International Journal of Research in Advent Technology, Vol.6, No.6, June 2018
E-ISSN: 2321-9637
Available online at www.ijrat.org

A Survey on Hand Gesture Recognition Techniques, Methods and Tools

Nidhibahen Patel[1], Dr. Selena (Jing) He[2]
Masters of Computer Science, Kennesaw State University, Marietta, GA, USA[1]
Department of Computer Science, Kennesaw State University, Marietta, GA,USA[2]
Email: npate303@gmail.com[1], she4@kennesaw.edu[2]

Abstract- Hand gesture recognition is used enormously in the recent years for Human Computer Interaction (HCI). It's an efficient way of interacting with machines make it more popular and applicable for many purposes. The hands, arms, body and face are used for gesture recognition to perceive critical demeanors of movement by a human. Gesture recognition is mainly applicable for video conferencing, sign language recognition, distance learning and in some forensic identification. This paper gives a complete survey of different techniques for gesture recognition. Some tools are also involved here for recognition such as HMM, ANN, Particle Filtering and Condensation Algorithm and many more. Whole recognition process is also covered here. Comparison of various approaches related to recognition process is also included like glove-based approach, vision approach and color marker approach. The benefits and limitations of all methods for hand gesture recognition is covered efficiently.

Index Terms- recognition; gestures; vision based; glove based; HMM; ANN; tracking and segmentation

1. INTRODUCTION

For the past decade, highly skilled interpreters were used to make communication easy between deaf and hearing people. The main objective is to build correct and effective translation between spoken language and sign language. Because these kinds of interpreters are expensive, frequent use of interpreters are not sustainable. Pre-notification is to be required when someone like to use it. The key point is that to support the millions of deaf people, a significant number of well-trained interpreters will not be available.

To bridge this gap a Machine Translator (MT) [8] system which translates to sign language from spoken language and vice versa. For an example, it translates from English to American Sign Language (ASL) and Indian Sign Language (ISL). Visualization of sign language is possible by MT. This system is used for numerous applications, like telephony accessibility for deaf communities and sign language education.

Through this the Deaf can easily access a variety of public and information services.

1.1. Gestures

Gestures are one of the first forms of communication when people of different languages meet, and no one knows in which language they should express their feelings or needs. In gesture reorganization the receiver receives the gestures made by users. Gestures are used to clarify and support your words. Gestures are highly visible, and it effectively paints your ideas. In recent days many gesture recognition systems are used to communicate between the Deaf and hearing communities. Different kind of gestures are facial expressions, hand movements, eye movements etc.

Expressive, meaningful body motions are considered as Gestures. In gesture, various physical movements of the body, such as hands, face, arms, and fingers, are involved. Gestures are important to interact with the environment and convey meaningful information. Gesture recognition is utilized as a part of numerous applications [8], for example:

- To enable interaction of very young children with computers;
- Development of forensic identification techniques;
- For the hearing-impaired aids are developed;
- Virtual environments manipulation and navigation;
- Sign language reorganization;

- Distance learning/tele-teaching assistance;
- Continuously watching stress level of medical patients;
- Lie detection;
- Video conference interaction;
- Observing vehicle drivers' readiness/languor levels, and so on

Gestures are deficient and equivocal, so there is one-to-numerous calibrating from gestures to concepts and the other way around. For instance, one can utilize a thumb up to say “ok”, one can utilize thumbs up to say “done” or one can utilize thumbs up to say “good”. Dynamic and static are the two ways that the gestures can be distinguished. Predefining stances are called as static gestures. Dynamic gestures can be represented by various stages such as pre-stroke, stroke and post stroke [11]. Also, the significance of dynamic gesture can be determined by its geographic data: from which point the gesture is originating, pattern data; what kind of design/pattern/sign the gesture is having, range data ; how much range gesture occupies on space to occur and finish [23]. According to a user’s comfort, cost, accuracy, resolution, latency and range of motion, the sensing technologies may differ.

1.2. Types of Gestures

Frequently, Gestures can be dialect particular or culture-specific. More specifically gestures can be categorized as mentioned below [23]:

1. Hand and arm gestures: utilizing hand and arm signals individuals can associate with virtual condition. These sorts of signals are required close by postures acknowledgment, in communications via gestures, and stimulation applications [23].

2. Face & head gestures: Some of the illustrations are provided here: a) head rotation; b) head moving up and down c) eye rotation d) eyebrows raising e) winking the eye f) To talk by opening the mouth, g) flaring the nostrils; and h) human emotions like Amazement, illness, fear, outrage, misery, and so on [23].

3. Body gestures: Full body movement is involved in it, as in [23]:

- Tracking interaction between people.
- Dancer movement analysis and
- Human gaits recognition for medical treatment and athletic training.

1.3. Sign language

Sign languages are being used widely in international sign use of the Deaf and mentally handicapped communities. An efficient human interaction performs most important role in our daily lives. To make it more effective gesture recognition is the step towards it. Throughout the world a simple gesture performing with one hand has the same significance as either ’hi’ or ‘goodbye’. Without any knowledge about official language of that country, people travels to different parts of the world.

Using gestures and with the help of Sign languages, they able to manage and communicate. using gestures and sign language. These example shows significance of gestures. They can be considered international and used almost all over the world. Around the world, Gestures are mainly using to communicate in number of jobs. The directions to the pilots regarding the take-off on an airport runway, a predefined set of gestures is used even in any sport, referee uses gestures to tell his decisions. Figure 1 shows A to Z alphabets used for Indian and American Sign Language. The Deaf people have to communicate using these alphabets.

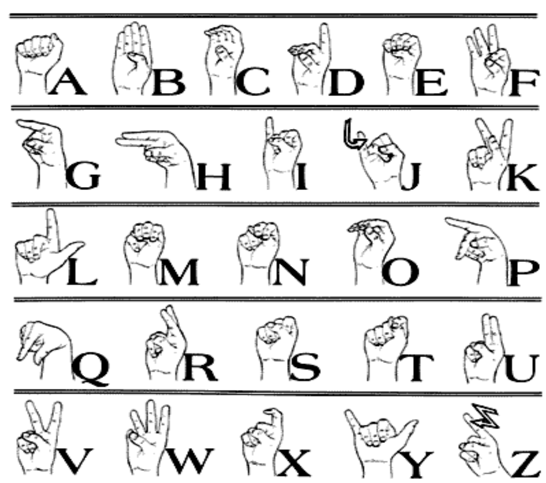

Fig.1 Hand gesture Sign Alphabets [8]

2. HAND GESTURE TECHNOLOGIES

To complete any task, first step would be to collect sufficient data. Recently, for recognition system for hand gestures, various technologies are used to acquire input data. All these technologies are categorized as following types [15]:

- Vision based approaches
- Depth based approaches
- Instrumented(data) Glove approaches and
- Colored Markers approaches.

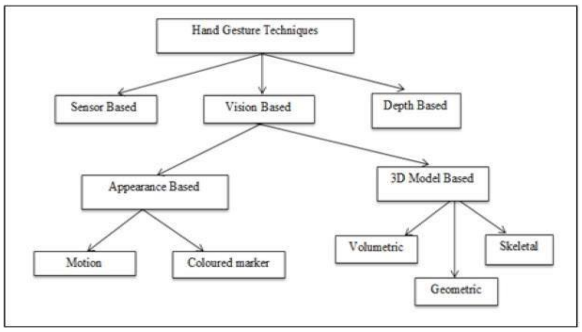

Fig 2 General Classification of Hand Gesture Recognition Techniques [17]

2.1. *Vision Based approaches*:

Pictures are taken to interface between human and PCs. Camera is the main gadget which is used to capture pictures in vision-based approach. There is no necessity for other gadgets. To capture pictures, computerized cameras are used. Acquired pictures are additionally prepared and broken down by utilizing the vision-based procedures. Cameras such as fish eye, infrared, monocular and time of flight are various divers sort ones [35]. Through vision-based techniques, recognition of the represented alphabets and numbers are becoming easier.

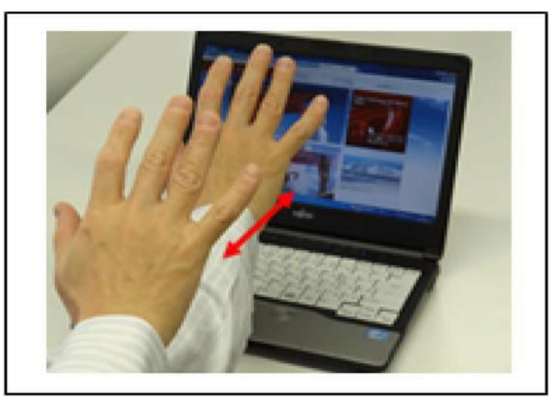

Fig. 3: Vision based Recognition [17]

This approach is simple and common to utilize so, this approach is extremely famous [16]. In this approach, there is a direct connection and interaction to human and computer devices. Conversely, numerous challenges for gestures need to be considered. Varieties of lights, different objects which are of skin color, brightening changes, and complex foundation are such difficulties in this approach. Other than that, recognition time, strength, speed and computational effectiveness are additionally challenging issues of this approach.

2.1.1. Appearance based recognition

In this approach highlights are separated from visual appearance of information hand pictures. Examination is done between effectively characterized layout and this extricated picture. The primary advantage is its compelling execution of continuous and effective examination using 3D displays. This approach can additionally recognize distinctive skin hues. The issue of impediment is overwhelmed by this approach [17]. Fundamentally two classes are there as 2D static based model and movement-based model.

2.1.2. Motion based recognition

For object recognition model based approach is used. Object motion is calculated by sequence of the images. Adboost framework is used for the learning action model. Additionally, histogram of local motion is framed for object detection. Moreover, object description that is modelling of motion and motif recognition is the necessity for the gesture recognition and that has very high complexity [17].

2.2. Instrumented Glove approaches:

For capturing hand position and motion, sensor devices are used in the instrumented data glove approach. Sensors such as vision based, multi touch screen and mount bases are utilized. Using this approach, the fingers and palm location, their orientation, configuration are calculated precisely in this approach [12]. Reaction speed is fast, and it has high accuracy rate. The main requirement of this approach [12] the user must be physically connected with the computer. This limitation is a barrier to interaction between users and computers. These devices are very costly [12] and not very efficient for virtual reality.

Fig. 4: Data Glove [17]

2.3. Colored Markers approaches:

To locate palm and fingers and to direct tracking process, human hand is worn with some colors in colored marker approach [13]. By forming hand shapes, necessary geometric features are extracted using this functionality [13]. The color glove shape has small regions with different colors. A wool groove is used when three different colors are to indicate the finger & palms in [14] where a wool grove is used. It's very easy to use and quite cheap compared to instrumented data glove [14]. Still human computer interaction is not that much natural in this technology [13]. That is the principle impediment of this method.

2.4. Recognition using skeletal based

To overcome the problem of volumetric model Skeleton model is used. Using sparse coding this model provides higher efficiency. Complex features are optimized here. Compressive sensing is used for sparse signal recovery which reduces resource consumption [17].

Fig 5: Recognition using color based [17]

Fig. 6: 3D Skeletal based Model [17]

It's mandatory to wear a lumbering gadget in glove-based gesture interface. One can need to convey a heap of such a significant number of cables associated with gadget to PC. This likewise place leaps in simplicity and expectation of client's connection. While vision based system beat this constraint [10]. There is a need to contend other impediment related issues. At the point when the client's hand is moving, quick and swoon developments of fingers are identified by tracking devices [9]. Vision based system have an incredible sense to legitimize the sort of finger movement. Tracking devices can't deal with texture and color properties while vision-based approach ready to deal with it exceptionally well. Vision based systems likewise differ [9] as per:

- Speed and latency
- Environment structure
- Specific user requirement
- The number of cameras used
- Which low level feature is used
- Which representation is used either 2-D or 3-D etc.

When 3D information is projected to 2D plane, an inherent information loss is occurring. Simple movements of a people's hand are considered as a Hand gestures. They are tiny components are used for communication which express the feelings of person. Sometimes it is used to convey the thoughts also [1]. For human computer interaction (HCI) [1] hand gestures are considered as efficient and most dominant method.

3. 3D HAND MODELLING

Hand poses and its motion are key technologies in many HCI systems [2]. 3D hand modelling and estimation of its tracking is also involved. This technology is mainly used in virtual keyboard, for recognition of sign language and robot surgery etc. According to the taxonomy discriminative approaches and generative approaches are main two approaches for hand pose estimation. Hybrid 3D modelling approach which combines both of these approaches has also been proposed. 3D hand modelling techniques are categories in following three classifications.

3.1. Discriminative 3D Hand Modelling

3D hand model with explicit Depth of Field (DoF) could not be constrained by discriminative approach. Rather than emphasizes on classifier's training, inverse mapping is performed. These mapping is between unknown hand parameters like labelling of different parts, parameter related with pose and appearance specific hand pixel features.

3.1.1 Part-Based Modelling

Human pose recognition technique is proposed by Shotton in which Kinect sensor is utilized to skeletal body parts. The body part name for every pixel inside and out area must be resolved. To make it more productive development of choice forests is done which is irregular. These forests contain 300k synthetic images per tree.

3.1.2. Efficient Part Labelling [3]

Rather than classify parts forest of shape classification is used to classify hand shapes as proposed by Keskin et al. [4]. Computation for pixel feature is same as part-based modelling. For every hand image election is done. In this voting every pixel of input depth image can participate, and they vote for label of single shape. Depth features in part-based modelling is simple [5], But it's susceptible to background changes. Yao and Fu [5] combined three highlights to overcome this limitation and make it background independent. For part labelling, following features are included into random forest (RF) classifier.

- Surface features;
- Position features using depth cameras and RGB
- Shape features;

Huge amount of training data with high quality are required in this approach. Self-occlusions and motion is not properly managed by this approach.

3.2. Generative 3D Hand Modelling

Generative approach is also likewise a model based hand following methodology [6]. 3D DoF model is explicitly trying to fit in Observed hand data [6]. To appraise the 3D hand demonstration 2D hand modelling or 3D hand modelling can be utilized.

3.2.1. Modelling of hand using 2D images

Images such as 3D hand are being analyzed and estimated using molecular 2D images by De La Gorce [6] and de La Gorce and Paragios. This model made hand by kinematic tree. Articulated kinematic tree contains 28 DoF. Hand surfaces and hand silhouette could be extracted. Synthesized silhouette matched with the observed frame to get maximum likelihood part of foreground background segmentation. This process is performed simultaneously to generate optimal hand pose. Shading and texture is used to handle self-impediment and time-shifting brightening. However, it computational cost is quite higher so it's not preferable for further appliances.

3.2.2. Single 3D Hand Modelling

As indicated by Oikonomidis offered to recoup 3D hand posture coordinating of 26-DoF hand display is done with different views of multiple cameras. Separation between the watched 3D point cloud and the rendered show was needed for the quick cost work. Quick and hearty hand following is accomplished by moreover joint angle based and stochastic enhancement strategies with without the utilization of GPU.

3.2.3. Modelling of a 3D hand using a manipulated object:

Instead of assessing promising outcome as the hand is solely seen in confinement for all the given methodologies of 3D hand tracking. At the point when a question is controlled by the hand they may not function admirably. Hamer proposed another plan in which during object manipulation, recuperation of the enunciated 3D structure of the hand is performed. Here solid impediment is taken care of by manipulating the object. 27 bones comprised by human hand model are utilized. Furthermore, they render each hand section by a chamber with a work like the skin. Similarity conceivability between the hand arrangement and the examined RGB-D information are assessed. In this both data identified with hand division and profundity were utilized.

3.2.4 3D Interactive Hand Modelling:

Detecting interactive hands is very much difficult to handle in 3D hand modelling. To track full articulation of powerfully interacting hands new parametric model is used. It contains joint kinematics of two hands for tracking. When hand interact with each other and with another object, it's very much difficult to find motion of two hands. Due to simple local optimization and multi-camera setup it generates fewer errors for pose estimation.

3.3. Hybrid 3D Hand Modelling

Combination of discriminative and generative approaches is called hybrid 3D Hand modelling. Three steps pipeline is proposed by Xu and Cheng. Hough forest regression model is used in first two steps. Discriminative depth features are dominated parts here for estimation of initial hand pose. Final step is for verification of 27 DoF hand model. This step is related with generative approach. Better accuracy and high efficiency is achieved here.

4. VARIOUS TECHNIQUES FOR HAND GESTURE RECOGNITION SYSTEM

4.1. Comparison of Pixel to Pixel

From image database, pixel to pixel comparison is done with frames in this method. Accuracy of this method is very low. However, implementation is easy. For comparing image pre- requisite image is used, which is segmented from virtual environment. For segmentation thresholding is used. Through Otsu's method pre-requisite image is segmented [18].

4.2. Edges Method

The target of this technique is to locate the highest gradient in a picture. Highest gradient is calculated by applying threshold in gradients. Through deciding correct threshold value, low magnitude gradients will be removed. Summation of derivatives of x and y directions is considered as magnitude of gradients [18].

4.3. Using Orientation Histogram

A feature vector decides the orientation of histogram. Edges of image form resultant histogram. Training phase is performed first. Conversion to grey scale image is done due to image captured by webcam. This grey scale image is used to create histogram which is used as training pattern. Following steps are performing in recognition process [19]:

- Capturing image
- Convert it in grey scale image and
- Calculating histogram

This method is quite fast and robust. Additionally, it's invariant to translation. However, this method is rotation dependent [18].

4.4. Thinning Method

In thinning method center of image is taken as a reference to find image histogram. At center of the image one window is assumed. This window is in RGB format. RGB to YCbCr conversion is required. For Cb and Cr some specific range is decided. If pixel lies within that range, then it mark with white color and others are converted as black color. Output of this process is grey scale image and this step follow by the conversion of grey image to binary image. This conversion requires thresholding and one have to choose proper threshold for Otsu's method. During this experiment noise and unwanted segments are added to it, which watered down the binary image [19]. Thus, noise removal is essentially applied before completing this whole process.

5. GESTURE RECOGNITION MODEL

Now-a-days, computer is controlled using mouse pad which is a common human interaction. Rather than this, recently touch free technologies are developed. For various technological application, gestures are very popular. Non-verbal communications which include hand movement are considered as Gesture. Mainly two types of gestures [22] are there 1) static gestures and 2) dynamic gestures. Involvement of motion does not exist in static gestures and the dynamic gestures involve body parts movement.

Uses of external devices are compulsory in sensor-based technology for caught out the movement of hand. Use of this technology is restricted. To overcome this limitation vision-based technology come in picture. Only image of gesture is required for communication in this approach [22]. Sensor based technique uses data-oriented approach, so it doesn't involve any activities related to image processing. But every algorithm follows the same step [22] like

- Detecting hand
- Extracting features and
- Recognizing gestures

A background subtraction methodology is used here to increase accuracy level [22]. The effective and simplest algorithm is selected just because complex algorithm makes the system more complex and decrease execution speed. Following key steps are required for every gesture recognition system [22]:

5.1. Tracking of Hand and Segmentation

For further processing, image frames are gathered by dividing captured input video to get dynamic gestures. Here main objective is tracking Hand and segmenting it from other objects or from background. For this mainly two approaches of gesture recognition are used. Initial one is sensor based and the second one is vision based.

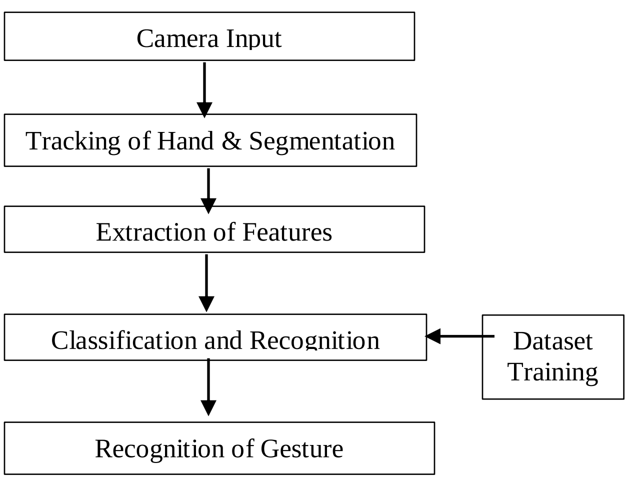

Fig 7. process of recognition

Skin detection is used in vision-based techniques. The result of this technique is varying because of illumination condition of different lights. Detecting skin HSV and YCbCr color models are very proficient. For detecting skin over here, conversion from RGB color model to HSV color model is compulsory. RGB color model is quite sensitive to light illumination. While in other techniques for hand detection color marker is used and one must wear data gloves. Due to this technique are very protective.

Dr. Johnny Chung Lee proposed a motion tracking method in which they used Wiimote to track hands [20]. Afterwards a new improved version of Wiimote was come up for image manipulation. Wilmot's camera is very sensitive to bright IR light source. So that array containing 100 LED's is used here [20]. The user who performs the gestures is in front of LED and Wiimote. Light is emitted on users' hand by LED and Wiimote camera captures that reflected gestures. Background subtraction is also used to remove jumble background.

HAAR suggested methods [20] which are used for face detection and background removal. The wrist is also discarded as an unwanted region of the hand. For that, wrist points are calculated [20] or straight skin color regions are detected. From each frame centroid of the hand is calculated which were used for hand motion tracking. To form the gesture path all these points are connected [21].

5.2. Extraction of Feature

Above step hand distinguishes hand effortlessly. Next task is to separate features from that hand to perceive motions related with that. Proficiency, accuracy and exactness of gesture recognition is determined by the quality of extracted features. Important factor over here is orientation, location and velocity.
Most prominent factor which emphasizes more is orientation. With the help of classified feature numerous algorithms are utilized to compute impeccability of orientation [21]. According to [21], feature pints are extracted using Camshift algorithm. Finger position and orientation are used to create feature points in model-based approach. They are extracted from image [21].

5.3. Recognition of Gesture

Hand gestures are recognized by creating hand models of captured images. This hand model contains a number of defects are present in capture images. The hand model is compared with a prototype model [20]. In the bag of features method only regions which contain a hand portion is considered, so size of the capture image is reduced. Key points are extracted from this hand image. SIFT algorithm [20] used that point for further tasks. Detection of the object which have same orientations and views are performed using Scale Invariance Feature Transform (SIFT). K-means clustering [20] algorithm is used to find a match between key points and the Key points of the bag of Features algorithm. In some contexts, regions with contours of hand [20] are also used for recognition of gestures.

Table 1: Segmentation and Tracking Techniques [19]

Strategies/ Methods /Approach	**Availability of Literature**	**Comments and Efficiency(Accuracy) (Exactness)**
HSI, YCbCr and morphological operations	convoluted environment and Diverse fluorescence	Gives better execution, exactness, contrasted with the HSI and YCbCr techniques
Gloved Hand Tracking	Various parts of body and convoluted environment	RGB to HSV and the preferable the Hue.
Free Hand Tracking	Various parts of body and convoluted environment	RGB to HSV, RGB to YCbCr. Parallel change
Tracking by Viola jones method	distinctive lighting conditions and plain environment	Accurate and fast learning-based method for object detection
Features such as Haar-like and adaptive skin color model for two hand segmentation	convoluted environment	Exactness(Accuracy): For the constrained hand movement, there is an accuracy of 89 to 98 percent
Segmentation using color based by using HSV and La*b* color spaces	convoluted environment along with plain environment	Better outcomes are obtained by utilizing Camshift that exploits HSV
Industrial CT image segmentation algorithm	Any kind of environment	Exactness: 95% High accuracy can recognize the crack of modern devices
Tower method for tracking	plain environment	2x-5x speedier than Camshift algorithm
Camshift algorithm for Tracking	plain environment	Quicker than mean shift algorithm
Alignment of RGB image for segmentation and on depth Using Kinect	convoluted and plain environment	Hereditary calculation to extricate the key focuses with both profundity and RGB.

Table 2: Feature Extraction Techniques [19]

Strategies/ Methods /Approach	Availability of Literature	Efficiency (Accuracy)% (Exactness)%
Finger angle characteristics	Static gesture	96.8
K-Curvature algorithm	Different background	92
Paraments like velocity, angle, motion pattern and location P2DHMMS	Complex background	98
Tracking of finger using laser tracking with HMM	Convoluted environment	95
Orientation Histogram	convoluted environment	90
Parameters such as angle, distance from endpoint of the hand	convoluted environment	92.13
A dense sampling and HOG3D descriptor	Diverse environment	97
Curvelet characteristics	Diverse environment	Up to 100

Table 3: Recognition Techniques

Strategies/ Methods /Approach	Learning hypothesis	Comments and Efficiency(Accuracy) (Exactness)
Fusing frame images	Hierarchical identification model	Accuracy: 90% Rough and Noisy hand gesture recognition
Support vector machine	Statistical learning theory	Accuracy:99.2%
Hidden Markov model	Machine learning Theory	Accuracy:93.7%
Non-linear SVM and Bag-of features (BOF)	Statistical learning Theory	Accuracy:97% Hand segmentation and tracking not required.
Superpixel Earth movers distance (SP-EMD)	Novel distance metric to overcome the partial matching problem	Accuracy:98.8% Separation estimation between two hand motions in view of shape and surface
Self-organizing Maps	Automatic learning Theory	---
Random Forest	Machine learning model	Accuracy: 94.33% Basic leadership trees for preparing tests of the preparation sets
MRS-CRF algorithm	Probabilistic models	Accuracy:90.45% Effective for dynamic hand signal acknowledgment Framework Accurate when contrasted with CRF calculation (85.3%)

6. ALGORITHMIC TECHNIQUES FOR RECOGNIZING HANDPOSTURES AND GESTURES

Information which is gathered by different approaches of gesture recognition techniques like glove based or vision based are analyzed. Check that any kind of gesture have been recognized by that or not. Many algorithms are used for that as follows [27].

6.1. Template Matching

Template matching [70] is the simplest method for recognition of hand postures. This method mainly contains two steps. In first step template is created. In posture data sets collection of data values for each gesture is considered as template.

Next step is related with matching process. Comparison is done between the recent data reading get by sensors and the data which are already stored in records. Aim is to find closely matching postures from dataset. In sort it's used for classification. Through that one can check that the data which is given by sensor are classified as a member of already stored dataset or not.

6.2. Feature Extraction Analysis

To generate higher level information related to semantic one has to collect and analyze low level information from raw data. Feature extraction means using this higher-level information for recognition of gestures. Gestures are recognized with 97% accuracy [24]. Simple hand gestures are easily recognized by this method. Additionally, complex gestures are also perfectly recognized. So, this recognition technique is robust [24] for gestures as well as for postures.

6.3. Active Shapes Model

Active shape model is also known as smart snake model. This technique is used for locating features in still images. The shape of the features which are used for tracking is considered as a contour on the image.

This contour is continuously moved nearby edges. Through this management of contour is complete that features are easily fit in that. On every frame active shape model is applied. For the next frame the initial approximation is done using the feature's position in current frame [27].

6.4. Principal Component Analysis

Principal Component Analysis is utilized to diminish the dimensionality of dataset. Many interrelated variables are also there in this technique. This technique holds on the variation of dataset [25]. Dataset reduction is the key point here. For that old data are transforming to new set of variables.

Most of the different present in original variables is reflected in first few variables. Eigen value and Eigen vectors of data sets are calculated. Transformation of original dataset is done by computing covariance matrix. When for image data this technique is used, one has to take care about orientation, scaling and position of hand in picture.

6.5. Linear Fingertip Models

Assumption is done here that the movement of fingers linear and contain very small movement related to rotations. Here input data only contains fingertips. Simple vector [26] is generated by trajectory of these fingertips. Model is used to represent this thing. Trajectories of detected fingertips are calculated by corresponding motion.

A small training set is used to model postures. This training set includes the name of gesture, code of motion and vectors which contains direction and magnitude for every fingertip. Matching is done between vector of direction and magnitude with already stored gesture training set.

The posture is recognized if this matching value lies within some threshold value. 90% accurate result is achieved by the testing system [26]. System can't perform correctly in real time. To discover the strength of the framework the stance and signal set ought to be extended.

6.6. Casual Analysis

In casual analysis vision-based recognition is used. Using the high-level knowledge, this technique extracts information from video frames. This information also contains action in the scene and its relationship with each other. For normalization of features, gesture filters are used. Fundamental knowledge is used to recognize gestures. It's additionally used to know how human interact with objects in physical world.

Information about elbow, shoulder and wrist joints position is captured by system. On image plane this capture information are gathered. Features set are extracted from these positions. Information about speed up of wrist and its deceleration, task done besides gravity, gesture's size, area between arms, area between forearms and body closeness are include in feature set [27].

Some categories of gestures which are recognized here are opening, stopping, pushing, patting, lifting and clustering. Accuracy of this method is not that much clear. Finger data are not used here is proving as disadvantage here. Further research work is requiring proving its robustness or checking whether it's applicable for any nontrivial application or not.

Table 4. Comparison of various algorithms [27]

Techniques	Usage	Advantages	Disadvantages
Template Matching	VB and GB	• Simplest • Accurate for small set of postures • Small amount of calibration	• Not for hand gestures • Does not work for large posture sets
Feature Extraction	GB	• Both postures and gestures • Layered architecture	• Computati o-nally expensive
Active Shape Models	VB	• Real time recognition • Both hand postures and gestures	• Tracks only the open hand
Principal Components	Both	• Recognize on the order of 25 to 35 postures	• More training needed
Linear Fingertip Models	VB	• Simple • Good recognition accuracy	• Not real time. • Recognizes small set of postures
Causal Analysis	VB	• Uses information about how humans interact	• Limited gestures • No orientation

*GB-Glove Based Technique.
*VB-Vision Based Technique

7. TOOLS FOR GESTURE RECOGNITION

So many tools are used for gesture recognition. PC vision (Computer vision), connectionist frameworks (Connectionist system)and Example acknowledgment (Pattern recognition), Measurable demonstrating(Statistical Modeling) and image processing approach are used in various tools. Such tools are HMM [28], PCA, more advanced particle filtering [30], Kalman filtering [29] and condensation algorithms [31]. For modelling human gestures FSM [32] is also being used effectively.

Template matching, standard pattern recognition and neural network are used to deal with static posture recognition [11]. Techniques such as HMMs, dynamic time warping [11], TDNN [11] and time compressing template are used for dynamic gesture recognition [11]. Backgrounds of some prominent apparatuses utilized for gesture recognition are discussed here.

7.1. Hidden Markov models (HMM)

Mainly double processes are performed in HMM [11] model. Initial one is to create Markov chain with limited number of states and second is an arrangement of arbitrary capacity related with one state. Corresponding to current state observation symbols is generated. According to the random functions this symbol is generated for one process. Probabilities pair is defined as follows for each transition between all the states [11].

- Probability identified with change: which make accessible the likelihood for experiencing the progress;
- Probability identified with yield: which depicts the confined likelihood of delivering a yield image from a limited letter set when given a state

Spatio-temporal information are effectively and naturally managed by HMM just because of its numerous mathematical structures. Here one only able to see the sequences of observations so its termed as "hidden" [11]. For evaluation, learning and decoding process the well- designed and well-organized algorithm developed by Baum–Welch and Viterbi [11] is also involved here.

HMM is expressed as $\lambda = (A,B,\Pi)$ where it can be explained as below:

An observations which has set of strings $O = \{O1, \ldots, OT\}$, where t varies from 1 to T

A set with N states can be given as $\{s1, \ldots, sN\}$;

For a k discrete observation, a set of symbols can be given as $\{v1, \ldots, vk\}$;

A state transition matrix $A = \{aij\}$, where aij gives the transition probability from one state si at time t to another state sj at time t + 1

$A = \{aij\} = Prob(sj \text{ at } t + 1|si \text{ at } t)$, (1)

for $1 \leq i, j \leq N$.

Probability of observation symbol matrix can be given as $B = \{bjk\}$, where bjk gives probability of generating symbol vk from state sj ;

The initial probability distribution for the states $\Pi=\{\pi j\}$, j=1, 2, . . .,N, where $\pi j = Prob(sj \text{ at } t = 1)$.

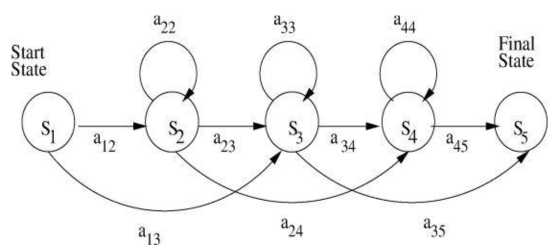

Fig. 8. Five-state left-to-right HMM for gesture recognition [11].

The ergodic model is fully connected structure of universal topology of HMM. In this structure, any state can be traversed from any state of origin. For dynamic gesture recognition, related to time variation of state index also travels left to right as above figure delineates. In this figure starting state is indicated as S1 and ending state is indicated as SN where N=5.

Here, if $j < i$, then the state transition coefficients aij = 0. The Viterbi [11] algorithm is additionally utilized for assessing an arrangement of HMMs. Deciphering process is finished by permitting just the greatest way at each progression as opposed to for all ways. Some key issues related to use of HMM are:

- Assessment: the model generates the probability of observed sequences. That probability is decided at this stage by using Forward-Backward algorithm;
- Training or approximation: here model is

altering for probability maximization. (Baum–Welch algorithm);

- Decipher: state sequences are recovered at this stage (Viterbi algorithm) [11]

Feng-Sheng also discuss about markov model [34]. Assumption about unobserved status of markov process is done here. It's considered as an uncomplicated network of dynamic Bayesian. HMM is mainly identified for its application related to temporal pattern recognition. This acknowledgment incorporates hand composing, motions, discourse grammatical feature labelling, melodic score and so forth. Here figure 9 demonstrates the stream of HMM. It chiefly incorporates highlight extraction, thresholding, skin shading recognition, edge location and nearby following of hand motion locale.

7.1.1. Feature Extraction

For extracting features and localizing object significant and functional details are provided by the motion of object. Assumption about input gestures are done to find information related to movement. In this assumption input gestures are considered as non-stationary. Moving object is tracked by observing the local grey-level changes. Motion detector is used to track this object when they move in spatial-time space.

7.1.2. Thresholding

After extracting the region of moving object thresholding is applied. Selected threshold value is applied on frame differences and from that we able to get moving region surrounded by complex background. To detect difference related to motion, Otsu method is not appropriate. Hence, more inclination is given to that thresholding which is more applicable with acknowledgment of skin pixels with RGB esteems.

7.1.3. Skin color detection

To discover the skin shading districts, the condition R>G>B is utilized. Ample ranges of color like red, brown, orange and pink color are included in skin color. To acquire the dynamic scope of skin pixels, the preparation model of skin areas with the band esteems for R, G and B pixels are utilized. Comparison of pre-determined sample color is dome to find the region which contains skin color. If we find similarity, then that region is labelled as skin area.

7.1.4. Edge detection

To isolate arm area from hand locale edge detection algorithm is enforced. To discover arm area is simple since arm locale contain fever edges then the palm district. Kirsch edge detector is used to acquire edges with different direction. To outline edge image absolute maximum value is choose for each pixel.

7.1.5. Local tracking of the hand gesture region

Foreground region information is used to discover a more accurate hand region. Motion information is used to find hand position. Additionally, skin color and edge information is also used. Occasionally, the hand position being built up by edge, movement, and skin are not at the focal point of genuine hand locale. To conquer it hand traverse is presented here. It incorporates two phases. One is identified with movement data and another is identified with frontal area data.
Following steps are included in local tracking.

- Determination of closer view and skin shading area near the uncertain focus.
- Choosing outskirt focuses in frontal area locale.
- Finding the focal point of that perimeter focuses as new focal.

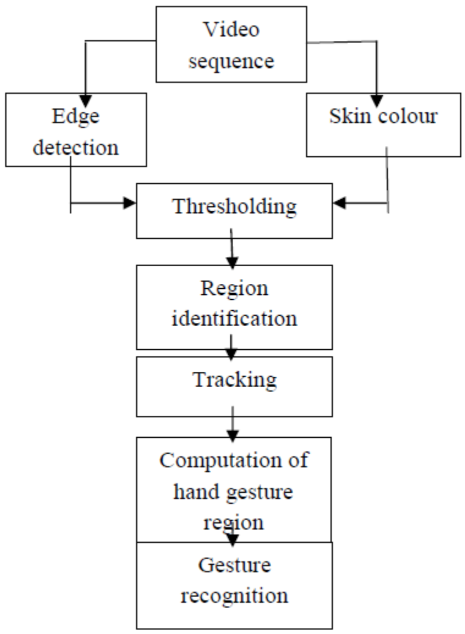

Fig 9 Flow of HMM [34]

7.2 Particle Filtering and Condensation Algorithm

In recent times particle filtering based tracking [31] and prosecution of this algorithm are become much admired for gesture recognition. From the sensor information the state of dynamic system is effectively estimated by this particle filters. Using symbol's set to signify probability density is the main idea. To track object successfully in clutter [31] this technique was developed.

State of object which is traced is described by vector Xt where t delineates time. Hence, Xt vector is used to illustrate the condition of a traced object at a time t.Xt Observation of all symbols are denoted by vectore {y1,y2,----,yt}. The observation density is calculated as P(Yt|Xt) and the posterior density is as P(Xt|Yt) . These densities are often non-Gaussian.

Using weighted sample set the approximation of probability density distribution is done. That's the main idea for finding PDD. In accordance with sampling procedure particle filter realizes recursive Byes filter. This procedure is called sequential importance sampling with re-sampling (SISR) [31].

Sample set with repetitive assessment is described by samples. As indicated by a framework demonstrate engendering test are portrayed. State of object is represented by each sample. During tracking process these samples are tracked e.g. its location and velocity. Using predictive model assumption about the next state is made. Such as the new state St +1 is assumed at time t+1 by the randomly selected sample St at time t.

In [11] the expansion of condensation algorithm is described to automatically switch between various prediction motion models. Multiple models are used for different type of motion prediction of the objects. Performance of the tracker is drastically improving because of the usage of such models. The observed motion at given time t is accurately represented by this. This property is very useful and significant for recognition of gestures. In the event that each difficult model symbolizes a solitary gesture, at that point the most expected framework conjectures the gesture which is getting experimented [11].

7.3. FSM Approach

In a spatio–temporal configuration space an ordered sequence of states is modelled as gestures in FSM approach. As per applications, the quantity of states in FSM are changed. From un- segmented sample trajectory is recognized as gesture. With the Incessant sensor information stream, trajectories group is created. Points cluster in 2D space are used to signify the trajectories of gestures.

From partial set of dynamic hand gestures, the chronological signature of hand motion is extracted. For this motion energy perception is used. This is afterwards examined and understands by a deterministic FSM. For temporal signature determination the comparative change of direction is measured. This motion is such as slowly or quickly. FSM is redefined for accomplishing additional cross-culture gestures. This redefinition is done in accordance with applicable rules of society.

7.4. YUV color space and camshift algorithm

Hand gestures are recognized by this method. For that following five steps [33] are performed.

- First, video stream of hand gestures is recorded by a digital camera. All the frames are taken into deliberation.

- Skin color-based segmentation is achieved by using YUV color space. Intensity and chrominance are isolated. For that the YUV color system is used. Y signifies intensity and UV indicates chrominance part.
- To separate hands, CAMSHIFT algorithm is utilized. Segmentation of hand from body is needed as hand is the largest connected region.
- Then location of hand centroid of is evaluated in each frame. For that zeroth and first moments are calculated, and that information is used to calculate other centroid.
- Presently the different centroid indicates are joined shape a direction. This direction plots the method for the hand rotations. Subsequently, tracing of hand methodology is unfaltering.

7.5. Using Time Flight Camera

As a part of this approach, X and y- projections of the photo are utilized. Furthermore, for gesture classification, [27] discretionary profundity features are likewise utilized. 3-D time- of-flight (TOF) sensor is utilized by this structure. Advantage of utilizing TOF sensor is an ease of hand fragmentation. The motions used in technique exhibit a pervasive division conceivable alongside the two picture tomahawks. That is the reason, the projection of the hand onto the x-and the y-focus point is utilized as highlights for the portrayal [27].

7.6. Naïve Bayes' Classifier

For recognizing static gestures Naive Bayes' Classifier is an effectual and swift method. This classification is done based on geometric invariants. After segmentation process from image data these can acquire. As a result, distinct from numerous other acknowledgment strategies, this technique is independent from skin color. Here extraction of gesture from each casing of video is done along with static background [35].

To start with this technique, segmentation and labelling of object with interest are done. In addition, geometric invariants are also extracted. This process is followed by classification. For classification K-nearest neighbor algorithm aided with distance weighting algorithm (KNNDW) is used. It offers appropriate information for a nearby weighted Naïve Bayes' classifier. Input of classifier is invariant of each region of intrigue. Type of gesture is output of this classifier. Lastly, to situate the precise properties of the gesture that are required for preparing in the framework [35].

7.7. Artificial Neural Networks (ANN)

Many researchers already worked on gesture recognition by artificial neural network. ANN has been utilized as a classier by most of the researchers whilst ANN also has been used for to extricate shape of hand by other researchers. [35]. Hand tracing and gesture recognition framework is presented by Tin H. [35] to perceive Myanmar Alphabet Language (MAL).

For finding edges of input image adobe Photoshop filter is used. For getting image feature vector histogram of local orientation is calculated. These vectors are considered contributing to the managed neural networks system. Two repetitive neural network models are used for recognizing Arabic Sign Language (ArSL) [35]. Partial and fully recurrent neural networks are used separately. For input image data a colored glove model is used and HSI color model is used for segmentation.

Six color layers are attainment during segmentation. From that one is for wrist and other five are for fingertips. To represent single image 30 highlights are extricated and assembled. To represent angle between the fingertips and wrist and between them 15 elements are used. Moreover, fifteen components are utilized to speak to separations amongst fingertips; and amongst fingertips and the wrist [35].

This information incorporates vector is the commitment to both neural frameworks systems. As setting up the set 900 shaded pictures were used and for system testing 300 tinted pictures were utilized. Totally dreary neural framework structure (with affirmation rate 95.11%) better than the Elman neural framework (with 89.67% affirmation rate) [35] were exhibited by the results.

To perceive static hand motion Self- Growing and Self-Organized Neural Gas (SGONG) Network [35] are utilized. Moreover, to recognize hand area, YCbCr shading space is utilized. In the wake of applying YCbCr color space, to recognize skin shading, threshold value is utilized to.

Vying Hebbian learning algorithm is incorporated in SGONG which generates training process. Only two neurons are contributing at starting position of learning process. Process continues growing until full construction of grid of neurons is build. Then full hand object is covered through which the shape of the hand will capture. Three geometric features are extracted from the resultant hand. Distance from the palm center and two hand slop angles were determined. To conclude the number of raised fingers all these features are used.

Gaussian distribution model is used to recognize fingertip. In that model fingers are classified into five classes and for each class features are calculated. Through this system we get recognition rate as 90.45% [35] with 31 predefined gestures. 1.5 seconds are the processing time for that. Shweta also introduced the system for gesture recognition using Neural Networks. Slow rate samples are taken here between 15-25 frames per seconds. For capturing input image at slower rate web-cams are used. On this input captured image some pre-processing is done.

Using MATLAB input image is converted into an arrangement of (x,y) facilitates. Neural classifier will take this sequence as an input and neural network will segregate the gesture to which class it belongs among various classes. These classes are predefined classes of system. Sidney and Geoffiey [35] introduced Glove-Talk introduced a system through which hand gestures can be plotted in speech synthesizer. This framework is called as Glove-talk that also runs through neural network. Through adaptive interface, gestures are translated to speech. It's also an important class of neural network application.

7.8. Histogram Based Feature

Many researchers work on histogram-based technique. In this approach feature vector is used based on orientation histogram. Orientation histogram has been implemented in gesture recognition firstly by William F. and Michal R [35]. Then, they used orientation histogram for recognition method which follows pattern recognition. Black and white input video was used for digitized input image. For computing the histogram of local orientation of each image some transformations were made on the image. To blur the histogram filter is applied on that and plot them on polar coordinates. Mainly two stages are there in that system: training stage and running stage. Diverse input gestures for various training sets are stored with their histograms in training stage. Running stage forms feature vector for new image which is an input image to PC. Orientation histogram mapped to entire picture set in training phase will now be compared to feature vector of incoming images with the assistance of Euclidean distance metric. Out of all this comparison the minimum error histogram will be selected. Per frame total processing time is 100 msec [35].Hanning Z has presented Hand gesture recognition system which is relying on local orientation histogram feature distribution model [35]. To discover a veil for the hand district skin color-based segmentation algorithm were used. Here input image is converted into HIS color space from RGB image. H part of HSI image is map with a nearest ratio image L. Using threshold value, the hand region is segmented. In local orientation histogram 128 elements were used. Histogram features with local orientation is implemented with inclusion of picture direction to the sub-window. For the compressed features demonstration, to augment local orientation histogram, K-means clustering has been reinforced [35]. Euclidean distance is used in recognition stage to calculate the precise coordinating score between the input picture and. Stored stances. For finding approximate nearest neighbors and to reduce computational cost Locality Sensitive Hashing (LSH) is used for image retrieval.

Rotation invariant static-gesture recognition approach is introduced by Wysocki [35]. In this approach they used boundary histograms. First filters are used to detect skin color. Erosion and dilation are performed as pre-processing operation. To find groups in the image, clustering is performed. Circumference of each cluster is obtained utilizing a common contour-tracing algorithm. Image is divided into grids and boundary size is normalized. Camera and hand's unique framework distance is obtained by normalization. Uniform foundation was connected, and the fringe is symbolized as harmony's size chain. The picture was isolated into N number of areas. The locales were partitioned in a spiral frame as per a particular point.

The limit harmony's size of histogram was computed. Thusly, a consecutive chain of histograms was spoken to as an element vector. For characterization, Multilayer Perceptron (MLP)

Neural Networks and Dynamic Programming (DP) coordinating [35] were utilized. From American Sign Language 26 static stances were taken and for each stance, 40 pictures were taken. Out of this 40, 20 pictures are utilized for preparing and 20 pictures are utilized for a test. Different quantities of histograms were utilized with various histogram resolutions. That range fluctuates from 8 to 36 expanding by two.

Language 26 static postures were taken and for every posture 40 pictures were taken. Out of this 40, 20 pictures are used for training and 20 pictures are used for test. Various numbers of histograms were used with different histogram resolutions. That range is varying from 8 to 36 increasing by two.

7.9. Fuzzy Clustering Algorithm

Clustering is defined as to partition the given sets of sample data into subsets or clusters. These clusters are formed based on some measures between grouped elements. The example with same qualities is gathered together to shape a cluster. Clustering has the ability to group complicated data in to regularly clusters. Because of this ability clustering is widely spread. In fuzzy clustering, fuzzy method is used to partition the sample data into groups. In that even the data pattern is single then it might belong to different data groups. This is the distinct feature of fuzzy clustering from other clustering algorithms.

To recognize hand gestures in remote mobile, Xingyan [35] introduced new fuzzy c-means clustering algorithm. To capture input raw images a camera was used. Input RGB pictures are changed over into HSV color model to proceed execution on HSV color picture. To remove noise and unwanted object some pre-processing is required. After that thresholding was used for segmenting shape of hand from the image. Here as feature vector 13 elements were used, from that the first one is used to find aspect ratio of the hand's bounding box, and the rest 12 parameters are used to represent grid cell of the image. Each cell symbolizes the mean grey level in the 3 by 4 blocks partition of the image. Normal brilliance estimations of those pixels are considered as the mean estimation of every cell. To classify gestures FCM algorithm is used. Various surroundings like complex background and invariant lighting conditions are used in the system. To create training set mainly 6 hand gestures are used. 20 samples are taken for each gesture in the vocabulary. Here we get recognition accuracy as 85.83% [35].

8. ALGORITHMS FOR HAND GESTURE RECOGNITION

For technical computational MATLAB is considered as a very productive and highly efficient language. Joining of calculation, perception, and writing computer programs is effectively done [36] in that. It's considered as easy to use as issues and elucidation are conveyed in perceptible mathematical representation. It's one kind of tool box which provides the facility to learn and apply specialize technology. Main two algorithms are implemented in MATLAB for gesture recognition. First one is Edge detection based and the second one is Skin detection algorithm.

8.1. Edge Detection

Edge detection algorithm performs following steps [36].

- Capturing picture utilizing a webcam. Here front camera of the cell phone is likewise permitted to utilize.
- This captured image is then converted into frames.
- Histogram equalization is used for image pre- processing.
- Edge is detected from hand through several algorithms. Here Canny Edge detector is used for that.
- Dilation is performed on twofold picture to amplify the edges of locales of frontal area pixels. This is done to acquire a constant edge.
- The object which is encased by the edges is then filled.
- Linear array is used to store the boundary of object which we got from above process.
- Operation related to Vectorization is performed on each of the pixel which locates at the boundary.
- Then fingertips are detected.
- To determine the motion, tracing of the fingertips is performed on successive casings.
- Based on that motion gestures are identify.
- Input stream is embedded into the ordinary incoming way of the processing gadget.

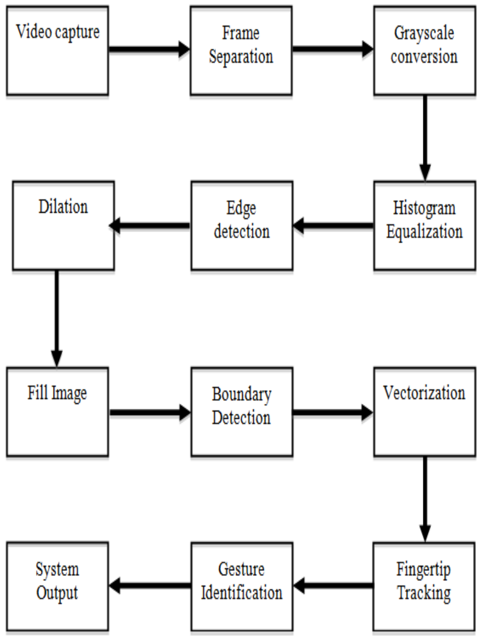

Fig. 10 Edge Detection Approach.

In sort captured input images by webcam are separated into frames and then grey-scale conversion is done. Using Histogram Equalization method, contrast is improved. Canny edge detection is reinforced to discover the edges and then dilation is used to fill up the broken edges. Using bw boundaries function of MATLAB boundary pixels are detected, and linear array is used to store them. Vectorization technique is used to detect fingertips. Relative movement of fingertips are collected through that gestures are recognized.

8.2. Skin Detection

For Skin detection following steps is performed [36].

- Capturing picture utilizing a webcam. Here front camera of the cell phone is likewise permitted to utilize.
- This captured image is then converted into frames.
- RGB image is converted into HIS color space.
- Hue and immersion estimations of different conceivable skin tones are utilized for Skin Detection.
- Linear array is used to store the boundary of object which we got from above process.
- Operation related to Vectorization is performed on each of the pixel which locates at the boundary.
- Then fingertips are detected.
- To decide the movement, tracing of the fingertips is performed on back to back edges.
- Based on that motion gestures are identify.
- Input stream is embedded into the ordinary incoming way of the processing gadget.

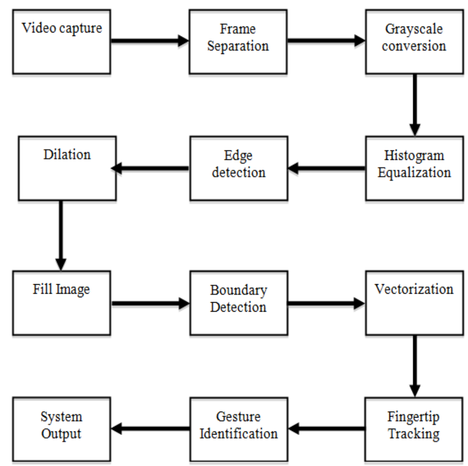

Fig 11 Skin Detection Algorithm [36].

9. CONCLUSION

Human machine interaction process works very smoothly just because of gesture recognition. For sign language recognition, for controlling robots and for graphical editor control gesture recognition is very important. Main challenge of this recognition process is to collect input raw data. For that two different techniques are used which we already discuss as glove based and vision-based approach. Both approaches have their own advantages and limitation. Recognition process includes tracking hand and segmenting it from background, then features are extracted using various techniques and finally, after applying classification gestures are recognized. Many methods and tools are used for hand recognition. The Hand gesture recognition models such as Hidden Markov model, YUV color space model, 3D model and Appearance model, YUV color space model, 3D model and Hidden Markov Model FSM approach and ANN distinguish the input and process them for acknowledgment.

REFERENCES

[1] Hong Cheng, Zhongjun Dai, and Zicheng Liu, "Image-to-class dynamic time warping for 3D hand gesture recognition," IEEE International Conference on Multimedia and Expo (ICME), July 2013, pages 1–6.

[2] Yuan Yao, Yun Fu "Real-time hand pose estimation from RGB-D sensor," IEEE International Conference on Multimedia and Expo (ICME), July 2012, pages 705– 710.

[3] Cem Keskin, Furkan Kıraç, Yunus Emre Kara and Lale Akarun, "Hand pose estimation and hand shape classification using multi- layered randomized decision forests," European Conference on Computer Vision, 2012, pages 852–863.

[4] Cem Keskin, Furkan Kıraç, Yunus Emre Kara, and Lane Akarun, "Real time hand pose estimation using depth sensors," IEEE International Conference on Computer Vision Workshops (ICCV), November 2011, pages 1228–1234.

[5] A. Erol, G. Bebis, M. Nicolescu, R. D. Boyle, and X. Twombly, "Vision-based hand pose estimation: A review," Computer Vision Image Understanding, volume 108, Issue 1–2, pages 52–73, October-November 2007.

[6] Y. Wu, J. Y. Lin, and T. S. Huang, "Capturing natural hand articulation," 8th IEEE International Conference on Computer Vision Workshops (ICCV), July 2001, pages 426– 432.

[7] Hong Cheng, Lu Yang, and Zicheng Liu, "Survey on 3D Hand Gesture Recognition", IEEE Transactions on circuits and systems for video technology, volume 26, Issue 9, September 2016, pages 1659-1673.

[8] Sagar P.More and Abdul Sattar, "Hand Gesture Recognition System Us s, NSF Science and Technology Center for Computer Graphics and Scientific Visualization, USA.

[9] D. M. Gavrila, "The visual analysis of human movement: survey," Computer Vision and Image Understanding, volume 73, pages 82–98, 1999.

[10] J. K. Aggarwal and Q. Cai, "Human motion analysis: A review," Computer Vision and Image Understanding, volume 73, pages 428–440, 1999.

[11] Sushmita Mitra and Tinku Acharya, "Gesture Recognition: A Survey", IEEE Transactions on systems, man, and cybernetics - part c: applications and reviews, Volume 37, Issue 3, May 2007, pages 311-324.

[12] Joseph J. LaViola Jr., "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Brown University, NSF Science and Technology Center for Computer Graphics and Scientific Visualization, USA

[13] Mokhtar M. Hasan, and Pramod K. Mishra, "Hand Gesture Modeling and Recognition using Geometric Features: A Review", Canadian Journal on Image Processing and Computer Vision, 2012, Volume 3, Issue 1.

[14] Luigi Lamberti& Francesco Camastra, "Real-Time Hand Gesture Recognition Using a Color Glove", Springer 16th international conference on Image analysis and processing: Part I (ICIAP'11), 2011, pages 365-373.

[15] Noor Adnan Ibraheem and RafiqulZaman Khan, "Survey on Various Gesture Recognition Technologies and Techniques", International Journal of Computer Applications (0975 – 8887), July 2012 Volume 50 – Number 7.

[16] Orasa Patsadu, Chakarida Nukoolkit, Bunthit Watanapa, "Human Gesture Recognitions Using Kinetic Camera", 2012, 9th International Join Conference on Computer Science and Software Engineering.

[17] Harpreet Kauri and Jyoti Rani, "A Review: Study of Various Techniques of Hand Gesture Recognition", IEEE International Conference on Power Electronics, Intelligent Control and Energy Systems (ICPEICES), 2016.

[18] Kamalpreet Sharma, Naresh Kumar Garg; "Hand Gestures Recognition for Deaf and Dumb" International Journal of Computer application and Technology (s), May - 2014, pages 10-13.

[19] Jayesh S. Sonkusare, Nilkanth. B. Chopade, Ravindra Sor and Sunil.L. Tade, "A Review on Hand Gesture Recognition System", International Conference on Computing Communication Control and Automation 2015, pages 790-794.

[20] Nasser H. Dardas and Nicolas D. Georganas, Fellow, IEEE, "Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques" in IEEE Transactions on Instrumentation and Measurement, Volume 60, Issue 11, November 2011.

[21] Prashan Premaratne, Sabooh Ajaz, Malin Premaratne, "Hand gesture tracking and recognition system using Lucas–Kanade algorithms for control of consumer electronics in

Neurocomputing" 2013, pages 242–249.

[22] Vaibhavi S. Gandhi, Akshay A. Khond, Sanket N. Raut, Vaishali A. Thakur,Shabnam S. Shaikh , " A Review of Various Gesture Recognition Techniques", International Journal of Electric and Computer Science, Volume - 3 Issue -9 September 2014, pages. 8202-8206.

[23] S. Mitra, T. Acharya. "Gesture Recognition: A Survey", IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, pages 311-324, 2007.

[24] YinghuiZhou, LeiJing, JunboWang, ZixueCheng"Analysis and Selection of Features for Gesture Recognition Based on a Micro Wearable Device" Graduate School of Computer Science and Engineering, University of Aizu Wakamatsu, Japan,(IJACSA) International Journal of Advanced Computer Science and Applications, Volume 3, Issue 1, 2012.

[25] Dardas, N.H.; Petriu, E.M., "Hand gesture detection and recognition using principal component analysis", Computational Intelligence for Measurement Systems and Applications (CIMSA), 2011 IEEE International Conference on 19-21 September 2011.

[26] K. Oka, Y. Sato and H. Koike, "Real-time tracking of multiple fingertips and gesture recognition for augmented desk interface systems", IEEE International Conference on Automatic Face and Gesture Recognition, 2002.

[27] R.Pradipa, Ms S.Kavitha, "Hand Gesture Recognition – Analysis of Various Techniques, Methods and Their Algorithms ", International Journal of Innovative Research in Science, Engineering and Technology(ICIET'14), pages 2003-2010.

[28] F. Samaria and S. Young, "HMM-based architecture for face identification," Image Vision Computer, volume 12, pages 537– 543, 1994.

[29] G. Welch and G. Bishop, "An introduction to the Kalman filter," Department Computer Science, University of North Carolina, Chapel Hill, Tech. Rep. TR95041, 2000.

[30] S. Arulapalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line nonlinear/non- Gaussian Bayesian tracking," IEEE Transaction and Signal Processing, Volume 50, Issue 2, pages 174–188, February 2001.

[31] A. Doucet, N. de Freitas, and N. Gordon, Eds., Sequential Monte Carlo in Practice. New York: Springer-Verlag, 2001.

[32] P. Hong, M. Turk, and T. S. Huang, "Gesture modeling and recognition using finite state machines," in Proc. 4th IEEE International Conference Automatic Face Gesture Recognition, Grenoble, France, March 2000, pages 410–415.

[33] Aarti Malik and Ruchika, "Gesture Technology: A Review", Department of Electronics and Communication Engineering International Journal of Electronics and Computer Science Engineering, ISSN- 2277-1956.

[34] R.Suriya, Dr.V.Vijayachamundeeswari, "A Survey on Hand Gesture Recognition for Simple Mouse Control ", International Conference on Information Communication and Embedded System 2014.

[35] Noor Adnan Ibraheem and RafiqulZaman Khan, "Survey on Various Gesture Recognition Technologies and Techniques" International Journal of Computer Applications, July -2012 pages 39-44.

[36] Shweta. K. Yewale and Pankaj. K. Bharne, "Hand Gesture Recognition Using Different Algorithms Based on Artificial Neural Network", IEEE International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016, pages 671-675.