International Journal of Science and Research (IJSR)
ISSN: 2319-7064
ResearchGate Impact Factor (2018): 0.28 | SJIF (2018): 7.426

Survey on Static and Dynamic Hand Gesture Recognition Techniques

Ophelia M[1], Keerthijith P[2]

[1, 2]Department of Computer Science and Engineering, Vimal Jyothi College of Engineering, Chemperi, Kerala-670632, India

Abstract: ***A hand gesture recognition system provides a natural way of non-verbal communication. Human Computer Inter-action mostly involves hand gestures. Vision-based hand gesture recognition techniques have many advantages over traditional devices, giving users a comfortable and more intuitive way of communication between a human and a computer. Hand gestures are of two types: Static hand gestures and Dynamic hand gestures. Hand gestures which can be either static or dynamic, for human computer interaction is an area of active research and with many numerous possible applications. This survey describes different systems used for gesture recognition. This paper presents a literature review on various gesture recognition methods.***

Keywords: Cyber-Glove, Dynamic gesture, Human-Computer Interaction, Gestures, Posture, Skin color detection, Static gestures

1. Introduction

Gesture is a form of non-verbal communication using various body parts such as hands and fingers. Gesture is the oldest method of communication among humans. Primitive men used to communicate about the information of food/ prey for hunting, information about their enemy, request for help etc. within themselves with the help of gestures. This mainly includes human-robot interaction, sign language recognition, interactive games, vision-based augmented reality etc. For communication by the people at a visible distance, but not audible distance and by the physically challenged people like the deaf and dumb gesture is the only method.

Hand gesture recognition system is built to create a natural interaction between human and computer. The target of this method is the proposition of a real time vision system for its application inside visual association situations through hand gesture recognition, utilizing broadly useful equipment and minimal effort sensors, similar to a straightforward PC and a USB Webcam, so any client could make utilization of it in his/her office or home Posture is another word often confused with gesture. Posture refers to a single image corresponding to a single command (such as stop), whereas a sequence of postures is called gesture (such as move the screen to left or right). They are either static posture and dynamic posture. When compared to postures is simple and needs less computational power, but gesture mainly dynamic one is complex. Despite the fact that occasionally face and other partof the body is used along with single hand or double hands, hand gesture is most popular among all and is used in wide variety of applications.

Hand gesture recognition technology have a place in wide variety of applications such as virtual environments, smart surveillance, sign language translation, medical systems etc. Hand gestures are used for analyzing and annotating video sequences of technical talks. Such a system is presented in [1]. Gestures like pointing or writing are automatically tracked and recognized to provide a rich annotation of the sequence that can be used to access a condensed version of the talk. Given the constrained domain a simple "vocabulary" of actions is defined, that can easily be recognized based on the active contour shape and motion. The recognized actions provide a rich annotation of the sequence that can be used to get a condensed version of the talk from a web page. Gesture recognition technique recognizes static or dynamic hand gestures or combinations of both. Static hand gesture restricts the movement of hands. In the case of dynamic hand gesture user can interact in a more comfortable manner.

There are various approaches used to recognize gestures like vision-based gesture recognition, glove-based gesture recognition, marker-based gesture recognition etc. Vision-based gesture recognition helps to create a more user-friendly interface restricting a user from wearing gloves and other external devices. Glove-based devices such as the most common CyberGlove have been used to capture human hand motions. However, the gloves and its attached wires are still quite cumbersome and awkward for users to wear those gloves during the interaction, and moreover, the cost of the glove is often too expensive for regular users. In the current state-of-the-art vision-based hand tracking and gesture classification methods, the research is more focused on tracking the bare hand without the help of any type of gloves and recognizing hand gestures. Whereas, the vision-based hand gesture recognition system also needs to meet the requirements including real-time performance and accuracy improving recognition rate.

2. Hand Gesture Recognition Techniques

Vision based hand gesture recognition creates a natural interface between human and computer. This approach mainly uses webcam as the camera. This approach do not require the user to wear anything i.e.bare hands are used in this approach. Video cameras are used to capture the images of hands, which are then processed and analyzed using computer vision techniques [2]. This type of hand gesture recognition is simple, natural and comfortable for users. These are the most popular methods for gesture recognition. However, there are several challenges to be addressed, for example, illumination change, background clutter, partial or full occlusion etc. Vision based hand gesture recognition can

be further classified into appearance based approaches and 3D model based approaches. Appearance based approaches involves extracting the low-level features from 2D images and comparing them with predefined template gestures. They are relatively simple in computation but the loss of the depth information make it vulnerable to background disturbance. The 3D model based approaches can exploit the depth information and are much more computation-ally expensive when compared to appearance based approach but can identify hand gestures more effectively. The main challenge of vision-based hand gesture recognition is to deal with the large variety of gestures. Another main challenge is inaccurate hand segmentation. Vision based approach doesn't operate well in cluttered environments.

Creation and implementation of such efficient and accurate gesture recognition systems are done through two types of enabling technologies for human computer interaction namely contact based and vision based devices. Contact based devices employed for gesture recognition systems involve physical interaction of user with the interfacing device i.e. the user needs to be familiar with the usage of these devices, hence not adaptable to the naive users. These devices are usually based on technologies such as data glove, accelerometers, multi-touch screen etc that uses several detectors. This class of contact based devices for gesture is further classified as mechanical, haptics, ultrasonic, inertial and magnetic.

Contact based system includes glove based systems whereas non-contact based system includes vision based techniques. Nowadays vision based techniques are found to be of great importance.Glove-based systems have an important role in the field of gesture recognition. They are designed to recognize meaningful expressions of human hand motion. In our daily life, we use our hands for interacting with the environment around us in many tasks. Glove-based gesturerecognition technique concentrate on developing technologies for studying the hand motion and interaction with a data glove which can augment the capabilities of some users to perform some specific tasks. This idea is relevant in many research areas, for example information visualization, robotics, sign language interaction medicine and health care.

Apart from vision-based and glove-based gesture recognition systems, sensors are also used in hand gesture recognition. There are different types of depth cameras used for gesture recognition but the most prominently used sensor is the Kinect from Microsoft and leap motion sensor due primarily to its low cost. Every sensors and camera systems produce a sequence of depth images that are then used for hand localization. Recent development of depth sensors provides a robust solution to hand segmentation. Kinect sensor developed by Microsoft containsa QVGA (320x240) depth camera and a VGA (640x480) video camera, both of that produce image streams at 30 frames per second (fps). The most popular depth sensors used are Time of Flight (ToF) cameras and stereoscopic cameras. ToF cameras will determine pixel depths in one of two ways: by measuring the round-trip flight-time of light projected onto the scene and then reflected back to the sensor. ToF cameras produce accurate depth images at a high frame rate of 50 fps, but at a relatively low resolution (144x176).

3. Hand Gesture Recognition System

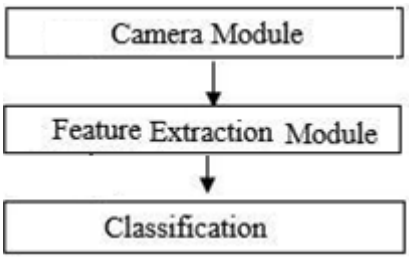

Figure 1: Block Diagram of Gesture Recognition System

Hand gesture recognition system have an important role in human-computer interaction. Main components of a basic gesture recognition system is shown in fig 1. These three components play a major role in hand gesture recognition system. Camera module may involve webcam or other depth cameras. In the first step, various preprocessing steps such as gray scale conversion, morphological operation, thresholding etc are done. In the feature extraction step, all relevant features like hand contour, convex hull of hands etc are extracted. These features are extracted and are used for later recognition phase.

4. Literature Survey

There are two types of hand gestures: Static hand gesture and Dynamic hand gesture. Both of these gestures play an important role in our day to day life. Whether static or dynamic, is recognized using different techniques. This section provides a brief literature survey on static and dynamic gesture recognition techniques.

4.1 Static Hand Gestures

Hamid et.al[3] proposed static hand gesture recognition for human-computer interaction. The objective of this work was to propose a vision based hand gesture recognition algorithm using both wavelet network for images feature extraction and supervised feed-forward neural network with back training algorithm for the classification of various hand gestures. The proposed system of hand gesture recognition consists of three basic steps: preprocessing, feature extraction using wavelet network and classification using neural network. This recognizes gestures using two methods. Firstly, hand segmentation is done by using wavelet network for feature extraction. After feature extraction neural network is used for gesture recognition. The main limitation of this system is that neural network takes large amount of time to classify the data which will add delay to the system for classifying various gestures.

An approach was made by Trong-Nguyen, Nguyen et.al[4] with low computational cost was proposed to recognize static hand gestures. A vision-based hand gesture recognition consists of three main steps: preprocessing, feature extraction and identification. In this paper, the first step involves following two sub-stages: segmentation which locates hand using color information and extracts its

silhouette; separation that separates arm, the part with less information, based on geometrical properties. In the second step, features which are extracted from hand-without-arm are general (ratio of width to height, wrist angle and number of fingers) and detailed (calculated based on fingertips and cross sections) characteristics. Finally, the support vector machine model with max-wins voting strategy is used to classify the hand gestures. The testing is conducted on color image dataset of Polish Ministry of Science and Higher Education, and obtained a result of 89.5% classification accuracy.

Hui Li et.al [5] proposed a hand gesture recognition system using the Kinect sensor using HOG features. This approach involves choosing HOG feature with both the geometric moment invariant features and adapted to the light transform by identifying the features of hands characteristics. The extraction of hand features shouldbe insensitive to light and rotation. It draws the histograms of oriented gradient (HOG) which has wide application on target detection in recent years to extract hand features. And then adopt Adaboost training methods to train the models of hand gestures. Finally it forms an accuracy and efficient static hand gesture recognition system. HOG have both the characteristics of geometric invariant moments and also adapt to the change of the light, so it is suitable for using in gesture recognition. In the first step image is divided into overlapping blocks, and then each block is divided into four cells. Then the gradient value and direction of each pixel is calculated. At last each block obtains the 4*9 dimensional feature vector, and normalized the feature vector of the block. Based on the HOG features and Adaboost training algorithm, the experiment demonstrated the detection results is great, but for the some challenging situations like hands covered in front of body or objects kind of similar to hands, there is still some high missing and false rate.

Xingyan Li et.al[6] have proposed gesture recognition based on fuzzy c-means clustering algorithm. The Fuzzy C-Means algorithm provided efficient speed and sufficient reliability to perform the desired task. This system uses preprocessing for detecting hand in the image. This method uses HSV channel to detect the skin-like regions in an image. Then from the segmented hand shape, it is converted into feature vector containing contours, edges etc. The main disadvantage of this system is that if the objects larger than the hand then the problem of wrong object extraction occurs.
Mokhtar et.al [7]HSV Brightness Factor Matching for Gesture Recognition System. This system is composed of four main phases. In the first phase, input gesture is given into the system. The second phase is image preprocessing which consists of segmentation, edge detection and normalization. Segmentation is used to segment the hand area in the input gesture and isolate it from the background. Most of the gesture recognition systems depend on the perfect segmentation of the hand gesture region, there are two main techniques for segmentation: first method is by using HSV model; which deals with the color pigment of the human skin. The second method used for segmentation operation is by using clustering algorithms or thresholding techniques. In the normalization phase, the gesture is trimmed to get rid of the unnecessarily area that surrounding the gesture area. Next feature extraction takes place which results in feature vectors. These feature vectors are stored in database and is used for further classification of gestures.
Yika et.al [8] have proposed a real-time and gesture recognition method. The whole process of gesture recognitionin this system includes three steps. In the first step hand detection with Adaboost is used to trigger tracking and recognition. In the second step adaptive hand segmentation is executed during detection and tracking with motion and color cues. Finally, scale-space features detection is applied to extract palm-like and finger-like structures. Type of hand gesture is determined by palm-finger configuration.

Ashfag et.al[9] proposed classification of hand gestures using gabor filter with bayesian and naive bayes classifier. In this first a skin filter is applied to detect the hand region. Next, for representing the input image,the L.A.B color space is used. The mask corresponding to skin regions is done by thresholding. The noise removal is then performed on the resulted binary image using morphological operations such as erosion and dilation etc. In the stage of feature extraction, a collection Gabor filters is employed to emphasize hand characteristics in different orientations.

Padam et.al[10] proposed static gesture recognition using moments. This system identifies the hand region through skin color identification and obtains the binary silhouette. These images are normalized for rotation and scale changes. The moment features of the normalized hand gestures are classified using a minimum distance classifier.

4.2 Dynamic Hand Gestures

Chenyang Zhang et.al [11] have proposed an edge enhanced depth motion map framework to model different hand gestures from their visual effects. They have designed a new dynamic temporal pyramidorganization approach to capture temporal structure to compensate the information loss due to building energy map by integrating discrete energy from each frame along temporal dimension. For classification, they apply a Support Vector Machine with linear kernel.

Somayeh Shiravandi et.al[12] proposed a combinatorial method for hand gesture recognition using Bayesian dynamic network was designed. This method includes two main subdivisions namely: hand posture recognition and dynamic hand gesture recognition. In the first step, after hand segmentation by using a method based on histogram of direction and fuzzy SVM classifier, we train the gesture recognition system. In the second step, after skin detection and face and hands segmentation, their tracing were carried out by means of Kalman filter. Then, by tracing the obtained data, the positions of hands were achieved. For combining both the achieved data and output of hand posture recognition unit we utilize Bayesian dynamic network.

HsiangYueh et.al[13] proposed a real-time dynamic hand gesture recognition system. The eleven types of dynamic hand gestures have been recognized, which represent the number from one to nine. The dynamic images are caught by a dynamic video. We use the YCbCr color space transformation to detect the skin color and then find the hand contour from the complex background. All hand gestures are

established under the dynamic video and OpenCV. The three convex defect character points of the hand contour are defined to calculate the angles between the fingers, and the fingertip positions are calculated to recognize the hand gestures.

Hong-xiang et.al[14] presents a new hand gesture model combining starting hand posture, middle motion trajectory and end hand posture. By detecting hand gesture's changes in movement rate the start and end of hand gesture can be determined. It gain starting hand and ending hand shape based on motion and color information and real-timely track the trajectory of hand gesture by using target detection of combining Mean shift algorithm. The experiment results show that the technique of the proposed-dynamic hand gesture modeling and extraction of motion hand gestures has good real-time and robustness.

Rajat et.al [15] proposed a hidden markov model based dynamic hand gesture recognition using opencv. In this method, HMM is used for hand gesture recognition. This system is divided into three stages detection and tracking, feature extraction and training and recognition. The first stage uses a more non-conventional approach of application of L color space for detecting hands. Feature extraction phase is the combination of Hu invariant moments and hand orientation. In the training phase, Baum-Welch algorithm using Left-Right Banded (LRB) topology is applied and recognition is done by Forward algorithm with an average recognition rate above 90% for isolated hand gestures. Due to the use of OpenCV's inbuilt functions, the system is easy to develop, its recognition rate is very fast and so the system can be practically used for real-time applications.

Joyeeta et.al[16] proposed a vision-based approach is used to build a dynamic hand gesture recognition system. Various challenges such as cluttered background, change in illumination and occlusion make the detection and tracking of hand difficult in many vision-based approaches. To overcome such challenges, a hand detection method is developed by combining three-frame differencing and skin filtering. The three-frame differencing is done for both colored and grayscale frames. The hand is tracked using modified KanadeLucas-Tomasi feature tracker where the features are selected using the compact criteria. Velocity and orientation information are added to remove the redundant feature points. Finally, color cue information is used to find the final hand region in the tracked region. In the feature extraction phase, 44 features were selected from the existing literatures. Combining all features could lead to over fitting, information redundancy and dimension disaster.

Xiaoyan et.al[17] presented a hidden markov model based dynamic hand gesture recognition. Here,Adaboost algorithm is used to detect the user's hand and a contour-based hand tracker is formed by combining condensation and partitioned sampling. Cubic B-spline is used to approximately fit the trajectory points into a curve. Invariant curve moments such as global features and orientation as local features are computed to represent the trajectory of hand gesture.

4.3 Static and Dynamic Hand Gestures

Guilluame et.al[18] presents a static and dynamic hand gestures recognition in the depth data using dynamic time warping. In this method, depth data is collected using Kinect sensor. The interest space corresponding to the hands are segmented first based on the assumption that the hand is the closest object in the scene to the camera. A novel algorithm is used to improve the scanning time in order to identify the first pixel on the hand contour within this space. Starting from this pixel, a directional search algorithm is performed for the identification of the entire hand contour. Then k-curvature algorithm is employed to locate the fingertips over the contour, and dynamic time warping algorithm is used to select gesture candidates and also to recognize gestures by comparing an gesture with a series of prerecorded reference gestures.

Subash et.al [19] presents a real time static and dynamic hand gesture recognition system. Hand gesture recognition system provides us natural, innovative, user friendly way of computer interaction. This system recognizes both static as well as dynamic gestures. Static system is divided into four major parts: normalization, skin color detection, skin color model, hand gesture recognition. Normalization is a technique that changes the range of pixel intensity values. In skin color detection image is captured using webcam and then convert the captured images into frames. For dynamic hand gesture there are two stages: training stage and testing stage. Training stage generates a transformation subspace for each hand gesture. Testing stage project the test image into each of the subspaces to find the subspace with nearest perpendicular distance. This subspace will be representative of one particular hand gesture. In testing stage, real time dynamic hand gesture recognition is implemented. Testing stage dynamic hand gesture detection and tracking.

5. Conclusion

This paper discusses various methods used for gesture recognition. There are two types of gestures involved here i.e., static as well as dynamic gestures. These methods include from Neural Network, HMM besides using orientation histogram for features representation. For recognising dynamic gestures HMM tools are perfect and have shown its efficiency especially for robot control. SVM, Neural network etc. are used as classifiers and for capturing hand shape. This paper also includes different gesture recognition techniques like contact based systems and non-contact based systems. Contact based system includes glove based systems whereas non-contact based system includes vision based techniques. Nowadays vision based techniques are found to be of great importance

References

[1] Ju, S., Black, M., Minneman, S., Kimber, D, Analysis of Gesture and Action in Technical Talks for Video Indexing, IEEE Conf. on Computer Vision and Pattern Recognition, CVPR 97 ,1997.
[2] Zhu, Yanmin, Zhibo Yang, and Bo Yuan. "Vision based hand gesture recognition." 2013 International Conference on Service Sciences (ICSS). IEEE, 2013

[3] Jalab, Hamid A."Static hand Gesture recognition for human computer interaction." Information Technology Journal 11.9 (2012): 1265-1271..
[4] Nguyen, T.N., Vo, D.H., Huynh, H.H. and Meunier, J., 2014, December. Geometry-based static hand gesture recognition using support vector machine. In 2014 13th International Conference on Control Automation Robotics and Vision (ICARCV) (pp. 769-774). IEEE.
[5] Li, Hui, Lei Yang, Xiaoyu Wu, Shengmiao Xu, and Youwen Wang. "Static hand gesture recognition based on HOG with Kinect." In 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, vol. 1, pp. 271-273. IEEE, 2012.
[6] Li, Xingyan. "Gesture recognition based on fuzzy c-means clustering al-gorithm." Department Of Computer Science The University Of Tennessee Knoxville (2003).
[7] Hasan, Mokhtar M., and Pramod K. Mishra. "HSV brightness factor matching for gesture recognition system." International Journal of Image Processing (IJIP) 4, no. 5 (2010): 456-467.
[8] Fang, Yikai, Kongqiao Wang, Jian Cheng, and Hanqing Lu. "A real-time hand gesture recognition method."In 2007 IEEEInternational Conference on Multimedia and Expo, pp. 995-998.
[9] Ashfaq, Tahira, and Khurram Khurshid. "Classification of handgestures using Gabor filter with Bayesian and naïve Bayesclassifier." *Int. J. Adv. Comput. Sci. Appl.* 7.3 (2016): 276-279.
[10] Priyal, S. Padam, and Prabin K. Bora. "A study on static hand gesture recognition using moments."*2010 International Conference on Signal Processing and Communications (SPCOM)*. IEEE, 2010.
[11] Zhang, Chenyang, and Yingli Tian. "Edge enhanced depth motion map for dynamic hand gesture recognition."In Proceedings of the IEEE Conference on Computer Vision
[12] Shiravandi, Somayeh, Mohammad Rahmati, and Fariborz Mahmoudi. "Hand gestures recognition using dynamic Bayesian networks." 2013 3rd Joint Conference of AI and Robotics and 5th RoboCup Iran Open International Symposium. IEEE, 2013.
[13] Lai, Hsiang Yueh, and Han Jheng Lai. "Real-time dynamic hand gesture recognition." 2014 International Symposium on Computer, Consumer and Control. IEEE, 2014.
[14] Duan, Hong-xiang, Qiu-yu Zhang, and Wei Ma. "An approach to dynamic hand gesture modeling and real-time extraction." In 2011 IEEE 3rd International Conference on Communication Software and Networks, pp. 139-142. IEEE, 2011.
[15] Wang, Xiaoyan, Ming Xia, Huiwen Cai, Yong Gao, and Carlo Cattani."Hidden-markov-models-based dynamic hand gesture recognition." Mathematical Problems in Engineering 2012 (2012).
[16] Singha, Joyeeta, Amarjit Roy, and Rabul Hussain Laskar. "Dynamic hand gesture recognition using vision-based approach for human-computer interaction." Neural Computing and Applications 29, no. 4 (2018): 1129-1141.
[17] Wang, Xiaoyan, et al. "Hidden-markov-models-based dynamic hand gesture recognition." *Mathematical Problems in Engineering* 2012 (2012).
[18] Bao, Jiatong, et al. "Dynamic hand gesture recognition based on SURF tracking." *2011 international conference on electric information and control engineering*. IEEE, 2011.
[19] Plouffe, Guillaume, and Ana-Maria Cretu. "Static and dynamic hand gesture recognition in depth data using dynamic time warping." IEEE transactions on instrumentation and measurement 65, no. 2 (2016): 305-316.
[20] N. Subhash Chandra, T.Venu, and P. Srikanth. "Real Time Static and Dynamic Gesture Recognition."International Journal of Engineering Inventions Volume 4, Issue 12 [Aug.