Natural Hand Gestures Recognition System for Intelligent HCI: A Survey

Vishal Nayakwadi
Department of Computer Science
TSSM's BSCOER College of Engineering
University of Pune, India

N. B. Pokale
Department of Computer Science
TSSM's BSCOER College of Engineering
University of Pune, India

Abstract: Gesture recognition is to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. Hand Gestures have greater importance in designing an intelligent and efficient human–computer interface. The applications of gesture recognition are manifold, ranging from sign language through medical rehabilitation to virtual reality. In this paper a survey on various recent gesture recognition approaches is provided with particular emphasis on hand gestures. A review of static hand posture methods are explained with different tools and algorithms applied on gesture recognition system, including connectionist models, hidden Markov model, and fuzzy clustering. Challenges and future research directions are also highlighted.

Keywords - Hand gesture interface, HCI, Computer vision, Fuzzy Clustering, ANN, HMM, Orientation Histogram.

1. INTRODUCTION

Gestures and face expressions easily used for daily humans interactions [1] while human computer interactions still require understanding and analyzing signals to interpret the desired command that made the interaction sophisticated and unnatural [1]. Recently the designing of special input devices witnessed great attention in this field to facilitate the interaction between humans and computers [2], and to accomplish more sophisticated interaction through the computer [2]. It is worth to mention that the window manager is the earlier user interface to communicate with computers [3]. The combining of traditional devices mouse and keyboard with the new designed interaction devices such as gesture and face recognition, haptic sensors, and tracking devices provides flexibility in Tele-operating [2] [24], text editing [4], robot control [2] [50] [42], cars system control [2], gesture recognition [4], Virtual Reality (VR) [5], and multi-media interfaces [4], video games [4] [44]. Gesture considered as a natural way of communication among people especially hear-impaired [6]. Gestures can be defined as a physical movement [6] of hands, arms, or body that delivers an expressive message [6], and gesture recognition system used to interpret and explain this movement as meaningful command [6][7].

Gesture recognition has been applied in a large range of application areas such as recognizing sign language [6] [8], human computer interaction (HCI) [6] [9], robot control [6], smart surveillance [6], lie detection [6], visual environments manipulating [6], etc. Different techniques and tools have been applied for handling gesture recognition, vary between mathematical models like Hidden Markov Model (HMM) [6][10] [51] [52] and Finite State Machine (FSM) [6][11] to approaches based on software computing methods such as fuzzy clustering [12], Genetic Algorithms (GAs) [13] and Artificial Neural Network (ANN) [14]. Hand posture recognition sill an open research area [15], since the human hand is a complex articulated object with many connected joints and links, which forms the 27 degrees of freedom [16] for the hand. Typically the implementation of gesture recognition system required different kinds of devices for capturing and tracking image/ video image [6] such as camera(s), instrumented (data) gloves, and colored marker [6]. Those devices are used for modeling the communication between human and environments rather than traditional interface device such as keyboards, and mice which are inconvenient and unnatural for HCI system. Vision based technique also different according to some system environments such as number of cameras used [6], speed [6], and illumination conditions [6]. The major difficulty in gesture recognition system is how to identify a specific gesture meaning by the machines (computers/ robot) [17].

The purpose of this paper is to present a review of vision based hand gesture recognition techniques for human computer interaction, and to explain various approaches with its advantages and disadvantages. Although recent reviews [1] [6], [7] [18] [17] [19] [20] 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 represent good starting for interested persons in hand gesture recognition area.

2. HAND GESTURE TECHNOLOGY

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. Present technologies for recognizing gestures can be divided into vision based, instrumented (data) glove, and colored marker approaches. Figure 1 shows an example of these technologies.

2.1 Vision Based approaches:

In vision based methods the system requires only camera(s) to capture the image required for the natural interaction between human and 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 [7] [17].

Fig. 1 Vision based

2.2 Instrumented Glove approaches:

Instrumented data glove approaches use sensor devices for capturing hand position, and motion. These approaches can easily provide exact coordinates of palm and finger's location and orientation, and hand configurations [17] [21] [22], however these approaches require the user to be connected with the computer physically [22] which obstacle the ease of interaction between users and computers, besides the price of these devices are quite expensive [22], it is inefficient for working in virtual reality [22].

Fig. 2 Data Glove

2.3 Colored Markers approaches:

Marked gloves or colored markers are gloves that worn by the human hand [5] with some colors to direct the process of tracking the hand and locating the palm and fingers [5], which provide the ability to extract geometric features necessary to form hand shape [5]. The color glove shape might consist of small regions with different colors or as applied in [23] where three different colors are used to represent the fingers and palms, where a wool glove was used. The amenity of this technology is its simplicity in use, and cost low price comparing with instrumented data glove [23]. However this technology still limits the naturalness level for human computer interaction to interact with the computer [5].

Fig. 3 Color Makers

3. GESTURE RECOGNITION TECHNIQUES

The recognition of gesture involves several concepts such as pattern recognition [19], motion detection and analysis [19], and machine learning [19]. Different tools and techniques are utilized in gesture recognition systems, such as computer vision [38] [55], image processing [6], pattern recognition [6], statistical modeling [6].

3.1 Artificial Neural Networks

The use of neural networks for gesture recognition has been examined by many researchers. Most of the researches use ANN as a classifier in gesture recognition process, while some others use it to extract the shape of the hand, as in [25]. Tin H. [26] presents a system for hand tracking and gesture recognition using NNs to recognize Myanmar Alphabet Language (MAL). Adobe Photoshop filter is applied to find the edges of the input image and histogram of local orientation employed to extract image feature vector which would be the input to the supervised neural networks system. Manar M. [27] used two recurrent neural network architectures to recognize Arabic Sign Language (ArSL). Elman (partially) recurrent neural networks and fully recurrent neural networks have been used separately. A colored glove used for input image data, and for segmentation process, HSI color model is applied.

The segmentation divides the image into six color layers, one for the wrist and five for fingertips. 30 features are extracted and grouped to represent a single image, fifteen elements used to represent the angles between the fingertips and between them and the wrist [27], and fifteen elements to represent distances between fingertips; and between fingertips and the wrist [27]. This input feature vector is the input to both neural networks systems. 900 colored images were used as training set, and 300 colored images for system testing. Results had shown that fully recurrent neural network system (with recognition rate 95.11%) better than the Elman neural network (with 89.67% recognition rate). Kouichi M. in [28] presented Japanese sign language recognition using two different neural network systems. Firstly, back propagation algorithm was used for learning postures of Japanese alphabet. For input postures, data glove is used, and normalization operation was applied as a preprocessing step on the input image. The feature extracted from input images was 13 data items, ten for bending, and three for angles in the coordinates.

The output of the network was 42 characters. The network consists of three layers, the input layer with 13 nodes, the hidden layer with 100 nodes, and the output layer with 42 nodes which corresponds 42 recognized characters. The recognition rate for learning 42 taught patterns was 71.4%, and for unregistered people 47.8%, while the rate improved when additional patterns added to the system, it became 98.0% for registered, and 77.0% for unregistered people. Elman Recurrent Neural Network was the second system applied for recognition gestures. The system could recognize 10 words. The data item have been taken from data glove and the same preprocessing applied for input image. Features extracted are 16 data items, 10 for bending, 3 for angles in the coordinates, and 3 for angles in the coordinates. The network consists of three layers, the input layer with 16 nodes, the hidden layer with 150 nodes, and the output layer with 10 nodes which corresponds 10 recognized words.

Some improvement in the positional data and filtering data space are added to the system [28]. Integration of these two neural networks, in a way, that after receiving data from data glove, a determination of the start sampling time and if the data item considered a gesture or a posture is sent to the next network, for checking the sampling data and the system hold a history, which decide the end of sign language. The final recognition rate with the encoding methods was 96%. Stergiopoulou E. [25] 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. Shweta K. in [29] introduced gesture recognition system using Neural Networks. Web-cam used for capturing input image at slow rate samples between 15-25 frames per second. Some preprocessing have been made on the input image which convert the input image into sequence of (x, y) coordinates using MATLAB, then passed into neural classifier, in which it will classify the gesture into one of several classed predefined classes which can be identified by the system. Sidney and Geoffiey [30],[31] used neural networks to map hand gestures to speech synthesizer using Glove-Talk system that translated gestures to speech through adaptive interface which is an important class of neural networks applications [31].

3.2 Histogram Based Feature

Many researchers have been applied based the histogram, where the orientation histogram is used as a feature vector [32]. The first implementation of the orientation histogram in gesture recognition system and real time was done by William F. and Michal R. [32]; they 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.

Hanning Z., et al. [33] presented hand gesture recognition system based on local orientation histogram feature distribution model. Skin color based segmentation algorithm were used to find a mask for the hand region, where the input RGB image converted into HSI color space, and then map the HSI image H to a likelihood ratio image L the hand region is segmented by thresholding value, 128 elements in the local orientation histogram feature were used. The augmented of the local orientation histogram feature vector implemented by adding the image coordinates of the sub-window. To compact features representation, k-means clustering has been applied the augmented local orientation histogram vectors. In Recognition stage, Euclidean distance used to compute the exact matching score between the input image and stored posture. Then Locality Sensitive Hashing (LSH) used to find the approximate nearest neighbors, and reduce computational cost for image retrieval. Wysoski et al. [34] presented a rotation invariant static-gesture recognition approach using boundary histograms. Skin color detection filter was used, followed by performing erosion, dilation as preprocessing operation, and clustering process to find the groups in the image. For each group the boundary was extracted using an ordinary contour-tracking algorithm.

The image Divided into grids, and normalized the boundary in size, which give the system invariance distance between the camera and hand. Homogeneous background was applied, and the boundary is represented as chord's size chain. The image was divided into a number of regions N. And the regions were divided in a radial form [34], according to a specific angle as shown in the Figure. The histogram of boundary chord's size was calculated. So the whole feature vector consists of a sequential chain of histograms. Multilayer perceptron (MLP) Neural Networks and Dynamic Programming (DP) matching were used as classifiers. 26 static postures from American Sign Language, for every posture, 40 pictures were taken, 20 pictures for training and 20 for test. Different number of histograms were used varies from 8 to 36 increasing by two, with different histogram resolutions.

3.3 Fuzzy Clustering Algorithm

Clustering algorithms is a general term comprises all methods that partitioning the given set of sample data into subsets or clusters [35] based on some measures between grouped elements [12]. According to this measure the pattern that share the same characteristics are grouped together to form a cluster [12]. Clustering Algorithms have been widely spread because of their ability of grouping complicated data collections into regularly clusters [35]. In fuzzy clustering, the partitioning of sample data into groups in a fuzzy way are the main difference between fuzzy clustering and other clustering algorithm [12], where the single data pattern might belong to different data groups [12].

Xingyan L. In [12] 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%.

3.4 Hidden Markov Model (HMM)

Many researches were applied in the field of gesture recognition using HMM. HMM is a stochastic process [6] [52], with a finite number of states of Markov chain, and a number of random functions so that each state has a random function [6]. HMM system topology is represented by one state for the initial state, a set of output symbols [6] [22], and a set of transitions state [22] [8]. HMM contained a lot of mathematical structures and has proved its efficiency for modeling spatio–temporal information data [6]. Sign language recognition, are one of the most applications of HMM [8], and speech recognition [10]. In [9] Keskiin C., et. al. presented HCI interface based on real time hand tracking and 3D gesture recognition using hidden Markov models (HMM) [54]. 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 [9] [52]. Markers used to segment the hand from complex backgrounds under invariant lighting conditions.

The markers are distinguished using marker detection utility, and connected components algorithm was applied to find marker regions using double thresholding. For fingertip detection, simple descriptors were used, where the bounding box and four outmost points of the hand that defining the box is determined [9]. The bounding box in some cases needs to be elongate to determine the mode of the hand, and the points used to predict the fingertip location in different modes of the hand. Kalman filter was used for filtering trajectory of the hand motion. For 3D reconstruction of finger coordinates, calibration utility was implemented for specific calibration object [9]. Least square approach used to generate fingertip coordinates, and kalman filter applied for smoothing the trajectory of 3D reconstructed coordinates. To eliminate coordinate system dependency, the 3D coordinates are converted into sequences of quantized velocity vectors. HMM interprets these sequences [9], which are directional characterizing of the motion [9]. The system designed for game and painting programs application. Hand tracking is utilized to imitate the movements of the mouse for drawing, and the gesture recognition system used for selecting commands. Eight gestures have been used for system training, and 160 for testing, with 98.75% recognition performance.

Table 1. Comparisons between various gesture recognition systems

<table>
<tr><th>Method</th><th>Type Of Input Device</th><th>Segmentation Type</th><th>Features (Geometric Or Non)</th><th>Feature Vector Representation</th><th colspan="2">Classification Algorithm</th><th>Recognition Rate</th></tr>
<tr><td>Tin H. [26]</td><td>Digital camera</td><td>threshold</td><td>Non geometric</td><td>Orientation histogram</td><td colspan="2">supervised neural network</td><td>90%</td></tr>
<tr><td rowspan="2">Manar M. [27]</td><td rowspan="2">Colored glove, and Digital camera</td><td rowspan="2">HSI color model</td><td rowspan="2">N/A</td><td rowspan="2">Available Features from resource</td><td rowspan="2">Two neural network system</td><td>Elman recurrent network</td><td>89.66%</td></tr>
<tr><td>Fully recurrent network</td><td>95.11%</td></tr>
<tr><td rowspan="2">Kouichi M. [28]</td><td rowspan="2">Data glove</td><td rowspan="2">threshold</td><td rowspan="2">Non geometric</td><td>13 data item (10 for bending, 3 for coordinate angles)</td><td rowspan="2">Two neural network system</td><td>back propagation network</td><td>71.4%</td></tr>
<tr><td>16 data item (10 for bending, 3 for coordinate angles, 3 for positional data)</td><td>Elman recurrent network</td><td>96%</td></tr>
</table>

Stergiopoulou E. [25]	Digital camera	YCbCr color space	geometric	Two angles of the hand shape, compute palm distance	Gaussian distribution	90.45%
Shweta in [29]	Web-cam		Non geometric		supervised neural network	N/A
William F. and Michal R. [32]	Digital camera		Non geometric	Orientation histogram	Euclidean distance metric	N/A
Hanning Z., et al. [33]	Digital camera	threshold	Non geometric	augmented of the local orientation histogram	Euclidean distance metric	92.3%
Wysoski, et al. [34]	Digital camera	skin color detection filter	Geometric	the histogram of radial boundary	MLP+ DP matching	98.7%
Xingyan L. [12]	Digital camera	threshold	Non geometric	One dimensional array of 13 element	Fuzzy C-Means algorithm	85.83%
Keskiin C., et al. [39]	Two colored cameras and marker	connected components algorithm with double thresholding	Geometric	sequences of quantized velocity vectors	HMM	98.75%

4. RECOGNITION SYSTEM METHODOLOGY

Many researchers have been suggested on gesture recognition system for different applications, with different recognition phases but they all agree with the main structure of the gesture recognition system. These phases are segmentation, features detection and extraction, and finally the classification or recognition phase. One of these structures illustrated in Figure 4.

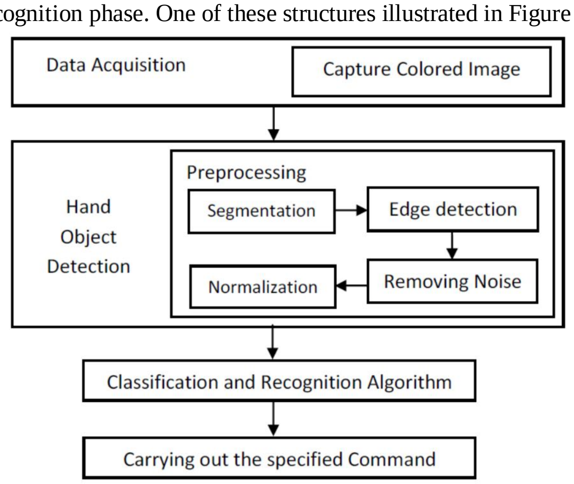

Fig. 4 The flow of gesture recognition system [73]

4.1 Segmentation

Segmentation phase plays an important role in the system recognition process. Perfect segmentation effects on the accuracy of the recognition system [74]. For any segmentation process, some image processing operations are required for hand shape detection [7] [74].

Segmentation image algorithms can be classified into two types according to image gray level properties as explained in [78]:

4.1.1 Discontinuity:

Which tries to find a mass change in the contrast?

4.1.2 Similarity:

Which computes the similarity between neighbor pixels?

When the input gesture acquired form colored camera, instrumented glove device or colored glove as shown in Figure 1. The first step is segmentation, to extract the hand region from the input image and isolate it from the background [74]. There are two main methods for object segmentation, first method depends on the color model that can be extracted from the existence RGB color model which could be HSV color model [74] [75] [77] or YCbCr color space [25] which deals with the pigment of the skin of the human hand [74], 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 [74]. Then, the hand area is isolated from the input gesture with some threshold value. Some normalization for the segmented image might require for obtaining the gestures database which should be invariant against different perturbations like translation, scaling and rotation [74]. The database created with many samples per single gesture, the relation between the number of samples and the accuracy is directly proportional, and between number of samples and the speed is inversely proportional [74].

Hasan [74] used HSV color model to extract the skin-like hand region by estimating the parameter values for skin pigment, and used Laplacian filter for detection the edges. Stergiopoulou [25] used YCbCr color model to segment the hand. Maraqa (2008) used color glove for input gestures and HSI color space for the segmentation process. Ghobadi (2008) treated the segmentation process as clustering method by grouping the image pixels among image objects. Lamberti [23] used HSI color model to segment the hand object. Table 3 shows some applications of the segmentation methods used in the discussed method.

4.2 Features Detection and Extraction

The features are the useful information that can be extracted from the segmented hand object by which the machine can understand the meaning of that posture. The numerical representation of these features can be obtained from the vision perspective of the segmented hand object which forms the feature extraction phase [76]. Many researchers have been applied to form this feature vector which takes different sizes as well as meanings. Hasan [74] extracted the features vector by dividing the segmented hand object into fixed block size 5x5 brightness value moments; this produce 625 features vector size and only 98 are stored as actual features vector. Stergiopoulou [25] applied Self-Growing and Self-Organized Neural Gas (SGONG) network to extract the exact shape of the hand region and determine three characteristics as the features; Palm region, Palm center, and Hand slope. Compute angle between the finger root and the hand center named RC Angle, and the joints the fingertip and the hand center named TC, and angle and distance from the palm center. Li [12] defined a grid of fixed size with 12 blocks gray scale features vector, and each grid cell represents the mean value of the average brightness of the pixels in the block. Lamberti [23] defined the distance d from the palm to the fingers d_i(i = 1, ..., 5), and computed the angle β between the line connecting the centroids of the palm and The fingers, which produces four angles β_i(i = 1, ..., 4), so the hand represented by nine numerical features vector [23]. Table 4 demonstrates features vector representation of these methods.

4.3 Recognition

Recognition or classification of hand gestures is the last phase of the recognition system. Hand gestures can be classified using two approaches as mentioned in [7]

4.3.1 Rule based Approaches:

Which represents the input features as manually encoded rule, and the winner gesture is the one that matched with the encoded rules after his features has been extracted. The main problem of this technique is that the human ability to encode the rules limits the successfulness of the recognition process [7]

4.3.2 Machine Learning based Approaches:

The most common approaches that considered the gesture as result of some stochastic processes [7]. Most of the problems that based on machine learning have been addressed based on the statistical modeling [16], such as PCA [79], FSM [80]. Hidden Markov Models (HMMs) [9] [54] have been paid attention by many researchers [7], Kalman filtering [77], Artificial Neural networks (ANNs) [27] [28] [30] [31] which have been utilized in gesture recognition as well. Some researchers used Gaussian distribution for gestures classification [25] and Euclidian distance metric [74].

5. APPLICATIONS

Lately there has been a great emphasis on Human-Computer Interaction (HCI) research to create easy-to-use interfaces by facilitating natural communication and manipulation skills of humans. Among different human body parts, the hand is the most effective interaction tool because of its dexterity. Adopting hand gesture as an interface in HCI will not only allow the deployment of a wide range of applications in sophisticated computing environments such as virtual reality systems and interactive gaming platforms, but also benefit our daily life such as providing aids for the hearing impaired, and maintaining absolute sterility in health care environments using touch less interfaces via gestures [38].Gesture recognition has wide-ranging applications [49] such as the following:

5.1 Virtual Reality:

Gestures for virtual and augmented reality applications have experienced one of the greatest levels of uptake in computing. Virtual reality interactions use gestures to enable realistic manipulations of virtual objects using ones hands, for 3D display interactions [56] or 2D displays that simulate 3D interactions [57].

5.2 Robotics and Telepresence:

Telepresence and telerobotic applications are typically situated within the domain of space exploration and military-based research projects [47] [24]. The gestures used to interact with and control robots are similar to fully-immersed virtual reality interactions, however the worlds are often real, presenting the operator with video feed from cameras located on the robot [58]. Here, gestures can control a robots hand and arm movements to reach for and manipulate actual objects, as well its movement through the world.

5.3 Desktop and Tablet PC Applications:

In desktop computing applications, gestures can provide an alternative interaction to the mouse and keyboard [59] [43]. Many gestures for desktop computing tasks involve manipulating graphics, or annotating and editing documents using pen-based gestures [60].

5.4 Games:

When, we look at gestures for computer games. Freeman et al. [61] tracked a player’s hand or body position to control movement and orientation of interactive game objects such as cars [44]. Konrad et al. [62] used gestures to control the movement of avatars in a virtual world, and Play Station 2 has introduced the Eye Toy, a camera that tracks hand movements for interactive games [63].

5.5 Sign Language:

Sign language is an important case of communicative gestures [15]. Since sign languages are highly structural, they are very suitable as test beds for vision algorithms [64]. At the same time, they can also be a good way to help the disabled to interact with computers. Sign language for the deaf (e.g. American Sign Language) is an example that has received significant attention in the gesture literature [65] [66] [67] [68].

5.6 Vehicle Monitoring:

Another important application area is that of vehicle interfaces. A number of hand gesture recognition techniques for human vehicle interface have been proposed time to time [69] [70]. The primary motivation of research into the use of hand gestures for in-vehicle secondary controls is broadly based on the premise that taking the eyes off the road to operate conventional secondary controls can be reduced by using hand gestures.

5.7 Healthcare & Medical Assistance:

The healthcare area has also not been left untouched by this technological wave. Wachs et al. [71] developed a gesture based tool for sterile browsing of radiology images. Jinhua Zeng, Yaoru Sun, Fang Wang developed a wheelchair with intelligent HCI [40] [45].

5.8 Daily Information Retrieval:

Sheng-Yu Peng implemented an approach that provides daily information retrieved from Internet, where users can operate this system with his hands' movements [41] [42].

5.9 Education:

Zeng, Bobo, Wang, Guijin presented a system using hand gestures to control poerpoint presentations [48].

5.10 Television Control:

Last application for hand postures and gestures is controlling Television devices [22]. Freeman [72] developed a system to control a television set by hand gestures. Using an open hand and the user can change the channel, turn the television on and off, increase and decrease the volume, and mute the sound.

6. IMPLEMENTATION TOOLS

Many implementation hardware and software tools have been utilized for recognizing gestures depending on the application fields used.

6.1 Hardware Implementation Tools

Input devices used for gesture recognition systems are various and different according to the system and application used for recognition process. Single camera can be used for postures recognition since this environment might be inconvenient for other types of image-based recognition [24]. Stereo camera which consists of two lenses with an isolated sensor for each lens [24], which imitates human visual system therefore, the 3D effect of views is created [24]. Stereo camera can be used to make 3D pictures for movies [24], or for range imaging [24]. Tracking device such as instrumented data gloves measure finger movements through various types of sensors technology [21], [22]. It can provide accurate information about the position and orientation of the hands using magnetic or inertial tracking devices [24]. For more details about various types of glove-based input device refer to [21], [22].

Controller-based gestures, in this type of input gestures, controllers represent a complement of the human, so that when the body moves to perform some gestures [24], these motions are captured using some software [24]. Mouse gestures are an example of such controllers [24]. Other systems based on accelerometers to measure hand movements [36] [37].

6.2 Software Implementation Tools

Software tools are the programming language and windows system used for implementing the gesture recognition system. Some researches applied programming languages like C, C++, and Java language. To simplify the work especially when image processing operations are needed, MATLAB® with image processing toolbox is used. Tin H. [26] used MATLAB® for hand tracking and gesture recognition. Manar M. [27] use MATLAB6 and C language, MATLAB6 used for image segmentation while C language for Hand Gesture Recognition system. Kouichi [28] use SUN/4 workstation for Japanese Character and word recognition. Also Stergiopoulou [25] used Delphi language with 3 GHs CPU to implement hand gesture recognition system using SGONG network. Shweta [29] used MATLAB® for hand recognition. Freeman and Michal Roth [32] used HP 735 workstation was used for implementing the system. Hanning Zhou et. al. [33] used C++ implementation costs only 1/5 second for the whole preprocessing, feature extraction and recognition, when running on a 1.3G Intel Pentium laptop processor with 512MB memory.

7. CONCLUSION

Building an efficient human-machine interaction is an important goal of gesture recognition system. Many applications of gesture recognition system ranging from virtual reality to sign language recognition and robot control. In this paper a survey on tools and techniques of gesture recognition system have been provided with emphasis on hand gesture expressions. The major tools surveyed include HMMs, ANN, 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. In this paper a literature review on gesture recognition has been reviewed and analyzed; the major tools for classification process include FSM, PCA, HMMs, and ANNs are discussed. Descriptions of recognition system framework also presented with a demonstration of the main three phases of the recognition system by detection the hand, extraction the features, and recognition the gesture.

8. REFERENCES

[1] John Daugman, Face and Gesture Recognition: Overview, IEEE transaction on pattern analysis and machine intelligence, vol. 19(7).

[2] Sanjay Meena, A Study on Hand Gesture Recognition Technique, Master thesis, Department of Electronics and Communication Engineering, National Institute of Technology, India. 2011

[3] Myers, B.A., A Taxonomy of User Interfaces for Window Managers, IEEE Transaction in Computer Graphics and Applications, 8(5), pp. 65-84.Doi; 10.1109/38.7762 , 1988

[4] Myers B. A., A Brief History of Human Computer Interaction Technology, ACM interactions. pp. 44-54, Vol. 5(2). Doi: 10.1145/274430.274436, 1998

[5] Mokhtar M. Hasan, and Pramod K. Mishra, Hand Gesture Modeling and Recognition using Geometric Features: A Review, Canadian Journal on Image Processing and Computer Vision 2012 Vol. 3, No.1.

[6] S. Mitra, and T. Acharya, Gesture Recognition: A Survey, IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 37 (3), pp. 311-324, doi: 10.1109/TSMCC.2007.893280.

[7] G. R. S. Murthy & R. S. Jadon, A Review of Vision Based Hand Gestures Recognition, International Journal of

Information Technology and Knowledge Management, 2009 vol. 2(2), pp. 405-410.
[8] Thad Starner and Alex Pentland, 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. 1996.
[9] C. Keskin, A. Erkan, L. Akarun, Real Time Hand Tracking and 3D Gesture Recognition for Interactive Interfaces using HMM, In Proceedings of International Conference on Artificial Neural Networks 2003.
[10] Lawrence R. Rabiner, A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, vol. 77 (2), pp. 257 – 286. 1989.
[11] Pengyu H., Matthew T., Thomas S. Huang, Constructing Finite State Machines for Fast Gesture Recognition, IEEE Proceedings, 15th International Conference on Pattern Recognition (ICPR 2000), vol. 3 ,pp. 3691-694, 2000, doi:10.1109/ICPR.2000.903639.
[12] Xingyan Li, Gesture Recognition based on Fuzzy C-Means Clustering Algorithm, Department of Computer Science. The University of Tennessee. Knoxville. 2003.
[13] David E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Pearson (2002) Edition 1
[14] Ben Krose, and Patrick van der Smagtan, An Introduction to Neural Networks, the University of Amsterdam, eighth edition. 1996.
[15] Sara Bilal, RiniAkmeliawati, Momoh J. El Salami, Amir A. Shafie, 2011. Vision-Based Hand Posture Detection and Recognition for sign Language - A study, IEEE 4th international conference on Mechatronics (ICOM 2011), pp. 1-6.
[16] Vladimir I. Pavlovic, Rajeev Sharma, and Thomas S. Huang, Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review, IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 19(7), pp. 677- 695. 1997.
[17] Pragati Garg, Naveen Aggarwal and Sanjeev Sofat, Vision Based Hand Gesture Recognition, World Academy of Science, Engineering and Technology 49, pp. 972-977, 2009.
[18] Thomas B. Moeslund and Erik Granum, A Survey of Computer Vision-Based Human Motion Capture, Elsevier, Computer Vision and Image Understanding 81, Ideal, pp. 231–268, 2001.
[19] Ying Wu, Thomas S. Huang, Vision-Based Gesture Recognition: A Review, Lecture Notes in Computer Science, Gesture Workshop, proceedings of the international Gesture Workshop on Gesture-Based communication in Human-Computer interaction, vol.(1739), pp. 103-115, 1999.
[20] Ali Erol, George Bebis, Mircea Nicolescu, Richard D. Boyle, Xander Twombly, Vision-based hand poses estimation: A review, Elsevier Computer Vision and Image Understanding 108, pp. 52–73, 2007.
[21] Laura Dipietro, Angelo M. Sabatini, and Paolo Dario, A Survey of Glove-Based Systems and their applications, IEEE Transactions on systems, Man and Cybernetics, Part C: Applications and reviews, vol. 38(4), pp. 461-482, doi: 10.1109/TSMCC.2008.923862, 2008.
[22] Joseph J. LaViola Jr. A Survey of Hand Posture and Gesture Recognition Techniques and Technology, Master Thesis, NSF Science and Technology Center for Computer Graphics and Scientific Visualization, USA, 1999.
[23] Luigi Lamberti & Francesco Camastra, 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, 2011.
[24] MacLean J, Herpers R, Pantofaru C, Wood L,Derpanis K, Topalovic D, Tsotsos J, Fast Hand Gesture Recognition for Real-Time Teleconferencing Applications, IEEE Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, Digital Object Identifier: 10.1109/RATFG.2001.938922, Publication Year: 2001 , Page(s): 133 – 140
[25] E. Stergiopoulou, N. Papamarkos, Hand gesture recognition using a neural network shape fitting technique, Elsevier Engineering Applications of Artificial Intelligence 22, pp. 1141–1158, 2009.
[26] Tin Hninn H. Maung, Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks, World Academy of Science, Engineering and Technology 50, pp. 466- 470, 2009.
[27] Manar Maraqa, Raed Abu-Zaiter, Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks, IEEE First International Conference on the Applications of Digital Information and Web Technologies, ICADIWT, Aug. 2008, pp.478-48, doi: 10.1109/ICADIWT.2008.4664396
[28] Kouichi Murakami and Hitomi Taguchi, Gesture Recognition using Recurrent Neural Networks, ACM, pp. 237-242., 1999.
[29] Shweta K. Yewale, Artificial Neural Network Approach For Hand Gesture Recognition, International Journal of engineering Science and Technology (IJEST), vol. 3(4), 2011.
[30] S. Sidney Fels, Geoffrey E. Hinton, Glove-Talk: A Neural Network Interface Between a Data-Glove and a Speech Synthesizer, IEEE transaction on Neural Networks, vol. 4(1), pp. 2-8, doi: 10.1109/72.182690
[31] S. Sidney Fels, Geoffiey E. Hinton, Glove-Talk II—A Neural-Network Interface which Maps Gestures to Parallel Formant Speech Synthesizer Controls, IEEE transactions on neural networks, vol. 9(1), pp. 205-212, doi: 10.1109/72.655042
[32] William T. Freeman and Michal Roth, Orientation Histograms for Hand Gesture Recognition, IEEE International Workshop on Automatic Face and Gesture Recognition, Zurich.
[33] Hanning Zhou, Dennis J. Lin and Thomas S. Huang, Static Hand Gesture Recognition based on Local Orientation Histogram Feature Distribution Model, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04). 2004.
[34] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi , Akira Iwata, A rotation invariant approach on static-gesture recognition using boundary histograms and neural networks, IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura, November 2002.
[35] James C. Bezdek, Robert Ehrlich, William Full, FCM: The Fuzzy C-Means Clustering Algorithm, Computers & Geosciences vol. 10(2-3), pp. 191-203.

[36] Damien Zufferey, Device based gesture recognition, ACM Second International Conference on Tangible and. Embedded Interaction (TEI'08), 2008.

[37] Marco Klingmann, Accelerometer-Based Gesture Recognition with the iPhone, Master Thesis in Cognitive Computing, Goldsmiths University of London, 2009.

[38] J. P. Wachs, M. Kolsch, H. Stern, and Y. Edan, Vision-based hand gesture applications, Commun. ACM, vol. 54, pp. 60–71, 2011.

[39] Haoyun Xue, Shengfeng Qin, Mobile Motion Gesture Design for Deaf People, IEEE Proceedings of the 17th International Conference on Automation & Computing, September 2011

[40] Jinhua Zeng , Yaoru Sun , Fang Wang, A Natural Hand Gesture System for Intelligent Human-Computer Interaction and Medical Assistance, Intelligent Systems (GCIS), 2012 Third Global Congress on Digital Object Identifier: 0.1109/GCIS.2012.60 Publication Year: 2012 , Page(s): 382 – 385

[41] Sheng-Yu Peng, Kanoksak, Wattanachote, Hwei-Jen Lin, Kuan-Ching Li, A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet, IEEE DOI 10.1109/U-MEDIA.2011.45, Publication Year: 2011

[42] Chun Zhu, Weihua Sheng, Wearable Sensor-Based Hand Gesture and Daily Activity Recognition for Robot-Assisted Living, IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, Volume: 41 , Issue: 3 Digital Object Identifier: 10.1109/TSMCA.2010.2093883, Publication Year: 2011 , Page(s): 569 - 573

[43] Stern H I, Wachs J P, Edan Y, Human Factors for Design of Hand Gesture Human -Machine Interaction , IEEE International Conference on Systems, Man and Cybernetics, SMC '06. Volume: 5 Digital Object Identifier: 10.1109/ICSMC.2006.384767 Publication Year: 2006 , Page(s): 4052 – 4056

[44] Ashwini Shivatare, Poonam wagh, Mayuri Pisal,Varsha Khedkar, Hand Gesture Recognition System for Image Process Gaming, International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 3, March - 2013ISSN: 2278-0181

[45] Yoshinori Kuno, Teruhisa Murashima, Mobutaka Shimada and Yoshiaki Shiraia, Interactive Gesture Interface for Intelligent Wheelchairs, IEEE Conference on Multimedia and Expo. ICME 2000,Vol. 2, 2000

[46] Juan Wachs, Helman Stern, Yael Edan, Michael Gillam, Craig Feied, Mark Smith, Jon Handler A Real-Time Hand Gesture Interface for Medical Visualization Applications, Advances in Intelligent and Soft Computing Springer 2006

[47] Jithin Thampi, Muhammed Nuhas, Mohammed Rafi, Aboo Shaheed, Vision based hand gesture recognization: medical and military applications, Advances in Parallel Distributed Computing Communications in Computer and Information Science Volume 203, 2011, pp 270-281

[48] Zeng, Bobo, Wang, Guijin ; Lin, Xinggang , A hand gesture based interactive presentation system utilizing heterogeneous cameras, IEEE Tsinghua Science And Technology ISSN 1007-0214 15/18, 17(3):329-336, June 2012, DOI: 10.1109/TST.2012.6216765

[49] Michael J. Lyons, Julien Budynek and Shigeru Akamatsu, Automatic classification of single facial images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 12, December 1999.

[50] Asanterabi Malima, Erol Özgür, and Müjdat Çetin, A Fast Algorithm For Vision-Based Hand Gesture Recognition For Robot Control, IEEE Signal Processing and Communications Applications, 2006 IEEE 14th, DOI: 10.1109/SIU.2006.1659822, Publication Year: 2006 , Page(s): 1 - 4 Cited by: Papers (9)

[51] J. Yamato, J. Ohya, and K. Ishii, Recognizing human action in time sequential images using hidden Markov model, in Proc. IEEE Int. Conf. Comput. Vis. Pattern Recogn., Champaign, IL, 1992, pp. 379–385.

[52] F. Samaria and S. Young, HMM-based architecture for face identification, Elsevier, Image Vision Computing, vol. 12, pp. 537–543, 1994.

[53] Stefan Reifinger, Frank Wallhoff, Markus Ablassmeier, Tony Poitschke, and Gerhard Rigoll Static and Dynamic Hand-Gesture Recognition for Augmented Reality Applications, Springer LNCS 4552, pp. 728–737, 2007.

[54] Mahmoud Elmezain, Ayoub Al-Hamadi, Jorg Appenrodt, and Bernd Michaelis, A Hidden Markov Model-Based Isolated and Meaningful Hand Gesture Recognition, International Journal of Electrical and Electronics Engineering 3:3 2009

[55] Ashutosh Samantaray, Sanjaya Kumar Nayak, Ashis Kumar Mishra , Hand Gesture Recognition using Computer Vision, International Journal of Scientific & Engineering Research, Volume 4, Issue 6, June 2013 ISSN 2229-5518

[56] Sharma, R., Huang, T. S., Pavovic, V. I., Zhao, Y., Lo, Z., Chu, S., Schulten, K., Dalke, A., Phillips, J., Zeller, M. & Humphrey, W, Speech/Gesture Interface to a Visual Computing Environment for Molecular Biologists, In: Proc. of ICPR’96 II (1996), 964-968.

[57] Gandy, M., Starner, T., Auxier, J. & Ashbrook, D, The Gesture Pendant: A Self Illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring, Proc. of IEEE Int. Symposium on Wearable Computers. (2000), 87-94.

[58] Goza, S. M., Ambrose, R. O., Diftler, M. A. & Spain, I. M, Telepresence Control of the NASA/DARPA Robonaut on a Mobility Platform, In: Proceedings of the 2004 Conference on Human Factors in Computing Systems. ACM Press, (2004) 623–629.

[59] Stotts, D., Smith, J. M. & Gyllstrom, K. Facespace: Endoand Exo-Spatial Hypermedia in the Transparent Video Facetop. In: Proc. of the Fifteenth ACM Conf. on Hypertext & Hypermedia. ACM Press, (2004) 48–57.

[60] Smith, G. M. & Schraefel. M. C, The Radial Scroll Tool: Scrolling Support for Stylus-or Touch-Based Document Navigation, In Proc. 17th ACM Symposium on User Interface Software and Technology. ACM Press, (2004) 53–56.

[61] Freeman, W., Tanaka, K., Ohta, J. & Kyuma, K. Computer Vision for Computer Games, Tech. Rep. and International Conference on Automatic Face and Gesture Recognition, (1996).

[62] Konrad, T., Demirdjian, D. & Darrell, T. Gesture + Play: Full-Body Interaction for Virtual Environments. In: CHI ’03 Extended Abstracts on Human Factors in Computing Systems, ACM Press, (2003) 620–621.

[63] Website: University of Chicago: Mcneill Lab for Gesture and Speech Research. Electronic Resource, (2006).

[64] Valli, C. & Lucas, C. Linguistics of American Sign Language: An Introduction, Washington, D. C.: Gallaudet University Press, (2000).

[65] Martinez, A., Wilbur, B., Shay, R. & Kak, A. Purdue RVLSLLL ASL Database for Automatic Recognition of ASL, In IEEE Int. Conf. on Multimodal Interfaces, (2002) 167–172.

[66] Starner, T., Weaver, J. & Pentland, A. Real-Time American Sign Language Recognition using Desk and Wearable Computer Based Video, PAMI, 20(12) (1998) 1371–1375.

[67] Vogler, C. & Metaxas, D. A Framework for Recognizing the Simultaneous Aspects of American Sign Language, Computer Vision and Image Understanding, 81(3) (2001) 358–384.

[68] Waldron, M. Isolated ASL Sign Recognition System for Deaf Persons, IEEE Transactions on Rehabilitation Engineering, 3(3) (1995) 261–271.

[69] Dong Guo Yonghua, Vision-Based Hand Gesture Recognition for Human-Vehicle Interaction, International Conference on Control, Automation and Computer Vision, 1998

[70] Pickering, Carl A. Burnham, Keith J. Richardson, Michael J. Jaguar , A research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction, 3rd Conference on Automotive Electronics, 2007

[71] Juan P. Wachs , Helman I. Stern, Yael Edan, Michael Gillam, Jon Handler, Craig Feied, Mark Smith, A Gesture-based Tool for Sterile Browsing of Radiology Images, Journal of the American Medical Informatics Association (2008; 15:321-323, DOI 10.1197/jamia.M24)

[72] Freeman, W. T., & Weissman, C. D, Television control by hand gestures, IEEE International Workshop on Automatic Face and Gesture Recognition. Zurich.

[73] Moni, M. A. & Ali, A. B. M. S, HMM based hand gesture recognition: A review on techniques and approaches, 2nd IEEE International Conference on Computer Scienceand Information Technology, (ICCSIT 2009).

[74] Hasan, M. M., & Mishra, P. K., HSV brightness factor matching for gesture recognition system, International Journal of Image Processing (IJIP), 2010 vol. 4(5)

[75] Hasan, M. M., & Mishra, P. K., Gesture recognition using modified HSV segmentation, IEEE International Conference on Communication Systems and Network Technologies.2011

[76] Hasan, M. M., & Mishra, P. K. , Brightness factor matching for gesture recognition system using scaled normalization, International Journal of Computer Science & Information Technology (IJCSIT), 2011 3(2).

[77] Mo, S., Cheng, S., & Xing, X, Hand gesture segmentation based on improved kalman filter and TSL skin color model, International Conference on Multimedia Technology (ICMT), Hangzhou.2011

[78] Peter, H. P. (2011). Image Segmentation. (1sted.). India. (Part 2). Image Segmentation Methods Image Segmentation through Clustering Based on Natural Computing Techniques.

[79] Kim, J., & Song, M, Three dimensional gesture recognition using PCA of stereo images and modified matching algorithm, IEEE Fifth International Conference on Fuzzy Systems and Knowledge Discovery, FSKD '08, pp. 116 – 120

[80] Verma, R., & Dev A, Vision based hand gesture recognition using finite state machines and fuzzy logic, IEEE International Conference on Ultra Modern Telecommunications & Workshops, ICUMT '09, pp. 1-6.