KIRIFUKI: INHALING AND EXHALING INTERACTION WITH VISUAL OBJECTS

Soichiro Iga, and Fumito Higuchi
Keio University, Keio Research Institute at SFC
{igaiga, wenren}@sfc.keio.ac.jp

Abstract: This paper proposes a novel system called "*Kirifuki*" which has an ability of operating a computer by breathing in and blowing out. The principle of this system is as follows; the computer screen is projected onto the physical desk, and when the user blows onto it, the open visual objects on the screen can be manipulated by breathing without using hands and any of physical pointing devices. In this paper, we describe our prototype system and demonstrate its application sub-systems which especially applied to entertainment fields.

Key words: Interaction technique, multimodal interaction, inhaling/exhaling user interface, input device for interactive system

1.   Introduction

Recently, in the field of interactive multimedia including video games and entertainment systems, progress has been made it possible for the people to interact with computer generated visual objects such as life-like characters or artistic graphics [6,7]. In order to bring out a potential ability of a broad range of visual representation on computing systems for games and entertainment, user-friendly interfaces and a wide range of variations in novel interaction techniques are greatly expected [2,3,4].

This paper proposes a novel interactive system called "*Kirifuki*" which has an ability of operating computer system by breathing in and blowing out. The computer screen is projected onto the physical desk, and when the user blows onto it, the open visual objects on the screen can be manipulated by breathing without using hands and any of physical pointing devices.

The system can be applied to wide variety of virtual/mixed reality systems or entertainment systems as the novel interaction device.
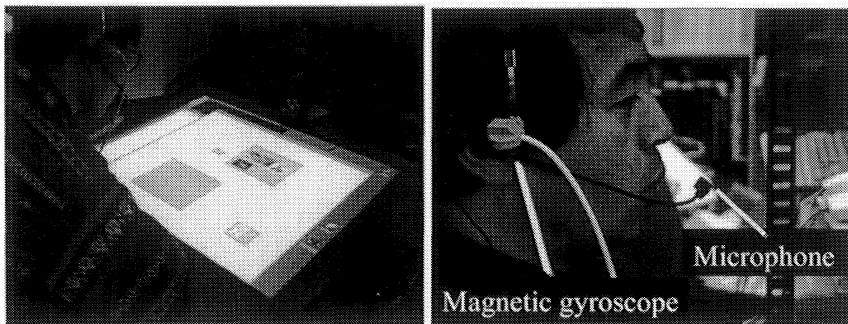
2.   Kirifuki System

We developed a novel prototype system called "*Kirifuki*" which has an ability of operating computer system by breathing in and blowing out. In this section, we describe implementation details of the system.

2.1   System Architecture

The system consists of a workstation(or a PC), a liquid crystal projector, a magnetic gyroscope, and a breath microphone switch. Software part is implemented in C and Basic.

Figure 1(a) shows the overview of the system. The computer screen is projected onto the physical desk by the liquid crystal projector. The user wears a head set device with a small breath microphone switch and a polhemus sensor. The polhemus sensor determines where the user is looking at. A small microphone detects the sound of the breathing and discriminates between inhaling and exhaling by means of the audio signals of the sound. The reason why we employed a microphone device approach is because a non-contact sensing of the user's breathing can be implemented.

(a) Overview of the system   (b) Device configuration
Figure 1. Kirifuki system

The user's head position is mapped to the manipulation of the pointing device, and the inhalation/exhalation will be the trigger to invoke the action of the graphical objects.

Figure 1(b) shows the head set of the breath microphone switch and the magnetic gyroscope. The breath microphone switch is configured to point

the user's lips. The magnetic gyroscope is located on the side of the user's head.

2.2   Pointing Device Manipulation

The user can manipulate a mouse pointer in the screen which is projected on to the physical desk by the movement of his or her head position. The magnetic gyroscope which is set up on the user's head recognizes the azimuth and elevation information. Horizontal movement of the mouse pointer is calculated from the azimuth information, and vertical movement is from the elevation information. Currently, the azimuth information is restricted to 90 degrees, and the elevation information is restricted to 20 degrees, not to exceed the user's field of vision. Mouse position is revised to move to the center of the screen when both azimuth and elevation information become 0 degrees.

2.3   Exhalation/Inhalation Recognition

The system recognize the exhalation and inhalation of the user's breath by the acoustic signals. The acoustic signals from the breath microphone switch that we used in the system are sampled at the rate of 16kHz with 8-bit resolution.

Generally, features of the acoustic signals like phases are difficult to be detected from the exhaled/inhaled acoustic signals. The randomness of the exhale signals are lower compared with the inhale signal, so we aimed at the randomness of the acoustic signals and we recognize the exhale and inhale signals by the local peak numbers of the acoustic signals. Local peak numbers are detected through the formula (1), where P stands for the acoustic power, t stands for time.

Each detected local peak numbers are counted by the buffer per unit (1024 byte) and the exhalation/inhalation is recognized by means of threshold.

$$\frac{dP}{dt} = 0, \frac{d^2P}{d^2t} < 0 \tag{1}$$

Thresholds of the local peak numbers are manually decided to separate three situations: normal, inhale, exhale. The *sign* function can be described as follows.

$$sign(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0. \end{cases} \tag{2}$$

Utilizing this function under the condition where the local peak numbers of normal, exhale, inhale are respectively described as N, E, I, and the candidate of the threshold is q:

$$F_1(q) = \sum_i \{sign(q - E(i))\}$$

$$+ \sum_i \{sign(I(i) - q)\}$$

$$+ \sum_i \{sign(N(i) - q)\}, \qquad (3)$$

$$F_2(q) = \sum_i \{sign(q - E(i))\}$$

$$+ \sum_i \{sign(q - I(i))\}$$

$$+ \sum_i \{sign(N(i) - q)\}. \qquad (4)$$

Data number F which satisfies the conditions can be counted by the formula described above. Thresholds P_1 and P_2 can be determined by finding the maximum q.

$$P_1 = P while_q^{max} F_1(q) = F_1(P) \qquad (5)$$

$$P_2 = P while_q^{max} F_2(q) = F_2(P). \qquad (6)$$

2.4   Recognition Trigger

From the nature of microphone switch, the system might recognize the user's utterance as the user is manipulating the system. To avoid manipulations which the user is not intended to do so, we combined the movement of the user's head with the trigger of the recognition procedure. Magnetic gyroscope detects the vertical height of the user's head movement and when the user bent forward to control objects on the physical desk top, which means that the height of the user's head become less than a threshold, the system starts triggering for the recognition of the microphone switch.

3.   Application

In this section, we provide some example applications which can be performed using our current prototype system.

3.1   Operating Visual Objects

The system can be applied to the manipulation of visual objects (*Figure 2*). The example here demonstrates the operation of graphical user interface objects such as moving and resizing objects and cutting and pasting objects. GUI objects such as a root window, windows, icons can be manipulated in the current example. For example, when the user exhales to the root window, the window objects near the mouse pointer would be diffused. On the contrary, the windows would be assembled when the user inhales.
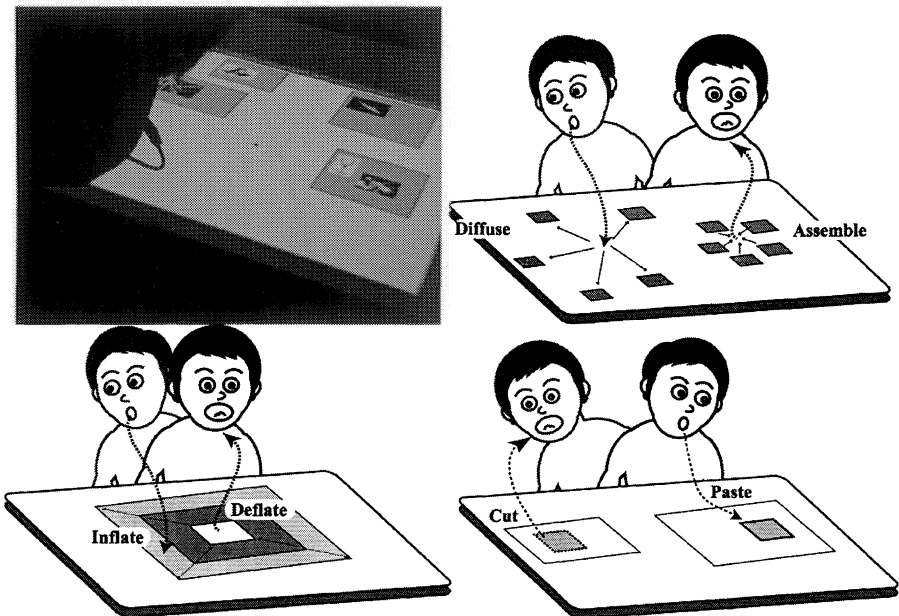

Figure 2. Manipulation of the GUI objects: Blowing windows by exhalation and gathering windows by inhalation

3.2   Drawing Picture

The system can be applied to draw a picture (*Figure 3*). The user can draw pictures like spraying by exhaling to the physical desktop. The user can erase his or her drawn picture by inhalation. Figure also shows the example pieces which has been performed by several users.
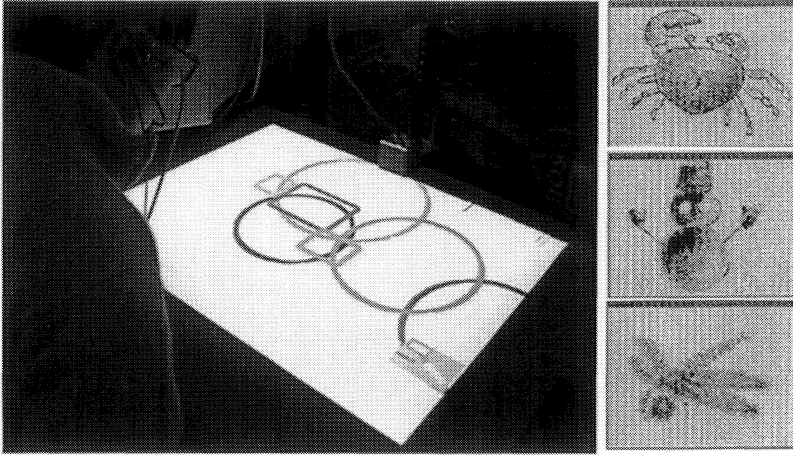
Figure 3. Picture drawing application and example pieces.

3.3   **Manipulating Interactive Movies**

Interactive movies can be manipulated by our system. Example application shown in *Figure 4* realizes the control of an interactive movie by breathing. For example when the user inhales to the system, animation image of coffee in a cup would decrease (the user would feel like drinking coffee). Inversely, when the user exhales to the system, coffee in the cup would increase.

Figure 4. Manipulating interactive movies: For example, images of coffee in the cup or images of a human face thumb through according to the breath operation.

3.4 Video Games

Our approach can be utilized as the input device for controlling video games. We implemented a simple video game called "Party time" which the user plays the game by both pointing device and breath input device (*Figure 5*). Capability of the exhalation and inhalation to be reacted reversible operation has some advantages in manipulating characters which performs reversible actions. For example, you can drink juice when you perform inhale action on a glass of juice and the glass would be filled up with juice when you exhale on a glass.
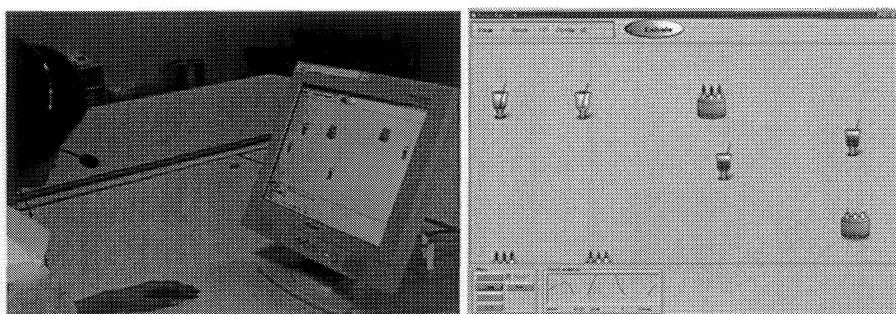

Figure 5. Example video game application in use.

4. Related Work and Discussion

Notion of breath switch or pneumatic switch is not new in the field of assistive technology for aged and disabled people [1]. We applied breath input technologies not only for an assistive support device but also as a general input device so that it could expand the variation of the application field of virtual and augmented reality domain. We also proposed a novel user interface technique to manipulate visual objects which fits to exhalation / inhalation input method.

Mouthesizer system controls the electric musical instruments by the shape of the mouth [5]. The user can manipulate digital musical equipments according to the shape of the mouth which is recognized by a CCD camera attached to the user's head. Although this approach may not be robust enough in the natural physical environment, a recognition accuracy of our current breath input can be improved by combining computer vision technique.

In our current prototype system, the user have to wear magnetic gyroscope on his or her head. However, the user who has some handicap in

upper half of the body would have some difficulty in operation pointing device. In the future system enhancement, techniques like gaze input [8] can extend the limits of our current pointing device manipulation.

5.   Conclusion

We proposed the novel system called "*Kirifuki*" which has an ability of operating computer system by breathing in and blowing out. We applied this technique for manipulating computer generated graphical objects such as GUI objects, picture drawings, interactive movies, and games. We believe that our approach can be applied to wide variety of virtual reality systems, games and entertainment systems as the novel interaction device.

References

[1] Adaptive Technologies web page [on-line], http:// www.adaptive-technologies.com/Products/MnSwitch.html

[2] Harrison, B.L., Fishkin, K.P., Gujar, A. ,Mochon, C., and Want, R., Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces, In *Proc. of CHI'98*, pp.17-24, ACM, 1998.

[3] Iga, S., Approximate Interaction: User interface for the ambiguous world, In *Proc. of SCI'99/ISAS'99*, Vol.2, pp.252-257, 1999.

[4] Iga, S., Blow Off Your Window Screens: Toward Ubiquitous Inhaling and Exhaling Interaction, In *Ubicomp2001 Informal Companion Proceedings*, Technical report:GIT-GVU-TR-01-17, pp.1-2, 2001.

[5] Lyons, M.J., Tetsutani, N., Facing the Music: A Facial Action Controlled Musical Interface, In *Proc. of CHI 2001*, pp. 309-310, ACM, 2001.

[6] Sparacino, F., Larson, K., MacNeil, R., Davenport, G., Pentland, A., Technologies and methods for interactive exhibit design: from wireless object and body tracking to wearable computers, In *Proc. of ICHIM 99*, pp.22-26, 1999.

[7] Tosa, N., Nakatsu, R., Life-like Communication Agent - Emotion Sensing Character 'MIC' and Feeling Session Character 'MuUSE', In *Proc. of the International Conference on Multimedia Computing and Systems*, pp.12-19, 1996.

[8] Zhai, S., Morimoto, C., Ihde, S., Manual And Gaze Input Cascaded(MAGIC) Pointing, In *Proc. of CHI'99*, pp.246-253, ACM, 1999.