

# Using Kinect to Develop a Smart Meeting Room

Hui-Huang Hsu

Dept. of Computer Science & Information Engineering  
Tamkang University  
Taiwan  
h\_hsu@mail.tku.edu.tw

Yu-Ren Chen

Dept. of Computer Science & Information Engineering  
National Central University  
Taiwan  
yrchen@ATCity.org

Yi Chiou

Dept. of Computer Science & Information Engineering  
National Central University  
Taiwan  
joeychiou@mine.csie.ncu.edu.tw

Timothy K. Shih

Dept. of Computer Science & Information Engineering  
National Central University  
Taiwan  
TimothyKShih@gmail.com

**Abstract**—Observe the most progressive conference system in the market, a higher meeting quality is always being pursued with better equipment, no matter on the personalized laptop, touch pad, direct microphone, or other high quality recording video equipment. It is known that to establish this system with low price is difficult. This paper aims to solve this predicament and contributes a higher quality to the smart meeting room. Hence, we combine Microsoft Kinect and Bluetooth techniques to build a smart conference system with personalized Bluetooth supported equipment to identify each participant's identity, and use these IDs to search the central database in order to retrieve his/her contact information (phone number, email, web storage space and text, or meeting records). Consequently, the system uses Kinect as a gesture recognition device to detect each person's skeletons with multi functions (Controlling the computer, sending information, auto-uploading files, or recording personal meeting records in database). Smart meeting room can be established with simplified equipment in this way.

In this smart meeting room, participants don't need to bring additional equipment but only to use his/her Bluetooth supported cell phone. As for the arrangement of the meeting room, besides the central computer, we only need Kinect to provide the gesture recognizing and detecting works, and to record the meeting records (text and video). This simplified smart conference room can provide great benefits for the users; this is the main purpose of this paper.

**Keywords**—Kinect; Mobile Device, Bluetooth Technology, Smart Meeting Room, Gesture

## I. INTRODUCTION

Conferencing system development abound on the market now, but looking ahead to the conference system, seems to be a hardware competition, manufacturers tend to launch a higher standard conferencing equipment in order to get better results of the meeting, but the fact provided on participants convenience real purpose is the conference system.

A simple analysis of the purpose of the meeting, the participants in the meeting space to discuss with each other, at the same time the relative audio, video, of the records of

meetings or conference participants to share with each other the contents of the file, either by mail or Internet space, to file sharing this is today's meeting the most basic of processes and purposes. Therefore, a simplified meeting space, you can take advantage of at least equipment to achieve the easiest of performance is the best policy! Simplify the equipment and environment. How do I know the identity of the participants as well as information, how will measure the sitting position when participating in the meeting of different systems. After this, the interaction of the participants of the proceedings, the process the record, these will be an important issue facing the conference system.

As technology advances, mobile devices and somatosensory technology to flourish, build anywhere immediately to participate in the meeting, as well as during the conference somatosensory interactive manipulation of the environment of the conference system. Kinect equipped with the devices and their characteristics, creating multiple skeletal information system environment can be retrieved through the Bluetooth devices paired with the user's mobile device, in order to obtain the identity of the participants and personal information, which can conferencing systems will crawl labeled corresponding to the identity of the skeleton as the purpose of the conference system of identity and minutes of meetings. Right to be heard when someone uses a relatively gesture, it can obtain the system of control of it, whether the user wants to report the slide or the operation of the system. Even more of the system since learned to crawl as a backbone, so participants in the meeting discussion, the conference video record retrieved through Kinect's RGB camera, but also the way in the film marked identity or cutting individual speak fragment stored in the central personal database, in other meetings in the future want the temporary scheduling corresponding to the statement, can also be easy to obtain.

In addition, if the participants even if not in the site, but also through the video camera of the mobile device, as well as far-reaching mobile network, reaching anywhere instant meeting demand, and finally, the participants of the transmission of the file stored in this system also has the function of a substantial portion of the share of the flat file

can be shared through the following ways, SMS, Email, network Neighborhood. If the meeting is in progress, the participants asked to share its files or messages, you can use the previous screen positioning via Bluetooth device and retrieve personal position and information to achieve file sharing Kinect required for identifying shared identity, whose contact information is automatically placed in the message file messages shared for the purpose of sharing groups, you can quickly reach.

This study will be launched in the system configuration settings at any meeting space, participants only need to carry a mobile device with a Bluetooth device using a simple Bluetooth communication device to determine the identity of participants, as well as scheduling personal information from the database. The meeting space is also just a central computer and a Kinect needed to achieve the minimum system requirements, you can at anytime, anywhere to open a highly convenient interactive meeting.

## II. RELATED WORK

We hypothesized that the participants are with their respective mobile devices to replace today's personal computer (PC) or laptop (Laptop) participating in the meeting, which must identify each mobile device corresponding to the actual conference system which position to construct virtual conferencing systems scenes, then combined the Kinect somatosensory recognition and gesture recognition, to complete a system suitable for small and medium-sized video conferencing meetings. Therefore, this study explores the related applications and research, to find the most suitable for application in this study on which the technology.

### A. Bluetooth Technology

In the literature [1], the use of sound sensors on a mobile device, to sense the measured ambient sound signal, via distance and filter background noise to determine the current environment as well as the position corresponding to Fig.1a [1] sound sensing system architecture diagram. [2] using a mobile phone and your computer's Bluetooth device to communicate, implement a peer-to-peer (Peer-to-Peer) data transmission system, by way of Bluetooth (Bluetooth Transmission) messages or files on the phonetransfer between computers. [3] is passed through the wireless network, so that two phones can pass on to each other, one's own GPS address information is sent to the other party, and each other's position is marked on one's own digital map, Fig.1b provided by the system structure from literature [3].

In the above-mentioned study, the proposed transfer information between mobile devices and mobile devices or positioning. Proposed above, this study uses the Bluetooth technology "for the identification of the identity of participants and the main message the way. Bluetooth devices (e.g. Smartphone with Bluetooth support) as the main message across for the following reasons:

1. At present, the development of mobile devices increasingly sound, smart phones and tablet PCs and

other portable embedded systems, have switched equipped with Bluetooth devices.

2. Bluetooth devices compared to the infrared device, RFID, such as lower cost, and have been widely used in the many 3C products.
3. The Bluetooth transmission without complicated wiring, wireless communication and transmission range of up to 100 meters, suitable for small and medium-sized conference venue.
4. Bluetooth security is quite good, you can set encryption to protect its hopping spread spectrum technology, and not subject to electromagnetic interference.
5. Bluetooth transmission can be one-to-one or one-to-many transmission, public messages or private messages sent through its characteristics, maximum support 8 Bluetooth Device, in line with the needs of small and medium-sized video conferencing.
6. Bluetooth, no direction, compared to infrared aligning the transmission port transmission, the power consumption is lower than the infrared transmission.

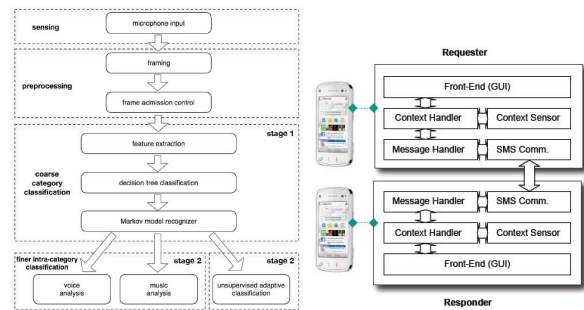


Fig. 1. (a) The Sound Sensing system Architecture form [1]. (b) Phone-To-Phone Context Sharing Architecture from [3].

Bluetooth device characteristics As described above, the system will use the Bluetooth device to construct a good interoperability and low cost and power consumption of network transmission architecture; each participant owned by Bluetooth device Bluetooth ID (Bluetooth ID), as the identity of the participants in the conference system representatives to allow participants to send private messages, transfer files to other specific needs of participants in the meeting.

### B. Kinect Interaction Technology

In 2010, Microsoft released "Kinect", applied to the XBOX360 console peripherals. Kinect replaces the use of hand-held or foot game controller, replaced by detecting the user's voice commands or gestures, movements to operate the game interface.

In addition to the somatosensory (Body Sensor) concept of the game is booming, with the Microsoft Windows version of Kinect in 2012, a wide range of applications in

response to born. The Kinect be different from the general camera, its ready the three photographing lens, the left and right sides of the lens for the 3D structure of the infrared emitter and an infrared camera CMOS "optical depth sensors, as well as the middle of a RGB color photographic lens, in addition to coordinates the recovery coke technology, the base motor rotation along with the focus moving objects. Kinect has more features, allow developers to through this device to capture color images with depth information and analysis through in-depth information to get captured by the lens of the human backbone, and then dynamically tracking the movement of people in the picture situation.

Traditional camera can only get the plane color information, so to obtain the body of information to be analyzed; you must go to a lot of math in order to remove the shooting posture, node information. Kinect makes the emergence of human information can be easily extracted and achieve the purpose of real-time analysis and feedback. Kinect is one to obtain the depth information of the captured content of the photographing device, the depth of the sensor can obtain the actual 3D position in the 2D image, which can help us to quickly extract from the human body movements of the camera will be analyzed using.

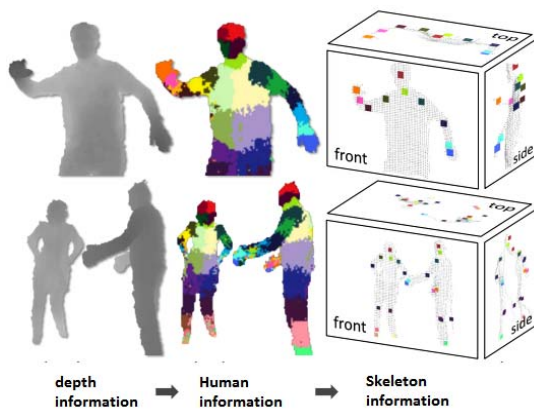


Fig. 2. Get the skeleton diagram [4].

Shotton et.al. [4] proposed a fast and accurate method, accurate predictions of the 3D position of the body joints through a large number and variety of movements and training to analyze a variety of different body structure, body size, posture, action, and even clothing and classified integration record. Finally, the screen will get the record of human information do than to obtain a more accurate and recognized in a variety of different body posture, and the results are presented to the user. Fig.2 [4] actual human skeleton diagram, in this study, are able to achieve the purpose of the skeleton retrieved to obtain the depth of information and the results of the human information is both accurate and fast.

Latest Kinect for Windows SDK 1.5 in not only grab the skeleton of the body, even the face information can also retrieve, you can catch the location of the various parts of the human face, construct a polygon face (it should be a total of 87point), and thus can determine the angle of the head, the

face of the read parameter [2], and added for the seated skeleton Tracking (skeleton Tracking) only through the upper skeleton tracking node, so that greatly enhance the performance of the information analysis.

Manipulation of computer punch card machines by the early evolution through the era of the mouse, keyboard, to the advent of Kinect somatosensory technology, Kinect is directly receive the body's most natural body language (Body Language) tool, so there are many studies define diverse gestures, expected to provide an interactive environment makes human-computer interaction more natural, with user-friendly operation.

In the Fig. 3, FFAST: The Flexible Action and Articulated Skeleton Toolkit [6] Integration of a variety of body movements correspond to the various acts in the virtual reality space or on any possible manipulation of the computer (keyboard, mouse, etc.).



Fig. 3. (a)When the user assumed a posture. (b)The PrimeSense NITE software will be able to recognize its posture, and the skeleton of the screen corresponding to its posture. (c)Using FFAST users can manipulate a virtual foot color in the system.

### III. PROPOSED METHOD

More and more conferencing system starting use the Kinect skeleton retrieve and somatosensory detection, the same technologies are also widely been used in a variety of different fields. Using simple body movements to replace the original mouse operation. In addition, with multiple skeleton detection technology, which can detect more than one person's actions, to achieve the purpose of multiplayer operation. Typically, the original purpose conference systems for the people are involved and focus the same screen. The traditional method of using a central control is not convenient enough to meet the needs of the participants. In this study, we use skeleton detection and Bluetooth identification to identify and tracking participants. To support remote meeting, we develop a central synchronization system to maintain the process of meeting. The flowchart of system is shown in Fig.4. The details of how we use Kinect to track skeleton and using hand gestures to control the meeting system will be discussed in below sections.



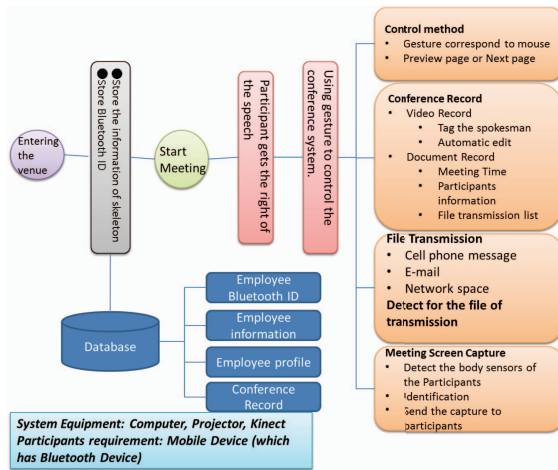


Fig. 4. The flowchart of the system.

### • Skeleton tracking improvements

As shown in the Fig. 5, the skeleton tracking of Kinect can identified 20 joint points in same time.

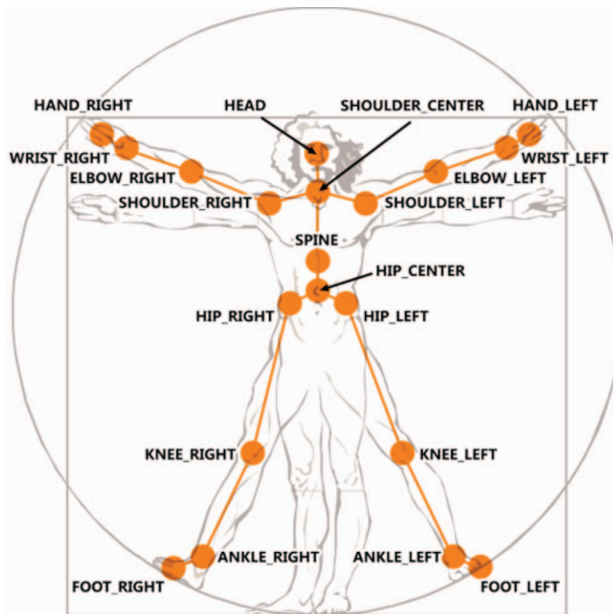


Fig. 5. Kinect for Windows SDK track 20 joints of the body

In addition to the Kinect SDK originally skeleton key points Location, absolute position of the system using the new Kinect SDK 1.5 development environment for the joint direction of information, this information can be divided in space and the relative position of the two modes for the human skeleton. With these directions information in gesture discrimination and somatosensory detection, more layers of auxiliary determine to a significant reduction in the program's misjudgment, or to avoid the constant use of the direction determination function to reduce this conference

system overall burden for the subsequent gestures determination and control, this improvement is to promote the progress of the performance.

In addition, one computer can use four Kinect sensors at the same time. Although no obvious help in the general application, in this study conferencing system, which is an integral part of the technology. That the proposed smart meeting system, the tracking portion of the number of participants, a Kinect is the most appropriate track number four, more will be a significant reduction in its effectiveness. If a computer on the next four Kinect, means the number of people participating in the meeting can be increased to more than 16 people, more than one Kinect synchronization technology developed by our laboratory can be achieved at the same time to retrieve the information of multiple Kinect and ready switch mastership to speak to the end where Kinect device, the next picture shows developed in our laboratory of a computer connected to four units Kinect, facing four different directions sync screen.

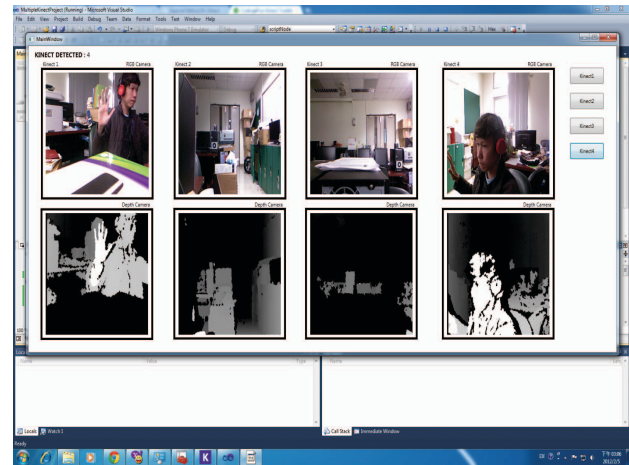


Fig. 6. Multiple Kinect synchronization program developed in the laboratory can be divided into capture images as well as the depth of information.

### • Kinect gesture control

In previous versions, the Kinect for Windows SDK only supports user standing skeleton track, the original if you need to use the sitting position tracking method, developers need to develop its own program, but for now, the effects are not so ideal. Today SDK version 1.5 finally joined the sitting skeleton tracking, under this mode, the system keeps track of only ten key points of the upper body, As for the lower body joints are ignored, so, for the skeleton of the user sitting Tracking will have a better effect.

For the proposed conference system, skeleton tracking for participants, sitting skeleton tracking a large number of improved performance, all sitting around a conference room in the meeting participants, there will be a better the skeleton track the effect, the system determines that the gesture will have better recognition rate.

Know the principle and mode of operation of the operation of the Kinect, Kinect recognize the user's movement and gesture, for the purpose of manipulating the central computer, the research and development of a series of computer guided by the body movements corresponding to the operation, its use in meeting, the report is not to be toward the front fixing computers, not to be armed with any hardware, just need to face the Kinect device can use a specific gesture control. .

Fig.7 report by the time of reporting, the need to control the slide on the screen, only the left and right swing, you can reach the slide up and down the page scroll. .

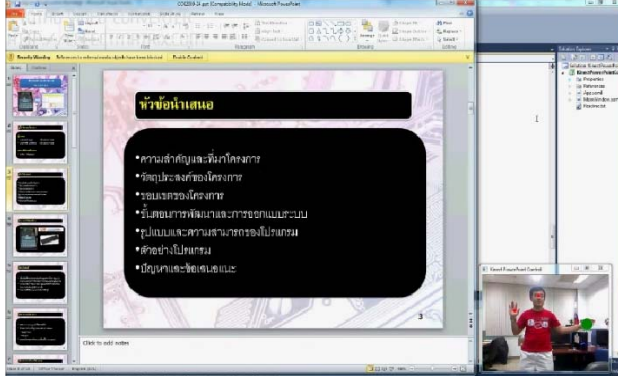


Fig. 7. Developed in our laboratory actual manipulation slide screen, the lower-right corner of the manipulator, waving gesture control slide meeting space equipped with: central computer, projector machine, Kinect.

Typically the user is to use the mouse to control the computer. For the meeting, each person who wants to control the computer needs to pick up the mouse to control the computer is not convenient. So we use Kinect and hand gestures to manipulate the computer mouse, and with specific gestures to trigger click event.

#### IV. EXPERIMENT

In this study, we proposed to build a smart meeting room. The next section explains the simulation progress of experiment, including the information of participants sensed by Kinect. Each Kinect can tracking up to four people at same time. In our simulation, we develop a synchronize system to support four Kinects in one room. As shown in Fig.8, we setup one computer connecting with one Kinect. Each participant has one own smartphone with Bluetooth support. .

Process and results of the experiment are shown in Fig.9. First step, after the equipment of meeting room been setup complete, the system will try to pair with the first participant's smartphone via Bluetooth. We use Bluetooth ID to identity user and download relative information of user from database. Furthermore, the system will use Kinect to retrieve the skeleton information of participant and binding with Bluetooth ID, stored in the system. Second, repeat the process to identity and track the rest participants.



Fig. 8. Venue, a single Kinect, a central computer, three participants and three Bluetooth-enabled mobile phone.

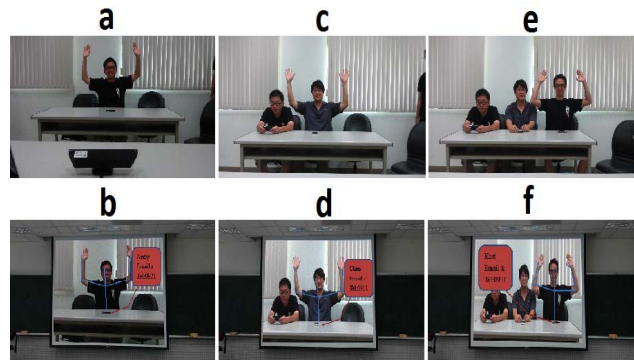


Fig. 9. Participants into the stadium, the operation of the system.

After all participants entered the meeting room, all information including Bluetooth ID, skeleton and relative information of users will project on local and remote screen. All participants can watch the screen and stored as a part of record for meeting. .

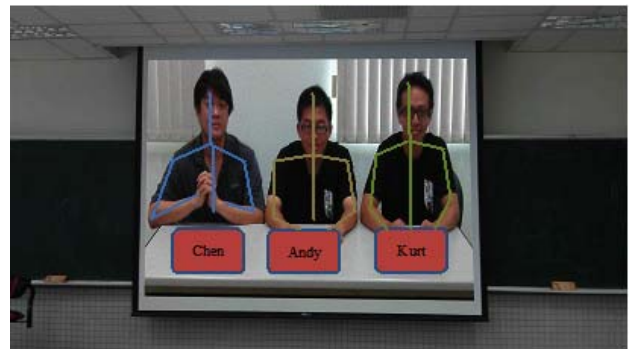


Fig. 10. Admission to all participants, the complete identities after the identification of the skeleton conference screen.

In the meeting, each participant can record by entering the venue of the skeleton information through the gesture to control the conference system. As shown in Fig.11, Fig.11a and Fig.11b. One of the users do a special movements to get the obtain control of the speech. Also user can control the slide show through gestures. Fig.11c and Fig.11d show how to switch previous and next slides through gestures. Fig 11e

shows the presenter transmission the file to specific user, which raised hand. Fig.11f shows the gesture of control cursor movement. The right hand control the position of cursor, the left hand trigger the click event of cursor.

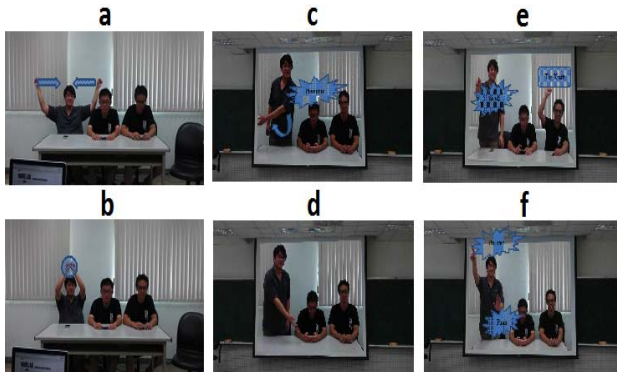


Fig. 11. The actual use of gesture control case conference system.

## V. CONCLUSIONS

In this study, the successful development of an intelligent meeting space, whether for any meeting space where a user can still use the same simple equipment, to achieve wisdom, interactive session results. This study is only needs one simple Kinect, one computer and one mobile device. You can make an environment through the method for intelligent interactive, not only improved the way that information is retrieved skeleton status, making the gesture control the skeleton of information to improve application performance and equipment made easy. It's low cost, combined with interactive technology to create a successful life with the wisdom of the cloud computing system.

## ACKNOWLEDGMENT

This research was supported in part by National Science Council of Taiwan under the grant NSC 101-2622-E-032-005-CC3.

## REFERENCES

- [1] Hong Lu, Wei Pan, Nicholas D. Lane, Tanzeem Choudhury and Andrew T. Campbell, "SoundSense: Scalable Sound Sensing for People-Centric Applications on Mobile Phones", MobiSys '09 Proceedings of the 7th international conference on Mobile systems, applications, and services, pp. 165-178, 2009.
- [2] Mohammed A. Qadeer, Rahul Agrawal, Amit Singhal, Sarosh Umar, "Bluetooth Enabled Mobile Phone Remote Control for PC", ICACC '09 Proceedings of the 2009 International Conference on Advanced Computer Control, pp.747-751, 2009.
- [3] Juong-Sik Lee, Umesh Chandra, "Mobile phone-to-phone personal context sharing", Communications and Information Technology, Proceedings of the 9th international conference on Communications and information technologies, pp. 1034-1039, 2009.
- [4] 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:Supplementary Material", Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR),2011.
- [5] Lu Xia, Chia-Chih Chen and J. K. Aggarwal, "Human Detection Using Depth Information by Kinect", International Workshop on

Human Activity Understanding from 3D Data in conjunction with CVPR (HAU3D), Colorado Springs, CO, June 2011.

- [6] E. Suma, D. Krum, B. Lange, S. Rizzo, and M. Bolas, "FAAST: The flexible action and articulated skeleton toolkit", In IEEE Virtual Reality, pp. 247-248, 2011.
- [7] Anandram Sundar, Adarsh Kowdle, "3D interactive interface using a Kinect sensor", a design project report presented to the engineering division of the graduate school of Cornell University, May, 2011.