figure, the structure is in two layers to allow the pins to be positioned more closely together. The screw in the pin relates rotation to the vertical movement. At the bottom of each pin is a tact switch, similar to that used in card type calculators that records which pins are pushed.

The selection key consists of three push buttons, and is similar in size to a computer mouse. It is used when the user wants to hear the voice message or change into another display mode (see Interface Section).

2.3 Voice Synthesizer

The voice synthesizer, which is another output device, is used to add more information to the tactile display. We used a "VC-1" voice synthesizer made by the RICHO Corporation of Japan. It is used for reporting the results of recognition and the concepts.

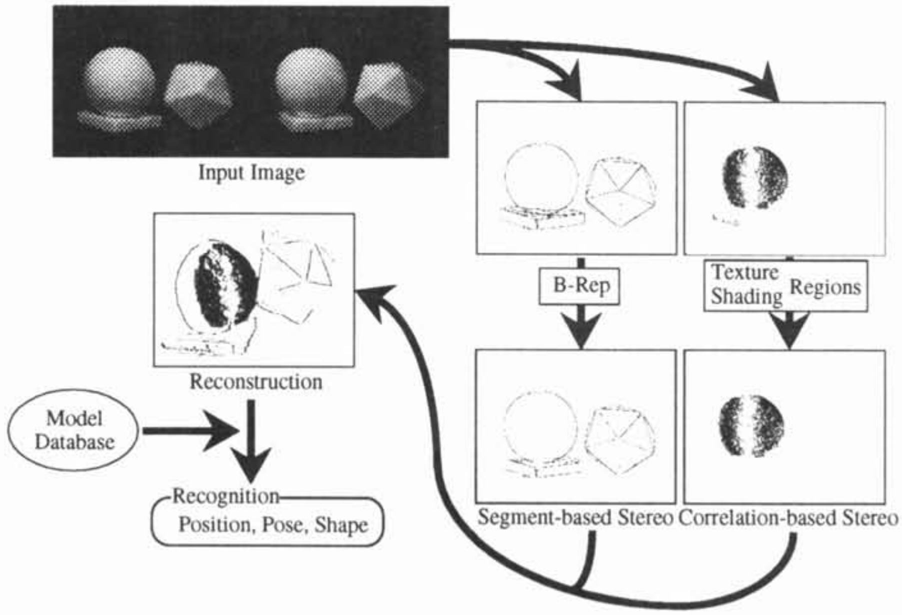

Figure 6: Stereo.

3 Stereo Vision

We use a stereo vision system to obtain 3D data of observed objects [10]–[13]. One reason why we don't use a conventional range-finder system is that it can detect only a limited distance. The main reason that we have developed a stereo vision system which can obtain 3D data as well as or better than range-finder systems is that the stereo vision system is more flexible and versatile. The stereo process is shown in Fig. 6. Two methods are used to obtain 3D data of objects: a segment-based method and a correlation-based method.

The segment-based method obtains 3D boundary data of objects [10],[11]. First, left and right image boundary edges are detected and converted into B-Rep (boundary representation). This B-Rep is segmented by some feature points. For each segment in the left image, the similarities with segments in the right image are calculated by the segment connectivity, figure, and brightness. Then one pair of segments which has maximum similarity is selected from the candidates of corresponding segment pairs of left and right images. By this process multiple correspondence is deleted. Three dimensional positions of segments are calculated by the difference between left and right segments.

The correlation-based method is used to obtain surface data, though it is not adapted for all points. First, in the pre-processing stage, suitable regions for this method are detected, such as texture or shading regions. In short, only characteristic regions are targets of this method. Some texture and curved surface data are reconstructed through these processes. Furthermore, each bit of 3D data measured by the two methods is integrated without contradiction.

After obtaining 3D data, the process of object recognition follows. We use a model-based method [12],[13]. Three dimensional data thus obtained is searched in an object model database and the correspondence between data and models is evaluated. By this process, we know what objects exist and what their position, pose and shape are.

4 Interface

Since the resolution of the tactile display is not enough to represent a lot of information at one time, we have developed a multi-level display mode (see Fig. 7) consists of position mode, boundary shape mode and surface shape mode. In the position mode, the user can recognize the relative position of each object by indicating the object position with a single pin. By pushing the voice guide button, he/she can hear an auditory explanation about the number of objects. The user can also hear an auditory explanation about the name of the object by depressing the corresponding pin. By pushing the information selection key to change the mode into the boundary shape mode or the surface shape mode, the user can learn the shape of the object. In these modes, the size and color of object are explained by voice. In the boundary shape mode, the user can feel the wire-frame shape, and in the surface shape mode, scan the convex or concave shape. He/She can select one of the three modes easily by pushing the information selection keys, and can recognize each object and environment by repeating these processes.

5 Discussion

Our new display system can represent more complex patterns than characters by making it possible

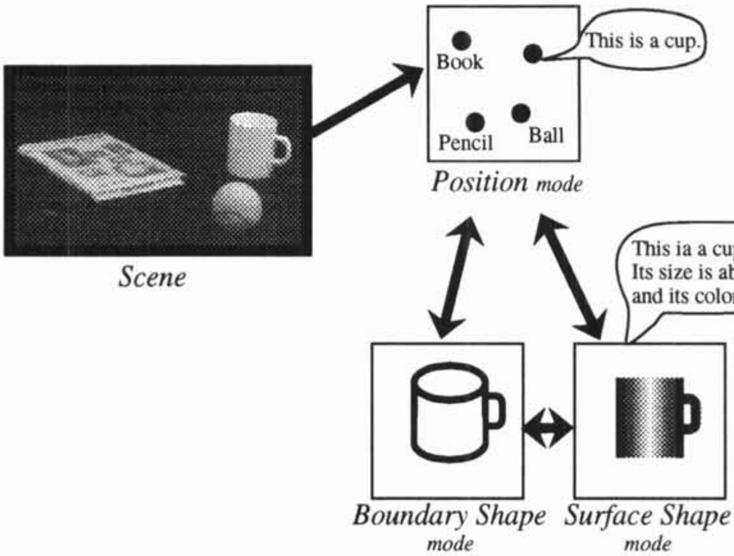

Figure 7: Interactive interface.

to adjust the pin height more than conventional tactile display systems. Although the resolution of the tactile display may not be enough to represent complex shapes, a higher resolution display would be much more expensive to produce and maintain due to current technical limitations. Thus, our concern is much more with how to use the tactile display effectively in the total supporting system than with developing the tactile display itself.

Each of the three display modes in the interactive interface serves a specific purpose. The position mode is suited to know the relative position of each object. The surface shape mode is useful for perceiving the approximate 3D shape of an object. For example, a ball is displayed as a smooth, convex hemisphere. However, it might be difficult to know the exact shape because of the low resolution. The interface has another shape mode called the boundary shape mode, which represents the shape by a wire frame. In both shape modes, it is possible to display objects from any viewpoint by using 3D models.

6 Conclusions

We have developed an interactive tactile display system from the viewpoint of 3D computer vision. The pin height of this tactile display can be set to several levels to increase the touch information and to display a three-dimensional surface shape. The system also has a digitizer function. Thus, it is possible for a user to communicate with this system, which we expect will help the visually disabled recognize 3D objects and environments by themselves.

We have to develop and refine the system through experiments with some visually disabled people. At present, the interface of this system is still at the prototype stage and we hope to further improve its function through more experimentation. The first experiment we are going to do is 3D object recognition by the user. We will determine which presentation is better for the visually disabled to display only one object. After that we will examine the suitability of the position mode for position recognition of objects. In the future we hope to develop a tactile display with higher resolution, further enhance the interface, and include a 3D auditory system.

References

[1] L. Kay, "Electronic Aids for Blind Persons : an Interdisciplinary Subject", IEE Proc., vol.131, part A, No.7. pp.559-576, 1984.

[2] K. A. Koczmarec, J. G. Webster, and W. J. Tompkins, "Electrotactile and Vibrotactile Displays for Sensory Substitution Systems", IEEE Trans. Biomed. Eng., vol.38, pp.1-16, 1991.

[3] R. Kowalik and I. Postawka, "The Concept of a Full Screen Tactile Display (FSTD) Driven by Electrochemical Reactions", Proc. of 4th Int. Conference on Computers for Handicapped Persons, pp.455-460, 1994.

[4] J. Fricke and H. Bähring, "A Graphic Input/Output Table for Blind Computer Users", Proc. of 3rd Int. Conference on Computers for Handicapped Persons, pp.172-179, 1992.

[5] J. Fricke and H. Bähring, "Displaying Laterally Moving Tactile Information", Proc. of 4th Int. Conference on Computers for Handicapped Persons, pp.461-468, 1994.

[6] Y. Kawai, N. Ohnishi, and N. Sugie, "A Support System for the Blind to Recognize a Diagram", Systems and Computers in Japan, vol.21, No.7, pp.75-85, 1990.

[7] M. Shinohara et al., "Development of a 3-D Tactile Display for the Blind: System Design", Proc. of 3rd Int. Conference on Computers for Handicapped Persons, pp.422-430, 1992.

[8] G. Jansson, "3-D Perception from Tactile Computer Displays", Proc. of 3rd Int. Conference on Computers for Handicapped Persons, pp.233-237, 1992.

[9] H. Takahashi and T. Suehiro, "Stereo Vision for the Mimic Robot", Computer Vision and Visual Communication, pp.75-78, 1994.

[10] T. Ueshiba et al., "An Efficient Method for the Correspondence Path Search in Segment-based Stereo Vision", Proc. of IEICE General Conf., Vol.2, pp.354, 1996. (in Japanese)

[11] Y. Kawai et al., "Searching for Stereo Correspondence based on Connectivity of Segments", Proc. of 51st IPSJ General Conf., Vol.2, pp.217-218, 1995. (in Japanese)

[12] Y. Sumi and F. Tomita, "Three Dimensional Object Recognition Using Stereo Vision", IPSJ SIG Notes, CV95-2, 1995. (in Japanese)

[13] Y. Sumi and F. Tomita, "Recognition of Complex Conical Objects Using Stereo Vision", IPSJ SIG Notes, CV98-7, 1996. (in Japanese)