Avatar Enriched User Interfaces for Older Adults

Christoper Mayer, Miroslav Sili,
Matthias Gira, Martin Morandell
AIT Institute of Technology GmbH
H & E Department, Biomedical Systems
Vienna, Austria
Email: christopher.mayer@ait.ac.at,
miroslav.sili@ait.ac.at, matthias.gira@ait.ac.at,
martin.morandell@ait.ac.at

Sascha Fagel, Andreas Hilbert
zoobe message entertainment GmbH
Berlin, Germany
Email: fagel@zoobe.com,
hilbert@zoobe.com

Christian Schüler, Ioan Cernei
weTouch e.U.
Vienna, Austria
Email: christian.schueler@wetouch.at,
ioan.cernei@wetouch.at

***Abstract*—Needs and wishes regarding the interaction with ICT solutions change over time and vary between older adults. They depend on the user's physical and mental capabilities and preferences. In particular usability, accessibility, as well as the freedom of choice concerning the interaction with such systems are the crucial points for acceptability and applicability. The aim of the AALuis project is to offer a practical solution for adapting user interfaces and services to changing needs and wishes of older adults in a flexible way by providing various devices and I/O modalities. The paper describes the approach for the integration of an avatar into the generated user interfaces. The user interfaces aim to provide consistent look and feel for different services to interact using the user's preferred modality. In the first development cycle, the focus was on covering the whole transformation process from abstract task descriptions to renderable UIs to be displayed on various I/O devices. First evaluations in a lab setting have been performed to detect usability issues. The participants gave very positive feedback about the identical layout of the application on different I/O devices, but nevertheless some usability issues have been identified. Solutions to tackle these issues are presented in this paper and especially the avatar integration will be a big step towards an increased acceptance of AAL services and their user interfaces.**

***Keywords*—Ambient Assisted Living, multimodal user interfaces, avatar enriched user interfaces, adaptability**

I. Introduction

Older adults benefit from information and communication technology solutions and services in the Ambient Assisted Living (AAL) domain. Needs and wishes regarding the interaction with ICT solutions change over time and vary between older adults. They depend on the user's physical and mental capabilities and preferences. Many currently available user interfaces (UIs) for ICT solutions for older adults often do not take these factors into account. And that is problematic, since the user interface has to be considered critical to the success or failure of an ICT product or service [1]. In particular usability, accessibility, as well as the freedom of choice concerning the interaction with such systems are the crucial points for acceptability and applicability. Furthermore, the benefit of such systems for the user him- or herself, for the society and also for other stakeholders depends on these issues [2]. Coherence (i.e., a seamless control of different services and appliances within the user interface), task orientation (i.e., the user interface should allow to start functionalities but not expose how it will be achieved), scalability (i.e., the possibility to reduce and/or expand the functionality of the user interface) and accessibility (i.e., accessible for a wide range of users) are important features when talking about user interfaces [3].

The aim of the AALuis project [4] is to offer a practical solution for the adaptation of user interfaces and services to changing needs and wishes of older adults in a flexible way by providing user interfaces for various devices and with multiple, possible I/O modalities. These objectives are based on the following three major challenges to improve the way older adults interact with AAL services: (1) Older adults are a very heterogeneous target group. A change of capabilities and needs over time is normal and typical within the process of aging. (2) Exchangeability, flexibility and usability of user interfaces and their standardized integration are of uttermost importance. (3) Freedom of choice regarding the user interface helps with an assisted living lifestyle approach and avoidance of care-stigmatized services are crucial challenges for scalable services.

The generated user interfaces are intuitive and follow recent market trends as the focus is on multi-touch tables, portable devices and smart-TV based solutions. The rendered interfaces offer new intuitive ways of interaction and avatar interfaces simulating a "face to face" communication. Using an avatar as a virtual presenter of information creates additional value, since the addition of a visual display to verbal information can increase the intelligibility and enhance the robustness of the information transmission [5] as known from natural speech [6]. The paper describes the approach for the integration of the avatar into the generated user interfaces, which aim to provide consistent look and feel for different services with the possibility for interactions using the user's preferred modality.

The paper is structured in the following way. After the introduction (Section I) and a brief description of the state of the art in (semi-)automatic user interface generation (Section II), an overview over the used methodologies and approaches is given (Section III). Thereafter, some intermediate results (Section IV) and an outlook for improvements (Section V) are presented. Finally, the presented solution is discussed (Section VI).

II. State of the art

Currently, user interfaces are customized for the device they are implemented for. There are approaches to separate the appearance and control of the user interface from the

 ISBN: 978-1-61208-314-8

device. However, these solutions are focused on the device-independent display of multimedia content and hence are too complex with respect to content delivery and not suitable for the device-independent control of services in the AAL context. To overcome those issues User Interface Description Languages (UIDL) [7] have been invented. Instead of creating an user interface for a specific platform directly, the user interface is modelled in a more abstract format. The model describes the necessary interaction elements of the user interface for all tasks to be carried out in a machine-readable form, which can be used for the (semi-)automatically generation of the renderable user interface.

In recent years, there has been research going on regarding the (semi-)automatic generation of user interfaces. universAAL aims to develop an open platform to ensure technically feasibility and economically viability to conceive, design and deploy innovative AAL services [8]. The approach is to detach the UI bus fully from the service and context bus and to create the user interface using XForms. Another approach is the URC, which is an international standard (ISO/IEC 24752) defining a way to control arbitrary electronic devices or services (i.e., hardware or software) with interoperable, pluggable user interfaces. The UCH [9] realizes the URC standard as a middleware server component providing connection points to existing, non URC compliant entities. Another approach is to use a task model for the creation of the renderable user interface [10] [11], which is the basis for the presented work.

III. Methodology and Approaches

A. User interface generation and adaptation

In Figure 1, the overall user interface generation and adaptation process is depicted. The generated user interfaces and the services are connected by the intermediate OSGi based AALuis layer, which performs the whole dialog management and transformation process. The middleware layer can either run stand-alone or on top of another OSGi based AAL middleware (e.g., HOMER [12]–[14]).

The services are integrated into the AALuis layer either as an OSGi bundle or as a loosely-coupled web service. To ensure that the service developers just need to take care about their services, but do not need to worry about the final user interface, they have to provide a description of the task flow of the service in concur task tree (CTT) notation [10]. A simple binding file in XML format connects concrete service methods to its corresponding CTT tasks.

The user interface is generated at run-time based on the task flow description. In the transformation process from the CTT to an abstract user interface (AUI) in MariaXML [11] to the renderable user interfaces (RUIs) in HTML5 different information is taken into account: (1) the user preferences, (2) the available devices and their capabilities, (3) service and task context, and (4) environmental context.

A user preference can, for example, be that the user prefers to use the animated talking avatar as an additional output modality to display text information.

B. Generation of the avatar

The avatar generation is carried out on a server with 3D graphics hardware and custom server software components. These components comprise a control server software that also contains the logic for animation automation, an audio speech synthesis module, a lip-sync component, a 3D render engine with additional video compression functionality, and a webserver that provides the generated video files. The components communicate via TCP/IP, the control server is queried over the internet via the WebSocket [15] protocol.

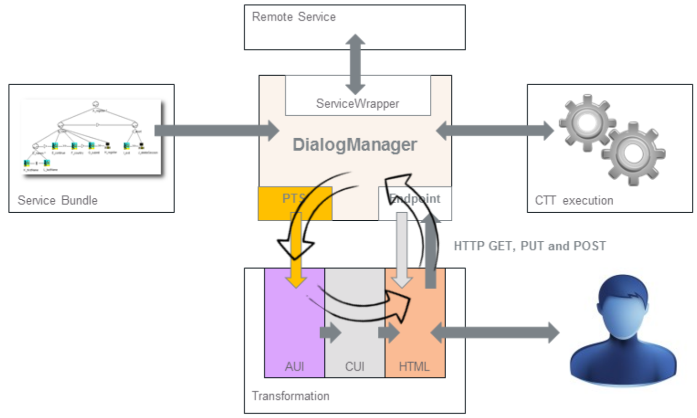

Fig. 1. The AALuis user interface generation process

A given text is sent to the control software and converted to audio speech by CereProc cServer [16], a commercial test-to-speech synthesizer. As intermediate data of this step a phonetic transcription, i.e., a sequence of phones to be spoken with their durations, is generated. This infromation is used to create control commands for jaw opening, lip opening, lip spreading, and tongue tip raising of the avatar by a numerical articulation model [17] implemented in the lip-sync module. A unique animation script is composed at each request. It layers the speech animation and an adequate sequence of body movements from a large set of animation clips. Words that are prominent in the synthesized audio speech signal are assigned to animation clips that visually emphasize the respective part of the spoken text by hand gestures, head or body movements. The animation script is executed, rendered by a modified 3D computer game engine, and converted by FFmpef [18] into an h.264 video that shows the avatar speaking the given text with speech-accompanying hand gestures and head and body movements.

C. Integration of the avatar

An interface to communicate with the avatar server is directly integrated into the OSGi framework running the AALuis layer. Character and scene settings for the avatar are defined by the service. The text and the settings (e.g., language) are passed on as call parameters and the avatar creation engine returns a reference to the generated and cached video file (Figure 2).

```
VideoObject video =
    zoobeService.getVideoFromText
    (sInputText, sLang, ...)
```

Fig. 2. Code snippet for the avatar creation

Each task can have several input and output parameters. As mentioned in Section III-A, the binding.xml file binds these task parameters to concrete service methods by using a special

 ISBN: 978-1-61208-314-8

AALuis data format (AALuisData). This allows the transport of different data types and is internally realized as a kind of key/value map, which enables the multi-modality support. By the fact that each task is able to serve multiple and different modalities, the transformation process places the current used modality set into the transformation context variable. In Figure 3, one can see a snippet of the context variable representing text and avatar based output.

```
eu.aaluis.context.service.data: {
  "/aaluis/task/AALuisService/Greet": {
   "text/plain":"Welcome to AALuis!",
   "application/x-mpegURL":
   "http://.../cacheManager/0254668.mpeg"
  }
}
```

Fig. 3. Code snippet for the context variable

Depending on the device and user preference either text/plain and/or the application/x-mpegURL (video) will be rendered. If so, the generated video file is directly embedded in the renderable user interface. The video file is streamed from the server and played when the user interface is displayed. The video file is cached locally and played from the file system, if the same utterance is requested with the same parameters again [2].

IV. Intermediate Results

In the first development cycle, the focus was on covering the whole transformation process from abstract task descriptions to renderable user interfaces to be displayed on various I/O devices. In Figure 4, one can see a generated user interface of the first prototype. Usability and the integration of avatars are the main goals of the second development iteration.

First evaluations in a lab setting have been performed to detect usability issues. The participants (two groups - UG1 ($N = 5; 67 \pm 7$ years) and UG2 ($N = 4; 76 \pm 4$ years)) gave positive feedback about the identical layout of the application on different I/O devices, but nevertheless, some usability issues have been identified. These issues covered disappearing navigation possibilities when scrolling down, missing titles on the different screens which led to confusions regarding the actual functionality of the application and the grouping of interaction items (buttons). Concerning the grouping, participants were irritated by the placement of buttons for confirmation/activation (e.g., send message) and navigation jointly at the top of the application.

All participants in user group 1 had general knowledge of technical devices, their own TV, and a mobile phone. All participants in user group 2 were living at home but needed some sort of support. In user group 2, all but one participants had no experience with touch screens and refused to use them or were not able to use them anymore due to health problems.

V. Outlook

The focus in the second development cycle will be to solve the detected usability issues and on enriching the generated user interfaces. Thus, the specification for the second prototype has been updated based on the evaluation results and the

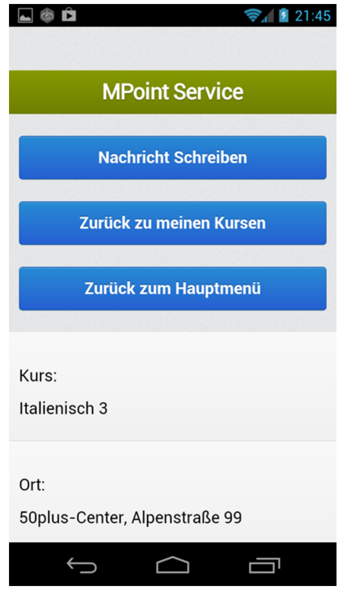

Fig. 4. Generated UI

integration of images and especially the avatar. As described in Section III-C, the AALuis layer is already capable of avatar integration, but the integration is not yet realized in the transformation process. This will be done in the final prototype. In Figure 5, one can see a sketch for the placement of the avatar as an additional modality for the presentation of information. It is planned to enable the user to start, stop and replay the information presented by the avatar. The figure also illustrates the dual modality by presenting the same information once as text and once as an avatar video as channel for audio-visual speech.

The issue of misleading grouping of interaction items will be tackled by using control types in the abstract user interface which are specified in the MARIA language, namely ACTIVATOR control and NAVIGATOR control. The idea is not to extend the CTT with additional semantic, but just to use the only available interaction task of type CONTROL and to distinguish between NAVIGATION (i.e., to navigate back to the previous rendered user interface), ACTIVATOR (i.e., to execute a POST method on the current rendered user interface) and option controls (i.e., all other controls) in the first transformation step creating the abstract user interface by analyzing the CTT of the service in detail. Option controls are not defined in the MARIA language, thus a grouping mechanism to summarize all controls that will be rendered in the ”option” menu will be performed. The title for each presentation task set, which is a group of all enabled tasks at the same time, will be derived from the currently active ACTIVATOR command and be displayed as title for the active screens.

VI. Conclusion and future work

As mentioned in Section IV, the results of the first evaluations are promising and the overall feedback was positive. Nevertheless there are still some usability and acceptance issues to be solved. Some solutions, which are mainly based on the user feedback, are presented in this paper and especially the avatar integration will be a big step towards an increased acceptance

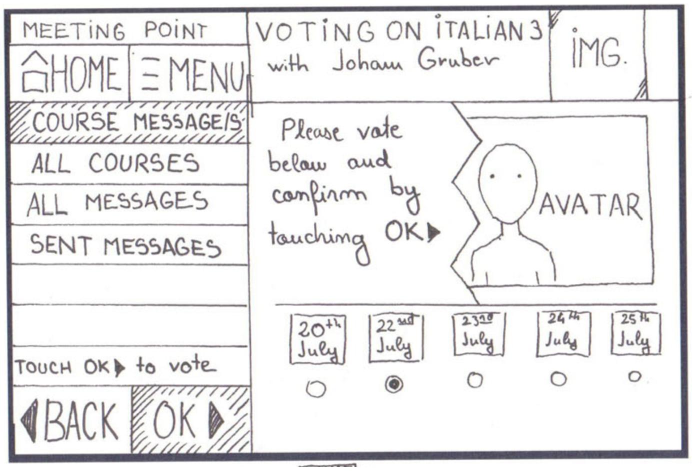

Fig. 5. Mock-up with integrated avatar for a tablet

of AAL services and their user interfaces. The improvements are all based on the evaluation results, interviews, experiences from other research projects and a former study [19] [20]. Further evaluation tests with older adults, which are planned in the course of the project, will bring deeper insights on acceptability, likeability and usability of avatars within AAL environments.

Generated user interfaces tend in general to be not very user-friendly. AALuis tries to change this using a layered template approach. One of the lessons learned from the first evaluations is that one has to find a good compromise between flexibility and user-friendliness. When dealing with older adults, the usability of the user interface is of uttermost importance. Typically, more user-friendly interfaces are less flexible from the middleware point of view and very flexible user interfaces (being almost completely automatically generated) tend to be less usable. Using the approach of taking user-friendly templates and merging them with accessibility needs and content seems to be promising. However, this approach might be not flexible enough or too complicated and thus has to be evaluated in the future.

The AALuis layer will be released as an open source module and thus can contribute to speed up the development process of AAL services and user interfaces and as a consequence to reduce related costs. Using this strategy, AALuis can be a driving force for the development of new and innovative user interfaces and services for older adults. For this purpose, the acceptance of service developers is of importance; thus, a standardized and relatively easy integration and usage of the AALuis layer in other AAL systems is one of the key objectives.

Acknowledgment

The project AALuis is co-funded by the AAL Joint Programme (AAL-2010-3-070) and the following National Authorities and R&D programs in Austria, Germany and The Netherlands: bmvit, program benefit, FFG (AT), BMBF (DE) and ZonMw (NL).

References

[1] Etsi user interfaces. [Online]. Available: http://www.etsi.org/index.php/technologies-clusters/technologies/human-factors/user-interfaces [retrieved: 09, 2013]

[2] S. Fagel, A. Hilbert, M. Morandell, and C. Mayer, "The virtual counselor – automated character animation for ambient assisted living," in *Proceedings of 6th International Conference on Advances in Computer-Human Interactions (ACHI'13)*. ARIA XPS Press, 2013, pp. 184–187.

[3] J. Alexandersson, G. Zimmermann, and J. Bund, "User interfaces for aal: How can i satisfy all users?" in *Proceedings of Ambient Assisted Living: 2. Deutscher Kongress, Berlin, Germany, January 27-28*, 2009, pp. 5–9.

[4] C. Mayer, M. Morandell, S. Hanke, J. Bobeth, T. Bosch, S. Fagel, M. Groot, K. Hackbarth, W. Marschitz, C. Schüler, and K. Tuinenbreijer, "Ambient assisted living user interfaces," in *Everyday Technology for Independence and Care, AAATE 2011*, ser. Assistive Technology Research Series, G. G. et al., Ed., vol. 39. IOS Press, 2011, pp. 456–463.

[5] S. Ouni, M. M. Cohen, H. Ishak, and D. W. Massaro, "Visual contribution to speech perception: measuring the intelligibility of animated talking heads," *EURASIP J. Audio Speech Music Process.*, vol. 2007, no. 1, Jan. 2007.

[6] W. H. Sumby and I. Pollack, "Visual contribution to speech intelligibility in noise," *The journal of the acoustical society of america*, vol. 26, pp. 212–215, 1954.

[7] J. Guerrero-Garcia, J. Gonzalez-Calleros, J. Vanderdonckt, and J. Muoz-Arteaga, "A theoretical survey of user interface description languages: Preliminary results," in *Web Congress, 2009. LA-WEB '09. Latin American*, 2009, pp. 36–43.

[8] S. Hanke, C. Mayer, O. Hoeftberger, H. Boos, R. Wichert, M.-R. Tazari, P. Wolf, and F. Furfari, "universaal–an open and consolidated aal platform," in *Ambient assisted living*. Springer, 2011, pp. 127–140.

[9] G. Zimmermann and G. Vanderheiden, "The universal control hub: An open platform for remote user interfaces in the digital home," in *Human-Computer Interaction. Interaction Platforms and Techniques*, ser. Lecture Notes in Computer Science, J. Jacko, Ed. Springer Berlin Heidelberg, 2007, vol. 4551, pp. 1040–1049.

[10] F. Paternò, C. Mancini, and S. Meniconi, "Concurtasktrees: A diagrammatic notation for specifying task models," in *Human-Computer Interaction INTERACT97*. Springer, 1997, pp. 362–369.

[11] F. Paterno, C. Santoro, and L. D. Spano, "Maria: A universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 16, no. 4, p. 19, 2009.

[12] C. Mayer, J. Kropf, A. Hochgatterer, S. Hanke, and P. Nieke, "Flexible and standard-compliant framework for sensor based aal smart home applications," in *Intelligent Buildings and Smart Homes Conference 2009 (IBASH2009,) Nov. 18-20, 2009, Taipei, Taiwan*, 2009, pp. 200–203.

[13] T. Fuxreiter, C. Mayer, S. Hanke, M. Gira, M. Sili, and J. Kropf, "A modular platform for event recognition in smart homes," in *2010 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom)*, 2010, pp. 1–6.

[14] T. Fuxreiter, M. Gira, L. Roedl, J. Kropf, and S. Hanke, "Rule-based indoor localization using non-intrusive domotic sensors and the homer platform," in *Proceedings of the 3rd AAL Forum: Partnerships for Social Innovations in Europe, Lecce, Italy, September 26 - 28, 2011*, 2011, pp. 569–576.

[15] Ietf: The websocket protocol. rfc6455, 2011. [Online]. Available: http://tools.ietf.org/html/rfc6455 [retrieved: 09, 2013]

[16] Cereproc, cserver text-to-speech server. [Online]. Available: http://www.cereproc.com/en/products/server [retrieved: 09, 2013]

[17] S. Fagel and C. Clemens, "An articulation model for audiovisual speech synthesisdetermination, adjustment, evaluation," *Speech Communication*, vol. 44, no. 1, pp. 141–154, 2004.

[18] Ffmpeg. [Online]. Available: http://ffmpeg.org/ [retrieved: 09, 2013]

[19] M. M. Morandell, A. Hochgatterer, S. Fagel, and S. Wassertheurer, "Avatars in assistive homes for the elderly," in *HCI and usability for education and work*. Springer, 2008, pp. 391–402.

[20] M. M. Morandell, A. Hochgatterer, B. Wöckl, S. Dittenberger, and S. Fagel, "Avatarshome," in *HCI and Usability for e-Inclusion*. Springer, 2009, pp. 353–365.