Scientific Journal of Impact Factor (SJIF): 3.134
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 3, Issue 1, January -2016

Survey paper on Hand Gesture Recognition

Ms. Hardi Desai [1], Mr. Yask Patel[2]

[1]Department of Computer Engineering, PIET Vadodara,
[2]Department of Information Technology, PIET Vadodara,

Abstract *- Now–a-days, computer become more extensive in society, facilitating natural human-computer interaction (HCI) will have a positive impact on their use. Till now human can interact with computer through input devices like mouse and keyboard or technologies like Graphical User Interface, Command Line Interface, touch user interface, motion tracking interface. Gesture recognition system could be used for improving human-machine interaction, this will allow a human user to communicate non- verbally using different body parts like finger, hand, face, eye, arm, palm etc. Hand gesture recognition approach works on the three main phases i.e. object detection, tracking of object, and gesture recognition. Each phases have different techniques which has its own benefits and limitations. In this survey paper different hand gesture recognition methods and issues are explained.*

Keywords *- Hand Gesture recognition, Human Computer Interaction, Segmentation techniques, Classification techniques.*

I. INTRODUCTION

Body language is an important way of interaction among humans, adding emphasis to voice messages or even being a complete message by itself [1]. Gesture is a form of non-verbal communication using various body parts, mostly hand and face. Gesture recognition systems could be used for improving human-machine interaction. This type of interfaces would allow human user to control remotely through hand postures a wide variety of devices. Gestures are used widely for different applications like human-robot interaction, sign language recognition, interactive games, vision-based augmented reality, making passengers aware about the safety features during driving etc. For communication at the people at a visible, but not audible distance, gesture is the only method [2].

Hand gesture can be classified in two types i.e. Static and Dynamic gestures. Static gesture is also called posture which refers only a single image corresponding to a single command. Stop sign is an example of static gesture. Static gesture is simple and need less computational power. Sequence of postures are called as gesture which is called as dynamic gesture [2]. Dynamic gesture is intended to change over a period of time and it is complex. A waving hand means goodbye is an example of dynamic gesture. Hand gestures are the meaningful body motions that are movements of hands, arms or fingers. Hand gesture identification ranges from the static gesture with complex background or dynamic gestures that express the human feeling and communicate with computer or humans [14].

The rest of this paper is organized as follows. Section II presents basic architecture of gesture recognition. Section III presents approaches of gesture recognition. Segmentation techniques are discussed in section IV. Section V presents different classification techniques of gesture recognition. Comparison of classification techniques are given in section VI. Section VII present different issues related to hand gesture recognition. Finally, conclusion is presented in section VIII.

II. ARCHITECTURE OF GESTURE RECOGNITION

Figure 1 shows the basic architecture of gesture recognition. It has four main phases to find out the gesture. They are image acquisition. Image pre-processing and segmentation, feature extraction and finally classification and recognition.

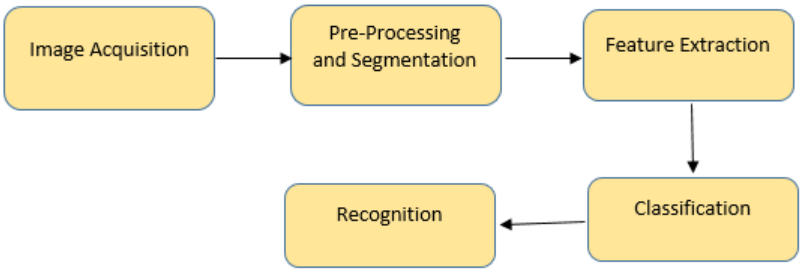

Figure 1. Basic Architecture of Gesture Recognition

Image acquisition is the primary step in gesture recognition system. A set of image frames are captured by a cheaper camera [14]. Image pre-processing is the process of noise removal, edge enhancement and normalization of image which improve the quality of image. In segmentation, image is extracted from foreground to background.Thresholding can be

used for differenciate object from its unnecessary background in image segmentation process. Features are an important elements for hand gesture recognition. Large number of features, such as, shape, orientation, textures, contour, motion, distance, centre of gravity etc. can be used for extracting feature of object [2]. There are three approaches available for extraction of features i.e. Model based, View based and Low-level-features based approach. Once the appropriate features have been extracted from the images and suitable data set have been selected, the gestures can be recognized using standard machine learning techniques or a special-purpose classifiers.For recognition of gesture, classification methods can be used such as Hidden markov model (HMM), Fuzzy C-means clustering algorithm, K-nearest neighbor (KNN), Principal component analysis (PCA) and so on.

III. GESTURE RECOGNITION APPROACHES

In these section, different approaches of hand gesture recognition can be describing with its benefits and drawbacks. There are mainly four approaches which can be given below:

A. Colored glove approach

Colored gloves are also known as marked gloves. For recognition of hand, tracking and locating the hand, palm, and fingers, user has to worn the color glove [4]. At the recognition time, user has to set some threshold value of that color because of that user can easily recognize the hand gesture. Due to shadow and complex background, hand gesture can not recognize all the colors.

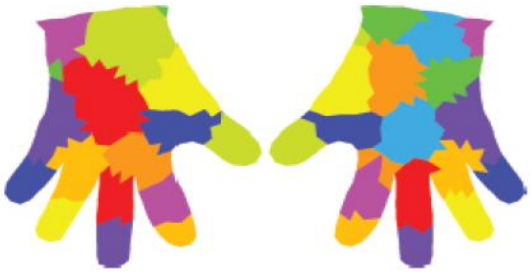

Figure 2. Glove design consist of patches [4]

Advantages

I. Colored gloves are inexpensive and no sensors are embedded in or outside the gloves.
II. It is robust method for hand gesture recognition.

Disadvantages

To recognize the hand, every time user has to wear the glove. So, it is inconvenient.

B. Data glove approach

Data glove are also known as Instrumented glove. For recognizing and tracking of hand, user has to worn the data gloves. These gloves consist of sensor device to capture hand position and motion. Data gloves can easily provide exact coordinates of palm and finger location, orientation and hand configuration [4]. It reduces the natural level of interaction with the computer.

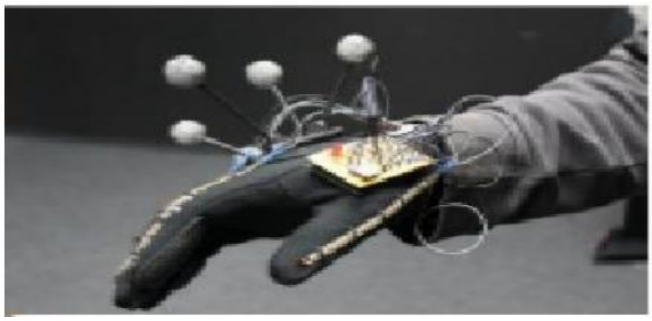

Figure 3. Data glove [4]

Disadvantage

It is expensive because of sensor node.

C. Vision based approach:

Vision based approach needs only high quality camera to capture the images. It does not require any external device or hardware. It deals with texture and color properties. The image gives natural interaction between human and computer. This approach is simple but raised many challenges such as complex background, number of camera used by those techniques can be different, speed and latency, lighting condition and skin color objects with the hand object, system requirements such as velocity, recognition time, robustness and computational efficiency.

Advantage
It is robust method for hand gesture recognition.

Disadvantage
Easily affected by complex background.

Figure 4. Vision based [4]

D. Depth camera approach:
Depth camera is also called Kinect. Kinect is nothing but RGB-Depth sensor introduced by Microsoft for human computer interaction. Kinect is used in many applications like video games, virtual reality [4]. In RGB camera, user only recognized the gesture where as in Kinect, user can recognize the depth of gesture.

Figure 5. Using Kinect recognized hand gesture (Measure the depth)

Advantage
Robust than any other approach because it can measure the depth.

Disadvantage
Kinect device costs more than any other devices.

IV. SEGMENTATION METHODS

Segmentation methods can be classified into four kinds, and the selection of proper method depends on the specific application and working environments. These methods are pixel, edge, region, and model based segmentation.

A. Pixel-Based segmentation

Pixel-based or point based segmentation is simple method which depends on gray-level values to segment image pixel. It is also called as thresholding [7]. Bayesian classifier, Piece wise, Gaussian models, thresholding, Histogram thresholding, NNs, Gas, HMMs all are pixel based segmentation techniques.

B. Edge-Based segmentation

Edge-Based or boundary-based segmentation method commonly refer to segmenting an image based on the edges, by searching for edge pixels and connect them to form image contour. For applying such methods two approaches are found that is manually by using mouse to draw lines that represent image boundaries among regions and automatically by implementing some edge detection filters [7]. Canny edge detector, Laplacian of Gaussian edge filter, Hough transform, Prewitt's filter all are techniques of edge-based segmentation.

Figure 6. Edge-Based segmentation

C. Region-Based segmentation

These segmentation partition an image into regions or groups of similar pixel depending on some properties. Its principle depends on an idea that neighboring pixels within the same region have same value [7]. Region growing, Split and merge, Region splitting, Region merging, NNs all are Region-Based segmentation techniques.

Figure 7. Region-Based segmentation

D. Model-Based segmentation

These methods require more information about image such as objects geometrical shape or repetitive form geometry of a specific ROI. These include Gaussian model and Decoded system [7].

V. GESTURE CLASSIFICATION TECHNIQUES

The task of assigning a feature vector or a set of features to predefined classes in order to recognize any gesture is called as classification. After analysis and modelling of input hand image is done, to recognize the gesture, classification methods can be used. There are number of classification methods that can be used for recognizing gesture from which some of them can be describe below.

A. Support Vector Machine (SVM)

SVM is a non-linear classifier which is reported as producing superior classification results compared to other methods. The idea is to non-linearly map the input data to some high dimensional space, where the data can be linearly separated, thus providing great classification (or regression) performance [8]. One of the bottleneck of SVM is the large number of support vectors used from the training set to perform classification (regression) tasks. To performing linear classification, SVMs can efficiently perform non-linear classification using what is called the kernel trick, implicitly mapping their inputs into high-dimensional feature space.

B. Hidden Markow Model (HMM)

A HMM is defined as a set of states of which one state is the initial state, a set of output symbols, and a set of state transitions [6] [8]. Each state transition is represented by the state from which the transition starts, the state to which transition moves, the output symbol generated, and the probability that the transition [8]. The state transitions represent the probability that a certain hand position transitions into another: the corresponding output symbol represents a specific

posture and sequence of output symbols represent a hand gesture. The HMM with the highest forward probability determines the user's most likely gesture.

C. Dynamic time warping (DTW)

It has long been used to find the optimal alignment of the signals. The DTW algorithm calculates the difference between each possible pair of points out of two signals in term of their associated feature values [6] [8]. By using these distances they calculate a cumulative distance matrix and find the least expensive path through this matrix. This path represents the ideal warp- the synchronization of the two signals which causes the feature distance between their synchronized points to be minimized. The signals are normalized and smoothed before the distances between points are calculated [6]. DTW has been used in various fields, such as speech recognition, data mining, and movement recognition. DTW mainly focused on speeding up the algorithm whose complexity is quadratic in the length of the series.

D. Principal Component Analysis (PCA)

PCA is a statistical procedure that uses an orthogonal transformation to convert a set of observation of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The limitation of this algorithm is that it depends on the scaling of the variables.

E. K-Means

The k-means problem is to determine k points called centers so as to minimize the clustering error, defined as the sum of the distances of all data points to their respective cluster centers [6]. This classification finds statistically similar groups in multi-spectral space. The most commonly used k-means algorithm are Lloyd's k-means and Mac-Queens k-means algorithm.

F. Mean Shift Clustering

It is nonparametric clustering technique which does not require prior knowledge of the number of clusters, and it is not constrain the shape of the clusters. The main idea behind means shift is to treat the points in the d-dimensional feature space as an empirical probability density function where dense regions in the feature space correspond to the local maxima or modes of the underlying distribution [6]. The data points associated with the same stationary point are considered the members of the same cluster.

G. K-Nearest Neighbor (KNN)

It is a method for classifying objects based on closest training examples in the feature space. In this algorithm, to identify neighbors, the objects are represented by position vectors in a multidimensional feature space. The KNN algorithm is very sensitive to the local structure of the data [6].

VI. COMPARISION BETWEEN GESTURE CLASSIFICATION TECHNIQUES

Table 1: Comparison of classification techniques [8]

Technique	Parameter	Advantage	Disadvantage
K-means	Cluster center locations	Computationally faster, produces tighter clusters	Prediction of K is difficult for fixed number of clusters.
KNN	Class of nearest neighbor	Easy to implement, Lowest complexity, carefully chosen features gives good results	Sensitive to arbitrary attributes.
Mean Shift clustering	Empirical probability density function	Spherical clusters not assumed works on a single parameter (window size), Robust to outliers.	Ambiguity in optimal parameter selection, computationally expensive, does not scale well with dimension of feature space.
SVM	High-dimensional feature space	Higher prediction accuracy, Robust for errors in training examples, Fast evaluation of the learned target function.	Long training time, Complex learned function (weight), Domain knowledge incorporation not easy.
HMM	Pixels in a vision-based input	Easily extended to deal with strong TC tasks, Embedded re-estimation is possible, easy to understand.	Large assumptions about the data, Huge number of parameters needs to be set, Large training data can be required.
DTW	Shape characteristics	Reliable time alignment, Robust to noise, Easy to implement.	Complexity is quadratic, Distance metric needs to be defined.

Finite State Machine	Feature vector such as trajectories	Easy to implement, Efficient predictability, Low processor overhead.	Not robust, Ridged conditions of implementation.

VII. ISSUES IN HAND GESTURE RECOGNITION

A. Change in illumination

When there is a change in the illumination condition, the system fails to recognize properly. Even if there is variation in lighting condition between the training dataset and inputs, some system fails to recognize.

B. Difficult background

Most of the work can be done in uniform background for recognizing hand. But in real-time gesture recognition, uniform background is not desirable or available.

C. Rotation or orientation limitation

An HGR system fails to recognize if the hand is oriented in a different angle with reverence to that in database.

D. Scaling problem

The problem of scaling arises due to different field of applications, hand size of the users, perspective.

E. Translation problem

The variation of hand positions in different images also leads to erroneous representation of features.

F. Skin-like-colored objects

Sometimes objects with similar color that of human skin may be present in the environment and this leads to confusion of the recognition systems.

G. Special hardware

A number of special hardware, like Range camera, 3D depth sensor, Data gloves have been used.

VIII. CONCLUSION

Hand gesture recognition finding its application for non-verbal communication between human and computer. With the increase in applications, the gesture recognition system demands a lots of research in different directions. Hand gesture can be recognized easily, and actions performed depends on gesture movement are the primary focus of many researchers. In this paper, various approaches of hand gesture recognition can be given with its advantages and disadvantages. Also different segmentation and classification methods for gesture recognition can be discussed in this work. Comparative study of various classification techniques used for recognizing gesture are shown with its advantages and disadvantages. Issues in hand gesture recognition can also discussed.

References

[1] Narendra V. Jagtap,Prof. R. K. Somani, Prof. Pankaj Singh Parihar, "Hand Gesture Recognition System for Human-Computer Interaction with Web-Cam", in IJSRD - International Journal for Scientific Research & Development| Vol. 1, Issue 9, 2013 | ISSN (online): 2321-0613.

[2] Arpita Ray Sarkar, G. Sanyal and S. Majumder, "Hand Gesture Recognition System: A Survey", International Journal of Computer Applications (0975 – 8887) Volume 71– No.15, May 2013.

[3] Surve Pranjali, Ubale V.S, "Hand gesture recognition system: A Survey" , in International journal of Inventive Engineering and Science (IJIES), ISSN: 2319-9598, Volume-3 Issue-3, February 2015.

[4] Dr. K.S.Kinage and Samara Mutha, "Study on Hand Gesture Recognition", in International Journal of Computer Science and Mobile Computing, 2015.

[5] Surve Pranjali, Prof. Ubale V.S, "Recognition of Hand Gesture using PCA, Image Hu's Moment and SVM" , in IPASJ International journal of Electronics and communication (IIJEC), ISSN 2321-5984, Volume 3, Issue 4, April 2015.

[6] Sukhdip Singh, Yogita Bhardwaj, "Hnad Gesture Recognition Techniques: A Review" , in National conference on Innovative Trends in Computer Science Engineering (ITCSE), ISSN:2349-7688, April 2015.

[7] Mokhtar M. Hasan, Noor A. Ibraheem and Rafiqul Z. Khan, "Comparative study of Skin Color based Segmentation Techniques", International Journal of Applied Information Systems (IJAIS), Volume 5, No. 10, August 2013.

[8] Anupam Agrawal, Siddharth S. Rautaray," Vision based hand gesture recognition for human computer interaction: a survey", Springer/Science Business Media Dordrecht, 2012.

[9] Gashree R N , Stafford Michahial , Aishwarya G N , Beebi Hajira Azeez , Jayalakshmi M R ,and R Krupa Rani, "Hand Gesture Recognition using Support Vector Machine", in International Journal Of Engineering And Science (IJES), Volume 4, Issue 6, Pages PP.42-46, June – 2015, ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805.

[10] Sushmita Mitra and Tinku Acharya, "Gesture Recognition: A Survey", in IEEE Transactions on systems, VOL. 37, NO. 3, MAY 2007.

[11] E. Ohn-Bar and M. M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations," IEEE Trans. Intelligent Transportation Systems, vol. 15, no. 6, pp. 2368–2377, Dec 2014.

[12] W. Heng, A. Kl¨aser, C. Schmid, and C.-L. Liu, "Dense trajectories and motion boundary descriptors for action recognition," Intl. Journal of Computer Vision, vol. 103, no. 1, pp. 60-79, May. 2013.

[13] P. Doliotis, A. Stefan, C. McMurrough, D. Eckhard, and V. Athitsos, "Comparing gesture recognition accuracy using color and depth information," in ACM Conf. Pervasive Technologies Related to Assistive Environments, 2011.

[14] Jayesh s. sonkusare, Nilkanth B. chopade, Ravindra sor and Sunil L. tade, "A review on Hand Gesture Recognition System", in International Conference on Computing Communication Control and Automation, 2015.

