Long-range Touch Gesture Interface for Smart TV

DoHyung Kim, Woo-han Yun, Jaeyeon Lee, Jaehong Kim, and Joochan Sohn
Electronics and Telecommunications Research Institute
Robot/Cognitive System Research Department
Daejeon, KOREA
{dhkim008, yochin, leejy}@etri.re.kr

Abstract- This paper proposes a long-range touch interface for interacting with various types of contents and services provided through Smart TV. The proposed interface provides the same functionality as the touch interface adopted on common smart phones by recognizing bare hand gestures of the user at a long distance, about three meters. To verify and test the functionality, usefulness, and performance of the proposed system, we built a demonstration system of a smart phone connected with a TV. We could verify that most of the smart phone operations can be accomplished by the proposed interface. We expect that the proposed interface can be promising natural user interface for Smart TV.

I. INTRODUCTION

Smart TV is the phrase used to describe the current trend of integration of the internet and Web 2.0 features into modern television sets, as well as the technological convergence between computers and these television sets. The technology that enables Smart TVs is not only incorporated into television sets, but also devices such as set-top boxes, Blu-ray players, game consoles, and other companion devices. These devices allow viewers to search and find videos, movies, photos and other content on the web, on a local cable TV channel, on a satellite TV channel, or stored on a local hard drive [1].

As traditional remote controls are not suitable for interacting with various types of contents and services provided through Smart TVs, new user interfaces are strongly required. Some products as remote controls for Smart TVs are recently introduced as shown in fig. 1 [2, 3].

Also, many attempts to control Smart TVs only using viewers’ body gestures without any handheld devices have been carried out widely since Microsoft introduced “Xbox Kinect”, a product from Xbox that uses spatial gestures for interaction instead of a game controller [4].

Since natural user interface (NUI) based on users’ hand and body gestures are very intuitive and users don’t have to keep some devices on their hands, it becomes increasingly popular.

With this trend, this paper proposes a long-range touch interface for interacting with Smart TV. The proposed interface provides the same functionality as the touch interface adopted on common smart phones by recognizing bare hand gestures of the user at a long distance, about three meters. Fig. 2 shows the overview of the proposed gesture interface.

II. THE PROPOSED INTERFACE

The proposed interface is developed by using 3D Depth camera of PrimeSense as shown in fig. 3 [5] and of course, other products such as Kinect and ASUS Xtion can be used. We implement our gesture recognition engine by using only 3D depth information from the camera. We don’t utilize any open libraries for helping to recognize hand gestures such as OpenNI/NITE[6], because we finally plan to load our gesture recognition engine on an android operated set-top box with low computational power.

Fig. 4 shows the overall flow of the gesture recognition system for the proposed interface for Smart TV.
The 3D depth image and RGB color image captured from the 3D depth camera are fed into the gesture recognition system. Those images correspond with each other, which are 640x480 pixel-sized.

A face detection module firstly runs in the RGB color image to find the position of a user. Our face detector is designed to detect faces as small as 20x20 pixels. The adopted face detection algorithm is implemented based on the AdaBoost procedure [7], which uses a modified census transform (MCT) to alleviate inference of illumination shown in fig. 5 [8]. MCT converts the pixel values into one of 511 patterns in a 3x3 neighborhood. These patterns represent the local spatial information of edges, junctions, and line segments. The transformed structure is robust in terms of illumination variation.

The proposed system then starts to recognize meaningful hand gestures by analyzing the change of the 3D hand position. We define three gestures to provide the same functionality as the touch interface on smart phones, which are clicking, dragging, and swiping. As shown in fig. 8, when a user pushes his/her hand forward, and then immediately pull the hand backward, the system recognizes the hand movement as a click gesture. A swipe gesture is a short movement in a specific direction either up, down, left, or right, followed by the hand resting. The gesture recognizer utilizes information of the hand movement including a period time of movement, a velocity of movement, a moving angle, and so on. The recognizer also adopts some stabilization methods to increase the recognition rate and lower false alarm rate.

If a certain hand gesture is recognized, the system sends a relevant command message to a GUI.

Fig. 4. The overall flow of the gesture recognition system.

Fig. 5. Example patterns of MCT (left) and example of illumination invariance of MCT (right).

Fig. 6. Virtual touch panel in front of the user.

Fig. 7. The 3D depth image (left) and the extracted hand blob (right).

Fig. 8. The click gesture (left) and swipe gesture (right).
III. DEMONSTRATIONS

As mentioned before, our goal is providing the same touch interface as smart phones at a long distance. To verify and test the functionality and performance of the proposed system, we built a demonstration system of a smart phone connected with a TV as shown in Fig. 9.

The interlocking device controller intercepts the screen of the smart phone and displays that on the 55 inches TV screen through HDMI. The hand movement and recognized gestures are interpreted as mouse events and are transferred to the smart phone in the device controller like touch commands through RS-232C. The used smart phone is a Samsung GALAXY S (SCH-M110S).

We could easily control most of the smart phone applications such as a map, a web browser, a calculator, and so on by the proposed interface. As showing the interaction with a smart phone at a long range, we could prove that the proposed touch interface is useful enough to interact with devices operated by touch interface. In the web browser, it was also possible to use the screen keyboard for keyword input, which demonstrates the high accuracy of pointing. Fig. 10 shows that the device controller which includes a smart phone inside and the user interacting with a mobile YouTube application on a smart phone through the demonstration system.

Fig. 9. The demonstration system of a smart phone connected with a TV.

Fig. 10. The device controller having a GALAXY S inside (left) and the user interacting with a smart phone through the demonstration system (right).

IV. CONCLUSIONS

This paper proposed a long-range touch interface for interacting with Smart TV. The proposed interface provided the same functionality as the touch interface adopted on common smart phones by recognizing bare hand gestures of the user at a long distance. To verify and test the functionality and performance of the proposed system, we built a demonstration system of a smart phone connected with a TV. We could verify that most of the smart phone operations can be accomplished by the proposed interface. We are planning to develop some semantic gestures based on the hand shape recognition for more intuitive and convenient interaction.

ACKNOWLEDGMENT

This work was supported by the ETRI R&D Program of KCC(Korea Communications Commission), Korea [11921-03001, "Development of Beyond Smart TV Technology"] and the IT R&D program of MKE & KEIT [KI001813, “Development of HRI Solutions and Core Chipsets for u-Robot”].

REFERENCES