Enhancing the Flexibility of a Multimodal Smart-Home Environment

Thimios Dimopulos[2], Sahin Albayrak[2], Klaus Engelbrecht[1], Grzegorz Lehmann[2], Sebastian Möller[1]
[1] *Deutsche Telekom Laboratories* [2] *DAI-Labor, TU Berlin,*
{*sebastian.moeller, klaus-peter.engelbrecht*}*@telekom.de*
{*thimios.dimopulos, grzegorz.lehmann, sahin.albayrak*}*@dai-labor.de*

Introduction

The smart home technology should increase communications, awareness, and functionality while avoiding usability problems caused by the technological complexity of electronic home appliances. Despite its advantages, strictly spoken interaction is less effective than multimodal interaction combining speech, gestures, eye gaze, body posture and traditional appliance controls (remote control, nods). The EU-funded IST project INSPIRE [5], developed a "home assistant" which hides the complexity of home appliances through an intuitive interface. During the INSPIRE project, system development focused on natural language control of domestic devices, e.g. TV, video recorder, answering machine, lamps, blinds etc. It has since then evolved into a framework for multimodal application design, by the work of Cenek, Melichar and Rajman [2].

In a collaboration between the Strategic Research Laboratories-Usability of Deutsche Telekom Laboratories and the DAI-Labor, Berlin University of Technology, we augmented the modularity and flexibility of the system, by exploiting recent advances in consumer electronics, wireless communication and user interface technologies. We enhanced the user interface with a VoiceXML platform providing remote or local speech interaction in a unified, standards-based fashion. We leveraged on the system's modularity with a protocol-independent device controlling mechanism. In Section 2, an overview of the current system architecture is presented. The main parts of the system along with our recent enhancements are discussed in Sections 3 and 4. Finally, conclusions and plans for future work are presented.

System Architecture - Latest Enhancements

The INSPIRE system is designed after a distributed component architecture, enabling flexible integration and the reuse of heterogeneous software components. An overview of the system is presented in Figure 1.

INPIRE Core

The INPIRE core is designed after a distributed component architecture paradigm. The dialogue (or interaction) manager is the central component that controls the interaction flow. Its input is a formal representation of the user's expression, e.g. a spoken utterance, a gesture, or a combination of them. The dialogue manager output is a formal representation of a response to the user, meant

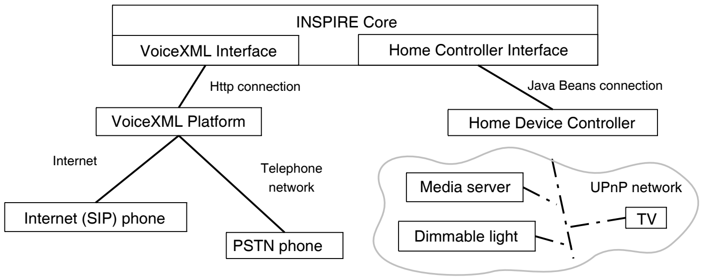

Figure 1: System architecture overview.

to be consumed by an output component. In addition, it communicates action-taking commands to the action-performing components. Bui et al. [1] have published a detailed description of the dialogue manager.

Further core components bridge the dialogue manager with the human interaction modalities, and with the devices of the smart home. Input components are concerned with gathering the input from the possible modalities, semantically interpreting it and feeding it to the dialogue manager. Output components drive the output of the dialogue manager towards the user, and action components realize the actions that the dialogue manager decides to take.

The newly implemented VoiceXML interface is both an output and an input component. For each dialogue exchange, it produces a VoiceXML document, which describes the next spoken interaction with the user. This description is then forwarded to the Voice Platform which handles the actual user interaction. Once the user response has been recognised, the Voice Platform forwards it to the VoiceXML interface component.

The Home Controller Interface is an action component that maps the actions emitted by the dialogue manager to the abstract device description provided by the Home Device Controller. Special effort is put into the flexibility of this mapping. For example, the dialogue manager might decide to display a video recording to the user. The home controller interface will instruct the controller to find the video-rendering-capable screen which is closest to the user's current position in the smart home, and render the video on that one.

Voice Platform

During the INSPIRE project, the system was thoroughly tested with locally-connected speech recognisers and TTS systems [5]. With the advent of sophisticated voice platforms based on the VoiceXML specification, developing a VoiceXML interface for the INPIRE system emerged as an appealing advancement. VoiceXML is currently a widely accepted standard for building speech interfaces. By enhancing the core of the INSPIRE system with a VoiceXML interface component, we gained a vendor-independent connection to any voice platform.

In the current setup, a commercial Voice Platform is used. It incorporates a speech recogniser which we have trained using a natural language modelling approach based on bi-grams. For speech output, one option is to use the TTS system of the Voice Platform. To cope with the lack of naturalness of TTS systems, the INSPIRE core includes a component that generates prompts by concatenating prerecorded speech. The prompts are then served to the Voice Platform which can be accessed either by an IP-phone through the internet, or by mobile and land line phones.

Despite its strong points, VoiceXML was mainly designed for call centre applications. In our system, it is used in the simplest way, namely as an interface to the Voice Platform. The purpose of this is to keep all crucial functionality like the dialogue logic or the semantic interpretation of user utterances in the Java components that have already been developed. This way, the modularity and flexibility of the system is preserved, while a new way to interact with it is incorporated.

Home Device Controller

In order to evaluate the INSPIRE system in a real-world smart-home environment, we have integrated it with the Home Device Controller (HDC). The HDC architecture has been developed during the "Seamless Home Services" project at DAI-Labor, and has been designed to serve as the back-end for smart-home systems [3]. It provides the means to easily access and use devices and services in the home environment. The HDC defines its own abstract, ontology-based description of devices, which is independent of any standard or vendor. This approach is necessary to cope with the heterogeneity of today's home appliances. Moreover, it also makes the HDC architecture easy to use for developers.

Internally, each device discovered in the environment by the HDC is managed by a controller, which is responsible for the standard-specific communication with it. If a new standard needs to be supported, a new controller must be implemented. This can be done in an easy way, because a lot of effort has been put to assure good extensibility of the HDC. Often this is not necessary though, because multiple controllers have already been created during the works on the HDC. The development focused on the UPnP technology, which is currently the most established standard for device discovery and control in the home environment.

By using the HDC architecture, the integration of the INSPIRE system within the home environment has been greatly simplified. As already described, the INSPIRE core has been enhanced with an action component that maps the actions emitted by the dialogue manager to the abstract device description provided by the HDC. This way our system can fully benefit from the flexibility of the HDC and its already implemented controllers. Furthermore, new device types can be integrated using the mechanisms provided by the HDC without any changes in the rest of the system.

Conclusions and Open Problems

We have leveraged the flexibility and extensibility of the INSPIRE system, by extending the core with a VoiceXML interface component that enables coupling it to any VoiceXML enabled platform. We are currently using a commercial Voice Platform so that the system can be accessed using any kind of speech telecommunication device. We have extended the smart-home environment with a controller framework that enables control of home appliances in a unified way, regardless of protocol or vendor. We already use UPnP appliances including media servers and various media renderers, along with non-UPnP based blinds and lights.

An open problem is the auto-generation of the, currently manually-produced, task and dialogue models. Ideally, connecting a new appliance in the smart home should automatically extend the dialogue to include the interaction with the new appliance. Our interests also go beyond the system itself. We would like to establish a framework where the overall quality of multimodal systems could be formally evaluated. In the MeMo project [4], we aim in evaluating and predicting the usability of interactive systems in an automated way.

References

[1] Bui, T.; Rajman, M.; Melichar, M.: Rapid Dialogue Prototyping Methodology. Lecture Notes in Computer Science, Volume 3206, Jan 2004.

[2] Cenek, P.; Melichar, M.; Rajman, M.: A Framework for Rapid Multimodal Application Design. Lecture Notes in Computer Science, Volume 3658, Aug 2005.

[3] Feuerstack, S.; Blumendorf, M.; Lehmann, G.; Albayrak, S.: Seamless Home Services. AmI.d Ambient Intelligence Developments Conference, September 20-22, 2006.

[4] Möller, S.; Engelbrecht, K.; Englert, R.; Hafner, V.; Jameson, A.; Oulasvirta, A; Raake, A.; Reithinger, N.: MeMo: Towards Automatic Usability Evaluation of Spoken Dialogue Services by User Error Simulations. Proceedings of INTERSPEECH 2006 - ICSLP, International Speech Communication Association (ISCA).

[5] Möller, S.; Krebber, J.; Raake, A.; Smeele, P.; Rajman, M.; Melichar, M.; Pallotta, V.; Tsakou, G.; Kladis, B.; Vovos, A.; Hoonhout, J.; Schuchardt, D.; Fakotakis, N.; Ganchev, T.; Potamitis, I.: INSPIRE: Evaluation of a Smart-Home System for Infotainment Management and Device Control. Proc. LREC 2004, Lisbon, 1603-1606.