The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

T-less : a Novel Touchless Human-Machine Interface based on Infrared Proximity Sensing

Dongseok Ryu, Dugan Um, Philip Tanofsky, Do Hyong Koh, Young Sam Ryu, and Sungchul Kang

***Abstract*— In today's industry, intuitive gesture recognition, as manifested in numerous consumer electronics devices, becomes a main issue of HMI device research. Although finger-tip touch based user interface has paved a main stream in mobile electronics, we envision touch-less HMI as a promising technology in futuristic applications with higher potential in areas where sanity or outdoor operation become of importance. In this paper, we introduce a novel HMI device for non-contact gesture input for intuitive HMI experiences. The enabling technology of the proposed device is the IPA (infrared Proximity Array) sensor by which realtime 3 dimensional depth information can be captured and realized for machine control. For the usability study, two different operating modes are adopted for hand motion inputs: one is a finger tip control mode and the other is a palm control mode. Throughput of the proposed device has been studied and compared to a traditional mouse device for usability evaluation. During the human subject test, the proposed device is found to be useful for PC mouse pointer control. The experimental results are shared in the paper as well.**

I. Introduction

HUMAN computer interface is continually evolving to more intuitive and convenient ways. Hand gesture is regarded one of the best ways to interact with computer. Currently, touch-less computer interfaces are not an option in many industrial environments, such as, hospital operation rooms, clean rooms, petroleum plants, war environments, and etc., because most equipment and PCs cannot be operated by a wet hand, oily hand or dusty hand. An effective touch-less interface would alleviate this problem, allowing anyone under any of these conditions to input data into a PC workstation. A viable touch-less interface has been studied for many years, however, a well working device has not been yet commercialized due to the high cost and lack of operational reliability.

A typical approach to integrate a touch-less interface is to utilize vision sensors. Hand gestures can be analyzed from the camera image, and the predefined motion is detected and converted to specific command [1-3]. In an actual environment, the detection conversion struggles to distinguish the hand in an image from the background. To resolve the problem, most existing interfaces use markers or a specially designed glove [4-6]. However, the results are sensitive to different light conditions. As long as the approach uses a 2D image, the basic information is limited in 2D.

To utilize a 3D image, a stereo camera is adopted to develop a touch-less interface [7-9]. The calculation disparity from two 3D images requires a large computational burden. Therefore, the touch-less interface requires a high performance computer.

The easiest method to acquire a 3D distance map is to use a Time of Flight (TOF) camera [10]. Since the camera directly provides a distance map and functions free from computational burden, the TOF camera delivers an adequate solution for fabricating a touch-less user interface. The problem remains the high cost to commercialize the interface.

In this research, a new interface using IPA is proposed. IPA can directly measure the distance map, and yield robust results at various light conditions by using infrared light. The cost of the IPA design is much less than the TOF camera, so that the touch-less interface by IPA can be implemented at a reasonable price. This research shows how to implement touch-less user interface using IPA, and the performance is

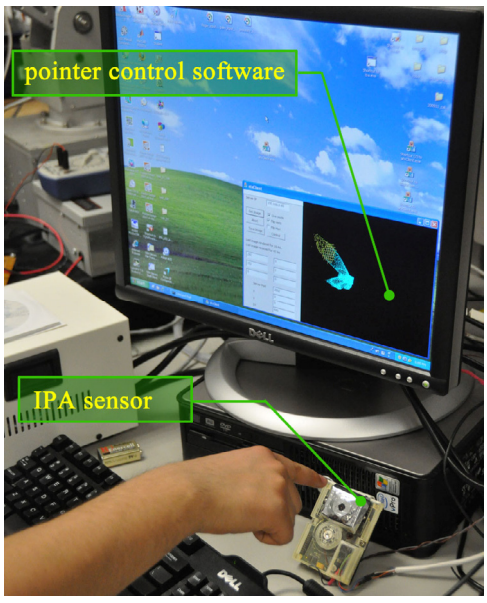


Fig. 1. T-less : The proposed touch less user interface by using IPA

Dongseok Ryu is with the Texas A&M University-Corpus Christi, Texas, USA (phone: 1-361-825-3446; fax: 1-361-825-5629; e-mail: dongseok.ryu@tamucc.edu).

Dugan Um is with the Texas A&M University-Corpus Christi, Texas, USA (e-mail: dugan.um@tamucc.edu).

Philip tanofsky is with the Texas A&M University- Corpus Christi, Texas, USA (ptanofsky@islander.tamucc.edu)

Do Hyong Koh is with the Texas State University-San Marcos, Texas, USA (dk1132@txstate.edu)

Young Sam Yoo is with the Texas State University-San Marcos, Texas, USA (yryu@txstate.edu)

Sunchul Kang is with the Korea Institute of Science and Technology, Seoul, Korea (kasch@ kist.re.kr).

also evaluated compared with other input devices.

The remainder of this paper is organized as follows. Section 2 introduces the working mechanism of the IPA device and its specifications. In section 3, the touch-less interface using IPA and its input method are described. Section 4 offers some experimental results for evaluation of the developed interface. Section 5 concludes the research results.

II. Infra red Proximity Array

The industrial infrared sensor has a simple working mechanism – to measure the intensity of reflection light. The sensor emits infrared light to sensing area. If an object exists in the sensing area, it reflects the light. The sensor detects the reflected light and measures the intensity.

The typical characteristics of an infrared sensor are applied to 3D image construction technology. IPA is composed of an infrared light source, an infrared filtering lens, and a charge coupled device (CCD) array. Since each pixel of CCD array is utilized as an infrared sensor, the IPA can be, therefore, considered as highly integrated infrared sensor array [11].

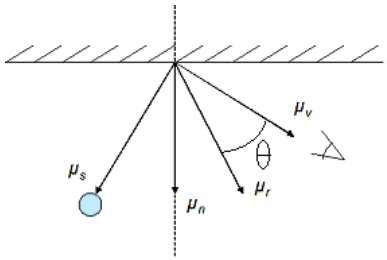


Fig. 2. Phong's illumination model

Based on the Phong's illumination model, as shown in Fig. 2, and the photometry theory [12], an infrared sensor is known to have reflected light intensity I given by:

$$I = C_o(\vec{\mu}_s \cdot \vec{\mu}_n) + C_1(\vec{\mu}_r \cdot \vec{\mu}_v)^n + C_2 \quad (1)$$

The variables C_o and C_1 are two coefficients (diffusivity and specularity) that express the reflectivity properties of the surface being sensed and μ_s, μ_n, μ_r, and μ_v are the light source, surface normal, reflected, and viewing vector, respectively.

Although, a pixel of the IPA sensor receives multiple reflected light beams with differing intensities from a single light source due to directivity, as shown in Fig. 3, we consider the light beam when θ is zero the most intensive reflection. Assuming $n = 1$ and $C_2 = 0$, equation (1) becomes:

$$I_p = C_o \cos\alpha_p + C_1 \quad (2)$$

The index p represents a pixel among sensor array, and α_p is the incidence angle of light with respect to the surface. Therefore the typical energy equation for a photo sensor becomes:

$$E_p = \frac{I_p A}{(2l)^2} = \frac{I_p A}{\left(2d/\cos\alpha_p\right)^2} \quad (3)$$

where d is the minimum distance between the sensor and a surface. E_p stands for the energy absorbed by a pixel p from a single light source, and corresponds to the actual sensor reading. A in the above equation is the area of an object, and l, the travel distance of the light.

Therefore, the distance from a pixel p to an obstacle is,

$$d_p = \sqrt{\frac{A}{2E_p}\sum_{m=1}^{8}\left[C_o\cos^3\alpha_p^m + C_1\cos^2\alpha_p^m\right]} \quad (4)$$

After calibrating the constant C_o and C_1 by a series of experiments for human skin, the intensity at each pixel can be converted to distance. From the arrayed distance data, a distance map over a sensing range can be obtained.

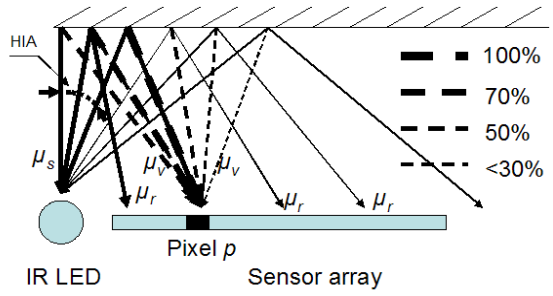


Fig.3. observation of reflecting light at pixel array (% : reflection rate)

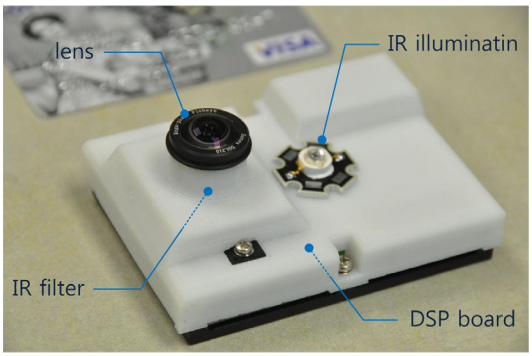


Fig. 4. Infrared Proximity Array

Figure 4 shows the recently developed compact version of the IPA sensor which size is similar to that of a credit card. A small DSP board equips a 640 by 480 CCD array, and a wide angle board mount lens is installed. With these elements, a compact size is achieved. The DSP has enough computational power for standalone operation. The DSP has enough computational power for standalone operation. High powered LED emits infrared light to the entire sensing area, resulting in a sensing range of about 450mm. The measured position

resolution is 1mm at 640 by 480 resolution images. Depth measuring, surface color, texture, roughness, and transparency of the target object affect the sensing results. Fortunately, human skin is well detected by IPA, and precise calibration enables a measuring depth with a 5mm error bound. Figure 5 shows an example distance map for a human face.

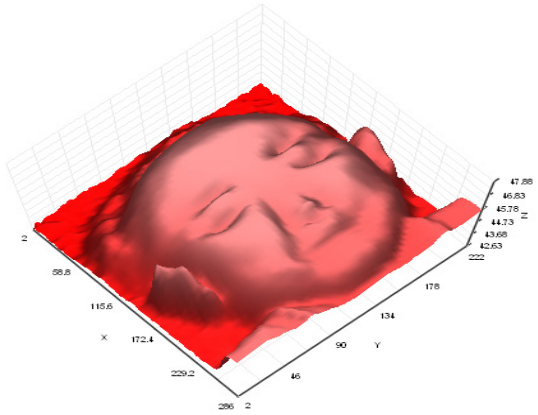
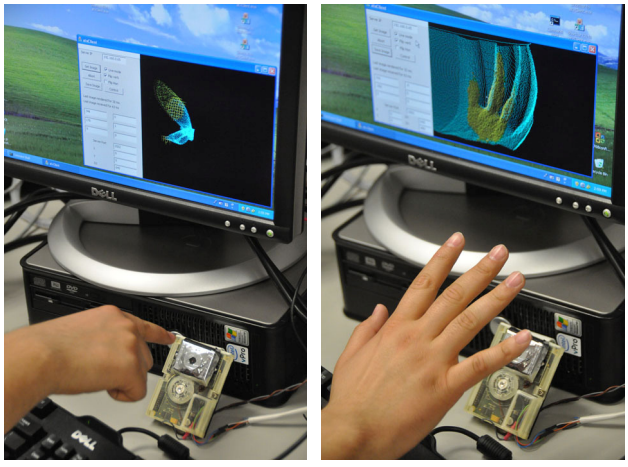


Fig. 5. Depth map for human face by IPA

III. Touch less User Interface

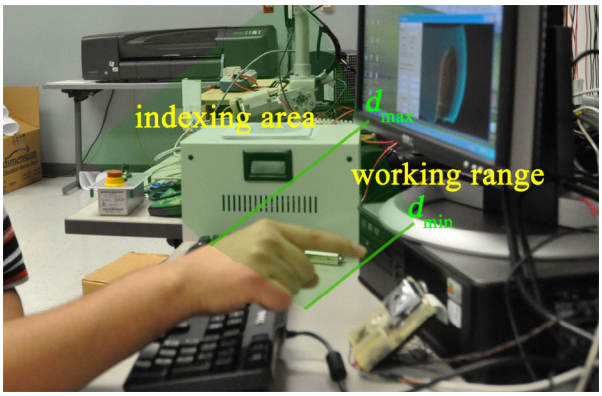
Figure 1 shows the developed user interface, T-less. In front of a PC, the 3D sensor is located instead of mouse, and the user controls the cursor by hand motion. While the interface recognizes the hand motion and generate command to control cursor, the distance information detected by IPA is fully utilized to distinguish the hand gestures. The interface provides two different operation modes to control the cursor. First, the finger input mode utilizes the position information of an active index finger, as shown in Fig 6 (a). In the finger input mode, the user rest his hand in front of camera and just point out the direction where s/he want to move the cursor, and then the cursor is exactly controlled to move along that direction. The other mode is the palm input mode shown in Fig. 6 (b). The palm input mode measures the angle of the palm while the user rotates his or her wrist. For the click operation, same motion is applied for both operation modes as shown in Fig. 8. The user unclasps his or her hand and dash toward the sensor.



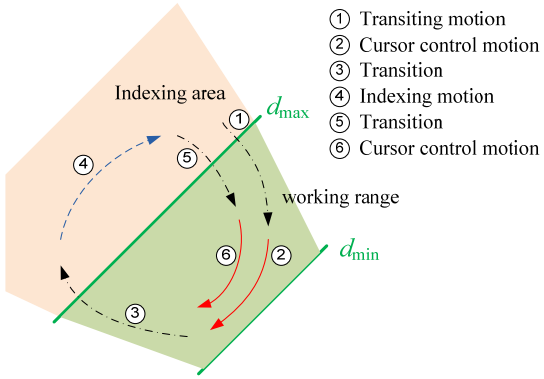
(a) Finger input mode (b) palm input mode

Fig.6. Finger input mode and palm input mode

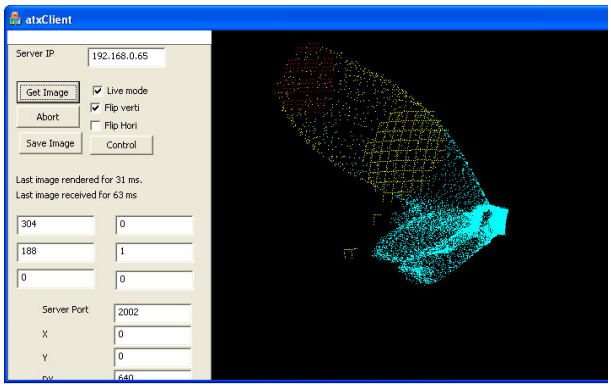
To detect the fingertip or palm angle, three stages of the signal process are performed. In the first stage, a prefiltering process reduces noise, and the 640 by 480 depth array is divided into 10 by 10 sensing windows. The average intensities are calculated for 64 by 48. Second, the process obtains the position of the finger tip or the angle of the palm. The last stage is a post process for converting the motion data to a position command for the cursor on the PC monitor.



(a) Finger motion for pointer control



(b) example trajectory for indexing motion

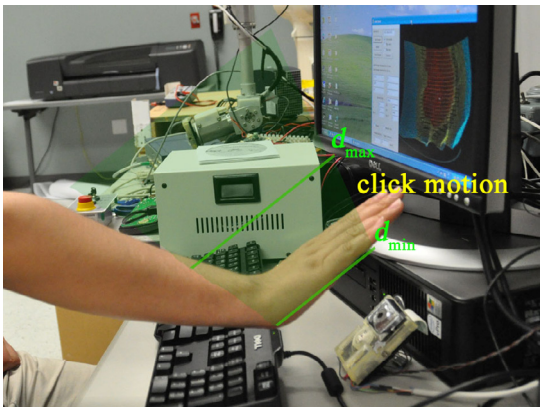


(c) distance map when finger input

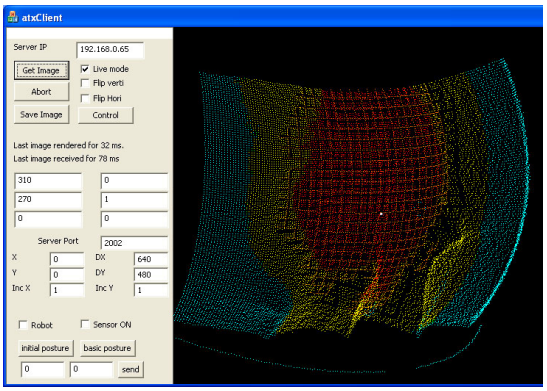
Fig. 7. Finger input motion and 3D sensing data

A. *Finger input mode*

To detect the fingertip, the process's algorithm searches for the window with the shortest distance. The working range is defined from d_{min} to d_{max} (i.e. from 50mm to 150mm in this research). If no object is detected in the working range, it implies no meaningful motion, and the command for moving the cursor is fixed. This limitation of the working range is also vital, in order to allow an index motion. The indexing is a widely used method in a typical mouse interface to enlarge working space by lifting the mouse off the table and repositioning it after the mouse has reached the edge of its available workspace, while the cursor remains fixed in position. In the developed interface, the user moves the finger tip back from the working range to the indexing area, as shown in Fig 7 (a). The user shifts his finger tip in indexing area and dives again into working range at the repositioned location. Figure 7(b) shows typical motion trajectory while control the cursor. The user get into the working range (① in Fig. 7(b)) and control the cursor by lowering the finger tip (② in Fig. 7(b)). If the user want to lower down the cursor more at the end of the working range, then move back to indexing area (③ in Fig. 7(b)) and reposition the finger for indexing (④ in Fig. 7(b)). While indexing. the cursor wait at the last position. After indexing, finger tip come into the working range again (⑤ in Fig. 7(b)), and finally control motion is continued (⑥ in Fig. 7(b)). Empirically, the fingertip encompasses 4 to 8 windows in the working range, as shown the example distance map in Fig. 7 (c). Therefore, the nearest 4 to 8 windows in the working range are searched, and the centroid of the windows is calculated as the location of the finger tip. If the more windows are found, the click motion is checked with different criteria. The case has less windows is regarded as indexing motion. the click motion is defined to dash to sensor with unfold palm, as shown Fig 8, and the distance map would be appeared as like Fig. 8 (b). When the click motion occurred, more than 80% of windows shows the distance in the working range, and the most distance values in those windows are over the predefined threshold.

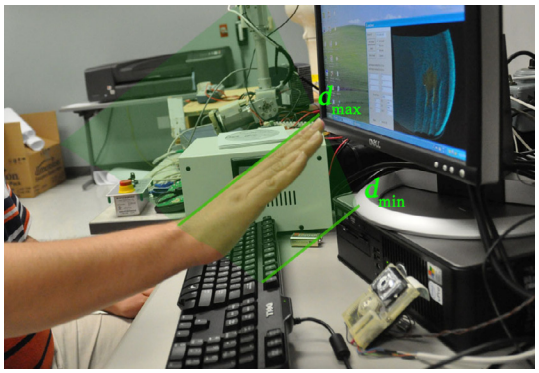


(a) click motion

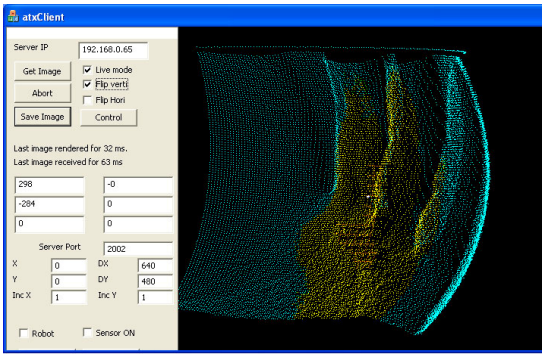


(b) distance map when click

Fig. 8. Click motion and 3D sensing data



(a) Palm motion for pointer control



(b) distance map when controlling pointer with palm

Fig. 9. Palm input motion and 3D sensing data

B. *Palm input mode*

The first step of the palm input mode requires determining whether or not the user's hand is properly located in front of sensor. The sensor checks each distance value for all the windows so that it can measure the distance from hand and roughly guess the angle. When the hand posture is located in the working range and ready to input, as shown in Fig 7,

About 80% of windows measure to a distance value in the working range. By averaging the values of those windows in the working range, the distance of the hand can be calculated. Even if the entire hand is located in the range, the higher distribution of the values of those windows implies a large angle from the hand. The windows in working range are regarded as effective windows, and ineffective windows are

excluded for measuring.

After checking the proper posture, the palm angle is precisely measured with the effective windows. In this research, least square multiple regression is used to find a best fit plain. Each window's array index implies location x, y, and the distance value at the window's implies distance z. The plain equation of palm can be defined as

$$z = a + bx + cy \tag{5}$$

For given data set (x_1, y_1, z_1), (x_2, y_2, z_2), …, (x_n, y_n, z_n), where n>3, the best fitting curve $f(x)$ has the least square error,

$$\Pi = \sum_{i=1}^{n} [z_i - f(x_i, y_i)]^2 = \sum_{i=1}^{n} [z_i - (a + bx_i + cy_i)]^2 = \min \tag{6}$$

To obtain the least square error, the unknown coefficient a, b, and c must yield zero first derivatives.

$$\begin{aligned} \frac{\partial \Pi}{\partial a} &= 2\sum_{i=1}^{n} [z_i - (a + bx_i + cy_i)] = 0 \\ \frac{\partial \Pi}{\partial b} &= 2\sum_{i=1}^{n} x_i [z_i - (a + bx_i + cy_i)] = 0 \\ \frac{\partial \Pi}{\partial c} &= 2\sum_{i=1}^{n} y_i [z_i - (a + bx_i + cy_i)] = 0 \end{aligned} \tag{7}$$

Expanding the eqn (7) as

$$\begin{aligned} \sum_{i=1}^{n} z_i &= a\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i z_i &= a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i z_i &= a\sum_{i=1}^{n} y_i + b\sum_{i=1}^{n} x_i y_i + c\sum_{i=1}^{n} y_i^2 \end{aligned} \tag{8}$$

Finally, a, b, and c can be obtained from the linear Eqn. (8), and the polynomial constant defines the normal vector of the plain.

In the palm input mode, the command for cursor control is mapped to the normal vector, and the click operation is defined in the same manner as the finger input mode.

IV. Performance Evaluation

A. Experimental results

To evaluate the developed interface, experiments for measuring the throughput and the completion time were carried out. Throughput is a well known performance index for measuring usability of input devices, such as a mouse and a track ball. The goal of the designed experiments was to measure the throughput and the completion time of the developed interface and compare the performance with the other input devices. Eight students participated in this experiment. There were 6 males and 2 females between ages of 24 and 32. None of the participants reported any sensory difficulties nor trouble on hands.

The participant sat in front of the developed interface and pointed the 3D sensor with index finger to input the motion. When the experiments started, a target button appeared at random location on the screen. The experiments were similar to the one dimensional Fitt's law task described by Douglas et al. [13]. The participant should control the cursor to move to the target button and successfully click the target button as soon as possible. Each trial started with a target appearing at a random horizontal location on the screen.

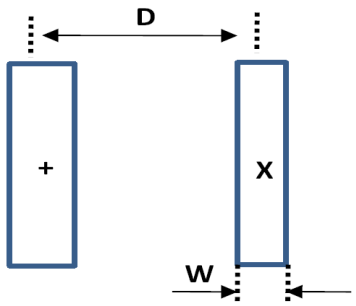
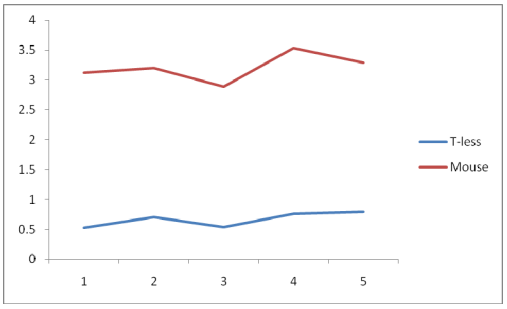
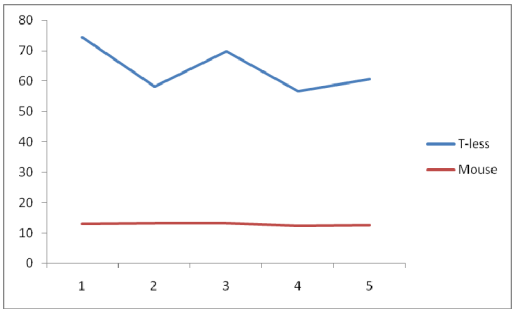


Fig. 10. Selection trial example

Figure 10 presents an example of a selection trial. The cursor was initially located on the left button, and the target button, whose width was w, was located in right side at the distance D. A successful movement and click allow to finish the trial and induced another target button to click. 20 buttons were shown in random order. The experimental results were as shown in Fig 11. The throughput is lower than a typical



(a)Through put



(b)completion time

Fig. 11. Experimental results

mouse, and completion time also much greater than a mouse. This discrepancy is due to the low gain and the errors at the transition motion between movement and click. If transition motion is clearly washed out, the performance will be increased. There is room for enhancing the performance by tuning up the gain fit the finger motion. All trials completed without a miss in order to prove practical reliability of the developed design. Further research to improve controllability and experiments for evaluation are in progress.

V. CONCLUSION

This research proposed a new touch-less interface, T-less, by using an IPA sensor. Since the distance map was directly obtained from the sensors, the position of fingertip and also the normal direction of palm were easily determined. The developed interface had two different operation modes. In finger input mode, a cursor can be controlled by a pointing motion. Another way to input was by rotating the wrist and changing direction of the palm. The way to figure out the palm angle by calculating best fittng plain was explained The developed interface and all the operation methods were designed to fully utilize the distance information so that the recognized hand gesture is highly reliable.

To evaluate the performance of the T-less, experiments for operating the cursor and click function were performed. The results show that the users reliably controlled the cursor on the screen and clicked on the given targets. The evaluated performance based on throughput and work time was lower than a typical mouse, but practically it can be utilized in the field requiring touch-less interfaces.

REFERENCES

[1] Alberto Tomita and Rokuya Ishii, "Hand Shape Extraction from a Sequence of Digitized Gray-Scale Images," in Proc. of International Conference on Industrial Electronics, Control and Instrumentation, pp. 1925-1930, 1994

[2] Wong Tai Man, Sun Han Qui, and Wong Kin Hong, "ThumbStick : A Novel Virtual Hand Gesture Interface," in Proc. of IEEE International Workshop on Robot and Human Interactive Communication, pp. 300-305, 2005

[3] Alvin T.S. Shan, Hong Va Leong, and Shui Hong Kong, "Real-time Tracking of Hand Gestures for Interactive Game Design," in Proc. of IEEE International Symposium on Industrial Electronics, pp. 98-103, 2009

[4] Luc Vlaming, Jasper Smit, and Tobias Insenberg, "Presenting using Two-Handed Interaction in Open Space," in Proc. of IEEE International workshop on Horizontal Interactive Human Computer Systems, pp. 29-32, 2008

[5] Yoichi Sato, Yoshinori Kobayashi, and Hideki Koike, "Fast Tracking of Hands and Fingertips in Infrared mages for Augmented Desk Interface," in Proc. of IEEE International Conference on Automatic Face and Gesture Recognition, pp.462-467, 2000

[6] Yoshio Iwai, Ken Watanabe, Yasushi Yagi, and Masahiko Yachida, "Gesture Recognition using Colored Glove," in Proc. of International Conference on Pattern Recognition, pp.662-666, 1996

[7] Akira Utsumi, Tsutomu Miyasato, Fumio Kishino, and Ryohei Nakatsu, "Hand Gesture Recognition Sytem using Multiple Camera," in Proc. of International Conference on Pattern Recognition, pp.667-671, 1996

[8] Kiyofumi Abe, Hideo Saito, and Shinji Ozawa, "3D Drawing System via Hand Motion Recognition from Two Camera," in Proc. of IEEE International Conference on Systems, Man, and Cybernetics, pp.840-845, 2000

[9] Koichi Ishibuchi, Haruo Takemura, and Fumio Kishino, "Real Time Hand Gesture Recognition using 3D Prediction Model," in Proc. of IEEE International Conference on Systems, Man, and Cybernetics, pp.324-328, 1993

[10] Stefan Soutschek, Jochen Penne, Joachim Hornegger, and Johannes Kornhuber, "3D Gesture-Based Scene Navigation in Medical Imageing Applications using Time-Of-Flight Cameras," in Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 1-6, 2008

[11] Dugan Um, "Infrared Photometry for 2D Proximity Sensing and 3D Geometry Visualization," Journal of Engineering and Technology, vol. 1, 2007

[12] Novotny, P.M. and Ferrier, N.J., "Using Infrared Sensors and the Phong Illumination Model to Measure Distance," IEEE Int. Conf. on Robotics and Automation, pp. 1644-1649, 1999.

[13] Douglas, S.A., Kirkpatrick, A.E., and MacKenzie, I. S. Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. The ACM Conference on Human Factors in Computing Systems – SHI'99, pp.215-222, 1999