2015 International Conference on Computational Science and Computational Intelligence

Pointing Gesture Recognition using Robot Head Control

Akira Nakamura, Masashi Ukai, Xiuming Wu, and Hideo Furuhashi
Dept. Electrical and Electronics Engineering
Aichi Institute of Technology
1247, Yachigusa, Yakusa-cho, Toyota, JAPAN
v14710vv@aitech.ac.jp

***Abstract*— We have developed a robot-head control system that recognizes a person's pointing gestures by a motion sensor. The robot head comprises two servo motors and a motion sensor and moves similar to a human head. The orbit of the person's pointing gesture is calculated from the hand and elbow positions measured by the motion sensor. The robot head moves such that the motion sensor orients along the pointing orbit. If an object measured by the motion sensor is within a certain space on the orbit, it is recognized as the target object of the pointing gesture. Object targeting by this system was successfully demonstrated.**

Keywords—pointing; motion sensor; gesture; robot; range sensor;

I. INTRODUCTION

Demand for communication robots has increased in recent years, and communication robots have become commercially available. In human–robot communications, both the conversation and the information acquired from camera images play important roles. As people use gestures in their conversations, our study focuses on pointing gestures.

Pointing gestures during conversations are mainly used to clarify an object. Therefore, such gestures are thought to be important when establishing communications with a robot. Humans find the target object by directing their head along the pointing orbit. A robot can realize the same motion by moving its camera along the pointing orbit measured by a motion sensor and finding the target by the range (motion) sensor.

In this study, we construct a head control system that recognizes a person's pointing gesture using a motion sensor. It then identifies the object being mentioned in the conversation. The accuracy of the system is studied.

II. CONTROL SYSTEM OF THE ROBOT HEAD

Figure 1 shows the mechanism of the robot head. The head comprises two servo motors and a motion sensor (Microsoft Kinect ver. 2.). The yaw angle θ ranges from $-135°$ to $135°$, and the pitch angle φ can be rotated from $-13°$ to $90°$.

Figure 2 is a flowchart of the system. The application was constructed by robot middleware (OpenRTM-aist1.1). First, a person starts a conversation and the motors move the motion sensor toward the person. The angle of the sound is recognized by a microphone array in the motion sensor. Second, the motion sensor detects two joints (in this case, the hand and the elbow) of the person pointing toward a target object. The pointing orbit is calculated from the positions of the two joints. Third, two servo motors move the motion sensor, directing it along the orbit. The motion sensor measures the range within a certain viewing area and compares it with the calculated distance from the sensor to the orbit. If the distances are equal, the motor stops and the target search terminates.

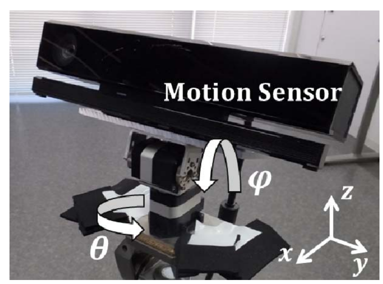

Fig. 1. Motion sensor and head mechanism

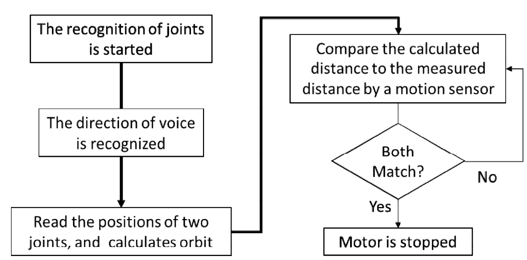

Fig. 2. The system flowchart

III. CALCULATION OF THE POINTING GESTURE

The local coordinates of the motion sensor (x_k, y_k, z_k) are converted to world coordinates (x, y, z) by the following equation:

978-1-4673-9795-7/15 $31.00 © 2015 IEEE
DOI 10.1109/CSCI.2015.38

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & r_{1z} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\varphi & 0 & \sin\varphi & r_{2x} \\ 0 & 1 & 0 & 0 \\ -\sin\varphi & 0 & \cos\varphi & r_{2z} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \\ 1 \end{pmatrix} . \quad (1)$$

Here, x , y , and z denote the depth, width, and height, as shown in Fig. 1, respectively.

The world coordinates of the pointing orbit are given by

$$\begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = \begin{pmatrix} at + x_0 \\ bt + y_0 \\ ct + z_0 \\ 1 \end{pmatrix} . \quad (2)$$

Here, (a, b, c) is the pointing vector calculated from the positions of the two joints, and (x_0, y_0, z_0) represents the hand position.

Therefore, the center of the motion sensor can be directed along the pointing orbit by moving the two motors through the following angles.

$$\theta = \tan^{-1} \frac{y}{x}$$
$$\begin{cases} 0 \le \theta \le 180 & : y \ge 0, \\ -180 \le \theta \le 0 : else \end{cases} \quad (3)$$

$$\varphi = -\tan^{-1} \frac{z - r_{2z} - r_{1z}}{\sqrt{x^2 + y^2} - r_{2x}}$$
$$\begin{cases} 0 \le \varphi \le 90 : z - r_{2x} - r_{1z} \le 0, \\ -90 \le \varphi \le 0 : else \end{cases} \quad (4)$$

The distance from the motion sensor to the pointing orbit is calculated by

$$x_k = \sqrt{(z - r_{2z} - r_{1z})^2 + \left(\sqrt{x^2 + y^2} - r_{2x}\right)^2} . \quad (5)$$

When the calculated depth equals the measured depth of the motion sensor, the search for the target object ends.

IV. RESULTS

Two objects are located as shown in Fig. 3. A person was positioned (x, y) = (3.5 m, 1 m) in front of the robot head (Kinect) and was required to point to one of the two objects. If the system identified the target object, the case was considered successful ("ok"). If the system identified the wrong object, the search result was "false". When no object was found, the result was "unrecognized".

The viewing area was varied as 1 × 1, 30 × 30, and 50 × 50 pixels, and the depth tolerance was varied as 0.3, 0.4, and 0.5 m. The recognition rates are presented in Table 1. Good results were obtained at viewing angles of 30 × 30 or 50 × 50 pixels and a depth tolerance of 0.5 m.

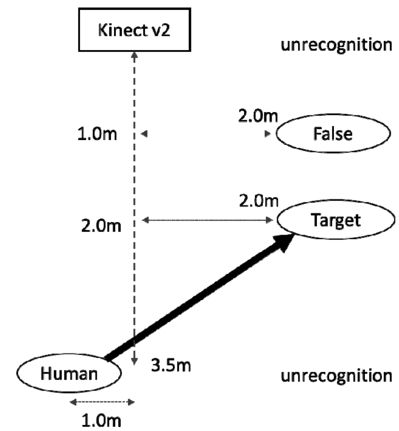

Fig. 3. Experimental setup.

Table 1. Recognition rate

Tolerance / Space	0.3 [m]			0.4 [m]			0.5 [m]		
	ok	false	un	ok	false	un	ok	false	un
1×1	0	0	100	40	6.7	53.3	26.7	0	73.3
30×30	0	0	100	80	6.7	13.3	100	0	0
50×50	26.7	0	73.3	73.3	6.7	20	100	0	0

V. CONCLUSIONS

We have developed a robot-head control system that recognizes the pointing gestures of a person by a motion sensor. At appropriate viewing angles and depth tolerances, the system successfully identifies the human's pointing target. Further studies with different conditions are required to clarify the effectiveness of the system.

REFERENCES

[1] K.Yasui, et al, "Estimation of the finger pointing position for a communication robot", Tokai Section, IEICE. , 2014, I2-2.

[2] A.Nakmura, et al, "A Pointing gesture recognition with a motion sensor and control of a robot head", Tokai Section, IEICE. , 2015, P2-9.

[3] Xu Xu, et al, "The validity of the first and second generation Microsoft Kinect for identifying joint center locations during static postures", Applied Ergonomics. 49(2015), 47-54.

[4] N. Jojic, et al, "Detection and Estimation of Pointing Gestures in Dance Disparity Maps," Proc. of IEEE FG2000, pp. 1000-1007, 2000.