Contents-Aware Gesture Interaction Using Wearable Motion Sensor

Tsukasa Ike
Toshiba Corporation
1, Komukai-Toshiba-cho, Saiwai-ku,
Kawasaki 212-8582, Japan
tsukasa.ike@toshiba.co.jp

Toshiaki Nakasu
Toshiba Corporation
1, Komukai-Toshiba-cho, Saiwai-ku,
Kawasaki 212-8582, Japan
toshiaki.nakasu@toshiba.co.jp

Yasunobu Yamauchi
Toshiba Corporation
1, Komukai-Toshiba-cho, Saiwai-ku,
Kawasaki 212-8582, Japan
yasunobu.yamauchi@toshiba.co.jp

Abstract
Gesture interaction has become a major role as intuitive control of remote devices. Motion-based hand gesture recognition using a wearable motion sensor equipped on the wrist-band helps decreasing recognition errors compared with that of video-based recognition systems. However, the user interaction is often interrupted even in the low-error conditions because an inappropriate gesture recognition mode is applied on the system or user can’t find out how to handle it using gestures. We propose a novel gesture interaction technique using visual attention to suggest user an appropriate gesture on the condition of selectable contents on the screen. We applied this interaction method on the content navigation interface for the TV and found it essential to realize natural and intuitive gesture interactions.

Author Keywords
Gesture interaction; Visual attention; Wearable sensor

ACM Classification Keywords
H.5.2 User Interfaces

Introduction
Gesture interaction has been developed for intuitive operation without any remote control devices. In recent years, it is available not only on consumer products [1]...
but also on professional use such as surgical operation [2]. Natural and intuitive motion as well as precise recognition of gesture is important for the comfortable gesture operation. Gesture recognition using motion sensors wore on user’s body helps decreasing recognition errors compared with video-based recognition. Wearable interaction technique is developed for a long time [3], and is becoming major [4] with the growth of wearable devices. However, gesture interaction is still not comfortable enough because it is difficult to define gestures intuitive for various users. Our challenge is to develop a novel gesture interaction for intuitive remote control. We focused on contents style and defined proper gesture for each style. Moreover, we introduced the visual attention, which reminds users the available gesture easily without learning.

**Concept**

The purpose of our work is to develop a gesture interaction which users can perform easily. For the purpose, we defined simple hand behaviors as gestures, and related each of them to the operation adequately from the viewpoint of contents style. Moreover, we introduced visual attentions which help users perform gestures without learning.

Figure 1 shows the defined gestures for interaction. Pointing means the horizontal and vertical motion with pointing toward the screen. Tapping means tapping one’s fingers. Flipping means moving one’s hand rapidly in horizontal direction. Rotation means rotating one’s wrist clockwise or counterclockwise. Waving means moving one’s hand rapidly from side to side.

**Figure 2:** Contents style and visual assistance.

Figure 2 shows the gestures and the visual attentions related to the contents style. We focused on the number of contents for defining gestures. In tile menu style, users can select all contents directly by pointing, so it is intuitive. In scrolling list style, many contents are not initially on the screen, so easy way of scrolling is required. Rotation is a good solution because it is able to express not only the direction but the scalar value for adjusting scroll speed, which helps users perform easily. In full screen style, easier way of changing contents is desirable. Flipping is one of the simplest gesture and adequate for the purpose.

Applying adequate gesture based on contents style makes the interaction easier, however, it confuses users how to do for manipulating contents. To solve this problem, we introduced the visual attention at a time to show users how to do. Visual attention is designed for making users recall the motion.
corresponding to the available gesture, so users can perform easily without learning gestures.

**Gesture Recognition**

Our system recognizes user’s gesture using motion data obtained from an inertial measurement unit (IMU) worn on the user’s wrist as shown in Figure 3.

*Pointing*

When the user points to the screen and tap one's finger, the cursor position is set to the center of the screen and our system begins recognizing gesture. The hand pose is calculated from motion data and stored as an initial pose. The differences of vertical and horizontal angle from the initial pose are calculated and are converted to the X and the Y coordinate of the cursor.

*Tapping*

Tapping detection is performed by analyzing Z-axis of acceleration. To reduce computational cost, we use a simple detection method. We calculate the difference between the sensor value and its average, and detect the local maximum peak as shown in Figure 4. The peak which confirms the equation below is detected as tapping:

\[
|A(t_{\text{MAX}})| > A_{T1}, \text{ and }  \\
|A(t)| < A_{T2} \text{ when } |t-t_{\text{MAX}}| > \text{dt},
\]

where \( A(t) \) means the calculated value in timestamp \( t \) and \( A_{T\text{MAX}} \) means the timestamp of the peak.

*Rotation*

When the cursor moves onto the scrolling list, the rotation angle is stored as an initial angle. During the cursor is on the list, the difference of angle from the initial angle is calculated. The scroll speed is controlled based on the angle.

**Flipping and Waving**

Hand flipping and waving is detected by analyzing hand motion in horizontal direction. Figure 5 shows the trail of the hand flipping motion to the left. We detect turn-rounds and evaluate length and velocity of each section. When the velocity is over the predefined threshold, the sequence is detected as flipping. Its direction is defined by the direction of the longest section and the other sections are ignored as unconscious motion. When turn-rounds are detected continuously at short interval, the sequence is detected as waving.

**Visual Attention**

Figure 6 shows screenshots of the content navigation interface for TV which the proposed interaction method is applied. Root menu is initially appeared on the screen and each tile is related to the functions such as TV program viewer, recommended program list, calendar,
picture viewer, and so on. Users can click these tiles by pointing and tapping gesture.

The recommended program list (Figure 6(b)) appears on the screen when the user clicks the corresponding tile of the root menu. The pane in the right is scrolling list, which users can operate using rotation and tapping. When the user moves the pointer cursor onto the list, visual attention of rotary controller appears on screen to notify user visually that rotation motion is available as shown in Figure 7. The visual attention also gives feedback of rotation angle to user so that the user can operate scrolling list intuitively.

The user can also call TV program viewer (Figure 6(c)) from the root menu. A TV program is displayed as a full screen contents, and corresponding visual attention is appeared at a time (Figure 8(a)). It notifies user visually that flipping motion is available. When the user performs flipping, it gives feedback to user so that the user can change contents by flipping intuitively (Figure 8(b)). When the user lowers the hand, it disappears automatically.

References

Implementation
We developed a prototype demonstration system as shown in Figure 9. The gesture recognition application and the content navigation interface application run on a tablet computer (Intel® Core™ i5-3317U CPU @ 1.70GHz). A 47-inch LCD monitor is connected to the computer and the UI application is displayed on the monitor. The user wore the wristband which a 9-axis inertial sensor unit “x-IMU” is attached. The computer receives measurement data stream from the sensor unit via Bluetooth™ wireless connection. The gesture recognition application recognizes user’s motion and controls the interface application based on the recognition result.

Conclusion
We propose a novel gesture interaction method. Intuitive gestures are defined based on the contents style. In addition, we provide users the visual attention which makes users recall gestures. It helps users perform gestures easily without learning.