ARGI: Augmented Reality for Gesture-based Interaction in Variable Smart Environments

Jonas Sorgalla, Jonas Fleck and Sabine Sachweh
Institute for Digital Transformation of Application and Living Domains,
University of Applied Sciences and Arts Dortmund, Otto-Hahn-Str. 23, Dortmund, Germany

Keywords: Augmented Reality, Assistive Technology, Gesture Interaction, Smart Environment, Human-Computer Interaction.

Abstract: Modern information- and communication technology holds the potential to foster the well-being and independent living of elderly. However, smart households which support older residents often are overwhelming in their interaction possibilities. Therefore, users demand a single and remote way to interact with their environment. This work presents such a way using gestures in free space to interact with virtual objects in an augmented reality to control a smart environment. For expandability and reliability the implementation of the approach relies on Eclipse SmartHome as a prevalent open source framework for home automation and the Microsoft HoloLens.

1 INTRODUCTION

The age group 60 years and older is predicted to increase from 901 million in 2015 to 1.4 billion people in 2030 which makes it the globally fastest growing age group (United Nations, 2015). This demographic shift makes it inevitable to develop new solutions in areas such as health care, housing or social protection and foster the well-being as well as independent living of elderly.

Modern Information and Communications Technology (ICT) plays a crucial role in the development of such solutions (Rashidi and Mihailidis, 2013). One domain, where ICT is applied, is smart living environments respectively smart homes (Ghaffarianhoseini et al., 2013). While smart homes are generally used to support Activities of Daily Living (ADLs)(Reisberg et al., 2001), scientific research particularly fosters the usage as assistive technology for an independent living of older and ill adults (Morris et al., 2013)(Ziefle et al., 2009).

However, the application of ICT for older adults holds several special accessibility and acceptability issues, which need to be addressed in particular (Leonardi et al., 2008). Especially in the context of smart living environments, interaction possibilities to configure and control devices can be overwhelming. Therefore, elderly demand a single way of remote control to interact with such environments (Koskela and Väänänen-Vainio-Mattila, 2004).

Therefore, several approaches like (Wobbrock et al., 2009) or (Kühnel et al., 2011) focus on gesture control and design to interact with a smart environment as a ubiquitous remote control. Such approaches mostly rely on camera input devices like Microsoft's Kinect[1] or mobile phones. We argue that such gesture approaches do not emphasize interaction feedback and easy learnability enough, although both have proven to be important for an adoption by elderly users (Zhang et al., 2009)(Mennicken et al., 2014). Additionally, such approaches are complex to expand because they tend to pair gestures and devices in a static way.

In this paper, we propose a novel approach to control a smart living environment fusing augmented reality (AR) with gesture recognition. We build upon prevalent solutions like Eclipse SmartHome as smart environment platform and Microsoft HoloLens as mixed reality smartglasses to provide easy extendability in a dynamic way. We developed the approach together with elderly participants in several participatory design sessions (Muller, 2003).

The remainder of this paper is structured as follows. Section 2 provides a brief introduction to our underlying participatory design methodology. Section 3 describes the general approach, while its prototype implementation is presented in the follo-

[1] https://developer.microsoft.com/en-us/windows/kinect

Sorgalla, J., Fleck, J. and Sachweh, S.
ARGI: Augmented Reality for Gesture-based Interaction in Variable Smart Environments.
DOI: 10.5220/0006621301020107
In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018) - Volume 2: HUCAPP*, pages 102-107
ISBN: 978-989-758-288-2

wing Section 4. We discuss related work in Section 5 and conclude in Section 6. Finally, we describe our plans for future work in Section 7.

2 METHODOLOGY

The initial motivation of our work emerged from a discussion with participants in a participatory design (Muller, 2003) workshop in the interdisciplinary research project QuartiersNETZ[2] (translates to network of neighborhoods). The project aims to enable an independent and self-determined life in old age. Among other measurements, e.g. the development of a social neighborhood platform or the promotion of volunteer teachers for technology, the project focuses on the development of low-threshold interaction possibilities for home environments. We use the participatory design methodology described in (Sorgalla et al., 2017) to take diversity and representativeness of the project's target group into account. The applied methodology leverages qualitative data from workshops, interviews, and neighborhood meetings as well as quantitative data from surveys and literature.

During the workshop we presented gesture-based human-computer interaction for a smart home. The 15 participants were encouraged to use a Kinect to dim and change the color of a Philips Hue lamp. The used gestures were simple and built based on the hand position recognition of the Kinect SDK. While the feedback was generally positive, our participants criticized missing feedback whether a gesture was recognized successful. Additionally, it was remarked that it may be a lot of effort to learn all interaction possibilities and corresponding gestures in a larger smart living environment. Based on the group discussion, we developed the idea of an environment which is controllable with a very reduced gesture set by using AR technology. In the following, we discussed our implementation process iteratively with participants in three additional meetings.

3 AUGMENTED REALITY IN SMART ENVIRONMENTS

AR describes an integrated view where virtual objects are blended into a real environment in real time (Milgram et al., 1994). While the idea of enriching a real environment with virtual objects goes back to the middle of the 19th century, first real consumer applications have only been available in recent years because of growing device capabilities. New commercial products with better display and tracking functionalities like Microsoft's HoloLens are prototypes towards a future in which vision is pervasively augmented with smart contact lenses or glasses (Perlin, 2016).

Our approach to integrating AR into smart living environments is depicted in Figure 1. We aim to augment the normal view with virtual interactive control elements to enhance interaction with the environment. Therefore we rely on a smart environment platform (component `a`) which has at least a twofold accessible interface. Once, to accomplish a variable configuration without the need to bind controls and devices in a static way, the platform needs an interface which provides meta information regarding the connected devices, i.e. the device type, name, and its capabilities (`Meta Information Interface`) for each device. Twice, the platform needs an interface to control such devices (`Command Interface`).

It is mandatory that the supported device types of the platform are closed which enables us to permanently assign a virtual control element to each device type. This is realized with a generator (component `b`) which dynamically generates visual identifiers (artifact `c`), e.g. QR code labels, which each hold meta information of a device as well as necessary information to send commands to the device through the command interface. Therefore, for the integration of new devices, it is simply needed to check for newly available devices through the meta information interface and generate the corresponding visual identifier rather than write static code. The visual identifiers need to be placed where the virtual control element of the real world device should appear.

Next, we use an AR capable head-mounted display which runs an application to visualize the control elements (component `d`) and is able to perform gesture recognition. The application is able to encode the previously generated labels including the device type and command interface information upon seeing. Based on the read data we (i) load a 3D model (artifact `e`) corresponding to the device type for each device in the view, (ii) locate each model close to the scanned visual identifier, and (iii) associate the command interface to certain parts of the 3D model, i.e. we determine which command should be sent when interacting with different parts of the model.

When the AR application recognizes an interaction with one of the 3D models, it sends a command to the associated device through the smart environment platform leveraging the device related command interface. Based on which part of the model was focused by the gesture, a different command is sent. For example, the 3d model control element of

[2] http://www.quartiersnetz.de

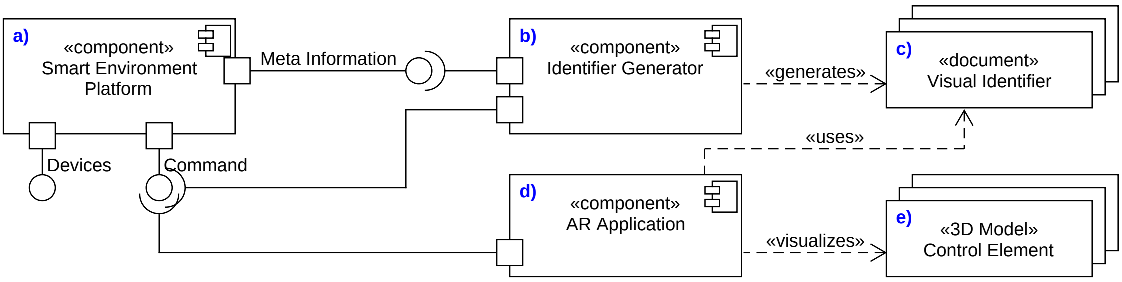

Figure 1: Component diagram visualizing the overall system.

a lamp could be a virtual light bulb with which can be interacted using a simple tap gesture to turn it on or off. Accordingly, when the AR application detects a tap gesture which turns the virtual bulb on, a command which turns the lamp in the real world on is sent to the device related command interface of the smart environment platform. For visual feedback, the 3d model could change color or carry a text label.

4 ARGI PROTOTYPE

Based upon our approach we implemented the ARGI[3] prototype. As AR device we rely on a Microsoft HoloLens because it already contains simple gesture recognition to interact with virtual objects. Overall ARGI comprises of two components: a JavaFX tool to generate QR codes (*Identifier Generator*), and a Windows Mixed Reality application (*AR Application*). To provide the mandatory environment interfaces we rely on the Eclipse SmartHome[4] (ESH) platform (*Smart Environment Plattform*). In the following, we introduce the underlying concepts of ESH for a better understanding of the smart home interfaces as well as present the JavaFX tool to generate the labels according to the interface information and, finally, the ARGI HoloLens application.

4.1 SmartHome Interface

ESH is a modular designed smart home framework by Eclipse. The ESH platform integrates devices with existing connectivity features. To integrate a device ESH needs a connector software which is called a *binding*. Conceptual, ESH dismembers each connected device based on its functionalities into *items*. The possible items are restricted to certain *item types* shown in Table 1 (Eclipse Foundation, 2017). For example, a Philips Hue lamp is connected through the Hue binding and can be dismembered as a color, dimmer and switch item. Each item can be controlled separately and manage different *command types* (cf. Table 1).

All integrated items and their meta information can be retrieved with a RESTful HTTP API. Additionally, the API provides the possibility to send commands encoded as simple text strings to items. Therefore, ESH fulfills the mandatory requirements for our approach.

4.2 Environment Configuration

Based upon an existing ESH instance, our implemented Java tool is able to create visual identifiers. The so called "SmartHome-Labeler" comes with a JavaFX GUI, which is depicted in Figure 2, and is able to fetch all integrated items from an ESH instance using its RESTful HTTP API. A user is able to select items from the fetched list and generate QR codes accordingly. A generated QR code encodes the following JSON string scheme:

```
{
  "name":"example_item",
  "link":"ip_address/rest/items/example_item",
  "label":"label_name",
  "type":"item_type"
}
```

Each code encapsulates the ESH internal item name and a human readable label name. Additionally, the code contains the item type and the direct link to the device API provided by ESH, which both are necessary for the AR application. For good usability, the QR codes are generated as PDF files whose format is matched to a label printer.

4.3 ARGI App

The Mixed Reality App for the HoloLens is implemented using the Mixed Reality Toolkit[5] which uses

[3] **A**ugmented **R**eality **G**esture **I**nteraction

[4] https://www.eclipse.org/smarthome

[5] https://github.com/Microsoft/MixedRealityToolkit-Unity

Table 1: Eclipse SmartHome item types and corresponding commands according to (Eclipse Foundation, 2017).

Item type	Description	Command Types
Color	Color information (RGB)	OnOff, IncreaseDecrease, Percent, HSB
Contact	Item storing status of e.g. door/window contacts	OpenClose
DateTime	Stores date and time	-
Dimmer	Item carrying a percentage value for dimmers	OnOff, IncreaseDecrease, Percent
Group	Item to nest other items / collect them in groups	-
Number	Stores values in number format	Decimal
Player	Allows to control players (e.g. audio players)	PlayPause, NextPrevious, RewindFastforward
Rollershutter	Typically used for blinds	UpDown, StopMove, Percent
String	Stores texts	String
Switch	Typically used for lights (on/off)	OnOff

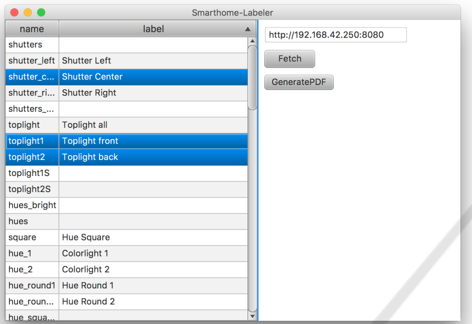

Figure 2: GUI of the SmartHome Labeler tool.

the game engine Unity[6]. As depicted in Figure 3 the app is able to scan the QR codes created with the SmartHome Labeler. In our implementation, the scan needs to be initiated with the speech command "scan".

Once a QR code is recognized, the app visualizes the corresponding 3d model according to the ESH item type. Currently, the prototype supports the `RollerShutter`, `Dimmer` and `Color` item types (cf. Table 1). The virtual control element models are shown in the Figures 4 and 5 from the perspective of the HoloLens. After their initial setup based upon the location of the QR code, the models can be moved freely with the built-in manipulation gesture of the HoloLens. Accordingly, the tap gesture enables the interaction with the model. For example, a tap on the red button depicted in Figure 4 sends a `StopMove.STOP` command to the roller shutter item using the RESTful HTTP API. Tapping the green arrows sends a `UpDown.UP` or `UpDown.DOWN` command. At last, ARGI only needs the QR codes for the first setup of the virtual control elements, after that, the HoloLens memorizes the location of every 3D model and reproduces it when starting the ARGI app.

Each of our 3D models gives visual feedback when an interaction occurs. For example, the bulb model for dimmer items changes its caption and the roller shutter controls have a visual push animation.

Figure 3: Scanning of a QR code using the ARGI app.

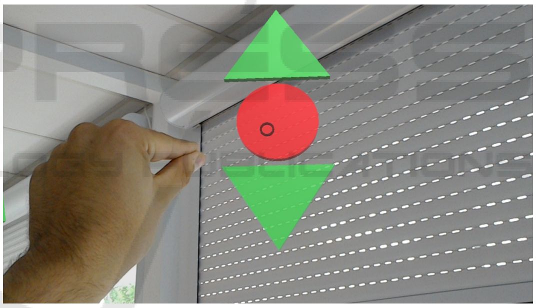

Figure 4: Control element for the roller shutter item type.

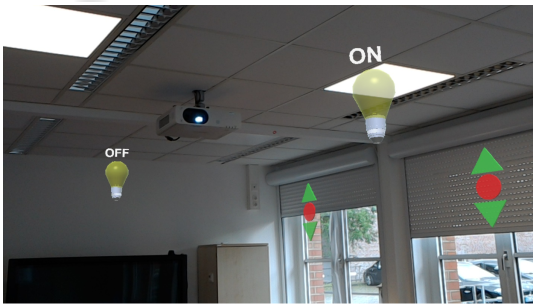

Figure 5: Room with multiple virtual control elements.

[6]https://unity3d.com

5 RELATED WORK

There are many other works which aim to implement a remote control for a smart living environment. For example, Seifried et al. presented the CRISTAL tabletop system to control devices in a living room with gestures (Seifried et al., 2009). While such systems provide a central remote control, they can not be used in free space but are bound to a certain location.

Other approaches like (Iqbal et al., 2016), (Kühnel et al., 2011) or (Budde et al., 2013) use motion sensing input devices like the Kinect camera to realize environment interaction with gestures in free space similar to our initial participatory workshop presentation. In comparison, our AR-empowered approach grants more possibilities to visualize the interaction capabilities of a device and visual feedback. Although, we admit that a user has to wear an AR device like the HoloLens to be able to use our system which can get uncomfortable the longer a user wears the device. However, we are sure that AR devices are shrinking in the future and our approach becomes more usable in everyday use.

The interaction with smart environments for a better assistance in old age is not among the traditional application domains of AR like medical visualization, maintenance and repair, or teaching and learning (Azuma, 1997)(Wu et al., 2013). Though, Kriesten at al. purposed an AR approach to interact with a smart environment using a mobile phone which is comparable to ours (Kriesten et al., 2010). However, our approach goes a step further because it enables free space gestures with the help of a head-mounted AR device rather than a mobile phone. Additionally, ARGI relies on Eclipse SmartHome as a proven and freely accessible smart environment solution which is beneficial to extendability and reliability.

6 CONCLUSION

In this paper we presented an approach to use AR to enhance the interaction in a smart environment. Based upon feedback in a participatory design workshop where we evaluated gesture-based interaction, we developed together with elderly participants the idea of an augmented smart environment which (i) gives feedback using the capabilities of 3D visualization, (ii) reduces the mandatory gestures, and (iii) is easy to configure. For an expandable yet reliable realization, our approach relies on the already existing ESH framework as a smart living environment platform with accessible interfaces to interact with connected devices and retrieve device meta information. According to our approach we implemented the ARGI prototype which comprises ESH as smart environment platform, the SmartHome-Labeler tool for producing visual identifiers to configure the environment for our AR application, and, finally, the ARGI application for Microsoft's HoloLens.

We have briefly presented and discussed the ARGI prototype in another participatory design workshop. While the feedback for the general approach was very positive, the participants stated the prototype character of our solution and pointed out in particular the uncomfortable fit of the HoloLens. However, based upon the feedback we are sure that if AR devices are able to shrink to the size of contact lenses or normal glasses, especially older adults are able to benefit from this development.

The ARGI HoloLens app[7] as well as the corresponding SmartHome Labeler[8] (Release 1.0a1) are accessible on GitHub. A tutorial how to setup openHAB2 which is the open source reference implementation for ESH can be found on the project website[9].

7 FUTURE WORK

In the future we plan to add 3D model support for the missing item types which will enable ARGI to fully represent every possible ESH-based smart living environment. Therefore, we would like to discuss and (further) develop the visual representations, e.g. the currently used light bulb or the roller shutter control, together with our participants in another participatory design workshop.

Finally, we are aware that ARGI currently lacks a proper usability evaluation besides the user feedback in our participatory design workshops. Therefore, we intent to present ARGI at multiple neighborhood meetings and gather quantitative data using the System Usability Scale (SUS) (Brooke et al., 1996) and in addition, perform guided interviews centering around the user experience.

ACKNOWLEDGEMENTS

This work has been partially funded by the German Federal Ministry of Education and Research (BMBF) under funding no. 02K12B061 as part of the QuartiersNETZ project.

[7]https://github.com/SeelabFhdo/ARGI-Holo
[8]https://github.com/SeelabFhdo/SmarthomeLabeler
[9]http://docs.openhab.org/tutorials

REFERENCES

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385.

Brooke, J. et al. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7.

Budde, M., Berning, M., Baumgärtner, C., Kinn, F., Kopf, T., Ochs, S., Reiche, F., Riedel, T., and Beigl, M. (2013). Point & control – interaction in smart environments: You only click twice. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 303–306, New York, NY, USA. ACM.

Eclipse Foundation (2017). Eclipse smarthome documentation. https://www.eclipse.org/smarthome/documentation.

Ghaffarianhoseini, A. H., Dahlan, N. D., Berardi, U., Ghaffarianhoseini, A., and Makaremi, N. (2013). The essence of future smart houses: From embedding ict to adapting to sustainability principles. *Renewable and Sustainable Energy Reviews*, 24:593 – 607.

Iqbal, M. A., Asrafuzzaman, S. K., Arifin, M. M., and Hossain, S. K. A. (2016). Smart home appliance control system for physically disabled people using kinect and x10. In *2016 5th International Conference on Informatics, Electronics and Vision (ICIEV)*. IEEE.

Koskela, T. and Väänänen-Vainio-Mattila, K. (2004). Evolution towards smart home environments: Empirical evaluation of three user interfaces. *Personal Ubiquitous Comput.*, 8(3-4):234–240.

Kriesten, B., Mertes, C., Tünnermann, R., and Hermann, T. (2010). Unobtrusively controlling and linking information and services in smart environments. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI 2010*. ACM Press.

Kühnel, C., Westermann, T., Hemmert, F., Kratz, S., Müller, A., and Möller, S. (2011). I'm home: Defining and evaluating a gesture set for smart-home control. *International Journal of Human-Computer Studies*, 69(11):693 – 704.

Leonardi, C., Mennecozzi, C., Not, E., Pianesi, F., and Zancanaro, M. (2008). Designing a familiar technology for elderly people. *Gerontechnology*, 7(2):151.

Mennicken, S., Vermeulen, J., and Huang, E. M. (2014). From today's augmented houses to tomorrow's smart homes. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM Press.

Milgram, P., Takemura, H., Utsumi, A., Kishino, F., et al. (1994). Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, pages 282–292.

Morris, M. E., Adair, B., Miller, K., Ozanne, E., Hansen, R., and et al. (2013). Smart-home technologies to assist older people to live well at home. *Journal of Aging Science*, 01(01).

Muller, M. J. (2003). Participatory design: the third space in hci. *Human-computer interaction: Development process*, 4235:165–185.

Perlin, K. (2016). Future reality: How emerging technologies will change language itself. *IEEE Computer Graphics and Applications*, 36(3):84–89.

Rashidi, P. and Mihailidis, A. (2013). A survey on ambient-assisted living tools for older adults. *IEEE Journal of Biomedical and Health Informatics*, 17(3):579–590.

Reisberg, B., Finkel, S., Overall, J., Schmidt-Gollas, N., Kanowski, S., Lehfeld, H., Hulla, F., Sclan, S. G., Wilms, H. U., Heininger, K., Hindmarch, I., Stemmler, M., Poon, L., Kluger, A., Cooler, C., Bergener, M., Hugonot-Diener, L., Robert, P. H., Antipolis, S., and Erzigkeit, H. (2001). The Alzheimer's disease activities of daily living international scale (ADL-IS). *Int Psychogeriatr*, 13(2):163–181.

Seifried, T., Haller, M., Scott, S. D., Perteneder, F., Rendl, C., Sakamoto, D., and Inami, M. (2009). Cristal: a collaborative home media and device controller based on a multi-touch display. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, pages 33–40. ACM.

Sorgalla, J., Schabsky, P., Sachweh, S., Grates, M., and Heite, E. (2017). Improving representativeness in participatory design processes with elderly. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '17, pages 2107–2114, New York, NY, USA. ACM.

United Nations (2015). World population ageing 2015. (ST/ESA/SER.A/390).

Wobbrock, J. O., Morris, M. R., and Wilson, A. D. (2009). User-defined gestures for surface computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1083–1092, New York, NY, USA. ACM.

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., and Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*, 62:41–49.

Zhang, B., Rau, P.-L. P., and Salvendy, G. (2009). Design and evaluation of smart home user interface: effects of age, tasks and intelligence level. *Behaviour & Information Technology*, 28(3):239–249.

Ziefle, M., Röcker, C., Kasugai, K., Klack, L., Jakobs, E.-M., Schmitz-Rode, T., Russell, P., and Borchers, J. (2009). ehealth–enhancing mobility with aging. In *Roots for the Future of Ambient Intelligence, Adjunct Proceedings of the Third European Conference on Ambient Intelligence*, pages 25–28.