Skeleton Tracking using Kinect Sensor & Displaying in 3D Virtual Scene

[1]Chanjira Sinthanayothin, [2]Nonlapas Wongwaen, [3]Wisarut Bholsithi
*1, First Author & Corresponding Author *National Electronics and Computer Technology Center, NSTDA, cephsmile@gmail.com, chanjira.sinthanayothin@nectec.or.th*
[2]*National Electronics and Computer Technology Center, NSTDA, nonlapas.wongwaen@nectec.or.th*
[3]*National Electronics and Computer Technology Center, NSTDA, wisarut.bholsithi@nectec.or.th*

Abstract

Current research on skeleton tracking techniques focus on image processing in conjunction with a video camera constrained by bones and joint movement detection limits. The paper proposed 3D skeleton tracking technique using a depth camera known as a Kinect sensor with the ability to approximate human poses to be captured, reconstructed and displayed 3D skeleton in the virtual scene using OPENNI, NITE Primesense and CHAI3D open source libraries. The technique could perform the bone joint movement detections in real time with correct position tracking and display a 3D skeleton in a virtual environment with abilities to control 3D character movements for the future research.

Keywords*: Skeleton Tracking, 3D Virtual Scene, OpenNI, NITE Primesense, CHAI3D, Kinect*

1. Introduction

Computer and animation technology application in three dimensional (3D) games has made significant progress. Detecting the movement of people in 3D can be applied, but many limitations exist for detecting the motion of the bones and joints in real-time. Since the tracking of bones and joints is advantageous in a character rigging for animated 3D movement, most 3D games record the movement of bone joints in the database. This recorded data will be applied for changing the character's position and movements and display quickly. Thus, this paper gives an overview of the state of the art and currently available literature and techniques related to body or skeleton tracking techniques. Some techniques are used to detect the body movement in 2D. Other segmentation techniques are designed to be used in bone and joint detection. Many schemes are proposed for tracking the body skeleton in 3D. The common aim of all these techniques is to develop an automated tracking body or skeleton movement which can help create a digital character animation in 3D or control through some devices such as TV, robot and so forth. This paper also presents a discussion on depth camera and libraries that can be applied for skeleton tracking as well. Then the paper proposes the process for developing the prototype to detect the movement of the bones in 3D with a Kinect sensor and display in 3D virtual scene.

2. Summarization of reviewed motion and skeleton tracking techniques

The researchers have surveyed a number of approaches and application examples for body motion and skeleton tracking techniques. Body tracking or skeletal tracking techniques using an ordinary camera are not easy and require extensive time in developing. The summarizations of reviewing body or skeletal tracking techniques are presented by the technique used along with advantages, disadvantages and illustrations as shown in Table 1 for the technique referred in [1] to [8], Table 2 for the technique referred in [9] to [15], and Table 3 for the technique referred in [16] to [23] respectively.

3. Implementation

From the body or skeleton tracking survey above, to detect the bone joints of human movement is still a major problem because the depth of the human body cannot be determined through the use of a typical camera. However, we have tried to use more than one video camera to detect and determine the depth of the object, but the consequence is that the cost increases and the process ability have been slowed down due to the increased data processing. Fortunately, with the availability of an infrared or depth camera such as a Kinect sensor makes it possible to acknowledge the depth of the object.

Real time human pose recognition technique on depth image is based on the paper by Shotton and et. al. [23] combined with the unsupervised learning technique by Weber [24] and segmentation images with the Layout Consistent Random Field Technique by Winn and Shotton [25].

Table 1. A survey table of advances in vision-based human motion capture and analysis along with Pros, Cons and illustrations of different techniques.

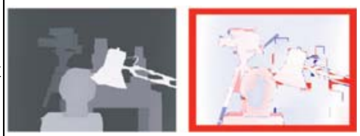
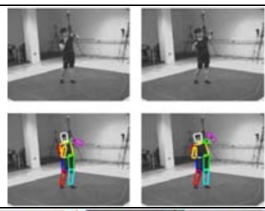
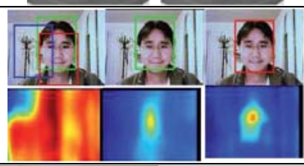
Techniques	***Pros***	***Cons***	***Illustrations***
Real-time Decentralized Articulated Motion Analysis and Object Tracking from Videos [1]	Fast and Easy to Implement.	Fails to provide satisfactory results for the case of self-occlusion due to the fact that it cannot handle pose relation between two adjacent parts.	
Binocular stereo algorithm to estimate scene Structure [2]	1. Reconstruction with sub-pixel disparity. 2. Precise boundary localization.	Poor result on the rough surface that has to be solved by adaptive parameters in the stereo algorithm to adjust segmentation coarseness & the amount of surface details.	
Kalman and Particle Filters constrained by human biomechanics to track human position [3]	Ability to handle any bipedal motion without being constrained to specific activities.	The pivot points are generally occluded for an extended period of time which is undesirable.	
Gaussian Process Annealed Particle Filter (GPAPF) algorithm [4]	1. Less error and More Robustness than Hierarchical Annealed Particle Filter (H-APF). 2. High Accuracy on Motion Classification.	1. Wrong motion classification in the case of hugging. 2. Need cross validation to classify ambiguous types of motion.	
Hierarchical space-time model (HSTM) with Histogram of Gabor Orientations (HIGO).[5]	Computational is efficient for searching human action in the video.	The query longer than two seconds has not been tested yet.	
Fast nonparametric belief propagation for real-time stereo articulated body tracking [6]	Fairly robust on arm movements to various human tracking position.	1. Failed to cope with complex occlusion 2. Slow processing at 10 frames per second.	
Visual tracking via dynamic tensor analysis with mean update (DTAMU) [7]	Effective and robust algorithm for human movement tracking, especially in the changing environments.	This algorithm has not been tested on 3D image environment yet.	
Real-time Body Tracking Using a Gaussian Process Latent Variable Model [8]	The algorithm of this system can track the body movement very well.	The algorithm will have a hard time to track the movements which has a no prior motion model.	

Table 2. A survey table of advances in vision-based human motion capture and analysis along with Pros, Cons and illustrations of different techniques (extended 1).

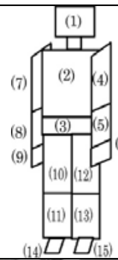
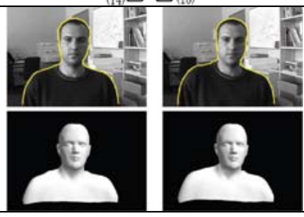
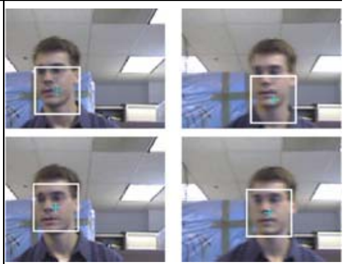
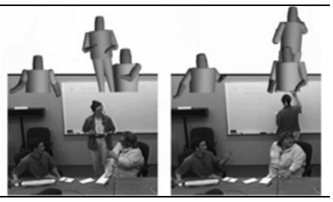
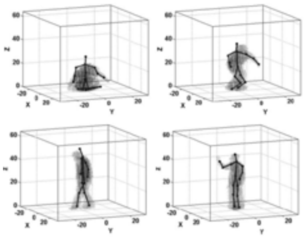
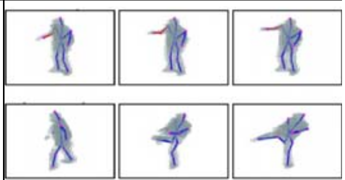
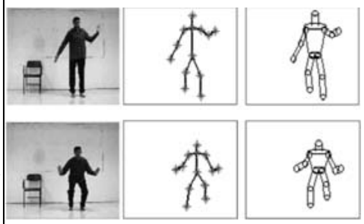
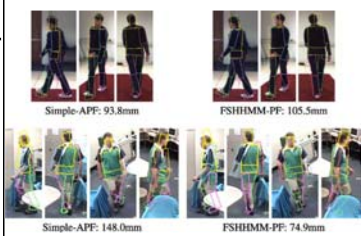
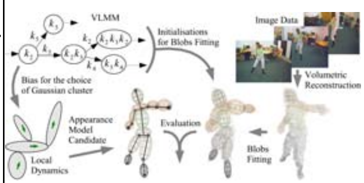
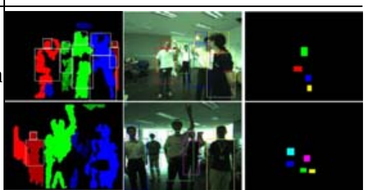
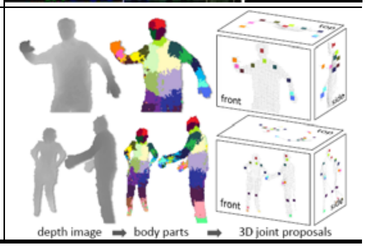
Techniques	***Pros***	***Cons***	***Illustrations***
A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package [9]	Accurate tracking could be achieved by applying Kalman filter to eliminate drift error while making the angular rate measurements from the rate gyro which enhances the gyro dynamic and linear acceleration disturbance from accelerometers.	1. Time lag generated by Kalman filter 2. The conflict between time lag vs. random error; the reduction of time lag has caused larger random error.	
Classification-based approach to finding texture boundaries [10]	1. Ability to deal with occlusion on highly texture objects for 3D pose estimation and 2D delineation. 2. Optimal performance and speed with robustness.	Not achieving an automatic algorithm for adjusting the classification window size for better tracking of texture boundary on 3D pose.	
Fast Global Kernel Density Mode Seeking [11]	1. Converge faster with smaller number of iteration than regular mean shift (MS) tracking algorithm. 2. No overshoot during the converging process. 3. Ability to handle the large displacement of video frame which regular mean shift (MS) failed.	Need further discriminative features other than color histograms for better localization and tracking performance.	
Human Pose Tracking in Monocular Sequence Using Multilevel Structured Models [12]	Full pose inference giving less position error and weighing average error than blob tracking alone.	Long processing time for rendering so it is unsuitable for real time application.	
Action and Gait Recognition From Recovered 3-D Human Joints [13]	Exemplar-based hidden Markov model (EMHMM) has given the average 3D pose recognition rate about 94% for testing data, better results than previous researches.	Having a problem on the recognition of punch action due to the differences between punch action by male and female subjects.	
Adaptive particle Filter with Body Part Segmentation for Full Body Tracking [14]	Ability to handle the fast movements with changing viewpoints from different actors.	1. Inability to perform real time tracking due to the algorithm complexity. 2. Need a fine human model to replace the old one to minimize the errors on movement tracking.	
Self-Organizing Map (SOM) for behavior detection of gesture [15]	Ability to track moving palm by using SOM and skin color for adaptive background update.	Still having a problem on tracking the palm movements of multiple subjects.	

Table 3. A survey table of advances in vision-based human motion capture and analysis along with Pros, Cons and illustrations of different techniques (extended 2).

Techniques	***Pros***	***Cons***	***Illustrations***
Human pose modeling and body tracking using Discrete Wavelet Transform (DWT) for Skin Color Segmentation [16]	The initialization and segmentation problem of the vision based markerless motion capture of the human body can be solved by DWT to eliminate background clutter with the average of 89.5% accuracy on 3D pose correction.	1. This algorithm has been tested only for frontal poses while the lateral poses have not been testes yet. 2. This algorithm has not been tested on multiple subjects.	
2D and 3D Upper Body Tracking with One Framework by Chain Graph (CG) and Dynamic Bayesian Network (DBN)[17][18]	1. Good tracking results 2. No drift problem in all sequences. 3. Generic algorithm for different body motion. 4. Ability to handle self-occlusion of body part.	1. Need algorithm for temporal smoothness of the motion. 2. Need at least 5 states of appearance on human activities to ensure the good tracking results.	
Non-rigid body object tracking using fuzzy neural system based on Multiple ROIs and Adaptive Motion Frame Method [19]	An adaptive motion frame method using panoramic images is applied to improve the tracking ability.	High computational complexity which requires modification for real time applications on 3D stereo images.	
Tracking-as-Recognition for Articulated Full-Body Human Motion Analysis [20]	1. The factored-state hierarchical hidden Markov model (FS-HHMM) can handle the cases of occlusions, poor segmentation & reduced resolution by fusing human body tracking with action recognition. 2. Auto-initialization and self-correction to handle the tracing failure is possible.	1. Tracking results are not so accurate tracking as annealed particle filter (APL) for the case of clean observation and sedate motion. 2. Difficult to track the Fridge action. 3. Need better particle filter in high-dimension human body tracking.	
Real Time Markerless Human Body Tracking by Colored Voxels & 3-D Blobs with learnt models of behaviors [21]	Broad range of body movement captured while eliminating jitters by Monte–Carlo Bayesian framework and a variable length Markov model for self-recovery.	1. Tracking for multiple subjects has not tried yet. 2. Need dimensionality reduction on the learning cluster 3. Need an online learning to handle unseen sequence.	
ML-Fusion base Multi-Model Human Detection and Tracking for Robust Human-Robot Interfaces [22]	Fast and Easy to Implement.	Fails to provide satisfactory results for the case of self-occlusion due to the fact that it cannot handle pose relation between two adjacent parts.	
Real-time human pose recognition in parts from single depth images [23].	1. Quickly and accurately predicts 3D positions of body joints from a single depth image, using no temporal information. 2. Ability to run the classifier in parallel on each pixel on a GPU to increase the speed.	Using large and highly varied training dataset to estimate body parts invariant to pose, body shape, clothing, etc. to pose the relation between two adjacent parts.	

3.1. Libraries to Control Kinect Sensor

Table 4. A Comparison Table for Different NUI libraries along with Pros and Cons.

#	*Techniques*	*Pros*	*Cons*
1.	OpenNI/NITE	1. Supports both Microsoft Kinect and Xtion Pro LIVE. 2. Several methods ready to use. 3. Very popular, with many applications in various fields. 4. Have skeleton tracking. 5. Available for most languages. 6. Any OS compatible application.	1. Relatively difficult to install. 2. Calibration Posture is in need.
2.	Libfreenect	1. Several applications. 2. Available for most languages. 3. Any OS compatible application.	1. Very difficult to install for beginner. 2. No skeleton tracking.
3.	CL NUI	1. Ability to capture the broad range of body movements. 2. Small jitter due to the camera noise can be filtered with ease.	1. It cannot perform motion prediction. 2. No learned soft constraints to handle the case of severe occlusions.
4.	Microsoft Kinect SDK	1. Widely known in the community of robotic systems. 2. Easy to install, fairly widespread.	1. Poor High level API. 2. Support for Windows only. 3. No skeleton tracking. 4. Available for only C/C++ and C#.
5.	Evoluce SDK	1. Easy to install. 2. Ready to use methods for gesture recognition. 3. Have skeleton tracking.	1. Support for Window7 only. 2. Calibration Posture is in need. 3. Available for only C/C++ and C#

For our implementation, outputs from Kinect have been applied through the Natural User Interface (NUI) Library. The five known NUI libraries include OpenNI (Open Natural Interaction) [26], NITE Primesense [27] which is used in the paper, libfreenect [28], CL NUI [29], Microsoft Kinect SDK [30], and Evoluce SDK [31]. A comparison for five different NUI libraries along with advantages and disadvantages are described in Table 4.

In this paper, we introduce the skeleton tracking using Microsoft Kinect sensor. Kinect sensor generates a depth map in real time, where each pixel corresponds to an estimate of the distance between the Kinect sensor and the closest object in the scene at that pixel's location. Based on this map, the Kinect system allows further developing applications to accurately track different parts of the human body in three dimensions.

OpenNI is designed such that applications can be used independently of the specific middleware and therefore allows further developing codes to interface directly with OpenNI while using the functionality from NITE Primesense Middleware. The main purpose of NITE Primesense Middleware is an image processing, which allows for both hand-point tracking and skeleton tracking. The technique for 3D skeleton tracking using Kinect sensor and displayed in 3D virtual scene can be explained by the flowchart in Figure 1.

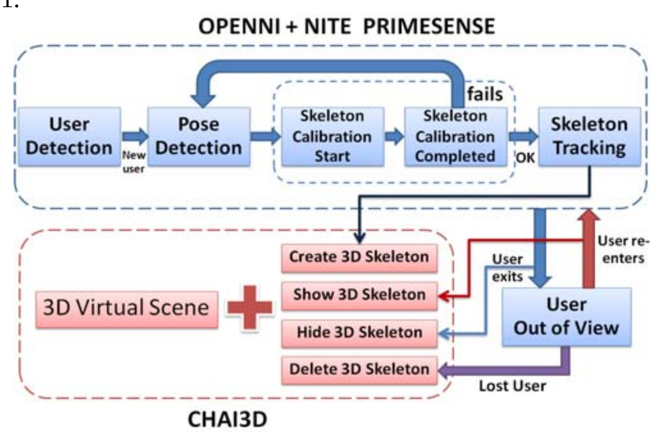


Figure 1. Flowchart of skeleton tracking and displayed in three dimensions.

3.2. CHAI3D Library[32][33]

When a new user is detected, the standard user pose is detected so the user's skeleton can be calibrated and tracked. The start of tracking triggers the creation of a new 3D skeleton using CHAI3D (Computer Haptics & Active Interfaces 3D), and its attachment to the scene graph. The skeleton is made invisible when the user leaves the Kinect's view, and visible again when it returns. If the user remains out of sight for several seconds, then the user is considered lost and his/her skeleton is deleted from the scene graph. For the creation of a 3D skeleton and 3D virtual scene is handled by a skeleton 3D object, which creates the 3D joints and limbs (i.e. spheres and cylinders) using an open source library named CHAI3D. The scene graph by CHAI3D is shown as in Figure 2.

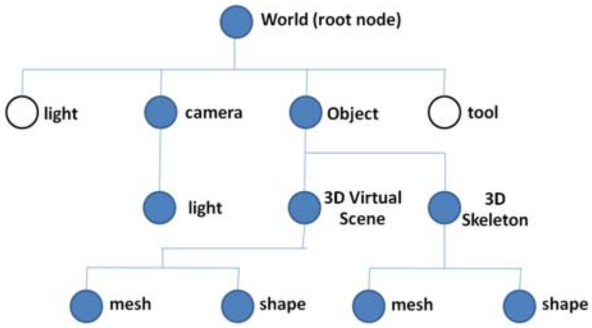


Figure 2. Scene graph by CHAI3D.

3.3. Kinect Sensor [34]

Kinect is capable of producing depth, texture, user information and skeleton information which can be explained in detail. The depth map uses the input from the IR cameras; The texture map is the RGB color map of the scene that can be recorded just like any RGB cameras. The user map is an output of binary images that would include the detected people in the scene. If the configuration is done, Kinect is able to detect the limbs of a person and extract them as information of the scene. However to be able to do that, the person has to stand in front of the Kinect, with the whole body visible, and has to stand for a few seconds with both hands in the air. This process might takes from 10 seconds or a little bit more depending on the positions of the Kinect sensor. Once the calibration is done, Kinect tracks the joints and limbs position as can be seen in Figure 3.

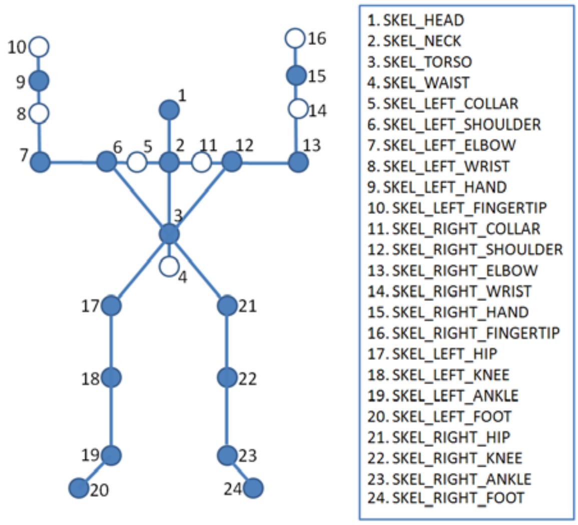
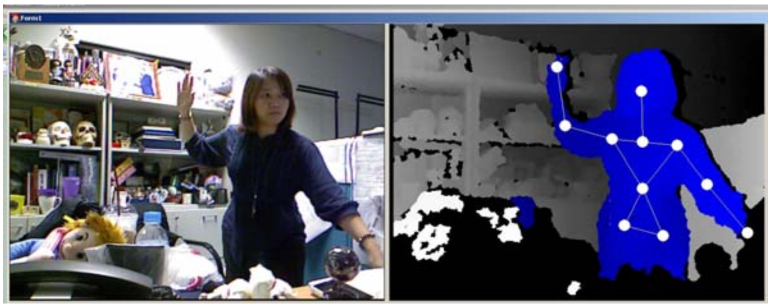


Figure 3. The Skeleton model of Joints and Limbs.

The joint circles without color have not been applied in our prototype. If the person goes out of the frame but comes in really quick, the tracking continues. However, if the person stays out of the frame for too long, Kinect recognizes that person as a new user once she/he comes back, and the calibration needs to be done again. Although Kinect is highly talented, many some limitations exist as well such as Kinect is not capable of detecting the skeleton of the person without doing the calibration.

People in the scene also need to be in the field of view of the camera as well. Kinect does not require to see the whole body if the tracking is configured as the upper-body only as seen in Figure 4 with the color image from Kinect shown in Figure 4(A) while the 2D skeleton tracking on depth image from Kinect is shown in Figure 4(B).



(A) (B)

Figure 4. Process for 2D skeleton tracking and display on top of depth image using Kinect sensor.

However, Kinect is not indeed introduced for developers by Microsoft but for the Xbox gaming console. Therefore, it is indeed impossible to get the mentioned data from Kinect. However, third part companies have developed software to make Kinect work with a PC and develop applications on it as well. Microsoft approves those open source frameworks. In our experiments we use the OpenNI framework designed for Kinect and similar sensor that can make it works when connected to a PC and receive the data it generates. On top of that, we have produced our own prototype for saving the data and processing it. Once the skeleton has been tracked, the 3D skeleton joints are displayed in the 3D virtual scene as shown in Figure 5 with the use of the open source named CHAI3D.

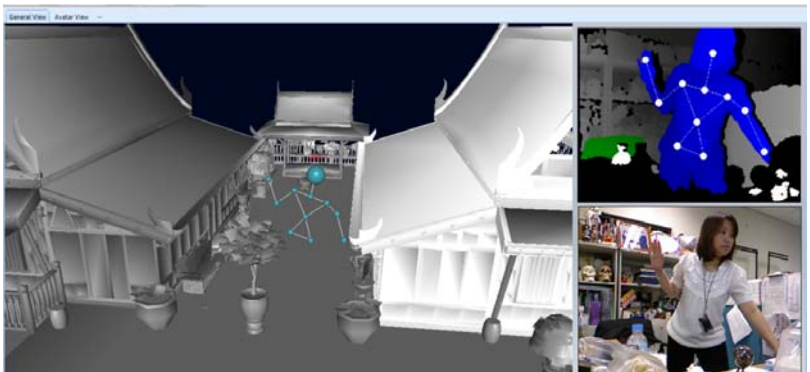


Figure 5. 3D skeleton tracking in 3D virtual scene.

Normally, CHAI3D can be used as a high level tool to quickly create haptic scenes. In our prototype, the base class world can be populated by various virtual objects (based on implicit surfaces or on meshes) whose physical characteristics (mass, texture, color, initial position and velocity, etc) can be saved in 3DS and OBJ formats. OpenGL-based graphic rendering of objects and tools is automatically taken care of by the CHAI3D engine.

4. Results

We have implemented and tested our prototype on the 3.0GHz CPU with a single core. Our system can run at real-time with 30 frames per second. The results of 3D skeleton tracking in 3D virtual scene has been demonstrated and tested in the NECTEC-ACE 2011 Conference on IT at the Thai event. Many people have been interested and tried to detect the movement of their own. Our prototype can detect the movement and display the user skeleton in 3D virtual environment in real time and the positions of the joints are presented accurately. Our prototype can detect many users' movements simultaneously as can be seen in Figures 6, 7 and 8. Figure 6 shows the graphics user interface for 3D skeleton tracking in 3D virtual scene. Figure 7 shows the 3D skeleton tracking in different virtual scenes in real time. Figure 8 shows the virtual scene in different views when the user has interacted with the 3D scene by changing the camera direction.

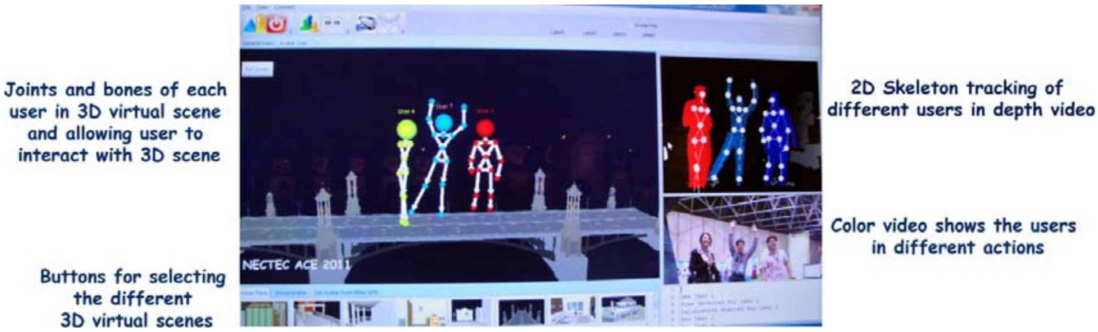


Figure 6. 3D skeleton tracking in 3D virtual scene.

Figure 7. 3D skeleton tracking in 3D virtual scene.

Figure 8. Virtual scene in different views when the user interacts with the scene by changing the camera directions.

5. Discussion & Conclusion

After conducting a variety of techniques associated with the review body/skeleton tracking using typical cameras, we found that the development of these techniques takes extensive development. Thanks to the recent development in depth-camera technologies with a single consumer-grade depth-camera, approximated human poses can be captured and reconstructed in real-time. Hence, we consider the application development with a Kinect Sensor. Our application/prototype can be concluded as follows:

1. Can displaying the video of RGB and depth during the user movement.
2. Can detect the movement of the bone structure in 2D with a Kinect sensor and display on top of depth image.
3. Can detect the movement with 3D structure of the bones and display in 3D virtual scene.
4. Can detect the movement of 3D skeleton in real time with multiple users simultaneously.
5. Bones and joints can be displayed in 3D model in different colors with the name of users on top of head joint.
6. Users can interact with the 3D virtual scene with rotation and zoom (zoom in & zoom out) functions while the user can see the 3D skeleton in a variety of perspectives.
7. Can display 3D virtual environment in a variety of formats (3DS and OBJ). This virtual environment can be adjusted without an interruption of the motion tracking.

This application is also important in virtual reality. To realize the interaction between humans and the virtual environment, the body movement needs to be tracked. In interactive games, motion can be captured and used to drive game characters to give the game player a new experience of participation, such as the Kinect for Xbox 360. Motion capture is very useful in physical training as it makes the traditional training based only on experience enter into the digital era. The Microsoft Kinect brought these capabilities to a new level, eliminating the need of a hand-held controller and allowing a precise position detection of the user in 3D space. This evolution extended greatly the possibilities of such a device from simple video games to much more, including its usage in virtual reality simulation. For further development, our 3D skeleton tracking will be used to control the 3D character in real time. Furthermore, many users can have a meeting or interact with each other from different places via the internet. For the matters of applications other than online games, it can be applied for sport training simulation on different events as mentioned in [35], risk accident simulation as a prevention measure in [36][37], training simulation in [38] and so on.

6. Acknowledgments

This research is under the project “MC-Avatar: Multi-Camera Reconstruction and Augmentation of Human in Virtual Reality”, which is financially supported by Discovery and Development Grant (DD grant) from National Science and Technology Development Agency (NSTDA), Thailand.

7. References

[1] A Wei Qu, Dan Schonfeld, “Real-Time Decentralized Articulated Motion Analysis and Object Tracking From Videos”, IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2129-2138, August 2007, DOI: 10.1109/TIP.2007.899619.

[2] Michael H. Lin, Carlo Tomasi, “Surfaces with occlusions from layered stereo”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 8, pp. 1073-1078, August 2004 , DOI: 10.1109/TPAMI.2004.54.

[3] Jesús Martínez del Rincón, Dimitrios Makris, Carlos Orrite Uruñuela, Member, and Jean-Christophe Nebel, “Tracking human position and lower body parts using Kalman and particle filters constrained by human biomechanics”, IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics, vol. 41, no. 1, pp. 26-37, February 2011, DOI: 10.1109/TSMCB.2010.2044041.

[4] Leonid Raskin, Michael Rudzsky, Ehud Rivlin, “Dimensionality reduction using a Gaussian Process Annealed Particle Filter for tracking and classification of articulated body motions”, Computer Vision and Image Understanding, vol. 115, no. 4, pp. 503–519, April 2011, DOI: 10.1016/j.cviu.2010.12.002.

[5] Huazhong Ning, Tony X. Han, Dirk B. Walther, Ming Liu, Thomas S. Huang, “Hierarchical Space–Time Model Enabling Efficient Search for Human Actions”, IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 6, pp. 808-820, June 2009, DOI: 10.1109/TCSVT.2009.2017399.

[6] Olivier Bernier, Pascal Cheung-Mon-Chan, Arnaud Bouguet, “Fast nonparametric belief propagation for real-time stereo articulated body tracking”, Computer Vision and Image Understanding, vol. 113, no. 1, pp. 29-47, January 2009, DOI: 10.1016/j.cviu.2008.07.001.

[7] Xiaoqin Zhang, Xinghu Shi, Weiming Hu, Xi Li, Steve J. Maybank, “Visual tracking via dynamic tensor analysis with mean update”, Neurocomputing, vol. 74, no. 17, pp. 3277–3285, October 2011, DOI: 10.1016/j.neucom.2011.05.006.

[8] Shaobo Hou, Aphrodite Galata, Fabrice Caillette, Neil A. Thacker, Paul A. Bromiley, “Real-time Body Tracking Using a Gaussian Process Latent Variable Model”, In Proceedings of the 11th IEEE International Conference on Computer Vision (ICCV 2007), pp. 1-8, 14-21 October 2007, DOI: 10.1109/ICCV.2007.4408946.

[9] Rong Zhu and Zhaoying Zhou, “A real-time articulated human motion tracking using tri-axis inertial/magnetic sensors package”, IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 12, no. 2, pp. 295-302, June 2004, DOI: 10.1109/TNSRE.2004.827825.

[10] Ali Shahrokni, Tom Drummond, François Fleuret, Pascal Fua, “Classification-Based Probabilistic Modeling of Texture Transition for Fast Line Search Tracking and Delineation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 3, pp. 570-576, March 2009, DOI: 10.1109/TPAMI.2008.236.

[11] Chunhua Shen, Michael J. Brooks, Anton van den Hengel, “Fast Global Kernel Density Mode Seeking: Applications to Localization and Tracking”, IEEE Transactions on Image Processing, vol. 16, no. 5, pp. 1457-1469, May 2007, DOI: 10.1109/TIP.2007.894233.

[12] Mun Wai Lee, Ramakant Nevatia, “Human Pose Tracking in Monocular Sequence Using Multilevel Structured Models”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 1, pp. 27-38, January 2009, DOI: 10.1109/TPAMI.2008.35.

[13] Junxia Gu, Xiaoqing Ding, Shengjin Wang, Youshou Wu, "Action and Gait Recognition From Recovered 3-D Human Joints", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, vol. 40., no.4, pp. 1021-1033, August 2010, DOI: 10.1109/TSMCB.2010.2043526.

[14] Junxia Gu, Xiaoqing Ding, Shengjin Wang, Youshou Wu, “Adaptive particle filter with body part segmentation for full body tracking”, In Proceedings of The 8th IEEE International Conference on Automatic Face & Gesture Recognition (FG '08), pp. 1 – 6, 17-19 September 2008, DOI: 10.1109/AFGR.2008.4813346.

[15] Youfu Wu , Gang Zhou, Jing Wu, “A Monitoring System for Supermarket Based on Trajectory of Palm”, IJACT: International Journal of Advancements in Computing Technology, vol. 2, no. 1, pp. 7 - 15, March 2010, DOI: 10.4156/ijact.vol2.issue1.1.

[16] Karuppanan Srinivasan, K. Porkumaran, Gopala Sainarayanan, “Skin Colour Segmentation based 2D and 3D Human Pose Modelling Using Discrete Wavelet Transform”, Pattern Recognition and Image Analysis, vol. 21, no. 4, December 2011, pp. 740-753, DOI: 10.1134/S105466181104016X

[17] Lei Zhang, Zhi Zeng, Qiang Ji, “Probabilistic Image Modeling With an Extended Chain Graph for Human Activity Recognition and Image Segmentation”, IEEE Transactions on Image Processing, vol. 20, no. 9, September 2011, DOI: 10.1109/TIP.2011.2128332.

[18] Lei Zhang, Jixu Chen, Zhi Zeng, Qiang Ji, “2D and 3D upper body tracking with one framework”, Proceedings of the 19th International Conference on Pattern Recognition (ICPR 2008), pp. 1-4, 8-11 December 2008, DOI: 10.1109/ICPR.2008.4761484.

[19] Hyunsoo Lee and Amarnath Banerjee, “Non-rigid body object tracking using fuzzy neural system based on multiple ROIs and adaptive motion frame method”, In Proceedings of The 2009 IEEE International conference on Systems, Man and Cybernetics, San Antonio, Texas, pp. 3871-3876, 11-14 October 2009, DOI: 10.1109/ICSMC.2009.5346633.

[20] Patrick Peursum, Svetha Venkatesh, and Geoff West, “A Study on Smoothing for Particle-Filtered 3D Human Body Tracking”, International Journal of Computer Vision, vol. 87, no. 1-2, pp. 53-74, March 2010, DOI: 10.1007/s11263-009-0205-5.

[21] Fabrice Caillette, Aphrodite Galata, Toby Howard, “Real-time 3-D human body tracking using learnt models of behaviour”, vol. 109, no. 2, February 2008, pp. 112–125, DOI: 10.1016/j.cviu.2007.05.005.

[22] Li Yuan Li, Jerry Kah Eng Hoe, Shui-Cheng Yan, Xin Guo Yu, “ML-fusion based multi-model human detection and tracking for robust human-robot interfaces”, The 2009 Workshop on Applications of Computer Vision (WACV 09), pp. 1-8, 7-8 December 2009, DOI: 10.1109/WACV.2009.5403083.

[23] Jamie Shotton, Andrew Fitzgibbon, Mat Cook, Toby Sharp, Mark Finocchio, Richard Moore, Alex Kipman, Andrew Blake, “Real-time human pose recognition in parts from single depth images”, In Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1297 – 1304, 20-25 June 2011, DOI: 10.1109/CVPR.2011.5995316.

[24] Markus Weber, Unsupervised Learning of Models for Object Recognition, Thesis for the Degree of Doctor of Philosophy, California Institute of Technology, Pasadena, California, May 2000, DOI: 10.1.1.90.8680.

[25] John Winn, Jamie Shotton, “The layout consistent random field for recognizing and segmenting partially occluded objects”, In Proceedings IEEE Computer Vision and Pattern Recognition (CVPR), New York, pp. 37-44, June 2006, DOI: 10.1109/CVPR.2006.305.

[26] PrimeSense Ltd., Willow Garage, Side-Kick Ltd., ASUS Inc., AppSide Ltd. OpenNI^{TM}: Introducing OpenNI, *http://www.openni.org/*, 2010.

[27] PrimeSense Ltd. NITE Primesense Middleware, *http://www.primesense.com/en/nite*, 2011.

[28] The OpenKinect Community. OpenKinect (libfreenect), *http://openkinect.org/wiki/Main_Page*, 2011.

[29] Code Laboratories Inc. CL NUI Platform - Kinect Preview, *http://codelaboratories.com/nui/*, 2010.

[30] Microsoft Inc. Kinect for Windows: Develop What's Next, *http://www.microsoft.com/en-us/kinectforwindows/develop/*, 2012.

[31] WIN&I Team; Evoluce AG., INCREON GmbH. Evoluce SDK, *http://www.evoluce.com/_win-and-i/en/software/overview/index.php?we_objectID=55*, 2011.

[32] François Conti, Federico Barbagli, Remis Balaniuk, Maurice Halg, Charity Lu, Dan Morris, Luis Sentis, Elena Vileshina, James Warren, Oussama Khatib, Kenneth Salisbury, “The CHAI Libraries”, Proceedings of Eurohaptics 2003 (Eurohaptics '03), Dublin, Ireland, pp. 496-500, 6-9 July 2003, DOI: 10.1.1.83.1087

[33] François Conti, Federico Barbagli, Dan Morris, and Christopher Sewell, “CHAI 3D: An Open-Source Library for the Rapid Development of Haptic Scenes”, Proceedings of the First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (IEEE World Haptics 2005 - WH05), Pisa, Italy, 18-20 March 2005.

[34] Tommer Leyvand, Casey Meekhof, Yi-Chen Wei, Jian Sun, Baining Guo, “Kinect Identity: Technology and Experience”, Computer, vol. 44, no. 4, April 2011, DOI: 10.1109/MC.2011.114.

[35] Xiang LIU, Jinhai SUN, Yaping HE, Yimin LIU,Li CAO, “Overview of Virtual Reality Apply to Sports”, JCIT: Journal of Convergence Information Technology, vol. 6, No. 12, pp. 1 - 7, 2011

[36] Cai Linqin, Zheng Xuesong, Qu Hongchun, Luo Zhiyong, "Risk Accident Simulation Using Virtual Reality and Multi-agent Technology ", JDCTA: International Journal of Digital Content Technology and its Applications, Vol. 5, No. 2, pp. 181 - 190, February 2011, DOI: 10.4156/jdcta.vol5. issue2.21.

[37] Cai Linqin, Pan Yuyou, Cen Ming, Yu Jimin, "Multi-agent Based Simulation for Underground Safety Accidents", JCIT: Journal of Convergence Information Technology, Vol. 6, No. 11, pp. 309 - 316, November 2011, DOI: 10.4156/jcit.vol6.issue11.35

[38] Gang Chen, “Design and Realization of Equipment Training Simulator”, Journal of Convergence Information Technology Volume 5, Number 4, pp. 15 - 22, June 2010, DOI: 10.4156/jcit.vol5.issue4.2