A Study of Vision Based Hand Gesture Recognition for Human Machine Interaction.

Prof. Rajeshri Rahul Itkarkar[1]   Dr.Anil Kumar Nandy[2]
Research Scholar : BVB COE, Hubli   *Professor: BVB COE, Hubli*
(Affiliated o VTU, Belgum)   (Affiliated to VTU, Belgum)

Abstract : With rapid development of 3-D applications & virtual environments in computer systems the need for a new type of interaction device arises. In other words, evolution of user interfaces shapes the change in the Human-Computer Interaction (HCI). "Hand Gesture" concept in human computer interaction, which has become popular in recent years can be used to develop such an interaction device. This paper presents an overview of real time , vision based hand gesture recognition system that works precisely on a relatively small-restricted gesture space for single user robot control, track a player's hand or body position to control movement and orientation of interactive game , and human-computer interaction.

Index Terms: Hand Gesture, Human machine Interaction, vision based etc.

1. INTRODUCTION.

Hand gesture recognition. It is a relatively new field .Now a day's much research is going on in the field of Artificial Intelligence in Natural language processing. Hand gesture, body postures are also the natural languages. The use of hand gestures provides an attractive alternative to the cumbersome interface devices for human-computer interaction (HCI). User's generally use hand gestures for expression of their feelings and notifications of their thoughts. In particular, visual interpretation of hand gestures can help in achieving the ease and naturalness desired for HCI. Recent researches in computer vision have established the importance of gesture recognition systems for the purpose of human computer interaction. The primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control. A gesture may be defined as a physical movement of the hands, arms, face, and body with the intent to convey information or meaning. An translator is required for an deaf person to interact with an normal person. Sign language is the best media to interact between them. There different sign languages used all over the world such American sign language(ASL), British sign language, Italian sign language, Chinese sign language, Indian sign language etc.. similarly Indian sign language is developed for Indians. This paper presents the gesture recognition for Indian sign language. Figure 1 shows the chart used for Indian sign language.
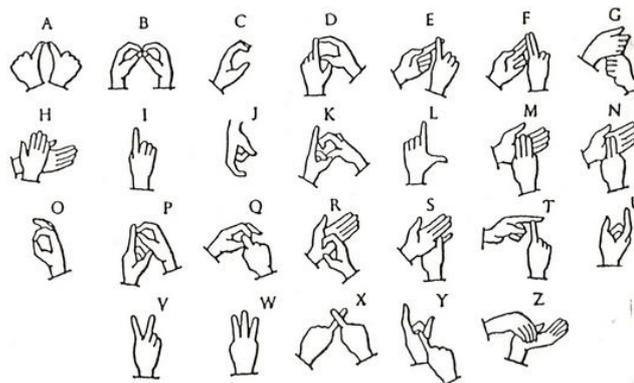
Fig.1. Indian Sign Language Gesture Chart.

Gesture recognition, then, consists not only of the tracking of human movement, but also the interpretation of that movement as semantically meaningful commands [3]. This paper presents an review of real time vision based hand gesture recognition system.

OUTLINE

Section 2 reviews the technique details. Section 3 gives vision based techniques in detail. Section 4 describes the application. Section 5 explains the detail of software's to be used and finally section 6 gives performance measurement parameters.

_____

2.   TECHNIQUES USED FOR GESTURE INTERPRETATION.

Two approaches are commonly used to interpret gestures for Human Computer interaction. They are (a) *Methods Which Use Data Gloves:* This method employs sensors (mechanical or optical) attached to a glove that transduce's finger flexions into electrical signals for determining the hand posture. This approach forces the user to carry a load of cables which are connected to the computer and hinders the ease and naturalness of the user interaction. [3]

(b) *Methods which are Vision Based:* Computer vision based techniques are non invasive and based on the way human beings perceive information about their surroundings. Although it is difficult to design a vision based interface for generic usage, yet it is feasible to design such an interface for a controlled environment.[3] This approach is easier as the person doesn't need to wear hardware.

Further the vision based hand gesture recognition system is modeled as

1.   3D Model based
2.   Appearance based techniques

Figure 2 [4] shows the classification of Vision based hand gesture recognition system. 3D model based methods make use of 3D information of key elements of the body parts. Using this information, several important parameters, like palm position, joint angles etc., can be obtained. This approach uses volumetric or skeletal models, or a combination of the two. In computer animation industry and for computer vision purposes, volumetric approach is better suited. This approach is very computational intensive and also, systems for live analysis are still to be developed [2].
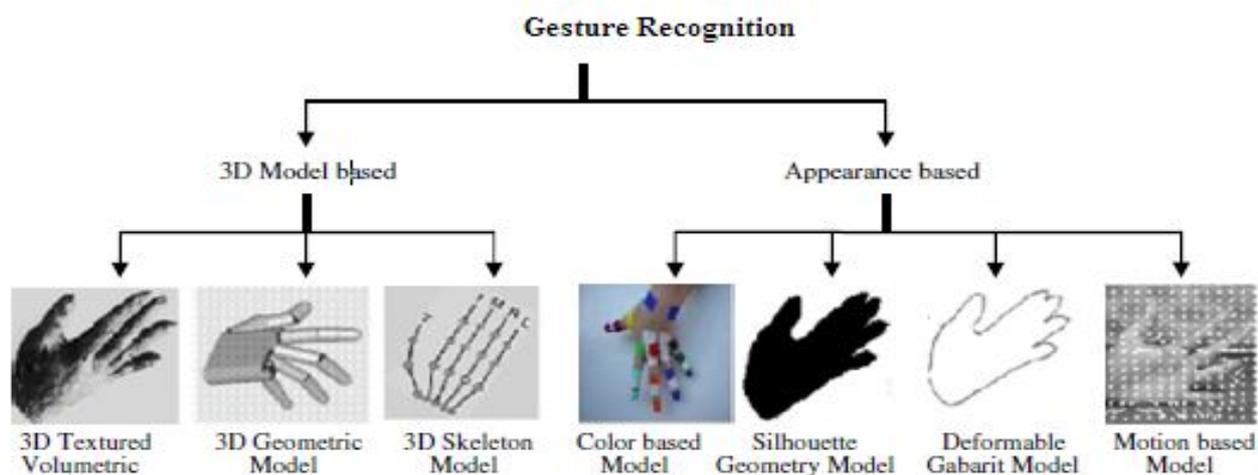
Fig.2 Classification of Vision based Hand gesture Recognition System.

Appearance-based systems use images or videos as inputs. They directly interpret from these videos/images. They don't use a spatial representation of the body. The parameters are derived directly from the images or videos using a template database. Some templates are the deformable 2D templates of the human parts of the body, particularly hands. Deformable templates are sets of points on the outline of an object, used as approximation. One of the simplest interpolation functions is linear. It performs an average shape from point sets, point variability parameters and external deformations. These template-based models are mostly used for hand-tracking, but could also be used for simple gesture classification. A second approach in sign language gesture detection using appearance-based models uses image sequences as the gesture templates. Either the images themselves, or certain features derived from these images can be used as the parameters [2].

Color based model uses body markers to track the motion of body or body part. The hand gesture recognition employs multi-scale color features, hierarchal models and particle filtering [7].

• Silhouette geometry based models include several geometric properties of the silhouette such as perimeter, convexity, surface, bounding box/ellipse, elongation, rectangularity, centroid and orientation. The geometric properties of the bounding box of the hand skin were used to recognize hand gestures [8].

_____

• Deformable gabarit based models: they are generally based on deformable active contours (i.e. snake parameterized with motion and their variants used snakes for the analysis of gestures and actions in technical talks for video indexing [9].

• Motion based models are used for recognition of an object or its motion based on the motion of object in an image sequence. Local motion histogram was introduced which uses an Ad boost framework for learning action models [10]. Recognizing gesture is a complex task which involves many aspects such as object detection, object description, motion modeling, motion analysis, pattern recognition and machine learning even psycholinguistic studies also required [5].

3. VISION BASED HAND GESTURE RECOGNITION TECHNIQUE.

There basically three steps in which Vision based Gesture recognition is achieved.
Step-1 Detection of hand and segmentation of image.
Step-2 Hand Tracking.
Step-3 Gesture recognition.

3.1 DETECTION OF HAND AND SEGMETATION OF IMAGE

Detection of hand gestures and segmentation is done on the basis of Color, Shape, Motion, Pixel and 3D model[4].
Color: This stage mainly does the work of hand detection and background removal from the image. The primary requirement before identifying and classifying the hand signs is to locate the hand in the frame, subtract the background and to be insensitive to lighting conditions. In order to obtain the location of the hand, skin colour recognition is used. The algorithm used for the skin colour identification is to convert the image obtained, which is an image in RGB colour space, to YCbCr model or HSV model.
Shape : The shape of the hand is used to detect the hand in many ways.
i) Contour-based shape representation and description methods are chain Code, Polygon, B-spline, Perimeter, Compactness, Eccentricity, Shape Signature, Handoff Distance, Fourier Descriptor, Wavelet Descriptor, Scale Space, Autoregressive, Elastic matching.
ii) Region-based shape representation and description methods are Convex Hull, Media Axis, Area, Euler Number, Eccentricity, Geometric Moments, Zernike Moments, Pseudo-Zernike Moments, Legendre Moments[11].
Pixel value: Significant work has been carried out on finding hands in grey level images based on their appearance and texture [4]. It is based on the principle that a highly accurate or "strong" classifier can be derived through the linear combination of many relatively inaccurate or\weak" classifiers. In general, an individual weak classifier is required to perform only slightly better than random [4].

3D model: In the current state of the art, due to the limitation of data glove/sensor based approaches, vision based (appearance based) and 3-D hand model based approaches are being used [13]. The employed 3D models should have enough degrees of freedom to adapt to the dimensions of the hand present in an image [4].

Motion: Motion is utilized by a few approaches to hand detection. The reason is that motion based hand detection demands for a very controlled setup, since it assumes that the only motion in the image is due to hand movement. Indeed, early works (e.g. assume that hand motion is the only motion occurring in the imaged environment. In more recent approaches, motion information is combined with additional visual aspects [4].

3.1.2 Hand Tracking
Tracking hands is difficult since they move very fast and their appearance can change within a few frames. Tracking can be defined as the frame-to-frame correspondence of the segmented hand regions or features towards understanding the observed hand movements. It works on two approaches
i) *Correlation-based feature tracking:* correlation-based template matching is utilized to track hand features across frames. Once the hand has been detected in a frame, the image regions in which they appear is utilized as the prototype to detect the hand in the next frame.
ii) *Contour based tracking:* The boundary of the image region is determined by intensity or color gradient. The technique is initialized by placing a contour near the region of interest. The contour is then iteratively deformed towards nearby edges to better fit the actual hand region [4].
This class of tracking is called Template based tracking. Similarly other tracking methods are optimal estimation which achieved by using kalman filter and next is particle filtering and Cam shift method based on principle of mean shift algorithm.

_____

3.1.3 Recognition.

The overall goal of hand gesture recognition is the interpretation of the semantics that the hand(s) location, posture, or gesture conveys. Vision based hand gesture recognition techniques can be further classified under static and dynamic gestures. To detect static gestures (i.e. postures), a general classifier or a template-matcher can be used. However, dynamic hand gestures require techniques that handle this dimension like Hidden Markov Models[4]. Some of the common techniques used for static and dynamic hand gesture recognition are :

K-mean: It 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. The algorithm starts by randomly locating k clusters in spectral space. Each pixel in the input image group is then assigned to the nearest cluster centre and the cluster centre locations are moved to the average of their class values. This process is then repeated until a stopping condition is met [14].

K-nearest neighbor: An object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst its k nearest neighbors; k is a positive integer, typically small. If k = 1, then the object is simply assigned to the class of its nearest neighbor. In binary (two class) classification problems, it is helpful to choose k to be an odd number as this avoids tied votes [4].

Mean Shift Clustering: The main idea behind mean 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. For each data point in the feature space, one performs a gradient ascent procedure on the local estimated density until convergence [15].

Support Vector Machine: The idea behind the method 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.

Dynamic time warping: The DTW algorithm calculates the distance between each possible pair of points out of two signals in terms of their associated feature values. It uses these distances to calculate a cumulative distance matrix and finds the least expensive path through this matrix.

Time delay neural networks: These are special artificial neural networks which focus on working with continuous data making the architecture adaptable to online networks hence advantageous to real time applications.

4 APLLICATIONS OF HAND GESTURE RECOGNITION.

Hand Gesture recognition has a wide range of application in real life & real time scenarios.
- Some advanced applications include tablet PC, games, medicine environment, and augmented reality.
- In desktop applications, hand gestures can offer a substitute interaction medium for mouse and keyboard. Many hand gestures for desktop computing tasks involve manipulating graphic objects or annotating and editing documents using pen-based gestures .Also use of pen gestures, make marking menu selections using stroke gestures.
- Virtual reality is one of application.
- *Robotics and Telepresence:* Telepresence and telerobotic applications are typically situated within the domain of space exploration and military-based research projects.
- In games
- In Sign Language Recognition.

5 USE OF SOFTWARES

When implementing a technique/algorithm for developing an application which detects, tracks and recognize hand gestures, the software used are the important platforms. Some software's which can be used are

OpenCV: It is an open source computer vision programming functions library aimed at developing applications based on real time computer vision technologies.

MATLAB: Matrix laboratory (MATLAB) is a numerical computing environment and fourth generation programming language. It was developed by Math Works. MATLAB allows implementation of algorithms, matrix manipulations, plotting of functions and data, creation of user interfaces and interfacing with programs written in other languages such as c, c++ and java.

6 PERFORMANCE ANALYSIS

Performance evaluation for Hand Gesture recognition can be done by evaluating % Error in recognition or rate of recognition

_____

and the time required for Recognition.

7. DISCUSSION AND CONCLUSION

The main objective of this paper is the vision based method used for gesture recognition with the different approaches. The model based methods is the method which is very computational intensive and also, used for live analysis.i.e. real time. During implementation the static and dynamic methods can be implemented and results can be compared. But the support vector machine from study gives the better performance. The study is done for the Indian sign language as more research till now is carried fro American Sign Language and Chinese sign language. OpenCV software is preferred as it is applicable for real time and execution is faster. Application can be in gaming, robotic control, computer/laptop interaction for deaf people, 3D interaction etc.

REFERENCES

[1] Reza Hassan pour, Asadollah Shahbahrami, *"Human Computer Interaction Using Vision-Based Hand Gesture Recognition"*, Journal of Advances in Computer Research, 2010.

[2] Anuja V. Nair, Bindu V, "A   Indain Sign language Recognition", *International Journal of Computer Applications (0975 – 8887) Volume 73– No.22, July 2013*

[3] G. R. S. Murthy & R. S. Jadon *"A Review of Vision Based Hand Gestures Recognition"* International Journal of Information Technology and Knowledge Management, July-December 2009, Volume 2, No. 2, pp. 405-410

[4] Siddharth S. Rautaray Anupam Agrawal *"Vision based hand gesture recognition for human computer interaction: a survey"* DOI 10.1007/s10462-012-9356-9 Springer Science+Business Media Dordrecht 2012.

[5] Dasgupta, Shulka, S. Kumar, D. Basu, "A Multilingual Multimedia Indian Sign Language Dictionary Tool", The 6th Workshop on Asian Language Resources, 2008, pp. 57-64.

[6] Cote M, Payeur P, Comeau G (2006) Comparative study of adaptive segmentation techniques for gesture analysis in unconstrained environments. In: IEEE international workshop on imagining systems and techniques, pp 28–33

[7] Bretzner L, Laptev I, Lindeberg T (2002) Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering. In: Fifth IEEE international conference on automatic face and gesture recognition, pp 405–410.

[8] Birdal A, Hassanpour R (2008) Region based hand gesture recognition. In: 16th International conference in central Europe on computer graphics, visualization and computer vision, pp 1–7.

[9] Ju SX, Black MJ, Minneman S, Kimber D (1997) Analysis of gesture and action in technical talks for video indexing, Technical report, American Association for Artificial Intelligence. AAAI Technical Report SS-97-03.

[10] Luo Q, Kong X, Zeng G, Fan J (2008) Human action detection via boosted local motion histograms. MachVis Appl.

[11] D. Zhang, C. Lu, "Review of shape representation and description techniques", The Journal of the Pattern Recognition Society, Elsevier, 2004, pp. 1-19.

[12] WuY, Huang TS (2000)View-independent recognition of hand postures. In: Proceedings of the IEEE computer vision and pattern recognition (CVPR), vol 2. Hilton Head Island, SC, pp 84–94.

[13] P. Garg, N. Agrawal, S. Sofat, "Vision based Hand Gesture Recognition", Proceedings of world Academy of Science, Engineering and Technology, Vol.37, 2009, pp. 1024-1029.

[14] Lu W-L, Little JJ (2006) Simultaneous tracking and action recognition using the pca-hog descriptor. In: The 3rd Canadian conference on computer and robot vision, 2006. Quebec, pp 6–13

[15] Derpanis KG (2005) Mean shift clustering, Lecture Notes.

[16] Oi Mean Foong, Tan Jung Low, and Satrio Wibowo, " Hand Gesture Recognition: Sign to Voice System S2V"Proceedings Of World Academy Of Science, Engineering And Technology Volume 32 AUGUST 2008 ISSN 2070-3740

[17] Yu Huang , Monekosso, D. ; Hui Wang ; Augusto, J.C. "A Concept Grounding Approach for Glove-Based Gesture Recognition", 7th International Conference on Intelligent Environments (IE), 2011, ISBN: 978-1-4577-0830-5

[17]Asanterabi Malima, Erol Özgür, and Müjdat Çetin," A FAST ALGORITHM FOR VISION-BASED HAND GESTURE RECOGNITION FOR ROBOT CONTROL*",* 2006 IEEE 14th Signal Processing and Communications Applications (SIU),**ISBN:** 1-4244-0238-7

[18] Hyosun Kim, Georgia Albuquerque, Sven Havemann,,Dieter W. Fellner "Hand Gesture Interaction for Immersive 3D Modeling"

[19]Reza Hassanpour, Asadollah Shahbahra- Journal of Advances in Computer Research 2 (2010) 21-30"Human Computer Interaction Using Vision-Based Hand Gesture Recognition