GENIO: An Ambient Intelligence application in home automation and entertainment environment

Alfonso Gárate[1], Nati Herrasti[2], Antonio López[2]

(1) Fagor Electrodomésticos, Barrio San Andrés 18, 20500 Mondragón, Spain
agarate@fagorelectrodomesticos.com

(2) Ikerlan, Paseo José Mª Arizmendiarrieta 2, 20500 Mondragón, Spain
nherrasti@ikerlan.es; alopez@ikerlan.es

Abstract

As one of the main aims of "Ambient Intelligence" is natural human interaction with the environment and one the most suitable is the home environment. Fagor has been working several years to develop a power-line network where all its household appliances, security sensors and actuators, heating systems and antiintrusion systems are connected and managed by a central controller named Maior-Domo. As a result of GENIO project, the user can dialog with his home and asks the services and functionalities he wants by talking as he was talking to a friend. The controller Maior-Domo has a human representation that the user can see and can interact with. When the user talks the Maior-Domo extracts the different commands from those vocal orders and controls the home devices. These orders are not specific commands that the user has to learn but natural speaking language without any need of learning. In the same way, any event or information from any device of the network is transmitted to the user by voice. In order to achieve a demonstrator of this Ambient Intelligence application, a real kitchen and sitting room have been built where the users can command the home talking naturally (in Spanish). Possible actions are: reading e-mails, programming the washing machine, checking the goods in the fridge, creating the shopping list, doing shopping with a PDA in the supermarket, activating the dishwasher, being guided on how to prepare a recipe for the oven checking if there are the needed goods to do it, listening some music stored at home, watching some photos, watching some selected video and so on. Every user has a wireless microphone in his/her shirt's pocket. This microphone captures his/her voice and all the sounds around him/her and sends them to a developed board which filters the voice frequency range from other sounds. From here the voice recognition system "understands" the pronounced sentence and process it. A quite extended number of sentences, called grammar, make up the possible dialogue between the person and the whole system. The user can address the whole system in different ways using a lot of expressions, talking naturally and spontaneously and dialoguing to the home. The defined grammar is so extended that almost the total speaker independence has been achieved.

1. Introduction

An environment may be termed as "Ambient Intelligence" when it is not intrusive, where diverse technologies complement each other so that, surrounding the users with said environment [15], offering as many services and features as required or as predictable, in as many environments as the users may have available.

Thus, an Ambient Intelligence environment, with a technological network surrounding those inhabiting it, will be able to:

- Surround to the users in a non-intrusive way.
- Recognise users and their circumstances (activities, state of mind, etc.) and operate consequently, i.e., be sensitive to human presence.
- Have a predictive behaviour based on knowledge of the environment (context awareness), of the habits of those who it is "serving" and of the specific activities of the same when acting.
- In real time, produce new services in fields such as entertainment, security, health, housework, the work environment, access to information, computing, communications, etc., to improve the quality of life by creating adequate atmospheres and functions.
- Allow access to as many services and features as it can carry out, regardless of where the user is located, of the position from where the user demands said services and the artefacts available at that particular moment (ubiquity).
- Relate, in a natural manner, to the users by means of multi-modal voice-based interfaces; by reading movements and gestures; by generating, emitting and projecting images; by generating holograms, etc. (natural relationship).

Home is the perfect place to apply Ambient Intelligent precepts and technologies for giving high-level services to the user and allowing him to address and command his home by natural human interfaces like the voice and dialogue [6], mainly taking into acount that one of the greatest human aspirations is to relate naturally with his surroundings, and this includes devices and machines ascribed to said environment. Being there no other more natural means of relating than with speech, voice-processing technologies become considerably important in the development of Ambient Intelligence applications [2][5].
Fagor Electrodomésticos, with the technological support of Ikerlan, has developed a first "Ambient Intelligence" application implementing two of these previously mentioned preceptors: ubiquity and natural relationship [7].
In this application the most relevant activities in a home (household appliances control, entertainment and telework) are implemented by several new services that are controlled by means of a conversational voice based interface.

2. Application

As main infrastructure, a network with the most sophisticated Fagor household appliances, based on power-line as communication medium, has been developed.

The main product in this network is the Domotic Controller, called Maior-Domo, which is able to receive natural voice orders from the users and to dialogue with them in a conversational way [8].
The Maior-Domo [1] extracts the different commands from those vocal orders and controls the home devices. In the same way, after receiving any event or information from any device of the network, the Maior-Domo generates a speech message to the users in order to inform them about it.
In order to achieve a demonstrator of this Ambient Intelligence application, a scenario has been built where the users can command his home talking naturally. A demo has been implemented to show some actions executed during a day by a user at home. These actions are: reading e-mails, programming the washing machine, checking the goods in the fridge, creating the shopping list, doing shopping with a PDA in the supermarket, activating the dishwasher, being guided on how to prepare a recipe for the oven checking if there are the needed goods to do it, listening some music stored at home, watching some photos, watching some selected video and so on.

2.1. Infrastructure: Devices and elements involved

The infrastructure and devices involved in the application are the following:

- The prototype of a new oven, called Conect@, which has a database of recipes inside and allows to prepare a recipe just putting the food inside because it knows how to command the sequence of temperatures, times and the corresponding heating method. The user can download new recipes from Fagor's web using the computer of his home, which sends a small adaptation of the recipe to the oven. This oven has a small display that shows the process and menu's tree to the user.
- A common fridge, which has a RFID antenna and a reader inside to read the goods stored in it. Each product has got a smart label attached to it.
- A big panel which contains the electronic board of a washing machine and a dishwasher. These boards are the same as having the real appliances.
- A computer working as the Maior-Domo, which can communicate by power line with the oven, the washing machine and the dishwasher, by WiFi with some PDAs and by radio with the pocket microphone (it seems a pen like a fluorescent marker) that the user carries.

Figure 1: Microphone

- A RFID reader and the voice reception system are attached to the Maior-Domo. In this computer the following modules are located: a voice recognition software ASR (Automatic Speech Recognition), a TTS (Text To Speech) and an application to command and control the home.

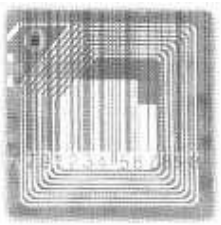

*Figure 2:*RFID tag

- The voice application is able to recognize complete sentences from a wide grammatical range and has been developed by means of VXML, JAVA, JSP and JavaBeans. Also, this computer stores all digital information like songs, videos and photos and the context awareness information needed for controlling the whole application.
- A 63" Plasma TV to show different screens to enforce the voice user interface, mainly in the entertainment domain; a projector located in the kitchen to project on the worktop or on the furniture surfaces.
- All these home devices are part of a demonstrator room where the ubiquity and the natural relationship, both of them as Ambient Intelligence preceptors, have been implemented in a real life application [4]. The ubiquity because the users are able to access to different services from any place at home, thanks to the wireless connection between the pocket microphone and the voice recogniser.

Figure 3: Demonstrator set

2.2. Services: Features and benefits

This paper presents the partial results of the project GENIO (Genie in Spanish, with the meaning of being the server of your home, who executes user's desires). The services implemented in this demonstrator are:

Reading e-mails

The user is able to ask for how many new e-mails he/she has, using the most common sentences for that such as "Maior-Domo, tell me how many e-mails I have" or "Maior-Domo, do I have any new e-mails", among many others. The Maior-Domo answers telling how many messages there are and the

sender of everyone. The user can choose which one he would like to listen to and the Maior-Domo will read it.

Activation, deactivation and programming of an electronic washing machine and a dishwasher

The user is able to switch on and switch off any one of these appliances, talking to the Maior-Domo by means of most usual expressions ("turn on" or "switch on" or "put on" the washing machine; "please, wash the clothes", etc.). In the same way the user can decide a programming time to turn on anyone of these appliances holding a dialogue like this:

- User: "Maior-Domo".
- Maior-Domo: "Tell me".
- User: "Wash my clothes by the time when I come back from work".
- Maior-Domo: "When are you going to come back?".
- User: "At eight o'clock in the evening".
- Maior-Domo: "Which washing program?"
- User: "Soft, 30ºC."
- Maior-Domo: "Washing machine programmed."
- User: "Show the programmed process in the TV".
- Maior-Domo: "Ok".

Checking the goods in the fridge, creating the shopping list and doing shopping with a PDA in the supermarket

Every goods used in this application have a RFID label in order to identify it. A RFID antenna has been installed inside the Fagor's fridge. A RFID reader is part of the Maior-Domo architecture. The user, always talking in a natural way, can ask for the list of the goods in the fridge.
The user defines his basic shopping list with the goods he always wants to purchase. When the Maior-Domo detects that some of this basic goods is not present in the fridge, it adds automatically this good to the shopping list. So, the user doesn't have to be aware of this. Moreover, at any moment, the user can address to the Maior-Domo asking for adding other goods or removing any current ones.
Before the user goes to do shopping, he asks the Maior-Domo to download the shopping list to his PDA, which will be used as an assistant in the supermarket. The PDA shopping application is also supported by images and voice interface in order to guide the user. This fridge application can be extended to any kitchen storage furniture.
The most common dialogues are possible in this context, for example:

- User: "Tell me the shopping list."
- Maior-Domo: "Right now there are chicken, carrots, champagne, olive oil, softener and detergent. Anything else?"
- User: "Add paper napkins to the list."
- User: "Tell me the shopping list."
- Maior-Domo: "Right now there are chicken, carrots, champagne, olive oil, paper napkins, softener and detergent. Do you want anything else?"
- User: "Download it."
- Maior-Domo: "Shopping list downloaded."

Preparing a recipe

The oven Conect@ has a database of recipes inside it. When the user wants to cook a recipe, he asks the Maior-Domo to check if all the needed elements are present in the fridge. If it's so, the user and Maior-Domo start a dialogue to help the user in the preparation of the recipe before cooking it. The user is guided in each step in the elaboration of the recipe at the pace he requires.

An example of this is:

- User: "MaiorDomo."
- Maior-Domo: "What do you fancy?"
- User: "What do I have in the fridge?"
- Maior-Domo: "You've got apples, onions, yoghurts, milk, boiled ham and champagne in the refrigerator. That's all. Anything else?"
  A piece of chicken and some carrots are put into the fridge by the user.
- User: "What do I have in the refrigerator?"
- Maior-Domo: "You've got chicken, apples, onions, carrots, yogurts, milk, boiled ham and champagne in the fridge. That's all."
- Maior-Domo: "Would you like anything else?"
- User: "Show me the recipe for roast chicken."
- User: "I want to see the next page."
- User: "Go on."
- User: "Show me the ingredients."
- Maior-Domo: "For how many people?"
- User: "Four people."
- User: "Go step by step through the cooking process."
  (Maior-Domo reads the recipe's first line)
- User: "Done."
  (Maior-Domo reads the recipe's second line)
- User: "OK."
  (Maior-Domo reads the recipe's third line)
- User: "Sorry?"
  (Maior-Domo repeats the recipe's third line)
- User: "Done."
  (Maior-Domo reads the recipe's fourth line)
- User: "OK."
  (Maior-Domo reads the recipe's fifth line)
- User: "All right."
- Maior-Domo: "This is the last step of the recipe. Turn the oven on whenever you want."
- User: "Go cook the chicken."

Entertainment

Of course, the home is getting more and more a place for leisure time. Therefore, Ambient Intelligence and more precisely natural voice interfaces have to be involved in this circumstance. In this context, the user can have a lot of digital information inside the home or coming from outside, like films, music, photos and so on. In this project this field has been covered using the database stored inside the home.
One of the implemented dialogues is:

- User: "MaiorDomo."
- Maior-Domo: "What do you fancy?"
- User: "Do I have champagne?"
- Maior-Domo: "Yes, you do."
- User: "I'd like to listen to some music."
- Maior-Domo: "What would you like to listen to?"
- User: "Something by Mozart."
- Maior-Domo: "You've got the following: "March alla Turca", "Symphony 40th" and "A little night music"."
- User: "A little night music."
- User: "MaiorDomo."
- Maior-Domo: "What do you fancy?"
- User: "I'd like to see some photos"
- Maior-Domo: "You've got photos from the Olympic games, from the wedding and from the Tibet. Which ones would you like to see?"
- User: "The ones from the Tibet."
- User: "Stop it."
- User: "Continue."

- User: "End it."
- User: "MaiorDomo."
- Maior-Domo: "What do you want?"
- User: "Switch off the music."
- User: "I want to watch the video of Harry Potter."
- User: "MaiorDomo."
- Maior-Domo: "What would you like me to do?"
- User: "Switch off the video."

3. User interfaces

The voice recognition system works perfectly if the voice arrives isolated from external noise. Every user has a wireless microphone in his/her shirt's pocket. This microphone captures his/her voice and all the sounds around him/her and sends them to an equalizer, which filters the voice frequency range from other sounds [9][10].

From here the voice recognition system "understands" the pronounced complete sentence and process it. A quite extended number of sentences, called grammar, make up the possible dialogue between the person and the whole system. The user can address the whole system in different ways using a lot of expressions, talking naturally and spontaneously and dialoguing to the home. The defined grammar is so extended that almost the total speaker independence has been achieved. This has been proved by several speakers (young and adult males and females) talking to the system without previous training.

Thanks to this natural dialogue, the user doesn't have to learn any kind of commands, doesn't need remote controllers for each one of the devices, feels his home more close to him and has the perception of having everything under control. We would like to stress that the class/style files and the template should not be manipulated and that the guidelines regarding font sizes and format should be adhered to. This is to ensure that the end product is as homogeneous as possible.

To complement the voice dialogues, a visual interface has been developed in order to get a multimedia and multi-modal environment. This interface can be displayed in several home screens, TVs, Tablet PCs, mobile phones, PDAs and by means of projectors.

4. Ubiquity

Thanks to both the voice capturing method used (a pocket microphone like a pencil) by each user involved in this application and the display network (screens and projectors distributed at home), the ubiquity is achieved. The user's voice orders will be handled by the Maior-Domo and the system information will be received by the user regardless where he/she is located at home.

5. Future activities on this application

Following the current developments, these are the next steps:

- To develop a wireless network of microphones integrated in the home environment (some of them in each room) as the way to capture the voice of the users and to isolate it from the stationary and the accidental noise [3]. This will be done thanks to microsystem technology.
- To get the total speaker independence enlarging the grammar and digitalizing any kind of human voice in order to offer to the system a non rejectable and easily understandable voice signal.
- To obtain an absolute speaker dependence for some specific services like security and safety applications. The challenge is to avoid that any user, for example a child, to be able to turn on the oven or to activate the anti-intrusion system. Voice verification techniques will be implemented.
- In this future home groups (families, students, …) will live in and will test the final prototypes with the purpose of getting the impressions and suggestions from real users.

6. Ambient Intelligence immediate challenges

The total implantation of Ambient Intelligence systems in the domestic environment described above involves obtaining products (product networks) based on their rules. Developments designed to obtain this should focus on the following initial steps:

- Development of hardware devices (generically termed as artefacts) able to detect whatever is relevant of a given environment (physical magnitudes, presences, locations, states of mind, etc.) These devices will be effective if, apart from implementing the features that define them, they are non-intrusive or, what is the same, lack a physical presence (small dimensions, easily mimicked and able to be integrated into everyday items, self-feeding, with capacities to be integrated into wireless networks, etc.).
- Obtain computational artefacts with "physical" characteristics similar to those described above.

Figure 4: Maior-Domo representation. The avatar.

- Improve performance in voice recognition systems and understanding of the spoken language in order to obtain natural, intuitive man-environment interfaces, with capacities for conversation.
- Development of software (middleware, framework) to facilitate the Ambient Intelligence paradigm and which works as a platform on which to install diverse application (different suppliers and manufacturers; different types of product, etc.) those make up an Ambient Intelligence system.

7. Conclusions

The impact of Ambient Intelligence on developments in the technological products over the next few decades promises to be colossal. To the same extent as the influence that these new products will exert on the lives of those who have them. Their proliferation in the West will generate, or at least will inspire,

a new social change of unpredictable proportions, although it may be supposed that it will have a bearing on the very concepts of globalization and on extending access to welfare.
The standard of living, taken as a natural access to numerous features and services that relieve us of tiresome tasks, giving us more leisure time while drawing us closer to culture and entertainment will increase, in a manner that we cannot imagine for the moment, and will be attainable for practically all social classes. The welfare and leisure society will be nearer removing, or at least, toning down the social stratifications that still persist.
Although the society as a whole will benefit from the technological advancements brought by Ambient Intelligence, it will especially be people with disabilities and the growing third age population who will most improve their standard of living, thanks to future.

References

[1] A.J.N. van Bremen, K. Crucq, B.J.A. Kröse, “A User-Interface Robot for Ambient Intelligent Environments”. 1st International Workshop on Advances in Services Robotics, Bardolino (Italy) 2003.
[2] Berry Eggen, Gerard Hollemans, Richard van de Sluis, “Exploring and Enhancing the Home Experience”. Cognition, Technology and Work. Springer-Verlag, 2002.
[3] Emile Aarts, Raf Roovers, “IC Design Challenges for Ambient Intelligence”. 2003 IEEE.
[4] Evert J. van Loenen, “On the role of Graspable Objects in the Ambient Intelligence Paradigm”. Media Interaction Group Philips Research Labs, Eindhoven, The Netherlands.
[5] ITEA Ambience project, http://www.extra. research.philips.com/euprojects/ambience/.
[6] Julio Abascal, Antón Civit, Klaus Fellbaum,” Smart Homes - Technology for the Future”.
[7] K. Ducatel, M. Bogdanowicz, F. Scapolo, “That’s what friends are for. Ambient Intelligence (AmI) and the IS in 2010”.
[8] K. Ducatel, M. Bogdanowicz, F. Scapolo, “SCENARIOS FOR AMBIENT INTELLIGENCE IN 2010”. European Commission Information Society Directorate-General. Febrero 2001.
[9] L. J. Rodríguez, I. Torres, A. Varona, "Annotation and analysis of disfluencies in a Spontaneous speech corpus in Spanish".
[10] M. I. Torres, "Reconocimiento del habla en sistemas de diálogo. El procesamiento del lenguaje y del habla en los sistemas de diálogo”. J. Llisterri ed. FDS, 2003.