The Thirteenth International Symposium on Artificial Life and Robotics 2008(AROB 13th '08),
B-Con Plaza, Beppu, Oita, Japan, January 31-February 2, 2008

A Human-machine Cooperative Interface of a Virtual 3-D Space Using Hand Gestures

*Keisuke MORI, **Joo Kooi TAN and **Seiji ISHIKAWA

*Kyushu Kyoritsu University,
1-8 Jiyugaoka, Yahatanishi-Ku,Kitakyushu-Shi, Fukuoka, 802-8585, Japan.
(Tel : 81-93-693-3251; Fax : 81-93-693-1553)
(kmori@kyukyo-u.ac.jp)
**Kyushu Institute of Technology

Abstract:
In this paper, we propose a new human-machine interface for manipulating the contents in an image on a screen provided by an LCD. Motions of a hand of a user are taken images by a simple USB camera. The motions are recognized and classified into respective corresponding instructions by PC. The result of the manipulation is displayed in the screen. Experiments were performed using GoogleEarth, 3-D CG and the proposed system successfully manipulated its rotation and magnification by some hand motions of the user. In addition to the method before, in this method, intentional movement and unconscious movement are separated, and a smoother operation feeling is obtained. It is expected that a more natural space operation is done by using this system.

Keywords: Human-machine interface, Image processing, Pattern recognition, Computer graphics, USB camera.

I. INTRODUCTION

The support service of work with a computer has come to be used in various scenes. Moreover, it has become more and more popular to handle a three-dimensional space in a computer. On the other hand, the coordinates operation on computer is generally done by using the mouse. As other devices, there are track ball that turns ball and a pen tablet, etc., by which the coordinates on the screen are absolutely specified. As for such devices, natural operations can be done according to usage. But certain skills are requested for the operations. So, they are not always a natural interface for many purposes.

To operate computer more freely, various methods are proposed. They include a method of detecting the movement of the radiant of a laser pointer[1] and a method of detecting gesture of an arm[2]. The use of a laser pointer is basically limited to the movement of the point. On the other hand, gesture of an arm is limited to the registered movement for instructions. Those who operate it cannot naturally operate as they must get accustomed to the gesture.

In this paper, we propose a new man-machine interface [3], [4] for manipulating the contents in an image on a screen provided by an LCD. Motions of a hand of a user are taken images by a simple USB camera. The motions are recognized and classified into respective corresponding instructions by PC. In this way, a user can handle an object in the display dynamically by hand.

II. METHOD

Figure 1 is an environment and equipment that detects hand gestures. A user sits in front of the display and operates the system. The movement of the user's hand is detected with the camera in front of the display. The camera set up is simple like the one used in web chat. The image range of the camera is adjusted to the area which includes a face and a hand. Even if a screen is not fixed, the hand can be detected stably. The parameter of the camera is dynamically adjusted even if the camera and the background move, and it is assumed that the system always detects the hand area.

Figure 1: Environment and equipment that detects hand gestures.

Two operation methods are examined from the detected hand area. The first is a method that decides a standard position of the hand like a joystick of computer (See Figure 2), and operates by detecting its displacement from the standard position. Such an operation as movements two dimensional looks like a usual joystick.

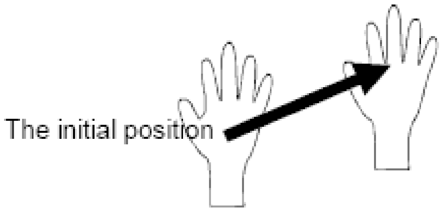

Figure 2: A method that decides a standard position of the hand like a joystick.

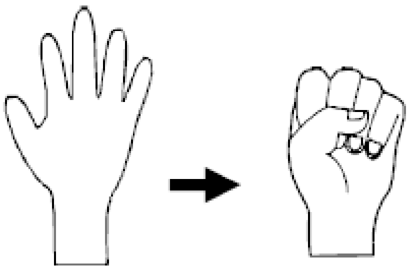

Figure 3: Mouse click is assumed to be the operation that opens and shuts the palm.

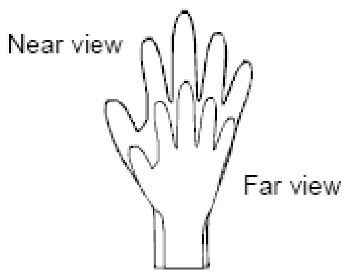

Figure 4: The movement toward the depth direction.

Mouse click is assumed to be the operation that opens and shuts the palm (See Figure 3). This shape change can be detected by calculating roundness (or complexity) of the hand. The roundness is defined by the ratio of peripheral length of the hand to its area. The movement toward the depth direction can be extracted by the change of the area of the hand. Obviously the area is larger in a nearer view, whereas it is smaller in a farther view (See Figure 4).

The allocation of the movement toward the depth orientation is normally different according to application. Ordinary 3-D software often allocates the movement to the rotation of a mouse wheel. The proposed system also allocates the movement toward the depth orientation in a screen to the rotation of the mouse wheel. The detected hand area is approximated in the form of an ellipse. Since the ellipse is enlarged or reduced its magnitude according to the movement of a hand toward the depth orientation, we extract the movement by evaluating the magnitude change.

Figure 5(a) is an example of a camera image. The area of the hand is extracted from this image. To do the installation and the operation easily, the numerical value of a prior parameter is detected while taking an image. The background image is detected real-time and the area with movement is obtained by differencing the background image from the original image. The movement image is separated into small areas of 8×8 pixels. The vector representing the small area of a present input image is denoted by F_n. Then the background image denoted by BF_n is given by the following formula;.

$$BF_n = (1-w)BF_{n-1} + wF_n \qquad (1)$$

Here w is a weight.

The area without movement is calculated by this expression. The weight is assumed to be 0 for the area with movement. The hand area is detected using this background together with the difference picture of a consecutive image.

The area of the hand is detected by using color information and a background subtraction algorithm. For expressing color information, the HSL (Hue, Saturation, Lightness) representation is employed. The parameter used to detect the hand is dynamically detected from the input image. The palm can be detected under the following conditions.

- The hand area occupies a large area in the image.
- The fingers point upwards.

Figure 5(b) shows the hand area detected from the movement area and the color information. The areas other than the hand area are also detected as shown in the figure. The palm can be detected under the following conditions.

Finally we obtain the hand area as shown in Figure 5(c).

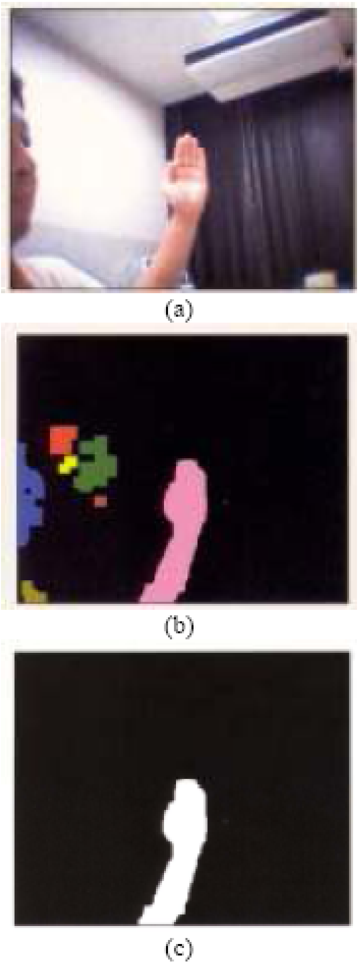

Figure 5: Area of a hand detected from movement area and color information: (a) Original image, (b) detected image, (c) hand area.

The palm is obtained from the detected hand area, and it is approximated as an ellipse. Its center of gravity is then detected by moment calculation. The position of the center of gravity is employed as an operation parameter.

Optical flow is also employed between successive image frames for smooth movement and for excluding unnecessary motion such as tremble.

Figure 6(a) and (b) show original successive image. Figure 6(c) shows movement vector detected by optical flow processing. Here, the hand and face movement have been detected.

Because the hand area has already been detected, only the movement of the hand can be separated from movement vector. The parameter used for detection is calculated from the image.

To achieve smooth movement, the value with greatly different value of the detected optical flow is excluded. Moreover, a median value of the detected movement vector of consecutive N frames is used to maintain the continuousness of operation.

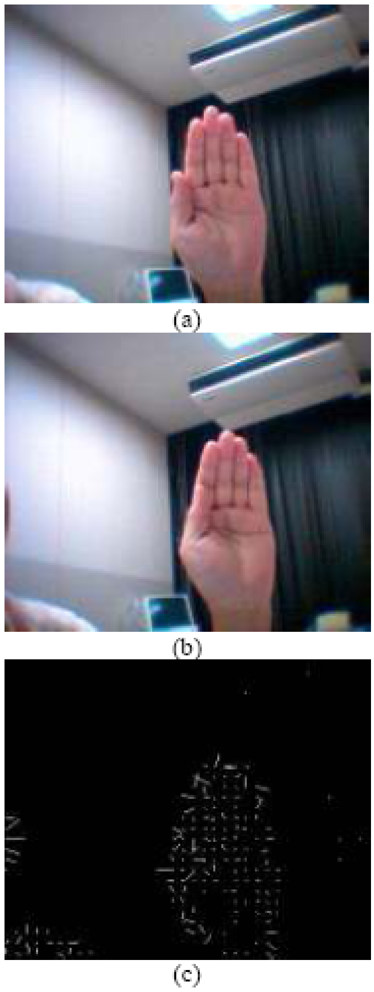

Figure 6: Movement area detected by optical flow processing. (a) Present image. (b) Past image, (c) optical flow image.

III. EXPERIMENT

In order to show performance of the proposed technique, we performed an experiment in which Google Earth in a PC display was operated remotely by a user's left hand. The Google Earth retrieves various kinds of information including maps, satellite and aerophotographs, etc., and allocates them on a virtual ball like a world globe. This virtual globe is rotated in a display using a mouse like we rotate a world globe by our hand. This Google Earth is going to be rotated and scaled by a hand operation of a user.

The allocation of the movement of the detected hand to the application software goes as follows. The upper, lower, right and left movements are allocated to the operation that turns the displayed globe in the same direction. The operations that move the hand backward and forward are allocated as those which bring the globe closer or nearer, respectively. In this way, the selected hand operations well imitate human actual movements.

The second experiment was done on the operations in three dimensional computer graphics using 3-D software called Metasequoia (See Figure 7). The objective of this experiment is to realize remote 3-D operation of an object in a display by a user's hand.

In the movement of the detected hand, upper, lower, right and left movements correspond to usual mouse operation, and the movement of the back and forth are allocated to the movement of mouse's wheel. Thus, by using the proposed system, an object in a display is put right or left, moved upper or lower, and placed closer or distant by a user's hand operation done in front of the display, as if we grasp the object directly and transforms it to arbitrary position. Note that object rotation is not installed at the moment.

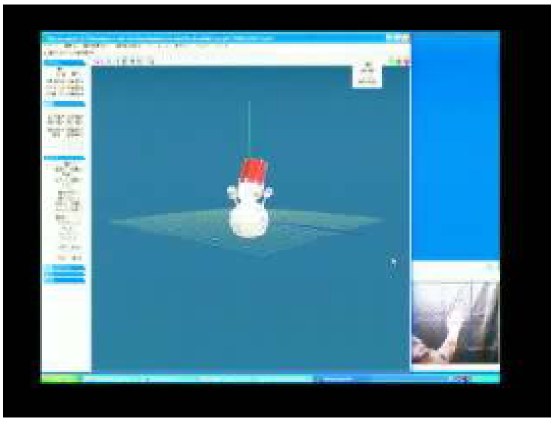

Figure 7 : The experiment using the software Metasequoia in a 3-D graphics environment.

One is able to use a cursor, a mouse or a track-ball, for example, in order to specify some operation to the object interested in the 3-D space in a display. It is then necessary to interpret actual operation to the operation of these devices. This interpretation is not necessary in the proposed system as similar hand motions are employed irrespective of in the real world or in the virtual world. This therefore realizes a human-friendly interface between computer and a human. In a practical sense, however, we can choose an appropriate device among a mouse, a track ball, the proposed system, etc., in order to realize the most convenient way of accessing an object in a 3-D virtual world computer provides.

Since the proposed system defines hand motions in place of a usual pointing device, it can be employed in any software without altering it.

IV. CONCLUSION

Three-dimension man-machine interface has been proposed employing a simple USB camera and a PC. The result of the manipulation was displayed in the screen. Experiments were performed using Google Earth and the proposed system successfully manipulated its rotation and magnification by some hand gestures of the user. It is, however, still necessary to adjust the system in the point that the speed and the amount of the change of movement should fit a user's own sense.

As a next step of the study, we have been groping for adopting more natural hand operation. The hand motions employed in this particular paper only realizes almost parallel displacement of the hand. On the hand, we often do a repeated hand motion when we call a person or turn the world globe right or left, for example. These movements will be taken into account in the next version of the system. It is also requested to realize a man-machine interface that allows the hand motions of not only adults but also kids or elderly people.

REFERENCES

[1] http://www.lleedd.com/afterglow/top.html
[2] N. Shimada, K. Kimura, and Y. Shirai: "Real-time 3-D hand posture estimation based on 2-D appearance retrieval using monocular camera", Proc. Int. Workshop on RATFG-RTS, pp. 23--30, 2001.
[3]K. Mori, et al.: "A study on the operation of a virtual 3-D space using hand motions", Proc. of the 25th SICE Kyushu Branch Annual Conference, pp.123-126 (2006). (in Japanese)
[4] K. Mori, et al.: "A study on the operation of a virtual 3-D space using hand motions", Proc. of the 1st ISICE 2007. pp.285-290 (2007).
[5]http://www.iplab.cs.tsukuba.ac.jp/~shin/papers/satou-shijibou-jssst00.pdf
[6]http://earth.google.com/
[7]http://www.metaseq.net/