Virtual Touch Wall by Kinect: Review

Author
Ms. AshviniGaneshsinghBais
M.E. student, CSE Department,
Sipna college of engineering, Amravati-444701
{baisashu@gmail.com}

Guide
Prof. Pritish A. Tijare
CSE Department
Sipna college of engineering,
Amravati-444701

Abstract:- Computing advances and increased smart phone use gives technology system designers greater challenge in flexibility and exploiting computer vision to support visually impaired users. Understanding these user's needs will certainly provide insight for the development of improved usability of computing device.
As new technologies and computer applications prove to be powerful tools for children's with special needs in order to improve specific skills. However, there is still a gap between research development and its applicability in schools, based on their classroom education proposed framework with several activities. This project focus on HCI (Human Computer Interaction) and Face tracking method based on Kinect, Kinect has the advantage over ordinary camera because it has 2 sensor, an ordinary and depth sensor. In this project a method based on depth information is used for optimizing the face recognition combining with hand gesture which can switch automatically for students and teacher easy to operate the PPT screencast system and other teaching activities related to education domain.

Index words: *Kinect, Natural Interaction, Educational Content, etc.*

*****

1. Introduction:

1.1The Microsoft Kinect Sensor

Recently for capturing both depth images and color many consumer products are available, such as Microsoft's Kinect. Microsoft Kinect sensors created many opportunities for multimedia computing.

Microsoft announced the release of Kinect sensor for windows Software Development Kit (SDK). These SDK will potentially transform HCI (Human Computer Interaction) in multiple industries, education, healthcare, retail, transportation & beyond. The Kinect sensor incorporates different advanced sensing hardware. It has an RGB camera and a dual infrared depth sensor: a projector and an infrared sensitive camera on the same band as well as a microphone array to capture the surrounding environment [4]. It runs with proprietary software whichprovides full-body 3D motion capture, facial recognition, and voice recognition capabilities. The RGB camera has a resolution of 640 480 pixels, while the infrared camera uses a matrix of 320 240 pixels or 640 x 480 pixels.

Kinect is a horizontal device with depth sensors, color camera, and a set of microphones with everything secured inside a small, flat box. The flat box isattached to a small motor working as the base that enables the device to be tiltedin a horizontal direction. The Kinect sensor includes the following key components:

- Color camera
- Infrared (IR) emitter
- IR depth sensor
- Tilt motor
- Microphone array
- LED

1.2Structure of Kinect sensor

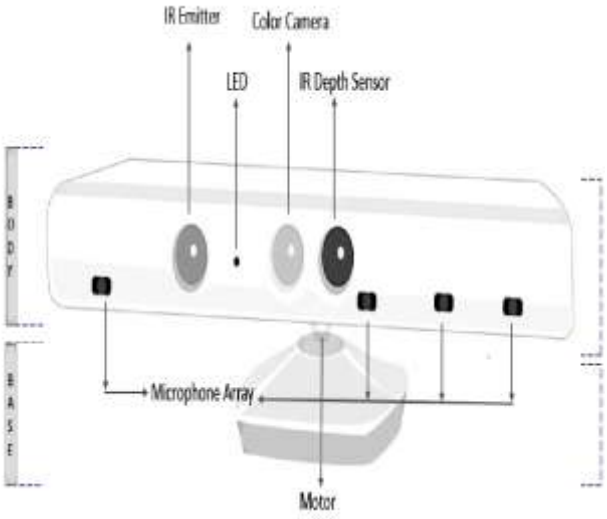

Figure1.2. Microsoft Kinect Sensor

1.3 Where can we use Kinect

- Capturing real-time videos using the RGB color sensor.
- Tracking a human body and then responding to its movements and gestures as a natural user interface
- Measuring the distances of objects and responding.
- Analyzing 3D data and making a 3D model and measurement about it
- Generating a depth map of the objects that tracked by Kinect
- Recognizing a human voice and developing hands-free applications which can be controlled by voice.

With this you can build a number of real-world applications that fall under various domains

1.4 Human Gesture Recognition

"Gestures are human body movements. It is non-vocal communication, which used in combination with verbal communication, intended to express meaning. It may be combined with the hands, arms or body & also can be a movement of the head, face & eyes." Human gestures are of different names such as human pattern, human behavior.

A gesture is nothing but an action that intended to communicate with emotions or intentions to the device. Gesture recognition has been a prime research area for a long time. However, in the last decade, a phenomenal amount of time, effort, and resources have been devoted to this field and in the wake of the development of devices. Gesture recognition allows people to interface with a device and interact naturally with body motion, as shown with the person in the following picture, without any device attached to the human body. Gesture recognition can be conducted with techniques from computer vision & image processing to interpret sign languages & identify human behaviors & postures [1]. It can build a richer bridge between machine & humans than primitive text, user interfaces.

1.5. Speech Recognition

The Kinect device consists of array of 4 microphones that supports a multitude of audio features. The Kinect device has four separate downward-facing micro-phones that are placed at the bottom of the Kinect device in a linear pattern.

The microphone array allows the following:

- Capturing better quality sound and providing inbuilt signal processing, including noise suppression and echo cancellation

The source direction of the incoming sound is identified

Based on the sound to each microphone in the array, it can automatically find out the direction that the sound is coming from and listen to a specific micro-phone by suppressing the other noises. Once the direction of the sound source is set, the Kinect sensor is intelligent enough to change the direction as and when the source moves. One of the common examples of such a scenario is playing a game using voice commands. If the player moves, the sound source's direction moves automatically. The Kinect sensor has an inbuilt audio-processing pipeline that takes care of the complete audio-processing capabilities. Another important aspect of the Kinect microphone array is speech recognition. The microphone array helps to recognize human speech very clearly by focusing only in a particular direction and canceling noises in the environment.

2. Related Work & Limitation

- **Vision based:**

Vision based technique is used to distinguish between fingers of both the hands which have been touch .In this method result can be achieved by using two cameras mounted above the surface, also hand image is also captured on surface to distinguish between the hand postures.

- **Pen & Touch based Interaction:**

As name suggest pen & touch surface is used in this method by pen provide input to make contact with surface.

- **Color based tracking:**

To detect human hand is to use the color based filter. Every human hand has a specific color. The basic idea is to use a set of threshold ranges for every image channel separately. The person has to be of specific skin color, constant lighting conditions and a captured image have same skin color characteristics with no other objects. Such as black people don't have the same range of threshold like white people.

As above related work concerned with touch gestures recognition:

Multi touch and multidimensional manipulations are focused on developing interaction gestures, exploring the manipulations of multi touch on multiple degrees. In this case direct touch manipulation have coordination and parallelism. It emphasizes more on the surface consistency. By jointly estimate pose & shape obtaining space results in more robustness.

As RGB color model, where every color is described by the intensity of these three basic colors. Red, Green, Blue. There are high correlation between these components and luminance mixing with the chromaticity makes this color model very sensitive to as light condition changes.

Most widely used vision based method encounter the drawback, it requires a black background for recognition & it also required additional hardware.as the combination of surface & pen which is used in technique, where pen as input to make contact with the surface and due to that increase the hardware required and it requires a lot of time for setup and background noise can cause problems while detecting the input.

3. Suggestions & future work

One to one mapping is presented is traditional multi touch surfaces between the physical location that is being touched and the digital allocation on the display. This Provides user with feeling of directly manipulating the digital content underneath their fingers.

Here going to propose a virtual wall, by using calibration technique on that wall and with the help of Kinect to calibrate the surface due to that is going to be virtual. To facilitate a similar interactive experience in the depth camera-based approach, map the depthcamera scene to the projected display area as shown in figure.

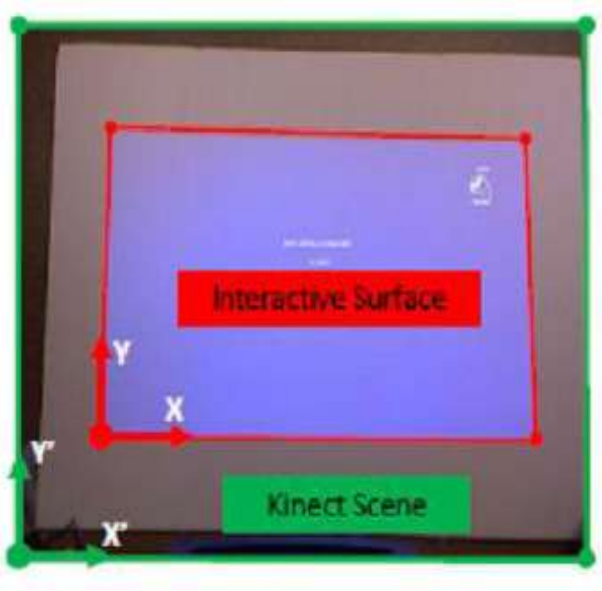

Fig 2.1.Callibration

Find the nearest neighbor between points in the frames and associate touch points in consecutive depth frames. Since the kinect can record at a maximum rate of thirty frames per second, not all depth frames are necessarily used in tracking touches. Our system intelligently discards outlier depth frames,Thereby increasing robustness.

The wall or surface will become computer vision and on that display area recovering gestures like fingers, wrist and hand postures as well as handiness (i.e. whether the user is using the left or right hand). Identifying anddistinguishing between different users who are interacting with surface.

As the computer vision provides a classroom education to motivate children with special needs to improving both motor and cognitive skills (beyond physical interaction). Classroom education includes data about students and teachers with respect to different class. Each class having their respective subject syllabus which is going to touch on that virtual wall as board for students with PPT's, Videos etc.We identify the working with body is important part as well as the key and specific competences based on the context of the school. As Microsoft Kinect is used this naturally works these competences by interacting with body. Proposed system having user's profile, once the system detects who is the player, according to user several interaction techniques able to begin for children's and it helps to improve skills of children.

As we talk about how to build an application that interact with human body motion, first of all we need to capture the information about the users standing in front of Kinect and from then on the skeleton tracking comes into the picture.

Activities tend to use the entire screen to place icons or simply to work, players introduce students and activities related to classroom.

Our objective was twofold. On the one hand, we wanted to provoke interest and motivation to children in working educational aspects, including the body. On the other hand, we wanted that almost all children would be able to interact with activities regardless their motor skills.

The user's study gave some insights that the interaction techniques are numerous enough and their implementation with the Kinect can be a powerful tool for improving specific skills & motivation.

References

[1] Kinect-Based Face Recognition and Its Application in MOOCs Production" Yuran Hu E-Learning Lab,ShanghaiJiao Tong University,Shanghai,China.IEEE 2015.

[2] "Virtual Touch Screen Using Microsoft Kinect Ekta Parwani1, Apeksha Pawar2, Chirag Ajwani3", Prof. Govind Pole4 Dept of Computer Engineering Modern Education Society's College of Engineering Pune-411001

[3] Human Gesture Recognition Using Kinect Camera" OrasaPatsadu, ChakaridaNukoolkit and BunthitWatanapa School of Information Technology King Mongkut's University of Technology Thonburi Bangkok, Thailand,JCSSE(2014).

[4] "Computer Vision for RGB-D Sensors:Kinect and Its Applications" IEEE TRANSACTIONS ON CYBERNETICS, VOL. 43, NO. 5, OCTOBER 2013

[5] Microsoft Kinect Sensor and Its Effect".Wenjumzeng,Universityof Missouri,IEEE(2014).
[6] Extended Multitouch: Recovering Touch Posture, Handedness, and User Identity using a Depth Camera" SundarMurugappana, Vinayaka, NiklasElmqvistb, KarthikRamania a) School of Mechanical Engineering b) School of Electric Engineering Purdue University, West Lafayette, IN 47907
[7] Gao, Q., & Sun, Q. (2015)"Examining the usability of touch screen gestures for older and younger adults" Human Factors: The Journal of the Human Factors and Ergonomics Society.
[8] Exploring and modeling unimanual object manipulation on multi-touch displays" JianZhaoR.WilliamSoukoreff(2015),Universityof Toronto,10King's CollegeRoad,Toronto,ON,Canada,M5S3G4
[9] Kinect for Windows SDK Programming Guide"(PDF).
[10] Hand Detecting and Positioning Based on Depth Image of Kinect Sensor" Van Bang Le, AnhTu Nguyen, and Yu Zhu International Journal of Information and Electronics Engineering, Vol. 4, No. 3, May 2014.