View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Institute of Advanced Engineering and Science

International Journal of Electrical and Computer Engineering (IJECE)
Vol. 7, No. 3, June 2017, pp. 1308~1315
ISSN: 2088-8708, DOI: 10.11591/ijece.v7i3.pp1308-1315

Real Time Interactive Presentation Apparatus based on Depth Image Recognition

Ahmad Hoirul Basori, Omar M. Barukab
Faculty of Computing and Information Technology Rabigh, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Article Info

Article history:

Received Jan 21, 2017
Revised May 1, 2017
Accepted May 14, 2017

Keyword:

Depth image recognition
Hand gesture
Presentation mechanism
Real time interactive

ABSTRACT

The research on human computer interaction. Human already thinking to overcome the way of interaction towards natural interaction. Kinect is one of the tools that able to provide user with Natural User Interface (NUI). It has capability to track hand gesture and interpret their action according to the depth data stream. The human hand is tracked in point of cloud form and synchronized simultaneously.The method is started by collecting the depth image to be analyzed by random decision forest algorithm. The algorithm will choose set of thresholds and features split, then provide the information of body skeleton. In this project, hand gesture is divided into several actions such as: waiving to right or left toward head position then it will interpret as next or previous slide. The waiving is measured in degree value towards head as center point. Moreover, pushing action will trigger to open new pop up window of specific slide that contain more detailed information. The result of implementations is quite fascinating, user can control the PowerPoint and event able to design the presentation form in different ways. Furthermore, we also present a new way of presentation by presenting WPF form that connected to database for dynamic presentation tools.

Copyright © 2017 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Ahmad Hoirul Basori,
Faculty of Computing and Information Technology Rabigh,
King Abdulaziz University,
Jeddah, Kingdom of Saudi Arabia.
Email: uchiha.hoirul@gmail.com

1. INTRODUCTION

Kinect is a camera that can provide user with depth stream image in real time using infrared and RGB lens. In the beginning, Kinect is designed for gaming console XBOX. Kinect can provide natural tracking control by combine with an algorithm for skeletal tracking. The Kinect device will make compute obtain the coordinate of the players in 3D and the environment, design the task much easier. It also identifies the users when they talk, detects the person identity, The it also able to track their activities and interpret them to position that creators can use to figure new immersion. The wireless motion sensing has been researched thoroughly for many years. Human tends to change the way of interaction as natural as they did in their daily life. However, technology evolution in the gaming industry has enabled motion sensing input devices to become a reality with the release of The Microsoft Kinect. Kinect become very popular recently, they have been used in many different fields such as: industries, medical or education. Kinect is one of stereo camera that provide depth image stream as tracking data. The key of Kinect is IR sensor that enables the tracking and poses recognition in real time.

Furthermore, Human Computer interaction has changing from time to time even since computers were invented. At the beginning, users only able to interact with computer using Command Line Interface (CLI) through keyboard. However it is also not good enough, human still not satisfied and they are looking for new interface which is trigger Graphical User Interface (GUI) born. GUI has been used more than

decades, it used widely and many people like this interface. Recently, a new idea has growing, human want to do interaction in natural ways that become beginning of Natural User Interface (NUI). Natural User Interface covers: Gesture, Brain Computer Interface or Voice/speech recognition. In this paper we will focus on gesture tracking as presentation control.

2. RELATED WORKS

The idea of interacting with computers in natural ways has existed since the old time when computer science needs some technology revolution. However, the idea stuck and remained as a dream for the past 40 years. Nowadays, these researches have triggered great process at universities, government and industries research in vide recognition and graphics, natural language and soft computing. At some stage, the research for NUI(Natural User Interface) have several approaches to starting with voice command, multitouch interface and ending with more unconventional methods like Microsoft Skinput project [1], muscle-computer interface [1] or brain computer interface using Electroencephalography (EEG) [2]. One of the first applications of NUI is the touch and multi-touch interface. They provide a more intuitive interaction with controls and applications than the standard cursor-based interaction. Instead of clicking on the graphical representation of an application the user can simply touch it. The Microsoft Kinect, the new revolutionary game controller for the Kinect was the real crucial moment for NUI, it is the first controller ever to enable to translate body movement into real time game actions without a need to hold any device in hands. In the beginning was only designed as part of xbox, however the device has been hacked few hours after its release, the depth data has been released to public and open a new ooportunity for natural interface application in some are asuch as military, medical, engineering, commercial and even education.

The multi-finger touch-interface give tangible sensation by putting the action through finger onto screen, so instead of the moving the cursor to an item and choose to open, user can touch the graphical representation directly in simply ways. Also, because of the ability to recognize more than one finger, the user can pinch two fingers to zoom in or out, providing a more natural and more intuitive control [3]. Additionally, through predefined motions and gestures, the user can interact with the system by tapping or dragging objects across the screen, or flicking the screen to scroll up or down. Furthermore, researcher has give more focus on touch-less interface where the invention of RGB and infrared sensor capable of recognizing the depth image and make computer to recognize the environment in three-dimensional space. This ability has overcome the regular sensor and open new prospect of recognizing the human body, face and even hand posture recognition. The body skeleton tracking will allow us to create an algorithm that can track gesticulation of human body while face recognition will give real time mimic expression. Finally, the hand recognition will enable the finger or hand posture recognition [3]. When the machine capable on recognizing the body movement and gesturesof human in real time, it will drive to new way of human-computer interaction called Touch-less Interface [4]. One of the first applications of Touch-less interface is moving the cursor across the screen using the user's hand. Combining Skeletal Tracking with Hand Detection the user can perform a click. First the user will take position in front of Kinect camera and raised the hands, then moves the cursor on the screen with his hand movements.

On the other hand, researcher also aims to improve realistic facial expression of virtual human to augment the interaction between human and virtual reality. Some numerous methods such as: fluid simulation for sweat and tears effect. the oxygenation concentration under skin also become consideration to adjust skin colour of avatar [12-16]. The interactivity between human and machine also can be performed through brain computer interface that read and analyze brain signal in accordance to the emotional condition of human. Subsequently, emotion state will be mapped to body action: walk, collision avoidance behavior and facial expression [17-20]. The previous statement has deliberate that Kinect has affect some sector such as medical, military or even engineering, they are producing an innovative ways of interaction for assisted surgery using new interface [21-23].

3. RESEARCH METHOD

This section will describe the methodology of the proposed project. The first stage is detecting the optimum object then it continues with depth image data stream processing. The next step is doing pattern classification with random decision forest. Afterward, the both hand will be tracked and synchronized simultaneously. This position will be mapped into controlling the presentation, refer to Figure 1.

Materials:

a. Material for experiments covers: Kinect Camera for depth image sensor
b. Computer with high end Processor and Graphics for processing

Methods:

a. Kinect Camera Positioning
b. Depth image data stream
c. Random Decision Forest Algorithm
d. Randomly choose a set of thresholds and features for splits.
e. Pick the threshold and feature that provide the largest information gain.

Repeat the process until a certain accuracy is reached or depth is obtained.The captured imaged data stream will be processed and analyzed random decision forest algorithm to decide what kind of action that human hand did during tracking process(refer to Equation 1).

$$P(c\backslash I,x) = \frac{1}{T}\sum_{t=1}^{T} P_t\,(c\backslash I,x) \quad (1)$$

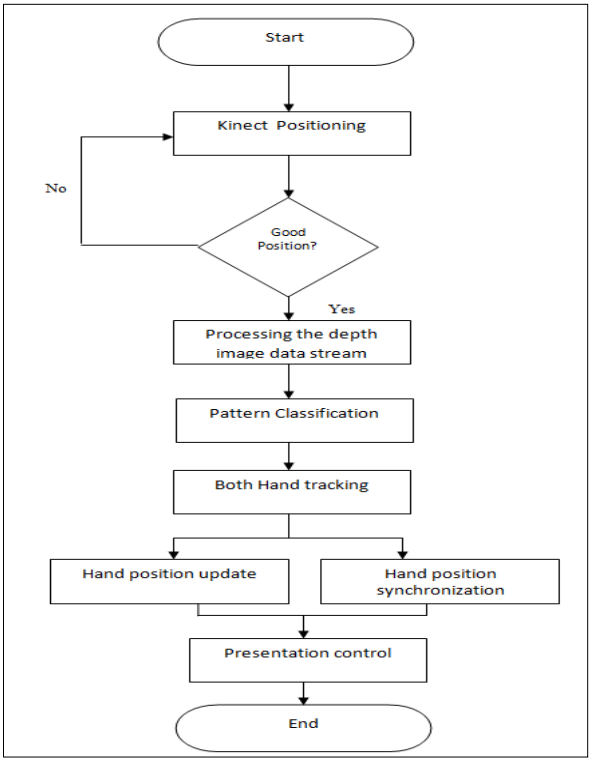


Figure 1. Methodology of the Research

4. RESULT AND DISCUSSION

The result has two main goals:

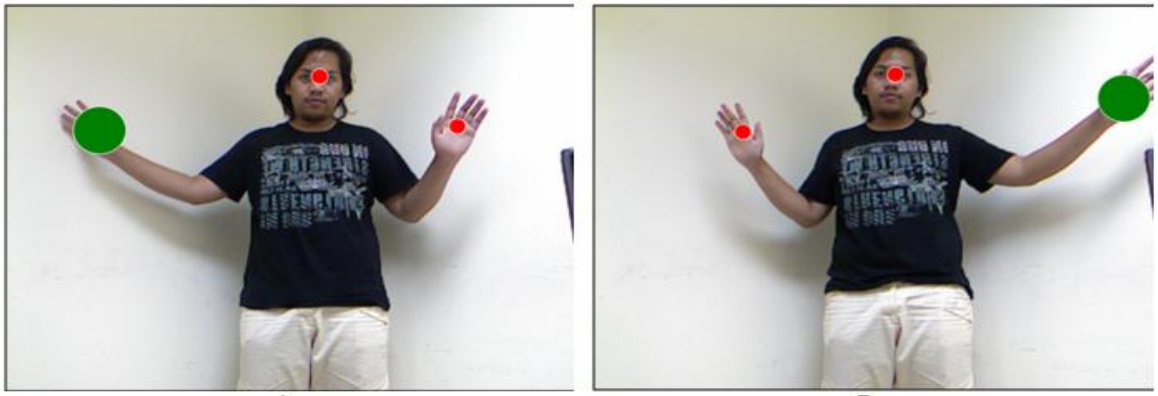
1. The first goal is controlling Microsoft Power Point. In this goal, our technique is work perfectly on navigating between slide and even stopping or play the PowerPoint just by waiving hand as human do in their daily life.
2. The second one is provide the new user interface for presentation media. The new interface can be treated as replacement of PowerPoint, because user can add and edit new information of their slide content through database. User also has menu to change the number of slide that they will used for their presentation.

The project has resulted two prototype, the first interface is only show the three main point tracked: head, right and left hand as shown in Figure 2.

In order to control the Microsoft PowerPoint user hand should be tracked properly. Once the user is in optimal distance in front of the camera, the Kinect can track the user's hands. When the user's hand move 45 degrees away from his head, the presentation will go to the next or previous slide depending on which hands the user used. If used the right hand, the camera will detect it and red circle on the right hand will turn green, it indicating the movement is confirmed and the presentation will go to the next slide, refer to Figure 3.

When user moves his left hand, the camera will detects the hand movement and the red circle on the hand will turn green and the presentation will go to the previous slide. (Figure 3B). WPF application is the second features of the proposed project that give user more interactive ways to control the presentation content. User can raise their hands, point at the camera with palms open, and interact with the system in different ways. The application prototype. The kinect presentation control can display the hand icon that move according to the user hand which is synchronized simultanoeusly.When user moves his hand, the cursor on the screen (shaped like a hand) will move corresponding with the movement and location of the user's hand, The hand is highlighted by red circle (Figure 4(a) and 4(b)).

Figure 2. Kinect presentation control with hand tracked



A B

Figure 3. (a). Moving the right hand-go to the next slide, (b) previous slide

A B

Figure 4. (a) Hand cursor that use for selection, (b) Selecting a thumbnail is indicated with cursor changes colours

If user move their hand until cursor hovers over a thumbnail, the cursor will change colours from white to magenta gradually indicating the selection of the thumbnail (Figure 4B). If the user wants to open the thumbnail, the user must make a small pushing motion with his hand; this action will locks the thumbnail to confirm the selection. When the user pulls his hand back, releasing the thumbnail; the thumbnail will open (Figure 5 and Figure 6).

Figure 5. Pushing on the Thumbnail

Figure 6. Selected Slide Content from Thumbnail is Displayed

The testing result is very interesting, users have more ways to interact with presentation, and they can move their hand over the thumbnail and make fast selection toward the navigation control. They can also control the Microsoft PowerPoint directly in efficient ways.

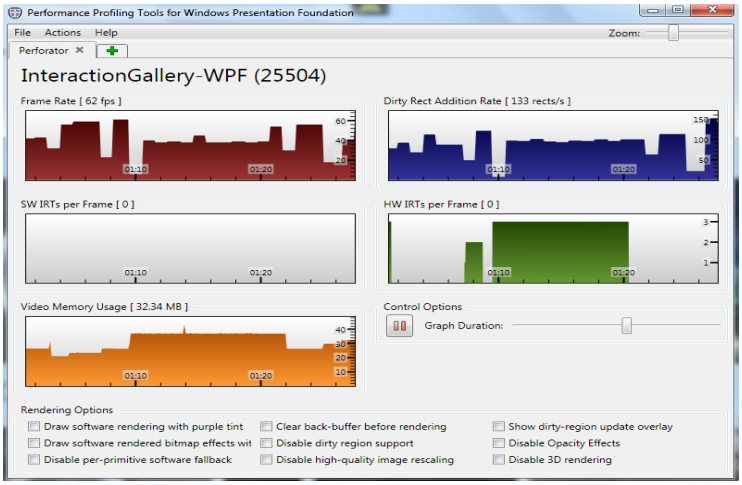


Figure 7. Performance Monitoring

The report of performance monitoring is measured through performance profiling tools for windows presentation foundation (Figure 7). The first graph reporting the rate at which the application is rendering to the screen (FPS), it is around 62 FPS and that more than normal fresh rate of monitor. The second measurement is Dirty Rectangle Addition rate that indicates how many rectangular regions that WPF needs to update each frame. The third performance indicator is SW IRTs Per Frame. It shows the number of software intermediate render targets required to render one frame of the application. The fourth one is HW IRTs Per Frame that Shows the number of hardware intermediate render targets required to render one frame of the application. Finally, the last part is video memory usage, it Tracks large allocations of video memory

made directly by WPF (textures and render targets). According to the reading, the video memory usage still low around 32.24MB. Overall, our proposed application is satisfying and can run perfectly even in lows spec of computer

5. CONCLUSION

Natural Interaction is where the future is headed, it's where science and fiction meet, what was once only seen as special effect in movies and games has now become a reality. It's one of scientists, gamers, developers, and average computer user's dream, to interact with computers and different technologies the same way we interact with other humans, by interacting naturally, However, The Kinect is still the beginning. By using The Kinect to give lectures and seminars, and using a more interactive system by utilizing the different features of The Kinect, we can give the students and audience a more visual and interesting experience in the meeting room, auditorium or even a class room. The Objective of this project has been achieved by presenting new ways of controlling the presentation, instead of using mouse or other presentation device, user only need to wave their hand and they can control everything in front of their eyes. This will drive people to interact with computer in more natural and interactive method. Although the technology almost looks ahead of its time, the device itself is still immature. For example, the device still needs to be plugged to an electrical outlet and it needs the user to stand close to the camera. Additionally, the camera is much bigger than a normal webcam so it won't fit in most laptop bags. In term of the prototype, there are still a lot of rooms for improvement such as improvement to control other Microsoft Office such as MS Word, Excel, etc. Hopefully in the future the device would be integrated in laptops so it's won't be a separate device, that way we can eliminate the size and cable issues. Or it can be wireless so it can be positioned in any part of the room. As for my software, I hope it can have more advanced ways of interaction, for example, we can zoom in on images or text by spreading both hands and then draw them together, and alternatively, we can spread our hands to zoom out. Another example is voice recognition, we can add any number of commands and it will be activated by user's voice, for example, saying "next slide" or "previous slide" will navigate through the slides of the presentation.

ACKNOWLEDGEMENTS

This work was supported by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah Saudi Arabia. The authors, therefore, gratefully acknowledge the DSR technical and financial support.

REFERENCES

[1] Donald A, Norman, 2010, "Natural user Interfaces are not Natural", Interactions 17, 3 (May 2010), 6-10. doi=http://dx.doi.org/10.1145/1744161.1744163.

[2] Webb, Jarrett, Ashley, James, "Beginning Kinect Programming with the Microsoft Kinect SDK", 2012, Apress, doi 10.1007/978-1-4302-4105-8.

[3] Andrea Corradini, 2001, *"Dynamic Time Warping for Off-Line Recognition of a Small Gesture Vocabulary",* In Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems (RATFG-RTS'01) (RATFG-RTS '01), IEEE Computer Society, Washington, DC, USA, 82-..

[4] Kenton O'hara, Gerardo Gonzalez, Graeme Penney, Abigail Sellen, Robert Corish, Helena Mentis, Andreas Varnavas, Antonio Criminisi, Mark Rouncefield, Neville Dastur, Tom Carrell. 2014. "Interactional Order and Constructed Ways of Seeing with Touchless Imaging Systems in Surgery", Comput. Supported Coop. Work 23, 3 (June 2014), 299-337. DOI=http://dx.doi.org/10.1007/s10606-014-9203-4

[5] Basori, AH, Qasim AZ, "Extreme Expression of Sweating in 3D Virtual Human", *Computers in Human Behavior*, 2014 Jan 1; 35:307-314. Available from, DOI: 10.1016/j.chb.2014.03.013

[6] Basori, A.H., et al., *"The Feasibility of Human Haptic Emotion as a Feature to Enhance Interactivity and Immersiveness on Virtual Reality Game",* in Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry. 2008, ACM: Singapore. p. 1-2.

[7] Alkawaz MH, et al., "Oxygenation Absorption and Light Scattering Driven Facial Animation of Natural Virtual Human", *Multimedia Tools and Applications*, pp 1-37. DOI=10.1007/s11042-016-3564-2

[8] Mohammed Hazim Alkawaz, Ahmad Hoirul Basori, Dzulkifli Mohamad, Farhan Mohamed, "Realistic Facial Expression of Virtual Human Based on Color, Sweat, and Tears Effects," *The Scientific World Journal*, vol. 2014, Article ID 367013, 9 pages, 2014. DOI=10.1155/2014/367013

[9] Alkawaz, M. H., Basori, A. H. & Mohd Hashim, S. Z. Multimed Tools Appl (2016), DOI=10.1007/s11042-016-3564-2
[10] Ahmad Hoirul Basori, "Emotion Walking for Humanoid Avatars using Brain Signals", *International Journal of Advanced Robotic Systems*, Vol.10, http://journals.sagepub.com/doi/abs/10.5772/54764, DOI=10.5772/54764
[11] Muhamed Abdul kareem Ahmed, Ahmad Hoirul Basori, *"The Influence of Beta Signal toward Emotion Classification for Facial Expression Control through EEG Sensors"*, Procedia Social and Behavioral Science, Elsevier, 6 Nov 2013, DOI: 10.1016/j.sbspro.2013.10.294
[12] Basori Ahmad Hoirul, Bade A, Sunar M. S., Daman D, Saari N, Hj. Salam, MD. S (2012), "An Integration Framework of Haptic Feedback to Improve Facial Expression", *International Journal of Innovative Computing, Information and Control (IJICIC)*, Vol.8, No.11, November 2012
[13] Nazreen Abdullasim, Ahmad Hoirul Basori, Md Sah Hj Salam, Abdullah Bade, "Velocity Perception: Collision Handling Technique for Agent Avoidance Behavior", *Telkomnika*, Vol. 11, No. 4, April 2013, pp. 2264 ~ 2270
[14] Yusman Azimi Yusoff, Ahmad Hoirul Basori, Farhan Mohamed, *"Interactive Hand and Arm Gesture Control for 2D Medical Image and 3D Volumetric Medical Visualization"*, Procedia Social and Behavioral Science, Vol.97, 6 November 2013, Pages 723–729, Elsevier
[15] Mohammad RiduwanSuroso, Ahmad Hoirul Basori, Farhan Mohamed, *"Finger-based Gestural Interaction for Exploration of 3D Heart Visualization"*, Procedia Social and Behavioral Science, Vol. 97, 6 November 2013, Pages 684–690, Elsevier
[16] John Hardy, Jason Alexander, 2012, *"Toolkit Support for Interactive Projected Displays"*, In Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM '12), ACM, New York, NY, USA, Article 42, 10 pages. DOI=http://dx.doi.org/10.1145/2406367.2406419.
[17] Ahmad Hoirul Basori, Fadhil Noer Afif, Abdulaziz S, Almazyad, Hamza Ali Abujabal, Amjad Rehman, Mohammed Hazim Alkawaz, 2015, "Fast Markerless Tracking for Augmented Reality in Planar Environment", 3D Res. 6, 4, Article 72 (December 2015), pp 1-11. DOI=http://dx.doi.org/10.1007/s13319-015-0072-5
[18] Ahmad Hoirul Basori, Daut Daman, Abdullah Bade, et, al., 2008, *"The Feasibility of Human Haptic Emotion as a Feature to Enhance Interactivity and Immersiveness on Virtual Reality Game"*, In Proceedings of The 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry (VRCAI '08), ACM, New York, NY, USA, Article 37, 2 pages. DOI=http://dx.doi.org/10.1145/1477862.1477910
[19] AH Basori, A Tenriawaru, ABF Mansur, "Intelligent Avatar on E-learning using Facial Expression and Haptic", *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, Vol. 9 (1), 115-124, 2011

BIOGRAPHIES OF AUTHORS

Ahmad Hoirul Basori, received B.Sc(Software Engineering) degree from Institut Teknologi Sepuluh Nopember Surabaya in 2004 and the Ph.D (Computer Graphics) from Universiti Teknologi Malaysia, Johor Bahru, Johor, in 2011. From 2004 to present, he was a lecturer with the Department of Informatics, Faculty of Information Teknologi, Institut Teknologi Sepuluh Nopember Surabaya, Indonesia. In 2011, he has appointed as Assistant Professor with the Department of Computer Graphics and Multimedia, Universiti Teknologi Malaysia. In 2016, he is promoted to Associate Professor rank in Faculty of Computing and Information Technology Rabigh, King Abdulaziz University. He is the member of Editorial board of some international journal, more than 60 articles, and also a member of professional membership IEEE, ACM SIGGRAPH, IAENG and Senior Member of IACSIT. His research interests include Computer Graphics, Facial Animation, Cloth Simulation, Medical Visualization, Haptic Interaction, Man Machine Interaction and Robotics.

Omar Obarukab receive Doctorate degree from Computer Engineering, College of Engineering, Florida Institute of Technology in 1999. Currently, he is Dean of Faculty of Computing and Information Technology Rabigh, He also active on research and teaching. His research interest: Logic - mobile agent - cryptography - data mining - Information security and audit.