CyPhy-UI: Cyber-Physical User Interaction Paradigm to Control Networked Appliances with Augmented Reality

Through Design, Implementation, and Evaluation of EVANS (Embodied Visualization with Augmented-Reality for Networked Systems)

Kenya Sato, Naoya Sakamoto, Shinya Mihara, and Hideki Shimada
Department of Information Systems Design
Doshisha University
Kyoto, 610-0321 Japan
e-mail: ksato@mail.doshisha.ac.jp, pooh708s@gmail.com, dezeckco@gmail.com, hshimada@mail.doshisha.ac.jp

***Abstract*—Many kinds of networked home appliances, which are connected by standardized control functions, have recently appeared and continue to increase. Because a general infrared remote control is for single-way communication from a remote control to a specific appliance, but not to receive signals from an appliance to a remote control, it is impossible to gather an appliance's information with a remote control. Meanwhile, unlike a general infrared remote control, it is difficult to control a specific appliance because users can simultaneously operate all of their appliances with a WiFi controller. In this paper, to control networked appliances with a smart phone or tablet computer as a WiFi controller, we propose a new interface paradigm called a cyber-physical user interaction that creates virtual (cyber) space, sends commands, and receives responses from networked (physical) appliances through space with augmented-reality (AR) technology. With the paradigm, which enables interconnectivity among appliances from various vendors, it is possible users with uniform and intuitive operation of home appliances. In addition, we implement and evaluate an Embodied Visualization with Augmented-Reality for Networked Systems (EVANS) that controls a system of home appliances and sensor devices through a cyber-physical user interaction (CyPhy-UI) paradigm by a web camera to retrieve information from real world environments and touch-screen display to show AR visualization and user interaction components to retrieve user input.**

Keywords-appliance; control; cyber-physical; user interface; augmented reality

I. INTRODUCTION

The appearance and popularity in homes of such audio/video appliances, as TVs, DVD players, hard-disk recorders, audio receivers, and digital speakers, continue to increase. Recently these appliances can be connected to each other through home networks with standardized network functions (e.g., DLNA [1]) to transmit audio/video stream data. In the near future, we expect that household electrical appliances, air conditioners, floor lamps, electric curtains, and sensor devices (for temperature, humidity, or illumination) will also be connected to home networks to enhance convenience in *smart homes* [2][3]. In general, since the controls over home appliances are used from exclusive infrared remote controls, as the number of home appliances increases, distinction by remote controls becomes more complicated. Because a general infrared remote control is for single-way communication (to send signals from a remote control to an appliance, but not to receive them from an appliance to a remote control), it is impossible gather an appliance's information with a remote control. Meanwhile, unlike a general infrared remote control, it is difficult to control a specific appliance because users can simultaneously operate all of their appliances with a WiFi controller.

In this paper, to control networked appliances with a smart phone or a tablet computer as a WiFi controller, we propose a new user interaction scheme called a cyber-physical user interaction (CyPhy-UI) paradigm in which virtual (cyber) space is created to send commands and receive responses from networked (physical) appliances through space with augmented-reality (AR) technology.

II. CYBER-PHYSICAL USER INTERACTION PARADIGM

A. Categorization of Control Types for Appliances

To analyze the control methods for networked appliances, we need to categorize user interaction models among a user and appliances. Suppose several appliances around a user, who chooses one of them and controls it by remote control. As described Figure 1, we categorized the situation into the following six types:

1) Type 1—Direct command and direct response: A user directly sends commands to Appliance A, and directly receives responses from it. For example, TVs are included in this type. When a user changes the channel with a remote control, she pushes a switch on it. The channel information is shown on the TV display, and she can know that follow the channel has been changed.

2) Type 2—Direct command and no response: A user directly sends commands to Appliance A without a response from it. For example, an air conditioner (without any display function on the appliance) is included in this type. When a user tries to turn down the temperature with a remote control, he pushes a switch on it. However, there is no direct response from the air conditioner. Although he may eventually feel that the temperature has been lowered, he

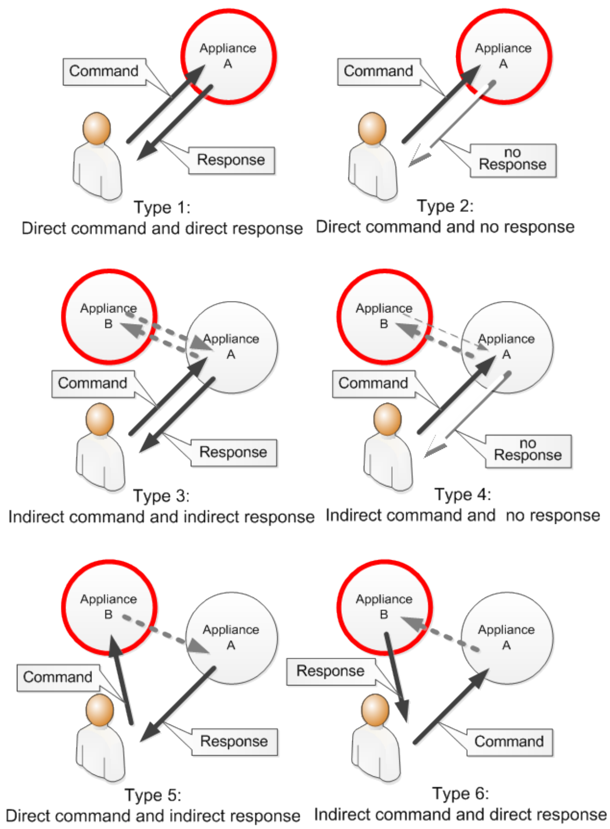

Figure 1. Six types of interactions among users and appliances.

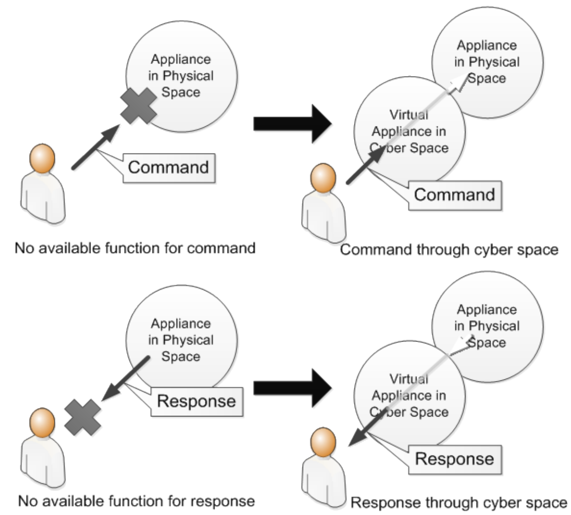

Figure 2. Request and response model with/without virtual appliances in cyber space.

cannot immediately check the air conditioner's status, because there is no response from Appliance A.

3) Type 3—Indirect command and no response: A user controls Appliance B through Appliance A without a response from Appliances A or B. For example, a speaker (without any display function) and an audio receiver are included in this type. When a user increases the volume of a speaker connected to an audio receiver with a remote control, he pushes a switch on it. The speaker becomes louder, but not the audio receiver. In addition, there is no response from the speaker or the audio receiver. A user may not directly understand which appliance to control.

4) Type 4—Indirect command and indirect response: A user controls Appliance B through Appliance A, and the response of Appliance B from Appliance A. For example, a TV set with separate speakers is included in this type. When a user increases a speaker connected to the TV set with its remote control, he pushes a switch on it. The speaker becomes louder, and the TV shows the volume status on its display. The user can realize that understand the volume has increased. However, suppose that six speakers (5.1 channels) are connected to the TV, and the user wants to decrease their volume. Identifying a certain one may be difficult without speaker identification on the TV display.

5) Type 5—Direct command and indirect response: A user directly controls Appliance B, and its response from Appliance A. For example, a DVD recorder is connected to a TV, and its status is displayed on the TV's display. When a user wants to identify the track of a DVD recorder connected to the TV with DVD recorder's remote control, she pushes a switch on the remote control. The DVD's tracks are displayed, and the TV shows their status. She can identify specific tracks. Even if both a DVD recorder and a hard disk video recorder are connected to her TV, she can differentiate between them and choose which to control.

6) Type 6—Indirect command and direct response: A user controls Appliance B through Appliance A and direct responses from Appliance B. For example, a speaker with volume indicators and an audio receiver are included in this type. When a user increases the volume of a speaker connected to an audio receiver with a remote control, she pushes a switch on it. The speaker becomes louder, and she can check the indicators on the speaker. However, it is difficult to identify a certain appliance among several only with the information on the audio receiver.

Among these six types, Type 1 direct control and direct response is the ideal case to control appliances. With it, users can directly control a target appliance and realize their request. However, in real situations, an appliance may have no display function and/or may have no direct control method with a remote control.

In addition, users might get not only an appliance's status information but also audio/video contents from the appliance as response messages without sending a request to it.

B. Approach

To send a request from a user to an appliance without functions to receive commands, and/or to receive responses

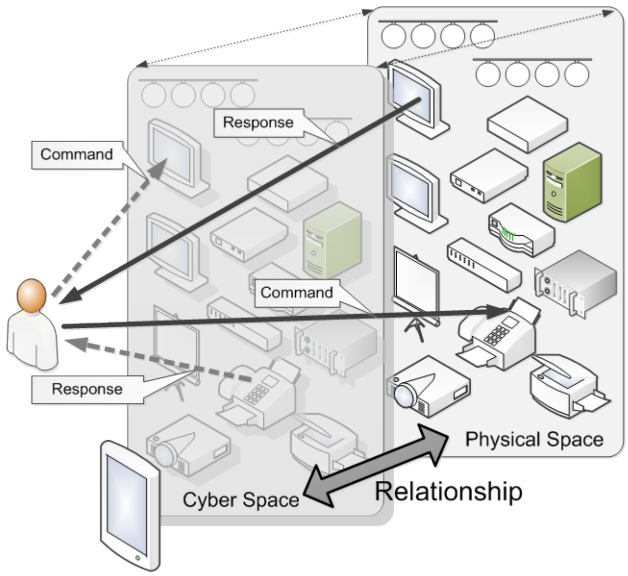

Figure 3. Relationship between appliances in cyber and physical space.

from an appliance to a user without functions to send a response, our approach creates a new paradigm. In it, a virtual appliance is defined in cyber space, and users can send a command to or receive responses from a virtual appliance in cyber space. Our cyber-physical user interaction (CyPhy-UI) paradigm is shown in Figure 2.

We apply such portable devices, as smart-phones and tablet computers with augmented reality technology to realize our CyPhy-UI paradigm. Portable devices have a touch display and a camera installed on its back, as well as a WiFi network function. With the camera, the appliances in the physical space are shown as camera images on the display, which means appliances in cyber space. The appliances in the physical space must be connected to the portable device. As described in Figure 3, a menu to control an appliance is shown on the portable device's display, and users can send a command by touching the menu. In addition, they can identify the appliance's responses, which are shown on the display.

III. RELATED WORKS

A. Tangible User Interfaces

Graphical user interfaces makes a fundamental distinction between input devices, such as the keyboard and mouse as control and output devices like monitors for the synthesis of visual representations. Tangible User Interfaces [4][5] proposed by Hiroshi Ishii et al., couple physical representations (e.g., spatially manipulable physical objects) with digital representation (e.g., graphics and audio), yielding interactive systems that are computationally mediated, but generally not identifiable as computers in itself. The design and selection of appropriate physical representations is an important aspect of tangible user interface design.

In our CyPhy-UI paradigm, virtual representations are defined in cyber space and users can send a command to or receive responses from physical objects through cyber space without using physical representations. In addition, virtual representations in CyPhy-UI paradigm are real camera images of physical objects, and users can easily identify physical objects and their functions.

B. Augmented Reality Systems

Many research projects make invisible information related to objects visible using augmented reality technology, such as Google's Project Glass [6], which is an outdoor mobile augmented reality street view application, MARA [7], which is a sensor-based augmented reality system for mobile imaging device. uMegane [8] which is a visualization system of sensor data with AR technology, easily acquire sensor data for users who are unfamiliar with sensor technology. Extate [9] is a visualization system of a wireless network with AR Technology that enables users to acquire such network status as packet data and network type. Sekai Camera[10] is a popular smart-phone application to enable users to view AR information about subjects of scenery. Our goal is not only to acquire appliance information but also to control sensor nodes and networks. Thus our target system is different from existing researches.

C. LED Visual Markers for Augmented Reality

In present AR technology, an image marker [11] is necessary to identify the AR graphics to display on the camera screen. However, some issues have presented, such as marker size, lighting environments, and distance from a controller to appliances. In related works using a LED for the AR marker, Visual Computer Communication (VCC) marker [12] was proposed. VCC markers use 16 LEDs as one AR marker, so cameras can receive the increased lighting information of LEDs, and display the AR graphics. However when operating in appliances, a VCC requires far more insertion space into appliances than our proposed LED visual marker. Since we only need the networked home appliance's ID information and the LED location displayed through the camera, few data are necessary. For this reason, VCC markers are not suitable for operating appliances. As described below, we propose an LED visual marker and a network home appliance operating system that intuitively and easily operates complex network appliances in dark, bright, or distant home environments.

IV. REALIZATION OF THE NEW PARADIGM

A. Design of EVANS

We designed and implemented EVANS, or Embodied Visualization with Augmented-Reality (AR) for Networked Systems, which is a system to provide users with a uniform and intuitive interface of home appliances and/or sensor networks using a CyPhy-UI paradigm.

AR technology can provide images generated by overlapping virtual information on real environmental

 ISBN: 978-1-61208-250-9

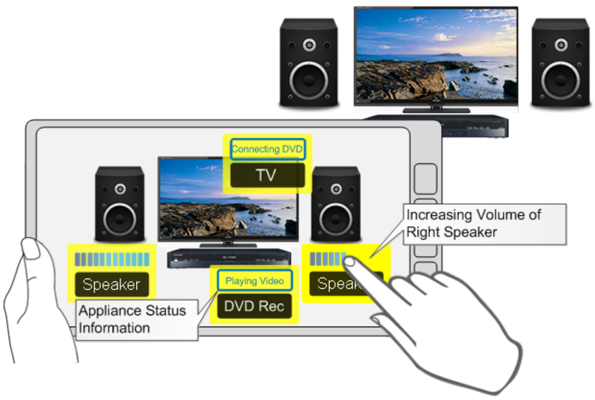

Figure 4. Using EVANS.

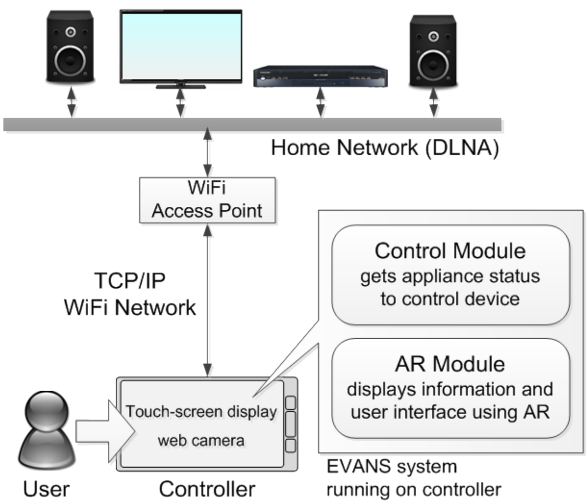

Figure 5. Implemented system architecture for EVANS.

Figure 6. Display image on a tablet computer for EVANS.

Figure 7. Volume control of right speaker with AR.

images captured by cameras. The virtual information in cyber space called an annotation includes the information associated with certain objects in real space. In general, AR technology uses an AR marker to detect the camera's position and orientation, which is two-dimensional code. In advance, the pattern data of AR markers are registered for the application, which can recognize the object in real space by tracking the AR markers with the camera.

For example, as shown in Figure 4, the user points the camera, which is installed on the back of the remote control (e.g., smart-phone or tablet computer) at the desired appliance, whose control interface for the specific appliance is shown on the controller's front display. To create annotations on the display, a specific AR marker is attached to each appliance. When the right speaker is chosen, the volume up/down control menu is shown as annotation information (Figure 5). The user then simply changes the volume. If a TV is chosen, appliances connected to it are shown; if a DVD recorder is chosen, the playback controls or content selection are shown on the controller's display. The user simply performs the operations. In addition, by providing a uniform user interaction, users can perform operations on any appliance connected to their home network.

Building upon standardized home networks that enable interconnectivity among appliances from various vendors, we aim to solve the appliance selection problem by providing a real world display augmented with appliance information and control interfaces.

B. Implementation of EVANS for AV Appliances

The system architecture for implementing EVANS to control and monitor AV appliances is shown in Figure 5. We used a web camera to retrieve information of the appliances in the real space, and a touch screen display to show AR visualizations and user interaction components and to retrieve user input. The entire system consists of AR and control modules.

The AR module displays appliance information and user interaction components after the web camera is pointed at an appliance. It also gathers input from the user and presents the intended control data to the control module. For example, the user may point the web camera at a speaker with a certain identifying marker. The control module communicates with

 ISBN: 978-1-61208-250-9

TABLE I. DMS RESPONSE TIME FOR ORIGINAL DLAN FUNCTION AND EVANS

Evaluation item	*Original DLNA function*	*EVANS*
Response time (msec)	0.33	0.39

the speaker to retrieve its volume value. The AR module places the user interaction components (in this case a volume value display and volume control buttons) on the marker. When the user operates the buttons, the AR module gives the control data to the control module, which sends the commands to the appliance, and the AR module updates the AR display.

This module communicates with such home appliances as TVs, HDD recorders and speakers and actually controls them based on user input and retrieves information about them. The control module uses the Digital Living Network Alliance (DLNA) to communicate with them.

We implemented EVANS using a note PC and a tablet computer. The note PC for the home appliances and the tablet computer were connected to the network, and to run DLNA Media Server applications (TVersity [13] and DiXiM [14]) that act as a DLNA Media Server (DMS) to provide video content. External speakers were attached to the note PC. The node PC and the speakers have identifying markers for use in AR. The following is the implementation setup:

- AR display: ARToolKit [15]
- DLNA controls: Cyberlink for C++ [16]
- Multimedia display: Simple DirectMedia Layer (SDL) [17]
- Graphics library: Freetype Graphic Library (FTGL)

Figure 6 shows the state of the AR display when the search button is pressed. It confirms that the appliance information was successfully discovered on the network; the names of the DMS applications and the details of the stored content for the note PC are overlaid on the marker by AR. Intuitive user interaction components are also displayed for the speakers attached to the note PC; the volume values and control buttons are displayed by AR. Figure 7 shows the AR displays when the volume is set to 50. We confirmed that user input against the system actually changed the volume of the speakers and that the AR display was updated accordingly.

To evaluate our implementation’s practicality, we measured the time before a response from the DMS is received after issuing a search for it. As a control value, we also measured the same response time for existing DLNA appliances. Table 1 shows the average results for 20 individual measurements. Since our implemented EVANS causes no major delay compared to the existing system, it is practical.

C. Implementation of EVANS for Sensor Networks

Sensor networks are currently used in such fields as home electronics, energy management, and security because they can immediately detect an event or a situation and automatically control an actuator. However, managing each sensor node connected to wireless sensor networks is difficult because sensor node status and wireless network topology are invisible. EVANS can display sensor data and such network information as the link status, the packet data, and the traffic in the sensor network as AR information in cyber space on the display.

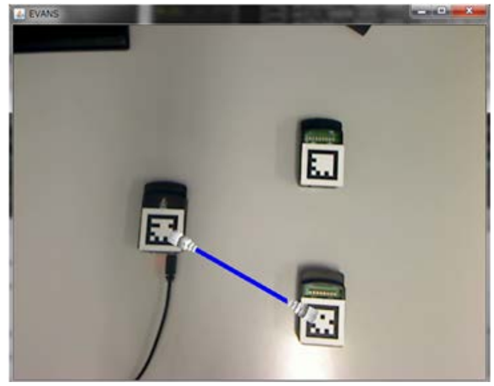

Figure 8. Displayed network connection information among sensor nodes.

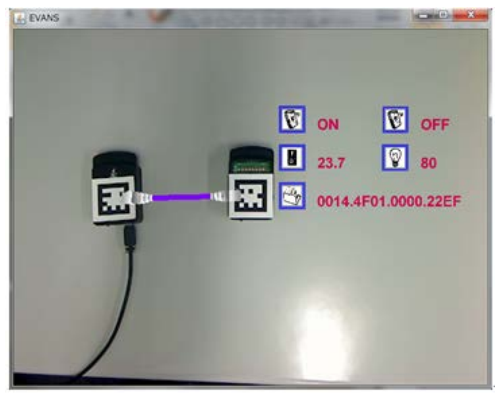

Figure 9. Displayed sensor information included on nodes.

We implemented EVANS [18] for a sensor network using Sun Small Programmable Object (SunSPOT) [19] as the sensor node and a Java version ARToolKit (NyARToolkit [20]) to generate AR images. SunSPOT, which is a wireless sensor network device, can measure temperature and illuminance and is also equipped with a push button switch.

Figures 8 and 9 show link status images of our prototype system captured by the sensor node’s camera. In real space images, we cannot directly see the connection and the sensor information between sensor nodes, but this system allows us to directly acquire such information through AR images.

This system has two operation methods. One shows the resource information. When users tap a sensor node on the touch screen, resource information is displayed. Another

 ISBN: 978-1-61208-250-9

controls the connection between sensor nodes. When users drag and drop between sensor nodes, the switches of both nodes are toggled. In the installed program, if the switch of each node is turned on, these nodes communicate with other sensor nodes. If the switch of each node is turned off, the connection of these nodes is interrupted. When the connection is interrupted, the virtual link cable of the AR annotation image disappears.

V. DISCUSSIONS

A. Current Issues with AR Solution

Presently, infrared remote controllers are the most popular way of controlling home appliances. Although some full-feature remote controllers allow users to control multiple appliances, in most cases each appliance comes with its own remote controller. These remote controllers send out infrared signals to control their appliances, but it is generally not possible to conversely retrieve appliance status information at the controller. This is partly compensated by displaying status information on the appliance itself if the target appliance is a television or is connected to a television, but most other appliances that are not equipped with a display only provide limited status information (most commonly using LEDs). Also, such additional detailed information as instruction manuals cannot be used directly on remote controllers.

Most people have at least one mobile phone and carry it when they are at homes [21], and smart-phones are also becoming more and more popular. Recent home appliances, especially multimedia appliances, can be easily connected to a local network. Based on these circumstances, traditional remote controllers are unnecessary if these smart phones could control home appliances. Furthermore, through smart-phone display, users can learn detailed information about their appliances and control them by touch-screen, providing further convenience.

ARToolKit, a popular AR framework also used in our above implemented system, uses image markers or graphic patterns that are usually printed on paper to identify objects and their placements. Figure 10 shows an example image marker used in ARToolKit. This framework poses a number of problems when applied to home appliance control. To address these problems, we proposed LED visible markers [21], where LEDs equipped on appliances blink at a fast rate and are used as AR identification markers (Figure 11). The following sections discuss the issues with image markers and the solutions that can be provided by LED visible markers.

B. Visual Attractiveness

Image markers tend to be rather large because the camera must recognize their patterns from a certain distance. They also need to be attached in places where they are always visible. These constraints make the appliances visually unattractive. LED visual markers, on the other hand, are hardly noticeable. They are also practical because the LEDs of most home appliances are lit up even when their power is turned off.

Figure 10. Image marker for ARToolKit..

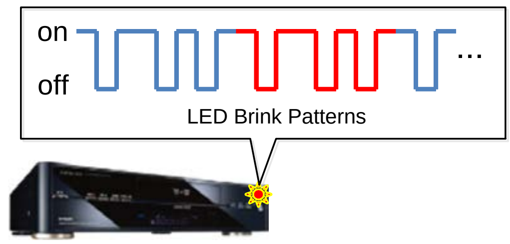

Figure 11. LED visible marker.

C. Recognition

Because home appliances are almost always used indoors, they must be controlled even when the room is dark. Image markers basically require that the room is lit up for them to be recognizable, but the LED visual markers are self-luminous and recognizable in complete darkness. Note that this does not imply that they are less recognizable in brighter rooms.

D. Dynamic Identification Changes

Because image markers are generally printed on paper, changing an appliance ID is cumbersome. Another maker must be printed to replace the old one. On the other hand, LED Visible Markers allow dynamic changes of appliance IDs. They are also more useful in networked setups because one could possibly use an appliance's MAC or IP address as part of the appliance ID.

E. Directional Recognition

One benefit of image markers is that they are two-dimensional, which lets the camera compute the direction and distance of the marker. LED visual markers, on the other hand, are zero-dimensional points, and do not yield such information. As a countermeasure, three LEDs of different colors can be placed in a triangular shape. For applications in appliance controls, since user interaction components merely need to be placed vertically and never in other directions depending on the marker position, this may not be an issue at all.

 ISBN: 978-1-61208-250-9

F. *Control of Legacy Appliances*

Some legacy appliances may not have network control functions and can be controlled only with an infrared remote control. In this case, we use an image marker stuck to an appliance instead of a battery-powered LED visual marker to identify it. We also developed remote control software on a smart-phone through iRemocon [23] (universal infrared remote control device connected to PC) to send control signal to an appliance.

VI. CONCLUSION AND FUTURE WORK

The popularity of home appliances that can interconnect with other home appliances through networks continues to increase. The operation and function of such appliances are complex since they can share contents and data with other network home appliances. However, it remains difficult for users to identify network home appliances, since their locations are not installed and cannot be displayed easily when operating them. Because of these problems, more obvious and intuitive operations are needed for controlling networked home appliances.

We proposed a new user interaction paradigm to control networked appliances. In our cyber-physical user interaction (CyPhy-UI) paradigm, cyber space is created and we can send commands to and receive information from physical appliances through it with augmented reality technology. With our paradigm, which enables interconnectivity among appliances from various vendors, we can provide users with uniform and intuitive operation of home appliances.

We also implemented an Embodied Visualization with Augmented Reality for Networked Systems (EVANS), which controls a system of home appliances and sensor devices through a CyPhy-UI paradigm, using a web camera to retrieve information from real world environments, and a touch screen display to show AR visualizations and user interaction components and to retrieve user input. We also evaluated our system using this method based on response times, and conclude that since our system does not introduce noticeable response delays, it is practical. We also discussed the issues about current AR image markers when applied to home appliance control, and discussed the practicality of LED visual markers to solve them.

Future research will implement the LED visual marker method and evaluate such characteristics as recognition rates. We also will produce a more general and practical implementation using handheld devices with limited computational resources such as smart phones.

ACKNOWLEDGMENT

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (KAKENHI) Grant Number 21500084.

REFERENCES

[1] DLNA: Digital Living Network Alliance, DLNA Networked Device Interoperability Guidelines, expanded, 2006.

[2] Scott Davidoff, Min Kyung Lee, Charles Yiu, John Zimmerman, and Anind K. Dey, Principles of Smart Home Control, Ubicomp 2006 LNCS Vol.4206, pp.19-34, 2006.

[3] Richard Harper, FromSmart Home to Connected Home, The Connected Home: The Future of Domestic Life, Part 1, pp.3-18, 2011.

[4] Hiroshi Ishii and Brygg Ullmer, Tangible bits: towards seamless interfaces between people, bits and atoms, Proceedings of the SIGCHI conference on Human factors n computing systems, pp.234-241, 1997.

[5] Brygg Ullmer and Hiroshi Ishii, Emerging frameworks for tangible user interfaces, IBM Systems Journal, Vol.39, Issue 3/4, pp.915-931, 2000.

[6] Google, Project Glass, Thougts, designs, and stories, http://www.google.com/+projectglass/ [Jan. 1, 2013].

[7] Markus Kähäri and David J. Murphy, MARA-Sensor based Augmented Reality System for Mobile Imaging Device, 5th IEEE and ACM International Symposium on Mixed and Augmented Reality, 2006.

[8] Takuya Imaeda, Kazunori Takashio, and Hideyuki Tokuda, uMegane: Visualization System of Sensor Data using AR Technology, IPSJ Technical Report Vol.2008, No.66, pp.39-44, 2008.

[9] Takuya Takimoto, Namatame Naoya, Jin Nakazawa, Kazunori Takashio, and Hideyuki Tokuda, Extate: Visualizing Wireless Networks by using AR Technology, The 3rd Asia-Europe Workshop on Ubiquitous Computing, 2010.

[10] Sekai Camera Beyond Reality, http://sekaicamera.com/ [Jan. 1, 2013].

[11] Hirokazu Kato and Mark Billinghurst, Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, pp.85-94, 1999.

[12] Yasue Kishino, Masahiko Tsukamoto, Yutaka Sakane, and Shojiro Nishio, Realizing a Visual Marker Using LEDs for Wearable Computing Environment, Proceedings of IEEE International Conference on Distributed Computing Systems Workshops (IWSAWC2003), pp.314-319, 2003.

[13] TVersity, the Leading UPnP & DLNA Media Server, http://tversity.com/ [Jan. 1, 2013].

[14] DigiOn, DiXiM Digital TV, http://www.digion.com/pro/dxtv/ [Jan. 1, 2013].

[15] Hirokazu Kato, ARToolKit: Library for Vision-based Augmented Reality, IEIC Technical Repor, Vol.101, No.652, pp.222-232, 2002.

[16] QNX, Simple Directmedia Layer, http://www.libsdl.org/ [Jan. 1, 2013].

[17] Cybergarage, CyberLink for C++ , http://www.cybergarage.org/twiki/bin/view/Main/CyberLinkForCC/ [Jan. 1, 2013].

[18] Naoya Sakamoto, Hideki Shimada, and Kenya Sato, Design and Implementation of Sensor Network Device Control System with AR Technology, Advanced Engineering Forum, Vols.2/3, pp.131-134, 2011.

[19] Oracle Labs, Sun SPOT World Home, http://www.sunspotworld.com/ [Jan. 1, 2013].

[20] NyARToolkit project, http://nyatla.jp/nyartoolkit/wp/ [Jan. 1, 2013]..

[21] TechBed.com, Thigns To Do With A Mobile Phone, http://www.techbed.com/things-to-do-with-a-mobile-phone/ [Jan. 1, 2013]..

[22] Shinya Mihara, Akira Sakamoto, Hideki Shimada, and Kenya Sato, Augmented Reality Marker for Operating Home Appliances, Proceedings of the 2011 IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC2011), pp.372-377, 2011.

[23] iRemocon, http://i-remocon.com/ [Jan. 1, 2013]