Proceedings of the World Congress on Electrical Engineering and Computer Systems and Science (EECSS 2015)
Barcelona, Spain, July 13 - 14, 2015
Paper No. 319

Gesture-Based Interaction Using Touchless Technology for Medical Application

Abdelkrim Belhaoua, Jean-Pierre Radouz, Florian Pereme
Altran Research Medica / ALTRAN EST
Parc d'Innovation, Boulevard Sébastien Brandt, Strasbourg, France
abdelkrim.belhaoua@altran.com; jeanpierre.radoux@altran.com;florian.pereme@altran.com;

Abstract- In this paper, we present a touchless human-machine interaction dedicated to a science museum for children, in the short-term, and surgeons in operating room in the long-term. In fact, surgeons need to review medical images during minimally invasive surgery without interact directly with the medical devices. Touchless technology is an interesting solution despite of the many challenges as low accuracy, bad lighting conditions, occlusion and real time computing. In order to performed such application, an educational game, so-called Anatomia, is developed using Kinect and interacting with a 3D model of a human body. Many tests and surveys are carried out from children, adults and surgeons to evaluate the touchless technology and improve its ergonomics.

Keywords: Touchless technology, Kinect, minimally invasive surgery, interaction.

1 Introduction

Minimally invasive surgery is becoming more and more common in hospitals. These procedures are performed through tiny incisions instead of one large opening. Because the incisions are small, patients tend to have quicker recovery times and less discomfort than with conventional surgery, all with the same benefits.

Despite its success, minimally invasive procedures remain a major challenge where it is impossible to palpate directly the organs which is performed through small incisions that limit free sight and possibility to palpate organs . In this context, our institution has the purpose of integrating innovative augmented reality technologies in operating room (OR) in order to improve minimally invasive surgery.

Using this technique, surgeons need to interact frequently with several medical imaging systems before and during surgeries in order to review medical images and records. In the other hand, it is strictly prohibited to the surgeon to interact directly with these medical devices because they are difficult to sterilize. Usually during a surgery, an medical assistant operates the mouse and keyboard for such interaction. To overcome this problem, sterility restrictions in surgical settings make touch-less interaction an interesting solution for surgeons to interact directly with their environment into the operating room (OR). One can distinguish two types of indirect interactions: voice recognition and gesture recognition. The voice based-approach is inadvisable in OR because one cannot distinguish between different people speaking in the same room. The gesture-based approach remains a "good" alternative despite of the many challenges, low accuracy, bad lighting conditions, occlusion and real time computing.

A literature survey shows that over the last years some efforts have been devoted to gesture-based approach. In [1], a study of people intuitively using gestures and speech to communicate with computers is presented. The analysis showed that people strongly prefer to use both gestures and speech for the graphics manipulation and that they intuitively use multiple hands and multiple fingers in all three dimensions. Recently, many researchers have investigated the gesture based approach [2, 3, 4, 5]. In [6], Kiliboz and al., propose a dynamic gesture recognition for human-computer interaction using finite-state machine (FSM)-based techniques. The experience shows an average of 73 % accuracy in real time for a vocabulary of eleven gestures from continuous hand motion.

In our institution Altran Research Medic@ (i.e., MEDical Image Computerized @nalysis), we are interesting to Kinect technology in order to validate its usefulness and relevance in the operating room during minimally invasive surgery. First, we have developed an application for children, so called Anatomia. It allows us to evaluate the robustness of this technology, because the children can afford to do different gestures without limits. In fact, many tests have been performed in different schools in order to study the complexity of gesture recognition. In this paper, we discuss these results.

2 Anatomia

Anatomia is an educational game for children (8-15 years old) using Kinect 2 and interacting with a 3D model of a human body, provided by IRCAD (Research Institute against Digestive Cancer). The basic idea is that the children tend to be more interested by medical filed. Can the children point out a heart on the 3D model of human body? This game bring them some entertaining and stimulating games to improve their medical knowledge. The quizzes include questions about the digestive system and the respiratory system (Figure 1). The selection of the organs is performed using the identification of the right hand-state (open/fristed). This application falls within the framework of a our research program on collaboration with IRCAD (Research Institute against cancers of the digestive system) and the LeVaisseau Museum of Strasbourg where the application will be installed on September 2015.

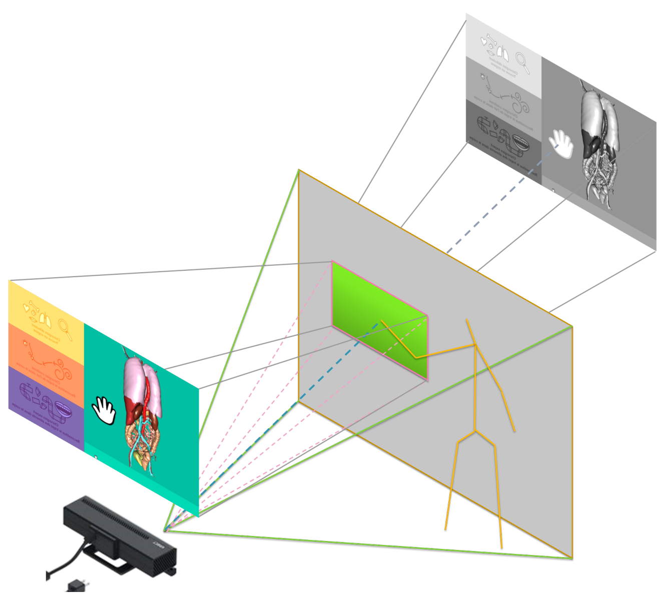


Fig. 1: Anatomia : Human interaction with 3D Model using Kinect

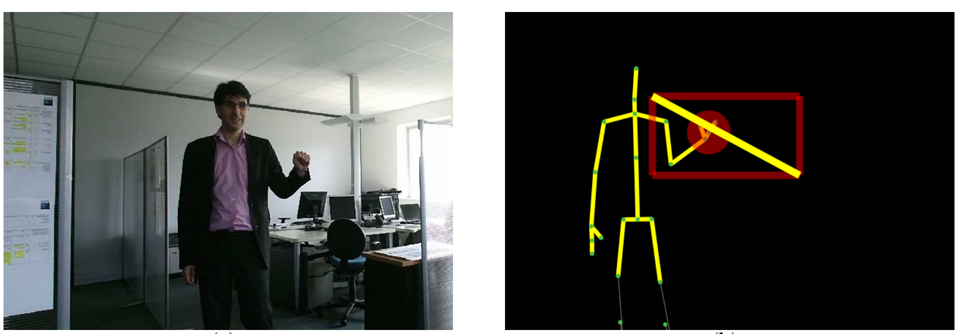
2.1 Kinect V2.0 for Windows

The Kinect sensor by Microsoft Corp. was introduced to the market in November 2010 as an input device for the Xbox 360 gaming console and was a very successful product. The Computer Vision society quickly discovered that the depth sensing technology in the Kinect could be used for other purposes than gaming and at a much lower cost than traditional 3D-cameras (such as time-of-flight based cameras). Kinect offers the potential to transform how people interact with computers and Windows embedded devices in multiple industries : Education, Healthcare, Transportation, Game. The next version of Kinect V2 for Windows hardware has been launched in the summer of 2014 with new and improved features, including increased depth-sensing capabilities, 1080p video, improved skeletal tracking and enhanced infrared technology (Figure 2).

Fig. 2: Kinect 2.0 for Windows 8.1: Motion sensing input device

Kinect device has a color camera, an infrared emitter, and a microphone array consisting of four microphones. Colour camera can save 30 RGB image frames with 1920 $\times$1080 resolution per second. The depth camera can save 30 image frames with 512 $\times$ 424 resolutions per second. Because data gathered by this sensor is used for depth-sensing it is also called as depth sensor. Angle of the view of the depth sensor is limited to 70 degrees horizontally and 60 degrees vertically. The range of the depth the sensor can measure is from 0.5 to 4.5 meters in the default mode.

In the scope of this project, skeleton tracking features of Kinect for Windows SDK is used. The SDK can process the raw data coming from the depth sensor in real time and track skeleton of human beings. It can track whole skeleton of two people at most within its view while position of six people can be tracked at the same time.



(a) (b)

Fig. 3: BodyBasics example using Kinect 2.0

Human skeleton can be divided into two parts as upper body and lower body. Kinect for Windows can track twenty five joints of human body, half of them belonging to the upper body while the other half belonging to the lower body (Figure 3).

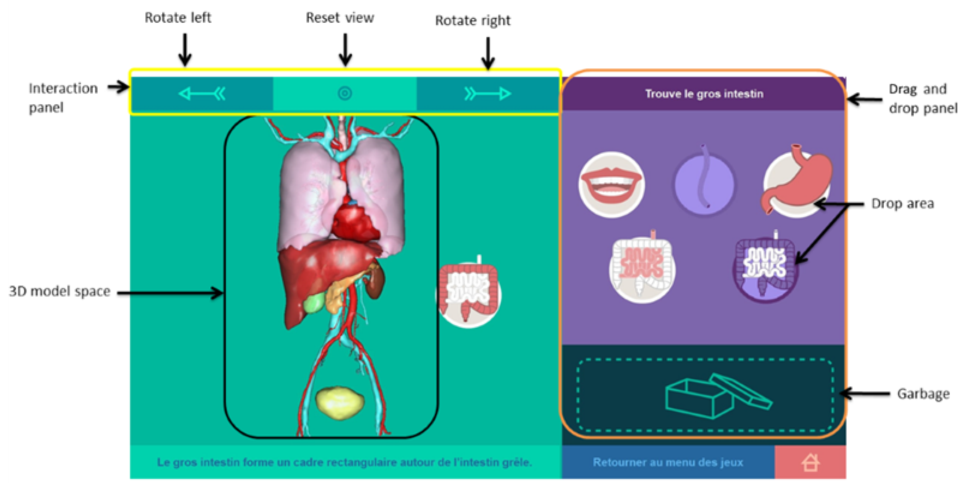


Fig. 4: Graphical User Interface for Anatomia application

2.2 Graphical user interface

The project is developed under the Framework FW4SPL ([7]) which is a component-oriented architecture with the notion of role-based programming. FW4SPL, developed by IRCAD, consists of a set of cross-platform C++ libraries and focuses on the problem of medical images processing and visualization. The graphical user interface (GUI) is managed with Qt 5.4.0 (Figure 4).

2.3 Software architecture

The Anatomia application is based on several modules, called bundles in the FW4SPL library [8]. Each of these modules operate independently from each other and communicate via messages (Figure 5).

- Kinect Bundle: It captures all the movements of potential users and forward the relevant information (virtual positions) to other bundles.
- AltranVisuVTKAdaptor: This Bundle is the link between the movements of the user (opened hand, closed hand, grab, drop, roll over) and interactions with the patient's 3D model used in the application. It also communicates to other bundles the selected organ by the user.
- AltranData: This is a bundle including multi-variables needed to run the application.
- UiVaisseau: It includes all of the GUI of Anatomia and contains the widgets (components).
- Managers: They deal with the information transfer between different Bundles. They are used to organize data flow and manage the specific messages of the application.

3 Result and discussion

The skeletal tracking system provides us joint positions of tracked persons' skeletons. These joint positions are the data consumed as position and pose. However, these positions data returned by tracking system has some noise due to many parameters as : room lighting; a person's body size; the person's distance from the sensor array; the person's pose; location of the sensor array; quantification noise; and rounding effects introduced by computations. Note that the joint positions are accurate (i.e., the joint position data is close

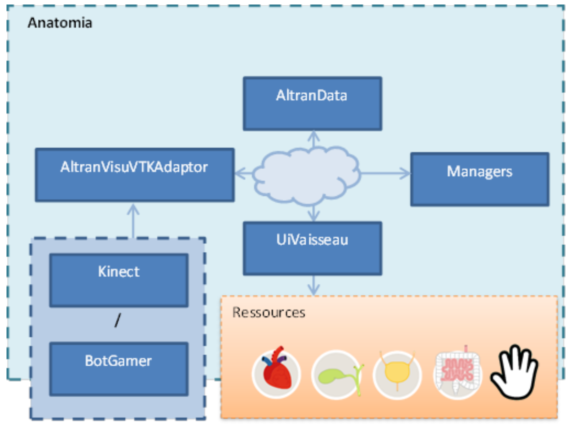
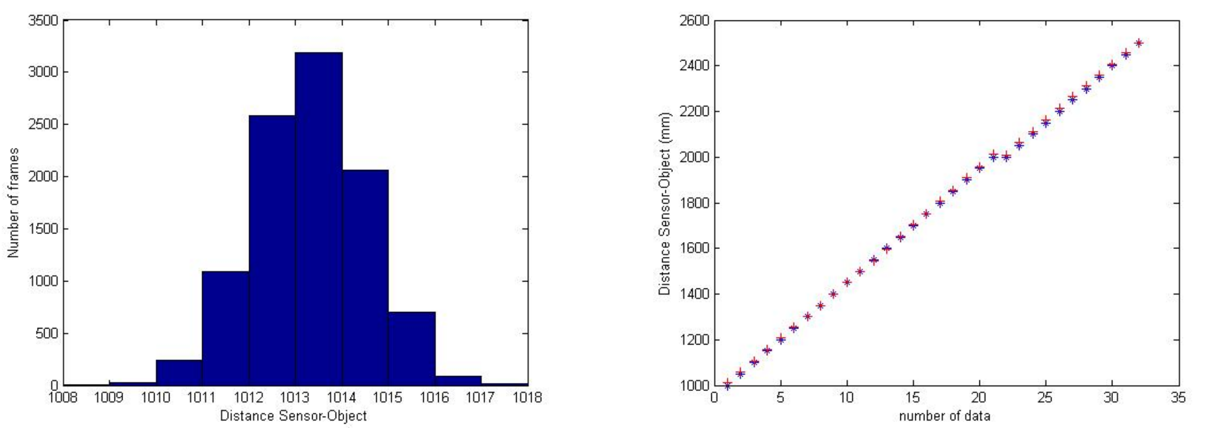


Fig. 5: Software architecture

to the positions in the real world) but not necessarily perfectly precise (i.e., the joint positions are scattered around the correct positions in each frame). Figure 6 illustrates some experiments to evaluate both accuracy and precision of the Kinect. In our application, only the right hand is tracked. This is a constraint imposed



(a) The occurrence of the distance sensor-object measured from 10000 frames

(b) The gap between the real distance (in blue) and the measured distance (in red) from Kinect

Fig. 6: Accuracy and precision of the Kinect

by the LeVaisseau Museum in order to simply the gesture interaction to the children. The principal problem encountered was the identification of the hand-state (open-fisted), that allows to select, drag and drop an organ. In fact, some children have frequently bracelets, watches and jewels in their handshake, and wear long-sleeved sweater falling on their hands. Furthermore, the hands of children are much smaller than the hands of adults. To overcome these problems, firstly a smoothing filter based on the Holt double exponential is used. In fact, Holt extended simple exponential smoothing to allow forecasting of data with a trend with less latency.This method can be adjusted via five parameters: smoothing, correction, prediction, jitter Radius, max deviation radius. All these parameters are studied but no method is developed yet to choose the parameters' value automatically.

Secondly and in order to maximize the adaptability of our application to the largest possible number of users according to their morphology (Figure 7), the detection zone is calculated dynamically based on a fluctuating measure depending on the individual (i.e., the distance between her neck and pelvis) . Many tests are carried out with children, adults and fellows (surgeons) to improve the ergonomics and validate the application. All experiments demonstrates that using one hand seems to be very promising as this would allow the user to interact without positional restrictions without too tiring.

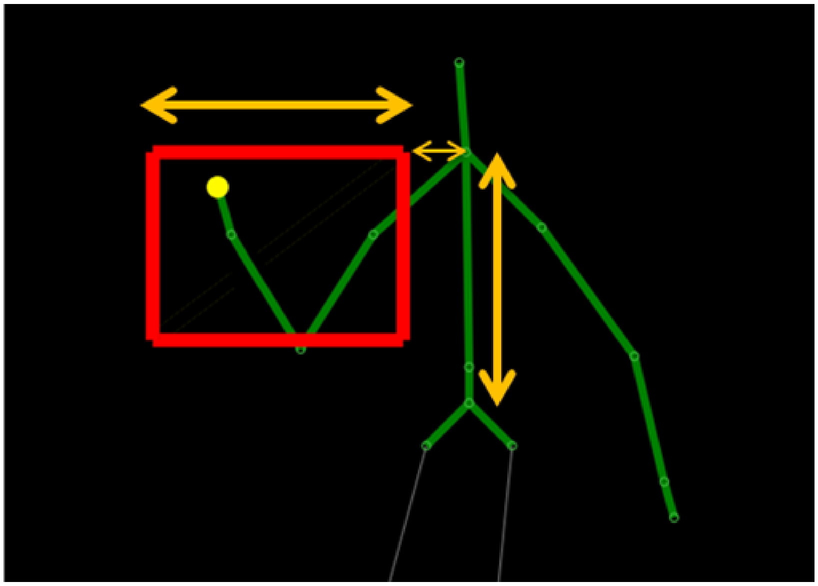


Fig. 7: The detection zone is computed dynamically based on the distance between her neck and pelvis

4 Conclusion

This paper presented touchless interaction human machine using Kinect, dedicated to operating room in the long term. Using Kinect sensor to capture gestures yields a system which requires more useful and simpler environmental setup which makes our application advantageous in terms of system setup and mobility. The development of an application for children allows us to first evaluate the robustness of this device. However, some drawbacks should be corrected in the future. The first of them is the fact that the system should be installed in non-reflective room and no mirror. The other drawback is the fact that the Kinect consumes a lot of resources of CPU and the influence of human tremor as potential source of jitter. Future works are devoted to these points.

Acknowledgments

The authors thank the developers : Jeremy Bernard, Bastien Scheurer, Yoann Chaumy, Lucas Saurel and Daniel Peyrebonne for their contributions during the project, without their motivation and their involvement, this research would not have been possible. We also thank Pr. Luc Solar (IRCAD) for constructing the human model and Johan Moreau (IRCAD) for his technical assistance and kindness. Thanks also goes out to Christel Le Delliou and Bruno Baechler (LeVaisseau Museum of Strasbourg) for pedagogical assistance.

References

Hauptmann, A. G. (1989). Speech and gestures for graphic image manipulation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 241–245.

Sinthanayothin, C., Wongwaen, N., Bholsithi, W. (2012). Skeleton tracking using kinect sensor & displaying in 3d virtual scene. *International Journal of Advancements in Computing Technology*, *4*(11), 213–223.

Jacob, M. G., Wachs, J. P. (2013). Context-based hand gesture recognition for the operating room. *Pattern Recognition Letters*.

Wena, R., Tayb, W.-L., Nguyena, B. P., Chng, C.-B., Chui, C.-K. (2014). Hand gesture guided robot-assisted surgery basedon a direct augmented reality interface. *Computer methods and programs in biomedecine*.

Krishnan, R., Sarkar, S. (2015). Conditional distance based matching for one-shot gesture recognition. *Journal of Pattern Recognition*, *48*, 12981310.

Kiliboz, N. C., Gudukbay, U. (2015). A hand gesture recognition technique for human-computer interaction. *Journal of Visual Communication and Image Representation*.

IRCAD. (2011). http://code.google.com/p/fw4spl/.

Medica, A. R. (2013-2015). Anatomia: Intgration Kinect dans le Framework FW4SPL. Master's thesis, Altran Est.