JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Study On Hand Gesture Recognition and Various Technologies

[1]Zeenat Sheikh, [2]Ashwini Deshbhratar, [3]Shreya Tode, [4]Amrin Sheikh [5]Prof.Rupatai Lichode

[1,2,3,4]B. E Student, department of C-Tech, Rajiv Gandhi College of Engineering Research and Technology, Chandrapur, India.

[5]Assistant Professor, dept of C-Tech, Rajiv Gandhi College of Engineering, Research & Technology, Maharashtra, India.

Abstract- *Hand gesture recognition system has developed excessively in the recent years, reason being its ability to cooperate with machine successfully. Gestures are considered as the most natural way for communication among human and PCs in virtual framework. We often use hand gestures to convey something as it is non-verbal communication which is free of expression. In our system, we used background subtraction to extract hand region. In this application, our PC's camera records a live video, from which a preview is taken with the assistance of its functionalities or activities.*

Keywords: Gesture recognition, OpenCV, human-computer interaction, python, machine learning.

1. INTRODUCTION

Gesture recognition is a computing process that attempts to recognize and interpret human gestures through the use of mathematical algorithms. Gesture recognition is not limited to just human hand gestures, but rather can be used to recognize everything from head nods to different walking gaits.

Gesture recognition is a growing field of computer science, with an international conference devoted to gesture and facial recognition. As the field continues to grow, so will the ways that it can be utilized. Gesture recognition computer processes are designed to enhance human-computer interaction, and can occur in multiple ways, such as through the use of touch screens, a camera, or peripheral devices.

Touch screen gesture recognition has become second nature to many people today. While some computers and operating systems allow for customized gesture recognition, most people today know that they can pinch-to-zoom on a touch screen when they want to get a closer look at something. This specific gesture transcends nearly all user interfaces, from smart phones to personal computers. Touchscreens allow for relatively easy interaction between humans and computers.

Gesture recognition technology that is vision based uses a camera and motion sensor to track user movements and translate them in real time. Newer cameras and programs allow for tracking of depth data as well, which can help improve gesture tracking. Through the use of real-time image processing, users can interact with the program immediately to achieve the desired results.

For example, the Xbox Kinect relied on a camera to translate players movements as part of different games. There have also been experiments performed around using a camera to track an individual's gait and then utilizing deep learning algorithms in order to assess their chance of falling, and to make recommendations on how to lower those chances. There have been devices created, such as those by Leap Motion, that use specialized cameras and programs specifically around hand tracking in order to optimize the motion- tracking results. By focusing only on hand gesture recognition, such programs can get increased accuracy, allowing users to interact with their systems easily and completely hands-off. Through integrating such technology into existing touch screen devices and kiosks, such gesture-based technology can allow multiple users to interact with the same device without the fear of spreading germs.

There are multiple different peripheral devices that allow for different gesture interfaces. For example, most virtual reality or augmented reality systems have some kind of glove or controller that users must utilize to detect hand gestures and translate their movements into the movements of the character in the game. Gesture recognition can be used to improve a variety of fields, such as

Public health: By removing the need for touchscreens on self-service kiosks, business and organizations could help reduce the number of germs being spread. This is especially helpful for mitigating the spread of infectious diseases such a Covid-19 or influenza.

Health diagnostics: Through analysis of movements, doctors can help diagnose patients with diseases or fall risks in order to improve their overall outlook. On top of analysing gaits, gesture analysis and machine learning can be used to identify small movements such as tics and spasms to help-recognize-potential-diagnoses.

Security: Programs can be set up to recognize hand gestures and to send alerts in response. For example, home security systems can be set up to recognize what hands look like when they are holding guns and then send an alert in response. Or an organization can set up and train employees around a specific hand gesture, so that when cameras pick up the gesture, they can silently alert law enforcement, such as in the event of a robbery. Hand gesture frame acquisition is to capture the human hand gesture by the computer.

Whereas hand tracking is the ability of the computer to trace the hand and separate it from the background and from the surrounding objects. The features needed to be extracted changes from one application to another, some of the features that could be taken into consideration are: fingers status, thumb status, skin color, alignments of fingers, and the palm position. In artificial intelligence, machine learning enables the computers to learn without being explicitly programmed. There are two types of learning; the process when algorithms reflect the gestures that has been learned in the past in training to new gestures is called supervised machine learning, and the process when algorithms draw inferences from the gestures is called unsupervised machine learning. Classification aims of building a model to classify new hand gestures based on previous training gestures. The research work in hand gesture recognition has been developing for more than 38 years. In 1977, a system that detects the number of fingers bending using a hand glove was proposed by Zimmerman. Furthermore, Gary Grimes in 1983 developed a system for determining whether the thumb is touching another part of the hand or fingers. In 1990, despite the limited computer processing powers, the systems developed then gave promising results. The field of hand gesture recognition is very wide, and a big amount of work was conducted in the last 2 to 3 years. The features extracted are then sent to training and testing the classification algorithm (such as Artificial Neural Networks (ANN), K-nearest neighbour (KNN), Naive Bayes (NB), Support Vector Machine (SVM), etc.) to reach the output gesture.

2.TIMELINE OF VARIOUS TECHNOLOGIES

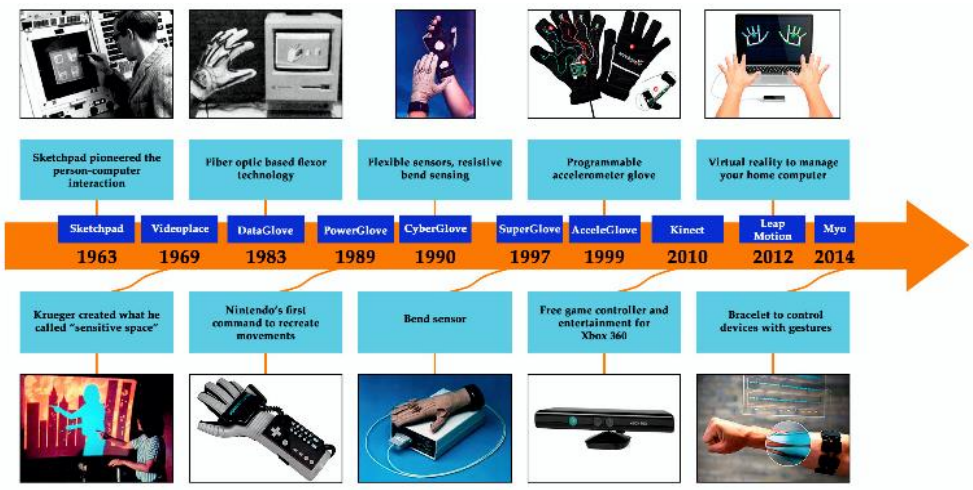

Fig. Timeline evolution of gesture recognition

3.LITERATURE SURVEY

Literature Survey on Glove Based Approach

In this approach we attach sensor to mechanical or optical gloves that convert inflection of fingers into electrical signals for hand posture determination and additional sensor for position of the hand. This approach is in utilization for hand gesture recognition method using magnetic field which is attached to the glove. The use of gestures among humans, both in the form of pantomime (ridiculous situation) or by using sign language, is closely linked to speech and represents an effective way of communication, used even prior to talking. The formality of the set of rules chosen in each case is related to the validity of the performed gestures, which means that a ridiculous situation gesture could be commending speech in an unplanned manner.

Table 1: Literature review on Glove Based Analysis

Authors	Year	Description
D. J. Sturman and D. Zeltzer (Sturman & Zeltzer, 1994)	1994	The authors proposed technologies such as position tracking, optical tracking, marker systems, silhouette analysis, magnetic tracking or acoustic tracking.
L. Dipietro and A. M. Sabatini and P. Dario (Dipietro, Sabatini, & Dario, 2008)	2008	The authors analyzes the characteristics of the devices, provides a road map of the evolution of the technology, and discusses limitations of current technology
Abhishek, K. S, Qubeley, L. C. Fai and Ho, Derek (Abhishek et al., 2016)	2016	The authors proposed a prototype that recognizes gestures for the numbers 0 to 9 and the 26 English alphabets, A to Z using capacitive touch sensor.

Literature Survey on Vision Based Approach

Vision based approach has the prospective to come up with natural and non-contact solutions and is built on the way humans explicate and interpret information about their surroundings. It is in all probability the most tough approach to execute (H. Hasan & Abdul-Kareem, 2014). A bare hand is used to extract data needed for recognition, and there is direct interaction between the user and with the system. For acquiring data needed for gesture analysis it uses some image characteristics like color and texture.

Table 2: Literature review on Vision Based Analysis

Authors	Year	Description
P. Garg, N. Aggarwal, and S. Sofa(Garg, Aggarwal, & Sofat, 2009)	2009	This paper is a review about Vision based Hand Gesture Recognition techniques for human computer interaction, combining the various available approaches, listing out their general advantages and disadvantages.
G. Murthy and R. Jadon,(Murthy & Jadon, 2009)	2009	The authors introduced the field of gesture recognition as a mechanism for interaction with computers.
M. K. Ahuja and A. Singh(Ahuja & Singh, 2015)	2015	The authors proposed a scheme using a database-driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching using PCA.

Literature Survey on Colored Marker approach

This approach uses marked gloves wore in the hand and be colored to be helpful during the hand snapping to capture the fingers and palm. This glove frames the shape of hand by using the geometric-features. In (Lamberti & Camastra, 2011) utilized a wool glove with three different color to represent the palms and fingers. This methodology considers basic and not costly whenever contrasted using Sensor or Data Glove (Lamberti & Camastra, 2011), however the basic interaction among human and computer still is insufficient.

Table 3: Literature review on colored marker approach

Authors	Year	Description
Wang, Robert Y Popovi, Jovan (Wang & Popović, 2009)	2009	The authors proposed an easy-to-use and inexpensive system that facilitates 3-D articulated user-input using the hands. Their approach uses a single camera to track a hand wearing an ordinary cloth glove that is imprinted with a custom pattern.
Lamberti, L Camastra, Francesco (Lamberti & Camastra, 2011)	2011	Their recognizer is formed by three modules. The first module, fed by the frame acquired by a webcam, identifies the hand image in the scene. The second module, a feature extractor, represents the image by a nine-dimensional feature vector. The third module, the classifier, is performed by means of *Learning Vector Quantization*.
Hasan, Mokhtar M Mishra, Pramod K (M. M. Hasan & Mishra, 2012)	2012	The authors had focused on the researches gathered to achieve the important link between human and his made machines, also they had provided their algorithms for overcoming some shortcomings existed in some mentioned algorithms.

There are various machine learning algorithms were used for the recognition of hand gestures and few of them are discussed as follows [1] proposes a real-time hand gesture recognition system with four modules such as Data acquisition, Pre-processing, Feature extraction, Gesture recognition, and the Real-time hand detection is done using Histogram of Oriented Gradients (HOG) feature in MATLAB and the k-Nearest Neighbour (KNN) algorithm is used to classify the input images. A technique which commands computer using six static and eight dynamic hand gestures is proposed in [2] with three main steps hand shape recognition, tracing of detected hand (if dynamic), and converting the data into the required command, the system uses VGG16 a CNN architecture is used as the pretrained model to recognize the gestures.[3] Introduces an effective Hand Gesture Recognition system where the feature extraction is done using Deep Convolutional Neural Network (DCNN) and the classification is done using a Multi-class Support Vector Machine (MCSVM). [4] realizes the segmentation of hand gestures by establishing the skin colour model, the Hand gesture segmentation is done using AdaBoost classifier based on Haar and the CamShift algorithm for hand gesture tracking and the hand gesture area is classified using a convolution neural network. The framework for hand detection and gesture recognition is proposed in [5], the raw data coming from the Kinect is used to get the depth information of an image and an algorithm is proposed to identify each point of a closed contour identified within a given depth interval. Starting from the regained contour points, the centre of the palm is identified. The fingertips are localized by employing the k-curvature algorithm.

The dynamic hand gesture recognition for video processing is reviewed in this. Recently there has been extensive research in the field of hand gesture recognition and video recognition models and has made great progress and achieved highly efficient recognition rates in almost all the domains. Primarily in the hand gesture recognition was implemented using a wearable gesture recognition system as proposed in the following [6],[7] used flex sensors coupled with accelerometers detect the gestures based on the resistance generated due to the finger movements. Paper [8] proposed switch-based digital gloves embedded with 5 light-emitting diodes of each finger to identify the gesture. Wearable technology is limited to small range applications and requires hardware such as gloves to be worn constraining and uncomfortable to use. Deep convolution networks are known for efficient gesture recognition from images. The papers [9] [10] have implemented video-based action detection however this model focus on short-term motions and lacks the capacity to incorporate long-range temporal structure which are important for dynamic action understanding.

Also, these ConvNets require large sets of labelled data to achieve high efficiency. 2D Convolution Neural Networks (2D CNN) is proposed using Temporal Segment Networks (TSN) which extracts short snippets randomly in a long video sequence using a sparse sampling scheme along the temporal dimension and aggregate the information to enable long-range temporal modelling.

TSN are computationally cheap compared to 3D CNN but not great at inferring complex temporal relations. In paper proposes 3D CNN model where it has high spatial and temporal modelling characteristics however 3D CNN suffers from high computational cost as the numbers of parameters used are more compared to 2D CNN and are under the risk of over fitting. Hence to implement a good video-based action recognition system that has both high accuracy as that of 3D CNN and having low computational cost as that of 2D CNN, we used a Temporal shift Model (TSM) presented in which is inserted into the 2D CNN back bone. Uni-directional TSM algorithm enables to do real time video recognition with low latency and high temporal modelling on high computation device.

4.TECHNOLOGIES

4.1. GESTURE SENSING TECHNOLOGIES

Gesture recognition is the process by which gestures made by the user are made known to the system. It can also be explained as the mathematical interpretation of a human motion by a computing device. Various types of gesture recognition technologies in use currently are:

4.2. Contact type

It involves touch based gestures using a touch pad or a touch screen. Touch pad or touch screen-based gesture recognition is achieved by sensing physical contact on a conventional touch pad or touch screen. Touch pads & touch screens are primarily used for controlling cursors on a PC or mobile phones and are gaining user acceptance for point of sale terminals, PDAs, various industrial and automotive applications as well. They are already being used for automotive applications, and PDAs. User acceptance of touch-based gesture automotive systems technologies are relatively easier for the public to accept because they preserve a physical user interface.

4.3. Non-Contact

Device Gesture Technologies Device-based techniques use a glove, stylus, or other position tracker, whose movements send signals that the system uses to identify the gesture. One of the commonly employed techniques for gesture recognition is to instrument the hand with a glove; the glove is equipped with a variety of sensors to provide information about hand position, orientation, and flex of fingers. First commercial hand tracker, Data glove, used thin fiber optic cables running down the back of each hand, each with a small crack in it. Light is shone down the cable so when the fingers are bent light leaks out through the cracks. Measuring light loss gives an accurate reading of hand poses. Similar technique is used for wearable suits used in virtual environment applications. Though gloves provide accurate measurements of hand shape, they are cumbersome to wear, and connected through wires.

Various other kinds of systems are reported in literature for intrusive hand gesture recognition. Some uses bend sensor on the index finger, an acceleration sensor on the hand, a micro switch for activation. Styli are interfaced with display technologies to record and interpret gestures like the writing of text. To reduce physical restriction due to the cables, an alternate technique used is to wear an ultrasonic emitter on the index finger and the receiver capable of tracking the position of the emitter is mounted on a head mounted device (HMD). To avoid placing sensors on the hand and fingers the “Gesture Wrist” uses capacitive sensors on a wristband to differentiate between two gestures (fist and point). Wearing a glove or suit is clearly not a practical proposition for many applications like automotive.

4.4. Vision-based Technologies

There are two approaches to vision-based gesture recognition: Model based techniques: They try to create a three-dimensional model of the users hand and use this for recognition. Some systems track gesture movements through a set of critical positions. When a gesture moves through the same critical positions as does a stored gesture, the system recognizes it. Other systems track the body part being moved, compute the nature of the motion, and then determine the gesture. The systems generally do this by applying statistical modelling to a set of movements. Image based methods: Image-based techniques detect a gesture by capturing pictures of a user’s motions during the course of a gesture. The system sends these images to computer-vision software, which tracks them and identifies

the gesture. These methods typically extract flesh tones from the background images to find hands and then try and extract features such as fingertips, hand edges, or gross hand geometry for use in gesture recognition.

4.5. Electrical Field Sensing

Proximity of a human body or body part can be measured by sensing electric fields; the term used to refer to a family of non- contact measurements of the human body that may be made with slowly varying electric fields. These measurements can be used to measure the distance of a human hand or other body part from an object; this facilitates a vast range of applications for a wide range of industries.

5.COMPARISION TABLE

Table I Comparision of different techniques for video processing

METHOD	ACCURACY	PURPOSE
Video inpainting	Medium	Filling the gap at the foreground
Graph based region segmentation, with priority function	High	Number of iterations of filling the gaps are reduced.
Best first algorithm, Block based algorithm	High	Filling the gap for large objects
Pix Mix algorithm Pixel based approach	High	To change the live video streams, to increase the accuracy in real time applications.
Background subtraction method, region based approach	Medium	Filling the gap by region based
Foreground and background method, classifiers	Medium	To reduce the false detection of the object.

Table II Comparision of different techniques for handgestures

METHOD	ACCURACY	PURPOSE
Hidden Markov Model for data glove.	98.7%	Spatio temporal variability is reduced.
Multiscale Gesture Model.	88%-96%	Segmentation and recognition of the hand.
Accelerometer-Based gesture recognition, k-means and Fast fourier transform algorithm.	Upto 95%	Recognition and normalization and filtering the gestures.
Novel hand gesture recognition scheme, SVM classifier.	95%	3D recognition of hand gestures.
CAMSHIFT algorithm, PCA algorithm.	93.1%	Recognition , Segmentation and normalization of hand gestures.
Discrete Hidden Markov Model.	ranges from 93.84% to 97.34%	Recognition for dynamic hand gesture
Finger Earth Movers distance metric method.	Upto 93.2%	Only fingers of the hand are recognized.

6. SYSTEM ARCHITECTURE

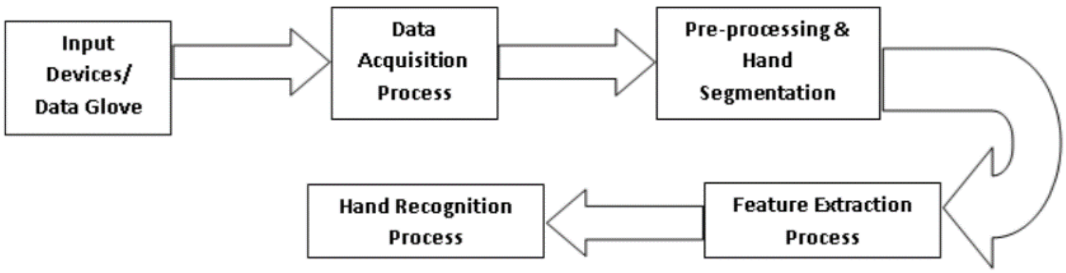

Fig. Generalized System Architecture for Hand Gesture Recognition

7. APPLICATIONS

1.Robot Control: Using gesture recognition we can control the robot easily. After hand gesture is recognized for performing the certain actions set the movement of robot for movement of hand or finger count of hand for e.g. “one” means “move forward”, “five” means “stop”, and so on.

2. Television Control: Hand postures and gestures are used for controlling the Television device. In a set of hand gesture or count of finger are used to control the TV activities, such as turning the TV on and off, increasing and decreasing the volume, muting the sound, and changing the channel using open and close hand.

3. Desktop and Tablet PC Applications: In desktop computing and PC applications, gestures can provide an alternative solution to the mouse and keyboard. Many gestures for desktop computing tasks involve manipulating graphics or annotating and editing documents using pen-based gestures.

4. Games: Gesture is used for computer games. Using gesture, we can easily interact with computer. In video game using gesture track and control the player’s movement or recognize the position of players. Using gestures control the movement of avatars in a virtual world and play station.

5. Sign Language: Sign language is an important part of communicative gestures. Sign languages are highly structural; they are very suitable for vision algorithms. At the same time, they can also be a good way to help the disabled people to interact with computers. Sign language for the deaf people has received significant attention in the gesture literature.

6. Healthcare & Medical Assistance: In healthcare and medical field also gesture technology is used. Using gesture patients can control the instrument which they require for their exercise. Gesture based tool used for sterile browsing of radiology images. Also, researcher developed a wheelchair with intelligent HCI.

7. Daily Information Retrieval: Researchers implemented an approach that provides daily information retrieved from Internet, where users can operate this system with his hands’ movements.

8. Education: In Education system we also used hand gesture recognition system. Example of a such system is using hand gesture control the power point presentations.

8. FUTURE SCOPES

The Hand Gesture recognition is moving at tremendous speed for the futuristic products and services and major companies are developing technology based on the hand gesture system and that includes companies like Microsoft, Samsung, Sony and it includes the devices like Laptop, Hand held devices, Professional and LED lights. The verticals include where the Gesture technology is and will be evident are Entertainment, Artificial Intelligence, Education and Medical and Automation fields. And with lot f Research and Development in the field of Gesture Recognition Field, the use and adoption will become more cost effective and cheaper. It's a brilliant feature turning data into features with mix of technology and Human wave. Smart phones have been experiencing enormous amount of Gesture Recognition Technology with look and views and working to manage the Smartphone in reading, viewing and that includes what we call touch less gestures. Google Glass has been also in the same cadre.

And the Technology has also been embedded into smart televisions nowadays as well, which can easily control and managed by Voice and Hand options. In the medical fields Hand Gesture may also be experienced in terms of Robotic Nurse and medical assistance. As the Technology is always revolving and changing the future is quite unpredictable but we have to be certain the future of Gesture Recognition is here to stay with more and eventful and Life touching experiences.

9. CONCLUSION

In the present world there are numerous was of providing input to the smart devices such as a smart tv, virtual environments used for various purposes, home automation systems etc. Although these are not feasible to the trending and upcoming technologies, thus to meet this standard in this paper we put forward a hardware efficient dynamic gesture recognition system, which could be a boon to the differently abled people as it lays a natural form to communications with hand movements to the smart devices. The model uses temporal shift model combined with CNN techniques to extract spatiotemporal features for efficient action detection. A data set of around 16 hand gestures can be predicted by the proposed prototype, also we are further planning to implement this model on a PowerPoint presentation where the presenter can control the movement of the slides it just by using hand gestures without any other support.

10. REFERENCES

[1] Sanmukh Kaur, Anuranjana (2018), "Electronic Device Control Using Hand Gesture Recognition System for Differently Abled", 8th International Conference on Cloud Computing, Data Science & Engineering.

[2] Soeb Hussain, Rupal Saxena, Xie Han, Jameel Ahmed Khan, Hyunchul Shin (2017), "Hand Gesture Recognition Using Deep Learning", International SoC Design Conference

[3] Md Rashedul Islam, Rasel Ahmed Bhuiyan, Ummey Kulsum Mitu, Jungpil Shin (2018), "Hand Gesture Feature Extraction Using Deep Convolutional Neural Network for Recognizing American Sign Language", 4th International Conference on Frontiers of Signal Processing.

[4] Jing-Hao Sun, Shu-Bin Zhang (2018), "Research on the Hand Gesture Recognition Based on Deep Learning", 12th International Symposium on Antennas, Propagation and EM Theory.

[5] Guillaume Plouffe and Ana-Maria Cretu (2016), "Static and Dynamic Hand Gesture Recognition in Depth Data Using Dynamic Time Warping", IEEE Transactions on Instrumentation and Measurement, Vol. 65, Issue: 2

[6] El Hayek, H., Nacouzi, J., Kassem, A., Hamad, M., & El-Murr," Sign to letter translator system using a hand glove", In the Third International Conference on eTechnologies and Networks for Development April,2014 IEEE.

[7] Tubaiz, Noor, Tamer Shanableh, and Khaled Assaleh. "Glove-based continuous Arabic sign language recognition in user-dependent mode." IEEE Transactions on Human Machine Systems 45, n o.4 (2015)

[8] Praveen, Nikhita, Naveen Karanth, and M. S. Megha. "Sign language interpreter using a smart glove." In 2014 International Conference on Advances in Electronics Computers and Communications, pp. 1-5. IEEE, 2014.

[9] Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Largescale video classification with convolutional neural networks. In: CVPR, pp. 1725– 1732 (2014)

[10] Zhang, B., Wang, L., Wang, Z., Qiao, Y., Wang, H." Real-time action recognition with enhanced motion vector CNNs," In CVPR, pp. 2718–2726 (2016)

[11]https://www.virtusa.com/digital-themes/gesture-recognition

[12]https://www.ijrte.org/wp-content/uploads/papers/v9i2/B3417079220.pdf

[13]https://www.mdpi.com/1424-8220/20/12/3571/htm

[14]https://www.engineersgarage.com/gesture-recognition-technology/

[15]https://www.semanticscholar.org/paper/Literature-Survey-on-Hand-Gesture-Recognition-for-Swapna-Shaikh/fe147092dbfb0c5ddb6a956cd98c13fcacc56a1f

[16]https://intellias.com/hand-tracking-and-gesture-recognition-using-ai-applications-and-challenges/