Ibero-American Symposium on Computer Graphics - SIACG (2006)
P. Brunet, N. Correia, and G. Baranoski (Editors)

Non-verbal communication by means of head tracking

C. Manresa-Yee[1] and J. Varona[1] and T. Ribot[1] and F.J. Perales[1]

[1]Unitat de Grà cs i Visió,
Dept. de Matemàtiques i Informàtica,
Universitat de les Illes Balears,
Palma de Mallorca, Spain

Abstract

Physical and mental disabled people are an important group of our society that haven't yet received the same opportunities as the others in their addition in the Information Society. Therefore their inclusion within the new technologies should be a must-do task. This paper describes the use of a vision-based user interface that replaces the mouse for interacting with a communication system of non-verbal children.

Categories and Subject Descriptors (according to ACM CCS): H.5.2 [HCI]: Input devices and strategies

1. Motivation

Information and Communication Technologies (ICT) are present in many of our daily activities. Even if information systems are regularly used in education, work, leisure or domestic purposes by the majority of modern societies' citizens, still some sectors such as disabled persons are at a disadvantage. But, nowadays our societies are paying more attention to physical and mental disabled people, as they form a big group of users that should be given the same chances as the others. Different efforts have been made in diverse areas, but the sector that concern us in this paper are the new technologies and their applications in bene t of all users. The *e-Accessibility* component of *e-Inclusion* is an important goal to accomplish for building societies where the guarantee of opportunities is equal for everyone. Research in new technologies directed to handicapped (physical or mental) people can help in improving their quality of life and to promote their integration in our society. Projects like *eEurope* remark that disabled persons should be taken into account and receive special attention for avoiding info-exclusion and for assuring their addition in the Information Society [eEu05]. Consequently, innovation in new technologies should have in mind these two issues: e-inclusion and e-accessibility. Challenges to beat are the development of new technologies and systems accessible for everyone and their application for offering assistive technology [RR03]. This can help people to lead more independent lives and to help them to play a more active role in our society.

The creation of non-invasive and more natural human-computer interfaces based on speech recognition or computer vision techniques can offer upper-body physical disabled persons an easier interaction with computers, rather than using a standard mouse or keyboard [KT05]. Mental handicapped people also can be included in the use of new technologies. One rst step to help them to get used to the new technologies is to involve the traditional non-digital systems that are used for education purposes into the new technologies.

This work presents a contribution of a system that can improve the social integration of children with physical disabilities and speech communication problems. The paper is about a multimodal user interface that replaces the functions of a standard mouse by means of head tracking and face gestures recognition working together with an educational application for teaching non-verbal children: *BlissSpeaker*.

This paper consists in three sections. The functionality of the visual-based interface is explained in the rst section. Although the feature tracking and gesture recognition system has been developed by the authors, this paper focuses on its application rather than in explaining the technical implementation details that can be found in other publications [MYVP06b]. Then its use in conjunction with *BlissSpeaker* is de ned in the next section. The last section concludes the paper and re ects the future work that can be done.

2. The visual-based interface: *HeadDev*

Part of the presented multimodal interface is a visual-based interface (VBI) that uses computer vision techniques to achieve a system that ful ls completely the functions of a standard mouse and replaces it by means of face feature tracking and face gesture recognition [MYVP06a]. It is focused on handicapped persons with limitations in the upper-body limbs, and it offers them a natural, non-invasive and low cost interface with the computer, that can help them in their integration within the society that involves them.

For its operation, only a standard USB webcam is needed for providing the images to process. Therefore it will allow the achievement of a low cost system and the user will not need any kind of extra device or cumbersome systems on him, that is, the user will sit in a comfortable position facing the screen and the webcam will be at the height of his face. See Figure 1. Other system's requirements is that the user's work environment conditions should be normal (of- ce, house or indoor environments), that is, with no special lighting or static background.

Figure 1: *Hardware con gur ation.*

In order to achieve an ef cient human-computer interaction, the system's feedback must be in real-time and it must be precise and robust. In this case, precision means that the user is able to place the mouse over a desired position and to carry out the intended event.

The implementation of this interface is based in a new mixture of several computer vision techniques, where some of them have been improved and enhanced to reach more stability and robustness in tracking and gesture recognition.

The system is divided in two modules: Initialization and Processing. In the Initialization module the system detects automatically the user's face and the best features in it. In order to achieve this process automatically, the user has to sit still for a few frames. In the Processing module the features found in the Initialization module are tracked for controlling the mouse's cursor position and by means of eyes' winks detection and recognition, different events of the mouse can be carried out. Due to illumination changes or fast movements of the head, the features can get lost, but the system will reinitialize itself if this occurs. In Figure 2 the process of each frame it is described and several images of the system are shown in Figure 3. Detailed technical information on this vision-based interface can be found in [MYVP06b].

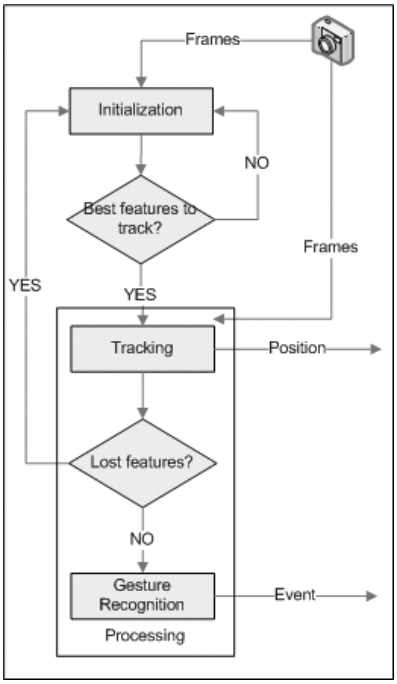

Figure 2: *HeadDev's diagram.*

Figure 3: *HeadDev in process.*

3. The non-verbal communication application

HeadDev was created as a result of an idea that came from an organization of disabled people. Its first intention was to assist people in their accessibility to a computer giving up all kind of use of cumbersome orand intrusive systems. Then, as accessibility was already achieved, a second aim was requested: to bene t other disabled users by combining *HeadDev* with other useful applications. With *HeadDev* the interface between human and computer was sorted out, so the next ambitious objective was to improve the human-to-human interface or communication for people with speech problems using the new technologies.

There are different augmentative communication systems for people with speech limitations, ranging from unaided communication such as American Sign Language, to computerized iconic languages with voice output systems such as Minspeak^{TM} [ACP98].

Our work is based in a symbolic graphical-visual system for non-verbal communication named Bliss [Bli06]. Bliss can be used as an augmentative system or for replacing completely verbal communication. It is commonly used by persons with cerebral palsy but with the next learning aptitudes requirements:

1. cognitive abilities,
2. good visual discrimination,
3. possibility of indicating the desired symbol and
4. good visual and auditory comprehension.

Some speech therapists use it in their sessions for communicating themselves with children with speech problems and to help in the prevention of linguistic and cognitive delays in crucial stages of a child's life.

The Blissymbolics language is currently composed of over 2000 graphic symbols and they can be combined and re-combined among them to create new symbols. Since the number of symbols is adaptable to the capabilities and necessities of the user, our system, *BlissSpeaker*, has 92 symbols de ned, that would correspond to the first set of Bliss symbols for the pre-school children [War85]. In Figure 4 a system's process diagram is shown. *BlissSpeaker* is an application that reproduces verbally the statements built using Bliss symbols, which allows a more "natural" communication between a Bliss child and a person that doesn't understand/use these symbols, in example relatives of the children. The application can work with any language, as long as there is an available compatible SAPI. The potential users of *BlissSpeaker* are children with speech problems; therefore its operation is very simple and intuitive. The users select the symbols for building a phrase and reproduce it. In Figure 5 part of the symbols de ned in *BlissSpeaker* are shown and a selection of them for reproducing it with audio.

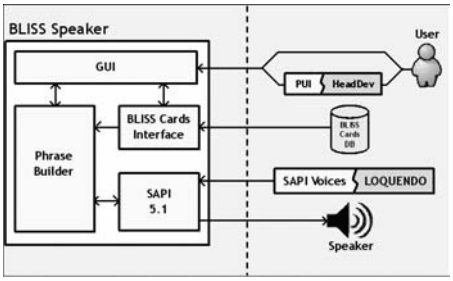

Figure 4: *BlissSpeaker's diagram.*

Figure 5: *Subset of Bliss symbols.*

Audio, vision and traditional graphical user interfaces combined together con gure a very appealing multimodal interface. Moreover, the use of *HeadDev* with *BlissSpeaker* will help to ful l the third requirement of a Bliss user, that is, the possibility of indicating the desired symbol. It will offer children with upper-body physical disabilities and speech dif culties a way to communicate themselves through an easy interface and their teachers or relatives will understand them better due to the symbols' reproduction. Furthermore, although the user is only in need of *BlissSpeaker*, the use of the new interface can make more enjoyable and entertaining the learning of Bliss language and it also promotes the children's coordination, because the interface works with face motion.

The main advantage of this digital Bliss system is to free the user of the physical Bliss cards and their inconveniences, such as the display or the need of repetition of a card. Furthermore, as audio is used to reproduce the cards, the under-

standing of the Bliss cards is easier and the range of listeners can grow.

The system was evaluated in a children's scienti c fair. The system was tested by more than 60 disabled and non-disabled children with ages raging from 6 to 14 years old. A short explanation on how it worked was given. They operated the application with surprisingly easiness and even if they had never seen Bliss symbols, they created statements with sense and reproduced them for their class mates. Children enjoyed interacting with the computer through the functionalities that *HeadDev* offered. Moreover, upper-body physical disabled children praised the opportunity of accessing a computer that it provided.

4. Conclusions and further work

In this paper a multimodal user interface using audio, vision and graphical user interfaces is presented. The combination of these techniques in the applications implemented allow non-verbal children to interact by means of head motion and face gestures with others, through a communication program that reproduces verbally the built statements.

HeadDev, has been tested by several disabled people (cerebral paralysis and physical disabilities) with encouraging results, but it can be improved with the recognition of more gestures (equivalent directly with BLISS symbols), sound (TTS and ARS) and adaptive learning capabilities for speci c disabilities.

BlissSpeaker is currently in an initial state, where the results are similar to the Bliss cards system. But possible enhancements are in process for enriching the communication process. The rst improvement is the extension of the symbols' data base, the possibility of saving combinations of symbols as phrases or the inclusion of new languages. As there is a verbal reproduction, emotions such as cheerfulness or anger could be inserted for contributing to a more natural interaction. Other important issue is the "inverse communication", that is, the development of a system that given a non-Bliss phrase, builds a corresponding bliss statement.

Although *HeadDev* and *BlissSpeaker* can work separately, they share in common the compromise of motivating the *e-Accessibility* for achieving goals in the *e-Inclusion*, and their combination can offer great bene ts for a sector of users.

5. Acknowledgements

This work has been subsidized by the national project TIN2004-07926 from the MCYT Spanish Government and TAGrv S.L., Fundación Vodafone and Fundación Integración Discapacidades Red. Javier Varona acknowledges the support of a Ramon y Cajal fellowship from the Spanish MEC.

References

[ACP98] ALBACETE P. L., CHANG S. K., POLESE G.: Iconic language design for people with signigcant speech and multiple impairments in assistive technology and arti cial intelligence: Applications in robotics, user interfaces, and natural language processing. *Lecture Notes in Computer Science 1458* (1998).

[Bli06] BLISSSYMBOLICS: Communication international. http://www.blissymbolics.org, 2006.

[eEu05] EEUROPE: Eu information society. http://europa.eu.int, 2005.

[KT05] KÖLSCH M., TURK M.: Perceptual interfaces. In *Emerging Topics in Computer Vision* (2005), Medioni G., Kang S., (Eds.), Prentice Hall.

[MYVP06a] MANRESA-YEE C., VARONA J., PERALES F. J.: Face-based perceptual interface for computer-human interaction. In *WSCG'2006 Short Papers Proceedings* (2006), Jorge J., Skala V., (Eds.), pp. 93–100.

[MYVP06b] MANRESA-YEE C., VARONA J., PERALES F. J.: Towards hands-free interfaces based on real-time robust facial gesture recognition. *Lecture Notes in Computer Science 4069* (2006), 504–513.

[RR03] RODRÍGUEZ A., RODRÍGUEZ C.: *Tecnologías de la Información y Comunicaciones y Discapacidad. Propuestas de Futuro*. Fundación Vodafone, 2003.

[War85] WARRICK A.: *Los símbolos BLISS en preescolar*. MEC-Fundesco, 1985.