Hand Gesture Recognition Using 3-D Depth Perception

Parth Sharma, Saru Dhir, Madhurima

Abstract: The paper provides information on hand gesture recognition through its various methods but primarily focused on 3D depth-perceiving sensor technology for hand tracking and gesture recognition. Hand gesture recognition provides new opportunities for human-computer interactions (HCI) and over the past decade has emerged in various fields and technologies. The paper lures an likeness amongst the present technologies in the market and the workings behind their operations. It differentiates between the old and newer mechanics employed in the products. Furthermore, the paper discusses the existing uses of gesture recognition with their limitations. It provides proposed ideas in various fields, their use and future developmental paths.

Keywords: Human-computer Interaction, Depth Data, Old and Newer Methods of Gesture Recognition.

I. INTRODUCTION

There is increasing emphasis recently on Human-Computer Interaction (HCI) with the aim to provide a more robust and international mode of interface. As a result, the focus on the gesture-based interface has enhanced and works to provide a touch-free experience for the user to interact with the computer, replacing the conventional mouse. This is where sensors come into play, to allow a computer to capture robust movements using motion-capture, vision, depth sensors, etc. This mode of interface has extensive application in the fields of robotics, design, gaming, security and virtual reality.

The computer software is used to capture a sequence of depth maps over a period of time. The depth data, allows for precise and fluid tracking of the parts of the body and small objects in free space within the peripheral's range, with sub-millimeter accuracy. It detects and analyses the motion of the body part in order to decode and classify the user's movements as the user interacts with the body display. These depth maps are analyzed and then compared with the predefined group of gestures. Depth maps are obtained from the data generated by the peripheral.
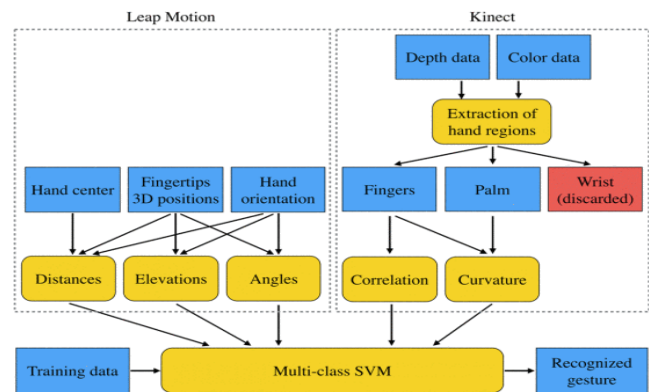
Fig.1 SVM approach [1]

There exist several hand gesture recognition systems and focus on the various palm and finger movements, positioning, angles, elevation, etc. to decode the subtle actions. Fig. 1 shows how these movements are picked up and categorized through the use of Support Vector Machines (SVM).

Finger-Earth Mover's Distance (FEMD) is another system where it conceives every finger as a assemblage and penalizes unmatched fingers [2]. It can distinguish the hand gestures of minute differences better as the aforementioned only matches fingers, not the whole hand.

X-Y plane is taken as the display screen with the depth as the perpendicular to the screen i.e. Z-axis [3]. Depth map includes all the object in the field of view, including the body parts. The depth map is also able to analyze the speed of the movement with the computer's internal clock or timestamp for each frame of the depth data, together with measuring the distance. By this X-Y plane method, the computer can capture small movements, unlike conventional motion capture devices. Motion-capture devices like Microsoft Kinect and Nintendo Wiimote mostly capture larges movements and are unable to detect small movements. [3]

The use of depth maps in the case of the leap motion device provides robust hand gesture recognition, this allows the device to capture small movements with sub-millimeter accuracy and allow for a range of control using these movements (e.g., hand and finger movements). This gesture-based interaction provides new opportunities in certain application areas (e.g., entertainment, robotics, health, engineering, etc). Through the use of depth-sensing sensors, the computer can even recognize the gestures of multiple users at a time.

II. LITERATURE REVIEW

Hand gesture recognition is a field that is getting additional attention over the years and has been continuously growing. Consequently, many different types of approaches have emerged.

Among the various Human-Computer Interactions (HCI), hand gesture based HCI is the most natural and user-friendly. It is also one of the prominent ways of expression and is used in day to day interaction. This makes hand gestured based HCI one of the best alternatives from the present user interface. It aims to provide a completely hands-free control, by building a robust hand gesture recognition system. Its natural and user-friendly aspects have widespread applications in robotics, virtual reality, health care, graphic designing, etc [4].

Support Vector Machines (SVM) were proposed by Giulio Marin, Fabio Dominio, and Pietro Zanuttigh as an approach to the gesture-based interface. It was incorporated in detecting both large and small motion when interacting with a computer. By the depth data, a silhouette is made using volumetric shape descriptors of the involved body part based on (SVM). This compares the data in the form of histograms of the distance of the hand center to the hand edge points to identify the gesture [1].

Another system was proposed, called Finger-Earth Mover's Distance (FEMD). Where it conceives every finger as a assemblage and penalizes unmatched fingers. The palm is treated as the hand center and recognizing the various movements according to the distance of these fingers from the hand center. These movements generated curves, which on analyzing gave the output [2].

A tool for fine finger tracking called Leap Motion was released in 2013, it allowed high-resolution tracking of the slight and specific finger and hand gestures along with more specialized movements. MRLeap, an external USB peripheral, allows a customizable interface between the Leap Motion and MAX/MSP program, allowing the users to make full use of the Leap Motion controller [5].
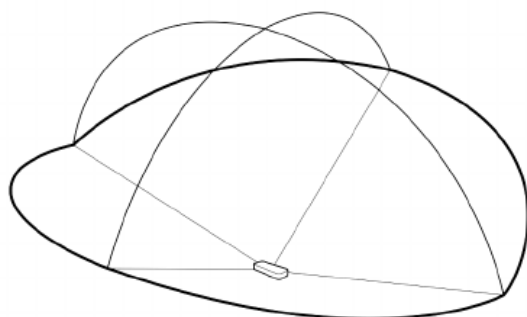

Interaction Area
2 feet above the controller, by 2 feet wide on each side
(150° angle), by 2 feet deep on each side (120° angle)

Fig. 2 Field of View of Leap Motion Device [5]

A. Developed Technology and Past Works

Hand gestures are a natural way of expressing one's thoughts and actions as such it is also an ideal way of interaction thus its applications and use is widespread. The following are the existing works of hand gesture technology in their various fields and its applications.

Gesture recognition for sign language is an example of the use of such a technology. It allows communication with the hard of hearing and deaf through the use of sign language. This involves many semantic aspects such as hand shape, hand location, palm orientation, and facial expression. With the help of Microsoft Kinect, the color and depth images of the signer are captured. Through analyzing this depth data and using a 3-D shape context to determine the hand shape, by comparing it to the database, the sign can be identified.

Gesture recognition for comprehending sign languages is still a compound problem due to the large number and variety of features defining each sign. This problem is then magnified when computing a continuous image stream, also known as Continuous Sign Language Recognition or Motion Epenthesis [6].

Using Hidden Markov Models (HMMs), hand gestures are useful to interact and give commands to robots. The hand gesture are captured over numerous frames, then classified and the recognized gestures are converted into control directives to operate the robot. Experimental data shows that the robust hand gesture interface can work efficiently in complex environments and real-time with an mean recognition rate of 98.4% [7]. This hands-free mode of interaction between human and robots enables everyone to interact with robots, even the elderly and disabled people.

Human-robot interaction is still immature with many issues yet to be addressed such as, the hand gesture detection methods based on skin complexion detection, 2D or 3D hand gesture detection isn't dynamic enough in complex environments and template recognition performance is usually heavily influenced by luminosity, varying color, and littered background [7].

B. Implementation

Over the past years, hand gesture technology has been introduced to the public and technological world. It ranges from various uses in different fields. The following implementations related to hand gesture recognition has been introduced in the market:

The most widespread use of hand gesture technology is seen in the gaming world. The various platforms such as PlayStation Move [8], Nintendo Wii [9], Microsoft Kinect [10], etc. have been using gesture-based technology for nearly the past decade.

The earliest most notable form of gesture recognition was in the form of Nintendo's Wii. Nintendo Wii provided new interaction methods on the path of HCI. Its aim was to recognize gestures to communicate with applications and present the design and

assessment for sensor-based gesture recognition. The Wiimote was used as an input device and had an autonomous acceleration sensor for gesture recognition. This allowed the training of capricious gestures by users which can then be recollected for interacting with systems. The developed library uses the Wii sensor data and utilizes a Hidden Markov Model (HMM) for training and identifying user-chosen gestures [11].

Another staggering implementation was in the form of Microsoft's Kinect. This device employed the use of depth cameras, through the camera a depth image of an object in the plane towards the camera was generated. This was then used to generate a depth profile of the subject after the removal of the background from the depth image. These profiles were taken across multiple frames and their differences provided the motion profile of the subject and were used for gesture recognition. The system was trained using a multiclass Support Vector Machine and allowed the successful recognition of multiple human gestures [12].

One of the latest introductions to gaming technology with the incorporation of hand gesture recognition is virtual reality (VR) [13]. Instead of using camera-based identification systems, which are expensive and very delicate to environmental changes another approach to this implementation is with the use of data gloves as an input device. It provides 18 measurement results for the angles of various finger joints. This glove performs alongside a neural network model, such as back-propagation and radial-basis functions, which are used by the identification system to recognize the gestures. Some of the network models attain a recognition rate of up to 100% over various numbers of test subjects. Because of its suitable operation, this recognition system is the foremost footstep towards virtual reality applications [13].

Another approach involves designing in Virtual Reality using gesture-based input. DesIRe (DESigning In virtual REality) is an interface where virtual reality is treated as a stage of virtual clay and enables real-time dynamic interaction with an object to perform design operations. In DesIRe, a hand-shaped cursor is controlled by the user by the help of a single infrared LED, connected to the data glove [14].

III. PROPOSED WORK

The following are the proposed inputs for hand gesture recognition technology in various fields:

1) The existing hand gesture technology can be further improved and refined with the addition of eye tracking. This can be used to pinpoint the location where the gesture is to be executed and thus improve the overall accuracy. For example, when zooming in or out, with the use of index finger and thumb moving towards or away from each other, the area to be affected on the

computer's screen can be accurately controlled. This would increase the accuracy as 3-D depth perception technology is able to detect the gesture and position of the hand it is unable to target precise and small areas due to the relatively large size of the hand when compared to the computer's screen. Thus using software to detect where a person is focusing at through capturing the movements of the eyes this barrier can be overcome and boost the accuracy of the product.
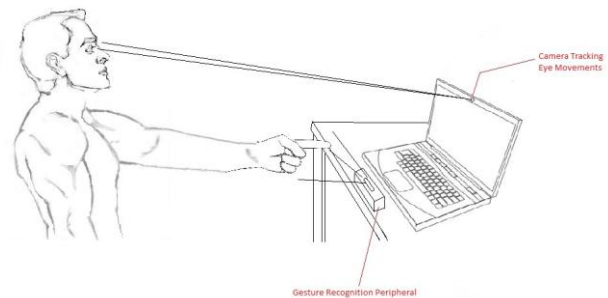

Fig. 3 Integration of Eye Tracking System With Hand Gesture Recognition

2) The hand gesture technologies use different and unique approaches to collect depth data and other information to analyze, decode and recognize a gesture. However there is often a case that not every type of a particular gesture can be directly added to the software for recognition and due to the everyone not necessarily making the same gesture in the same way it can lead to rigidity, unresponsiveness and hang or delay in response. By applying Machine learning in the process of detecting and recognition of gestures can lead to a solution. Through this technology the computer and learn relatively fast to recognize small alternations in the gestures and overtime learn bigger changes. For example, when zooming in or out, as mentioned above, some users tend to also move nearby fingers not involved with the gesture. This might also be recognized as an opening palm gesture and lead to an entirely different undesired result. Through machine learning, this problem can be resolved and lead to a better refined and structured database of gestures. These can be included on newer models that have yet to come into the market and be regularly updated by getting feedback and analyzing the problems users commonly face when using a gesture through user input and suggestions and looking through the data and the correct gesture that the user wants to make when this issue occurs. This can be accompanied by OTA [15] (over the air) updates to the gesture recognition devices.

3) Hand gesture recognition can be extensively applied in the field of Robotics. It can be used in a manner that fundamentally changes the way users interact with robots. Hand gestures can be used to control the movements of different parts as well as the

overall directional movements. Certain fixed gestures or series of gestures can be programmed to execute specific commands, this also helps in providing gestures for a detailed task multiple times through the steps, leading to security, anonymity, and efficiency.

4) Hand gesture recognition has been widely used in gaming from the emergence of the Nintendo Wii in 2006 to current virtual gaming technology [16]. The use and implementations of small motion capturing gesture recognition technology can greatly enhance the face of gaming and can also be used extensively. It would lead to more immersive, lively and interactive gameplay. Hand gesture technology can be developed into a depth based gaming method with full 360° movement paired with virtual reality and hands-free controls. This would elevate gaming to a new platform and improve the experience.

5) Hand gesture recognition technology can be applied in the field of security, by adding a program which, through the help of cameras, detects and learns the hand gestures one naturally makes over time thus enabling the camera to recognize the person. This can be further developed to accommodate body language and the way one moves when walking to enhance the security system and be able to identify those that have their movements analyzed and learned. This would lead to the great efficiency and proper monitoring for in areas where security is of utmost importance and can be used to detect hostile or abnormal movements as a sign of caution and reduce security mishaps and breaches.

6) Hand gesture technology can also be improved upon by the introduction of AI. Through the use of AI a model of human forelimbs can be developed with accurate joints and range of motion. With the use of this model and its integration with AI, the various gestures that a human can execute and the various motions can be categorized and saved. Leading to a robust database of gestures and motions that can be used to control elements and be customized to execute certain specific commands. This would eliminate the need for manually imputing commands one by one to grow the database for the actions that a notion recognition system would grasp and execute accordingly. These then can be provided to the masses through OTA [15] updates leading to a more in-depth use of the technology.

7) Hand gestures can be used to operate heavy machinery. Through pre-programmed actions, heavy machinery can execute a certain movement. This can be used in lifting, operation, and maintenance of various machinery. This would lead to a reduction in work as well as human injuries and mishaps at the workplace.

IV. RESULTS

(Results co-respond to the above points in proposed work)

1) Eye tracking along with hand movements can lead to an increase in accuracy for HCI. This would further enable the seem-less interface between the computer and the user. Through this, we can reduce the false interpretation of inputs and provide a more user-friendly experience.

2) With the help of machine learning at the stage of gesture input and recognition, we can recognize small alternations in the gestures and overtime learn bigger changes. When making various gestures not everyone may follow the exact same way and movements pre-designed to be recognized. Machine Learning overtime can eliminate this issue and accompany the variations in gesture but still provide the desired results. This would make the process more natural and efficient. Through user feedback and inputs this can also be distributed to others through OTA [15] updates.

3) In the field of gaming, hand gesture recognition technology has widespread and essential use. Through the incorporation of depth based gesture recognition technology, which is able to read and understand small movements the level of immersive gameplay can be elevated. When paired with VR and complete 360° recognition it would lead to truly hands-free, immersive and interactive gameplay and experience.

4) When effectively paired with robotics gesture recognition can do a variety of tasks from mechanical movement to performing a pre-programmed task with a specific set of gestures. By pre-programmed gestures, the time taken for the tasks as well as input after every step can be eliminated and it would lead to anonymity and security as only those who know the programmed task would be able to identify the robot's actions.

5) When paired with security devices hand recognition would be able to identify known individuals, whose data has been captured and keep tabs on them. This can be further enhanced to another level by adding in body language recognition. This would allow for a great degree of security as those who act against the interest of the user would be easily and efficiently identified. This would lead to a decrease in illegal activities, a breach of security and increase safety. This technology can also be used to analyze a person acting suspiciously and highlight them before acting by reading their body language.

6) Through merging AI with hand gesture technology a modal of human forelimbs can be made with restriction in the range of movements as humanly possible. This

can then be programmed to make various movements and gestures within the limited range and create a database. This database can then be used in gesture recognition without the need of manually adding in gestures one by one and also increase the number of gestures one can use for interaction. This can also be easily distributed to all with OTA [15] updates regularly.

7) Hand gesture technology can be used to maintain and operate heavy machinery that is designed to execute certain movements on the commands thus increasing use and efficiency and reducing the time take any injuries or mishaps in the workplace.

8) Hand gestures can be used to operate heavy machinery. Through pre-programmed actions, heavy machinery can execute a certain movement. This can be used in lifting, operation, and maintenance of various machinery. This would lead to a reduction in work as well as human injuries and mishaps at the workplace.

V. CONCLUSION

This paper discusses the upcoming hand gesture recognition technology, its use along with its applications and history of development. It covers the various research done in the field and its neighboring topics, with emphasis on the hand gesture recognition techniques and their implementations. The paper also shows the collations between the early and present technologies in the market through the comparison of the Kinect based gesture recognition and the more recent leap motion gesture recognition. This paper also presents proposed implementations of gestured based recognition and its future developments in the fields of security, robotics, gaming, industrial use, and methods to further refine the technology.

REFERENCES

1. Giulio Marin, Fabio Dominio, and Pietro Zanuttigh, "Hand Gesture Recognition with Leap Motion and Kinect Devices", IEEE International Conference on Image Processing (ICIP), 1565-1569, 2014.
2. Zhou Ren, Junsong Yuan and Zhengyou Zhang, "Robust Hand Gesture Recognition Based on FingerEarth Mover's Distance with a Commodity Depth Camera", MM '11 Proceedings of the 19th ACM international conference on Multimedia, 1093-1096, 2011.
3. Eran Guendelman, Aviad Maizels, Tamir Berliner and Jonathan Pokrass, "Gesture-Based User Interface", US20120204133A1, 2012, Accessed on 28th May 2019, https://patents.google.com/patent/US20120204133A1/en.
4. Zhou Ren, Jingjing Meng and Junsong Yuan, "Depth Camera Based Hand Gesture Recognition and its Applications in Human-Computer-Interaction", IEEE: 8th International Conference on Information, Communications & Signal Processing, 1-5, 2011.
5. Martin Ritter and Alyssa Aska, "Leap Motion As Expressive Gestural Interface", Proceedings ICMC|SMC|2014, 14-20 September 2014, Athens, Greece,659-662, 2014.
6. Lalit K. Phadtare, Raja S. Kushalnagar and Nathan D. Cahill, "Detecting Hand-Palm Orientation and Hand Shapes for Sign Language Gesture Recognition Using 3D Images", IEEE: Western New York Image Processing Workshop, 29-32, 2012.
7. Dan Xu, Yen-Lun Chen, Chuan Lin, Xin Kong, and Xinyu Wu, "Real-time Dynamic Gesture Recognition System based on Depth Perception for Robot Navigation", IEEE: 2012 IEEE International Conference on Robotics and Biomimetics (ROBIO), 689-694, 2012.
8. M.P. Jacob Habgood, David Wilson, David Moore, and Sergio Alapont, "HCI Lessons From PlayStation VR", ACM: CHI PLAY '17 Extended Abstracts Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play, 125-135, 2017.
9. Torben Schou and Henry J. Gardner, "A Wii remote, a game engine, five sensor bars, and a virtual reality theatre", ACM: OZCHI '07 Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces, 231-234, 2007.
10. Florian Kammergruber, Andreas Ebner and Prof. Willibald A. Günthner, "Navigation in Virtual Reality Using Microsoft Kinect", 12th International Conference on Construction Application of Virtual Reality, 350-359, 2012.
11. Thomas Schlomer, Benjamin Poppinga, Niels Henze, Susanne Boll, "Gesture Recognition with a Wii Controller", ACM: TEI '08 Proceedings of the 2nd international conference on Tangible and embedded interaction, 11-14, 2008.
12. K. K. Biswas and Saurav Kumar Basu, "Gesture Recognition using Microsoft Kinect®", IEEE: The 5th International Conference on Automation, Robotics and Applications, 100-103, 2011.
13. John Weissmann and Ralf Salomon, "Gesture recognition for virtual reality applications using data gloves and neural networks", IEEE: IJCNN'99. International Joint Conference on Neural Networks. Proceedings (Cat. No.99CH36339), 2043-2046, 1999.
14. Manolya Kavakli, Meredith Taylor and Anatoly Trapeznikov, "Designing in virtual reality (DesIRe): a gesture-based interface", ACM: Proceedings of the 2Nd International Conference on Digital Interactive Media in Entertainment and Arts, DIMEA '07, 131–136, 2007.
15. Nils Aschenbruck, Jan Bauer, Jakob Bieling, Alexander Bothe and Matthias Schwamborn, "Selective and Secure Over-The-Air Programming for Wireless Sensor Networks", IEEE: 2012 21st International Conference on Computer Communications and Networks (ICCCN), 1-6, 2012.
16. Mark Bozon, "Wii: The Launch Games", 18 MAY 2006 3:59 PM PDT, https://www.ign.com/articles/2006/05/18/wii-the-launch-games.

AUTHORS PROFILE

Mr Parth Sharma is pursuing B.Tech in Information Technology from Amity School of Engineering & Technology, Amity University, Noida, Batch (2018-22). His area of interests are Machine Learning, Artificial Intelligence, Deep Learning, Human Computer Interaction, Cloud Computing and Gesture Recognition. He has attended various workshops and seminars at national level.

Dr. Madhurima Hooda is currently working as an Assistant Professor in the Discipline of Information Technology at Amity School of Engineering and Technology, Amity University, Uttar Pradesh, Noida, India. She has approximately 14 years of Teaching and research experience. She has published many research papers in national and international journals and conferences of repute. She has published a book titled "Computer Networks" with Laxmi publications and contributed as an editor in global edition of the book "Absolute Java", Pearson Education. Her area of research is video object tracking, software testing, IoT, AJAX Applications and Machine learning. Her M.Tech work has been published as a book titled "Video Object Tracking" by LAP LAMBERT Academic Publishing GmbH & Co. KG, Germany. She has currently three patents and two copyrights on her name. She is an editor of Journal of Information Processing Systems (Scopus indexed). She is also a member of several professional bodies IETE, IFERP, IOAP and many more.

Dr. Saru Dhir has done her Ph.D in CSE, M.Tech in CS and M.Sc. She has more than 14 years of experience in teaching and research. Her research areas are: Software Engineering, Agile Development, Software Testing, Cyber crime. She has published many research papers in national and international journals and conferences of repute. She has completed Microsoft technology associate certification in DBA, C#, Networking and visual basic.