Abstract. In this paper, a set of new advanced interaction interfaces, will be shown. These prototypes have been integrated into our Ambient Intelligence Room, where the devices are controlled by these natural interfaces based in the interaction via vision and/or voice. The elements included in this first prototype of the AmI room are the following ones: a wearable gaze tracking system, an advanced gesture interface and a voice command module. Developments include the identification of the user based on different kind of biometric-systems as fingerprint hand-print and hand outline and face recognition. All these systems can be used independently or in a combined way. In summary, in this paper, the different elements composing the present AmI Room will be explained. Furthermore, it will be described the interaction among these elements, as well as the different applications that can use this kind of systems.

1 INTRODUCTION. The advances achieved in the Information Technologies field have had an important impact in the current way of living, working and in the way the users can interact with the computers and other electronic equipment. During last decade, we has been developing technologies intimately related to Ambient Intelligence such as Intelligent Interfaces, Biometrics, Image Processing , Digital Security, Knowledge Technologies, Production of Software, Wireless systems…

This previous know-how allows us to integrate these technologies into the vision proposed by Ambient Intelligence, for providing more natural and advanced interfaces which can make easier the communication between users and computers. Systems that will allow to use multimodal interfaces in a similar way to the human ones as speech or gestures.

Keeping this ideas in mind, we are working in the development of new gesture interfaces which allow the communication by an easy and natural way between the user and the computer. This communication can be both by means of gesture movements and by means of the voice. Initially, several prototypes have been developed which add different existing technology branches in this sector, working with both voice recognition systems, and machine vision systems capable of tracking the eye movements or recognizing the hand gestures, working in laboratory controlled environments. All of them are made keeping always in mind the way to be able to implement them in the industrial sector, optimizing these prototypes to provide them with the necessary robustness and effectiveness for this kind of industrial environments.

The layout of this paper is the following: it will start with a short description of the evolution of advanced interfaces. Then, a description of the developments made by us will be presented, finishing with the description of the AmI Room and the future work to be done in order to go forward to a more natural interaction between users and computers.
2 BACKGROUND. During the last years, different researching groups have been working in new communication methods between humans and computers. First steps were gave with the incorporation of several external elements, such as touchscreens, haptic gloves… Nowadays, new systems are designed to achieve human-machine interaction without any intermediate equipment. These new systems use different input and output systems, such as voice commands, hand movements, hand gestures or gaze tracking.

In the case of gesture, they have not been integrated into the usual devices, because recognizing gestures is a complex task which involves many aspects such as motion modelling, motion analysis, pattern recognition and machine learning. Indeed, most of the gesture recognition work is limited to understanding artificially imposed signs and gestures. This means, that gesture based systems are difficult to operate, because in the end, the user is using a predefined hand gesture grammar that have to be learnt. This training converts what would be a natural communicating system into a forced one.

So, a good gesture interface would be able to answer in a efficient way to the user commands that are given by natural movements. The interface would be the one which would have to adapt to the users communication way.

The gesture-based systems can be classified, attending to their relationship to the speech (Efron (1), Kendon (2), McNeill (3)), into:

- Deictic gestures, that are used to focus the user’s attention to a physical reference in course of a conversation.
- Iconic and metaphoric gestures, that are associated with abstract ideas, mostly peculiar to the subjective notions of an individual.
- Beats serve as gesture marks of speech pace.

Voice interaction systems developments have been carrying out for a long time. There have been developed several applications capable of recognizing a determinate set of commands to manage data, for example, text processors within the computer. However until now, these applications are very sensitive to environment noise and they require quite a lot of training to get a successful recognition.

3 DEVELOPMENT OF ADVANCED INTERFACES: EVOLVING APPROACH. FROM STAND-ALONE SOLUTIONS.

Most developments accomplished in the area of advanced interfaces are devoted to solve individual aspects of the human communication. This approach provides accurate solutions to specific problems but limit then application domain of such systems. Final users look for human-like interfaces where different communication methods can be used in an stand-alone or in a combined way, and independently of the situation or the used equipment.

For this reason, our developments have evolved from stand-alone solutions that focus on the technical problems to be solved, to a integrating approach, centered in user needs. This new approach has driven to the creation of an Ambient Intelligence Laboratory that provides a global solution that allows users to interact with any application using the same set of integrated advanced interfaces.

Following, the stand-alone solutions for different kind of interfaces are presented, to finalize in the next chapter with the description of the AmI Laboratory.

3.1 Gaze Tracking System. Within the diverse forms of human communication, the sight, constitutes one of the most developed senses. Before acting, human beings observe their surroundings looking for elements to get information from the environment. The sight focuses
in the object of interest before deciding the action to take. The process of focusing in an object in movement is, in the best case, more than twice faster than the same process using a classical electronic device such as a mouse.

Based on these premises, we are developing a system that approaches the analysis, study and development of a set of methods. It will allow to locate and to follow the gaze of the user in order to retrieve information, to analyze their points of interest and to give back augmented information of the environment to the user.

![FIGURE 1. Possible application for Gaze tracking](image)

This prototype (Renteria et al. (6) ) has been designed to greatly improve the portability of this sort of systems. This will allow to use them in open environments and also to increase the application fields of these type of systems.

The prototype consist on a mini wireless IR camera that sends the captured images to a processing PC, which process the image and locates the pupil. One of the possible applications that this kind of systems can have is its use as a alternative mouse for the computer. The user can move the cursor with his eyes and even perform the events associated to the left and right mouse buttons by a simple action of closing the eye.

At this moment, we are working in reducing the size of the system, adapting algorithm processing to run into smaller devices, such as a notebook or even, into a PDA that user can wear easily.

![FIGURE 2. Eye tracking program](image)
3.2 **Advanced Gestual Interface.** One of the first works developed by us in this field was made in collaboration with the Engineering School of University of the Basque Country in Bilbao and consisted on an static hand gesture recognition system (Isasi (7)) applying an Optical Character Recognition (OCR) algorithm. The idea of this method came from the fact that a static hand gesture has the same properties than a written character has. This system is formed by an image acquisition camera to obtain the hand gesture image, a processing computer, to analyze the hand position and provide the control command.

![Diagram](image)

**FIGURE 3.** Hand gesture recognition diagram

The different command hand gestures that the application is able to distinguish are:

![Hand gestures](image)

**FIGURE 4.** Command hand gestures

A computer interaction prototype based on hand movements has also been implemented. This prototype is the first step to get an interactive communication method between the user and the computer using a natural language based on body language and voice commands.

In this first step, an advanced user interface based on two different color hand markers has been developed. With this system it is possible to manage two different cursors in the computer, being one of the markers assigned to the right hand and the other to the left one.
In the next phase of this development, the commands given by the marker state changes are going to be changed by hand gestures, first in 2D, as the system works now, but thinking in the possibility of working in 3D spaces using stereovision.

Up to now, there are some research groups working with 2D spaces and video sequences analyze using Hidden Markov Models (HMM) (James et al. (9), Lenman et al. (10), Krum et al. (11), LaViola (12) ) but systems using 3D models of hand gestures need to use electronic sensors as haptic gloves or magnetic fields sensors (Corradini and Cohen (8) ).

FIGURE 5. Gesture interface diagram

3.3 Mouse emulator. Other modules developed by us are a hand-based mouse emulator. The Windows mouse can be controlled by the above described interface, using the hand movements. With this application the user is able to perform all of the events associated to the Windows mouse, using only his hand. The user can send any mouse event to the computer such as press the left mouse button, release it, press or release the right mouse button, move the mouse wheel… opening a new application field to the gestual interface system.

FIGURE 6. Gesture interface program

3.4 Voice recognition module. This system allows a “human-like” communication between the user and the whole AmI Room. When entering the room, the user authenticates himself using biometric systems within a PDA device and his voice profile is activated in the room. Once the
user privileges are checked, he can control the room by means of voice commands such as activating recognition methods, enrolling users, authenticating users, using different applications…

All the stand-alone applications developed within the AmI laboratory are being integrated with the voice command and gestual system so that a completely integrated system can be shown in the laboratory.

4 …TO THE AMBIENT INTELLIGENCE LABORATORY. As mentioned above, we have created an Ambient Intelligence Laboratory (AmI Lab), in which different kind of developments are integrated, from biometric-based security mechanism, to advanced smart interfaces, infomobility systems… This laboratory is growing everyday in a progressive way.

At the beginning, the biometric and security applications were the first components of the laboratory, counting with a fingerprint recognition application and a hand authentication system, both of them controlled by voice commands and totally own developed. After that, portable devices such as PDAs where used to to monitor all the events produced in the laboratory. In addition and in collaboration with the Engineering School of the University of the Basque Country a face recognizing system was added to the security system.

Now, all the previously described applications have been implemented in the AmI Lab so that all of them could be connected. A laboratory server application has been developed to listen to all the applications that are running in the AmI Lab. This program is attending all the messages that the different applications are sending, and it is in charge to distribute the message to the appropriate listener. With this method, it is easy to integrate new applications into the laboratory and connect them to the rest of the system, for example, to implement the advanced gestual interface with the rest of programs, without doing any change in the own laboratory architecture. Applying the same method, the warehouse video security system, and the mouse emulator program have been easily integrated in the Ambient Intelligence Laboratory that now has different applications running and being managed by voice (thanks to the voice recognition module) and by gestures (thanks to the gestual interface).

5 FUTURE WORK. Up to now different prototypes for new advanced interfaces using voice, gaze and gestures over different application scenarios, such as industrial warehouses, have been developed. However, one of the problems that have found during these prototypes development
is that users are very used to use the traditional devices, such as mouse and keyboards, and now these new devices, using the hand movements, are uncomfortable and strange for them. This is, firstly, because until now, the developed systems impose to the user a closed grammar to establish the communication with the computer, and this grammar is not natural or intuitive, having to do the hand gesture in a predefined order.

A long way has to be covered until this kind of interfaces are comfortable enough for being broadly used. Gestual interfaces must be as natural and intuitive as possible for people to use them and current solutions are still in its early stages.

For this reason, ongoing works try to create a more intuitive gesture grammar, in order to provide an easier to use and more comfortable operating interface to the user. Other issues to be solved are environmental “noise” when using this kind of systems, It is easy to control voice commanded systems or gestual interfaces when used in a laboratory, but a lot of problems arise when used in real scenarios such as workshops, markets or open spaces. Noise, changing light, multiple people in the interaction area and other issues have to be taken into account and controlled for the proper operation of the system.

Different scenario conditions are being studied to collect the needed requirements, so that prototypes can be improved to meet their expectations.

6 REFERENCES.


