OHAJIKI Interface: Flicking Gesture Recognition with a High-Speed Camera

Toshiki Sato, Kentaro Fukuchi, and Hideki Koike

Graduate School of Information Systems,
The University of Electro-Communications Choufu-shi, Tokyo , JAPAN 182-8585
{den,fukuchi,koike}@vogue.is.uec.ac.jp

Abstract. This paper describes a novel interaction technique that recognizes a finger flicking gesture for power adjustment input for a sports game, such as golf swing or hitting. Our system measures speed of the finger motion and direction of the gesture by using a high-speed camera and a high frame rate image processing technique. By using this system, users can adjust power and angle intuitively. We developed a 3D golf game using this interaction technique to provide an intuitive golf swing input.

1 INTRODUCTION

Conventional interfaces of power adjustment for video games employ common controllers (e.g. game pad) or dedicated input devices. For example, in a golf game, the player presses a button of a controller several times to shoot. Typically the pressure power on the button is ignored. However, the feeling from this action is very different from an actual golf swing. Since it is preferred that a video game provides a real and intuitive interaction, this input is not desirable.

Finger flicking, by the way, is a very familiar action to us, as anyone who had played marbles or flicked coins on a table. [1]

We consider that finger flicking has the same property with golf swing or hitting in respect of action that hits something to a distance. Hence, we propose to use the finger flicking for power adjustment, which can be useful for video games.

2 RELATED WORKS

There are two well-known finger tracking techniques. One uses special equipments or sensors, while the other uses a vision based recognition technique.

Data glove is a well-known glove-shaped input device for finger tracking. A data glove detects motion of the hand and the fingers, and it is used to recognize or record manual gestures, but a user has to put a sensor device on his hand that can act as an obstacle in the gesture.

Gokturk and Sibert introduced a finger mounted pointing device based on infrared sensing technique[1]. With the device, a user can point to a position in a display directly

[1] In Japan, playing marbles, or “Ohajiki”, is a popular game with children.

with his finger, but mounting a device on a finger is uncomfortable for the user. Besides, it restricts his hand motion physically.

By contrast, a vision based interaction technique can recognize a user's motion without using any special equipments or sensors, therefore the user can interact with the computer intuitively. Oka et al developed a real-time finger tracking system, which can detect positions of hands and fingertips[2].

However, most of vision-based finger tracking system has the same problems that are a tracking latency caused by the cost of image analysis and low scan rate of camera (30 frames per second). Because of that, those systems cannot track fast motion of fingers, such as finger flicking.

3 VISION BASED FINGER FLICKING RECOGNITION

3.1 Finger Flicking Motion

Finger flicking is a motion of hitting a small object with a quick motion of a forefinger or midfinger. This motion can be seen in our daily life (e.g. playing marbles or flicking a coin).

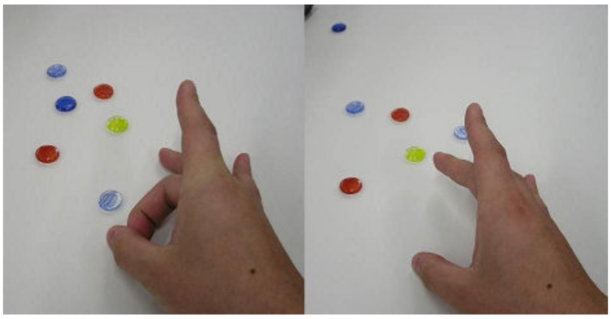

Fig. 1. Finger flicking: at first put a forefinger or mid finger on a thumb and add power to the finger (left), then release the finger (right).

Controlling the power on the finger enables to adjust the impact force of the flicking. Besides, its direction can be adjusted by changing the relative position or angle of the hand to the object.

This kind of motion can be found in various situations, such as hitting or kicking. In fact, we can find finger flicking motions in many kids plays, such as "finger flicking football". Therefore, we consider that finger flicking can be used as an intuitive input method of power in terms of using real force to control.

In addition, finger flicking is safer than a real action with a golf club or bat for home use, and it saves space. Besides, because it is familiar to us, it is unnecessary to learn how to do it.

3.2 Design Principle

It is difficult to put some sensors on a finger to track the quick motion of finger flicking gesture, because the size of the sensor is limited to the size of the finger and it may be different among users. Besides, because the motion is very quick, the sensor has to be fitted to a finger tightly, and its size and weight should be as small and light as possible.

For that reason, we used a vision based recognition technique to track finger flicking motion and measure the position and direction of a hand. We hypothesized that the power of the finger flicking can be recognized by measuring the speed of the finger motion which can be tracked by vision based techniques. However, its accuracy and scan rate are worse than the result by the hardware based recognition typically, because of the low frame rate of the camera. Because of that, it was difficult to track a quick motion by a camera. To solve this problem, we have employed a high-speed camera and developed a high throughput image processing system to track a quick finger motion.

4 IMPLEMENTATION

Figure 2 shows the hardware structure of our system. An image from the camera is processed by the image processing PC, and the results are sent to the application PC via Ethernet.

4.1 Finger Flicking Recognition

In a finger flicking gesture, while a flicking motion is done in very short time, positioning and directing are performed in relatively longer time. Therefore, the high-speed image processing is only required during a very short period of a flicking gesture. For that reason, our system consists of some recognition phases and each phase uses an optimal recognition mode. When the system tracks a quick motion, it scans the image in maximum speed of the camera with a simple image processing technique, while it uses rich image processing techniques with low scan rate when it is unnecessary.

We employed Point Grey Research's high-speed camera, Dragonfly Express, which has several scan modes and it can increase the scan rate by limiting the scan range. The size of an image, camera scan rate and image processing frame rate in each recognition phase is shown in the table below.

Table 1. Frame rate for each phase of recognition

Phase	Image Size	Frame rate	Scan rate
stance detection	640×480	200 fps	160 fps
window setting	320×240	320 fps	130 fps
fingertip detection	100×80	450 fps	450 fps

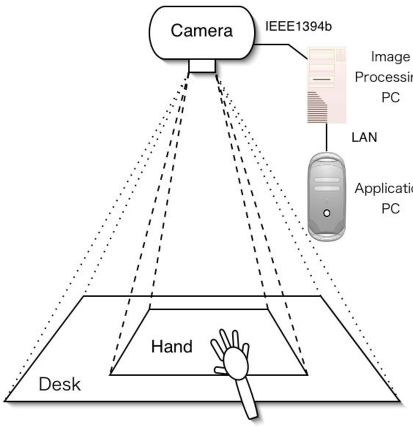

Fig. 2. System overview

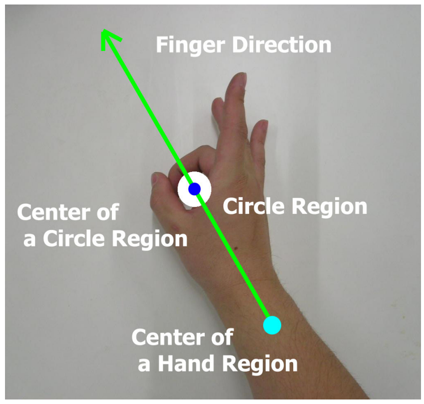

Fig. 3. A stance of finger flicking

4.2 Processes of Motion Recognition

The gesture recognition process consists of following three phases.

1. Hand detection phase
   At first, the system scans the table to find whether a user is going to flick with his finger or not. To recognize the gesture, the system detects a user's stance of flicking. We defined this stance that it makes a small circle region by a thumb and a flicking finger as seen in figure 3. The system detects this small circle region using a background subtraction and a region extraction process. When the region is detected, the system goes into the next phase.
2. Search window setting phase
   Next, the system sets a search window that detects a flicked finger. The size of this window is smaller than the original camera view used in previous phase, to increase scan rate of the camera and reduce the computational cost of the following image processing phase. The finger is tracked only in this window.
   The position of the windows is set according to the position and the angle of the user's hand. To set the window correctly, its position is determined in the following way: it sets the search window on the position shown in the figure 4 whose initial position is calculated from the positions of the circle region and the hand region. Then, it slides the window towards the circle region until the window conflicts with the hand region. Finally the window is set on this position.
   At the same time, the direction of the hand is calculated from the line between the center of the hand region and the circle region. When the user keeps his hand still for a while, the system goes to the next phase.
3. Fingertip Detection Phase
   The system changes its view to the search window and tracks the finger motion inside the view. This process uses background subtraction and template matching that uses a small circle-shaped template to detect a fingertip. When the fingertip

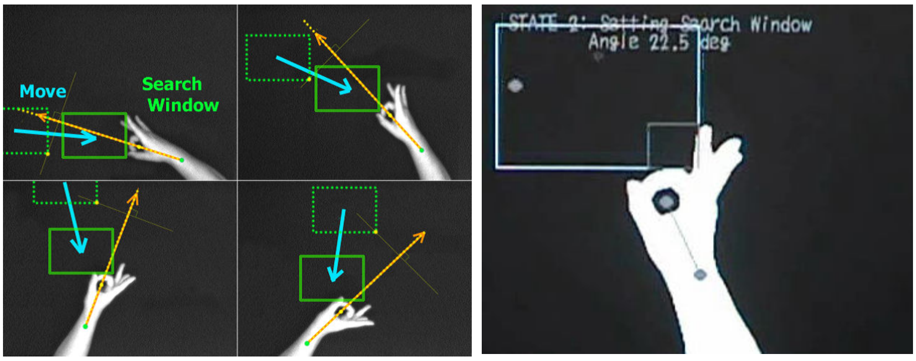

Fig. 4. Search window setting image(left), and the screen shot(right)

runs out from the camera view, the system stops tracking and calculates the velocity of the finger from the result.

4. Result Calculation Phase
   The velocity of the finger is calculated by the following steps: let T be a length of time (msec) from the first detection of a finger to the last, and L_n represents the distance in pixels between positions of fingers of n and $n+1$, and N is the number of detected fingertips. The velocity of the finger V (pixels/msec) is calculated as $V = \frac{\sum_{i=1}^{N-1} L_i}{T}$

5 EVALUATION

We ran an experimental evaluation to test that our system could recognize the power of the finger flicking. We should note that we did not measure the actual impact power of the finger flicking. Instead, we had four subjects to flick with their fingers five times with five grades of power, from "weakest" to "strongest". Figure 5 shows the result of the test. As seen in the graph, our system could distinguish those five grades of power. But the number of detected fingertips was 20–40 at the first (weakest) flicking, 10–14 at the second, 5–10 at the third, 4–9 at the forth, 2–4 times at the fifth (strongest). In this simple tracking, the result depends on the scan rate of the camera. Therefore, the precision of the result could decrease if a flicking speed is too fast. To solve this problem, we conclude that we should use a camera which has a higher scan rate to track the motion of a fingertip certainly, or to use a mathematical model of finger movement to estimate the speed precisely.

6 APPLICATION

We developed a simple 3D golf game that a player can shoot a ball by a finger flicking gesture. A player controls the direction of a shot by moving his hand on the input surface, and then controls the power of the shot by finger flicking.

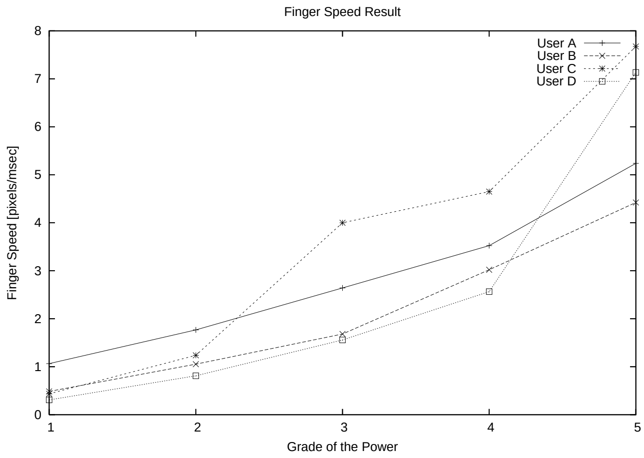

Fig. 5. The result of the evaluation test. The horizontal axis represents the grade of the power, and the vertical axis shows the measured velocity of the finger (pixels/msec).

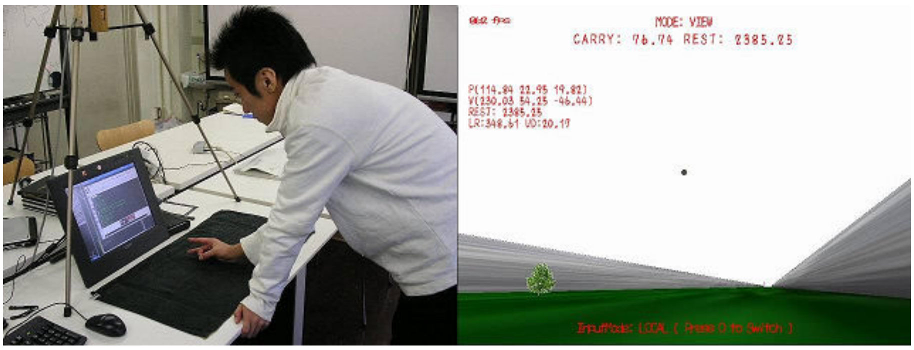

Fig. 6. A user playing the 3D golf game (left) and a screenshot of the game (right).

7 CONCLUSION

We introduced a finger flicking gesture recognition technique and the result from the evaluation test demonstrated that our system works sufficiently. At this time the estimation of the finger power is not accurate enough. We plan to evaluate the estimation closely.

References

1. Mehmet Gokturk and John L. Sibert, “An Analysis of the Index Finger as a Pointing Device”, ACM CHI99 Conference Companion, May 1999.
2. K. Oka, Y. Sato, and H. Koike, “Real-time Tracking of Multiple Fingertips and Gesture Recognition for Augmented Desk Interface Systems” in Proc. of 2002 IEEE International Conference on Automatic Face and Gesture Recognition (FG 2002), May 2002.