GESTURE-BASED INTERFACES FOR INTEROB: INTERACTING WITH INFORMATION AND ROBOTICS SYSTEMS

Radu-Daniel Vatavu, Stefan-Gheorghe Pentiuc

University Stefan cel Mare of Suceava
str.Universitatii nr.13, RO-720229 Suceava
vatavu@eed.usv.ro, pentiuc@eed.usv.ro

Abstract

We discuss in this paper several implementations of computer vision applications that were developed in the last two years in our laboratory and for which gesture-based interactions were introduced. The aim is to provide enhanced human-computer interfaces for several commonly encountered application scenarios: manipulating virtual objects and working inside virtual environments, playing computer games and interacting with robotic systems. We particularly focused on table-based systems which allow natural and intuitive interactions as they transform into comfortable and familiar interfaces.

Introduction

The INTEROB project (Interacting by Gestures with Information and Robotics Systems) was launched having in mind the idea of providing natural, comfortable and intuitive interfaces for a variety of application scenarios by the sole use of human gestures. It is a known fact that today's applications and systems require interaction techniques that will match their advanced requirements, tasks and workloads and which are not always easy to perform using standard input devices. Our project focused on several common application scenarios such as: virtual environments which require proper pointing, selection and manipulating techniques that do not importunate, add extra cognitive load or distract users from the actual task to accomplish [1, 2]; computer games that benefit from advanced interaction techniques beyond standard controllers [3, 4, 5]; controlling robots and working in collaboration with robotic systems.

The use of human gestures as natural interfaces is a very popular and legitimate research area which received much attention since the very first time Bolt demonstrated how to "put-that-there" by interacting with virtual objects on a large screen using solely voice and gesture commands [6]. When it comes to gestures, there is the immediate net advantage of familiarity and naturalness over standard interfaces due to the fact that the experiences people encounter are similar to what they are used with when carrying out their everyday work [7, 8]. Human sensing technologies and acquisition devices have developed in the recent years to complex levels of accuracy and sensitivity which allows their reuse and implementation in order to enrich specific interaction tasks in various environments (home, office, playgrounds) and for various user categories. Gestures may be acquired using sensors and gloves, IR devices, Nintendo Wiimote controllers, stylus, mouse or using video cameras and employing computer vision algorithms. The use of computer vision assures the naturalness of the interaction as the users are not required any longer to wear additional equipments or devices. The drawbacks of the video technology relate to the amount of processing required for analyzing video streams (especially when multiple video cameras are involved), dependency of the working scenario parameters such as lighting, amount of noise and extra motion, user skin color, etc. In order to get around this disadvantages several methods are commonly employed: control the scenario and the environment where the actual processing takes place (for example by restricting the interaction area to a pre-defined region on a table surface, using blank or homogeneous-colored walls as foreground, assuming that the user's hands are the only objects that generate motion in the scene, built skin color models from training pixels, hues and saturations, etc).

The paper discusses several implementation notes from applications that were developed during the INTEROB project with a focus on a specific implementation of a table-based interaction technique: hands and gestures are acquired above the surface of a common table that becomes "interactive".

Scenarios for capturing gestures

We used the same acquisition scenario for all our implementations and applications: gestures were captured above the surface of a common working table. Users sat in front of the table while a video camera permanently monitored the working area on its surface. Visual feedback of the user's actions was made available on the display monitor located at the opposite end of the interaction table. Various working scenarios that make use of the interactive table concept are illustrated in Figures 1 (single user interaction), 2 (single user working collaboratively with a robotic system) and 3 (multiple users that sit around the table in a CSCW scenario).

By allowing hands to rest on the surface of the table, we managed to reduce the fatigue factor that may intervene for longer working intervals. This also brings in another advantage: our scenario is similar to the ones that users are already accustomed with such as when working with objects on a desk or typing at a keyboard while watching the monitor screen. The video camera monitors the desk as well as the users' hands and the various gestures they may execute. Various objects may be placed, removed or translated on the surface of the

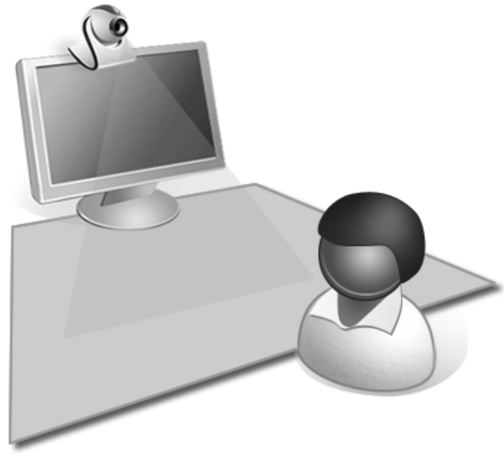

Fig.1. Single-user interaction scenario with information systems: the user sits comfortably in front of the working table facing the monitor screen while a video camera monitors a pre-defined interaction area on the surface of the table. Gestures are performed using one or two hands on the table surface

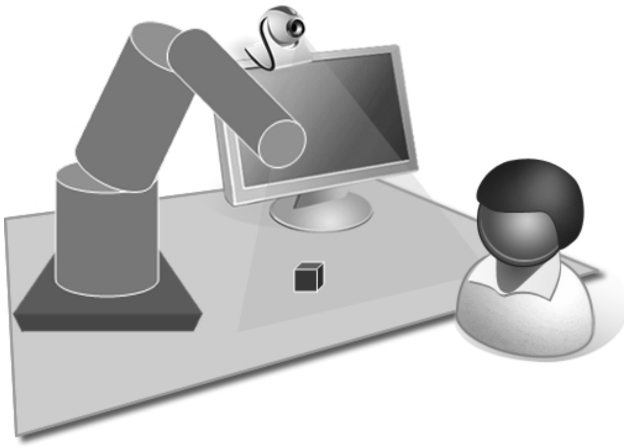

Fig.2. Single-user interaction scenario with robotics systems: the user works collaboratively with the robot by sharing the working region: the table

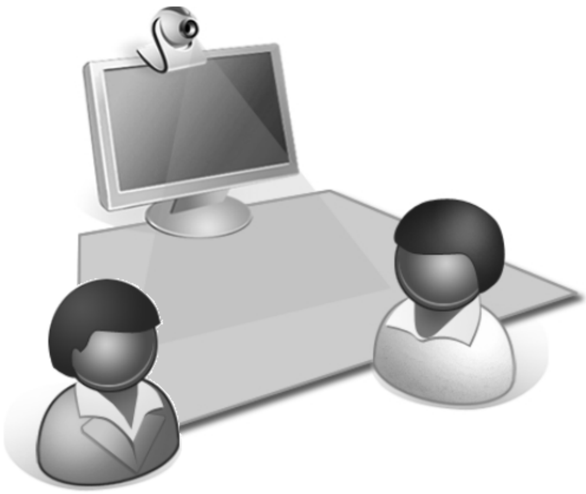

Fig.3. Multiple-users interaction scenario with information systems: users sit around the table sharing its surface in a CSCW scenario

desk with a direct correspondence within the application and with a permanently visual feedback on the monitor screen. Hands are detected above the desk by making use of a skin filter algorithm [9] and ensuring that a certain amount of contrast exists between the hands skin and the table color. Detected blobs may be further filtered by geometric constraints such as minimum and maximum width, height, area or aspect ratio [10] in order for only the hands objects to be identified in the end. Many techniques are available for posture and gesture recognition [8, 11] hence simple postures such as point (forefinger pointed), grab or pinch (thumb touching the forefinger) or hand open or closed may be easily recognized. Neural networks may be used in order to discriminate between previously learned postures [12]. Figure 4 illustrates an example of such an application that makes use of simple gestures for reconstructing a puzzle image.

Fig.4. Snapshots of a puzzle game. Top: constructing the puzzle pieces by dividing the image into a 4 x 6 matrix structure. Bottom:users interact with the pieces by translating and rotating them in order to re-construct the initial image

Conclusions

We presented in this paper several results which were obtained during the INTEROB project (Interacting by Gestures with Information and Robotics Systems). The project investigated the use of human gestures for interfacing information and robotics systems.

Our conclusion is that gestures captured in vision-based scenarios were able to enhance standard interfaces when designed carefully with respect to the environment working parameters. We presented some of our interaction techniques to several people and we received a very positive feedback in what concerned their reaction. During the "Open Doors" action organized at the Faculty of Electrical Engineering and Computer Science (Suceava, Romania) during 31 March – 4 April 2008, approximately 100 pupils from the terminal grade together with accompanying teachers visited our laboratory and we introduced them to a promotional video of our game prototypes. Their reaction was one of pleasant surprise due to the novelty of the interaction we proposed and one of excitement - “where can I get this from?” one of them asked. They showed themselves interested in the technology which motivates us to further continue our work as, in the end, they represent a considerable part of the target of the computer games industry (ages between 16 and 18).

Acknowledgements

This work was supported from the national funding grant “Research of Excellence”, Ref. No. 131/2006, acronym INTEROB.

References:

[1] Radu-Daniel Vatavu, Ştefan-Gheorghe Pentiuc, Christophe Chaillou, On Natural Gestures for Interacting with Virtual Environments, Advances in Electrical and Computer Engineering, 24(5), 2005, pp. 72-29.

[2] Radu-Daniel Vatavu, Ştefan-Gheorghe Pentiuc, Interacting with Gestures: An Intelligent Virtual Environment, 1st International Conference on Virtual Learning, ICVL, 2006, pp. 293-299.

[3] Microsoft Corp. XBox, www.xbox.com

[4] Sony Playstation, www.playstation.com.

[5] Nintendo. www.nintendo.com.

[6] Richard A. Bolt, Put-that-there: Voice and gesture at the graphics interface, ACM SIGGRAPH, 1980, ACM Press, pp. 262-270.

[7] Axel Mulder, Hand Gestures for HCI: Research on human movement behavior reviewed in the context of hand centered input, Technical Report TR 96-1, Simon Fraser University, 1996.

[8] R. Watson, A survey of gesture recognition techniques, Technical Report TCD-CS-93-11, Trinity College Dublin, 1993.

[9] Elli Angelopoulou, Understanding the color of human skin, Proc. SPIE Conf. on Human Vision and Electronic Imaging VI, 2001, pp. 243-251.

[10] William K. Pratt, Digital Image Processing: PIKS Scientific Inside, 4th Ed., Wiley-Interscience, 2007.

[11] J. LaViola, A survey of hand posture and gesture recognition techniques and technology, Technical Report CS-99-11, Department of Computer Science, Brown University, Providence RI, 1999.

[12] Radu-Daniel Vatavu, Ştefan-Gheorghe Pentiuc, Christophe Chaillou, Laurent Grisoni, Samuel Degrande, Visual Recognition of Hand Postures for Interacting with Virtual Environments, Advances in Electrical and Computer Engineering, Volume 6 (13), Number 2(26), 2006, University „Stefan cel Mare” of Suceava, ISSN 1582-7445, pp. 55-58 (republication in special issue from Development and Application Systems 2006)