Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 97 (2013) 723 – 729

Procedia
Social and Behavioral Sciences

The 9th International Conference on Cognitive Science

Interactive hand and arm gesture control for 2D medical image and 3D volumetric medical visualization

Yusman Azimi Yusoff[a,]*, Ahmad Hoirul Basori[a,b], Farhan Mohamed[a]

[a]IJN-UTM Cardiovascular Engineering Centre, VicubeLab, Faculty of Computing
Universiti Teknologi Malaysia, Skudai, Malaysia
[b]Interactive Media and Human Interface Lab,Department of Informatics, Faculty of Information Technology
Institut Teknologi Sepuluh Nopember Surabaya, Indonesia

Abstract

The study related to human machine interaction has been done intensively to improve the ways human communication with the machine. In this paper, we present the application that uses of hand and arm gesture control to interact with a 2D medical image 3D volumetric medical visualization. User is allowed to use hand and arm gestures to navigate and manipulate the visualization with the help of Microsoft Kinect for Xbox™ as the input device. We explained the gestures that can be used to control the visualization as well as the framework and the interaction flow of our application with Kinect sensor in order to perform these operations. User will be able to perform simple navigation and manipulation to the visualization by using this application.

Selection and/or peer-review under responsibility of the Universiti Malaysia Sarawak.

Keywords: Human machine interaction; Gesture; Control;

1. Introduction

The development of new technologies in the area of human-computer interaction are very encouraging .These technology will help in integrating technological solution in everyday live. Apart from that, new technology is also developed to facilitate people's job. Progress in human computer interaction field has also opened a wide opportunity for researchers and scholars to develop new technology. These phenomena have affected the field of human-computer interaction. Studies in this area have been increased with the mission to improve the way people interact with computers. Tools such as Microsoft Kinect, Wii Remote, PlayStation Move and PlayStation Eye have opened a new chapter in this field. Current technology trends nowadays are more towards interaction using gestures. Users might no longer need to use keyboard and mouse as input devices [1]. There are several ways users can interact with computer which is by using gestures of voice, touch and hand movement. Each gesture requires specific tools so that the computer can receive input from the user. The method used to interact also should be related with the task carried out by the user. Interaction using body movement is well suited for applications such as virtual reality application [2].

Research on human machine interaction is very important in improving the ways for human to interact with the machine. The process of learning new ways of interaction is one of the obstacles that faced by researchers since user

* Corresponding author. Tel.: +6016 2226982
E-mail address: yazimi2@gmail.com

Selection and/or peer-review under responsibility of the Universiti Malaysia Sarawak.
doi:10.1016/j.sbspro.2013.10.293

need to understand the process the technique which is time consuming. Human body gestures start to take part in the research after the idea of touch-less based interaction is introduced. Interaction using body gestures are much easier for the user to understand, making the learning process since it does not need an intensive training to learn the gestures [3]. This is because of gesture are chosen based on the significant relationship with the operation that will done. For example, user can rotate the visualization in any direction by pointing and dragging. It signifies that user can grab and turn the object in real world. Each gesture that used to navigate the visualization is studied to make sure the gesture signifies the user can understand the method is proposed.

2. Related Work

According to [4] human-computer interaction is fundamental to the discipline of information technology. Any computing system is said to be interactive when it involves one or more interface that allows users to give commands and get results. Development using the graphical user interface, which gives the users from varying levels of knowledge about computers, allows them to understand how to use the application.

For system developers who are experts in the field of human-computer interaction, Arrangements are necessary to take into account the needs of users in advance of the development process begins. This is important so that the information obtained can be presented to system engineers [5]. Latest achievements in the field of human-computer interaction have an impact on the development of natural interfaces that allow users to interact with computing devices in a more intuitive ways [6]. Development of natural interfaces is expected to facilitate the users to learn to use the interface as soon as possible.

The intention to develop natural interfaces has been around for tens of years. After World War II, academic groups have been established to improve the process of interaction between man and machine. After the formation of these groups, the field of human-computer interaction is growing. Many studies have been done to improve the way people interact with computers. At present, the concept of human-computer interaction is an important aspect in the process of system development.

Gesture control interfaces is an interface that receive and recognize the gestures made by the user. This kind of interface is hardly fits in everyday life. However, there are two main difficulties in the design of gesture control interfaces [7]. The first problem is about the temporal segmentation ambiguity. The definition for starting and ending points of continuous gestures is not ambiguous. The second problem is about the spatial-temporal variability. The understanding on gestures among individual are varies. Generally, there exist many-to-one mappings from concepts to gestures and vice versa, and hence gestures are ambiguous and incompletely specified [8].

This interface seems to be suitable for operating rooms. Unfortunately, there are no gesture control interfaces that match enough with the requirement of the operation room due to many reasons. This interfaces need to identify and trace the body position and movement. Devices that is used to obtain these information may equipped by the user such as magnetic field tracker, special glove or suite, or combination of these equipment. This equipment may have different outcome in term of accuracy, latency and user comfort.

A glove-based interface is presented that can be used for the medical image analysis [9]. They use a data glove and a tracker to track and recognize independent hand posture and trajectory and send the information to the host application. Improvement is made in [10] by replacing the tracking system with a Wii mote to allow exploration for medical data at a distance. The idea from these projects about the gesture control interface are very convincing but these solutions are unsuitable to be use in the medical field especially in the operating rooms where non-sterilizable devices are not permitted to avoid contamination of the patient, the OR and the surgeon. There are also vision-based interfaces have been proposed to be used in the operation room. A system named Gestix is presented in [11] which is a video-based hand gesture recognition system. It allow user to navigate and manipulate magnetic resonance images (MRI). The problem with the system is vision-based interfaces need to contend with other problems since the vision can be affected by room lighting, user movement and the background of the user.

There are many devices able to be used with the gesture-enabled application. Microsoft Kinect sensor is one of the robust sensor that can be used to accomplish different task since it ability to provide depth data and color image at the same time. Raheja in [12] use Kinect sensor to track the fingertip and center of palm by a series of process. The advantage of the depth image is that the depth information is useful in order to isolate the region of interest from other part of the image which allow faster detection process.

3. Hand and Arm Gesture Description

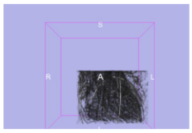
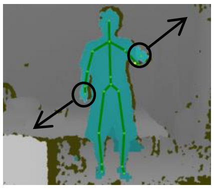
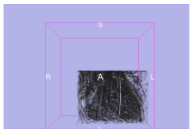
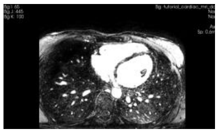
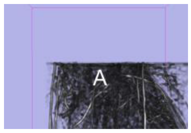
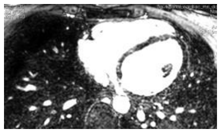
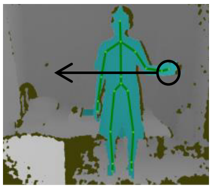
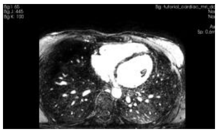
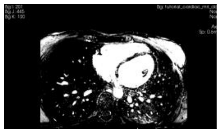
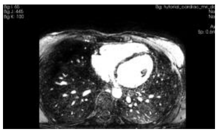
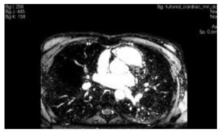
Name	Gesture/hand and arm movement	Effect on the medical image or Visualized model	
		Before	After
Rotation	1 hand point and rotate 		
Zoom	2 hand move in/out 	 	 
Windowing	1 hand swipe back and forth horizontally 		
Change Slice	1 hand swipe up and down vertically 		

Fig. 1. The effect of gesture to the medical image/visualization

Fig. 1 show the available gestures in the application used to control and manipulate the medical image and visualization model. Rotation can be done using 1 hand at the front, pointing at the start position and drag to rotate it based on the needs to rotate the visualize model. We can see that action signifies holding and turning the object. The model starts rotating when the user’s hand is inside the gesture recognition area. The movement of the arm is compute and the direction of rotation is sent to the mouse as input for rotation process.

Zoom are divided in two types that is zoom in and out. Both gesture are quite the same but different in direction. Zoom out require user to show both hand at a distance and move them closer. Zoom in require user to show both hand at a close distance and move them further from the other. The distance of the hand before and after the movement are used to simulate the input of the mouse scroll. Mouse scroll can be used to adjust the zooming capability of medical visualization. These gesture is signifies that user want to enlarge of shrink an object.

Windowing is the process of selecting some segment of the total pixel value range (the wide dynamic range of the receptors) and then displaying the pixel values within that segment over the full brightness (shades of gray) range from white to black. Windowing process can be done moving one hand back and forth horizontally to set the window size based on the user's need. The gesture is much like controlling slider of a window size.

Medical image contain many slices that can be process and visualize in 3D volumetric model. The medical visualization application allows users to change the image slice so that they can study the image from different position. Changing image slice can be done by swiping one hand in up or down direction to change the image slice. This gesture is only valid to be used on 2D medical image.

Table 1. List of full gesture

Name	Gesture/hand and arm movement
Rotation	1 hand point and rotate
Zoom	2 hand move in/out
Windowing	1 hand swipe back and forth horizontally
Change Slice	1 hand swipe up and down vertically
Close Application	Right hand wave
Reset view	Left hand wave
Change viewing type	1 hand push forward

Table 1 show the gesture that provided in the application that can be used in the application. There are also gestures that can be used to do operation that not related in navigate and manipulate the visualization. Each gesture has its own operation to avoid confusion.

4. System Overview

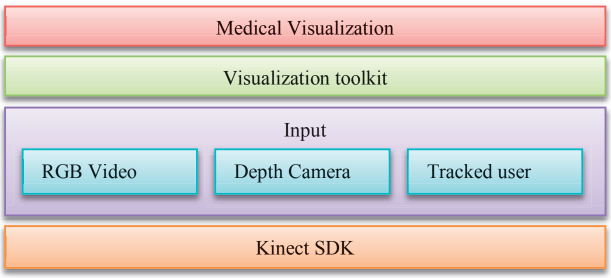


Fig. 2. System framework of the application

Fig 2 show the overview of the system framework with the essential component that support the system. This application is used with the 3D Slicer, an open source medical visualization software which able to load a 3D volumetric medical data. There reasons why 3D slicer is used to be used with the application. Medical area is well known for DICOM (Digital Imaging and Communications in Medicine) data format which is implemented in almost every radiology, cardiology imaging, and radiotherapy device (X-ray, CT, MRI, and ultrasound). 3D Slicer is able to load this data format and visualize it into 3D model. The result of visualized model is depending on the user interest area.

Kinect SDK are the foundation for the application since it provide the information that is required to identify the users hand and arm gesture. This SDK can provide image and users skeleton information obtained from depth camera and RGB camera. There are also microphone array that can be used to acquire audio data. The information that is required in the application is the user skeleton data because it contains the movement information that needs to be processed to identify and recognize the gestures.

Our application uses the user's skeleton data that is tracked by Kinect and monitors the hand and arm movement. The movement is compared with a set of movement criteria to identify which gestures are done by the user. Each movement has its own criteria that make its unique from other gestures. Some gestures require to use both hands and other only require one hand for the gesture to be recognize.

Every gesture has different ways of sending input to the software. Some gesture simulates keyboard input and sends to the software while others need to simulate the mouse movement and click event for the software. Each recognized gesture done by user will simulate the input and these inputs are sent to 3D Slicer depending on the operation assign for the particular gesture.

4.1 Interaction Flow

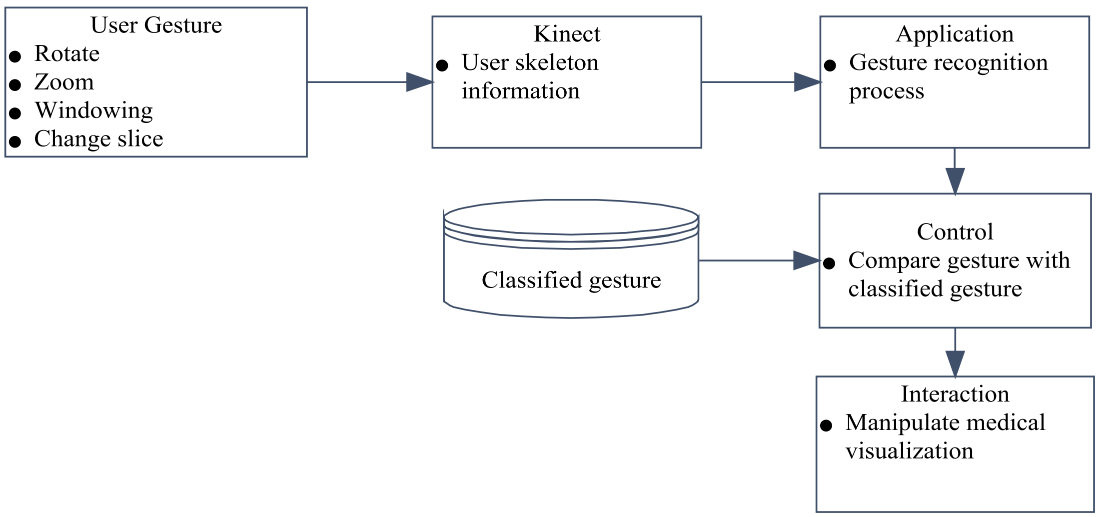


Fig. 3. Flow of interaction in the system

At first, user needs to setup the Kinect sensor and the medical visualization application so that the visualized model is ready to be used with the application. The interaction flow start with hand and arm gestures from user. Kinect will track the user movement and send the information to the application for the gesture recognition process. Application's output is the input simulated that is used to control the medical visualization.

4.2 Implementation

The essential hardware required for this application to run is the Kinect itself as the camera used to track the user. There are many camera that equipped with depth sensor by Kinect is among the best in term of price and performance. The availability of Kinect also is the one of the factor because it is easily to find all over the world.

Table 2. Hardware Specification during application development

Hardware	Specification
Processor	Intel Core i7 920
RAM	6 GB
Hard disk	640 GB
Graphic Card	NVidia GTX 670 2GB

The setup also requires a computer with decent specification. Table 2 shows the computer specification used during the development process. A computer with a standard graphic card is suitable enough to run the application since there are graphical operation need to be executed while user using the application. The amount of RAM also need to enough because Kinect has minimum requirement of 2 GB of RAM. Running a heavy application also require more RAM so that the computer can manage to run the application. Using projector as the display media is the best setting but LCD screen also can do the job well.

The operating system that we used to develop the application is Windows 7 64-bits. We used Microsoft Visual Express C# as the IDE since we develop the application in C# language. Kinect SDK is used to allow us to obtain the data from Kinect sensor. Kinect SDK also comes with Kinect driver which needed when connecting Kinect to the computer. The SDK version that we used during development is version 1.6 which is the latest version on that time. We belief that newer SDK version also allow the application to works.

Only one user is allowed to use the application at a time because the application is set to track and receive one input from a tracked user. User must stay at the predefine range based on the Kinect point of view so that the skeleton structure can be trace completely.

5. Conclusion

In this work, the user is allowed to interact with the medical image and 3D visualized model using a touch-less based interaction technique. This method is done by recognizing the gestures shown by the user. Gestures recognition process become much easier by using inexpensive sensor. Application is developed to recognize the gestures and simulate inputs corresponding to gestures. Each gesture has its own operation on the medical image or 3D volumetric visualization.

Acknowledgments

This research is supported by the Ministry of Science and Technology (MOSTI) and collaboration with Research Management Center (RMC), Universiti Teknologi Malaysia (UTM). This paper is financial supported by E-Science Grant Vot. No.: Q.J130000.2428.00G68.

Reference

[1] Lindley S., Couteur J. and Berthouze N. B. (2008). Stirring up Experience through Movement in Game Play: Effects on Engagement and Social Behaviour. Proceedings of ACM CHI 2008 Conference on Human Factors in Computing Systems. April 5–10. Florence, Italy. ACM: 511- 514.

[2] Nadia Berthouze(2010). Affective posture and body movement as modality for human-computer interaction. Perada Magazine, 10 September, p.1.

[3] K. Montgomery, M. Stephanides, S. Schendel, and M. Ross. User interface paradigms for patient-specific surgical planning: lessons learned over a decade of research. *Computerized Medical Imaging and Graphics*, 29(5):203–222, 2005.

[4] Lunt, B., Ekstrom, J., Reichgelt, H., *et al.* (2008), *Communications of the ACM*, Vol 53, Iss. 12, page 133.

[5] Maíra Greco de Paula, Simone Diniz Junqueira Barbosa and Carlos José P. de Lucena (2005). Conveying Human-Computer Interaction Concerns to Software Engineers Through an Interaction Model. *Latin American Conference on Human-Computer Interaction 05.* October 23-26. Cuernavaca, México: ACM 109 – 119.

[6] Norman Villaroman, Dale Rowe, and Bret Swan (2011). Teaching natural user interaction using OpenNI and the Microsoft Kinect sensor. *Proceedings of the 2011 conference on Information technology education* (SIGITE '11). October 20–22. New York, USA. ACM: 227-232.

[7] C. Shan. Gesture control for consumer electronics. In L. Shao et al., editors, *Multimedia Interaction and Intelligent User Interfaces*, Advances in Pattern Recognition, pages 107–128. Springer London, 2010.

[8] S. Mitra and T. Acharya. Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37(3):311–324, 2007.

[9] B. S. Tani, R. S. Maia, and A. v. Wangenheim. A gesture interface for radiological workstations. In *IEEE CBMS '07*, pages 27–32, Washington, DC, USA, 2007. IEEE Computer Society. 15

[10] L. Gallo. A Glove-Based Interface for 3D Medical Image Visualization. In R. J. Howlett et al., editors, *Intelligent Interactive Multimedia Systems and Services*, pages 221–230, Berlin Heidelberg, 2010. Springer-Verlag 2

[11] J. P. Wachs, H. I. Stern, Y. Edan, M. Gillam, J. Handler, C. Feied, and M. Smith. A gesture-based tool for sterile browsing of radiology images. *Journal of the American Medical Informatics Association*, 15(3):321–323, 2008. 16

[12] J.L. Raheja, A. Chaudhary, K. Singal. Tracking of Fingertips and Centre of Palm using KINECT. In *proceedings of the 3rd IEEE International Conference on Computational Intelligence, Modelling and Simulation*, Malaysia, 20-22 Sep, 2011, pp. 248-252.