International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 21 Issue 3 – APRIL 2016.

A SURVEY ON HAND GESTURE RECOGNITION SYSTEM

Neethu P S[1], Dr. R. Suguna[2]

[1] Research Scholar, Information and communication Engineering, Anna University, Chennai.

[2]Professor & Dean, Department of Computer science & Engineering, SKR Engineering College, Chennai

ABSTRACT: ***Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text user interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse. In this paper, we provide a review on the hand gesture recognition including segmentation; feature extraction and classification methods are discussed.***

KEYWORDS: **Gesture recognition, segmentation, hand gestures, vision based system.**

I. INTRODUCTION

A gesture is a form of non-verbal communication made with a part of the body, used instead of or in combination with verbal communication. The language of gesture allows individuals to express a variety of feelings and thoughts, from contempt and hostility to approval and affection. Most people use gestures and body language in addition to words when they speak. The use of gesture as language by some ethnic groups is more common than in others, and the amount of such gesturing that is considered culturally acceptable varies from one location to the next. Gesture-based interface holds the promise of making human-machine interaction more natural and efficient. Gesture-based interaction was firstly proposed by M. W. Krueger [10] as a new form of human-computer interaction in the middle of the seventies and there has been a growing interest in it recently. As a special case of human-computer interaction, human robot interaction is imposed by several constraints the background is complex and dynamic; the lighting condition is variable; the shape of the human hand is deformable; the implementation is required to be executed in real time and the system is expected to be user and device independent.Numerous techniques on gesture-based interaction have been proposed, but hardly any published work fulfills all the requirements .R. Kjeldsen and J. Kender [14][1][2]presented a real-time gesture system which was used in place of the mouse to move and resize windows. In this system, the hand was segmented from the background using skin color and the hands

Pose was classified using a neural net. A drawback of the system is that its hand tracking has to be specifically adapted for each user. The Perseus system developed by R. E. Kahn[15][3] was used to recognize the pointing gesture. In the system, a variety of features, such as intensity, edge, motion, disparity and color has been used for gesture recognition. This system was implemented only in a restricted indoor environment. In the gesture-based human-robot interaction system of J. Triesch [9] and C.Ven Der Malsburg,[4][5] the combination of motion, color and stereo cues was used to track and locate the human hand, and the hand posture recognition was based on elastic graph matching. This system is prone to noise and sensitive to the change of the illumination.

Gestures are mostly culture-specific. Pointing gesture is probably one of the exceptions. Some gestures closely coordinated with speech. Gestures are mainly of two types. It can be static or dynamic. Static gesture means the user assumes a certain pose or configuration. Dynamic [20] [23] means stroke, post stroke etc [1].

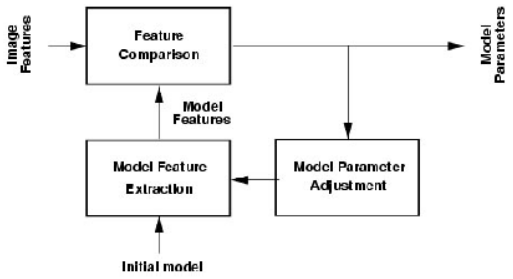

Fig 1 .Block diagram for a typical hand gesture recognition system

Hand gestures may be given as input either using data gloves, sense boards or using a bare hand.

This paper is organized as follows: section 2 provides gesture recognition technology. Section 3 deals with vision based hand gesture recognition approaches. Section 4 explains about segmentation. Feature extraction and classification are given in section 5.Conclusion is given in section 6.

II. GESTURE RECOGNITION TECHNOLOGIES

Hand gestures represent a powerful non-verbal context-dependent human communication modality. The expressiveness of hand gestures can be explored to achieve natural human-computer interactions in a smart habitat environment. Vision-based hand tracking and gesture classification, focusing on tracking the bare hand and recognizing hand gestures without the help of any markers and gloves. The existing hand gesture recognition system use either vision based [10][14] or sensor based solution.

A Vision based solutions

Vision-based[37] solutions collect human motion from one or multiple cameras [26]. Vision-based devices can handle properties such as texture and color for analyzing a gesture, while sensor cannot. A tracker [18] also needs to handle changing shapes and sizes of the gesture-generating object (that varies between individuals), other moving objects in the background, and noise.

Vision-based techniques can vary among themselves in 6 ways:
1) The number of cameras used. 2) Their speed and latency. 3) The structure of environment (restrictions such as lighting or speed of movement). 4) Any user requirements (whether user must wear anything special). 5) The low-level features used (edges, regions, silhouettes, moments, histograms). 6) Whether 2-D or 3-D representation is used. 7) Whether time is represented [31].

B Sensor or Data Glove based solutions

Sense board is a U-shaped device that fits around the hand parallel to the wrist. It contains three accelerometers to detect movement of the hand and two rotational sensors which detects wrist and elbow rotation [8]. This data is sampled at regular intervals of 25Hz and sent to a computer using Bluetooth.

The data transmitted is a raw stream of sensor outputs which has not been filtered in any way. The task of the gesture recognition software is to attempt to classify this input, and output the most probable gesture to the robot control software.

Sensor-based solution typically requires the user to wear a cumbersome device [6][37] nowadays. This hinders the ease and naturalness of the user's interaction with the computer. According to Moor's law, sensors are getting smaller and cheaper as time goes on. It will be a pervasive in future, we believe.

Researchers from UC Berkeley developed a wearable motion sensor network [7] to recognize human activities. They introduced distributed pattern recognition. Each mote (mote is a sensor network terminology means node interchangeably) can do classification locally and sends events to global classifier to recognize motion. Compare to centralized approach, it dramatically reduced the communication and energy consumption. It increases the robustness of system, user can active necessary node when needed on the fly. HMM is used to recognize gesture. They manually do the segmentation on data stream.

The interesting point is a procedure based on adding noise-distorted signal duplicates to training set is applied and it is shown to increase the recognition accuracy while decreasing user effort in training. With the help of noisy data, user can have only 2 training samples for each gesture.

Fig 2. Data Gloved based and vision based system

C Comparison and analysis

	Vision based solutions	Sensor based solutions
Cost	☺	😐
User Comfort	☺	😐
Calibration	☺	😐
Computing Power	☹	😐
Portability	☹	☺

Table 1: Comparison and analysis of existing hand gesture recognition system

III. VISION BASED HAND GESTURE RECOGNITION APPROACHES

A Appearance Based Approaches

In this approach [35], the visual appearance of the input hand image is modeled using the feature extracted from image and compare with the feature of the stored image. The advantages of this approach are simple, easier than 3D model based because of easier extraction of features in 2D [27] image. But this method is affected by changing illumination condition and other background objects.

B 3D Model Based Approaches

This approach use 3D model description [3] for modeling and analysis of the hand shape. It search for kinematic parameters and are requires by making 2D projection from 3D model [35][36] of the hand to correspond edge images of the hand. But lot of hand features might be lost in 2d projection. Different types of methods are there including Volumetric and Skeletal type. Volumetric model deals with 3D visual appearance of human hand and usually used in real time applications.

The problem is it deals with all the parameters of the hand which are huge dimensionality. To overcome volumetric hand parameter problem, skeletal method can be used. It limits the set of parameters to model the hand shape from 3D structure.

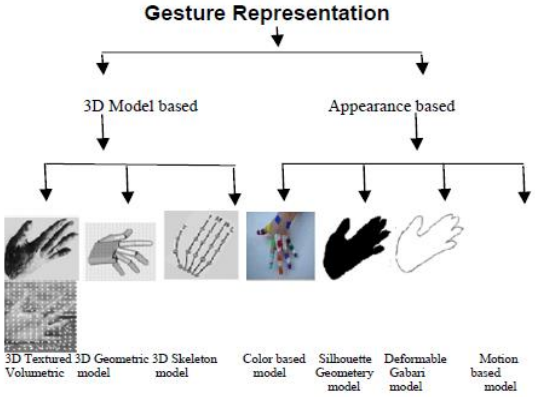

Fig 3 Gesture Recognition Approaches

IV. SEGMENTATION

The detection of the hand region can be achieved through color segmentation. The aim is to classify the pixels of the input image into skin color and non-skin color clusters [16] [17]. This can be accomplished by using a thresholding technique that exploits the information of a skin color [14] distribution map in an appropriate color space. It is a fact that skin color varies quite dramatically. First of all, it is vulnerable to changing lightning conditions that obviously affect its luminance. Moreover, it differs among people and especially among people from different ethnic groups. The perceived variance, however, is really a variance in luminance due to the fairness or the darkness of the skin. Researchers, also, claim that the skin chromaticity is the same for all races [19]. So regarding to the skin color, luminance introduces many problems, whereas chromaticity includes the useful information. Thus, proper color spaces for skin color detection are those that separate luminance from chromaticity components.

Wu [11] developed a hand gesture recognition system for media player control. The system firstly separated the left arm by background subtraction and detected the straight line by both Hough transform and Radon transform. The disadvantage of this method was the non-instinct of defined hand gestures. Lai [12] designed and implemented an interactive biped robot which could be controlled by hand gestures. The number of fingers and angles between fingers were used to classify nine types of static hand gestures. To overcome the effect of lighting, they utilized scroll bars to manually set the scope of skin color in *YCbCr* space [14]. Tu [13] presented a face based hand gesture recognition system for human-computer interaction by a single camera. Hand region was assumed to appear by the side of face. Eleven static hand gestures were defined to control the computer. Back propagation neural network was utilized for hand gesture recognition. However, users need to remember the meaning of each hand gesture which may be confused due to similar shape.

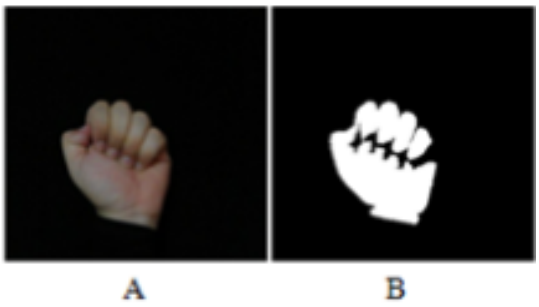

Fig.2. Segmented hand

The methods used for segmentation are color model [14][7], clustering, RGB to HSV[1] space method and YCbCr [14] color method. YCbCr color space [21] deals with the pigment of the skin of the human hand. The significant property of this color space is that the human different ethics group can be recognized according to their pigment concentration which can be distinguished according to some skin color saturation. HSV color model is used to extract the skin-like hand region by estimating the parameter values for skin pigment, and used Laplacian filter for detection of the

edges [21] [22].Clustering method [34] is used for segmentation by grouping the image pixels among image objects [3].

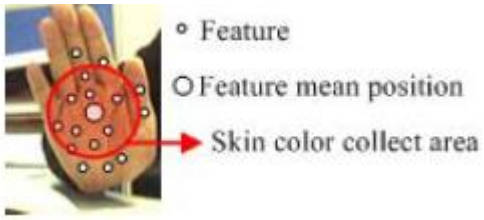

Figure 7 Skin color segmentation

Fig 8 Segmentation results

V. FEATURE EXTRACTION AND CLASSIFICATION

A Neural Network

A neural network is an information processing system loosely based on the operation of neurons in the brain. While the neuron acts as the fundamental functional unit of the brain, the neural network uses the node as its fundamental unit; the nodes are
connected by links, and the links have an associated weight that can act as a storage
mechanism. Artificial neural networks (ANNs) provide a general, practical method for learning real-valued, discrete-valued, and vector-valued functions from examples. Algorithms such as back propagation use gradient descent to tune network parameters to best fit a training set of input-output pairs. ANN learning is robust to errors in the training data and has been successfully applied to problems such as interpreting visual scenes, speech recognition, and learning robot control strategies. Most ANN software runs on sequential machines emulating distributed processes, although faster versions of the algorithms have also been implemented on highly parallel machines and on specialized hardware designed specifically for ANN applications. While ANNs are loosely motivated by biological neural systems, there are many complexities to biological neural systems that are not modeled by ANNs, and many features of the ANNs are inconsistent with biological systems. For example, ANNs whose individual units output a single constant value, whereas biological neurons output a complex time series of spikes.

Most of the researchers use ANN as a classifier in gesture recognition system. Stergiopoulou E [24] recognized static hand gestures using self growing and self organized neural gas (SGONG) network. YCbCr color space is used to detect the hand and SGONG network use competitive Hebbian learning algorithm for learning process. Two neurons are used and continuous growing till a grid of neurons is constructed and covers the hand object which will capture the shape of the hand. Manar M [25] used two recurrent neural network architectures to recognize Arabian Sign Language. The result shows it has a recognition rate of 95.11%.

B Hidden Markov Model

In describing hidden Markov models it is convenient first to consider Markov chains. Markov chains are simply finite-state automata in which each state transition arc has an associated probability value; the probability values of the arcs leaving a single state sum to one. Markov chains impose the restriction on the finite-state automaton that a state can have only one transition arc with a given output; a restriction that makes Markov chains deterministic. A hidden Markov model (HMM) can be considered a generalization of a Markov chain without this Markov-chain restriction [21]. Since HMMs can have more than one arc with the same output symbol, they are nondeterministic, and it is impossible to directly determine the state sequence for a set of inputs simply by looking at the output (hence the “hidden” in “hidden Markov model”).

Hidden Markov model a stochastic process [1] with a finite number of states of Markov chain, and a number of random functions. It is used in sign language recognition [27], speech recognition [28] etc. Keskiin C. [29] presented HCI interface based on real time hand tracking and 3D gesture recognition using hidden Markov models (HMM). Two colored cameras for 3D construction are used. To overcome the problem of using skin color for hand detection because of hand overlapping with other body parts, markers are used to reduce the complexity in hand detection process.

Schlenzig also used a single HMM to recognize hand gestures in a vision-based solution[31]. The state of the HMM represents the gestures and the observation symbols represent the current static hand posture. This HMM had three possible states and nine possible observation symbols so the number of recognizable gestures was limited. The system employs a recursive filter for updating estimates of the gesture being recognized based on the current posture information. This recursive estimator allows recognition of combined gestures and requires only one HMM.

Hidden Markov models provide a good way to perform hand posture [30] and gesture recognition, and can be used in both vision-based and glove-based solutions. The literature has shown that high accuracy can be achieved and the number of possible hand gestures or postures in a posture or gesture set can be quite large. Like neural networks, HMMs must be trained and the correct number of states for each posture or gesture must be determined to maximize

performance. If the number and types of hand posture and gestures are known beforehand, HMMs are a good choice for recognition. If the hand postures and gestures are determined as the system is developed, the development process can be more time consuming due to retraining. If one HMM is used for all gestures, as in Starner's work, then the single HMM must be retrained. If each gesture has an associated HMM, then only the new gestures HMM will have to be trained. Although HMMs require extensive training, and their hidden nature makes it difficult to understand what is occurring within them, they still may be the technique of choice since they are well covered in the literature and the accuracies reported are usually above 90 percent.

C Condensation Algorithm

The particle filtering in its basic form actually realizes the recursive Bayes filter according to a sampling procedure, often called sequential importance sampling with resampling (SISR) it has the ability to represent a wide range of probability densities [32], allowing real-time estimation of nonlinear, non-Gaussian dynamic systems. This technique was originally developed to effectively track objects in clutter. The condensation algorithm was developed based on the principle of particle filtering. It was originally applied effectively in tracking rapid motion of objects in clutter [32]. A mixed-state condensation algorithm has been extended to recognize a greater
number of gestures based on their temporal trajectories.

D FSM

A method to recognize human-hand gestures using a FSM model- based approach has been used in [33]. The state machine is used to model four qualitatively distinct phases of a generic gesture—static start position (static at least for three frames), smooth motion of the hand and fingers until the end of the gesture, static end position for at least three frames, and smooth motion of the hand back to the start position. The hand gestures are represented as a list of gesture vectors and are matched with the stored gesture vector models based on vector displacements.

E Fuzzy clustering

Clustering algorithms is a general term comprises all methods that partitioning the given set of sample data into subsets or clusters [34] based on some measures between grouped elements. Xingyan L. In [34] presented fuzzy c-means clustering algorithm to recognize hand gestures in a mobile remote. A camera was used for acquire input raw images, the input RGB images are converted into HSV color model, and the hand extracted after some preprocessing operations to remove noise and unwanted objects, and thresholding used to segment the hand shape.

VI. CONCLUSION

Hand postures and gestures are an interesting interaction paradigm in a variety of computer applications. Two principal questions must be answered when using them. The first question is what technology to use for collecting raw data from the hand. Generally, two types of technologies are available for collecting this raw data. The first one is a glove input device, which measures a number of joint angles in the hand. The second way of collecting raw data is to use computer vision. In a vision-based solution, one or more cameras placed in the environment record hand movement. The second question to be answered when using hand posture and gestures is what recognition technique will maximize accuracy and robustness. There are a number of interesting areas for future research in hand posture and gesture recognition. The field is by no means mature – we have a long way to go before this type of metaphor is robust enough to be seen in commercial, mainstream applications. Research into better hardware for data collection is important. Better joint angle bend sensors, tracking systems, and faster processors will benefit the field immensely.

VII. REFERENCES

[1] Susmitha Mitra, Tinku Acharya, "Gesture Recognition: A Survey," IEEE Transactions on system, man and cybernetics, vol 37, No. 3, May 2007
[2] Noor Adnan Ibraheem, " Survey on various gesture recognition technologies and techniques,"International Journal of computer Applications , vol 50, No. 7, July 2012
[3] Rafiqul Zaman Khan, " Survey on gesture recognition for hand image postures," www.ccsenet.org/cis, vol 5, No. 3, May 2012
[4] Manjunatha M B, Pradeep Kumar B P, "Survey paper on hand gesture recognition," International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol 3, No. 4, April 2014
[5] Harshith C, Karthik R Shastry, "Survey on various gesture recognition techniques for interfacing machines based on ambient intelligence," International journal of computer science and engineering survey, vol 1,No.2, Novemeber 2010
[6] Eshed Ohn Bar, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision based approach and evaluations," IEEE transactions on intelligent transportation systems, vol 15, No. 6, December 2014
[7] Zhou Ren, Junsong Yuan, "Robust part-based hand gesture recognition using kinect sensor," IEEE transactions on multimedia, vol 15, No. 5, August 2013
[8] Yuan Yao, Yun Fu, "Contour model-based hand gesture recognition using the kinect sensor," IEEE transactions on circuits and systems for video technology, vol 24, No. 11, Novemeber 2014
[9] J Triesch, C Ven Der Malsburg, "A System for person- independent hand posture recognition against complex backgrounds", IEEE computer society, December 2001.
[10]M W Krueger, "Videoplace- An artificial reality", SIGCHI Conference, CHI'85.
[11] Y. M. Wu, "The implementation of gesture recognition for media player system," Master Thesis of the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, 2009.
[12] I. H. Lai, "The following robot with searching and obstacleavoiding," Master Thesis of the Dept. of Electrical Engineering, National Central University, Chung-Li, Taiwan, 2009.
[13] Y. J. Tu and H. Y. Lin, "Human computer interaction using face and gesture recognition," Master Thesis of the Department of Electrical Engineering, National Chung Cheng University, Taiwan, 2007.
[14] R Kjeldsen , J Kender, "Automatic face and gesture recognition", Second international conference, IEEE 1996.
[15] R E Kahn, " Gesture recognition using the preseus architecture", IEEE Computer Society Conference 1996.

[16] Kwang Ho Seek, " A robot motion authoring using finger- robot interaction", IEEE 2009
[17] Xiaoming Yin, "Hand posture recognition in gesture based human robot interaction", IEEE 2006
[18] Dnyanada Jadhav, "A gesture recognition system to control slide show navigation", IJAIEM 2014
[19] Joyeeta Singha, Karen Das, " Hand gesture recognition based on Karhunen-Loeve Transform", Mobile and embedded technology, International Conference 2013.
[20] Anupam Agarwal, "Real time hand gesture recognition system for dynamic applications", IJU, vol. 3, No.1, January 2012.
[21] Hasan M M and Mishra P K, "Gesture Recognition using modified HSV segmentation", IEEE International conference on communication systems and network technologies. http://dx.doi.org/10.1109/CSNT.2011.75
[22] Sanjay Meena, 2011. A Study on Hand Gesture Recognition Technique, Master thesis, Department of Electronics and Communication Engineering, National Institute of Technology, India.
[23] Swapnil D. Badgujar, Gourab Talukdar (2014), "Hand Gesture Recognition System", International Journal of Scientific and Research Publications, Volume 4, Issue 2.
[24] Stergiopoulou, E., & Papamarkos, N. (2009). Hand gesture recognition using a neural network shape fitting technique. Elsevier Engineering Applications of Artificial Intelligence 22, 1141-1158. http://dx.doi.org/10.1016/j.engappai.2009.03.008
[25] Maraqa, M., & Abu-Zaiter, R. (2008). Recognition of Arabic Sign Language (ArSL) using recurrent neural networks. IEEE First International Conference on the Applications of Digital Information and Web Technologies, ICADIWT 2008, pp.478 http://dx.doi.org/10.1109/ICADIWT.2008.466(Fels, 1998)96
[26] Bilal, S., Akmeliawati, R., El, S. M. J., & Shafie, A. A. (2011). Vision-based hand posture detection and recognition for sign language - A study. IEEE 4th international conference on Mechatronics Proceedings, (ICOM 2011), pp. 1-6. http://dx.doi.org/10.1109/ICOM.2011.5937178
[27] Thad Starner and Alex Pentland, 1996. Real-Time American Sign Language Recognition from Video Using Hidden Markov Models, AAAI Technical Report FS-96-05, The Media Laboratory Massachusetts Institute of Technology.
[28] Lawrence R. Rabiner, 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, vol. 77 (2), pp. 257 – 286.
[29] C. Keskin, A. Erkan, L. Akarun, 2003. Real Time Hand Tracking and 3D Gesture Recognition for Interactive Interfaces using HMM, In Proceedings of International Conference on Artificial Neural Networks.
[30] Rubine, Dean. Specifing Gestures by Example. In Proceedings of SIGGRAPH'91, ACM Press, 329-337, 1991.
[31] Schlenzig, Jennifer, Edward Hunter, and Ramesh Jain. Recursive Spatio-Temporal Analysis: Understanding Gestures. Technical Report VCL-95-109,Visual Computing Laboratory, University of California, San Diego, 1995.
[32] M. J. Black and A. D. Jepson, "A probabilistic framework for matching temporal trajectories: Condensation-based recognition of gestures and expressions," in Proc. 5th Eur. Conf. Comput. Vis., vol. 1, 1998, pp. 909–924.
[33] J. Davis and M. Shah, "Visual gesture recognition," Vis., Image SignalProcess., vol. 141, pp. 101–106, 1994.
[34] Xingyan Li, 2003. Gesture Recognition based on Fuzzy C-Means Clustering Algorithm, Department of Computer Science. The University of Tennessee. Knoxville.
[35] Rama B. Dan, " Survey on hand gesture recognition approaches" ,(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 2014, 2050-2052
[36] Hong Cheng, "A Survey on 3D Hand gesture Recognition", IEEE Transactions on Circuits and systems for Video Technology, July 2015
[37] Lingchen Chen, " A survey on Hand Gesture Recognition", International Conference on Computer sciences and Applications, 2013
[38] Haitham Hasan, "Human computer interaction for vision based hand gesture recognition: A Survey", International conference on advanced computer science applications and technologies, 2012.
[39] Jayesh S Sonkusare, " A Review on Hand gesture Recognition system", International conference on computing communication control and automation, 2015.

Neethu P S received the BTech degree in Applied Electronics & Instrumentation from M G University in 2008, M.E. degree in Applied Electronics from Anna University, Chennai in 2012 and currently pursuing PhD in Anna University Chennai.

Dr. R Suguna received PhD degree from Anna University Chennai. Her area of interest are Networking, Image processing etc.