Design and Evaluation of a Hand-held Device for Recognizing Mid-air Hand Gestures

Kyeongeun Seo† · Hyeonjoong Cho††

ABSTRACT

We propose AirPincher, a handheld pointing device for recognizing delicate mid-air hand gestures to control a remote display. AirPincher is designed to overcome disadvantages of the two kinds of existing hand gesture-aware techniques such as glove-based and vision-based. The glove-based techniques cause cumbersomeness of wearing gloves every time and the vision-based techniques incur performance dependence on distance between a user and a remote display. AirPincher allows a user to hold the device in one hand and to generate several delicate finger gestures. The gestures are captured by several sensors proximately embedded into AirPincher. These features help AirPincher avoid the aforementioned disadvantages of the existing techniques. We experimentally find an efficient size of the virtual input space and evaluate two types of pointing interfaces with AirPincher for a remote display. Our experiments suggest appropriate configurations to use the proposed device.

Keywords : Remote Control, Mid-air Pointing Devices, Hand Gesture Recognition

1. Introduction

A large amount of effort has focused on replacing the traditional button-based interfaces with hand gesture-aware devices to control remote displays more intuitively and interactively. The hand gesture recognition approaches are divided into two categories, i.e., glove-based and vision-based[1]. The glove-based techniques utilize sensors embedded into gloves to read hand and finger gestures. The close proximity between hands and sensors enables recognizing various delicate hand poses and movements. However, it is cumbersome to wear gloves whenever to use[2] and sharing gloves with others could cause concern about hygiene. On the other hand, vision-based techniques utilize cameras which free users from aforementioned inconvenience. Especially, several depth cameras recently commercialized, e.g., Kinect, provide three-dimensional visualization of objects, which allows barehanded control for remote displays. A drawback of
these approaches is that the resolution performance of distinguishing small gestures depends on the distance between a camera and a user. In order to alleviate those drawbacks of both approaches for utilizing hand gestures as interfacing terminology, we consider using a new form factor, named AirPincher, in combination with Kinect. AirPincher here is a handheld device that is a slight modification of the previous apparatus in [3]. A handheld device embedded with a low-cost Infrared (IR) camera, IR emitters, and low-cost motion sensors. The proposed device, named AirPincher, allows a user to hold the device in one hand and to generate several finger gestures, e.g., pinching, rotating, etc., with a user’s thumb, index and middle fingers of the holding hand. The close proximity between AirPincher and a user hand makes its resolution performance independent from the distance between a user and a remote display and allows for accommodating several delicate hand gestures. In addition, the handheld devices are one of the common form factors for remote control and thus, we believe that it incurs less hesitation for use than glove-based devices.

We apply AirPincher for controlling a cursor on a remote display. Cockburn et al. classified air pointing interfaces into three categories, such as raycasting, 2D plane, and 3D volume [4]. They found that 2D plane interface where the positional movement of a user hand in an imaginary mid-air 2D plane is projected to the remote display outperforms the other air pointing interfaces in terms of speed and accuracy. Their observation motivated us to consider two types of air pointing interfaces with AirPincher based on 2D plane.

We also assume that AirPincher is used together with Kinect located by a remote display. AirPincher here is responsible of detecting delicate hand gestures while Kinect monitors the positional transition of AirPincher.

In this paper, we experimentally find the comfortable size of the imaginary mid-air 2D plane for AirPincher. Then, we evaluate two types of air pointing interfaces with AirPincher based on 2D plane to find out their tradeoffs over various conditions. Our experiments suggest appropriate configurations to use the proposed device.

2. Related Works

While various mid-air hand gestures have been studied as moving and selecting commands, the pinch gesture is found to have a number of attractive features [5-6]. Our hand is able to quickly bring a thumb and a forefinger together and apart and additionally, the tactile feedback from our pinch fingers makes us ensure whether the pinch is performed or not. It enables precise and stable interaction [7]. The pinch gestures have been used by various gesture command systems [5, 8]. For example, Disits was a wrist-worn device for recovering the full 3D pose of a user hand. Kim et al. introduced a use case that a user initiates zooming with a mid-air pinch gesture of a hand wearing Disits.

Cockburn et al. investigated three different air pointing techniques for spatial target acquisition, i.e., raycasting, 2D plane, and 3D volume [4]. They experimentally showed that raycasting relying on small angular movement of a hand was fast but inaccurate. 3D volume was found to be expressive but slow and inaccurate. On the other hand, 2D plane was found to be appropriate as the air pointing interface in speed and accuracy.

Casiez et al. investigated control of display gain (CD gain) that affects the performance of pointing tasks. CD gain is defined as the ratio of the pointer velocity to device velocity. If CD gain is 1, the display pointer moves at exactly the same speed as the control device does. In this paper, we test different CD gains with AirPincher in order to search for an appropriate value for our purpose.

3. AirPincher

Fig. 1A and Fig. 1B show our prototype of AirPincher and the posture of a user’s holding hand. The device is incorporated with a single low-cost IR camera, an array of IR emitters on both sides of the camera, and a low-cost motion sensor consisting of accelerometers, gyroscopes, and magnetometers. The IR camera and IR emitters are used to detect the pinch gesture of a user’s holding hand.