Applications for people with special needs using Kinect

Bruno Almeida (a21160603@alunos.isec.pt)
Joel Jegundo (a21171043@alunos.isec.pt)
César Páris (cparis@isec.pt)
Nuno Martins (ncmartin@isec.pt)
Anabela Gomes (anabela@isec.pt)
Álvaro Santos (ans@isec.pt)

Departamento de Engenharia Informática e de Sistemas
Instituto Superior de Engenharia de Coimbra - Instituto Politécnico de Coimbra

Abstract

This article describes a scientific project related to the Human-Computer Interaction (HCI), which was developed with the purpose of helping people with special needs, specifically with cerebral palsy.
The key points of this project are how a person's body movements and his/her voice are detected by a device (in this case the Kinect sensor) and how this is translated into actions on the computer. All the applications should facilitate the access to learning and entertainment.
This project is divided into three areas:

- Help in education
- Accessibility
- Training and entertainment

To develop this project we rely on the help of APCC (a Portuguese acronym for Coimbra Cerebral Palsy Association).

1 Introduction

This paper describes a research work related with the creation of new human-machine interfaces related with accessibility concepts. The work was done using motion capture techniques based on the use of vision sensors and infrared sensors, both available on one system, called Kinect.
Kinect is a mixture of the Greek "kinesis" (that means moving or move) to "connect". The name defines the sensor extremely well because the user is connected to a platform which enables him to go through the actions of his/her movements and gestures [1]. Kinect was developed by Rare (company that is part of Microsoft) and PrimeSense (which developed a system that can detect 3D object movements) [1]. In developing the Kinect, Microsoft demystified the paradigm of controllers to develop a new system of natural user interface that enables advanced human control, voice control, motion recognition among other things [2]. The Kinect uses multiple resources (image, sound, tilt motor, infrared, and depth) with high precision synchronization (in real time) in a single device.
Kinect has a parallel processing algorithm (programmed into the chip SoC – this designation is related to all elements of a computer that could be put in a integrated circuit), used for obtaining the depth map from the structured light received. [3] The images are aligned pixel by pixel so as to obtain more precise information of the sensors.
These innovative features provide plenty of opportunities for interaction between services, applications, and users.
Through the camera, Kinect identifies the user by creating a three-dimensional image, to accurately identify the movements of the body.
The depth sensor consists of an infrared laser projector combined with a monochrome CMOS sensor that captures 3D video in any ambient lighting conditions. The range of the depth sensor is adjustable through software, Kinect is able to automatically calibrate the sensor based on the user and accommodating the presence of obstacles (furniture ...).
The software allows Kinect to recognize gestures, faces and voices. According to the manufacturers it is able to store information of up to six people simultaneously [4]. All this is a highly innovative combination of sensors and software that transform the user's body into control applications.
The paper is organized as follows: The next section explains each of the applications developed. The third section describes the actual development and the options taken. In the fourth section, some conclusions are drawn.

2 Developed applications

We developed applications in three different approaches. One approach was to help in education specifically in the area of mathematics, another concerns the computers accessibility, and the third deals with training and entertainment.

2.1 Help in education

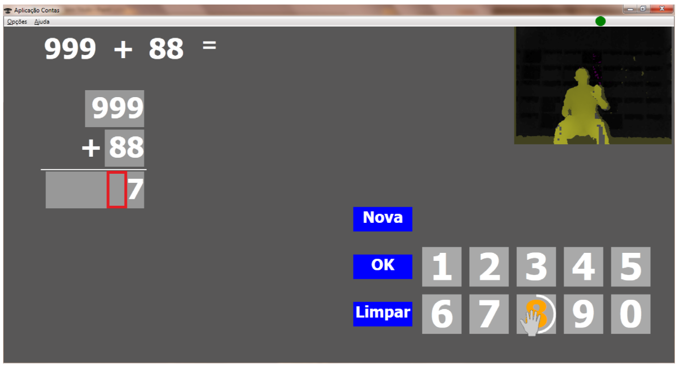

Figure 1: Application "Vamos fazer Contas" (Let's do Math)

This application aims to facilitate the learning process of basic math operations thinking mainly in children with cerebral palsy having difficulty in mobility.

We decided to widen the application interaction possibilities to its full extent, so that anyone could use it regardless of their restrictions. The user could control the application with his/her left or right hand, left or right foot or his/her head not limiting the application use. We also have an option to use both hands simultaneously. The size of the buttons and click time is configurable to adjust the application to each specific user. It also has an automatic mode that allows the result of each operation to be automatically verified at every insertion of a new digit in the result. If the result is correct a new math operation will be presented. Operations (addition, subtraction, multiplication and division) can also be configured by a user or a tutor through the specific menu. As can be seen in Figure 1 the operations are performed as in elementary school, in other words, the result is inserted from right to left.
The application has an initial calibration to verify the user interaction capabilities. If there are difficulties it is possible to change some configurations like the range of motion or the size of the buttons. In order to familiarize the user with the application, and since it is an application focused mostly on children, the calibration is done through the use of appealing images, such as animals. This is done by touching animal's images that reproduce the sound of the specific animal.
On top of the window there's a circle which can be red, indicating that no user is being detected, or green when the user is being detected. This will help to verify that the reason why the application is not responding to gestures is the fact that no skeletal is being detected.
In order to facilitate the interaction of a second person (tutor) a set of keyboard shortcuts were created to interact with the application as well as interaction via computer mouse.
The application allows the Kinect sensor to be restarted, if for some reason it is needed, without having the need to restart the whole application.

2.2 Accessibility

Computers used by a person with reduced mobility are often a real problem, so we decided to develop an application so that the user can control the mouse cursor through gestures and using speech recognition.
The user has three options to perform a click:

- Only voice - one word for the click and another for double-click;

- A gesture and voice - a gesture and through a specific word the user can select whether to make the normal click or double-click;
- Two specific gestures - one to click and the other to double-click.

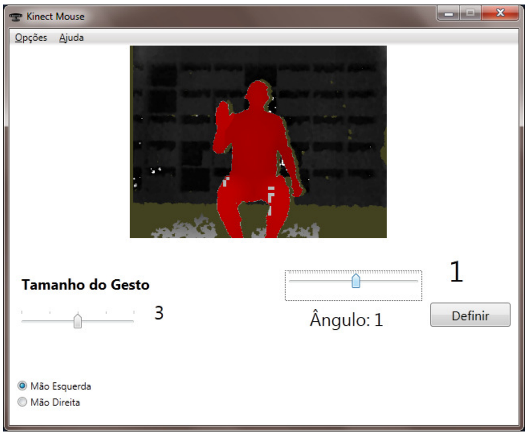

Figure 2: Kinect Mouse

In Figure 2 we can see a user controlling the application with the left hand, and we can also check some configurable options - range of motion (Tamanho do Gesto), tilt angle (Ângulo) of the Kinect and what hand (Mão Esquerda / Mão Direita) will control the cursor.

2.3 Training and entertainment

Interaction with the Kinect is not a very complicated task, but it can improve through some training, so the more we use it the better we can control it. For this reason we decided to create a game, which is a fun way of exercising with Kinect. Other games can be developed for exercising the body in a pleasurable way. So, while the children are working they are having fun.
We created a series of colored boxes and through gestures the user can throw a ball affected by the gravity to hit the boxes, in the shortest time possible, forcing the user to move (Figure 3).
The game camera can be controlled by one hand and the opposite hand is used to throw the ball through one of two distinct gestures: one consists in raising the hand above the shoulder and the other consists in stretching the hand forward. Throwing the ball can also be controlled by the computer mouse. There are three stages in the game each having three difficulty levels (easy, normal and hard). The increase of difficulty consists in an increase in the number of boxes as well as their weight.
This game has been developed in XNA because it is an existing framework oriented to games, including a physics engine making the games more realistic.

Figure 3: Training Game

3 Development and Usability

In this section we point out some of the problems inherent in applications that take advantage of Kinect capabilities and how we solve them.

3.1 How to use the application?

Most of the children that test our applications were in wheelchairs, making it difficult to detect the skeleton, creating an obvious need to find a solution to circumvent this problem.
In our implementations, users can choose if the Kinect needs to detect and treat twenty points of the skeleton (ten in the upper body and ten in the lower portion), or they can choose the seated mode which only detects and treats the ten points of the upper skeleton ignoring the ten points of the bottom.

3.2 Who's controlling the application?

We use a RGB camera in applications to identify if any skeleton is detected and, if there are different people, we use different colours for each skeleton detected. Only the first one to be detected will control the application, ignoring all the others. When the first person leaves the detection area, the second person caught becomes the first and thus controls the application.

3.3 Improving detection

We decided to create two modes of gesture tracking of the skeleton in the applications developed.
In the first mode we used a direct tracking in which the value of the position of the real world is placed in the position of the cursor. The position is recalculated according to the resolution of the window or application area where you want to use the Kinect. Despite being more fluid this mode is more susceptible to interference from sunlight.
In the other mode the last thirty second positions of the skeleton are saved, and after several tests we decided to treat only the last ten positions (corresponding to the last 1/3 second) and making a calculation of the weighted average to ascertain whether the movement is acceptable or if it is an incorrect reading. If it is acceptable the mean value is placed at the coordinates of the cursor.

3.4 Range of user interaction

- In the applications it is possible to change the size of the gesture. This affects the extent of the hand movement required to move the cursor from one side to the other of the window, scaling between the position of the real world and the window.
- The recommended distance to interact with the Kinect sensor is 1.2m (minimum) and 4m (maximum) [5], so we decided to check in real time the distance between the user and the Kinect sensor. To guaranty a more robust interaction, if the user is not between 1.2 and 2.5m the application warns the user if he/she is too close or too far away from the sensor.

4 Conclusion

This project has shown how it is possible to control any type of application using a set of gestures, a set of existing words in the Windows grammar or both. The same gestures can be used for different interactions, alternating the specific function by a word.

It was also shown that Kinect could be used to improve and facilitate the way people with limited mobility can interact with computer applications.

References

[1] http://www.xboxic.com/news/6071 accessed on 06/03/2012.

[2] http://www.microsoft.com/Presspass/press/2010/mar10/0331PrimeSensePR.mspx?rss_fdn=Press%20Releases accessed on 06/03/2012.

[3]http://social.technet.microsoft.com/wiki/contents/articles/6370.aspx accessed on 07/03/2012.

[4]http://seattletimes.com/html/technologybrierdudleysblog/2009296568_e3_new_info_on_microsofts_nata.html accessed on 07/03/2012.

[5]http://www.slideshare.net/marcosprod/apresentacao-kinect accessed on 08/07/2012