IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

GESTURE RECOGNITION TECHNOLOGY FOR INTERPRETING HUMAN GESTURE VIA MATHEMATICAL ALGORITHMS

Ms. Divya P, Rekha K

Assistant Professor, MCA Student Scholar

Department of MCA, Nehru College of Engineering and Research Centre,

Department of MCA, Nehru College of Engineering and Research Centre, Thrissur, India

Abstract: Gesture recognition is a topic in computer science and language technology with the goal of using mathematical algorithms to analyze human gestures. Gestures can come from any physical move or state; however, they are most typically made with the hands or the face. Emotion identification from the face and hand gesture recognition are two of the field's current hot topics. Cameras and computer vision algorithms have been used in a variety of ways to read sign language. Gesture recognition techniques, on the other hand, are used to identify and recognize posture, gait, proxemics, and human behaviors. Gesture recognition can be seen of as a means for computers to begin to grasp human body language, creating a richer bridge between machines and humans than simple text user interfaces or even graphical user interfaces (graphical user interfaces).
Gesture recognition technology aims to develop a system that can recognize certain human movements and utilize them to convey information or control devices. Gestures of the human body, especially hand movements, are used to interact with computers. A camera analyses the movements of the human body and sends the information to a computer, which then uses the gestures as input to operate objects or apps.

Index Terms **-** Hand Postures, Hand Gestures, Human Computer Interaction (HCI), Segmentation, Feature Extraction, Classification Tools, Neural Networks.

1. INTRODUCTION

Gesture recognition is a topic in computer science and language technology with the goal of using mathematical algorithms to analyze human gestures. Gestures can come from any physical move or state; however, they are most typically made with the hands or the face. Gesture recognition allows humans to intuitively interact with machines (HMI) without the use of mechanical devices. It is possible to point a finger at the computer screen and have the cursor move accordingly using the notion of gesture recognition. This might render traditional input devices like mice, keyboards, and even touch screens obsolete.
Computer vision and image processing techniques can be used to recognize gestures. Gestures of the human body, especially hand movements, are used to interact with computers. A camera analyses the movements of the human body and sends the information to a computer, which then uses the gestures as input to operate objects or apps. When a person claps his hands together in front of a camera, the sound of cymbals crashing together can be produced when the motion is sent into a computer.

The use of gesture recognition to assist the physically challenged in interacting with computers, such as interpreting sign language, is one example. The technology has the potential to transform the way people interact with computers by removing input devices like joysticks, mouse, and keyboards and allowing users to communicate with computers using gestures like finger pointing.

Gesture recognition, unlike haptic interfaces, does not require the user to wear any additional equipment or attach any devices to their bodies. Instead of sensors attached to a device such as a data glove, a camera reads body gestures. Gesture recognition technology can read facial and verbal expressions (i.e., lip reading) as well as eye movements in addition to hand and body movement. In the computer vision discipline, there is ongoing research on recording gestures or more general human position and motions using cameras attached to a computer.

II. LITERATURE SURVEY

Hasan [1] used a multivariate Gaussian distribution to recognize no geometric hand movements. Two methods are used to segment the input hand image: skin colour based segmentation using the HSV colour model and clustering based thresholding algorithms [2]. The updated Direction Analysis Algorithm is utilized to identify a link between statistical parameters (variance and covariance) from the data, and used to compute object (hand) slope and trend by finding the direction of the hand motion, as shown in Figure 1.

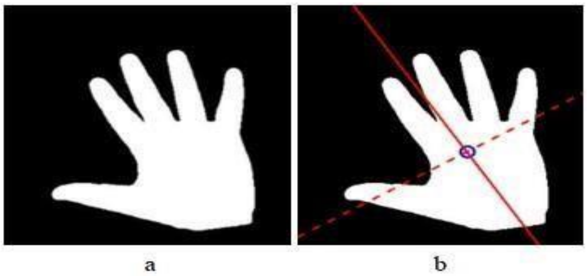

Figure1. Computing hand direction.

Then Gaussian distinction is applied on the segmented image, and it takes the direction of the hand as shown in figure 2.

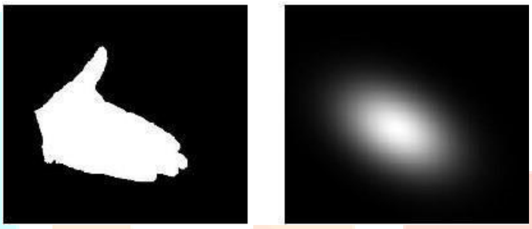

Figure 2. Gaussian distribution applied on the segmented image.

The image has been separated into circular sections, or in other words, those regions have been made in a terrace shape to eliminate the rotating effect. The shape is broken into 11 terraces, each of which is 0.1 width. The 0.1 width division produces nine terraces: (1-0.9, 0.9-0.8, 0.8-0.7, 0.7-0.6, 0.6, 0.5, 0.5-0.4, 0.4-0.3, 0.3-0.2, 0.2-0.1), and one terrace for the terrace with a value less than 0.1 and the last terrace for the external region that expanded out of the outer terrace [1][2]. Figure 3 shows how this divide is explained.

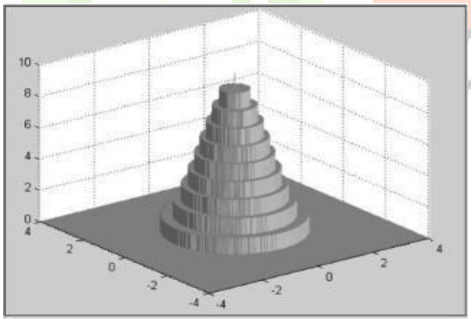

Figure 3. Terraces division with 0.1 likelihood.

Each terrace is divided into eight feature sections, with the number eight being found to be the most ideal for feature divisions. Re-estimation is conducted on the shape to fit capturing the hand object, then the Gaussian shape is matched on the segmented hand to prepare the final hand shape for extracting the features [1].

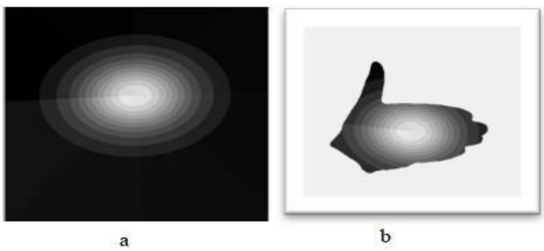

Figure 4. Features divisions. a) Terrace area in Gaussian. b) Terrace area in hand image.

To build the feature vector, two types of features are retrieved after capturing the hand shape: local features and global features. Local features based on geometric central moments (00,11), as indicated in equation (1):

$$\mu_{pp} = \sum_{x}\sum_{y}(x-\mu_x)^p(y-\mu_y)^p f(x,y) \quad (1)$$

$$\mu_{pp}^{(k)} = \sum_{y}\sum_{x}\left(x^{(k)}-\mu_x^{(k)}\right)^p\left(y^{(k)}-\mu_y^{(k)}\right)^p f\left(x^{(k)},y^{(k)}\right) \quad (2)$$

$$\forall\, k \in \{1, 2, 3, \ldots, 88\} \;\&\; \forall\, p \in \{0, 1\}$$

Where x and y are the coordinated, mean values for the input feature area, and

As detailed in equation, the input image is represented by 88*2 features [1]. While the global features are two features, the first and second moments that are computed for the entire hand features area, the global features are two features, the first and second moments. These feature areas are calculated by multiplying the intensity of the feature area with the position of the feature area on the map. Any input image has 178 features in this situation. The system used 20 different gestures, 10 samples for each gesture, 5 samples for training and 5 samples for testing, and had a 100% identification rate when the number of gestures was more than 14 gestures. Recognizes six movements with ten samples for each gesture. The feature was classified using the Euclidian distance.

Hasan [5] used scaled normalization to recognize gestures using brightness factor matching. The input image is segmented using the thresholding approach with a black background. Any segmented image is normalized (trimmed), and the image's Centre mass is computed, with the coordinates altered to match the centroid of the hand object at the X and Y axis origin. Because this method is based on the object's central mass, the resulting images are of varying sizes. As shown in Figure 9, a scaled normalization procedure is used to solve this problem while maintaining image dimensions and time, with each block of the four blocks scaling with a factor that is distinct from the factors of the other blocks. The features are extracted using two methods: first, edge mages, and second, normalized features, in which only the brightness values of pixels are evaluated and other black pixels are ignored to shorten the feature vector. The database contains six different gestures, with ten samples per gesture, five for training and five for testing. The recognition rate for the normalized feature problem was higher than the recognition rate for the normal feature problem, with 95 % for the former and 84% for the letter [5].

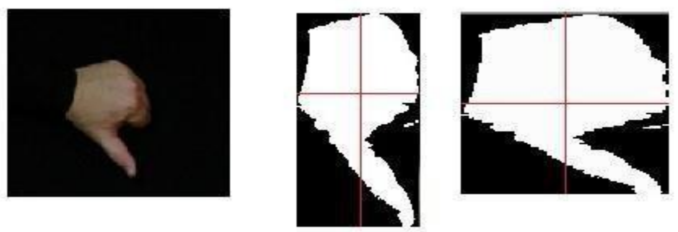

Figure 5. applying trimming process on the input image, followed by scaling normalization process.

Using a boundary histogram, Wysoski et al presented rotation invariant postures [8]. The input image was captured using a camera, and a filter for skin colour detection was applied, followed by a clustering procedure that employed the ordinary contour tracking algorithm to locate the border for each group in the clustered image. Grids were created from the image, and the boundaries were standardized. By splitting the image into number of regions N in a radial form, according to certain angle, the boundary was represented as chord's size chain, which was used as histograms. MLP Neural Networks and Dynamic Programming DP Matching were employed in the classification procedure. Many tests have been carried out using various feature formats, as well as different chord size histograms and chord size FFT. The investigations used 26 static postures from American Sign Language. The work was done on a homogeneous background. For hand gesture recognition, Stergiopoulou proposed a new Self-Growing and Self-Organized Neural Gas (SGONG) network. A colour segmentation technique based on a skin colour filter in the YCbCr colour space was used for hand region detection, and an approximation of hand shape morphology was detected using the (SGONG) network. Three features were extracted using the finger identification process, which determines the number of raised fingers and hand shape characteristics, and a Gaussian distribution model was used for recognition.

III. OBJECTIVE

The purpose of static gesture recognition is to categorize the hand gesture data, which is represented by various attributes, into a finite number of gesture classes. The major goal of this project is to investigate the utility of two feature extraction methods, namely hand contour and complex moments, in solving the problem of hand gesture recognition by identifying the fundamental benefits and drawbacks of each method. The back-propagation learning algorithm is used to create an artificial neural network for classification. The suggested system includes a recognition algorithm that can distinguish six different static hand gestures: Open, Close, Cut, Paste, Maximize, and Minimize. Pre-processing, feature extraction, and classification are the three processes that the gesture image goes through. Some operations are performed in the pre-processing stage to remove the hand gesture from its context and prepare the hand gesture image for feature extraction. The hand contour is employed as a feature in the first method to address scaling and translation issues (in some cases). In addition to scaling and translation, the complex moments algorithm is employed to define the hand motion and treat the rotation problem. The multi-layer neural network classifier uses the back-propagation learning process. The results demonstrate that the hand contour method has a recognition rate of (71.30 percent), whereas complicated moments have a higher recognition rate of (86.90 percent).

IV. RESEARCH METHODOLOGY

4.1 Extraction Methods and Extraction Features

After obtaining the input image from camera(s), movies, or even a data glove instrumented device, most studies divided gesture recognition systems into three processes. These are the steps: Figure 6 shows the extraction method, feature estimation and extraction, and classification or recognition.

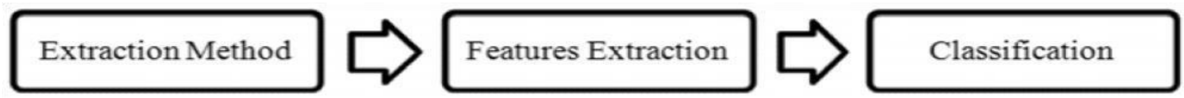

Figure 6. Steps in the gesture recognition system.

4.2 Extraction Method and image pre-processing

The first step in recognizing hand motions is to segment them. It is the process of splitting an input image (in this case, a hand gesture image) into sections with distinct borders. The segmentation procedure is determined by the type of gesture; if it is a dynamic gesture, the hand gesture must be found and tracked; if it is a static gesture (posture), simply the input image must be segmented. First, the hand must be located; generally, a bounding box is used to specify the location depending on the skin colour; second, the hand must be tracked; there are two main approaches for tracking the hand: either the video is divided into frames and each frame must be processed separately; in this case, the hand frame is treated as a posture and segmented; or using some tracking information such as shape, skin colour, and some tools such as Kalman filter.

The skin colour is a typical helpful indication for segmenting the hand since it is simple and invariant to size, translation, and rotation changes. To simulate the hand, many tools and methodologies employed skin and non-skin pixels. These techniques are parametric and non-parametric; parametric techniques include the Gaussian Model (GM) and Gaussian Mixture Model (GMM), while nonparametric techniques include histogram-based techniques. However, different races are affected by variations in lighting conditions. Some studies use a data glove and coloured markers to solve this difficulty, which provide precise information about the orientation and position of the palm and fingers. Others used infrared cameras and range data from specialized cameras. Although TimeofFlight (ToF) cameras can identify different skin colours against a cluttered background, they are influenced by changes in temperature degrees, in addition to their high cost. Segmentation is regarded as a problem in and of itself. The colour space utilised in a particular application is critical to the effectiveness of the segmentation process; however, colour spaces are sensitive to illumination changes, therefore most studies focus on chrominance components solely, ignoring luminance components such as r-g and HS colour spaces. Complex backgrounds, lighting fluctuations, and inadequate video quality are some of the variables that obstruct the segmentation process.

YCbCr colour space was utilized with the HSV colour model, which focuses on the pigments of the pixel. The colour space was normalized r-g. To improve the segmented hand picture, preprocessing techniques such as subtraction, edge detection, and normalization are used. Examples of segmentation methods are shown in Figure 2.

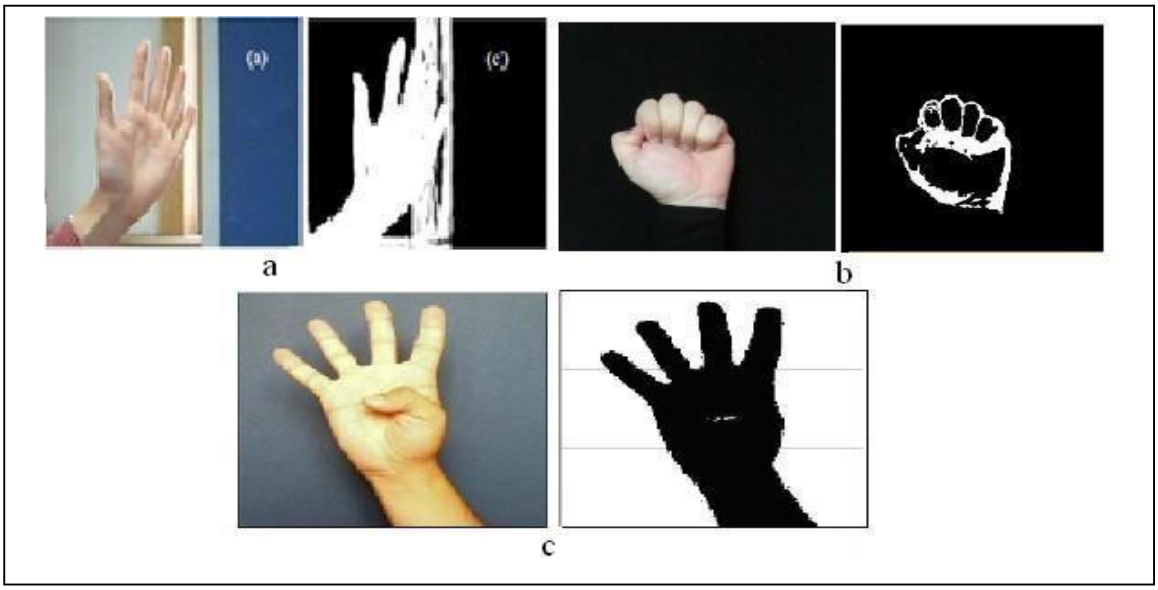

Figure 7. Segmentation method

4.3 Features Extraction

A successful recognition process begins with a good segmentation process, which leads to a great feature extraction procedure. The segmented image's features vector can be extracted in a variety of methods, depending on the application. Various approaches for representing and extracting features have been used. Hand contour and silhouette were used in some ways, whereas fingertips position, palm center, and other factors were used in others. The first parameter reflects the ratio aspect of the bounding box of the hand, while the next 12 parameters represent mean values of brightness pixels in the image. Calculated the segmented hand's Center Of Gravity (COG) and the distance between the COG and the furthest point in the fingers, then extracted a single binary signal (1D) to estimate the number of fingers in the hand region. The image was divided into multiple size blocks, each representing the brightness measurements in the image. Many studies were conducted in order to determine the best block size for achieving a high recognition rate. Geometric central moment was extracted as local and global features using Gaussian pdf. Figure 8 depicts various examples of feature extraction algorithms in operation.

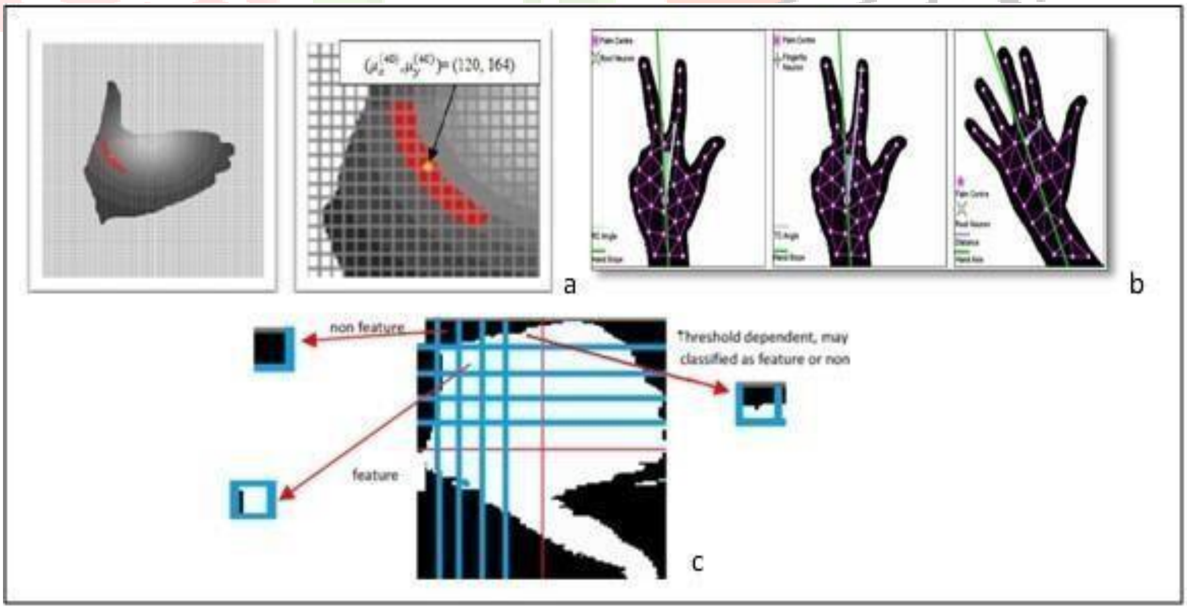

Figure 8. features representation. a) The segmented image is partitioned into 11 terraces with 8 regions per terrace to extract local and global geometric central moment. b) Three angles are extracted: RC angle, TC angle, and distance from the palm center. c) Segmented hand divided into blocks and the brightness factor for each block represents the feature vector (blocks with black area are discarded).

4.4 Classification of Gestures

To recognize the gesture, the gesture classification approach is employed after modelling and analysis of the input hand image. The right selection of feature parameters and a good classification method affect the recognition process. Edge detection or contour operators, for example, cannot be employed for gesture identification since numerous different hand postures are generated, resulting in misclassification. The gestures were classified using the Euclidean distance metric. Among the statistical techniques used for gesture classification, the HMM tool, in addition to the Finite State Machine (FSM), Learning Vector Quantization, and Principal Component Analysis (PCA), has demonstrated its ability to distinguish dynamic movements. The neural network has been widely used in the field of hand form extraction and hand gesture recognition. Other soft computing technologies, such as Fuzzy CMeans clustering (FCM) and Genetic Algorithms GAs, are also useful in this sector.

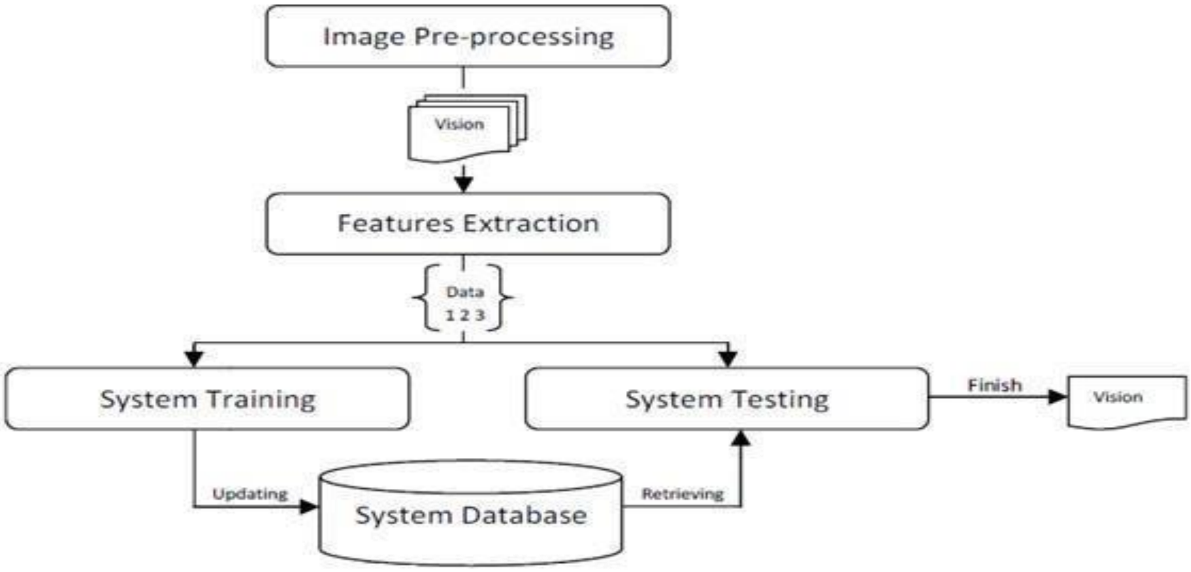

Figure 9. Architecture of gesture recognition system.

V. RESULT AND DISCUSSION

Some hand gesture recognition systems are summarized in the tables below. A comparison of recognition methods in hand gesture recognition systems is shown in Table 1. Table 2 summarizes some hand gesture recognition systems' application areas and invariant vectors. Table 3 shows the hand extraction technique, features vector representation, and recognition employed in the selected hand gesture recognition systems, as well as a summary of the extraction method, features representation, and recognition.

Method	**Recognized Gestures**	**Total number of gestures used in training and testing**	**Percentage of Recognition**	**utilized database**
1	26	1040	DP 98.8% MLP 98.7%	American Sign Language (ASL)
2	6	60	normal method 84% Scaling normalization method 95%	Own Database
3	26	208	92.78%	American Sign Language (ASL)

4	0-9 numbers	270 video sequence for continuous motions/ 298 video sequence for isolated gestures	90.45%	Recognize Arabic numbers from 0 to 9.
5	5 static/ 12 dynamic gestures	A total of 240 data points are taught and then tested.	98.3%	5 static gestures and 12 dynamic gestures.
6	31	130 for testing	90.45%	Own Database
7	6	60	100% for more than 4 gestures	Own Database
8	20	200	For 14 gestures, 100%, and for 15-20 gestures, >90%.	Own Database

Table1.Comparison of several hand gesture recognition algorithms.

Method	Application Area	Invariant factor
1	Real time system / control a computer graphic crane by hand gestures/ play games such as scissors/paper/stone	Lighting conditions / Translation
2	Sign Recognition	Lighting conditions / Translation
3	Sign Recognition	Rotation
4	Sign language	Rotation/ Translation/ Scaling
5	Robot control application	Translation/ Rotation / Scaling
6	Real-time system/ moderate computational resources devices e.g. netbooks	Rotation / Translation
7	Sign Recognition	Rotation/Translation/ Scaling

8	Sign Recognition	Rotation/Translation/ Scaling
9	Drawing graphical elements such as triangle, rectangular/ Editing graphical elements such as copy, paste, undo/ Mobile robot control/Virtual Reality.	-

Table 2. Some hand gesture recognition systems have an invariant vector and a summary of application areas.

Method	**Extraction method**	**Features Vector Representation**	**Classifier**
1	HSV color space	The first component in the feature vector indicates the ratio aspect of the bounding hand box, while the next 12 parameters provide the mean values of brightness pixels in the image.	Fuzzy C-Means (FCM) algorithm
2	-	Boundary Chord's size FFT	MLP Neural Network/ Dynamic Programming (DP) matching
		Boundary Chord's size	
		Boundary Chord's size histogram	
3	YCbCr color space	Three angles of the hand shape were determined using the SelfGrowing and Self-Organized Neural acquired hand form: RC Angle, TC Angle, and Distance from the palm center.	Gaussian distribution
4	Colored glove/ HSV and threshold based method	Distances from the palm to all fingers comprise nine numerical features, with four angles between them.	Learning Vector Quantization (LVQ)
5	GMM for skin	Orientation quantization	HMM
	color Detection and YCbCr color space		
6	Thresholding technique	13 data points (10 for bending, 3 for angles)/ 16 data points (10 for bending, 3 for angles)	Back propagation network / Elman recurrent network
7	Thresholding	Divide the normalized scaled hand picture into intensity feature Blocks.	Euclidean Distance Metric.

8	HSV color space/thresholding	178 features, for local and global features using moments	Euclidean Distance Metric.

Table 3. Summary of the hand gesture extraction method, features representation, and recognition of hand gesture recognition systems

VI. CONCLUSION

Various methods for gesture detection are explored in this paper, including Neural Networks, HMMs, fuzzy c-means clustering, and the use of an orientation histogram for feature representation. HMM tools are ideal for dynamic motions and have demonstrated their effectiveness, particularly in robot control. NNs are employed as a classifier and to capture the contour of the hands. Some methods and algorithms are necessary for feature extraction, even to capture the contour of the hand, such as the application of a Gaussian bivariate function for fitting the segmented hand, which was utilized to reduce the rotation affection. The recognition algorithm that is chosen is determined by the application. The application areas for the gestures system are presented in this paper.

REFERENCES

[1] Mokhar M. Hasan, Pramod K. Mishra, (2012) "Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition", International Journal of Computer Science & Emerging

[2] Mokhar M. Hasan, Pramod K. Mishra, (2012). "Robust Gesture Recognition Using Gaussian Distribution for Features Fitting', International Journal of Machine Learning and Computing, Vol. 2(3)

[3] S. Mitra, and T. Acharya. (2007). "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.

[4] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "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.

[5] Joseph J. LaViola Jr., (1999). "A Survey of Hand Posture and Gesture Recognition Techniques and Technology", Master Thesis, Science and Technology Center for Computer Graphics and Scientific Visualization, USA.

[6] Rafiqul Z. Khan, Noor A. Ibraheem, (2012). "Survey on Gesture Recognition for Hand Image Postures", International Journal of Computer And Information Science, Vol. 5(3), Doi: 10.5539/cis.v5n3p110

[7] Thomas B. Moeslund and Erik Granum, (2001). "A Survey of Computer Vision-Based Human Motion Capture," Elsevier, Computer Vision and Image Understanding, Vol. 81, pp. 231–268.

[8] Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "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.