Keyboard Input by Movement of the Finger and Pointer using a Smart Device

[1]Jungpil Shin, [1]Hiromasa Omote, [2]Cheol Min Kim

[1]School of Computer Science and Engineering,
The University of Aizu, Fukushima, Japan
E-mail: jpshin@u-aizu.ac.jp
[2]Department of Computer Education, Jeju National University, Korea
E-mail: cmkim@jejunu.ac.kr

***Abstract*— Nowadays, users of smartphones use keyboard input, flick input and multi-tap input. Moreover, there are keyboard input and handwritten character input using a pen tablet for the PC. There is some input methods based on handwriting of characters in the air using Kinect and a web camera in much research. However, character input by using devices such as Kinect and a web camera are limited by the environment. The user also needs to learn a new input method because many systems use the original character input method. We reduce the burden of users by using a smart device with a high penetration rate. Moreover, we use the keyboard input method because it is a general input method. Therefore, users do not need to learn a new input method. Our method can be used by the various environments by using a smart device. Since character input from long distance can be performed using a camera of a smart device moreover, our method can be used broadly.**

Keywords- Keyboard Input; Finger Recognition; Smart Device

I. INTRODUCTION

In recent years, the number of character input methods has increased significantly. For example, there is keyboard input, voice input and handwriting input on the PC. Table 1 shows the research about character input method. There is handwritten input in the air using Kinect [1] and character input using Leap Motion [2]. Moreover, there is smartphone and tablet keyboard input, multi-tap input, flick input and the voice input.

However, users need to learn the newly developed character input method. It is a burden for users.

Currently, there are many devices using character input. For example, there is the keyboard, web camera, Kinect, Leap Motion, the pen tablet and the microphone. These devices are used as accessory devices on a PC. Therefore, users are limited to a location to use the device.

Smart devices are evolving. Therefore, smart devices will be larger in size. Moreover, the spread of smart phones increases character input in the usage of the Internet. However, the conventional character input system cannot correspond to the difference among the possessors and the length of a user's finger. Therefore, the character input of a smart device does not have good operability.

There are a lot of recognition methods for the finger and the hand. Some examples are the hand recognition using a depth camera [10], Kinect recognition of the hand and fingertip [11] and recognition using a colour glove [12]. However, these methods require a special tool. Therefore, users cannot easily use these systems. Special tools do not have a high penetration rate. So, users do not use the recognition system.

We solve this problem by using a smart device. This paper has three purposes. First, users can input characters regardless of the location and the environment. This system uses only a smart device, so this system does not require a special tool. Therefore, users are not limited to a location and environment to use this system. Second, users can input characters without special tools. Last, users can input characters at a remote location. Users can input characters without touching the screen by using this system. Therefore, users can input characters at a remote location. Users can input by touching using the virtual keyboard that is displayed on the screen of the smart device by the camera. The method of inputting characters uses the keyboard input that is most used in the PC and smart device. Users do not need to learn a new character input method by using the keyboard input method. An accessory device of a PC is used in most study about character input. This system is an input of the character using the front camera of a smart device. Therefore, users can easily use this system because this system does not need an attached device.

This system recognizes fingertips by using hue recognition and template matching. Hue recognition gets the hue information of users in advance. Thereby, this

TABLE I. RELATED PAPERSR OF THE CHARACTER INPUT METHOD

Researchers	Using Device	Main Device	Method of a Character Input
M. Fujimoto et al. [3]	Original Device	PC	Keyboard
N. Matsui et al. [4]	RGB-Camera	PC	Original Keyboard
M. Weiss et al. [5]	SLAP	PC	Virtual Keyboard
Y. Nishida et al. [6]	Kinect	PC	Aerial Handwritten
M. Kaneko et al. [7]	Non	Smart Device	Original Method
S. Takahiro et al. [8]	Non	Smart Device	Original Method
R. Yamamoto et al. [9]	Earphone	Smart Device	None

system correctly recognizes the background and fingertips. In addition, only the hue recognition recognizes the fingertips. Therefore, this system is template matching using a fingertip image of users.

This system is available (1) in the any impossible place to touch such as surgery, (2) any environment without PC or Camera and (3) any remote place in the range of appropriate distance.

In Section 2 we show the approaches for keyboard input. In Section 3 we show the experiment and results. In Section 4 we show the discussions. In Section 5 we conclude the paper.

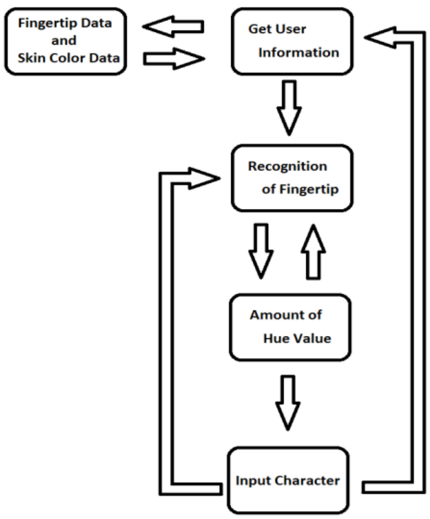

Figure 1. Flow chart of this system

II. KEYBOARD INPUT APPROACHES

This system targets a device, such as a smart phone, to operate it by touching the screen. This system uses a front camera of a device. So, users do not need to have a new device.

This system performs the character input by the step shown in Fig. 1. First, this system gets the information of the user's hand. It is the colour information of the skin and fingertip information. The colour of the skin and the shape of the hand of each person are different. We increase the accuracy of this system by corresponding it to the difference of shapes and colours. In addition, input is done by fingertip's touching the virtual keyboard using the camera. Therefore, we need the recognition with high accuracy. We explain these steps in detail.

A. System Simulation

This system gets the information of a user's hand before the character input. First, the screen of the device is projected as in Fig. 2. Users see their fingertip on the green square. Then, the user presses a button on the bottom centre of the screen. As a result, the screen changes to the character input screen. As shown in Fig. 3, on the character input screen, a virtual keyboard and a red square are projected. The user moves the fingertip to the position of the character. Then, the fingertip selects the virtual keyboard character by contact. This system can also recognize objects that are not a fingertip. However, this research is mainly performed on the use of fingertip.

B. Get User Data, Skin-color data and Fingertip data

This system gets the information of the user's hand before the character input. The user information is necessary to get correctly because it changes the accuracy of the character input system. The user's information for recognition is the colour of skin and information on a fingertip. The colour of the skin is different in a person respectively. It is the cause of recognition errors. In addition, the colour of the skin changes by a change in the environment like the brightness. Therefore, we increase the accuracy by getting the hue information at system start-up. A green frame is reflected on the screen at system start-up.

When this system starts, the information of the skin colour of the user can be acquired by putting a finger in the indicated green square. This can be detected by checking the occurring frequency of the hue value and using the threshold with the high occurring frequency. Thereby, the hue value of the background is not included.

In this system, the fingertips cannot be recognized only by hue information because an error in fingertip recognition might occur due to noise. The fingertip information is required to eliminate the error. The fingertip information is a fingertip image of the top of finger from the first joint of the finger. In this system, the information on a fingertip also can acquired when information with the colour of the user's skin is acquired.

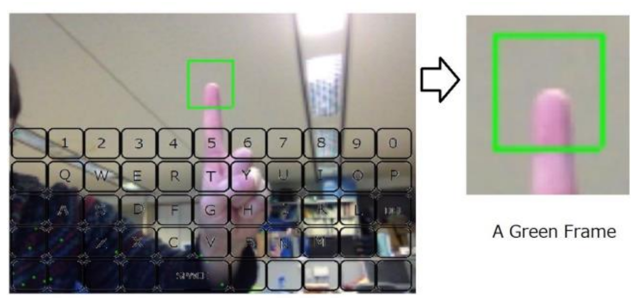

Figure 2. Screen of getting user data; A green frame is reflected on the screen at system start-up

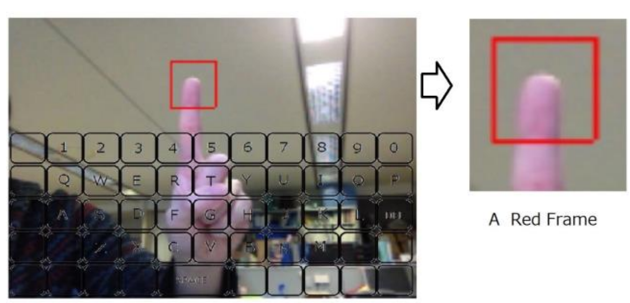

Figure 3. Screen of Recognized Finger of Character Input

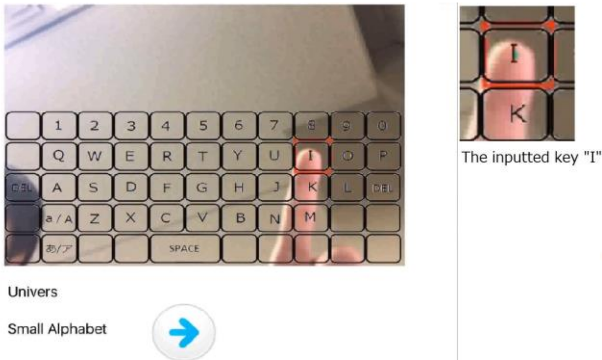

Figure 4. An example of inputted key "I" by fingertip recognition

C. Recognition of the Fingertip

This system uses two methods to recognize the fingertip. The first is template matching using the fingertip information. The second is how to find the hue of a certain area of skin in the region.

Fig. 4 shows an example of keyboard input by the recognition of fingertip. The "I" key was selected by the red square.

This system uses template matching to detect a user's fingertip. Template matching is a technique in digital image processing for finding small parts of an image, which match a template image. A template image uses the fingertip image of the user. The calculation uses zero mean normalized cross-correlation (Fig. 5). Therefore, it is possible to stably calculate variations in brightness. As a result of the calculation, the lowest point of the value is the fingertip of a candidate. However, the point does not determine the fingertip.

At the second step, the hue of the skin in the specific area can be searched. The part recognized as a fingertip can be searched by template matching. When the skin colour of the region exceeds the fixed value, its location is the fingertips.

In the present system, we must have the skin colour threshold for fingertip recognition. Below describes a method for determining the threshold value. First, we perform fingertip recognition in the 30–40% threshold for the fingertips candidate. We perform the fingertip recognition 10 times in each threshold. Next, the threshold is increased by 10%. The threshold is recognized up to 80–90%. Table 2 is the number of successful recognitions in each threshold. From Table 2, we set the threshold to 50%–60%.

D. Using a Keyboard for the Input of a Character

This system uses the keyboard input method. Originally, the keyboard input method can input characters by pressing the keys. In this system, a fingertip reflected by the camera inputs a character by the user touching the virtual keyboard. Therefore it's necessary to decide touching time between a fingertip and a virtual keyboard. A decision procedure of touching time of a fingertip and a virtual keyboard is described in next sentence.

The letters "abcdefghijklnmopqrstuvwxyz" are entered in this system. The touching time of character input is 0.5, 1.0, 1.5 and 2.0 seconds. We measure the number of erroneous inputs and the input time. Table 3 is the result of each number of seconds.

From Table 3, we set the touching time to 1 second. It's because in case of 0.5 seconds there is many input error, and input time is too long in case of 1.5 seconds and 2.0 seconds.

$$\bar{T} = \frac{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} T(i,j)}{MN} \qquad \bar{I} = \frac{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1} I(i,j)}{MN}$$

$$R_{ZNCC} = \frac{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1}\left(\left(I(i,j)-\bar{I}\right)\left(T(i,j)-\bar{T}\right)\right)}{\sqrt{\sum_{j=0}^{N-1}\sum_{i=0}^{M-1}\left(I(i,j)-\bar{I}\right)^2 \times \sum_{j=0}^{N-1}\sum_{i=0}^{M-1}\left(T(i,j)-\bar{T}\right)^2}}$$

n = Template image of height, m = Template image of height
Search pixel = (i , j)
I (i , j) = Brightness Value of Input Image,
T(i , j) = Brightness Value of Template Image

$\bar{I}$ = Average Brightness Value of Input Image
$\bar{T}$ = Average Brightness Value of Template Image

Figure 5. Zero mean normalized cross correlation

TABLE II. RELATED PAPERS OF CHARACTER INPUT METHOD

Skin Color Threshold (%)	30 -40	40 - 50	50 - 60	60 - 70	70 -80	80 - 90
No. of Successful Recognition	0	2	9	2	0	0

III. EXPERIMENT AND RESULTS

In this paper, we perform three experiments. The first experiment evaluates the accuracy of fingertip recognition. In this experiment, subjects performs the input 10 times for the letters from "a" to "z". In the second part of of the experiment, subjects input the phrase "university of aizu". After inputting, the subjects evaluate five items in five levels. We evaluate the utility of the system for each item. By two experiments, we evaluate the accuracy of fingertip recognition and keyboard input, the utility of the system. We use Apple's iPad 2 in the experiment. Device performance is described in Table 4.

A. Experiment of Finger Recognition

In this experiment, subjects perform the input 10 times for letters from "a" to "z". Experiments are performed in the situations of Fig. 6. Subjects are 30 cm away from the device. When a subject enters a character, the subject releases the finger from the virtual keyboard.

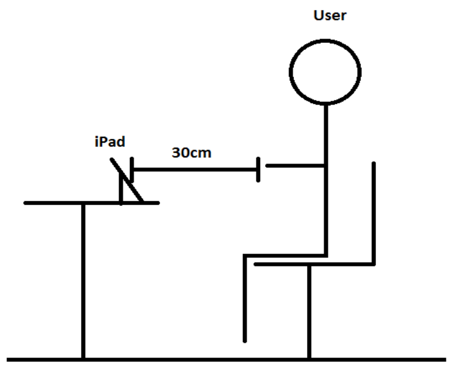

Figure 6. Simulation of Experiment

TABLE III. EXPERIMENT OF TOUCHING TIME THRESHOLD

Touching Time (second)	0.5	1.0	1.5	2.0
Input Time (second)	33	40	60	154
Erroneous Input	13	2	0	0

TABLE IV. PERFORMANCE OF IPAD 2

CPU	GPU	Memory	Camera
Apple A5 Dual Core Processor 1GHz	Power VR SGX543MP2 Dual Core Processor	512MB LPDDR2 RAM	VGA Image quality Max 30fps

TABLE V. EVALUATION ITEM

Evaluation Item	Explanation
Learnability	Users can easily use the System.
Efficiency	Users can achieve efficient purposes.
Memorability	Users do not forget how to use the system.
Errors	The system is a low incidence of error. Users can be easily addressed error.
Satisfaction	Users are satisfied with the system. Users can use without stress.

This experiment is to measure the recognition rate of each letter. As a result, we will evaluate the accuracy of fingertip recognition.

B. *Experiment of Performance Evaluation*

In the experiment, the subjects enter the phrase “university of aizu”. The experiment will be performed on the condition shown as Fig. 6. After entering the phrase, the use evaluates the system as 5 steps for each of 5 items such as Learnability, Efficiency, Memorability, Errors, Satisfaction. Evaluation items refer to Jakob Nielsen [13].

From this experimental result, we measure the utility of this system.

TABLE VI. RESULT OF FINGERTIP RECOGNITION

Character	a	b	c	d	e	f
Detection Accuracy (%)	97	96	98	100	100	100

g	h	I	j	k	l	n	m	o	p
100	100	100	100	100	98	96	97	100	99

q	r	s	t	u	v	w	x	y	z
100	100	100	99	100	97	100	98	100	96

IV. DISCUSSION

From Table 6, the average input accuracy of the characters is 98.96%. The highest recognition rate is 100%, and the lowest recognition rate of character (b, m, z) is 96%. For research of character input using a virtual keyboard [4], the recognition rate average is 66%. Moreover, research using a smart device [14] has a recognition rate of 87% to 93%. Research of the character input gesture in Kinect [1] is the character input of Japanese. As a result, the recognition rate of each character is 50% to 100%, and the recognition accuracy is 86%. Our research obtained the better results than these studies.

This system is used for recognition of colours of objects and faces. For this reason, this system can recognize objects that are not a fingertip. There are three causes of error. First, we get the user’s information in advance. However, there is a case that cannot be successfully acquired. The reason is that the image quality of the camera is low. This system can use a video image with the size of 1280 x 720 pixels. However, in this paper, we are using a video image of 192 vertical pixels × 144 horizontal pixels for the weight of the program. Therefore, there is a possibility of obtaining an incorrect user’s hue value.

Secondly, when the user moves the finger quickly, the camera may not recognize the movement of the fingertip. This is caused by the specs of the device. The device used in the experiment is the iPad 2. The OS initially installed is iOS 4.3. However, the OS used in the experiment is iOS9. Therefore, the device does not have enough performance. To solve this problem, it is necessary to change the device used in the experiments to a newer device.

The third reason is the face as reflected on the screen. The hue information of the fingertips and the face is almost the same value. When the fingertip and the face are far away, the fingertip is recognized correctly. However, when the fingertip overlaps with the face or when the fingertip is near the face, the fingertip is not recognized correctly (Fig. 7). To solve this problem, it is necessary to separately

 ISBN: 978-1-61208-502-9

recognize the fingertip and face by using depth. However, it is difficult to recognize depth in a smart device.

From Fig.8, the average of each evaluation item is 4.58. The highest item, “Memorability”, is 4.9. “Learnability” is a high item, second at 4.8. (These are very good results compared with the result of input method using cell-phone [15]) With this result, it is considered very easy to use system for users by employing the keyboard input method. It was able to reduce the burden of the input method.

However, “Efficiency” was 4.1. This is the lowest result for the evaluation items. This is because, when the user moves the finger quickly, the camera may not recognize the movement of the fingertip.

In addition, when the user inputs a character, the fingertip requires contact with the virtual keyboard for a certain period of time. Therefore, character input is slowed down. However, “Satisfaction” is 4.6. This value is not low. From this result, we understand that users use the system with less stress.

Figure 7. The error cases of finger.

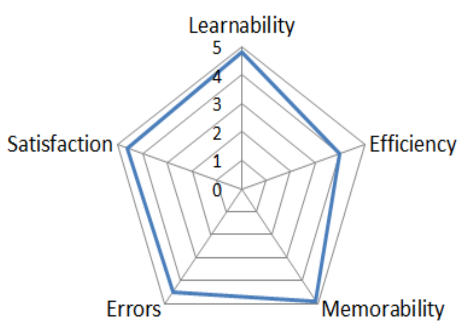

Figure 8. Result of Performance Evaluation

V. CONCLUSION

This paper has three purposes. First, users can input characters regardless of the location and the environment. Second, users can input characters without special tools. Last, users can input characters without touching the screen by using this system. Therefore, users can input characters at a remote location. Character input by fingertips by using the smart device has a high precision with an average of 98.96% for each character. In addition, users use this system and users evaluate five items in five levels. As a result, we obtain an evaluation of the average at 4.58.

We did fingertip detection using a smart device. Then, we made character input from a long distance possible by a virtual keyboard. We did not use special tools. Then, we input a character, and the character input did not depend on the environment and location. Therefore, we were able to obtain high practicality.

As future works, first the initial camera change mode will be developed for more comfortable use. Second the model to support multi-finger input (e.g., when using the two hands) will be developed, as users on smart devices strive to be more and more efficient in text messaging.

REFERENCES

[1] O. Masato and M. Sou, “KooSHO: Japanese Input Environment Based on the Handwriting Gestures in the Air,” The Association for Natural Language Processing, Proceedings of the NAACL HLT 2013 Demonstration Session, pp. 24–27, Atlanta, Georgia, June 2013.

[2] K. Hosono, M. Sasakura, H. Tanabe, and T. Kawakami, “A proposal of the character input method by gesture operation using Leap Motion,” The 28th Annual Conference of the Japanese Society for Artificial Intelligence, pp. 1-4, 2014.

[3] M. Fujimoto, M. Imura, Yoshihiro Yasumuro, and Yoshitsugu Manabe, Kunihiro Chihara, “AirGrabber : vlrtual keyboard using miniature camera and tilt sensor,” The Virtual Reality Society of Japan (TVRSJ), vol. 9, no. 4, pp. 413-422, 2004.

[4] N. Matsui and Y. Yamamoto, “Virtual Keyboard: Realization of Robust Real-Time Fingertip Detection from Video-Image,” Proc. SPA2000, Japan Society for Software Science and Technology, 2000-3, pp. 1- 7, 2001 (in Japanese).

[5] M. Weiss, R. Jennings, J. Wagner, R. Khoshabeh, J. D. Hollan, and J. Borchers, “SLAP: silicone illuminated active peripherals,” Medical Computing Group, ITS 08. 2008.

[6] Y. Nishida and T. Yoshida, “A research for aerial handwritten gesture recognition with Kinect interaction”, Memoirs of the Fukui Institute of Technology, No. 440, pp. 25-30, 2014.

[7] M. Kaneko and M. Iwata, “Proposal and implementation of T9-Flick Keyboard on tablet devices screen”, Tokyo Conference of Institute of Electronics, Information, and Communication Engineers, D-6, pp. 223, 2015.

[8] S. Takahiro, M. Yoshiaki, “Smartphones text input system”, SIG Technical Report, Information Processing Society of Japan, Vol. 2013, No.3, pp. 1-6, 2013.

[9] R. Yamamoto and H. Miyashita, “Operation and authentication of smart phone by earphone”, Internal Conference of Information Processing Society of Japan (Interaction 2013), pp. 626-631, 2013.

[10] T. Aoki, M. Nomura, H. Tamura, “Finger language recognition using the RGB-D camera”, Internal Conference of Information Processing Society of Japan, 4ZB-3, pp.187-188, 2013.

[11] Y. Orimo, Y. Tamakuni, Daisuke Takahashi, and Noriyoshi Okamoto, “An experiment on finger spelling recognition by using Kinect,” The Institute of Image Information and Television Engineers, ITE Technical Report, Vol. 38, No. 9, pp. 31-32, Feb. 2014.

[12] K. Watanabe, Y. Iwai, Yasushi Yagi, and Masahiko Yachida, “Gesture recognition using a colored glove,” Institute of Electronics, Information, and Communication Engineers, vol.1, No. 25, pp.662-666 Aug 1996.

[13] J. Nielsen, 2002, “Usability engineering,” Morgan Kaufmann, 1993.

[14] A. Tatsuya, T. Jirou, and Mimatsu Kazuo, “Smartphones input method based on magnetic,” University of Tsukuba, Master’s thesis, 2015.
[15] Y. Masato, “A study of the software key input method of the operation device for the information weak,” Thesis for a degree in University of Kouchi, 2004.