Virtual Bridge: AR-Based Mobile Interaction for Easy Multimedia Control of Remote Home Devices

DongJin Eun, Taik Heon Rhee, Seonghoon Kang, Minsuk Choi, Sangil Lee, and Hark-Joon Kim

Samsung Electronics Co., Ltd. 416 Maetan3, Yeongtong, Suwon, Gyeonggi 443-742, Republic of Korea
{dj.eun,taikheon.rhee,shoon.kang,minsuk1.choi, lignas.lee,harkjoon.kim}@samsung.com

Abstract. This paper proposes an interaction for controlling multimedia contents of remote devices using a mobile device based on AR technology. Using a real-time object recognition method, home devices detected by the camera of a mobile device are displayed on the camera preview screen along with the thumbnails of their own multimedia contents around the recognized positions. A user may drag a multimedia content which he or she wants to play, and drop it onto another target home device which he or she wants to play the content through. The user study showed that the proposed interaction expects higher usability since once a home device is registered with its device name when registering its image shown on the mobile camera for the object recognition, this matching process is no longer necessary when a user controls the device through the mobile device.

Keywords: Mobile Interaction, Augmented Reality, Mobile AR, DLNA, Contents Sharing, Real-Time Object Recognition.

1 Introduction

Recently, most of homes have several home devices holding multimedia contents, such as televisions, digital video recorders, photo frames or PCs, and the interaction issues between home devices have taken on great important. Although standardized technologies for transferring multimedia contents among home devices, including DLNA (Digital Living Network Alliance) [1], are spreading rapidly, the user interfaces to find other home devices or multimedia contents on them are still complex so that users not only need many steps to control the home devices. Moreover, users feel a lot of trouble in matching a device itself with its name or identifier displayed on the other device which is to connect with it.

Most of the commercial DLNA solutions use a list of home devices with text and icon, when a user needs to connect a server device with a display device. Figure 1 shows the example screenshot of selecting remote devices in commercial solutions, *Samsung AllShare* [2] and *Braintransfer DLNA remote* [3]. As shown in Figure 1, a user should always memorize the correct name of the target device whenever he or she wants to control remote devices. In addition, such a series of selection steps is

C. Stephanidis (Ed.): Posters, Part II, HCII 2011, CCIS 174, pp. 102–106, 2011.

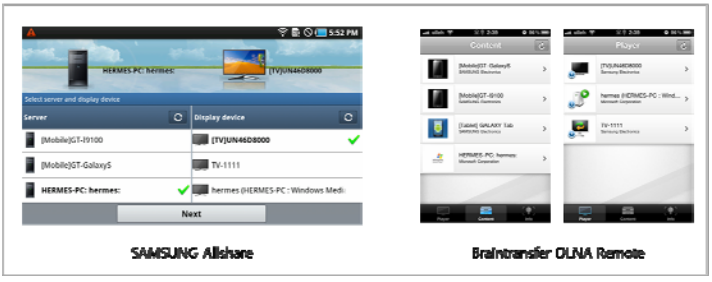

Fig. 1. Example screenshots of selecting remote devices in commercial solutions

required in the commercial solutions and it makes the user interface look complex and difficult to use.

This paper proposes an interaction for controlling multimedia contents of remote devices using a mobile device based on AR technology. When a user focuses on one or more home devices through the camera of a mobile device, they are detected by a real-time object recognition method on the camera preview screen of the mobile device. The position of each recognized device is displayed on the touch-enabled camera preview screen, along with the thumbnails of its own multimedia contents around the recognized position. Then the user drags the thumbnail of a multimedia content which he or she wants to play, and drops it onto the recognized position of another target device which he or she wants to play the content through.

2 *Virtual Bridge*: Interaction Design

Since Tani *et al.* [4] proposed the concept of using live video images to make users interact with a real-world object located at a distance, a number of studies have been utilized the metaphor in various usage scenarios. Boring *et al.* [5] mentioned a number of related works which adopted the "interaction through video" metaphor, while they proposed *Touch Projector*, a mobile interaction through video. In their work, the *Touch Projector* allows users to manipulate content shown on displays at a distance by touching the image of the content taken by a mobile camera.

Our proposed interaction called the *Virtual Bridge* focuses on virtually bridging two real-world home devices by a mobile controller. In other words, our interaction connects a *Digital Media Server* (DMS), which stores content, with a *Digital Media Renderer* (DMR), which play content from a DMS, by the touch interaction on live video images taken by the camera of a mobile *Digital Media Controller* (DMC), which finds the content on DMS and plays it on DMR. The definitions of DMS, DMR and DMC are introduced on the DLNA specification [1]. The name of the "*Virtual Bridge*" reflects the concept that a user's bridging interaction of two devices in the real world through the virtual space on a live video image. Figure 2 shows the walkthrough of the concept of the *Virtual Bridge*. Note that the mobile controller (DMC) itself may take on a role of content server (DMS) and directly connect its content to the remote device (DMR) for rendering.

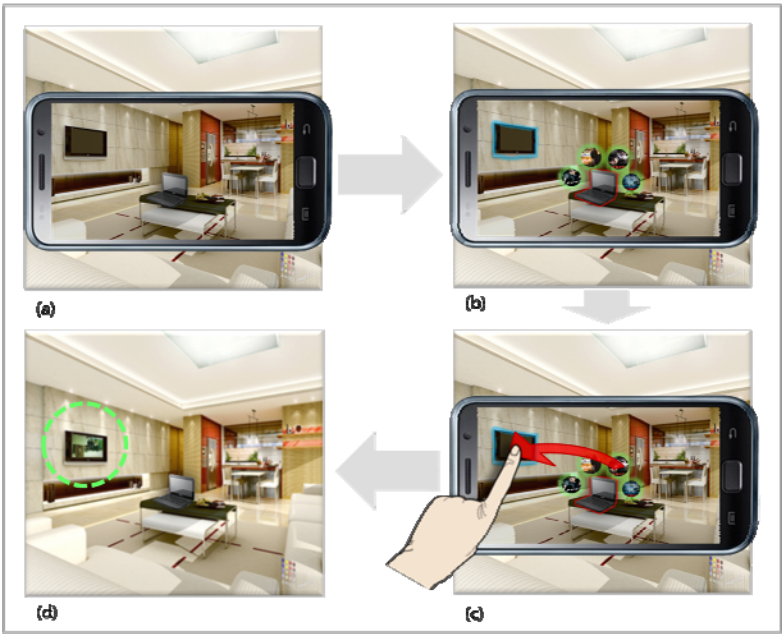

Fig. 2. User scenario flow. (a) A user focuses on one or more home devices through the camera of a mobile device. (b) The position of each recognized device is displayed on the touch-enabled camera preview screen, along with the thumbnails of its own multimedia contents around the recognized position. (c) The user drags the thumbnail and drops it onto the position of another target device. (d) The content corresponding to the thumbnail is played through the target device.

The Virtual Bridge interaction needs pre-registration procedure of the home devices as shown in Figure 3. Once a home device is registered with its device name when registering its image shown on the mobile camera for the object recognition,

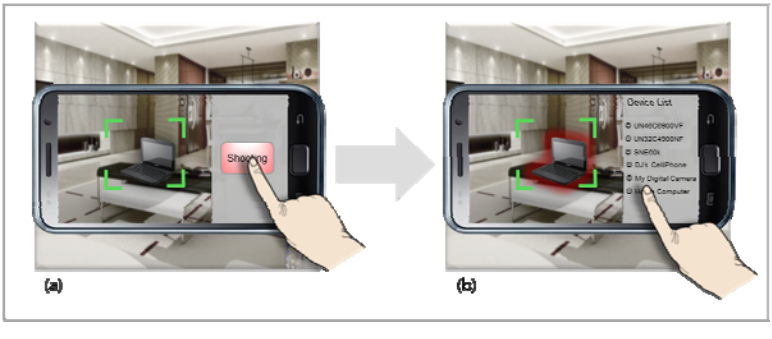

Fig. 3. Pre-registration procedures. (a) A user takes a snapshot of the target device to register and extract features of the image. (b) The snapshot with the extract features is matched with the device name for the real-time object recognition afterwards.

this matching process is no longer necessary when a user controls the device through the mobile device.

3 Implementation and Evaluation

To compare usability of the proposed interaction with *Samsung AllShare* [2], a conventional approach applied in commercial products, we implemented a prototype of *Virtual Bridge* interaction controller. The controller (DMC) prototype was implemented as an android application on a *Samsung Galaxy Tab* (SHW-M180S). A variant of *SIFT* algorithm [6] has been implemented for the real-time recognition of home devices on live video images. The DLNA functions in *Windows Media Player 7* application on a notebook PC and Samsung Galaxy-S, a smartphone with *Samsung AllShare* were used for the content server (DMS). Also, a television supporting *Samsung AllShare* was used for the content renderer (DMR). All the devices were assumed to be connected to a Wi-Fi network in advance. Figure 4 shows examples of the implemented prototype in use.

Fig. 4. Prototype implementation of the *Virtual Bridge*. (a) Searching content through a server (smartphone). (b) Moving content to a renderer (TV).

To measure user-satisfaction with the proposed interaction compared with the commercial solution, a questionnaire based on the IBM computer usability satisfaction questionnaire [7] was given to 10 participants, who were aged between 25 and 36 with a mean of 29.7 years. The questionnaire consisted of 7 items in total, using a 5-point Likert scale from strongly agree (5) to strongly disagree (1). Figure 5 presents a summary of the results.

In general, the proposed interaction gained better results than the conventional approach. Especially, the participants felt that it was more enjoyable, more intuitive, and easier to learn than the conventional approach. From the questionnaire results with comments of the participants, we may infer that the results are due to the direct matching between the device name listed in a controller and its image shown in the live video image.

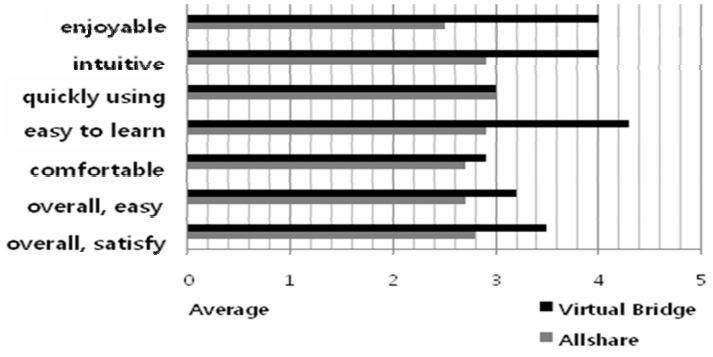

Fig. 5. Average user feedback regarding the interaction techniques

4 Conclusion and Future Work

This paper proposes an interaction for controlling multimedia contents of remote devices using a mobile device based on AR technology. While the conventional approach always requires matching a remote device and its name, the proposed interaction makes a user connect content from a server to a render intuitively on the camera preview screen of a mobile controller.

The proposed interaction relies on performance of the real-time object recognition algorithm to detect home devices on live video images. The currently implemented algorithm only detects a home device when its image is taken in the similar direction with the one taken at the registration time. In the future, we plan to overcome such a limitation by improving the recognition algorithm robust on direction variations, as well as introducing proximity sensors or indoor position determination technologies. Meanwhile, we also plan to find various user scenarios beyond the multimedia control, based on interaction among home devices.

References

1. Digital Living Network Alliance, http://www.dlna.org
2. SAMSUNG AllShare, http://www.samsung.com/au/tv/internet-tv.html
3. DLNA Remote, Braintransfer Inc., http://braintransfer.de/en/services/mobilesolutions/dlnaremote
4. Tani, M., Yamaashi, K., Tanikoshi, K., Futakawa, M., Tanifuji, S.: Object-oriented video: interaction with real-world objects through live video. In: Proc. CHI 1992, Monterey, CA, USA, pp. 2287–2295 (1992)
5. Boring, S., Baur, D., Butz, A., Gustafson, S., Baudisch, P.: Touch Projector: Mobile Interaction through Video. In: Proc. CHI 2010, Atlanta, GA, USA, pp. 2287–2295 (2010)
6. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Computer Vision 60(2), 91–110 (2004)
7. Lewis, J.R.: IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. Int. J. Hum.-Comput. Interact. 7(2), 57–78 (1995)