Pointing Angle Estimation for Human-Robot Interface

Nur Safwati Mohd Nor, Yoshio Maeda, and Makoto Mizukawa

Abstract—**In this paper, we propose an approach to track and estimate user pointing direction in 3D Space. In the area of human-robot interaction, user communicates with service robot in their daily life activities to give commands and execute the given task accordingly. Therefore, the ability of user to gives command to service robot naturally can provide an interactive user interface system for real 3D space environment. For this purpose, we aim to perform pointing gesture tracking and after that estimate the user's pointing direction. Our method of pointing direction estimation is based on 3D orientation of hand and shoulder center of user. We make comparison with our previous method to find the best hypothesis. Experimental results show the angular error for the estimation of pointing direction is successfully improved from our previous method. As a result, our natural user interface system can manipulate 3D objects in living room environment thus providing intuitive robotic service for human robot interaction.**

Index Terms—**Human-robot interaction, pointing gesture, user tracking.**

I. INTRODUCTION

Tracking of human gesture has been widely researched either in human-computer interaction or human-robot interaction. Besides voice, gesture is an example of platform to interfacing between human and robot intuitively and naturally. Pointing is one part of gesture which provides information and intention from human to service robot. Pointing-based user interface can easily indicate object and location for the robot delivering robot task in home environment.

However, in the real environment, some ambiguity can occur and affect the performance of robotics service. The ambiguity in pointing gesture comes from the angular error in estimation of user's pointing direction particularly when certain objects is closely located. Due to the ambiguous pointing command, correct object and location cannot be identified thus service robot surely unable to deliver task successfully. To overcome this problem, we conduct several experiments on tracking of pointing arm and make comparison study to find the best hypothesis.

In this paper, we present a robust tracking method to be used in pointing direction estimation. We measure the pointing angle from the position of hand with respect to the shoulder center. To obtain the 3D coordinates of pointing arm

and shoulder center, we use visual-based tracking that coming from single 3D Kinect sensor. From the skeleton image trajectory data, we estimate the pointing angle by using simple trigonometry function. Although the trigonometry calculation is very simple, we show through experiment that the accuracy of our user interface system is significantly improved to previous research. We evaluate our result based on angular error and mean squared error.

A. Related Work

Employing service robot at home to help human on the daily life activities is an example of the application for human-robot interaction. To achieve this scenario, service robot is placed in an intelligent space where the real living environment is designed with many sensors, computers and frameworks such that it becomes ubiquitous [1]. Human-robot interface is designed to pass useful information from human to system and vice versa. Niitsuma [2] presents a system to intuitively get command from human to system by using a vibrotactile interface at smart wheelchair in intelligent space. Whereas, a PDA-based interface also is introduced to get the input signal from user to the system for bring service application [3].

A user interface based on gestural input from human has been researched in [4]-[6]. By using hand gesture such as sign language, human or user can convey information naturally to the system. Ishii [7] introduces an approach to give instructions to object-delivery robot based on drag and drop operation from user's specific finger.

Furthermore, in direct manipulation tasks, pointing gesture can be interpreted very well to accurately identifying an arbitrary location in space [8]. Sato [9] uses pointing gesture to give command to service robot in virtual space. In her research, the pointing direction is estimated by using marker at user's cap and hand glove. On the other hand, research in [10] presents an HMM-based pointing detection to recognize pointing phases of dynamic hand pointing.

The problem of the system to indicate which target object is pointed by user is crucial in such a way that robot do not understand the ambiguous command from the user. Nickel [11] presents an approach to solve this ambiguous pointing angle by combining 3D head and hand tracking with head orientation. Meanwhile, Ngo [12] proposes a system combining pointing gesture and commonsense knowledge to correctly identify the target object.

B. Outline

This paper is organized as follows: In chapter 2, we explain the general system overview of our pointing-based user interface. Also, in this chapter, we introduce our methodology on tracking of user pointing gesture as well as estimate the pointing angle. In chapter 3, we show the experimental setup and result based on angular error in pointing estimation. Next,

in chapter 4, we discuss the accuracy of pointing gesture proposed in this paper by doing comparison with our previous tracking method. Finally, we conclude this paper by explaining the importance of our findings and present some future works.
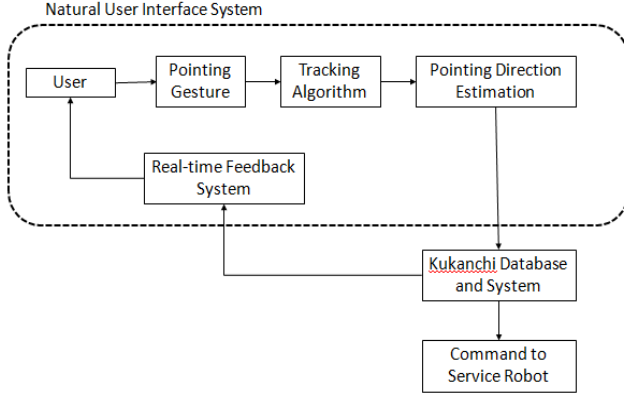

Fig. 1. System design for human-robot interface.

II. POINTING-BASED USER INTERFACE

A. System Overview

We design our pointing-based user interface such that user can easily using the system hence manipulate objects in 3D space. The user must not spend much time on using the system to give task to service robot. In addition, we design our user interface system to be interactive to allow continuous interaction between user and service robot. For that reason, we include a real-time feedback system to our pointing-based user interface. This feedback system is crucial in integrating information and confirmation on the robot tasks thus provides a closed loop interaction process.

Fig.1 shows the overview of our natural user interface system. The pointing-based user interface system communicates and gives command to service robot through Kukanchi (Interactive Human Space Design and Intelligence) database and system.

B. Tracking of Pointing Arm

In this study, we define pointing gesture by a movement of arm pointing to a target object. The line of arm including three joints which are shoulder, elbow and hand. In this paper, to track the pointing arm, user has to be at a predefined position in front of the 3D sensor Kinect XBox. The center of the Kinect must be aligned with the user's center of shoulder.

By using Kinect XBox API and OpenNI, the system will automatically detect user's 3D coordinates of the skeleton joints. Three joints in the skeleton image which describes our pointing gesture are shoulder center, shoulder and hand. The shoulder of a user is significant to be tracked because we differentiate the pointing posture, either standing or sitting based on the height of the shoulder. In order for the system to track the pointing gesture, user needs to hold arm for five seconds during pointing.

C. Estimation of Pointing Direction

To get pointing direction, we estimate the pointing angle from user's arm to center of shoulder during hold state. We

measure the pointing angle in horizontal axis. From the coordination of hand and shoulder center, the estimated angle is identified accordingly at every pointing direction. Next, we calculate the angular error at each pointing direction for the evaluation purpose.

III. RESULT

A. Experimental Setup

We have conducting an experiment to measure the angle of pointing arm by user in our intelligent space called Kukanchi (Interactive Human Space Design and Intelligence). In this experiment, we measure the pointing angle based on the predefined location of user and object. After that, we evaluate the pointing estimation by calculating the angular error. For this purpose, user needs to point his arm to six locations of an object, five times at each location. In addition, we conduct this experiment in two posture mode, standing and sitting. Fig. 2 shows the environment setup for our experiment.
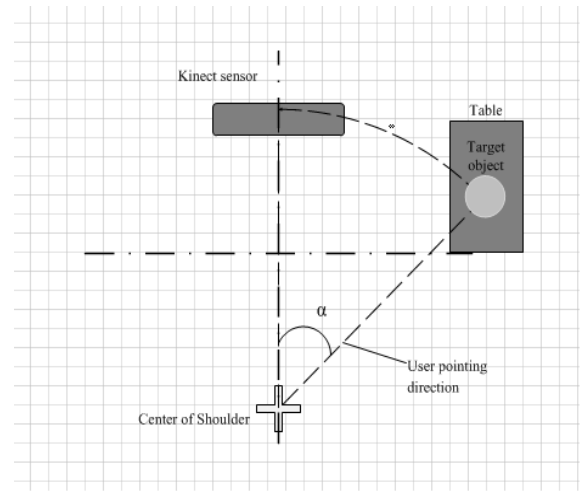

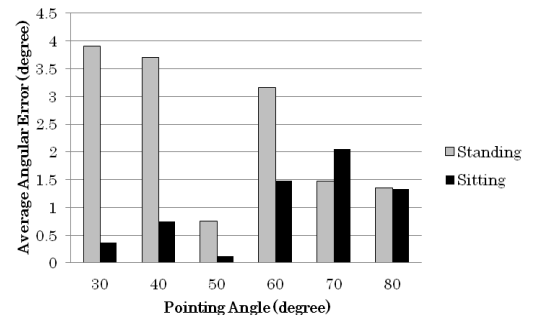Fig. 2. Setup to estimate pointing angle.

Fig. 3. The angular error for pointing direction estimation

As for the setup, we are using Kinect XBox to track the pointing gesture, thus providing the 3D coordinates of pointing arm. The pointing angle is defined by the angle between user's arm and shoulder center in the skeleton image. The estimated pointing angle, α is measured by using (1).

$$\alpha = \tan^{-1}\left(\frac{Hand.x - ShoulderCenter.x}{Hand.y - ShoulderCenter.y}\right) \quad (1)$$

B. Angular Error

For the analysis, we measure the pointing angular error both during standing and sitting posture. Afterward, we

calculate the average angular error for each predefined angle. Fig. 3 shows the angular error measured for all six locations of object position. As of calculated average angular errors, it is clearly seen that the pointing direction estimation has the accuracy of below four degrees of error.

IV. COMPARISON AND HYPOTHESIS

In order to evaluate and choose the best pointing gesture tracking algorithm, we make comparison and hypothesis on the experimental data. To achieve that, we calculate the Mean Squared Error (MSE) for both pointing posture and compare them with our previous work in [1]. The different method of tracking user pointing direction is evaluated by finding the least MSE. Fig. 4 shows the angular error for our tracking method presented in [1]. In [1], we estimate the pointing direction by finding the angle between pointing arm and center of the head. Table I shows the MSE for the estimation of pointing direction methods.

From the calculated MSE, the method presented in this paper has the least MSE for both standing and sitting mode. Therefore, the tracking algorithm and pointing direction estimation based on the center of user's shoulder is more accurate than the estimation made based on the center of user's head. Furthermore, the difference in MSE for standing and sitting posture is smaller for center of shoulder's method compared to center of head method. This small difference indicates the validity of the method and shows our current method gives better performance than our previous method. Hence, we plan to use shoulder's center as our reference point to detect the pointing angle from user's arm for our natural user interface system.
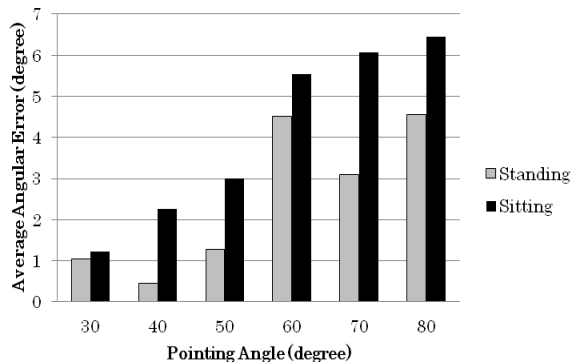

Fig. 4. Pointing angular error based on hand-head center tracking.

V. CONCLUSION AND FUTURE WORK

From the experimental result, we would like to concentrate on the tracking method of pointing gesture and also estimation of pointing direction. We make comparison of the respective methods to find the best hypothesis. This study is important because our natural user interface needs to employ the best pointing direction estimation so that the pointing angle can be specified to the correct object in 3D space. Accurate pointing angle estimation is further needed for service robot to bring the object which the user wants especially when multiple objects near to each other and small in size.

In future, we will consider multiple Kinect sensors to address occlusion problem. The limitation of single Kinect sensor such that, the camera projection is not wide enough to detect pointing gesture at certain angles in different pointing posture. Moreover, practically in intelligent space like Kukanchi, user should be moving and not static at one position only. Therefore, multiple 3D sensors are required to detect pointing gesture from many angles and postures. In addition, we plan to create an interactive feedback system by employing augmented reality technology.

TABLE I: COMPARISON OF MEAN SQUARED ERROR

Method	Standing	Sitting
Hand-Shoulder center	7.27	1.48
Hand- Head center	8.95	20.76

REFERENCES

[1] N. S. Mohd. Nor, N. L. Trung, Y. Maeda, and M. Mizukawa, "Tracking and Detection of Pointing Gesture in 3D Space," accepted for publication International Conference on Ubiquitous Robots and Ambient Intelligence, to be published.

[2] M. Niitsuma, T. Ochi, M. Yamaguchi, and H. Hashimoto, "An Approach of Human-Smart Electric Wheelchair Interaction in Intelligent Space," in *Proc. IEEE Workshop on Robotic Intelligence in Informationally Structured Space*, pp. 119-124, April 2011

[3] T. L. Ngo, Y. Tsukumori, and M. Mizukawa, "PDA-based Interactive HMI Device for Service Robotic System," in *Proc. International Conference on Ubiquitous Robots and Ambient Intelligence*, November 24-27, 2010.

[4] S. Eickeler, A. Kosmala, and G. Rigoll, "Hidden Markov Models Based Continuous Online Gesture Recognition," in *Proc. International Conference on Pattern Recognition*, vol. 2, pp. 1206-1208, Aug 1998.

[5] R. A. Bolt, "Put-That-There: Voice and Gesture at the Graphics Interface," *ACM SIGGRAPH Computer Graphics*, vol. 14, pp. 262-270, July 1980.

[6] H. K. Lee and J. H. Kim, "An HMM-Based Threshold Model Approach for Gesture Recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 961-973, Oct 1999.

[7] K. Ishii, Y. Takeoka, M. Inami, and T. Igarashi, "Drag-and-Drop Interface for Registration-Free Object Delivery," in *Proc. IEEE International Symposium on Robots and Human Interactive Communications*, pp. 228-233, Sept 2010.

[8] R. Kjeldsen and J. Hartman, "Design Issues for Vision-based Computer Interaction Systems," in *Proc. the Workshop on Perceptual User Interface*, pp. 1-8, 2001.

[9] E. Sato, T. Yamaguchi, and M. Harashima, "Natural Interface Using Pointing Behavior for Human-Robot Gestural Interaction," in *Proc. IEEE Industrial Electronics*, vol.54, no.2, pp. 1105-1112, 2007.

[10] K. Nickel and R. Stiefelhagen, "Visual Recognition of Pointing Gesture for Human-Robot Interaction," *Image and Vision Computing*, vol. 25, pp. 1875-1884, 2007.

[11] K. Nickel and R. Stiefelhagen, "Pointing Gesture Recognition based on 3D Tracking of Face, Hands and Head Orientation," in *Proc. International Conference on Multimodal Interface*, pp. 140-146, Nov 2003.

[12] T. L. Ngo, Y. Maeda, H. Lee, and M. Mizukawa, "Ambiguous Command Understanding with Commonsense)," accepted for publication in IEEE International Symposium on Robot and Human Interactive Communication, to be published.

Nur Safwati Mohd Nor received the B.E in Mechatronics (Honours) in 2008. She was working as Assistant Lecturer at Universiti Teknologi Malaysiai. In 2011, she received the M.E in Mechanical Engineering from Universiti Teknologi Malaysia. She is a member of Board of Engineer Malaysia (BEM).

Yoshio Maeda received the B.E in Electric Engineering from Shibaura Institute of Technology in 2011. He is a member of the Robotics Society of Japan.

Makoto Mizukawa is a professor of Department of Electrical Engineering, College of Engineering, Shibaura Institute of Technology, Japan. In 1975, he joined Nippon Telegraph and Telephone Corp. In 1981, he became a visiting scholar in Columbia University. He worked as Senior Research Engineer and Group Leader in NTT Human Interface Lab in 1992.

Makoto Mizukawa is a member of IEEE and The Robotics Society of Japan. He also is a fellow of the Japan Society of Mechanical Engineers and the Society of Instruments and Control Engineer.