International Journal of Engineering &Technology, 7 (3.12) (2018) 1175 -1181

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Non – Vision Based Sensors for Dynamic Hand Gesture Recognition Systems: A Comparative Study

Panduranga H T[1], Mani. C[2]

[1]*Department of EC, Government PolyTechnic, Turuvekere, Karnataka*
[2]*Department of EC, Maratha Mandal Engineering College, Belagavi, Karnataka*
**Corresponding Author E-mail: [1]ht_pandu@yahoo.co.in , [2]mani.pe12@gmail.com*

Abstract

Gestures are considered as a type of configuration associated with motion in concerned body part, signifying meaningful information or expressing motion or intending to command and control. Wide ranges of sensors working with different technology are available in market. Gesture recognition process involves steps like data acquisition from sensor, segmentation, an algorithm for taking gesture data as input, an algorithm to extract parameters and algorithm to classify hand gestures. Three - dimensional hand gestures have been widely accepted for advanced applications like creation of virtual world where in users can feel the naturality of interacting or playing a musical instrument without presence of any physical device. Techniques for dynamic finger gesture recognition can be classified as visual based and wearable sensor based. The purpose of this paper is to compare various non – vision based sensors with different tracking technologies, updating advantages and drawbacks helping investigators and researchers working on this area.

Keywords: *Human Computer Interaction; Motion Capture; Hand Gesture Recognition; Machine Learning;*

1. Introduction

Human hand is a flexible joint connecting 27 bones and 5 fingers. Each finger (index, middle, little, ring) has three joints. All the four fingers are structured connecting to wrist bones tied and there is a thumb at a distance. Human hand joints can be classified as flexion, twist, directive or spherical depending on type of movement or possible rotation axes. Gestures are language and culture specific. Variety of gestures can be generated with huge DOF (Degree Of Freedom). Therefore hand need to be modeled in a proper manner treating as interfacing unit for Human Computer Interaction[1]. The motivation behind hand gesture recognition is to give interaction naturally between human and computer using hands for variety of different applications like video games, robotic control, sign language interpretation or conveying information, creating virtual reality etc. A gesture with meaningful information has both static and dynamic gestures. A dynamic gesture is intended to change over a period of time. Hand gesture recognition is a process of recognizing and interpreting a continuous stream of input data obtained from the features of hand to identify or classify a particular gesture depending on the application[3]. Dynamic hand gesture recognition takes the parameters of the gesture. These data's are continuous and sometimes repetitive. A gesture is complete, when these data's are clubbed with user's hand and arm positions. Different users perform gestures at different speeds creating a non-linear wave track on the time line. Lot of work is carried out in elimination of such non-linear dynamic fluctuations[2]. To gather data for temporal gesture recognition, 2D motion paths have to be taken from image sequences. Although 2D tracking algorithms such as color / motion / blob tracking, and template matching helps in positional information of hand, few recognition application require more features of hand such as hand inclination and hand pattern.

3D tracking ways helps to place hand in 3D space[5]. The hand shape in 3D can be modeled into geometrical, skeletal or volumetric. Volumetric models are difficult in nature and have computational difficulty in reality. Skeletal model represents the hand with 3D structure with less set of parameters. Gestures identified in terms of position and inclination in space over time is called as 3D gesture. 3D gestural interaction provides a strong and real way to interact with computing systems using hands. Patterns has to be identified in a continuous stream of data in 3D space and with suitable signal and image processing techniques, has to be classified into specific categories using machine learning algorithms. Therefore sensors have been broadly classified as active and passive sensors. In Active sensing process the user has to hold the device or wear it to the fingers by some means, whereas passive sensing are basically vision based. Depth based sensors like Kinect and Leap motion controller have been more popular in this category. These cameras provide more information than traditional cameras supporting and enabling interaction through skeleton tracking of hands and fingers. Apart from advantages like naturality and less hardware complexity, vision based sensors limitations are related to camera's field of view, recognition under low lighting conditions, occlusion, low frame rate and removal of background noises etc. In many applications to reduce the complexity and improve the performance we prefer hybrid technology combining both active as well as passive (e.g. combining camera with accelerometer). The authors in this paper therefore have focused on providing an up to date information on active sensors and also the feasibility of using multiple sensors. Attempt is made to give overview of extracting important features from the data and using those features as input to a classification algorithm[6].

2. Features and Classification of Gesture Recognition

Hand gestures yield large richness in parameters related to shape motion and textures.. Therefore selecting proper features is very much essential for gesture recognition. Features extracted may be geometric or non-geometric. Geometric features are related to fingertips, direction of finger and hand contours; whereas non-geometric features are color, silhouettes and textures. Some features may be inadequate and some features may be hindered due to occlusion, low lighting conditions and may not be always reliable. The whole image or sometimes transformed image is to be fed as input to the gesture recognition system. Some gesture recognition applications require more features from the above 2D features such as hand inclination and position, especially in case of dynamic hand gestures recognition. Therefore 3D (Three – Dimensional) tracking methods places hand in 3D space in terms of hand position and orientation [4]. The methods to classify hand gesture are done using machine learning or rule based. The latter consists of manually encoded rules between features. Here extracted features are compared with encoded rules. The rule that matches with the gesture is resultant gesture. The major problem with the rule based approach is, there cannot be fixed rules and it depends on the ability of the person to frame the rule and may vary. In machine learning approach, the output of a gesture is treated as a stochastic process [7].

3. Motion Capture

Motion capture is use of proper sensing technique to pick and store the data. Commonly representations and tracking of motion is dealt in digital domain. Therefore a motion tracking system should be capable of sensing, processing and holding the processed data as shown in fig [1].

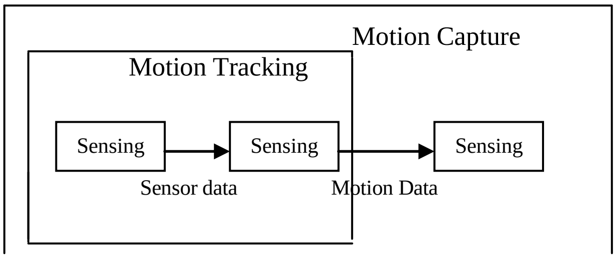

Fig. 1: Motion Capture Block Diagram

Storing data is optional. For real time interactive applications, the processed data is directly fed as input to the recognition system. The data is sensor specific. It may vary from measuring analog parameters like voltage, field, resistance etc., to obtaining color information from the camera pixels. The processing section deals with transformation of raw data into the parameter that describes the motion, for example: getting the derivatives such as rotation, acceleration and velocity from the position. Advanced systems are available providing joint angles. Term Degrees of freedom (DOF) denotes the number of rotational and linear joints in kinematic models. Types of tracking with the help of pointed objects e.g. spherical markers are able to find position. Several patterns will represent complete rigid object. Others can be rigid objects - non-deformable structures, which are capable of tracking both orientation and position. A kinematic model can be developed and rules can be defined for various rotations and translations occurring between them.

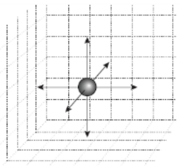

Fig. 2: Marker

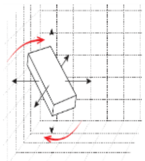

Fig. 3: Rigid Object

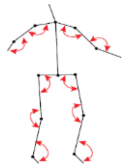

Fig. 4: Kinematic Model

Efficient motion tracking systems use higher sampling rates. Classification of tracking technologies is based on the physical medium of the technology. It can be through electric, acoustic, mechanical, magnetic, inertial and optical tracking means.

A. Electric Tracking

Example of popular touch screens like Resistive touch screen has conductive layers, causing to change the resistance proportional to pressure in the layers. Similarly in capacitive touch screens, change in capacitance is proportional to pressure on the capacitive surface that holds the charge. These compete for increased resolution and durability [8].

B. AcousticTracking

Tracking through acoustic means can be based on the concept of time of flight or through measurement of phase difference between the transmitted and received signal. Efficient tracking systems, to find out 3D position of the receiver use multiple transmitters. Tracking will be difficult due to variation of sound speed and temperature in air. Acoustic systems usually work in the ultrasonic range and therefore development of ultrasonic sensors is utilized for playing digital musical instruments [9].

Acoustic gesture recognition systems combine TOF (time Of Flight) along with triangulation method to determine the object. But the requirement is that both the object and the transmitter are to be in line of sight. Advantages of acoustic based systems are implementation with lower cost and small sensor size. Disadvantages are related to limited precision because of multiple reflections from the surface in the environment, low speed of response due to retardation of received acoustic signal and the need of pulse synchronization. Other factors to be considered which limit the performance include susceptibility of sound wave to pressure, temperature and humidity [8].

Yuri et al. uses HC-SRO4 ultrasound sensors. Sensors are designed to perform a highest rate of 200 fps, capable of detecting objects in the range 20 and 30 millimeters. Eight sensor matrix is driven by an arduino board, supporting each sensor through a trigger signal and waits for the response, and these times are calculated to obtain the distance. Actions are realized using library newPing. Estimation of velocity of fingers is done with the help of distance thresholds. Data from arduino supports to use distance estimation algorithm. Limitations include encumbrance, not much portable for user interface. Ultrasonic based systems amounts to about one hundred Euros in terms of hardware prototypes [10].

Paper from Cheng et al. describes about soundtrack- The methodology is to work with a ring embedded with small speaker on the finger transmitting sound signal at 11 KHz. Array of microphones placed captures the received signal having accumulated phase information. Area on which it can cover the spatial resolution is around 1.3cm. This sensing technique tracks the finger position in 3D space with a mean Euclidean distance of 1.305 cm within a volume of $3520cm^3$, using four MEMS microphone and one speaker. Physics based model is used to derive location of a tracked speaker and hence no machine learning techniques. Tracking is limited to only one finger. The phase of received multipath signal may contribute error to the system, strengths with multipath signals will be lower than the straight line signal. Speaker need to need to be powered separately. Detection is difficult, whenever there is block in audio signal or sometimes even the finger position may be out of range. When the audio signal blocks or when the finger moves out of range, detection becomes difficult. Calibration is required frequently. High speed in receiving the signal may cause large frequency deviation due to Doppler Effect and result in phase difference. Recent paper based on Low Latency Acoustic phase (LLAP) has been implemented to detect the sound signal reflected by the moving hand or finger. Phase variation of echo signals caused by hand is obtained and used in classification and detection of finger position [11].

Sangki Yun et al. present acoustic device free tracking. The phase from channel impulse response is estimated instead of receiving raw signal. Here to address the problems associated with multipath signal propagation, the channel impulse response is estimated and channel coefficient is calculated for each channel tap. Features include high accuracy within 0.3cm and low latency. Accuracy can be enhanced by calculation of relative distance and absolute distances [12].

Rajalakshmi Nandakumar et.al presents fingerIO - fine grained finger tracking system. Here the finger is free from instrumenting sensors and overcomes between the finger and the device. It tracks the finger movement around the smartphones achieving 2D accuracy of 8mm at 169 fps both on surfaces and on walls. The interaction space is $0.5 – 0.25m^2$ around the device. Smartphones ace as active sonar systems. An inaudible signal of 18-20Khz signal is transmitted from device speaker. The reflected signal from the finger is received as echo at device microphone. Algorithms use the features of OFDM to follow the changes in the echoes due to movement in finger. The echo profiles are generated for all reflections at the microphone and the actual echo related to the moving finger is cross correlated. Peaks in the profile are processed for finding distance for the object present between microphone and speaker. However detection limited to 2D can be improvised for 3d tracking by using additional microphones and triangulated to find the position of the finger for 3D finger tracking. Others being chances of receiving noisy echoes and signal obstructions, detection done on only one finger and device is static[13].

Paper implemented by Wenguang et al. presents a high precision acoustic tracker CAT. System uses audio signals to achieve mm level accuracy with the help of multiple speakers transmitting at different inaudible frequencies. Estimation of parameters like velocity and distance of mobile with respect to speakers are done by using distributed FMCW and Doppler shifts. Then the distances and velocities are fused in an optimization framework for tracking the movement. With the help of frequency shifts, speed of the mobile with respect to speakers can be estimated and position is estimated by integrating speed over time. Prototypes designed here give 5-7mm error in 2D and 8 – 9mm error in 3D. The mobile samples the received signal at 44.1 KHz and frequencies are extracted with the help of STFT. This is compared with the frequency of original sine wave and peak frequency is obtained and velocity is estimated using appropriate equations [14].

C. Magnetic Tracking

Magnetic systems works with the principle of electromagnetic induction. Magnetometers can be passive and active. Active magnetometers utilizing an electromagnetic source and sensor with multiple coils, determine the field strength produced in each sensor coil. Then the direction and position of each sensor coil is measured. Among available sensors, magnetic systems have capable of operating with higher sampling rates of about 200Hz [9].

Commercially available 3D device called Polhemus leveraging magnetic field sensing, can track objects with 6DOF. Commercially available high quality 3D magnetic tracking devices Polhemus provides rms error below 0.025m up to 3m, a wireless sensor technology giving a high update rate of 188 Hz with 1-8 markers. Microsensor 1.8 is a flexible embeddable into glove, worn under hat or helmet, can be inserted into traditional catheter, without LOS can provide 6DOF for hand and finger tracking.

Paper represented by Ke-Yu chen et al. titled Finexus is of magnetic sensing to track the precise motions of multiple fingertips. The fingertips are instrumented with four magnetic sensors. The electromagnets are driven using alternating current at different frequencies. To overcome interference of noise, each electromagnet is operated at different frequency and proper filters are used to distinguish signals from individual sensing points. The size of electromagnets used here is shrinked compared to traditional implementations and utilized 1D magnetic field for positioning. The sensing space is converted into beacon – like system and trilateration is used for positioning. The system has an average accuracy of 0.95mm (with fixed orientations) and 1.33mm (with random orientations) within a sensing distance of 120mm.. System first calculates the distance between electromagnet and magnetic sensors, trilateration is used to identify the electromagnet's 3D position [15].

D. Optical Tracking

Motion tracking with optical systems utilize cameras of type either video / IR /Depth. Cameras depending on their type, they sense light which is related to range in electromagnetic spectrum. Each pixel in the image gives a value related to amount of light sensed. IR (Infra-Red) cameras sense light related to infrared region of electromagnetic spectrum. IR cameras used for applications either captures light from the source directly or signal reflected from the objects.

IrMoCap is based on set of cameras placed around the object under test. Cameras on emitting IR light are reflected by the reflective markers attached to the object. Through triangulation technology, systems calculate absolute position in space with good resolution speeding above 500Hz [16]. Advanced technologies depending on the applications consist of a group of IR LED's positioned near IR camera and capturing the reflected light from the spherical markers. Depth cameras also called TOF cameras provide additional information along with information provided by 2D image. These cameras are enclosed with an IR emitter in which light is reflected from the objects field of view. Calculation of the distance towards each pixel is determined based on speed of light. The other method is to emit a pattern of light and to find the deformation in the pattern when it is been reflected from the objects [9].

The following camera specifications like range, resolution and accuracy, frame rate, camera computer interaction and lens can modulate the characteristics of gesture recognition system. For hand gesture recognition, devices commonly used are quantum detectors working in visible light spectral range (350 – 750 nm). Here detectors rely on light reflected by the object (human hand) towards the focusing lens. Digital cameras associated with sensor arrays are characterized by either CCD or CMOS type, each element size in the array, pixel resolution and color depth. Color

depth means the number of bits used to identify pixel color. Common frame rate (frames per second) standards for video recording are 30 for NTSC and 25 for PAL in U.S and Europe respectively. Low grade connections between computer and camera have limitation on transfer rate. USB 1.1 can support only up to 12Mbps whereas USB 2.0 standard allows up to 480 Mbps and 4.8Gbps for 3.0. Further frame rates are degraded by capabilities of computer and run time load. Cameras FOV is dependent on lens focal length while lens aperture is space for letting amount of light.
Optical Motion Capture (MoCap) are widely used for real time applications due to their reduced cost, flexibility and availability of tools, examples are Max (MSP/Jitter and Eyes web). Non – optical tracking systems have to go for sensors like magnetometers, accelerometers and gyroscopes. Here each sensor output's the relevant data in themselves. Optical motion capture systems combine with sensors called as sensor fusion to obtain raw data. Here it is required to combine data from individual sensors and processed to obtain position and orientation with least drift.

E. Inertial Tracking

Inertial tracking systems utilize gyroscopes and accelerometers. Whenever we apply force to the accelerometer, acceleration developed on small "proof mass" is measured. Gyroscopic instruments are required to measure changes in rotational movement. For applications like dynamic hand gesture recognition, the performance is measured in terms of degree of freedom. Therefore above devices are effectively utilized with sensors mounted placed at right angles.

- Inertial trackers are self- contained.i.e; they do not depend on external sources like ultrasound acoustic sensors or cameras.
- They are not disturbed by lighting and ferromagnetic materials.
- The sensors are of lesser weight and hence portable.
- Due to their low latencies, sampling rates will be higher.

Orientation is obtained by integrating the data from the gyroscopes. Position is determined by first adjusting for any change in the gravity vector and then integrating data twice. A small error from gyroscope or accelerometer data can produce a large error in positional estimation, and hence these trackers can work to the best when combined with other tracking technologies. Inertial measurement units (IMU's) contain accelerometer, gyroscope and magnetometers (compass).

F. Mechanical Tracking:

Here we design systems based on mechanically constructed parts which measures lengths or angles between the parts with the help of potentiometers or bend sensors [motion capture]. Exoskeletons have become popular for identification and tracking of finger motion to few degrees of freedom. These are rigid structures designed to follow up, to certain extent, the motion of hand. The sensors are implemented in an exoskeleton to derive model of the joint angles of hand. The main difficulty here is to create an adaptability of exoskeleton to variation in hand deformations [background]. Mechanical tracking has been in demand, particularly for the purpose of developing new musical interfaces.
Xiaochi et al. presents exoskeleton – Here the design is for motion capture working with the concept of passive force feedback especially for virtual reality applications. The construction consists of using multiple sensors, link rod structures and actuation units. Above Paper presenting Dexmo has components like main controller, feedback units, upper and lower connectors and finger caps. The rod structures exhibits motion whenever the user's hand are in motion. The controlling unit collects the data from the rotational sensors connects to the computer. The algorithm used here is kinematics regeneration algorithm. The device only tracks MCP joint of each finger sensed by the rotational sensor and the other two joint values, DIP and PIP are assumed. Dexmo offers ruggedness, reduced weight and lower power consumption, weighs less than 270 gm. The disadvantages with the system are associated with binary (on / off) haptic feedback system, delay with force feedback unit. Mean error rate is 61% in without-force feedback and 44% in force feedback condition[17].

Table[1][2] describes features, advantages and limitations of different tracking technologies. In order to improvise over performance and efficiency of hand gesture recognition system, latest researchers have involved in encumbering different sensors especially in applications like robotic control for surgery, creating virtual reality, Augmented reality, playing video games etc,. Hybrid recognition systems are developed by combining one or more sensors. Modern devices are instrumented by integrating different sensors. Data from different sensors connected to different parts of the body are either processed independently or integrated in a form required and then provided as input for recognition algorithms. The aim to fuse the data from different sensors (multisensory data) is to have improved accuracy, authenticated recognition, reliability and working- range. Dom Brown et al. combines inertial measurement units combined with leap motion optical sensor for musical environment. Markus et al. presents finger and hand motion capturing system using inertial and magnetic sensors. Jess et.al gives information in combining EMG and pressure sensor data at wrist, quoting classification accuracy of 96%. Similarly Kristian Nymoen et al. develops MuMYO which consists of combining EMG sensors with IMU. Data is sensed and processed wirelessly through reflective markers for calculating 3D position. Paper titled Glove sense MIDI – capable data glove by Mert et al. for musical applications implements piezoelectric flex sensors and inertial measuring units. Data gloves used here along with the Max software detects gestures with the help of piezoresistive flex sensors, accelerometers, gyroscopes and magnetometers. Data from IMU are in terms of Quaternions to describe the rotation data in 3D space (yaw, pitch and roll). The prototype was implemented with arudino mini and two flex sensors, sending data wirelessly through BlueSMiRF within 20m range. Implementation of Mi.Mu glove for gesture recognition lead to accuracy of 90%.
Eider et al. combines armband and Leap motion sensor for full arm tracking. MYO detects for rotation data from arm and leap motion tracks hand motion. Data used as input to algorithms are fed in probabilistic model to estimate real time body's limb position and orientation. Data here also is processed in quaternion and results are quoted using graphical way. Kalman filter is an estimate algorithm used to deal with real time measurements. 3D models utilized here, get transmitted data from sensors and simulate human hand movements. Rotational data from both sensors and data fusion results are converted from quaternion to euler angles. The Kalman filter establishes an average of values determined by both the sensors, but still noise like salt exists[20].
Xu Zhang et al. paper encourages implementing multichannel EMG sensors and accelerometers for hand gesture recognition. Here decision tree and HMM are used as decision level fusion to obtain the final results. The procedure is obtaining signals from accelerometer and EMG sensors, out of which active segments are identified. Statistical values like standard deviation and mean value is estimated for the data and are sent to decision tree classifiers through multi stream HMM classifiers for recognition. For training and classification of hand orientation, K-means clustering and linear discriminant classifiers are proposed. The overall accuracy with the recognition system implemented was 90.2%[21].
Paper titled Leimu from Dom Brown et al. combines leap motion optical sensor with IMU. Algorithm AHRS fusion gives estimated instantaneous values of hand orientation with reference to earth coordinate frame. The skeletal geometry of tracked hand

is converted to obtain joint angles and this combined with gestural events plus orientation data from IMU are used for analysis and mapped for audio and music parameters. The results are compared with glove consisting of flex sensors and it is found that the accuracy with the combination of leap motion and IMU is increased[22].

4. Hand Gesture Recognition

The first step towards gesture recognition system is extracting appropriate features from the hand or availability of data set in a proper form. The gestures having meaningful information can be recognized by using either machine learning algorithms or with the help of suitable classifiers. Popularly implemented methods for hand gesture recognition are template matching, dictionary look-up, statistical matching, linguistic matching, neural network and adhoc method. Parameters like gesture shape, duration, velocity and integrality keeps varying in time – space, due to which significant complexity arises and hence performance deviation and complexity is significant. HMM statistical model have good response because of its capability in modeling variable time – series gestures and retaining the identity of hand [23]. Tools for recognition can range from statistical modeling, computer vision and pattern recognition, and image processing etc. Learning algorithms like ANN (Artificial Neural Networks) and HMM are from the community of artificial intelligence. It is been popularly implemented by researchers as it is found that that the recognition accuracy can be increased from training.

A. Neural Network

Using the concept of active human brain, an artificial neural network is modeled. This network is formed with nodes acting as a fundamental component. The nodes are computationally capable of calculating and storing values. The main functions associated with the nodes are, to calculate the weighted sum of its inputs and with activation function to convert weighted summation into final result

Common activation functions are sign, step and sigmoid. Using any one of these activation functions, if weighted sum is higher than certain threshold, then the function returns a value ‘1’ indicating the particular node is been fired. Networks are trained in either supervised or unsupervised manner. In supervised learning, the network is trained by providing match between the input and output patterns. This trains the network in advance and as a result, network is not learned while running. In an unsupervised learning or self-organizing system, the network is trained to respond to a large cluster of patterns within the input. There is no training in well - before and the system should give its own representation of the input, since no matching is provided.

Algorithms are implemented depending on the model. Examples are feed – forward, back propagation algorithms – the network learns from the desired output.

Neural networks are capable of improving accuracy but dependent on network training. They are implemented for both wearable sensor and vision based sensor technologies. However the limitations are: Results vary with different types of configurations used in the network and therefore difficult to determine which is best, increased time consumption for network training, reframing entire network for new gesture.

B. Hidden Markov Model (HMM):

HMM’s statistically modeled networks. It recognizes patterns which are learned through observing sequences of a particular kind of data. The system is modeled which is considered to be one in set of particular number of states. The process is identified as markov process whenever it satisfies the conditional probability distribution of future states. The intermediate states in the HMM are not observable [24] but identified by other set of stochastic process producing sequence of observable symbols. We need to measure the maximum likelihood performance from the parameters chosen in model. The difficulties in HMM design is from three steps like probability evaluation, finding optimal state sequence and parameter estimation. HMMs provide solutions for wearable sensor and vision based sensing technologies, can be implemented for large set of gesture sets, and has registered for improved accuracy and performance with adequate training. However, implementation is complex due to time consumption for training and hidden nature [25]. Model is appropriate for dealing with statistical properties in gesture recognition. Researchers use HMM for data which are continuous in time series. Extensive use of HMM is because of its high recognizing rates. Few implementations also consists of combining with classifiers such as HMM with dynamic Bayesian network and conditional random fields, HMM with K-NN (K-Nearest Neighbor) and SVM (Support Vector Machine) etc., SVM is again classified under supervised learning operating on a set of input data and prediction of the output for every given input based on its training on standard available data set.

5. Conclusion

Data collection and representation in dynamic hand gesture recognition system can be either from the glove based or from vision based. The accuracy is more in glove based systems, but it depends on type of sensor technology and also expensive. The other part of difficulty is the type of recognition technique is to be implemented that can maximize the accuracy and enable robustness. A clear decision has to be taken in terms of feature extraction, statistical parameters estimation and network learning algorithms. An attempt is made to compare the distinct features involved and highlighting the parameters involved with each of the tracking technologies. The methods of implementation and conversion starting from the process of acquiring the data and their usage is documented and proves to be helpful for budding researchers working in this area for different applications

Table 1: Comparison of different tracking technologies

	Technology / Features	Advantages	Disadvantages
Magnetic tracker	• Uses source devices that generates electromagnetic field • Sensor reports its position and orientation relative to the source • Uses magnetic field generated by static transmitter to obtain the position of receiver in motion.	• Low cost, reasonable accuracy • No requirement of LOS for tracking purposes. • Higher speed response compared to acoustic.	• Disturbed by surrounding ferromagnetic or magnetic materials. • Transmitter and receiver should be free from metallic objects in working range. • Support tracking frequencies upto 100 Hz. • Inaccuracies due to the metallic interference in the magnetic field • Encumbered system
Ultrasonic tracker	• Uses ultrasonic signal generated by static transmitter to estimate the position of receiver in motion.	• Not disturbed by metallic interference	• Requirement of LOS from the source to the receiver • Acoustically reflective surfaces can

	• Uses triangulation method to determine the position of the source		cause interference in the system
Optical tracker	• Uses optical sensing to determine orientation and position of object. Optical tracking uses markers and silhouttes Requirement of more cameras to detect markers. Silhouette method use edge detection to extract silhouette of gesture. Methods to estimate orientation and position of object.	• Insensitive to metallic interference. • Give significantly higher update rates compared to acoustic trackers.	• Require direct LOS between transmitter and receiver • Sensitiveness to light reflections of surfaces in the environment. • Processing power required to make large calculations at that rate is large. • Accuracy depends on number of cameras used. • More cameras used leads to increased complexity in calculations. • Conventional video camera is inefficient for capturing finger details. • Occlusion problems cannot be solved with single camera. Frame rate (30 to 60 frames per second) is not sufficient for capture of rapid movements
Inertial tracker	A self-contained sensor for estimation of rate of change of object’s velocity	Source less operation. Unlimited range No LOS requirements Very low sensor noise	Position and orientation are second order derivatives of inertial tracker output parameters. These have sensitiveness to drift and bias of sensors.

Table 2: Comparison of available glove based systems

	Features	Authored By	Classifier	Algorithm	Recognition rate	Limitations
Data glove		Mario ganzeboom		Training HMM with K-means algorithm	93%	Wearable to one hand Restricted to single user since size of hand varies
Data glove(DG5 VHAND 2.0) using flex sensors	Wireless data glove Connectivity range – 10m	Piyush Kumar et al.	K-NN - classifies data based on Euclidean distance			
HiTEg Glove V4(Piezoresistive)	15 sensors placed at the joints 90° -highest value of bending for each finger Static gestures recognition	G. Saggio et al.	SVM , Mahalanobis Euclidean		90 % with SVM and Euclidean	Maximum angle detection is from 0 to 90°
Glove with pressure and bending sensors (printed polymeric)	Wireless using zigben for data transfer	Nattapong and Natthapol		Microsoft visual c# used to develop tracking software		Bend angle detection from 0 to 60°
Tyndall IMU glove (16 9-axis IMU) – MPU-9150	9 DOF inertial sensors are used Wireless interface	Brendan O’Flynn et al.			85.24 %	
Sensorized glove (flex resistive sensors)	10 sensors are used and placed at proximal joints Optical system measures joint angles with markers. Transmission of data through planar antenna ANT-433-S OOK modulation for transmission	Michela Borghetti et al.		Labview interface for conversion resistance angle	95.4%	System to be calibrated for new users
Data glove equipped with magnetic sensors	Measures 17 DOF of hand	Chin and Herman			50mV output with 45°	Angle variation estimation from 0 to 45°

References

[1] Arpita Ray Sarkar, G.Sanyal, S.Majumder , “Hand Gesture Recognition Systems: A Survey”, Internationational Journal of Computer Applications ,Vol.71 – No 15, ,pp: 0975-8887, May 2013.

[2] Shuai Yang, University of Wollongong, Research Online, “Robust Human Computer Interaction Using Dynamic Hand Gesture Recognition”, Doctor of Philosophy thesis, July-2016.

[3] Aarti Malik, Ruchika, “ Gesture Technology: A Review”, International Journal of Electronics and Computer Science Engineering”, ISSN – 2277-1956, Pages 2324 – 2327.

[4] G. Simion, V.Gui and M.Otesteanu.” Vision Based Hand Gesture Recognition: A Review”, International Journal of Circuits, Systems and Signal Processing”, issue – 4, Vol – 6, 2012, pages 275-282.

[5] Ying Wu, Thomas S. Huang, “Vision – Based Gesture Recognition: A Review”, Beckman Institute, University of Illinios at Urbana – Champaign, Lecture Notes in Computer Science, Gesture Workshop, 1999.

[6] Joseph J LaViola Jr, “3D Gestural Interaction: The State of the Field”, Review Article, ISRN Artificial Intelligence vol 2013, October 2013.

[7] G.R.S.Murthy & R.S.Jadon, “A Review Of Vision Based Hand Gesture Recognition System”,International Journal Of Information technology and Knowledge Management”, July – Dec 2009, vol-2, No.2, PP 405 – 410.

[8] Sigal Berman. Helman Stern, “ Sensors for Gesture Recognition Systems”, IEEE Transactions on Systems, Man and Cybernetics Part C(Applications and Reviews) Vol.42, No.3, May – 2012.

[9] Motion Capture – Chapter.

[10] Yuri De Pra Fausto Spoto, Federico Fontana and Linmi Tao, “ Infrared vs, Ultrasonic Finger Detection on a Virtual Piano Keyboard “, Proceedings ICMC|SMC|2014, 14 – 20 september 2014, Athens, Greece, PP 654 – 658.

[11] Cheng Zhang et al., “ SoundTrak: Continuous 3D tracking of a finger using active acoustics”, PACM on Interactive, Mobile, Wearable and Ubiquitious Technologies (IMWUT), Vol v, No. N, Article, Jan YY.

[12] Sangki Yun et al. “ Strata: Fine – Grained Acoustic – based Device – Free Tracking”, MobiSys’17, June 19 – 23, 2017, ISBN 978-1-4503-4928-4/17/06.

[13] Rajalakshmi Nandakumar et al. “ FingerIO: Using Active Sonar for Fine – Grained Finger Tracking”, CHI’16, May 7 -12, 2016, ISBN 978-1-4503-3362-7/16/05.

[14] Wenguang Mao, Jian He and Lili Qiu,”CAT:High – Precision Acoustic Motion Tracking”, MobiCom’16, October 03 – 07, 2016, ISBN 978 – 1 – 4503-4226-1/16/10.

[15] Ke-Yu Chen, Shwetak Patel,Sean Keller, “ Finexus: Tracking Precise Motions of Multiple Fingertips Using Magnetic Sensing”, CHI’16, May 07 – 12, 2016, San Jose, ISBN 978-1-4503-3362-7/16/05

[16] Stale A.Skogstad at al. “Using IR Optical Marker Based Motion Capture for Exploring Musical Interaction”, Proceedings of the 2010 Conferences on New Interfaces for Musical Expression (NIME 2010), Sydney, Australia.

[17] Xiaochi Gu et al. “ Dexmo:An inexpensive and Lightweight Mechanical Exoskeleton For Motion Capture and Force Feedback in VR”, www.unity3d.com.

[18] Laura Dipietro et al. “ A Survey of Glove – Based Systems and Their Applications”, IEEE Transactions on Systems, Man and Cybernetics – Part C : Applications and Reviews, Vol. – 38, No.4, July – 2008.

[19] Joao Lourenco, “ Lierature Review: Glove Based input and Three Dimensional Vision Based Interaction “, Rhodes University, B.Sc.[Honours],29th October 2010.

[20] Eider C.P.Silva et al. “ Sensor Data Fusion for Full Arm Tracking using Myo Armband and Leap Motion”, SBC- Proceedings of SBGames 2015|ISSN : 2179 – 2259.

[21] Xu Zhang et al. “ A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors”, IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans, Vol – 41, No.6, November 2011.

[22] Dom Brown et al. “ Leimu: Gloveless Music Interaction using a Wrist Mounted Leap Motion”, NIME’16, July 11-15, 2016, Griffith University, Brisbane, Australia.

[23] Xiaoyan Wang et al. “ Hidden-Markov-Models-Based Dynamic Hand Gesture Recognition”, Research Article, Mathematical Problems in Engineeering, Vol. 2012, Article ID 986134.

[24] Mario Ganzeboom, “ How hand gestures are recognized using a Dataglove”, Human Media Interaction MSc, University of Twente, The Netherlands.

[25] Deepali N. Kakade, Prof. Dr.J.S.Chitode,”Dynamic Hand Gesture Recognition: A literature Review”, IJERT, ISSN:2278 – 0181, Vol.1, Issue 9, Nov-2012.