

Depth Sensing Cameras – A Comparative Study

^{#1}Samruddhi Naik

¹samruddhi.naik17@gmail.com

^{#1}Department of Information Technology, Savitribai Phule Pune University,
Pune, India



ABSTRACT

Hand gesture recognition has become widely implemented as one of the features offered to customers for computer control. Software tools that come with these devices have predefined tracking patterns for specific operating system. Specific tools that offer customization are generally available for development. Kinect by Microsoft is one of the most popular devices available on market among developers. Kinect for Windows is a depth sensing and RGB camera – specifically it consists of depth camera, RGB camera, IR sensor, tilting motor and four microphones. Another successful product on the market that enables depth recognition is called Asus Xtion PRO LIVE, depth-sensing and RGB camera. Designers of device called Leap Motion took another approach to depth sensing. Its principal use is aimed towards hand tracking and gesture recognition. This paper provides insights for all the three devices used for gesture recognition with a focus on leap motion controller.

Keywords— Gesture Recognition, Indian Sign Language, Camera, Kinect, Leap Motion Controller, ASUS XTION PRO, Comparative Study.

ARTICLE INFO

Article History

Received: :4th January 2016

Received in revised form :

6th January 2016

Accepted: 7th January, 2016

Published online :

9th January, 2016

I. INTRODUCTION

This paper presents the early findings of an exploration in to the suitability of a range sensing camera for recognising Indian Sign Language. A camera is an optical instrument for recording images, which may be stored locally, transmitted to another location, or both. The images may be individual still photographs or sequences of images constituting videos or movies. The functioning of the camera is very similar to the functioning of the human eye. The camera is a device which is easily available and has an adequate level of accuracy in producing a system which could be used to recognise Indian signs. This functionality could then be incorporated into a system to help young deaf and hard of hearing children to learn Indian signs. The system would be able to demonstrate specific signs using videos and images, and provide feedback to the child on its own about their Indian Sign accuracy through the application developed. This project is aimed specifically for Indian Sign Language and the principles will be relevant to any sign based communication system. Gesture recognition is concerned with identifying human gestures using technology. This is an established research area with a

broad background and many gesture recognition systems have been developed. The Seek and Sign research project is specifically interested in gesture recognition technologies that may be suitable for recognising sign language. Research has been conducted in this area, with the most promising technology to date being glove technology, wrist sensors, 2D and 3D cameras, and the Kinect platform.

Objectives of Work.

- 1) To suggest and build a new gesture based system by using a camera.
- 2) To develop a Sign language interpretation software using Image Processing and Artificial Intelligence.
- 3) To provide real-time translation of sign language through computer processing.
- 4) To suggest and build a simple application for interpreting input through camera that is more suited to Indian Sign language than the pre-existing methods.
- 5) To contemplate the use of an innovative input method for computers that is different from the

conventional input devices such as keyboards, mice and touch screens.

II. LITERATURE SURVEY

Depth sensors are used in Human gesture recognition is a popular new way to input information in gaming, consumer and mobile devices, including smart phones and tablets. A majority of industrial applications for 3D vision, including industrial and manufacturing sensors, integrate an imaging system from as few as 1 pixel to several million pixels. Advertisements already bombard us on a daily basis, but with interactive digital signage, companies will be increasingly able to use pinpoint marketing tools to deliver the most applicable content. The medical field also benefits from the new and unprecedented applications that 3D vision offers. This technology will ensure that the best medical care is available to everyone, no matter where.

Sr. No.	Reference Name (IEEE/ACM/Springer/Any other journal)	Work description	Problems found	P u b l i c a t i o n y e a r
1	Approach to Hand Tracking and Gesture Recognition Based on Depth-Sensing Cameras and EMG Monitoring	Contemporary technology allows us to transfer almost any component of the real world into the digital form.	Accuracy of leap motion is greater than Microsoft Kinect and Asus.	J a n 2 0 1 4
2	On Depth Sensors Applications for Computational Photography and Vision.	There is an increasing interest in interaction devices that go beyond the keyboard and mouse, the so called Natural User Interaction.	Finally, the Kinect's built-in RGB camera has low image quality compared to DSLR cameras.	2 0 1 3

III. DEPTH MEASUREMENT TECHNIQUES

a. Structured light

Structured light based sensors share the active triangulation approach above mentioned. However, instead of scanning the surface, they project bi-dimensional patterns of non-coherent light, and elaborate them to obtain the range information for each viewed point simultaneously.

b. Laser Scan

One of the most significant advantages of laser triangulators is their accuracy, and their relative insensitivity to illumination conditions and surface texture effects. Single-point laser triangulators are widely used in the industrial field, for the measurement of distances, diameters, thicknesses, as well as in surface quality control applications.

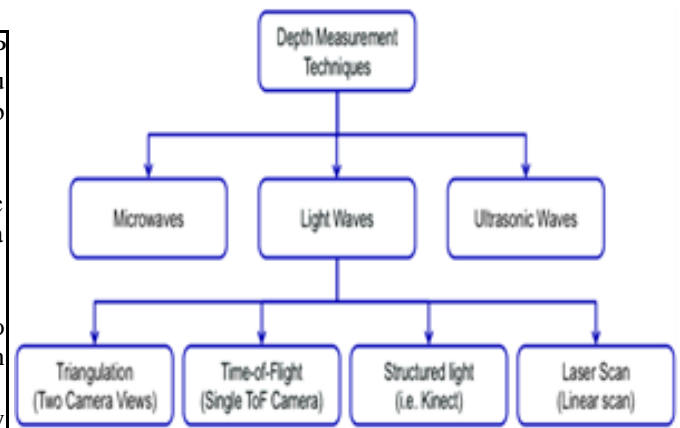


Fig 1. Techniques for Depth Measurement

HAND TRACKING AND GESTURE RECOGNITION

Gestures as such arise from a person's mental concept and are expressed through the motion of arms and hands. These expressions are then observed and recognized by the spectator (Pavlovic, Sharma, Huang, 1997). In our case the spectator is computational device that is able to recognize gestures through specific pre-learned models. Two conventional approaches towards gesture recognition are free-hand tracking and glove-based tracking. Both of these approaches have in common the need for specific mathematical model. First approach, free-hand tracking, enables control without any mechanical devices directly attached to user; the only input here is based on image processing. Second approach, glove-based tracking, uses mechanical device fixed on hand with sensors that track the movement. As for the free-hand tracking Pavlovic, Sharma & Huang (1997) further define two main models. First, 3D hand-based models are based on 3D description of hand. Further divided into volumetric model, where hand/arm postures are analysed by synthesizing of 3D model of human hand while its parameters are modulated

until compliance with real human hand is found, and skeletal model - comparable to volumetric model but this model deals with reduced set of equivalent joint angle parameters together with segment lengths. 3D-based model was used for sign language recognition by Mehrez & Jemni (2012) with recognition rate of 98.5 using 900 ASL. Second, appearance-based models, are not directly derived from 3D description of hand but are rather based on display of hands in sequence of images. Great numbers of such models use parameters that are derived from images in the templates such as contours and edges, image moments or image eigenvectors.

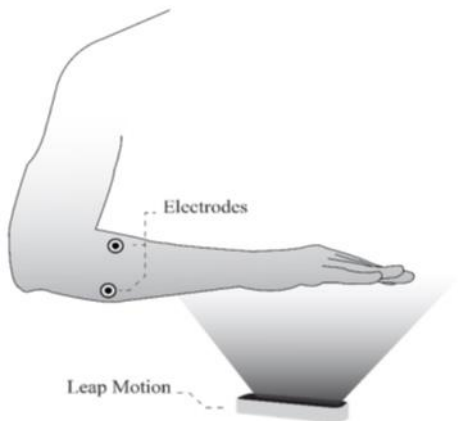


Fig. 2 Hand Positioning for gesture capturing

KINECT

Kinect (codenamed in development as Project Natal) is a line of motion sensing input devices by Microsoft for Xbox 360 and Xbox One video game consoles and Windows PCs. Based around a webcam-style add-on peripheral, it enables users to control and interact with their console/computer without the need for a game controller, through a natural user interface using gestures and spoken commands. The first-generation Kinect was first introduced in November 2010 in an attempt to broaden

Xbox 360's audience beyond its typical gamer base. A version for Windows was released on February 1, 2012. Kinect competes with several motion controllers on other home consoles, such as Wii Remote Plus for Wii and WiiU, PlayStation Move/PlayStation Eye for PlayStation 3, and PlayStation Camera for PlayStation 4.



Fig 3. Microsoft Kinect

Kinect builds on software technology developed internally by Rare, a subsidiary of Microsoft Game Studios owned by Microsoft, and on range camera technology by Israeli developer Prime Sense, which developed a system that can interpret specific gestures, making completely hands-free control of electronic devices possible by using an infrared projector and camera and a special microchip to track the movement of objects and individuals in three dimensions. This 3D scanner system called Light Coding employs a variant of image-based 3D reconstruction.

Kinect sensor is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise above or below the video display. The device features an "RGB camera, depth sensor and microphone running proprietary software", which provide full-body 3D motion capture, facial recognition and recognition capabilities. The depth sensor consists of an infrared laser projector combined with a monochrome CMOS sensor, which captures video data in 3D under any ambient light conditions. The sensing range of the depth sensor is adjustable, and Kinect software is capable of automatically calibrating the sensor based on game play and the player's physical environment, accommodating for the presence of furniture or other obstacles.

ASUS XTION PRO



Fig 4. ASUS Xtion Pro

Xtion PRO uses Infrared sensor and adaptive depth detection technology to capture users' real-time body movements, making body tracking more precise. The Xtion PRO development solution comes with a set of developer tools to make it easier for developers to create their own gesture-based applications without the need to write complex programming algorithms.

1. Gesture detection: The Xtion PRO development solution tracks people's hand motions without any delay, which turns your hand into a controller.

2. Whole body detection: The Xtion PRO development solution allows developers to track a users' whole body movement, which makes it ideal for whole body gaming, while also supporting multiple player recognition.

The Xtion PRO development solution allows developers to apply the latest motion-sensing technology in various

applications and industries to stand out from the competition. For example; Electronic Kanban is an ideal application for this technology and can be used widely for marketing, guidance or tourism and so on. Physical rehabilitation is also a good field for body detection applications. There are a number of fields that could benefit from motion-based technology, such as education, medical, conferences, games and many more.

The Xtion PRO development kit is widely open. You can create your own apps for business or for making people's lives more convenient and intuitive. Anyone can develop their own apps. A far-reaching application can be created and applied based on the developers' imaginations

LEAP MOTION CONTROLLER

The Leap Motion controller is a sensor device that aims to translate hand movements into computer commands. The controller itself is an eight by three centimetre unit that plugs into the USB on a computer. Placed face up on a surface, the controller senses the area above it and is sensitive to a range of approximately one metre. To date it has been used primarily in conjunction with apps developed specifically for the controller these apps consist of games, scientific and educational apps, and apps for art and music. While the potential for the technology is great, some early criticisms have emerged in product reviews in relation to app control, motion sensitivity, and arm fatigue. One factor contributing to the control issues is a lack of prescribed gestures, or set meanings for different motion controls when using the device (Metz 2013). This means that different motion controls will be used in different apps for the same action, such as selecting an item on the screen. Leap Motion are aware of some of the interaction issues with their controller, and are planning solutions. This includes the development of standardised motions for specific actions, and an improved skeletal model of the hand and fingers.

The Leap Motion controller is a small USB peripheral device which is designed to be placed on a physical desktop, facing upward. Using two monochromatic IR cameras and three infrared LEDs, the device observes a roughly hemispherical area, to a distance of about 1 meter (3.28084 feet). The LEDs generate a 3D pattern of dots of IR light and the cameras generate almost 300 frames per second of reflected data, which is then sent through a USB cable to the host computer, where it is analyzed by the Leap Motion controller software using "complex math" in a way that has not been disclosed by the company, in some way synthesizing 3D position data by comparing the 2D frames generated by the two cameras.



Fig 5. Leap Motion Controller

IV.EVALUATION AND COMPARISON

Device	Pros	Cons
Micro soft Kinect	<ul style="list-style-type: none"> High quality of device drivers Stable work with various hardware models Has motor that can be controlled remotely by iPi Recorder application this makes device positioning more convenient 	<ul style="list-style-type: none"> Bigger size (12" x 3" x 2.5" against 7" x 2" x 1.5") Higher interference with another Kinect sensor in "Dual depth sensor" configuration Lower RGB image quality in comparison with Kinect
ASUS Xtion PRO	<ul style="list-style-type: none"> More compact (7" x 2" x 1.5" against 12" x 3" x 2.5") Lighter weight (0.5 lb against 3.0 lb) Lower interference with another ASUS Xtion / Prime Sense Carmine sensor in "Dual depth sensor" configuration 	<ul style="list-style-type: none"> Less popular device Lower drivers quality Does not work with some USB controllers (especially USB 3.0)
Leap Motion	<ul style="list-style-type: none"> Inexpensive Extremely precise Easy to set up 	<ul style="list-style-type: none"> Range of detection is limited Limited to practical controller.

Table 2. Pros and Cons

DEPTH SENSING IN LEAP MOTION

1) EMG

A term EMG is known since the 1943 when introduced by Weddell et al. for the description of clinical application – in this process the needle was used for the examination of skeletal muscles (Katirji, 2007). Phenomena of electromechanical coupling in muscle is utilized, more specifically deep imaginations of muscle membrane inside the muscle cells enable conduction of electrical impulses in skeletal and cardiac cells, these are called t-tubules. When muscle contraction occurs it can be then graphically depicted as electromyogram and recording is realized using electromyography (Kumar, Mital, 1996). EMG enables diagnosis of some peripheral nervous system disorders. Traditional approach of EMG includes the usage of needle inserted into the muscle to enable recording. Less invasive or completely non-invasive approach is called surface electromyography (SEMG). Fig. 2 represents the difference between intramuscular and surface EMG. Sensor may be attached to specific muscles to monitor and provide capability for differentiation of the activity of the muscle. SEMG signal is non-linear and short-time stationary.

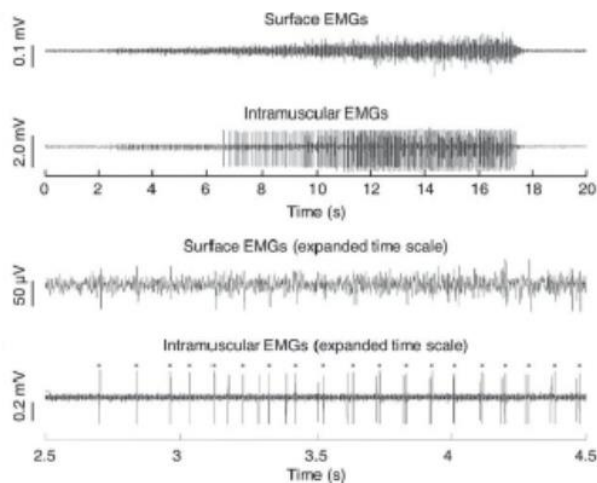


Fig 6. EMG Signal

2) VISION-SEMG-BASED TECHNOLOGY.

In this approach, we take combination of SEMG and depth-camera. Visual monitoring of hands realized, as proposed, by Leap Motion sensor device that is directly connected to the computer. Main drawback when using this device is requirement for a user to remain near the interaction zone, i.e. near computer. Hand and fingers position or general gestures were using functions available by Leap Motion API. Four disposable electrodes were utilized in SEMG monitoring, with Arduino microcontroller used to record the signals. Arduino enabled us to utilize its capabilities – we designed shield that was directly attached to microcontroller, this shield serves as an amplifier of EMG signals and connects Arduino board to electrodes via 3.5 mm jack. Electrodes were directly attached to a skin; conductive gel was used

as to enhance connection. All the processing of signal is accomplished on data processing machine, i.e. personal computer.



Fig 7. Vision based Processing

V. CONCLUSION

Thus study represents a new hand tracking approach for gesture recognition. This approach is indented for human computer interaction using SEMG technique in combination with visual-based tracking system with depth recognition. Monitoring is achieved through the hand muscles EMG signals tracking and visual image detection of hand movement. Leap Motion device proved ability to effectively recognize gestures, however issues with continuous finger tracking and certain hand positions are observed. This system is supposed to overcome shortcomings of solely visual-based system where tracking of hand is not always recognized due obstacles – e.g. hand covered by hand. For hand tracking with individual fingers the Leap sensor is a more suitable option than the Kinect and for close range tracking Intel Creative camera performs better than Kinect, but when full body motion is required, Kinect or Asus Xtion is appropriate. The Xtion Pro camera is faster than the Kinect and has a higher resolution.

VI. ACKNOWLEDGEMENT

It is my privilege to acknowledge with deep sense of gratitude to our project guide and Head of Department (Information Technology), Dr. N J Uke for his valuable

suggestions and guidance throughout my course of study and timely help given to me for my project. I would also like to thank my parents, group members and friends for providing suggestions, help and moral support.

REFERENCES

[1].Jihyun Han, Nicolas Gold “Lessons Learned in Exploring the Leap Motion ® Sensor for Gesture-based Instrument Design”, Proceedings of the International Conference on New Interfaces for Musical Expression, Goldsmiths University of London: London, UK, http://nime2014.org/proceedings/papers/485_paper.pdf, ISBN-13: 9781906897291, 2014-06-30 - 2014-07-04.

[2].Changa, Y.-J., Chenb, S.-F., Huang, J.-D. (2011). A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities*, 32(6), 2566–2570. doi: j.ridd.2011.07.002

[3].Dominio, F., Donadeo, M., Zanuttigh, P. (2013). Combining multiple depth-based descriptors for hand gesture recognition. *Pattern Recognition Letters*, in press, doi: j.patrec.2013.10.010 Gaoxiang, O., Xiangyang, Z., Zhaojie, J., Honghai, L. (2014). Dynamical Characteristics of Surface EMG

[4].Signals of Hand Grasps via Recurrence Plot. *IEEE Journal of Biomedical and Health Informatics*, 18(1), 257–265. doi: 10.1109/JBHI.2013.2261311. Garcia M.A. C., Vieira T.M. M. (2011) Surface electromyography: Why, when and how to use it. *Revista Andaluza de Medicina del Deporte*, 10(6), 17-28

[5].González-Ortega, D., Díaz-Pernas, F.J., Martínez-Zarzuela, M., Antón-Rodríguez, M. (2014). A Kinect-based system for cognitive rehabilitation exercises monitoring. *Computer Methods and Programs in Biomedicine*, 113(2), 620–631. doi: j.cmpb.2013.10.014