Gestural flick input-based non-touch interface for character input

Md Abdur Rahim [1], Jungpil Shin [1] and Md Rashedul Islam [1]
[1] School of Computer Science and Engineering, The University of Aizu, Aizuwakamatsu, Fukushima 965-8580, Japan.

Cite: Rahim MA, Shin J, Islam MR. Gestural flick input-based non-touch interface for character input. The Visual Computer. 2019:1-4.

Abstract

A non-touch character input is a modern system for communication between humans and computers that can help the user to interact with a computer, a machine or a robot in unavoidable circumstances or industrial life.

Therefore, this paper proposes a gesture flick input system that offers a quick and easy input method using a hygienic and safe non-touch character input system.

In the proposed model, the position and state of the hands (i.e., open or closed) are recognized to enable flick input and to relocate and resize the on-screen virtual keyboard for the user. In addition, this system recognizes hand gestures that perform certain motion functions, such as delete, add a space, insert a new line, and select language, an approach which reduces the need for recognition of a large number of overhead gestures for the characters.

Introduction

Humans communicate with computers in many ways, and the interface between humans and computers is an important factor in the convenience of this interaction. At present, the user is expected to use both touch and non-touch interfaces to interact with the machine, and various different devices are used in daily life, such as computers, tablet PCs, and smartphones.

In view of the above, this paper proposes a new character input system based on the gestural flick input method that considers the position, state and motion gestures of hands, and which can be used to facilitate human–computer interaction.

•In this study, we use the Kinect v2 sensor to detect the color and depth of an image, body skeleton information, and hand motion gestures.

Based on recognizing the position and the open or closed state of the hands, this paper proposes a gestural flick input system that can be used to input characters, delete characters, add a space, insert a new line, and select a language. As a result, users can perform non-touch input under conditions that do not require touching any devices.

Proposed Model

This section describes the overall process of the proposed system, which allows non-touch character input based on captured information about the hand state (open or closed) and gestures via the flick input system.

The input information is acquired by the Kinect sensor and processed using our proposed model. We divide our system into two parts: (a) hand state and gesture recognition; and (b) the flick input system.

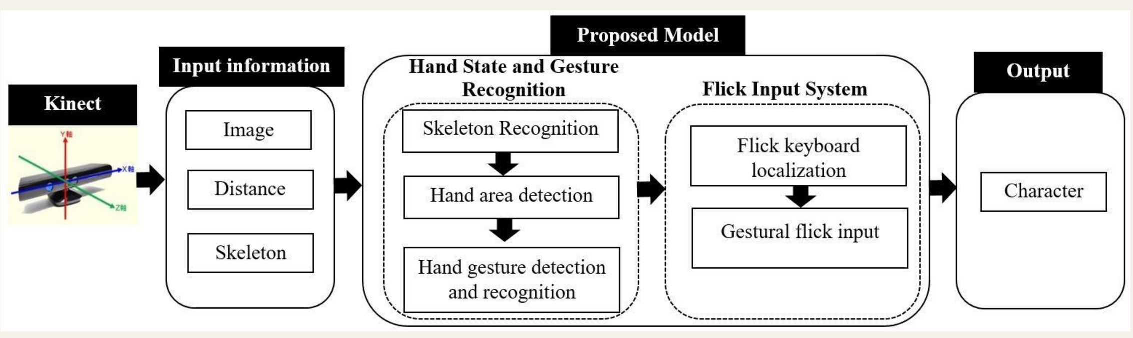

Figure 1. Overall scheme for the proposed system.

Detection of the hand area

The steps taken to identify the fingertip from the mask image in the dominant hand area in order to recognize the open or closed state of the hand are as follows:

Table 1. Motion-Function

Motion function	Description of hand gestures
Delete	Move the right hand from right to left
Add a space	Move the left hand from left to right
Insert a new line	Move both hands from top to bottom
Select language	Cross both hands

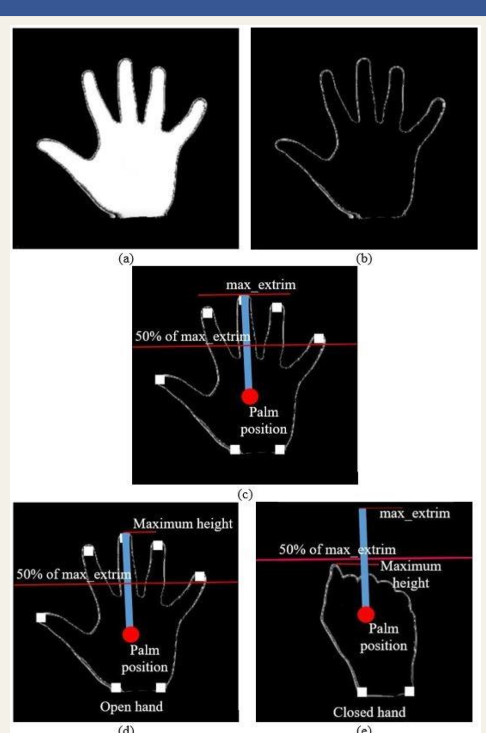

Fig. 2. Process of detection of the hand area and the open/closed hand state.

1. Identify the outline of the dominant hand. To do this, the system uses the FindContours function in OpenCV as features of the Suzuki85 algorithm. This function identifies the contour from a binary image called the "mask image," as shown in Fig. 2b. When the mask image of the dominant hand area has been detected, the outline and fingertip can also be detected. The position of the fingertip is at the extreme point from the hand center among each "long enough" convex part points of the effective hand outline. The distances (max_extrim) between the most extreme point at the top of the middle finger and the palm position are therefore calculated and are taken as the maximum finger length of the open hand. To detect whether the hand is open or closed, we extend a straight line to the (50%) point of max_extrim, as depicted in Fig. 2c.
2. To detect whether the hand is open or closed, the distances of all the top extreme points are calculated from the palm position in both states (open and closed).
   (a) If the maximum distance is more than 50% of max_extrim, the system identifies the location of the candidate fingertip and determines that the hand is open. An open hand is shown in Fig. 2d.
   (b) Alternatively, if the maximum distance is less than 50% of max_extrim, the system identifies that the candidate fingertip does not exist and determines that the hand is closed. A closed hand is shown in Fig. 2e.

Detection and Recognition of Hand Gestures

The steps used to recognize the motion function from the mask image are as follows:

1. Calculate the distance of the palm position of the dominant hand from both shoulders. The distances from the right shoulder (shoulder_right) to the palm position and from the left shoulder (shoulder_left) to the palm position are expressed as SRP1 and SLP1, respectively, for the right hand. The distances from the shoulder_right to the palm position and from the shoulder_left to the palm position are expressed as SRP2 and SLP2, respectively, for the left hand. When SRP1 is greater than SLP1, the "delete" function is executed.
2. When SLP2 is greater than SRP2, the "add a space" function is executed.
3. Calculate the distances for the palm positions of both hands, from the shoulder position to the hip position horizontally. When the distance of the palm position (at the same time, and for both hands) is greater than that from the shoulder position to the hip position, the "insert a new line" function is performed.
4. When SRP1 is greater than SLP1 and SLP2 is simultaneously greater than SRP2, the select language function is performed. Users can select any language using this method.

Flick Input System

The system uses a motion gestural flick input as a character input method. These motion gestures are simple enough that a user can remember and perform them.

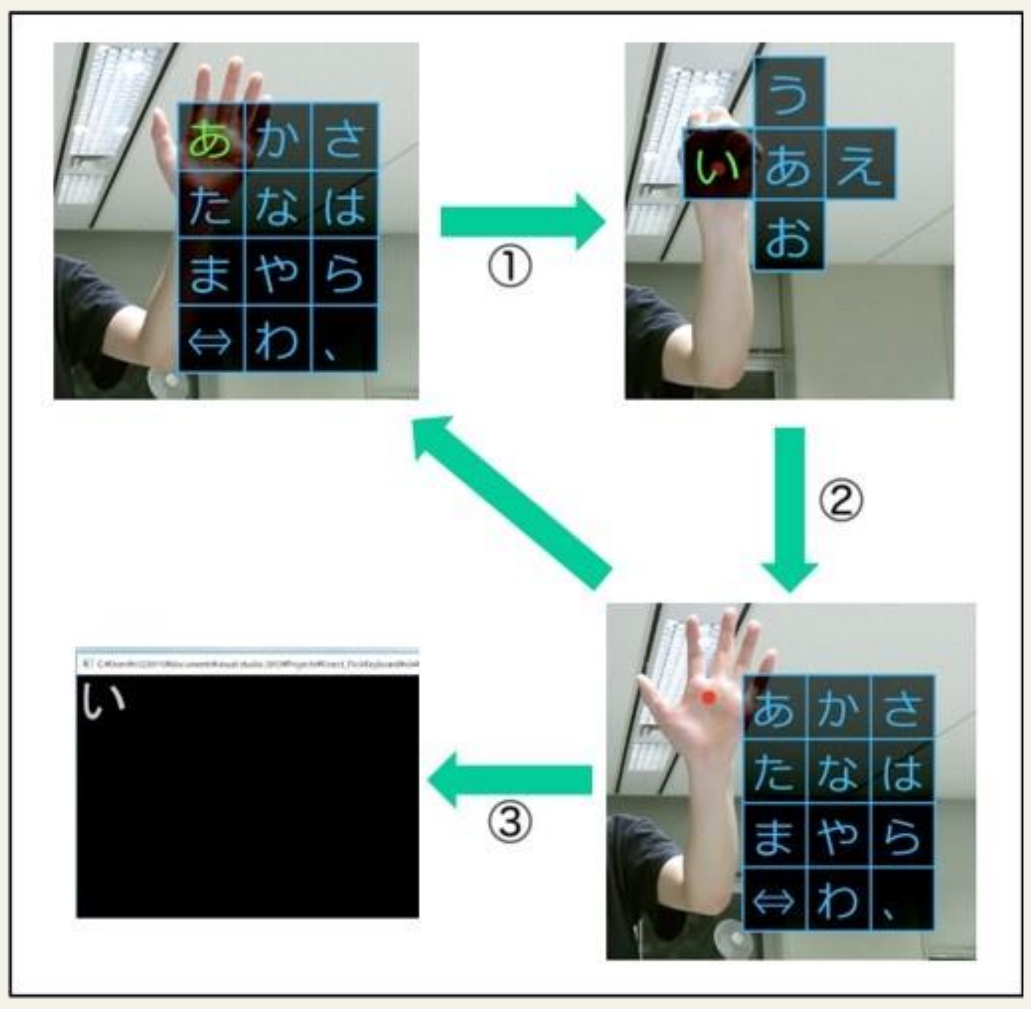

Fig. 3. Process of the gestural flick input system

The character input system shown in Fig. 3 uses the following procedures:

1. The user holds a hand open over the sequence for the first character that they want to enter as input, and then closes the hand. The flick keyboard is shown in the display monitor.
2. The user selects a suitable character by moving the closed hand on the flick keyboard to the key that they want to enter as input, and then opens the hand.
3. The selected character is input

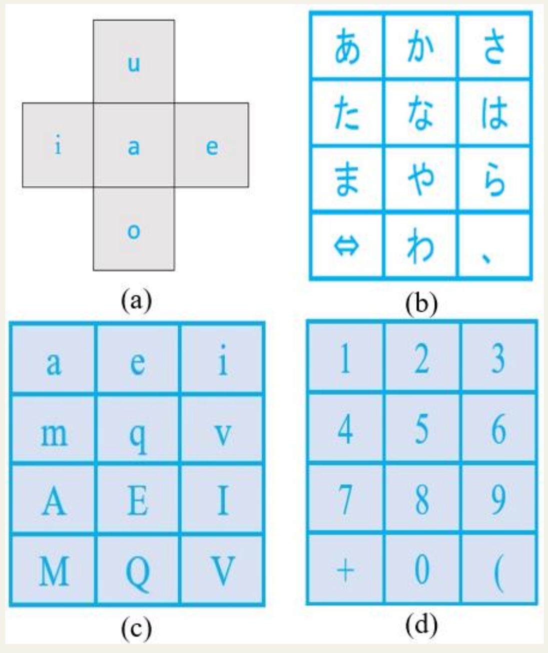

Fig. 4. Gestural flick input keyboard map

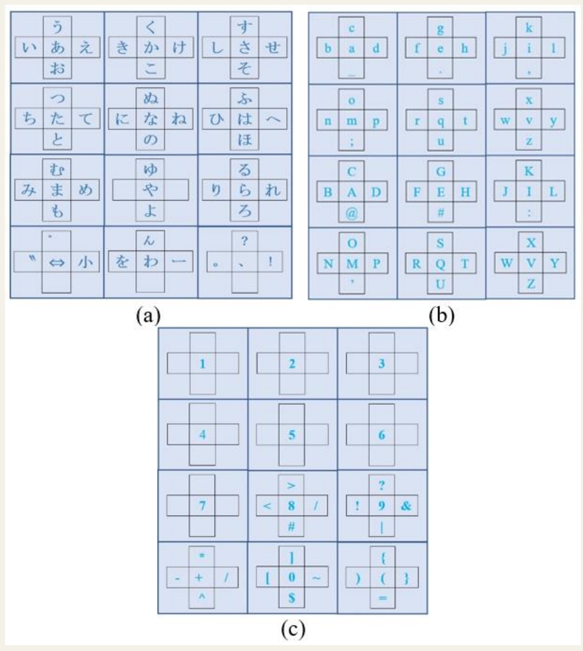

Fig. 5. Character sequences of the flick input.

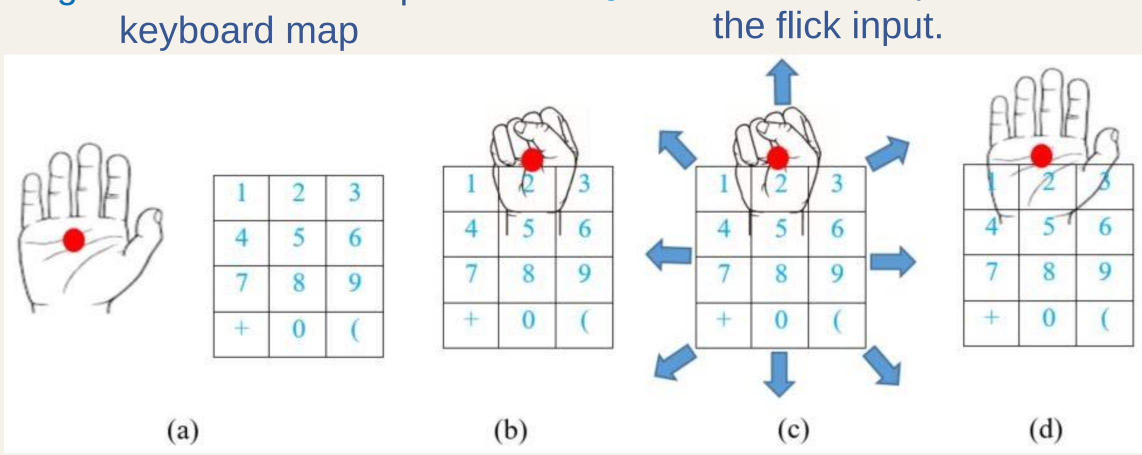

Fig. 6. Keyboard Localization.

Experiment, Results, and Evaluation

Table 2. Experiment results

User	Avg. character selection rate (%)	Avg. character recognition rate (%)
User 1	98.15	100
User 2	100	100
User 3	100	100
User 4	96.3	100
User 5	98.15	100
User 6	98.15	100
User 7	96.3	100
User 8	98.15	100
User 9	96.3	100
User 10	94.44	100
User 11	100	100
User 12	98.15	100
User 13	96.3	100
User 14	96.3	100
User 15	96.3	100
User 16	94.44	100
User 17	94.44	100
User 18	96.3	100
User 19	96.3	100
User 20	100	100
Average	97.22	100
Avg. recognition rate	98.61	

Fig. 7. Character input display screen

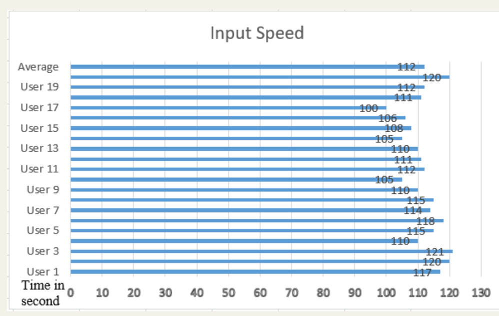

Fig. 8. Average speed of character input.

Table 3. Character selection accuracy

No. of users	No. of characters	Avg. input speed of whole experiment (in s)	Input speed (characters/min)	Avg. error rate of character selection
20	54	112	29.0	2.78%

Table 4. Recognition rate of motion function

Motion function	Avg. recognition rate (%)
Delete	97.0
Add a space	100.0
Insert a new line	95.0
Select a language	98.0
Average	97.5

Fig. 9. Flick keyboard display screen in the proposed system.

Conclusion

This paper proposes a non-touch character input system based on gestural flick input, using a Kinect sensor.

The proposed model recognizes the position, state (open or closed) and movements of a hand entering a character and performing other motion function inputs. A dynamic flick keyboard was developed for Japanese hiragana, English, and numbers, via which the user can perform character input quickly and easily.

Hand positions and movements are used to indicate the character of the flick keyboard for input, and the motion function is used for additional tasks in the character input system. Users can also adjust the size and location of the keyboard by hand motions.

The proposed system offers a natural and efficient non-touch interface technique that is specially designed to provide reliable performance for the most demanding environments and applications; this can reduce the gap between humans and computers, machines or robots and can resolve the underlying problems with an advanced input system, allowing the user to manage the system in a hygienic and safe way.

Contact: Jungpil Shin, Corresponding author. Phone/Fax: +81-242-37-2704, Email: jpshin@u-aizu.ac.jp
Web: http://www.u-aizu.ac.jp/labs/is-pp/pplab/

Pattern Processing Lab