STATIC HAND GESTURES RECOGNITION USING IMAGE PROCESSING

Poorva Arya
Research Scholar (Digital communication)
Govt. Engineering college ,Ajmer

Dr. U.S.Modani
Associate Professor
Govt. Engineering college ,Ajmer

Abstract

Hand gesture recognition system can be used in different area,for example ,HCI(human-computer interaction), remote control, robot control, computer generated reality and so forth. Hand gesture recognition system is for the most part the investigation of identification and acknowledgment of different hand gestures like American Sign Language hand gestures, Danish Sign Language hand motions and so on by a computer system. This work is centered on three fundamental issues in building up a motion acknowledgment framework. This work is centered on three fundamental issues in building up a motion acknowledgment framework. Human Computer Interaction (HCI) requires using various modalities (for example body position, speech, hand motions, Lip development, Facial articulations, and so on.) and coordinating them together for an increasingly vivid client experience. Hand signals are a natural yet ground-breaking correspondence methodology which has not been completely investigated for Human Computer Interaction (HCI). The most recent computer vision, image processing methods make vision based hand gesture recognition plausible for Human Computer Interaction (HCI).In this thesis sign language recognition system is presented in which a Hand gesture detection system is proposed based on shape context matching with ANN algorithm .The proposed work was implemented on MATLAB. To indicate the potency and effectiveness of the proposed system results performance are compared with existing work with 90% and it's been analyzed that the proposed algorithm had achieved highest accuracy with 95%.

Keywords: Human Computer Interaction, Gesture Recognition System, ANN, Hand Gesture Recognition System, Feature Extraction.

Introduction

Human computer interaction is the study of how people interact with computers and to what extent computers are or are not developed for successful interaction with human beings. Human interacts with computers in many ways to transfer information between human and computer. For instance, the most well-known method of HCI depends on straight forward mechanical gadgets—consoles and mice .These gadgets have become well-known yet intrinsically limit the speed and instinctive nature with which human can cooperate with the PC. This impediment has turned out to be considerably increasingly obvious with the rise of novel presentation

innovation, for example, augmented reality. In this way lately the field of PC vision has committed significant research exertion to the location and acknowledgment of countenances and hand gestures.

Human computer interaction (HCI) is characterized as the connection between the human and PC (machine), spoke to with the rising of PC itself [1]. Imperative undertaking of hand motion acknowledgment framework is to make a characteristic collaboration among human and PC when controlling and passing on important data. There are two primary attributes ought to be viewed as when planning a HCI framework as: usefulness and convenience.

Sign language recognition system

Communication with the physical world by non-verbal communication and gestures is progressively solid. Gestures or forms of non-verbal communication dialects can be communicated in different ways. These can be communicated by essentially waving hands, making a significant motion by hand, finger or body present or by an important outward appearance. In the middle of these gestures, hand gestures are those proficient intends to express important and note worthy data. In our genuine circumstance, we use hand gestures to speak with quiet & hard of hearing individuals by utilizing communications through signing, to check numbers and to express a feeling like ‘good bye’ or 'stop'. With the ongoing improvements in Artificial smart, Soft Computing and Neural Networks, hand gestures are turning into the most significant instrument to connect with PCs and machines. By and large, gestures are of two sorts: Static gestures and Dynamic gestures [2].

Static gestures are essentially communicated by some important body articulations. Basic 'Stop' motion is a static hand motion to express a note worthy information. Commonly static gestures are utilized to express powerful gestures moreover[3]. In ASL (American Sign Language), digits more noteworthy than 9 have been communicated by the hand development and with a static motion from 0 to 9. Dynamic gestures are communicated by body developments. Basically waving hand to express ‘good bye’ is a dynamic signal. Contrasted with other body parts, hands are progressively adaptable. The improvement of dynamic & static hand signal acknowledgment framework depends exclusively on the picture securing and preparing innovation. With the ongoing advancements in picture securing innovation and with the creation of very dependable cameras, both static and dynamic gestures are turning into the most significant apparatus for human PC cooperation. Hand gestures are enroute to supplant the normally utilized information gadgets like mouse, console ,joystick and some exceptional pens.Indeed,even it is believed that hand gestures will supplant the touch screen innovation of portable and different gadgets very soon. Numerous organizations have begun to create motion based PC control advancements, however a great deal improvement in this field is required. A few analysts have utilized gloves and comparable kind of equipment for static hand motion acknowledgment utilizing some costly sensors. Along these lines,these techniques are muddled progressively applications .Therefore, vision based static hand signal acknowledgment systems are generally utilized progressively applications. These techniques need no equipment with the exception of a camera, so these strategies are very less expensive and can be effectively acknowledged

by any industry[4].

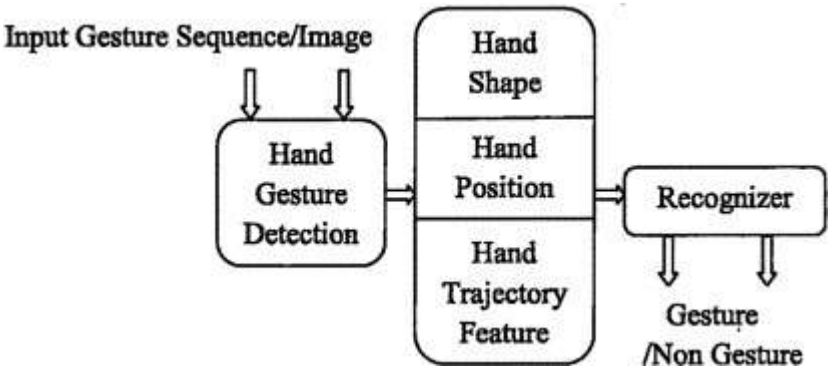


Fig 1 Basic Hand Gesture Recognition System

The paper presents a simple yet an efficient recognition system which will convert the static sign gestures of numbers 1to 5 into its feature vector matrix.Neural network is used for recognition and classification task. The rest of the paper will discuss different phases for the development of the system image acquisition, image preprocessing, feature extraction and classification in detail.

Related work

This will give an overview of various work studied to complete this research. This study shows various type of hand gestures detection like static hand gestures, real time hand gestures detection methods of feature extraction, feature detection and classification In [1] authorized has proposed the gradient based - key frame extraction scheme to solve the problemsarisingintheIndianSignLanguage(ISL)gesturerecognitionsystem.Thosekeyframes were used to split the gestures of sign language into series of signs, and further, each gesture was considered as an isolated gesture which also has removed the similar frames. The combination of Orientation histogram (OH)with Principal component analysis(PCA) has applied to extract the features of the preprocessed gestures. Finally, the experimental result was compared using classifiers Hidden Markov Model (HMM), Support vector machine (SVM) with various distance metrics like Euclidean distance, and have proved that Euclidean distance and correlation provides better accuracy.

In [2] authorized has adopted the multi- feature fusion based recognition algorithm for gestures. The initial process was the image segmentation stage, where it extracts the interested region of gestures in terms of color and depth through the mixture of depth data. Then the features such as weighted Hu invariant moments of depth map and Histogram of oriented gradients (HOG) of the

color image were extracted, and the fusion of the two features were performed. Further, the corresponding features were applied to the HMM classifier, and the performance were carried out. The experimental results

have proved that, the proposed method can adapt with the skin object influence, moving of multiple objects, background interference and performs the operation even in real time.

In [3] authorized has put forth a novel system for sign language gesture recognition. There were three main steps in this technique includes detection of region of interest, detection of key frames and recognition of gestures. From the uniform and non-uniform back grounds, the segmentation algorithm has distinguished the regions of interest. Further, the implementation results were analyzed, and have stated that it is a simple and fast algorithm that can handle temporal illuminations.

In [4] authorized Conceived a hand signal discovery furthermore, acknowledgment framework utilizing the BOF approach, and a multi class SVM classifier. They assembled a language structure that produces signal summons to control applications. Their framework is confined to following and perceiving static stances. Their punctuation was not able create sentences from their static stances. Their framework can accomplish an acceptable continuous execution, and additionally high order precision under variable conditions. Be that as it may, full DOF hand posture data is restricted by appearance-based techniques and may influence the consensus of this framework.

In [5] authorized has suggested the spatio-temporal feature extraction techniques for gesture recognition of Arabic Sign Language. Using forward, backward and bidirectional predictions, the temporal features of the video-based gesture were extracted. Then the motion of thesequence was represented from which the prediction errors are threshold and accumulated into a single image which was then followed by the feature extraction. Additionally, the linear severability of the extracted features were evaluated which was then applied to the classification techniques such as K nearest neighbor and Bayesian classifier. The experimental results have revealed the proposed method provides the requirements of the computation and storage of the classifier as the features are precise and concise.

In[6]authorizedProposedthesignalandhandactfollowingframeworkforNewZealandgesture based communication acknowledgment. They followed 13 signals without utilizing any marker. They perceived static stances of a hand and fingers to perceive communication via gestures. They Additionally utilized markers to achieve exact outcomes for stance acknowledgment. While they achievedhigh-determination input pictures for identifying finger poses, the prepare thus brought about a high computational cost.

In [7] authorized has developed the feature learning approach that consists of sparse auto-encoder (SAE) and PCA in order to recognize the human actions in case of RGB-D images. The components of the feature learning process were divided into two steps. In case of first component, the SAE with conventional neural network were learned the features from the image and depth channels where a sincase of second component, the final features were obtained by concatenating the feature which was then applied to the multi-layer PCA. The testing results have assured that the proposed method is effective by improving the recognition rate.

Artificial Neural Network

An ANN, a type of artificial intelligence, endeavors to mimic the conduct of human cerebrum. Artificial neural network is a framework made out of numerous straightforward handling components working in parallel whose work is controlled by system design, association qualities, and the preparing performed at processing components or hubs. Artificial neural networks are utilized for the most part when there does not exist any observational equation among input and output [8].

In [9] authorized has presented a comprehensive description of concepts involved in Neural Networks. Sandberg, A. [10] has developed a hybrid artificial neural network combining a Radial Basis Function and a Bayesian Classifier Network. The input data is passed to a Radial Basis network, which is trained using the Competitive Selective Learning algorithm, linked toBayesian neuralnetworkwhichdoestheclassification.RadialBasisfunctionseekstorecognizeposturesor features of signs. The resulting feature activation values are fed to Bayesian net, which calculates the estimated probability of the sign classes. The class with the highest probability estimate is chosen as the likely sign class.

Perceptron based neural networks have also been implemented for static sign recognition. Murakami, K., et. al. [11] have developed a back propagation learning algorithm based Recurrent Neural Network for finger alphabet recognition. Fels, S., [12] did extensive work with neural networks with his Glove-Talk II system ,in which three back propagation neural networks are used to translate hand signs into speech. The hand sign recognition process was broken up into three networks in order to increase speed and reduce training time. The consonant and vowel systems utilized standardized spiral premise initiation work [13] for the hidden inputs that tackled issues emerging from comparable sounding consonants and vowels [14].

The accompanying tables show synopses of some hand motion recognition frameworks. In Table 1 a correlation between recognition techniques close by signal recognition strategies utilized.

S.No.	Recognized Gestures	Total Gestures used ForTraining AndTesting	Recognition Percentage	Database used
[15]	26	1040	DP 98.8% MLP 98.7%	American Sign Language (ASL)
[16]	6	60	normal method 84% Scaling normalization method 95%	Own Database
[17]	26	208	92.78%	American Sign Language (ASL)
[18]	0-9 numbers	298 video sequence for isolated gestures/ 270 video sequence for continuous gestures	90.45%	Recognize Arabic numbers from 0 to 9.

[19]	5 static/ 12 dynamic gestures	Totally 240 data are trained and then the trained are tested	98.3%	5 static gestures and 12 dynamic gestures.
[20]	31	130 for testing	90.45%	Own Database
[21]	6	60	100% formore than 4gestures	Own Database
[22]	20	200	100% for 14 gestures,and >90 for15-20 gestures	Own Database

Table 1 Comparison b/w recognition methods used in hand gesture recognition

Proposed methodology

The system is designed on the principle of pattern recognition and a supervised training of neural network is done here. Pattern recognition is a process that takes raw data and perform some basic image processing steps on data and create feature vector matrix which will be given as input in neural network for its training.The flow diagram of the system is shown in Fig. 2.

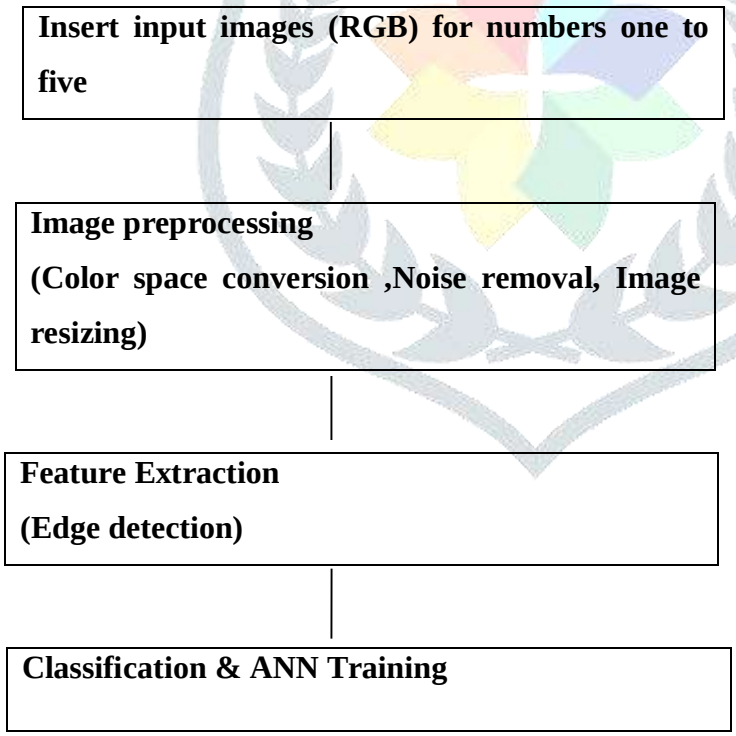


Fig. 2 Flow Diagram for Proposed Sign Gesture Recognition

Step 1: Image Input

In the first phase an image is taken from the webcam and stored in database. The system read input image from the database which contain RGB images for different hand gestures for numbers from 1 to 5. The database contains samples of 22 signs performed by different users for numbers 1 to 5. The images are with uniform background of dark and light color under different lightening conditions. The image database consists of total of 22 images.4 gestures for number 1,3 and 5 each and 5 gestures for number 2 and 4.

Step 2: Image Pre processing

Pre processing is performed using three steps: color space conversion, morphological operations to remove noise and image resizing for ease of feature extraction.

a) **Color space conversion**

We have our input in RGB color space for performing other pre processing steps further first we need to convert this RGB image to Gray scale image. The rgb2gray function converts RGB images to Gray scale image by eliminating the hue and saturation information while retaining the luminance.

b) **Morphological operations**

Image noise is random variation of brightness or color information in images and usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise is an undesirable by-product image capture that obscures the desired information.

Addition of salt and pepper noise

Salt and pepper noise is a form of noise sometimes seen on images. It is also known as impulse noise. This noise can be caused by sharp and sudden disturbances in the image signal.It presents itself as sparsely occurring white and black pixels.

Removal of salt and pepper noise

Use the median filter block to eliminate the black and white speckles in the image. Use the default parameters. The median filter block replaces the central value of the 3by3 neighborhood with the median value of the neighborhood. This process removes the noise in the image.

c) **Image resizing**

All images might be possible they all are not of same size so first ir is required to convert all input images in same size before going for feature extraction. Image interpolation occurs when you resize or distort your image from one pixel grid to another. Image resizing is necessary when you need to increase or decrease the total number of pixels, whereas remapping can occur when you are correcting for lens distortion or rotating an image. Zooming refers to increase the quantity of pixels, so that when you zoom an image, you will see more detail.

Step 3: Feature extraction

Feature extraction is a type of dimensionality reduction that efficiently represents interesting parts of an image as a compact feature vector. This approach is useful when image sizes are large and a reduced feature representation is required to quickly complete tasks such as image matching and retrieval. Feature detection, feature extraction, and matching are often combined to solve common computer vision problems such as object detection and recognition, content-based image retrieval, face detection and texture classification. MATLAB provides inbuilt functions to detect strong key points and then extract features from these key points.

Feature vectors are used widely in machine learning because of effectiveness and practicality of representing objects in a numerical way to help with many kind of analyses .They are good for analyses because there are many techniques for comparing feature vectors.

In image processing, features can be gradient magnitude ,color ,gray scale intensity ,edges areas and more. Feature vectors are used in classification problems, artificial neural networks and k-nearest neighbors algorithms in machine learning. Here we are using edge as a feature vector. For that we are using edge detection method for finding edge feature vector.

Edge detection includes a variety of mathematical methods that aim at identifying points in a digital image at which image brightness changes sharply or more formally has discontinuities. The points at which the image brightness changes sharply are typically organized into a set of curved line segments termed edges. Edge detection is a fundamental tool in image processing, machine vision particularly in the area of feature extraction and feature detection. There are various edge detection techniques for finding edges of objects in images using sobel,prewitt,roberts and canny edge detector.

If, for the method parameter, you select sobel, prewitt, or roberts, the edge detection block finds the edges in an input image by approximating the gradient magnitude of the image. The block convolves the input matrix with the sobel, prewitt, or roberts kernel. The block outputs two gradient components of the image, which are the result of this convolution operation. Alternatively, the block can perform a thresholding operation on the gradient magnitudes and output a binary image, which is a matrix of Boolean values. If a pixel value is 1, it is an edge.

If, for the method parameter, you select Canny, the edge detection block finds edges by looking for the local maxima of the gradient of the input image. It calculates the gradient using the derivative of the Gaussian filter. The Canny

method uses two thresholds to detect strong and weak edges. It includes the weak edges in the output only if they are connected to strong edges. As a result, the method is more robust to noise, and more likely to detect true weak edges.

STEP 5: Classification

An algorithm is created for finding number of finger tips for classification. We have applied masking and convolution operation in algorithm.

STEP 6: Artificial neural network training

The system uses feed-forward back propagation network for classification of sign gestures. Back propagation training algorithm is a supervised learning algorithm for multilayer feed forward neural network. Since it is a supervised learning algorithm, both input and target output vectors are provided for training the network. If there is an error, the Perceptron network will re-adjust the weights value until there is no error or minimized and then it will stop. Each pass through the input vectors is called epoch. Input vector is the 25x35 feature vector so thirty five input neurons are used. The target vector is also defined corresponding to each hand gesture. The performance of the training is evaluated with MSE.MSE set to 0.001, maximum validation failure set to 6 times ,the maximum number of epochs set to 10000.Sim function is used to simulate the model. Finally output from the neural network is trained for each classified hand gesture.

Results

For the implementation of the proposed system, image database is created for training and testing images. The image database consists of 22 static sign gestures of numbers 1 to 5 in '.jpg format'.The method is implemented using MATLAB R2015a. Skin region detection, image cropping resizing and feature extraction is performed using Image Processing Toolbox. Neural network toolbox is employed for training of neural network for classified hand gestures. The MATLAB built in function (sim) simulates network.Table 1,table 2 and table 3 shows implementation results of 5 gestures one for each number 1 to 5 .

S.no.	Input	Addition of Noise	Feature Extraction

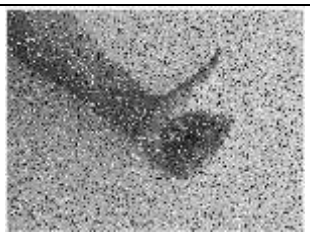
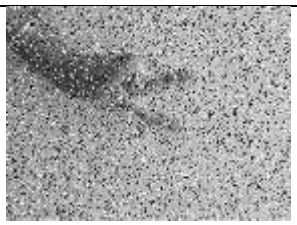
1			
2			
3			
4			
5			

Table 1 Edge detected features of images

Gesture	No. of image	Correctly classified image	Incorrectly classified image
One Finger	4	4	-
Two Finger	5	5	-
Three Finger	4	3	-
Four Finger	5	4	1
Five Finger	4	4	-
TOTAL	22	21	1

Accuracy= 22-1 /22 *100 95.4%

Table 2 Accuracy Measure for Gestures

S.no.	Input image	Finger Count
1		
2		
3		

4		
5		

Table 3 Finger count from input image

Conclusion and future scope

The system developed presents a simple yet an efficient method of gesture recognition using geometrical features based on the shape based properties of hand. Static hand gestures are recognized and classified then neural network is trained with feature vector matrix.. The accuracy is achieved 95.4% . Best training performance is .0053132 at epochs 10000 is achieved. In future the system can be extended to recognize dynamic hand gestures in an unrestricted environment for real life applications.

References

1. Yan, R. R., Tee, K. P. K., Chua, Y. Y., Li, H. H. and Tang, H. H. (2012), 'Gesture recognition based on localist attractor networks with application to robot control [application notes]', IEEE Computational Intelligence Magazine 7(1),64–74.
2. Liang, W., Guixi, L. and Hongyan, D. (2015), 'Dynamic and combined gestures recognitionbasedonmulti-featurefusioninacomplexenvironment',TheJournalofChina Universities of Posts and Telecommunications 22(2),81–88.
3. Rokade, R. S. and Doye, D. D. (2014), 'Spelled sign word recognition using key frame', IET Image Processing 9(5),381–388.
4. Dardas, N. H. and Georganas, N. D. (2011), 'Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques', IEEE Transactions on Instrumentation and Measurement 60(11),3592–3607.
5. Shanableh, T., Assaleh, K. and Al-Rousan, M. (2007), 'Spatio-temporal featureextraction techniquesforisolatedgesturerecognition inarabicsignlanguage', IEEETransactionson Systems, Man, and Cybernetics, Part B (Cybernetics) 37(3),641–650.

6. Akmeliawati,R.,Dadgostar,F.,Demidenko,S.,Gamage,N.,Kuang,Y.,Messom,C.,Ooi, M., Sarrafzadeh, A. and SenGupta, G. (2009), Towards real-time sign language analysis via markerless gesture tracking, in 'Instrumentation and Measurement Technology Conference, 2009. I2MTC'09. IEEE', IEEE, pp.1200–1204.
7. Li, S.-Z., Yu, B., Wu, W., Su, S.-Z. and Ji, R.-R. (2015), 'Feature learning based on sae– pca network for human gesture recognition in rgbd images', Neurocomputing 151, 565– 573.
8. Stephens, J., Adams, E., Huo, X., Hicks, J., and McCarty, D., "Use of neural networks in avalanche hazard forecasting" , Proceedings of International Snow Science Workshop, Utah, USA, 1994, pp327-340.
9. KroseBenJ.A.andP.PatrickvanderSmagt,"AnIntroductiontoNeuralNetworks",Fifth Edition, University of Amsterdam,1995.
10. SandbergA.,"SignRecognitionusingNeuralNetworks",Master'sThesis,Departmentof Computer Science, Stockholm University,1997
11. Murakami K. and H., Taguchi "Sign recognition using Recurrent Neural Networks" , Proceedings of CHI '91 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Kawasaki, Japan, 1991, pp237-242.
12. Fels S. and G. Hinton, "Glove-Talkll : An Adaptive Sign-to-Format Interface" , ProceedingsofCHI'95HumanFactorsinComputingSystems,ACMPress,1995,pp.456- 563
13. BellD.K.andLeeW.Lewi,"ArtificialNeuralNetworksforthemanagementResearchers: the State of the Art" , Technical paper, Department of Organizational Leadership and Strategy, Marriot School of Management, Brigham Young University, Provo, UT, November2003.
14. Fels S., "Glove-Talkll: Mapping Hand Signs to Speech Using Neural Networks - An Approach to Building Adaptive Interfaces" , Ph.D dissertation, University of Toronto, 1994.
15. Simei G. Wysoski, Marcus V. Lamar, Susumu Kuroyanagi, Akira Iwata, (2002). "A Rotation Invariant Approach On Static-Gesture Recognition Using Boundary Histograms AndNeural network.
16. Mokhtar M. Hasan, Pramoud K. Misra, (2011). "Brightness Factor Matching For Gesture Recognition System Using Scaled Normalization", International Journal of Computer Science & Information Technology (IJCSIT), Vol.3(2).
17. V. S. Kulkarni, S.D.Lokhande, (2010) "Appearance Based Recognition of American Sign Language Using Gesture Segmentation", International Journal on Computer Science and Engineering (IJCSE), Vol. 2(3), pp.560-565.
18. Shuying Zhao, Wenjun Tan, Shiguang Wen, and Yuanyuan Liu, (2008). "An Improved Algorithm of Hand Gesture Recognition under Intricate Background", Springer the First International Conference on Intelligent Robotics and Applications (ICIRA 2008),: Part I. pp. 786–794, 2008.Doi:10.1007/978-3-540-88513-9_85

19. Min B., Yoon, H., Soh, J., Yangc, Y., & Ejima, T. (1997). "Hand Gesture Recognition Using Hidden Markov Models". IEEE International Conference on computational cybernetics and simulation. Vol. 5, Doi:10.1109/ICSMC.1997.637

    Networks," IEEE Proceedings of the 9th International Conference on Neural Information Processing, Singapura.

20. E. Stergiopoulou, N. Papamarkos. (2009). "Hand gesture recognition using a neural network shape fitting technique," Elsevier Engineering Applications of Artificial Intelligence, vol. 22(8), pp. 1141 – 1158, doi:10.1016/j.engappai.2009.03.08.

21. Mokhar M. Hasan, Pramod K. Mishra, (2012) "Features Fitting using Multivariate Gaussian Distribution for Hand Gesture Recognition", International Journal of Computer Science & Emerging Technologies IJCSET, Vol.3(2).

22. Mokhar M. Hasan, Pramod K. Mishra, (2012). "Robust Gesture Recognition Using GaussianDistributionforFeaturesFitting',InternationalJournalofMachineLearningand Computing, Vol.2(3).

23. Bhumika Gupta , Pushkar Shukla, Ankush Mittal,(2016) "K-Nearest Correlated Neighbor Classification for Indian Sign Language Gesture Recognition using Feature Fusion" 978- 1-4673-6680-9/16/$31.00 ©2016 IEEE.