Magical Mirror: multimedia, interactive services in home automation

Luigi Ceccaroni and Xavier Verdaguer

:tmtfactory>
Edificio Forum de la Tecnología, c. Marie Curie s/n, 08042 Barcelona
`luigi@tmtfactory.com, xavier@tmtfactory.com`

Abstract. Several academic and industrial research groups are working to deploy home-automation systems and to improve their user interfaces. Based on user studies and prototype implementation, this paper presents the development of an innovating appliance that incorporates interactive services of leisure and information, offered through a high-quality user interface on the surface of a mirror. This appliance is thought to be both integrated in the programmable network of a home and used as an independent element. The prototype offers basic services, like interactive television, the presentation of personalized weather data, news and selections of musical tunes, and will incorporate some advanced functionalities, like an intelligent service of personal motivation and a reminder service.

1 Introduction

Modern man is basically sleepwalking if he is not multitasking. His lifestyle has evolved in such a way that optimizing time is the most important thing. The newest innovations show that we can -and therefore must- be online while watching TV, be e-mail—ready while driving the car, be taking calls while ascending Mount Everest, be watching television while shaving in front of the mirror. Most evenings most first world's kids do their homework on the computer while instant messaging friends and talking on the phone. The stereo is typically playing and they are probably downloading MP3s and DivX movies. Occasionally their foot extends to gently connect with the dog. There may be something wrong with this picture, but the kids are definitely not sleepwalking. The pressure to be wired -everywhere, always, fast- is hard to ignore [6].

Before the increasing interest of the industry and the market for the scope of home automation, the concept of intelligent human-habitat, managed by machines, has left the science-fiction domain to become, in the last years, a necessary investment-area for the companies that try to lead the sector in the mid term. But, is home automation really a necessity for consumers? It's too easy, of course, to be a skeptic, to regard each new invention with a sense of wise detachment and, while praising its state-of-the-art capabilities, wonder aloud if we really need it. After all, nobody wants to risk missing real progress.

2 Related work

Several projects and public initiatives worldwide are dealing with ambient intelligence (e.g., [1] [4] [7] [9] [10] [12] [13]) and a few of them are developing interactive mirror-displays on which, based on personal preferences, different applications can be run. In particular, the HomeLab [14] at Philips Research [15] is working on a project which is very similar to the subject of this paper: the Magical Mirror project.

Several other groups dealt with the development of dialogue systems in diverse domains, not only home automation, with projects such as: SmartKom [16] on the design of new interfaces, which aims to help reducing the hesitations people presently feel upon using information technology and thus to make a contribution to the user friendliness and user centeredness of technology in today's information society; TRINDI [17] on task oriented instructional dialogue, which built a computational model of dialogue moves; Jupiter [18], a conversational system that provides up-to-date weather information over the phone, and other projects at MIT's Spoken Language Systems; the Conversational Interfaces project [7], whose aim is to build a general purpose dialogue system which supports multi-modal (i.e., speech and gesture) activity-oriented dialogues with devices, applications and services, and other projects at Stanford University's Center for the Study of Language and Information; and August [4], a Swedish spoken dialogue system.

3 Overview

Often, really useful technologies are not considered a need in a first phase (e.g., telephones, computers, PDAs) and part of the challenge in this project is demonstrating a technology to help creating new needs. In the information society, the emphasis is shifting from traditional computer interfaces (screen and keyboard) to computational modules embedded in appliances, hand-held devices and information systems, which often have to interchange significant amounts of multimedia content.

To be sure, current technologies that transmit images and voice fall far short of anything remotely realistic [6]: at present, home automation generally involves jerky video and canned sound. In Magical Mirror we have to deal with these limitations in order to develop an interactive appliance, which is able to provide the following realistic services, controlled by voice, in Spanish (English and Catalan, in the future), and presented with the support of a hand-designed, high-quality, advanced visual interface (see Fig. 1):

S1. customized weather data
S2. customized information on traffic and public transport
S3. customized news (stock market, sports)
S4. customized music management
S5. interactive television (integration with the TVFinder service [2])
S6. Internet access and search
S7. personal agenda
S8. home automation (control of other devices: lights, heating, oven)
S9. communication among users (voice-message recording, email)

Fig. 1. Magical Mirror's high-quality interface (BCNMedia and **:tmtfactory>** are trademarks of the same company.)

4 Architecture

Magical Mirror is a meta-service, with the layered structure shown in Fig. 2, which is suitable to be integrated into a home automation network. Multiple infrastructures of communication exist that are candidates to form the backbone of a home-automation environment. It does not seem that, in the short term, any of them is going to constitute a broadly accepted standard and, in any case, it seems that there is a strong dissociation between the network of devices control and the network of data distribution (video, audio, and communications between computers). Nevertheless, it is important that the home-automation environment appears, before the user, as one interface. It would not be acceptable that he has to use an interface in order to handle home devices (lights, heating, alarms, oven), another one for communications devices (telephone, fax, email) and a different one for audio and video devices (TV, music). In order to obtain this unity in the interface, a high level of interaction, among the heterogeneous infrastructures of communications that will support it, is needed.

4.1 Device layer

At the physical-device level, interactivity is possible through a reception-emission system with the following characteristics:

- reception: microphone arrays (for voice recognition), micro-camera (for image recognition [future work]), touch screen (for user input [future work]).
- emission: retro-projection of images and videos, retro-projection of text, micro-speakers for voice synthesis and music.
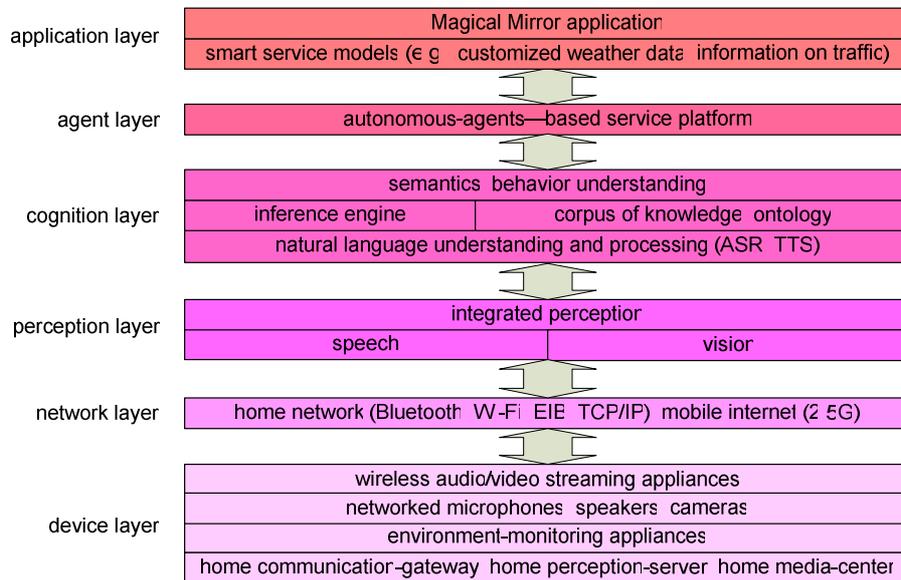
| application layer | Magical Mirror application |
| | smart service models (e.g. customized weather data, information on traffic) |

Fig. 2. Layered structure of the Magical Mirror service

4.2 Network layer

As home technologies expand to encompass computer and communication networks, standards and specifications are quickly multiplying. In Table 1 a list of the main ones developed by alliances and working groups (non proprietary) is presented. In Magical Mirror, the control of the different devices that compose the home-automation environment is, for the moment, simulated and will be mainly based on equipment with the following connection standards: Bluetooth (between PDAs and Magical Mirror), IEEE 802.11g and IEEE 802.15 (among Wi-Fi--enabled computers and devices). With the purpose of connecting all the devices that integrate the environment and to provide a suitable base of operation, a hardware architecture that is able to manage and integrate the different components, independently of the network to which they are connected, will be needed. In addition, it seems very suitable to provide means of connection between the Web world and the home-automation environment. As solution, we will work at the integration of different networks using the TCP/IP protocol. Among the advantages of having access by means of the TCP/IP protocol to elements connected by different networks, we can enumerate:

- Extending the scope of home automation to Web services and making available a global, remote access.
- Having access to the home-automation environment by means of WAP (wireless application protocol)-based applications, used in the mobile telephony, which will allow possibilities of ubiquitous control.
- Guaranteeing that we can have all kinds of connected equipment, without being restricted to a specific interconnection protocol.

Standard, specification	Media	Description
Bluetooth [http://www.bluetooth.com]	RF	Bluetooth is the codename for a technology specification for small form-factor, low-cost, short-range radio links between PDAs, PCs, mobile phones and other portable devices. The Bluetooth Special Interest Group is an industry group consisting of leaders in the telecommunications and computing industries that are driving the development of the technology and bringing it to market.
EIB [http://www.eiba.com]	Twisted Pair	The European Installation Bus is a home-automation system developed under the auspices of the European Union with the objective of balancing the imports of similar technologies that were taking place from the Japanese and the North American market. It links sensors and actuators to building systems that control HVAC (heating, ventilating and air conditioning), security, access, and life safety.
HomePlug Alliance [http://www.homeplug.com]	Power Line	Created to set a technology specification for home power-line networking and to promote its wide acceptance in the marketplace. The alliance's objective is to enable and promote rapid availability and adoption of cost effective, interoperable and specification-based home power-line networks and products enabling the connected home.
IEEE 802.15.4 [http://www.ieee802.org/ 15/pub/TG4.html]	Wireless	The IEEE 802.15 TG4 is chartered to investigate a low data rate solution with multi-month to multi-year battery life and very low complexity. It is intended to operate in an unlicensed, international frequency band. Potential applications are sensors, interactive toys, smart badges, remote controls, and home automation.
Konnex [http://www.konnex-knx.com]	All	The common association of EIBA (European Installation Bus Association), BCI (BatiBUS Club International) and HES (European Home System).
WECA [http://www.wirelessethernet.org]	RF	Wireless Ethernet Compatibility Alliance's mission is to certify interoperability of IEEE 802.11 high-rate products and promote that standard for the enterprise, small business and the home. IEEE 802.11 offers connectivity at speeds up to 11 Mbps.
ZigBee [http://www.zigbee.org]	Wireless	The ZigBee Alliance is an association of companies working together to enable reliable, cost-effective, low-power, wirelessly networked, monitoring and control products based on an open global standard.

Table 1. Main home-automation/network standards and specifications, developed by alliances and working groups, deployable in Europe

Through the protocols described, Magical Mirror can be connected to different devices that can be found in a home, becoming, in a flexible way, an integral part of a home-automation environment (see *modality c* below).

4.3 Perception and cognition layers

Voice and image feeds are collected in the perception layer and processed (voice only, at the moment) in the cognition layer. The main challenges during perception are:

- automatic speaker-identification;
- voice-silence detection;
- robustness with respect to the environmental noise (basically, water noise, echo and other voices).

Research and commercial tools are used for automatic speech recognition (ASR) and text to speech (TTS) processing. VoiceXML (VXML)[1] is used in service models to represent grammars, speech acts and dialogues. Here is a simple example of VXML encoding and a sample human-computer interaction (translated from Spanish to English) in the case of customized music management:

```
<form>
  <field name="artist">
    <prompt>Whom of these artists World you like to lis-
            ten to?: Anticappella? Aphex Twin? Eiffel
            65? Elio e le Storie Tese? o Leftfield?
    </prompt>
    <grammar       src="artist.grxml"
                   type="application/srgs+xml"/>
  </field>
  <block>
    <submit next="http://www.tmtfactory.com/art2.asp"/>
  </block>
</form>
```

A field is an *input* field. The user must provide a value for the field before proceeding to the next element in the form.

C (computer): Whom of these artists World you like to listen to?: Anticappella? Aphex Twin? Eiffel 65? Elio e le Storie Tese? o Leftfield?

H (human): Mogwai.

C: I didn't understand what you said. (default message.)

C: Whom of these artists World you like to listen to?: Anticappella? Aphex Twin? Eiffel 65? Elio e le Storie Tese? o Leftfield?

H: Aphex Twin

C: (continues in document art2.asp)

4.4 Agent layer

The highest layers of the architecture are based on independent agents which interact among them and with the user, and which accomplish the intelligent control of diverse devices, sensors as well as actuators (in phase of study for *modality c*). In the project, the design of the architecture of independent agents for the implementation of home automation has been carefully studied. Agent-based architectures present, potentially, a series of fundamental advantages, such as modularity, reusability, scalability and dynamic reconfiguration. Each agent has a basic functionality and in addition

---

[1] See [http://www.voicexml.org].

is equipped with certain *decision and adaptation* capabilities, in such a way that it can improve its efficiency on the basis of information that receives directly from other agents or indirectly through three kinds of central support knowledge-components, a *base of facts and knowledge* that is common to all the agents, a **domain ontology** (*DomOnt*) and a number of **service ontologies** (*ServOnt*). An ontology is a model of the world in which the different agents of the system work. The ontology is made up of a hierarchy of entities, which represent buildings, rooms, elements that compose the rooms, users, tasks that these users are accomplishing, and any other information of interest for the community of agents which belong to the home-automation environment. Agents use the *DomOnt* as a mechanism for obtaining data they need about the environment and, in turn, deposit in the ontology their results so that other agents can take advantage of them.

While the *DomOnt* contains a representation of the environment and (in its instances) of its state, the *ServOnt* models the information flow that is exchanged among the different agents part of the environment, influencing the language and the communication protocols between them. The agents channel the information according to the flow model contained in the *ServOnt*. For example, if an agent deduces that the user requires that a certain flow of video (the output of a camera) is transferred from the hall to the mirror of the bath, the only thing it has to do is to change some parameters of the model of information flow contained in the *ServOnt*. The agents controlling the camera in the hall and the computer in the bath respond to this change in the ontology one emitting the video flow towards the computer in the bath and the other beginning to emit this flow in the Magical Mirror.

An important element of the Magical Mirror project is to allow the addition to the home-automation environment of any sort of new agents and devices developed by third parties using the same standards. With this aim, we use OWL [3] and OWL-S [9] (both based on XML) as the languages of definition of the *DomOnt* and the *ServOnt*, respectively. This assures that the requirements that a new agent will have to fulfill to join the home-automation environment will be minimum, standardized and of easy implementation. With respect to the agent systems' architecture, the general requirements are: (1) compliance with the FIPA standard of communication and interaction, and (2) ability to use OWL ontologies.

5 Personalization

In Magical Mirror, due to the presence of a number of different services, personalization is a long and challenging process to be carried out. The personalization methodology that we use is defined by the following steps:

1. Creation of domain models with formal semantics, via ontologies:
   a. user domain;
   b. television-programs domain;
   c. music domain;
   d. weather domain;
   e. …

2. Creation of a recommendation algorithm, which associates instances of the user-domain ontology with instances of the service-domain ontologies (e.g., in the case of television, in the form of personal programming guides). This step includes a certain degree of learning and constitutes the core of the personalization process.
3. Creation of an exploration algorithm (a search engine) for the different services, which takes into account the user profile.

Domains are formalized using ontologies: the static part of each domain is represented by a subset of *DomOnt*, while the dynamic part of service domains is represented by a *ServOnt*. User profiles, i.e. instances of the user ontology, are created through a Component for Creation of User Models (CCUM), which reuses a module developed for the TVFinder system [2]. In Fig. 3, an example of the interface for the capture of user preferences is shown. In this particular case the user is asked to explicitly express his preference with respect to the *ingredients* (such as love, action and suspense) of his favorite TV series. Explicit preferences are just one kind of preferences used by the CCUM; other kinds include stereotypical and implicit preferences.
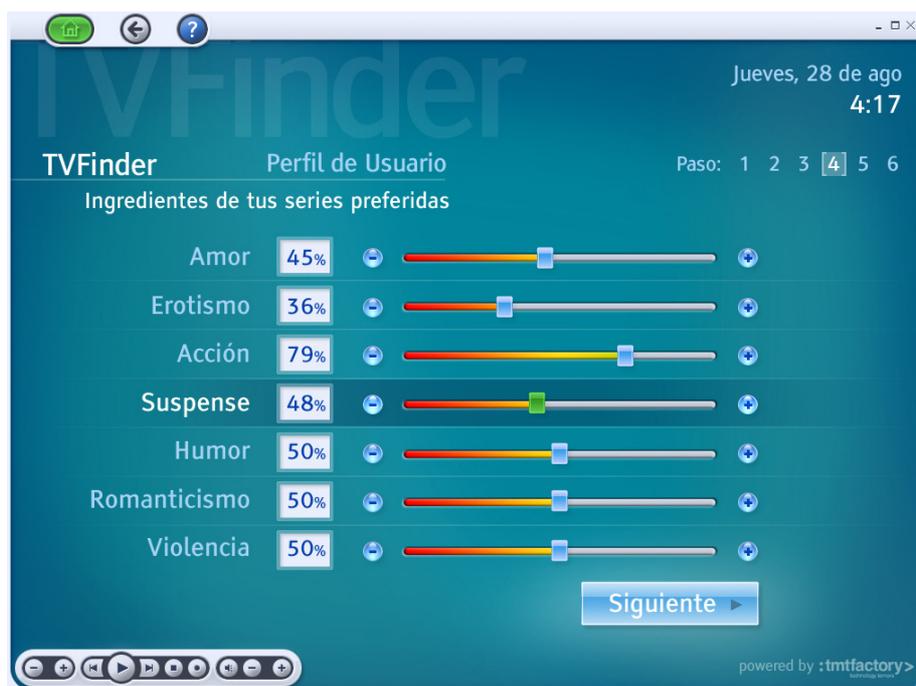

Fig. 3. Interface representing a step in the creation of the user profile (Copyright © 2003, :tmtfactory>)

A code is provided to each user in order to access his personalized interface with the Magical Mirror (e.g., music libraries and TV channels). The identification of the user is important also for improving the voice, image and gesture recognition.

6 Modalities

Magical Mirror can work in three modalities:
- [a] independent;
- [b] connected to the Internet (in particular, to an online data-base);
- [c] connected to the Internet and to a home-automation network.

The data-base includes audio-visual contents and preferences of the users. After a process of initialization and content upload, the way to update the system depends on the modality of the mirror:
- [a] learning;
- [b] updates from a content portal + learning;
- [c] updates from a content portal and the home-automation network + learning.

7 Natural-language dialogue management

Among the necessary technologies to be integrated in the Magical Mirror, so that it is possible to provide the type of personalized services described above, we are building a preliminary speech interface. To this aim, the main artificial-intelligence technique used is an agent for dialogue management, i.e. an agent for natural-language conversation, aware of the profile and preferences of the user.

One of the objectives of the project is that the developed system is able to maintain robust dialogues in natural language. By *robust dialogue* we mean one in which the user of the system can use general expressions (i.e., not restricted to a predetermined language of commands) to access the functions and services of the mirror. Naturally, this objective is excessively ambitious if we pretend that the scope of these dialogues is an ample domain without restrictions. Nevertheless, if the domain is sufficiently restricted, the complexity of the problem is tractable with the present technology of natural language processing. Therefore, a set of dialogue domains has been identified that corresponds with the services offered and is sufficiently restricted to allow such processing.

The architecture of the system is organized around the *VoiceXML framework* and the documents that comprise the *dialogue manager*. These documents are generated with dynamic content coming from the *data bases* of the system: one for the data of the services and one for the user profile. The choice of VoiceXML turns the dialogue manager independent of both the signal-processing technology and the logic used in the application.

The value of the technology being developed does not primarily consist of an advance in each one of the specific techniques of interaction (intelligent processing of the language, agents), but rather in their correct integration to form a coherent whole. Therefore, the process of design and development of a correct architecture of integration of all the modules constitutes an essential part of the project.

8 Task identification by means of artificial vision

As future work, we plan to integrate image-processing techniques which will allow more complex decisions; in particular: (1) identification of the context in which the user carries out her tasks and (2) gesture recognition. The first one consists of determining the number of individuals that are in the room and their arrangement. This information will have to be congruent with the rest of the information received by the home-automation network as well as complementary to it. Other agents can then use this information to adjust the behavior of the network. On the other hand, the analysis of gesture consists of making a suitable segmentation so that the user will be able to control the behavior of the system with clear, natural gestures. This capability of interaction will complement the agent for dialogues in natural language.

Furthermore, the personalization of the interaction will increase the robustness, because the mirror will more easily respond only to the voice and gestures of the person that is logged in to use the Magical Mirror.

9 Use case

A demonstration version of Magical Mirror was shown at various expositions around Catalunya (Spain), in 2003, to check the reaction of the public (see Fig. 4), and received a positive feedback. We hand-designed graphical interfaces on Magical Mirror for interactive television, a multimedia player, weather data and news services. Since we could not yet control the actual applications by voice commands, we simulated control with a laptop (connected to the interface of the mirror) that responded to commands and signals from the users by playing multimedia content that closely matched the response expected of the actual applications. Most people attending the expositions, including kids and seniors, interacted with the mirror speaking in complete comfort.

Fig. 4. Magical Mirror at a public exposition

The average time spent interacting with it was about 5 minutes per person; longer in the case of kids. The attribute that attracted the greatest deal of attention was not the function as information-services provider, but rather the capability of differentiating

among interlocutors. In other words, the high degree of personalization and humanity was the most appreciated feature identified by the public. Visitors could listen to their favorite songs and see images corresponding to their specific requests. The importance of the use of stereotypes was absolutely clear at the moment of the mirror's deployment. About 80% of the visitors requested the same artists and type of images; and these artists corresponded to the ones at the top of the charts of the season. Sport, and specifically soccer, was often the object of search requests.

10 Conclusions

A prototype of a device that incorporates interactive services of leisure and information, offered to the user through a natural interface in form of language and the visual superposition of images on the surface of a mirror, is being developed. The main services, controlled by voice, in Spanish, and presented with the support of an advanced visual interface are: customized weather data, customized news, customized music management, interactive television, Internet access and search. All services are in various phases of research and prototypical development. Nonetheless, a comprehensive prototype, called Magical Mirror, was deployed at various public expositions, where people could interact lively with it. In general, the public showed a positive disposition at adopting a similar mirror at home, given that the price were not much higher then the one for a traditional appliance, demonstrating that people did not consider the apparatus as something completely new and different, and that the additional technology was installed in a satisfactorily transparent way. Additional user studies are in progress to confirm this.

Acknowledgements

References

[1]   Aarts, E. H. L., Roovers, R.: IC Design Challenges for Ambient Intelligence. Proceedings of Design Automation & Test in Europe (DATE 2003), Munich, Germany (2003)

[2]   Ceccaroni, L., Verdaguer, X.: TV finder: una aproximación semántica a la televisión interactiva. Proceedings of the workshop on Ubiquitous computation and

ambient intelligence of the Conference of the Spanish Association for Artificial Intelligence (CAEPIA 2003), Donostia, Spain (2003)

[3]   Dean, M., Schreiber, G., Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D.L., Patel-Schneider, P.F., Stein, L.A.: OWL Web Ontology Language Reference. W3C Proposed Recommendation (15 December 2003)

[4]   Ducatel, K., Bogdanowicz, M., Scapolo, F., Leijten, J., Burgelman, J-C.: Scenarios for Ambient Intelligence in 2010. ISTAG final report for the European commission (2001) Available at [http://www.cordis.lu/ist/istag-reports.htm] (921kb), accessed 23 Apr 2004

[5]   Gustafson, J., Lindberg, N., Lundeberg, M.: The August spoken dialogue system. Proceedings of Eurospeech'99, Budapest, Hungary (1999)

[6]   Ignatius, A.: We have contact. Time, June 4 (2001)

[7]   IST Advisory Group: Ambient Intelligence: from vision to reality. ISTAG draft consolidated report for the European commission (2003) Available at [http://www.cordis.lu/ist/istag-reports.htm] (225kb), accessed 23 Apr 2004

[8]   Lemon, O., Gruenstein, A., Battle, A., Peters, S.: Multi-tasking and Collaborative Activities in Dialogue Systems. Proceedings of 3rd SIGdial Workshop on Discourse and Dialogue, Philadelphia (2002) 113-124

[9]   Lindwer, M., Marculescu, D., Basten, T., Zimmermann, R., Marculescu, R., Jung, S., Cantatore E.: Ambient Intelligence Visions and Achievements: Linking Abstract Ideas to Real-World Concepts. Proceedings of Design Automation & Test in Europe (DATE 2003), Munich, Germany (2003)

[10]   Marzano, S., Aarts, E.: The New Everyday: View on Ambient Intelligence. Uitgeverij 010 Publishers. ISBN: 9064505020 (2003)

[11]   McIlraith, S., Martin, D. Bringing Semantics to Web Services. IEEE Intelligent Systems, 18(1) (2003) 90-93 OWL-S's latest version available at [http://www.daml.org/services/owl-s/], accessed 23 Apr 2004

[12]   Myers, B. A.: Graphical User Interface Programming. In: Tucker, A. B. (ed.): Computer Science Handbook. 2nd edn. CRC Press, Inc., Boca Raton, FL, USA. To appear on 25 Jun 2004, ISBN: 158488360X (2004)

[13]   Nichols, J., Myers, B., Harris, T. K., Rosenfeld, R., Shriver, S., Higgins, M., Hughes, J.: Requirements for Automatically Generating Multi-Modal Interfaces for Complex Appliances. Proceedings of the Fourth IEEE Inter-national Conference on Multimodal Interfaces, Pittsburgh, PA, USA (2002) 377-382

[14]   Philips HomeLab [http://www.research.philips.com/InformationCenter/Global/FArticleDetail.asp?lArticleId=2753&lNodeId=931&channel=931&channelId=N931A2753], accessed 23 Apr 2004, is part of Philips Research Laboratories

[15]   Philips Research [http://www.research.philips.com], accessed 23 Apr 2004. Prof. Holstlaan 4, 5656 AA Eindhoven, The Netherlands

[16]   SmartKom project: Dialog-based Human-Technology Interaction by Coordinated Analysis and Generation of Multiple Modalities [http://www.smartkom.org], accessed 23 Apr 2004

[17]   TRINDI project: Task Oriented Instructional Dialogue [http://www.ling.gu.se/projekt/trindi/], accessed 23 Apr 2004

[18]   Zue, V., Seneff, S., Glass, J., Polifroni, J., Pao, C., Hazen, T.J., Hetherington, L.: Jupiter: A Telephone-Based Conversational Interface for Weather Information. IEEE Trans. on Speech and Audio Processing, 8 (1) (2000) 85-96