E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

A COMPREHENSIVE REVIEW ON VISION BASED HAND GESTURE RECOGNITION TECHNOLOGY

Ms. Chinnu Thomas [1], Asst. Prof. Ms. S.Pradeepa [2]
[1] *PG Scholar, Department of Computer Science and Engineering*
[2] *Asst. Prof, Department of Computer Science and Engineering*
[1 2] *Adithya Institute of Technology, Coimbatore.*
1 chinnuthomas13@gmail.com 2 praddeepa.s@gmail.com

ABSTRACT:
In this era of an incredibly fast and evoluting industry of usability the topic of User Experience and User interaction has been increasingly popular and widespread lately. As life is getting easier and more plentiful, people are demanding more. They are no longer satisfied with powerful features and solid quality, the more intuitive way to use the function and more decent user impression is what they need and prefer. In the present day framework of interactive, intelligent computing, an efficient human computer interaction is assuming utmost importance. Gesture based computing enables humans to interface with the machine (HMI) and interact naturally without any dedicated devices. Building a richer bridge between machines and humans than primitive text user interface or even (graphical user interfaces) GUIs, which still limit the majority of input to keyboard and mouse. In fact we are bridging this gap by bringing intangible, digital information out into the tangible world, and allowing us to interact with this information via natural hand gestures. Gesture Based Computing provides an attractive alternative for human computer interaction (HCI). This project proposes comprehensive review for a real-time system capable of understanding commands through a survey on tool, techniques and algorithms with emphasis on hand gestures which focuses on the analysis and comparisons from the review. Also discuss and highlights the challenges, applications and future scope.

Keywords: *Human Machine Interface; Human Computer interaction; Hand Gesture; Gesture based Computing.*

1. INTRODUCTION

The miniaturization of computing devices allows us to carry computers in our pockets, keeping us continually connected to the digital world there is no link between our digital devices and our interactions with the physical world. We are bridging this gap by bringing intangible, digital information out into the tangible world, and allowing us to interact with this information via natural hand gestures. Gestures considered as a natural way of communication among human especially for hear-impaired, since it is a physical movement of hands ,arms, or body which conveying meaningful information by delivering an expressive message. Gesture recognition then, is the interpretation of that movement as semantically meanings command.

Computerized hand gesture recognition has received much attention from academia and industry in recent years, largely due to the development of human-computer interaction (HCI) technologies and the growing popularity of smart devices such as smart phones. The essential aim of building hand gesture recognition system is to create a natural interaction between human and computer where the recognized gestures can be used for controlling a robot [15] or conveying meaningful information [1]. How to form the resulted hand gestures to be understood and well interpreted by the computer considered as the problem of gesture interaction.

Hand gestures recognition (HGR) is one of the main areas of research for the engineers, scientists and bioinformatics. HGR is the natural way of Human Machine interaction and today many researchers in the academia and industry are working on different application to make interactions more easy, natural and convenient without wearing any extra device. However, human-robot interaction using hand gestures provides a formidable challenge. For the vision part, the complex and cluttered backgrounds, dynamic lighting conditions and a deformable human hand shape, wrong object extraction can cause a machine to misunderstand the gesture. If the machine is mobile, the hand gesture recognition system also needs to satisfy other constraints, such as the size of the gesture in the image and adaptation to motion. In addition, to be natural, the machine must be person-independent and give feedback in real time.

The purpose of this paper is to presents a comprehensive study on the hand postures and gesture recognition methods, which is considered to be a challenging problem in the human-computer interaction context and promising as well., and to explain various approaches with its advantages and disadvantages. Although recent reviews [6], [7][8][15][19][20][1] in computer vision based have explained the importance of gesture recognition system for human computer interaction (HCI), this work concentrates on vision based techniques method and it's up-to-date. With intending to point out various research developments as well as it emphasis good beginning for interested persons in hand gesture recognition area. Finally, we give some discussion on the current challenges and open questions in this area and point out a list of possible directions for future work.

The paper is organized as follows. Section 2 explains the basic concepts and approaches of Hand Gesture Technology. The vision based hand gesture approaches are detailed in Section 3. Section 4 includes several related works. The applications and challenges are discussed in Section 5. Finally the conclusion is given in section 6.

2. HAND GESTURE TECHNOLOGY

To enable hand gesture recognition, numerous approaches have been proposed, which can be classified into various categories. For any system the first step is to collect the data necessary to accomplish a specific task. For hand posture and gesture recognition system different technologies are used for acquiring input data. A common taxonomy is based on whether extra devices are required for raw data collecting. In this way, they are categorized into data glove based hand gesture recognition, vision based hand gesture recognition [5], and color glove based hand gesture recognition [6]. Figure 1gives an example of these technologies.

2.1. Data glove based approaches

Data glove based approaches require the user to wear a cumbersome glove-like device, which is equipped with sensors that can sense the movements of hand(s) and fingers, and pass the information to the computer. Hence it can be referred as Instrumented glove approach. These approaches can easily provide exact coordinates of palm and finger's location and orientation, and hand configurations, The advantages of these approaches are high accuracy and fast reaction speed. However these approaches require the user to be connected with the computer physically which obstacle the ease of interaction between users and computers, besides the price of these devices are quite expensive.

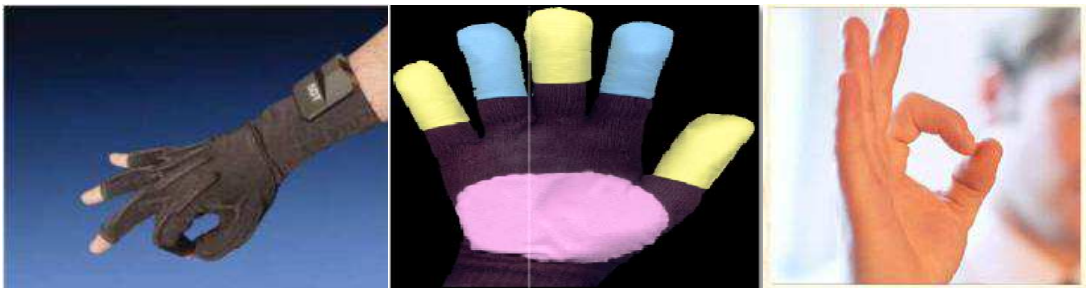
2.2. Colored Markers based approaches

Color glove based approaches represent a compromise between data glove based approaches and vision based approaches. Marked gloves or colored markers are gloves that worn by the human hand [6] with some colors to direct the process of tracking the hand and locating the palm and fingers [6], which provide the ability to extract geometric features necessary to form hand shape [6]. The amenity of this technology is its simplicity in use, and cost low price comparing with instrumented data glove. Intrinsically, they are similar to the latter, except that, with the help of colored gloves, the image preprocessing phase (e.g., segmentation, localization and detection of hands) can be greatly simplified. The disadvantages are similar to data glove based approaches: they are unnatural and not suitable for applications with multiple users due to hygiene issues.

2.3. Vision Based approaches:

Vision based approaches do not require the user to wear anything (naked hands). Instead, video camera(s) are used to capture the images of hands, which are then processed and analyzed using computer vision techniques. This type of hand gesture recognition is simple, natural and convenient for users and at present they are the most popular approaches to gesture recognition**.** Although these approaches are simple but a lot of gesture challenges are raised such as the complex background, lighting variation, and other skin color objects with the hand object, besides system requirements such as velocity, recognition time, robustness, and computational efficiency [1][10].

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org



(a) Data-Glove based. (b) Colored marker based (c) Vision based.

Fig 1: Examples of hand gesture recognition input technologies.

3. VISION BASED HAND GESTURE RECOGNITION APPROACHES

Only a vision-based approach will allow for freehand and barehanded interaction. Vision based technology deals with some image characteristics such as texture and color [6] for acquiring data needed for gesture analyze. In a vision-based gesture recognition system, a mathematical model of gestures is always established first. Two major approaches in gesture modeling have been utilized so far for detecting hand object after some image preprocessing operations. One is a 3D hand model which uses a set of joint angle parameters and segment lengths to model the hand. The other is an appearance-based model, which models the gesture by relating its appearance to the appearance of the predefined, template gestures

3.1. Appearance Based Approaches

Appearance based approaches where hand image is reconstructed using the image properties and extraction. Also known as View Based Approaches, which model the hand using the intensity of 2D images and define the gestures as a sequence of views. The visual appearance of the input hand image is modeled using the feature extracted from the image, which will be compared with the features extracted from stored image. Appearance based approaches considered easier than 3D model approaches, which led many researchers to search for alternative representations of the hand.

3.2. Model Based Approaches

Model based approaches where different models are used to model image using different models to represent in Computers. Model-based approaches [8][9][10] estimate the current hand state by matching a 3D hand model to the observed image features. Such approaches can achieve some good results, however they search the hand in a high dimensional space hence may not be suitable for real-time application. Although this approach reduces the searching dimension, the use of color glove does look natural. 3D Model can be classified into volumetric and skeletal models . Volumetric models deal with 3D visual appearance of human hand and usually used in real time applications [15][2]. The main problem with this modeling technique is that it deals with all the parameters of the hand which are huge dimensionality. Skeletal models overcome volumetric hand parameters problem by limiting the set of parameters to model the hand shape from 3D structure [9].

Soft Computing Approaches

Under the umbrella of soft computing principal constituents are Neural Networks, Fuzzy Systems, Machine Learning, Evolutionary Computation, Probabilistic Reasoning, etc. and their hybrid approaches.

4. LITERATURE REVIEW OF GESTURE RECOGNITION SYSTEMS

William F. and Michal R. [3] presented a method for recognizing gestures based on pattern recognition using orientation histogram. For digitized input image, black and white input video was used, some transformations were made on the image to compute the histogram of local orientation of each image, then a

filter applied to blur the histogram, and plot it in polar coordinates. The system consists of two phases; training phase, and running phase. In the training phase, for different input gestures the training set is stored with their histograms. In running phase an input image is presented to the computer and the feature vector for the new image is formed, Then comparison performed between the feature vector of the input image with the feature vector (oriented histogram) of all images of the training phase, using Euclidean distance metric and the less error between the two compared histograms will be selected. The total process time was 100 msec per frame. Problems include similar gestures might have different orientation histograms and different gestures could have similar orientation histograms, besides that, the proposed method achieved well for any objects that dominate the image even if it is not the hand gesture.

Xingyan L. In [5] 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. 13 elements were used as feature vector, first one for aspect ratio of the hand's bounding box, and the rest 12 parameters represent grid cell of the image, and each cell represents the mean gray level in the 3 by 4 blocks partition of the image, where the mean value of each cell represents the average brightness of those pixels in the image, Then FCM algorithm used for classification gestures. Various environments are used in the system such as complex background and invariant lighting conditions. 6 hand gestures used with 20 samples for each gesture in the vocabulary to create the training set, with recognition accuracy 85.83%. There are some restrictions with current system as The recognition accuracy drops quickly when the distance between the user and the camera is greater than 1.5 meters or when the lighting is too strong. The system cannot deal with an image that has two or more patches of skin with similar size. The system regards the arm as part of the hand.We will try to delete the arm from the image by checking the physical size and fingertips in our future work.

Stergiopoulou E. [1] recognized static hand gestures using Self-Growing and Self-Organized Neural Gas (SGONG) network. A camera used for acquiring the input image, and YCbCr color space is applied to detect hand region, some thresholding technique used to detect skin color. SGONG network use competitive Hebbian learning algorithm for learning process, the learning start with only two neurons and continuous growing till a grid of neurons are constructed and cover the hand object which will capture the shape of the hand. From the resultant hand shape three geometric features was extracted, two angles based on hand slope and the distance from the palm center was determined, where these features used to determine the number of the raised fingers. For recognizing fingertip, Gaussian distribution model used by classifying the fingers into five classes and compute the features for each class. The system recognized 31 predefined gestures with recognition rate 90.45%, in processing time 1.5 second, but it is time consuming and when the number of training data increase, the time needed for classification are increased too.

Hasan [4] applied multivariate Gaussian distribution to recognize hand gestures using nongeometric features. The input hand image is segmented using two different methods ; skin color based segmentation by applying HSV color model and clustering based thresholding techniques. Some operations are performed to capture the shape of the hand to extract hand feature; the modified Direction Analysis Algorithm are adopted to find a relationship between statistical parameters (variance and covariance) from the data, and used to compute object (hand) slope and trend by finding the direction of the hand gesture.

Lamberti [6] presents a real-time hand gesture recognizer based on a color glove. The recognizer is formed by three modules. The first module, fed by the frame acquired by a webcam, identifies the hand image in the scene. The second module, a feature extractor, represents the image by a nine-dimensional feature vector. The third module, the classifier, is performed by means of Learning Vector Quantization (LVQ). The recognizer, tested on a dataset of 907 hand gestures, has shown very high recognition rate. Used HSI color model to segment the hand object. The features vector formed by five distances from palm to all fingers and four angles between those distances.

Daeho Lee and SeungGwan Lee [11] presents a novel vision based method in which , fingertips are detected by a novel scale-invariant angle detection based on a variable *k*-cosine. Fingertip tracking is implemented by detected region-based tracking. By analyzing the contour of the tracked fingertip, fingertip parameters, such as position, thickness, and direction, are calculated. Finger actions, such as moving, clicking, and pointing, are recognized by analyzing these fingertip parameters.

Zhou Ren [20] propose a novel distance metric, Finger-EarthMover's Distance (FEMD), to measure the dissimilarity between hand shapes. By thresholding from the hand position with a certain depth interval, a rough

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

hand region can be obtained , represent it as a time-series curve. Can apply the height information in time-series curve to decompose the fingers.

Table 1 comparisons between different gesture recognition techniques

Paper	Segmentation type	Feature Vector Representation	Classification Algorithm	Recognition Rate
William F. and Michal R. - Orientation histogram	N/A	Orientation histogram	Euclidean distance metric	N/A
Xingyan Li – Fuzzy C-Means Clustering	threshold	One dimensional array of 13 element	Fuzzy C-Means algorithm	85.83%
Stergiopoulou E. neural network	YCbCr color space	Two angles of the hand shape, compute palm distance	Gaussian distribution	90.45%
Hasan - HSV brightness	HSV color model	5x5 geometric moments which is brightness value of each block separately.	Laplacian filter Euclidian distance metric	91 %
Lamberti – Color glove	HSI color space	Five distances from palm to all fingers and four angles between those distances.	Learning Vector Quantization	98%

5. APPLICATION AREAS AND CHALLENGES

5.1 Application areas

Hand gestures recognition system has been applied for different applications on different domains as alternative level of interaction. This section introduces a series of new applications of vision based hand gesture recognition in recent years to give an indication of its prospect in the future.

5.1.1. Alternative to Touch-Based Devices

E-ISSN: 2321–9637
Volume 2, Issue 1, January 2014
International Journal of Research in Advent Technology
Available Online at: http://www.ijrat.org

Tablet PCs and smart phones have driven our everyday life into the era of touch-to-control, by getting rid of traditional input devices such as mouse and keyboard. In the meantime, vision based hand gesture recognition techniques are pushing the user experience one step further by providing a touch-free solution.

5.1.2. *Sign Language Recognition*

Since the sign language is used for interpreting and explanations of a certain subject during the conversation, it has received special attention. A lot of systems have been proposed to recognize gestures using different types of sign languages.

5.1.3. *Robot Control*

Controlling the robot using gestures considered as one of the interesting applications in this field[2][15]. [15] proposed a system that uses the numbering to count the five fingers for controlling a robot using hand pose signs. The orders are given to the robot to perform a particular task [15], where each sign has a specific meaning and represents different function for example, “one” means “move forward”, “five” means “stop”, and so on.

5.1.4. Graphic Editor Control

Graphic editor control system requires the hand gesture to be tracked and located as a preprocessing operation. uses 12 dynamic gestures for drawing and editing graphic system. Shapes for drawing are; triangle, rectangular, circle, arc, horizontal and vertical line for drawing, and commands for editing graphic system are; copy, delete, move, swap, undo, and close .

5.1.5. Virtual Environments (VEs)

One of the popular applications in gesture recognition system is virtual environments VEs, especially for communication media systems. [9] provided 3D pointing gesture recognition for natural human computer Interaction HCI in a real-time from binocular views. The proposed system is accurate and independent of user characteristics and environmental changes [9].

5.1.6. *Numbers Recognition*:

Another recent application of hand gesture is recognizing numbers. Proposes an automatic system that could isolate and recognize a meaningful gesture from hand motion of Arabic numbers from 0 to 9 in a real time system using HMM.

5.1.7. Surgical System

In a surgical environment, hand gesture recognition systems can help doctors manipulate digital images during medical procedures using hand gestures instead of touch screens or computer keyboards

5.1.8. Television Control:

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

5.1.9. 3D Modeling

To build 3D modeling, a determination of hand shapes are needed to create, built and view 3D shape of the hand [9]. Some systems built the 2D and 3D objects using hand silhouette [9]. 3D hand modeling can be used for this purpose also which still a promising field of research [9][14].

5.2 Difficulties

The main difficulties encountered in the design of hand pose estimation systems include:

5.2.1. High-dimensional problem

The hand is an articulated object with more than 20 DOF. Although natural hand motion does not have 20 DOF due to the interdependences between fingers, studies have shown that it is not possible to use less than six

dimensions. Together with the location and orientation of the hand itself, there still exist a large number of parameters to be estimated

5.2.2. *Self-occlusions*

Since the hand is an articulated object, its projection results in a large variety of shapes with many self-occlusions, making it difficult to segment different parts of the hand and extract high level features.

5.2.3. Processing speed

Even for a single image sequence, a real-time CV system needs to process a huge amount of data. On the other hand, the latency requirements in some applications are quite demanding in terms of computational power. With the current hardware technology, some existing algorithms require expensive, dedicated hardware, and possibly parallel processing capabilities to operate in real-time.

5.2.4. Uncontrolled environments

For widespread use, many HCI systems would be expected to operate under nonrestricted backgrounds and a wide range of lighting conditions. On the other hand, even locating a rigid object in an arbitrary background is almost always a challenging issue in computer vision.

5.2.5. Rapid hand motion

The hand has very fast motion capabilities with a speed reaching up to 5 m/s for translation and 300_/s for wrist rotation. Currently, off-the-shelf cameras can support 30–60 Hz frame rates. Besides, it is quite difficult for many algorithms to achieve even a 30 Hz tracking speed. In fact, the combination of high speed hand motion and low sampling rates introduces extra difficulties for tracking algorithms (i.e., images at consecutive frames become more and more uncorrelated with increasing speed of hand motion).

6. CONCLUSION

A survey on tools and techniques of gesture recognition system has been provided with emphasis on hand gesture expressions. The major tools surveyed include HMMs, ANN, Orientation Histograms, Geometrical method, HSV and fuzzy clustering have been reviewed and analyzed. Most researchers are using colored images for achieving better results. Comparison between various gesture recognition systems have been presented with explaining the important parameters needed for any recognition system which include: the segmentation process, features extraction, and the classification algorithm. Even Though the recognition rate increases gradually the concerned algorithms experience several issues. The issues include wrong object extraction, complex and nonuniform background, partial-occlusion, background disturbance, object reappearance, illumination change etc. . The selection of specific algorithm for recognition depends on the application needed. In this work application areas for the gestures system are presented. Explanation of gesture recognition issues, detail discussion of recent recognition systems are given as well. Two-handed dynamic-gesture multimodal interaction is thus a promising area for future research. Customization of gestures also can be employed to make it more user friendly.

References

[1] 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

[2] Malima, A., Özgür, E., & Çetin, M. (2006). 'A fast algorithm for vision-based hand gesture recognition for robot control'. IEEE 14th conference on Signal Processing and Communications Applications, pp. 1- 4. http://dx.doi.org/10.1109/SIU.2006.1659822

[3] W. T. Freeman and Michal R., (1995) 'Orientation Histograms for Hand Gesture Recognition', IEEE International Workshop on Automatic Face and Gesture Recognition.

[4] M. M. Hasan, P. K. Mishra, (2011). 'HSV Brightness Factor Matching for Gesture Recognition System', International Journal of Image Processing (IJIP), Vol. 4(5).

[5] Xingyan Li. (2003). 'Gesture Recognition Based on Fuzzy C-Means Clustering Algorithm', Department of Computer Science. The University of Tennessee Knoxville.

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

[7] C. Keskin, F. Krac, Y. Kara, and L. Akarun, (2011). 'Real time hand pose estimation using depth sensors,' in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Nov. 2011, pp. 1228–1234.
[8] Z. Pan,Y. Li,M. Zhang, C. Sun,K. Guo,X. Tang, and S. Zhou, (2010) 'A realtime multi-cue hand tracking algorithm based on computer vision,' in IEEE Virt. Real. Conf., 2010, pp. 219–222.
[9] G. Hackenberg, R.McCall, andW. Broll, (2011). 'Lightweight palm and finger tracking for real-time 3-D gesture control,' in IEEE Virtual Reality Conf., Mar. 2011, pp. 19–26.
[10] A. Chaudhary, J. Raheja, and S. Raheja, (2012). 'A vision based geometrical method to find fingers positions in real time hand gesture recognition,' J. Software, vol. 7, no. 4, pp. 861–869.
[11] D. Lee and S. Lee, (2011). 'Vision-based finger action recognition by angle detection and contour analysis,' ETRI J., vol. 33, no. 3, pp. 415–422.
[12] J. Raheja, A. Chaudhary, and K. Singal, (2011). 'Tracking of fingertips and centers of palm using Kinect,' in Proc. Int. Conf. Computat. Intell., Modell. Simulat., pp. 248–252.
[13] V. Frati and D. Prattichizzo, (2011). 'Using Kinect for hand tracking and rendering in wearable haptics,' in IEEE World Haptics Conf. WHC, Jun. 2011, pp. 317–321.
[14] H. Liang, J. Yuan, and D. Thalmann,(2011). '3-D fingertip and palm tracking in depth image sequences,' in Proc. 20th ACM Int. Conf. Multimedia, pp. 785–788.
[15] Ghobadi, S. E., Loepprich, O. E., Ahmadov, F., Bernshausen, J., Hartmann, K., & Lo®eld, O. (2008).'Real time hand based robot control using multimodal images'. International Journal of Computer Science IJCS.Vol 35(4). Retrieved from www.iaeng.org/IJCS/issues_v35/issue_4/IJCS_35_4_08.pdf
[16] Garg, P., Aggarwal, N., & Sofat, S. (2009).'Vision based hand gesture recognition'. World Academy of Science, Engineering and Technology. Retrieved from www.waset.org/journals/waset/v49/v49-173.pdf
[17] K. Grauman and T. Darrell (2004) . 'Fast contour matching using approximate earth mover's distance'. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Washington DC.
[18] Zhou Ren, Junsong Yuan, Wenyu Liu, (2013). 'Minimum Near-Convex Shape Decomposition', IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 10, october.
[19] Marco Maisto, Massimo Panella, Member, IEEE, Luca Liparulo, and Andrea Proietti (2013). 'An Accurate Algorithm for the Identification of Fingertips Using an RGB-D Camera', In IEEE journal on emerging and selected topics in circuits and systems, vol. 3, no. 2, june.
[20] Zhou Ren, Junsong Yuan,(2013). 'Robust Part-Based Hand Gesture Recognition Using Kinect Sensor', IEEE transactions on multimedia, vol. 15, no. 5, August.