A Real Time Hand Gesture Recognition Using Feature Extraction

Shruti Wale
PG Student
Electronics & Telecommunication
Engineering
PHCET, Rasayani, India
Email: wale.shruti@gmail.com

Swapnil Patil
PG Student
Computer Engineering
PHCET, Rasayani, India
Email: swapnilpatil89@gmail.com

Prof. Neha Rai
Assistant Professor
Dept. of Electronics & Telecomm.
Engineering
PHCET, Rasayani, India
Email: nrai@mes.ac.in

Abstract - Gesture recognition is finding a meaning attached to some motions or actions. The main objective is of this paper is hand gesture recognition from the real time video. Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures by using some mathematical algorithms. Gestures are commonly originated from the face or hand. Current focuses in the field include emotion recognition from face and hand gesture recognition. Many approaches have been made using cameras and computer vision algorithms to interpret sign language. However, the recognition of posture and human behaviours is also related to gesture recognition techniques. The major problem faced in the gesture recognition is that a person will appear at different scales in videos. The movement of camera is another problem as the person holding it may shake it while shooting the video. In certain cases the camera may mounted on something which moves with the person performing the action. The another problem faced is occlusions, background clutter, human variation and action variation .Gesture recognition also various with person to person because every person has different skin texture and color. The non-parametric histogram based on RGB model is used for the skin detection. In this system the gesture recognition is mainly divided into two phases: training phase and recognition phase. Feature extraction is the main function of both the phases. The fuzzy logic is used in this system for gesture recognition.

Index Terms – Recognition phase, Training phase, Fuzzy logic, RGB model

_____*****_____

I. INTRODUCTION

The Gesture acknowledgment interface goes about as a correspondence channel amongst people and machines. The human-machine collaboration is like human-human connection, in which, the important data are conveyed utilizing the human organs like hand motion, head development, face demeanor, voice correspondence and general body stance. The configuration of a motion acknowledgment framework depends on basic equipment backing, for example, web-cams or portable incorporated cameras, to be pertinent to current PCs, cell phones, Digital Cameras, and so forth. While outlining the frameworks, certain parameters, for example, that it ought to be work under complex or non-uniform foundation, i.e., distinctive light force and loud environment, and so on are taken into contemplations. This anticipate presents a constant hand signal acknowledgment framework by utilizing webcam. This procedure includes recognizing the hand area, investigating the hand-position varieties. At that point the got movement data is been utilized as a part of the acknowledgment period of the motion.

II. LITERATURE SURVEY

We have concentrated on numerous past works done in this field by various scientists. There are numerous methodologies that were trailed by various scientists like vision based, information glove based, Artificial Neural Network, Fuzzy Logic, Genetic Algorithm, Hidden Markov Model, Support Vector Machines and so forth. A percentage of the past works are given beneath. Numerous specialists

[1][2][3][4][5][6][7][8][9][10][11] utilized Vision based methodologies for distinguishing hand motions First the skin hued district from the information picture caught and after that this picture with sought hand locale was power standardized and histogram was discovered for the same[1]. Highlight extraction step was performed utilizing Hit-Miss Transform and the motion was perceived utilizing Hidden Markov Model (HMM). Acknowledgment rate acquired was 98%. In [2] utilized YCbCr shading model to recognize skin hued pixels from the foundation. The required part of the hand was separated utilizing this shading display and sifted utilizing middle channel and smoothing channel. The edges were distinguished and highlights removed were hand border, viewpoint proportion, hand range after which Artificial Neural Network (ANN) was utilized as classifier to perceive a motion. Exactness rate acquired was 97.4%. In [3][8] fingertip identification was utilized for hand motion acknowledgment. In[3] this paper for gesture based communication acknowledgment initially changed over the RGB picture caught to paired and Canny Edge Detection Technique was utilized for removing edge of the palm. The fingertip positions of the fingers were distinguished from the extricated edge of palm by measuring their separation from a reference point which is taken to be at the base of the palm. Acknowledgment rate acquired was 98.125%. In this paper [8] examined the skin separated picture in all heading to discover the edges of the fingers and the tips of the edges were alloted the most noteworthy pixel esteem and in that capacity fingertip was recognized. In [4] utilized hand motion acknowledgment for

controlling the robot. The Red/Green proportion was discovered which was utilized for deciding the skin shaded areas. The focal point of gravity of the hand was discovered alongside the most distant separation from it and in this manner in such a way the fingertips were resolved. A circle was made around the focal point of gravity and number of white pixels past that circle was checked to know the fancied motion. Acknowledgment rate acquired was 91%. In [5] utilized the same strides as [4] aside from that the RGB info was changed over to HSV shading space before experiencing further steps. Just about 100% exactness was gotten. In [6] found the skin shaded pixels from the picture, after which the elements like introduction, spatial recurrence, spatial territory were removed for which Gabor channel and Principle part analysis(PCA) was utilized. Bolster vector machines (SVM) was utilized as classifier for this paper.Recognition rate acquired was 95.2%. In [7] utilized an additional progression toward the starting that is the Active Appearance model which considers the shape and shade of the picture. This model discovers the harsh shape of the hand. Acknowledgment rate got was 82.6%. In [9] changed over info picture to YCbCr skin shading model, utilizing appropriate limit technique craved hand part was extricated from the information picture. Guideline Curvature Based Region indicator (PCBR) and 2-D Wavelet Packet Decomposition (WPD) procedures were utilized for highlight extraction. Acknowledgment Rate got was 91.3% for static and 86.3% for element. In [10] found the skin shaded pixels and districts relating to such pixels were edited out. The introduction of the picture was discovered utilizing PCA. Highlights removed were distance,angle by which every focuses on the form was identified with each other by IDSC (Inner separation shape context)algorithm lastly motion was perceived by SVM. In [11] connected Adaptive Boost calculation for identifying hand from the information picture. The primary favorable position of utilizing this calculation is, it could distinguish a solitary hand as well as could identify covered hand. Highlights extricated were palm and finger structures which were controlled by drawing blobs and edges. Acknowledgment rate acquired was 98%. In [12][13][15][16] Data-Glove Based Approaches was utilized. In this paper method [12] utilized is DG5 VHand 2.0 information gloves for hand motion acknowledgment. The elements like position of fingers, hand was given by the gloves after which KNN classifier calculation was utilized. This is utilized it as a part of a few applications like air composing and picture program. In this paper the technique[13] utilized is KHU-1 information glove alongside Kinematic affix hypothesis to extricate the components like joints from hand. Glove comprised of 3 accelerometer sensor, a controller and a Bluetooth.Finally, run based calculation was utilized for motion recognizable proof. Some[15] utilized Cyberglove that considered points made by 18 joints of hand. Highlights extricated utilizing this glove were edge made between the neighboring fingers, wrist pitch, thumb revolution which was then prepared utilizing ANN. Another method [16] utilized glove where 18 markers were appended with it, of which 15 were for fingers and 3 for the reference taken. The picture caught was then arranged taking into account Singular quality disintegration (SVD).In [17][18][19][24] ANN based

framework was proposed for perceiving the signal. It was utilized in view of beneficial components like all inclusive statement, versatile learning, self arranging and ongoing operations. The ANN[17] is utilized to perceive American Sign Language. Firstly, the Skin shaded locales were extricated after which the minute invariant was acquired. ANN was utilized where the system has 58 neurons of which 34 were information neurons, 20 concealed neurons and 4 yield neurons and the dataset included of270 highlight vectors. In [18] utilized ANN for perceiving the Ethiopian Sign dialect of 34 letters. The neural system comprised of three layers, info layer containing 200 neurons giving the quantity of highlight vectors, yield layer with 34 neurons depicting the quantity of classifications which was to be perceived and shrouded layer with 100 neurons. Back proliferation calculation was utilized for preparing. Acknowledgment rate accomplished was 98.53%.In [19] utilized ANN for perceiving the hand signal utilized for insightful human-machine interface. The neural system involved 2 layers with 25 neurons, the shrouded layer with 20 neurons and the yield layer with 5 neurons. It was handled and prepared by Back spread technique alongside 1500x5 vector dataset.In [24] utilized ANN with back engendering calculation which had 20 or 24 neurons in information layer, 42 neurons each in the covered up and yield layer. The system after prepared by an info vector gave a solitary yield neuron giving the fancied acknowledgment. In [20][21] Genetic Algorithms (GA) was utilized for taking care of issues as a part of which steps were determination of guardian information, recombination and transformation. The explanation behind its regular use is it helps in getting ideal answers for the problem.In [20] utilized the above calculation as a part of perceiving Indian Sign Language and said that the arbitrariness of the specimens which was taken at the information can be overseen legitimately utilizing this methodology yet they could perceive just the couple of letter sets of Indian Sign Language which was one of the inconvenience. In [21] utilized GA to separate strong fingertip for connecting with the robot. YCbCr shading model was utilized yet just to foresee the likelihood of whether the pixels are skin hued or not. After the forecast, a critical step comes that is the limit choice which in this paper is done utilizing GA. Amid this calculation, beginning populace was 8 bits and capacity deciding the wellness was accomplished utilizing Otsu's method.Fuzzy rationale is a critical thinking approach in view of degrees of truth as opposed to the standard genuine or false i.e. 1 or 0. It incorporates 0 and 1 as amazing instances of truth furthermore incorporates the different conditions of truth in the middle. For instance insight can't be measured with ordinary 1 or 0. It must be contrasted and other's knowledge and result can be 0, 1or in the middle. In [22] said that the movement of the hand can be recognized by Finite state machine's (FSM) states. These states are accepted as bunches which are in reality framed by fluffy c-implies grouping. At that point the centroid of every bunches discovered numerically, and henceforth conditions of FSM was resolved lastly motion was recognized. In [23] utilized Fuzzy rationale for perceiving Korean Sign Language. With time, the position and speed of hand changes, these distinctive rates were thought to be the fluffy sets which were demonstrated scientifically as

401

_____

zero, little, medium, substantial, etc.In [24][25] Principal Component Analysis (PCA) was executed in their works. In [24] utilized PCA for separating highlights from the info picture in which mean, covariance, Eigen qualities and Eigen vector were discovered out.Mean depicting the position of finger, Eigen esteem portraying the state of the finger, Eigen vector demonstrating the bearing of the picture were the components utilized. In [25] in his paper utilized PCA as a classifier where Eigen face was removed from the picture to be tried and after that the Euclidean separation was found between classes.In [26][14] Hidden Markov Model (HMM) was utilized for hand signal acknowledgment as a part of various fields. In [26] utilized HMM for recognizing signal while connecting with the robot. Highlights separated like state of hand, optical stream were sustained to the HMM after which certain frameworks like state move networks were gotten which helped in perceiving the signal when a self-assertive hand was nourished. In [14] utilized HMM to perceive letters in order while hand in movement. Highlights extricated in this paper were area, speed whose vector was then sustained to the HMM. Left-Right Banded model alongside Baum-Welch calculation was utilized for perceiving the signal. Acknowledgment rate acquired was 92.3%

III. PROPOSED METHOD

This project basically deals with the design of a system that acquires a hand gesture and classifies it based on the predefined hand gestures, stored in a database. The system is mainly divided into two phases namely, training phase and recognition phase. The figure 1 shown below is the list of gestures that can be recognised by the system.
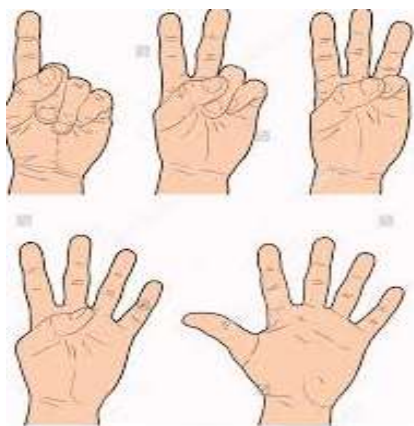

Fig. 1Gesture representing Numbers

The system uses web cam for capturing the hand gestures and an algorithm that processes the acquired images and then classifies the hand gesture correctly. The work mainly highlights on the feature extraction from the hand gestures and use that features in the recognition algorithms. Firstly, features are extracted in the training phase by using the suitable algorithms for different hand gestures. Once this feature extraction is completed in training phase the system is able to classify the given hand gesture based on the gained knowledge during training phase. The efficiency of the recognition phase totally depends on the training phase. The better the task performed in training phase, better will be the performance of

recognition phase. So, all the functions in training phase are carried out properly. The figure 2 shown below are the list of functions for training phase:
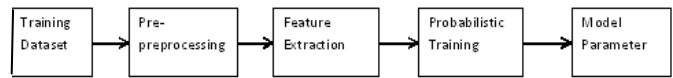

Fig. 2 Training Phase

The main purpose of the training phase is as follows:
1. Hand gesture is extracted from the input image.
2. Removal of noise and unwanted regions by suitable algorithms.
3. Process the extracted image to form a binary image and
4. Feature are extracted from the processed image to form a feature set or training dataset for classification.

In training phase,the dataset consists of hand gesture representing specific numbers between 1 to 5 with unique binary pattern.

The figure 3 shown below are the list of task for training phase:
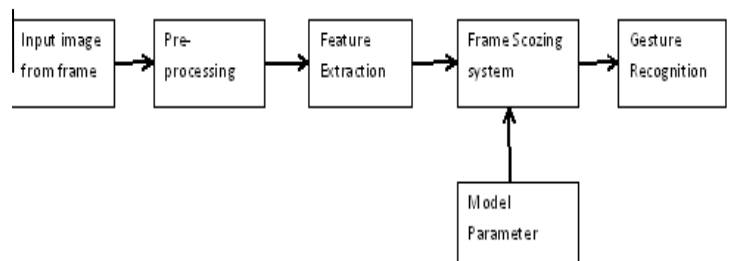

Fig. 3 Recognition Phase

As examined before in writing study, there are different techniques for grouping of signal that are as of now utilized. They are Gradient, PCA (Principal Component Analysis) and SVM (Support Vector Machine). In any case, this paper essentially concentrates on fluffy rationale for motion acknowledgment. The customary strategy is followed in preparing stage for highlight extraction reason from info picture. The commotion free picture is gotten with noteworthy components from various hand signals are extricated and spoke to in the grouping calculation. The exactness of arrangement relies on upon the quantity of components incorporated into the calculation.

After the picture securing just the hand bit is extricated from unpredictable or uniform foundation or uniform foundation with help of skin shade of hand with the assistance of non prametric histogram RGB model and change over the resultant picture to paired picture. The thresholding procedure was utilized to change over picture into twofold picture. . As framework is working progressively, there exists some clamor in a picture. The commotion is essentially because of low determination cameras that it utilized amid picture securing step or because of the natural factors.Therefore, the picture containing the extricated hand divide additionally contain some noises.So, the sifting strategies are utilized trailed by

402

_____

disintegration morphological operations to expel the clamor and undesirable little sharp detail in an image.The consequence of this operation is sans commotion picture with intrigued range as it were. . At last, the resultant picture is handled to decide the fingertips and comparing co-ordinate qualities are stored.The system used to discover the fingertips is skeletonization. The centroid of a hand locale is figured in next step. For deciding the effectively raised finger,the centroid computation is utilized. Outwardly, it can be resolved that the separation amongst centroid and the distinctive fingers of a hand are distinctive. Thus, this idea is utilized to make a twofold example for effectively brought finger up in a picture. The dynamic raised finger is stamped 1 and the in-dynamic fingers is checked 0. These diverse mixes of 0's and 1's speak to any one numbers between 1 to 5 . In this way, a database is made for various hand signal that speaks to numbers. The deviation of 10% to 20% is thought about in these framework for separation to decide effectively raised finger.

IV. EXPERIMENTAL SETUP

As examined before, the configuration of framework is divded into essentially into two parts:training pahse and acknowledgment stage. The vital part which plays for the arrangement of signal is preparing stage. Firstly, the picture securing is done in preparing stage which stores the hand part region in dataset. The imperatives mulled over amid picture procurement procedure is that the pictures for preparing and also testing were taken at equidistant level. The second step in preparing stage is to separate the hand parcel just from uniform or non-uniform foundation. Along these lines, for extricating the skin part in the acquired picture RGB model method is utilized. Also, yield got from the RGB displayed the relating paired picture is controlled by utilizing non-parametric histogram method.

To diminish the unpredictability of highlight extraction for hand motions, the yield picture of hand bit extraction procedure was changed over into twofold picture utilizing a non parametric histogram system technique.The picture may contain some noises.The nearness of such clamors is because of the be because of the environmental condition at which the pictures were taken furthermore the sort of source that is utilized for catching an image.For evacuation of commotion the middle channels are utilized. The morphological administrator specifically picture disintegration can likewise be utilized for expulsion of little sharp undesirable points of interest (i.e., commotion) from a picture contingent on the rate of clamor present in the binarised picture. The rightness of information hand motions relies on upon the degree for which clamor expulsion is straightforwardly corresponding to the degree to which a framework can be prepared accurately. There are different various types of recognized elements that can be removed from the separated picture, however the paper fundamentally concentrates on the dynamic and in-dynamic finger which is spoken to by 1 and 0 individually. From the fingertips the dynamic and latent fingers are determined.Skeletonization method is utilized for recognizing fingertips. In this strategy, the diminished picture was gotten that speaks to the finger partition just by separately associated

pixels, the co-ordinate values for the fingertip was put away. The pixel is considered as a fingertip that has stand out neighbor in 3x3 window mask.The particular co-ordinate values for dynamic fingers utilizing pixels.Now,the next stride subsequent to putting away the co-ordinate esteem for dynamic fingers is to discover the centroid of a hand. In the following step, the separation between the centroid and the fingertip is computed. For separation, the Euclidean separation is figured further.

Regardless of the fact that there is a little bents of finger in the motion to be perceived, a specific level of deviation is incorporated into the separation parameter to build the proficiency of the system.As the separation from centroid to various fingertip are different,this critical component can be utilized for distinguishing the dynamic or in-dynamic finger. The dynamic and in-dynamic finger are spoken to in paired code after the blend of various dynamic finger which are discovered taking into account separation from centroid. The dynamic fingers were spoken to by 1 and the in-dynamic finger was spoken to by 0. In this manner, for various types of signal speaking to various number, there will be an alternate mix of 0's and 1's. For instance, the double code for signal 1 is will be 01000, in which, the second finger is the main dynamic finger and rest of the fingers are in-dynamic. To keep up consistency for twofold representation of fingers, the numbering of finger is done frame left to right.

A model parameter is made for various motions by utilizing this technique for speaking to the hand motion that is as paired code.There will an alternate blend of twofold numbers for the diverse hand motion speaking to various numbers.Thus, the model will contain the discernable parallel pattern.Hence, the framework is prepared in view of these double examples.

With a specific end goal to separate the double example for the info hand signal to be tried in acknowledgment stage same steps are taken after. Furthermore, the separated parallel example for every info hand signal is contrasted and the substance of model parameter made in the preparation stage. At long last, if there is match of the information parallel example with any of the double example in model paraemeter, the file of the coordinated paired example from model parameter is shown which is considered as the coordinated motion that speaks to a number between 1 to 5 by the framework. The Binary yield is utilized for arranging, in which, the info picture is named either right or inaccurate. Along these lines, number of double yield correlation must be made with model parameter. Wherever there is a match with any double example in model parameter, the framework will show its comparing perceived motion number.

V. RESULTS

The result of the system is very simple, but the paper aims at the efficient way of representing data or information for comparison and focuses on the method to reduce the complexity during comparison for a hand gesture image to be tested and hence to increase the correctness of the system. The figure 4 shown below is the abstract view of result of the system.
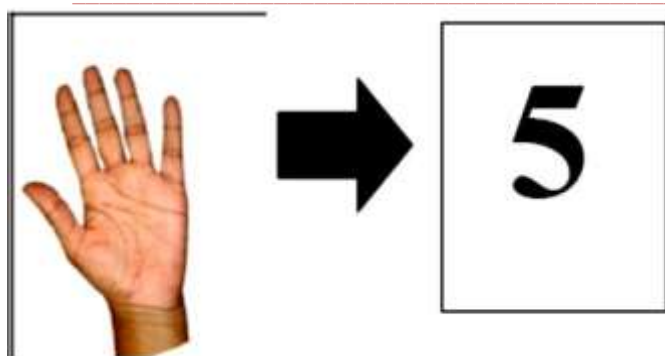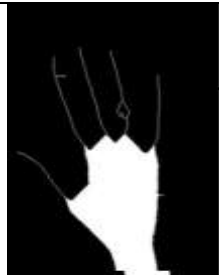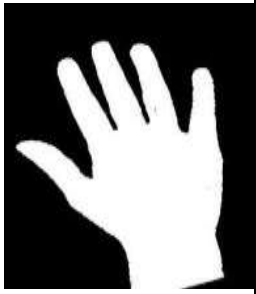
Fig. 4 Input Gesture and Output Gesture

Model Parameter		**11111**
Comparison	**11111**	5

The various steps are carried out to design a system that will correctly classify and give the above sample output for its corresponding input. The various steps that are carried out are as follows:

Table 1 Sample Results for Number 5

Process	Input	Output
RGB to Binary		
Noise Removal		
Skeletoniz-ation		

VI. CONCLUSION

The real time hand recognition system is designed by using non parametric histogram RGB skin detection & fuzzy type logic. The designed system is able to classify hand gesture in videos. Firstly, in training phase the dataset is created and labelled as a modelled parameter. In the recognition phase, the image is captured and gesture is recognised by comparing the results with the dataset in training phase. The different gestures are classified correctly. The system is able to classify successfully the hand gesture representing number. In future, work can be further extending to recognize alphabets, expressions, etc. And also it can be enhanced to increase the performance of the system.

REFERENCES

[1] T. Kapuscinski and M. Wysocki, "Hand Gesture Recognition for Man-Machine interaction", Second Workshop on Robot Motion and Control, October 18-20, 2001, pp. 91-96

[2] C. Yu, X. Wang, H. Huang, J. Shen and K. Wu, "Vision-Based Hand Gesture Recognition Using Combinational Features", IEEE Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 543-546.

[3] P.S. Rajam and G. Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf-dumb People", IEEE, 2011, pp. 737-742.

[4] A. Malima, E. Ozgur and M. Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", IEEE, 2006.

[5] Manigandan M and I.M Jackin, "Wireless Vision based Mobile Robotcontrol using Hand Gesture Recognition through Perceptual Color Space", IEEE International Conference on Advances in Computer Engineering, 2010, pp. 95-99.

[6] D.Y. Huang, W.C. Hu and S.H. Chang, "Vision-based Hand GestureRecognition Using PCA+Gabor Filters and SVM", IEEE Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 1-4.

[7] E. Koh, J. Won and C. Bae, "On-premise Skin Color Modeing Method for Vision-based Hand Tracking", The 13th IEEE International Symposium on Consumer Electronics (ISCE), 2009, pp. 908-909.

404

[8] J.L. Raheja, K. Das and A. Chaudhury, "An Efficient Real Time Method of Fingertip Detection", International Conference on Trends in Industrial Measurements and automation (TIMA), 2011, pp. 447-450.

[9] J. Rekha, J. Bhattacharya and S. Majumder, "Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition", IEEE, 2011, pp. 30-35.

[10] R. Gopalan and B. Dariush, "Towards a Vision Based Hand Gesture Interface for Robotic Grasping", The IEEE/RSJ International Conference on Intelligent Robots and Systems, October 11-15, 2009, St. Louis, USA, pp. 1452-1459.

[11] Y. Fang, K. Wang, J. Cheng and H. Lu, "A Real-Time Hand Gesture Recognition Method", IEEE ICME, 2007, pp. 995-998.

[12] P. Kumar, S.S. Rautaray and A. Agrawal, "Hand Data Glove: A New Generation Real-Time Mouse for Human-Computer Interaction", IEEE 1st Int'l Conf. on Recent Advances in Information Technology(RAIT), 2012.

[13] J.H. Kim, Nguyen Duc Thang and Tae-Seong Kim, "3-D Hand Motion Tracking and Gesture Recognition Using a Data Glove", IEEE International Symposium on Industrial Electronics (ISlE), July 5-8, 2009, Seoul Olympic Parktel, Seoul , Korea, pp. 1013-1018.

[14] M. Elmezain, A.A. Hamadi, G. Krell, S.E. Etriby and B. Michaelis,"Gesture Recognition for Alphabets from Hand Motion Trajectory Using Hidden Markov Models", IEEE International Symposium on Signal Processing and Information Technology, 2007, pp. 1192-1197.

[15] J. Weissmann and R. Salomon, "Gesture Recognition for Virtual Reality Applications Using Data Gloves and Neural Networks",IEEE, 1999, pp. 2043-2046.

[16] A. Cavallo, "Primitive Actions Extraction for a Human Hand by using SVD", IEEE 9th International Symposium on Intelligent
     Systems and Informatics (SISY), September 8-10, 2011, Subotica,Serbia, pp. 413-419.

[17] M.P. Paulraj, S. Yaacob, Mohd S.B.Z Azalan, R. Palaniappan, "A Phoneme Based Sign Language Recognition System Using Skin Color Segmentation", IEEE 6th International Colloquium on Signal Processing & Its Applications (CSPA), 2010, pp. 86-90.

[18] Y.F. Admasu and K. Raimond, "Ethiopian Sign Language Recognition Using Artificial Neural Network", IEEE 10th International Conference on Intelligent System Design and Applications, 2010, pp. 995-1000.

[19] Stefan Oniga, János Vegh and Ioan Orha, "Intelligent Human- Machine Interface Using Hand Gestures Recognition", IEEE.

[20] A.S. Ghotkar, R. Khatal, S. Khupase, S. Asati and M. Hadap, "Hand Gesture Recognition for Indian Sign Language", IEEE International Conference on Computer Communication and Informatics (ICCCI),Jan. 10-12, 2012, Coimbatore, India.

[21] L.K. Lee, S.Y. An, and S.Y. Oh, "Robust Fingertip Extraction with Improved Skin Color Segmentation for Finger Gesture Recognition in Human-Robot Interaction", WCCI 2012 IEEE World Congress on Computational Intelligence, June, 10-15, 2012, Brisbane, Australia.

[22] R. Verma and A. Dev, "Vision based Hand Gesture Recognition Using Finite State Machines and Fuzzy Logic", IEEE, 2009.

[23] J.B. Kim, K.H. Park, W.C. Bang and Z. Z. Bien, "Continuous Gesture Recognition System for Korean Sign Language based on Fuzzy Logic and Hidden Markov Model", IEEE, 2002, pp. 1574-1579.

[24] Marcus V. Lamar, Md. Shoaib Bhuiyan, and Akira Iwata, "Hand Alphabet Recognition Using Morphological PCA and Neural Networks", IEEE, 1999, pp. 2839-2844.

[25] M.S. Chang and J.H. Chou, "A Friendly and Intelligent Human-Robot Interface System Based on Human Face and Hand Gesture",

[26] IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Suntec Convention and Exhibition Center Singapore,July 14-17, 2009, pp. 1856-1861