International Journal of Engineering Research-Online
A Peer Reviewed International Journal
Articles available online http://www.ijoer.in

Vol.3., Issue.4., 2015
(July-Aug)

RESEARCH ARTICLE

ISSN: 2321-7758

FINGER SPELLING SIGN LANGUAGE RECOGNITION

JAGRUTI S. CHAUDHARI[1], Dr. P.M. MAHAJAN[2]

[1]M.E. IVth Semester, Student, J.T. Mahajan College of Engineering, Faizpur, India

[2]Associate Professor, J.T. Mahajan College of Engineering, Faizpur, India

JAGRUTI S. CHAUDHARI

ABSTRACT

One of the most precious gifts of nature to the human breed is the ability to express himself by responding to the events occurring in his surroundings. Every normal human being sees, listens and then reacts to the situations by speaking himself out. But there are some less fortunate ones who are deprived of this valuable gift. Such people, mainly the deaf and the dumb, rely on some sort of sign language for communicating their feelings to others. Sign language is a mean of communication among the deaf people. The sign language is the fundamental communication method between the people who suffer from hearing defects. It is deals with ASL, test image feature are computed by using PCA feature extraction method and the image is recognized by using KNN classifier.

Communication is the process of exchanging information, views and expressions between two or more persons, in both verbal and non-verbal manner. Hand gestures are the non verbal method of communication used along with verbal communication. A more organized form of hand gesture communication is known as sign language. In this language each alphabet of the English vocabulary is assigned a sign. The physically disabled person like the deaf and the dumb uses this language to communicate with each other. The idea of this work is to design a system that can understand the sign language accurately so that the less fortunate people may communicate with the outside world without the need of an interpreter. By keeping in mind the fact that in normal cases every human being has the same hand shape with four fingers and one thumb, this work aims at designing a real time system for the recognition of some meaningful shapes made using hands.

1. INTRODUCTION

Sign language is a more organized and defined way of communication in which every word or alphabet is assigned some gesture. In American Sign Language (ASL) each alphabet of English vocabulary, A-Z, is assigned a unique gesture. Sign language is mostly used by the deaf, dumb or people with any other kind of disabilities..No one form of sign language is universal as it varies from region to region and country to country and a single gesture can carry a different meaning in a different part of the world. Various available sign languages are American Sign Language (ASL), British Sign Language (BSL), Turkish Sign Language (TSL), Indian Sign Language (ISL) and many more. There are a total of 26 alphabets. Hand gesture recognition is of great

importance for human-computer interaction (HCI), because of its extensive applications in virtual reality, sign language recognition, and computer games [1]. Hand gesture recognition provides an alternative to these cumbersome devices, and enables people to communicate with computer more easily and naturally [2].For handling different hand gesture recognition many tools have been applied including mathematical models like Hidden Markow Model (HMM) [3],K-Nearest Neighbors (K-NN)[4], Finite State Machine (FSM)[5],software computing methods such as fuzzy clustering[6], Artificial Neural Network (ANN) [7].In the English vocabulary, each alphabet may be assigned a unique gesture.

Finger spelling is the process of spelling out words by using signs that correspond to the letters of the word. An ASL user would use the American Finger spelling Alphabet, (also called the American Manual Alphabet). Now, sign languages are being used extensively in international sign use of deaf and dumb, in the world of sports, for religious practices and also at work places [8]. Gestures are one of the first forms of communication when a child learns to express its need for food, warmth and comfort. It enhances the emphasis of spoken language and helps in expressing thoughts and feelings effectively. In order for an ordinary person to communicate with deaf people, a translator is usually needed the sign language into natural language and vice versa.

2. Literature Review

2.1 Existing work for PCA technique :

Dipali Rojasara and Nehal Chitaliya et.al [9] proposed to use PCA technique. She is used haar transform with Principle component Analysis. She is proposed a system using Euclidean distance as a classification technique for recognition of various Signs of Indian sign Language. She is see that all one handed alphabets images have been identified correctly even though 20% noise is added. Some images are identified correctly after adding noise up to 60 to 70%.Hence; the feature extraction facilitates to reduce the computational time. Shreyashi Narayan Sawant et.al [10] presented design and implementation of real time Sign Language Recognition system from the Indian Sign Language using MATLAB .The signs are captured by using web cam and this signs are preprocessed for feature extraction. The obtained features are compared by using Principle Component Analysis (PCA) algorithm. After comparing features of captured sign with testing database minimum Euclidean distance is calculated for sign recognition. Finally, The proposed method gives output in voice and text form that helps to eliminate the communication barrier between deaf-dumb and normal people.

M.Ashraful Amin et.al [11] proposed a system that is able to recognize American Sign Language (ASL) alphabets from hand gesture with average 93.23% accuracy. The classification is performed with fuzzy-c-mean clustering on a lower dimensional data which is acquired from the Principle Component Analysis (PCA) of Gabor representation of hand gesture images.

2.2 Existing work for K-NN technique:

Dewinta Aryanie et.al [12] presented finger-spelling recognition method for American Sign Language (ASL) Alphabet using k-Nearest Neighbors (k-NN) Classifier. Performance of k-NN classifier is better for finger-spelling recognition when finger pattern is represented using full dimensional feature rather than using reduced-dimensional feature. This research also examines the effect of PCA for dimensional reduction to k-NN performance. The results show that k-NN classifier achieves the highest accuracy 99.8 percent.

Er. Aditi Kalsh et.al [13] proposed a system that can understand the sign language accurately so that the less fortunate people may communicate with the outside world without the need of an interpreter. Six alphabets are chosen these are alphabet A, alphabet D, alphabet J, alphabet O, alphabet P, and alphabet Q. It is recognized easily at 100% rate to show that this project not only improves the recognition rates of lacking alphabets, but also maintains the 100% recognition rate of other alphabets. It was observed that the recognition rate was fairly improved and recognition time was reduced significantly.

Christopher Lee and Yangsheng Xu et.al [14] developed a glove-based gesture recognition system that was able to recognize 14 of the letters

from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a rate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power Glove.

3. Proposed Methodology

This work is carried out by using two approaches, Principal Component Analysis and Region Based Technique for finger spelling recognition.

PCA was introduced in 1901 by Karl Pearson [15]. The advantage of finding these patterns is that PCA reduces the number of dimensions by compressing the data without losing too much information [16]. Principle component analysis aims to catch the total variation in the set of the training set, and to explain the variation by a few variables. In the case of hand recognition, there is a fact that is not necessary to analyze the whole hand image, but it is enough to store and analyze only difference between individual hand and it is used for hand recognition. After application of the image processing technique, the resulting images are used for the training and recognition phases of the PCA model.

3.1 Sign recognition by Principal Component Analysis

A general block diagram of sign language recognition system is shown in Figure 3.1. The recognition process is alienated into two phases- training and testing. In the training phase, the classifier has to be trained using the training dataset. The database can be either created by the researcher himself or an available database can be used.

Most of the sign language recognition systems classify signs according to hand gestures only or in other words .The important steps involved in training phase are creation of database, preprocessing, feature extraction and training the classifier. The testing phase contains test image acquisition (input can be images), preprocessing, feature extraction and sign language recognition.

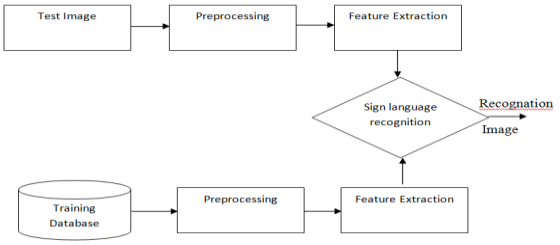

Figure 3.1: Generalized Block Diagram of Sign Recognition by Principal Component Analysis.

3.1.1 Data Acquisition

To achieve a high accuracy for sign recognition in sign language recognition system we use 5 user images for each letter and 100 images for single user. This is same for all 24 letters. These images are included in training and testing database. Data acquisition should be mostly perfect as possible for efficient hand gesture recognition. Suitable input device should be selected for data acquisition.

3.1.1.1 American Sign Language

American Sign Language (ASL) National Institute on Deafness and Other communication Disorders, 2005) is a complete language that employs signs made with the hands and other facial expressions and postures of the body. Use datasets for American Sign Language (ASL) finger spelling recognition is available online [17]. The datasets contain a set of RGB and depth images for each letter in the alphabet, organized by subject, for estimating generalization.

It is five user dataset. The first dataset comprises 24 static signs (excluding letters j and z because they involve motion). This was in five different sessions, with similar lighting and background. This dataset is use as it has availability of finger spelling alphabets, variety in signs, number of movement in single signs, hand shapes.

Figure 3.2: ASL Finger spelling Dataset

The five user dataset for various alphabets is as shown in figure 3.2.

3.1.2 **Pre-processing**

Preprocessing is very much required task to be done in hand gesture recognition system. We have taken ASL database [17] which is standard database in gesture recognition and use five user images for each letter and 100 images for single user. Preprocessing is applied to images before we can extract features from hand images. The respective images were passes through morphological operation.

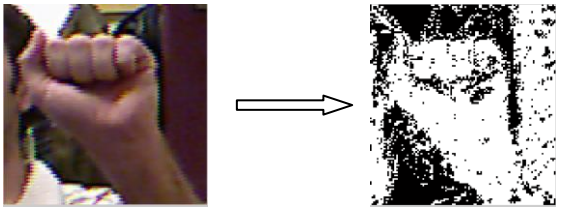

Fig 3.3 (a): Letter a input Image Fig 3.3(b): Skin Threshold

Figure 3.3(a) letter a input image is selected from American Sign Language database after preprocessing obtained the letter a skin threshold as shown in figure 3.3(b).The resulting threshold value for eigenvector is 0.0001 and skin detection is -2.

Preprocessing consist image morphological filtering methods. These features are further used for gesture recognition. Morphological filtering techniques are used to remove noises from images so that we can get a smooth contour. The preprocessing operation is done on the stored database. Preprocessing consist of Morphological filtering. Morphological techniques consist of operations dilation and erosion.

3.1.3 Feature Extraction

Feature extraction is a method of reducing data dimensionality by encoding related information in a compressed representation and removing less discriminative data. Feature extraction is vital to gesture recognition performance. Therefore, the selection of which features to deal with and the extraction method are probably the most significant design decisions in hand motion and gesture recognition development.

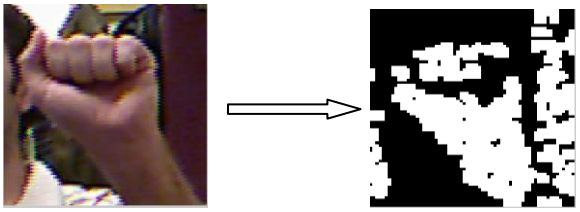

Fig 3.4(a): Letter a input image Fig 3.4(b): Extracted Feature

Figure 3.4 (a) letter a input image is selected from American Sign Language database after feature extraction obtained the extracted feature as shown in figure 3.4(b).Principal Component Analysis, by taking appropriate eigenvector of the covariance matrix. The letter a extracted feature is as shown in figure 3.4(b).

3.1.4 Sign Recognition

It is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data, the gesture recognition using PCA algorithm that involves two phases viz. training phase and testing phase.

During the training phase, each gesture is represented as a column vector. These gesture vectors are then normalized with respect to average gesture. Next, the algorithm finds the eigenvectors of the covariance matrix of normalized gestures. Lastly, this eigenvector matrix then multiplied by each of the gesture vectors to obtain their corresponding gesture space projections.

Later the test images are applied for recognition. In the testing phase, the PCA algorithm find the eigenvector of the image .These feature were fed to K-NN algorithm for recognition and use the value of K-NN=5. The K-NN is use for computing the distance. The minimum distance or nearest image recognize the K-NN. Finally, the performance of the classifier is measured.

The K-NN is used in finger spelling sign language recognition due to the reasons are the k-NN classifier is a simplest classifier in computer vision to solve classification problem. This classifier will find the k “closest” labeled examples in the training dataset and assign test data to the class that appears most frequently within the k-subset. K-NN is often categorized as a “lazy learning” classifier as it defers computation until it receives a request to

classify an unlabeled data. This characteristic makes it adaptive to new dataset. The k-nearest neighbor (K-NN) search helps us to find the k closest point or nearest point in X to a query point or set of points.

3.2 Region Based Technique:

Region Based parameter are Perimeter, area, Major axis length, Minor axis length. This parameter is used in this work which is some of the shape descriptors. This parameter is calculated as shown in table 3.1. Region-based analysis is invariant to translation, rotation and scale. They are also computationally simple.

The region based steps are going to be presented by the following sections.

Algorithm: Training

Input: Finger spelling training images.

Output: Finger spelling image features.

Method:

Step 1: Apply any of pre-processing techniques to the training images.

Step 2: Perform the feature extraction process for training images.
Region Based and Color Channel Process.

Step 3: Apply region based parameter for training images, such as Area, Major axis length, Minor axis length and Perimeter.

Step 4: For each channel divide the image in 3 color channels are R, G, B.

Step 5: Each channel find mean and standard deviation for each image.

Step 6: Show the extracted image for recognize the letter.

Training ends.

Following is the algorithm designed for recognition.

Algorithm: Recognition

Input: Finger spelling testing images.

Output: Recognize image, Classification label of input image.

Method:

Step 1: Apply any of pre-processing techniques to the testing images.

Step 2: Perform the feature extraction process for the testing images. Region Based and Color Channel Process.

Step 3: Apply region based parameter for testing images, such as Area, Major axis length, Minor axis length and Perimeter..

Step 4: For each channel divide the image in 3 color channels are R, G, B.

Step 5: Each channel find mean and standard deviation for each image.

Step 6: Apply the KNN and recognize nearest the letter.

Recognition ends. Calculation of region base parameter for individual letter

Table 3.1: Region base parameter of various letters

Image	Area	Major axis length	Minor axis length	Perimeter
c	2.309	0.100	0.039	0.293
h	2.567	0.147	0.047	0.537
m	672	45.79	20.35	129.19
i	701	49.49	31.35	161.68
f	2.64	0.099	0.065	0.480

Furthermore the feature of Region Based technique is R, G, B color channel. Region Based features R, G, B are used for human skin detection. In R, G, B calculated Mean and Standard for letter c, h, m, i, f in table 3.2.

Table 3.2: Region base feature of various letters

Image	R		G		B	
	Mean	Std	Mean	Std	Mean	Std
c	0.120	0.06	0.09	0.08	0.094	0.07
h	0.108	0.06	0.07	0.05	0.075	0.05
m	103.7	60.3	78.0	62.9	76.06	60.9
i	103.8	66.5	81.3	64.9	75.66	62.2
f	0.131	0.07	0.10	0.07	0.099	0.07

Region Based features R, G, B are used for human skin detection. In R, G, B here calculated the Mean and Standard for single user of letters c, h, m, i, f in table 3.2.

4. Results and Discussions

The proposed two algorithms Principal Component Analysis and Region Based technique were experiment on American Sign Language dataset and result in the form of accuracy.

After completing training parts of PCA, the image are selected from testing part. When image is

selected, then the test image is displayed. After selecting the test image, PCA by taking the appropriate eigenvector and then multiplied by each vector to obtained corresponding gesture and then recognize the image using K-NN classifier. The value of K-NN is five. The output of recognized image is shown in fig 4.1.

Figure 4.1: Recognize image using KNN classifier

After recognition of image, calculate the accuracy of PCA on confusion Matrix**:**

Confusion Matrix for PCA:

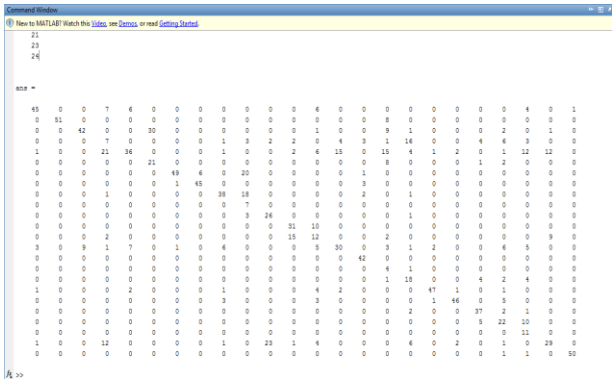

Figure 4.2: The Accuracy on Confusion Matrix

In this confusion matrix, we see the diagonally recognize number of images out of 50 images for PCA technique. Then, calculate the accuracy on confusion matrix and represent in the chart form as shown in fig 5.3.

Accuracy can be calculated out of 50 images,

Total no. Images=1200

Total no. of recognize images =736

Accuracy = (correctly recognize images / total no. of images)*100

= (736 / 1200)*100

=61.33%

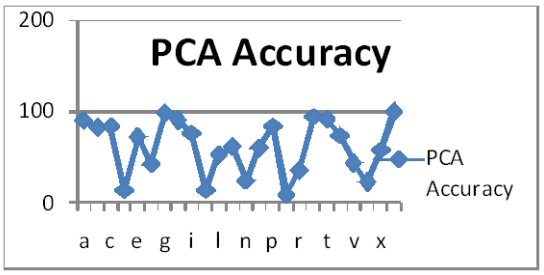

Figure 4.3: Accuracy of PCA for K=5.

4.1 Region Base Technique:

After completing training parts of region base, the image are selected from testing part. When image is selected, then the test image is displayed. After selecting the test image, features are computed by using region base feature extraction methods, and then recognize the image using K-NN classifier. It is shown in fig 4.4

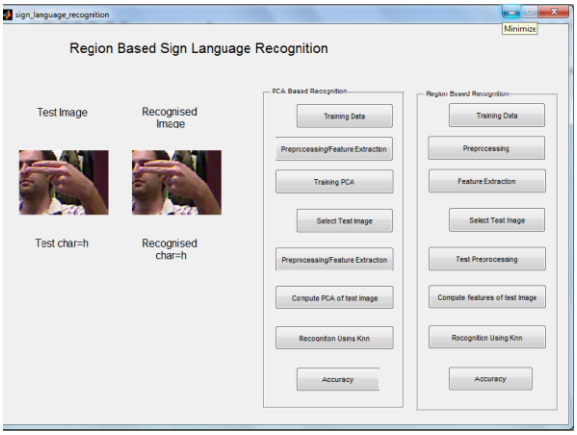

Figure 4.4: Recognize image using KNN classifier

After reorganization of image, calculate the accuracy of Region Base Technique on confusion Matrix**:**

We see in figure 4.5 the diagonally recognize number of images out of 50 images for Region Based.

Confusion Matrix for Region Based:

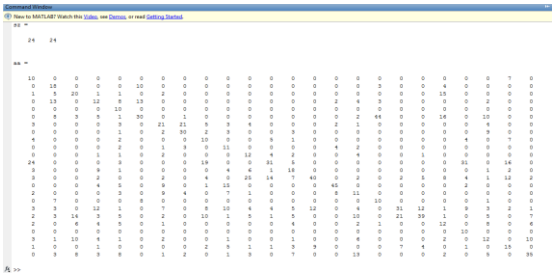

Figure 4.5: Confusion Matrix for Region Based

Then calculate the accuracy on confusion matrix for region base and represent in the chart form as shown in figure 4.6.

Accuracy can be calculated out of 50 images,

Total no. Images=1200

Total no. of recognize images =493

Accuracy=(correctly recognize images / total no. of images)*100

= (493 / 1200)*100

=41.08%

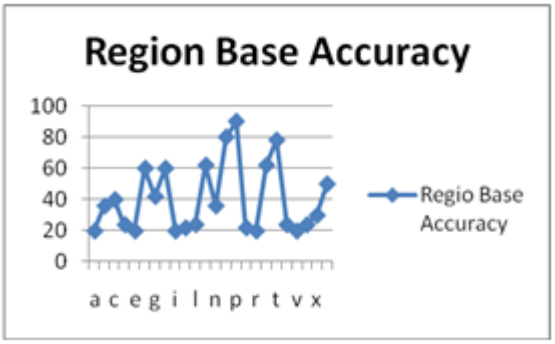

Figure 4.6: Accuracy of Region Base for K=5.

From the table 4.1 accuracy comparisons of PCA and Region Based contain letter, Principal Component Analysis accuracy in percentage and Region Based accuracy in percentage. Principal Component Analysis and Region Based were compared in terms of accuracy for individual letter (user 1) and using out of 50 images for each letter (user 1) to calculate the accuracy.

Table 4.1: PCA and Region Based Accuracy Comparison

Letter	*PCA Accuracy (%)*	*Region Base Accuracy (%)*
a	*90*	*20*
b	*82*	*36*
c	*84*	*40*
d	*14*	*24*
e	*72*	*20*
f	*42*	*60*
g	*98*	*42*
h	*90*	*60*
i	*76*	*20*
k	*14*	*22*
l	*52*	*24*
m	*62*	*62*
n	*24*	*36*
O	*60*	*80*
P	*84*	*90*
q	*8*	*22*
r	*36*	*20*
S	*94*	*62*
T	*92*	*78*
U	*74*	*24*
V	*44*	*20*
W	*22*	*24*
X	*58*	*30*
Y	*100*	*70*

These are the accuracy of PCA and Region Base both are combine in this chart as shown in figure 4.7.

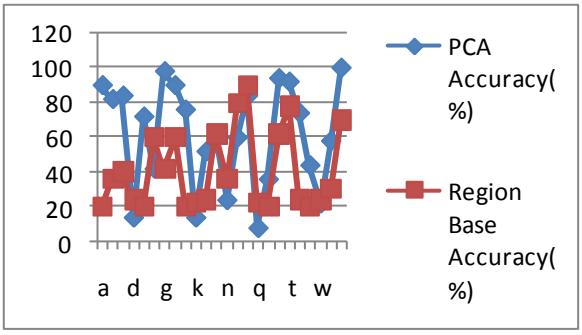

Figure 4.7: Compare the accuracy of PCA and Region Base

From the figure 4.7 show the on x-axis all the alphabets and on y-axis the accuracy in percentage. Some letters of PCA are a,b,c,e,g,h,p,s,t,u,x,y.This letters have highest accuracy as compared to Region Based.

5. Conclusions

Sign language is the medium of communication language generally used by deaf-dump community. This work is to identify some images signals deaf, dumb and gestures and try to facilitate communication between the deaf, dumb themselves and among normal people This language is not known by everyone. This system will be similar to the dictionary in which some signals deaf, dumb and gesture and what meaning to facilitate the communication between people. It uses gestures instead of sound patterns to convey meaning.

This work compares two feature extraction method i.e PCA and region Based. Most of the letters of PCA is the highest accuracy as compared to Region Based. Hence, the PCA accuracy is better than region base. The main advantage of PCA is very fast and consumes low memory as compared to

other methods. PCA is that reduces the number of dimensions by compressing the data without losing too much information. PCA results will come out with respect to standardized variables. They find it difficult to communicate with the normal people as the hearing or normal people are unaware of sign language.

REFERENCES

[1]. J. P. Wachs, M. Kolsch, H. Stern, and Y. Edan, —Vision-based handgesture applications, *Communications of the ACM*,vol. 54, pp. 60–71, 2011.

[2]. D. Y. Xu, —A Neural Network Approach for Hand Gesture Recognition in Virtual Reality Driving Training System of SPG, In Proceedings of the 18th International Conference on PatternRecognition, vol. 3, pp. 519-522, 2006.

[3]. Lawrence R. Rabiner, 1989. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, vol. 77 (2), pp. 257 – 286.

[4]. Pugeault, N., and Bowden, R. "Spelling It Out: Real-Time ASL Fingerspelling Recognition," In Proceedings of the 1st IEEE Workshop on Consumer Depth Cameras for Computer Vision, jointly with ICCV'2011, 2011.

[5]. Pengyu H., Matthew T., Thomas S. Huang. 2000. 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.

[6]. Xingyan Li, 2003. Gesture Recognition based on Fuzzy C-Means Clustering Algorithm, Department of Computer Science. The University of Tennessee. Knoxville.

[7]. Ben Krose, and Patrick van der Smagtan, 1996. An introduction to Neural Networks, the University of Amsterdam, eighth edition.

[8]. Theodore Brun. "Teckensprks Lexikon". Bokforlaget Spektra AB, Halmstad, 1974.

[9]. Dipali Rojasara,Nehal Chitaliya, International Journal of Soft Computing and Engineering,Vol.4,2014.

[10]. Shreyashi Narayan Sawant, International Journal of Computer Science and Engineering Technology.

[11]. M.Ashraful Amin ,Proceeding of the Sixth International Conference on Machine Learning and Cybernetics,Hong Kong ,19-22 August 2007.

[12]. Dewnta Aryanie,Yaya Heryadi,"American sign language-based finger-soelling recognition using k-nearest neighbours classifier", Researchgate conference paper,May 2015.

[13]. Er.Aditi Kalsh ,Dr.N.S.Garewal,"Sign Language Recognition System",International Journal of Computational Engineering Research,vol,03,Issue,6.

[14]. Christopher Lee and Yangsheng Xu. Online, interactive learning of gestures for human robot interfaces. Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania, USA, 1996.

[15]. K.Person, on lines and planes of closest fit to system of pointsin space ,in Philosophical Magzine,pp.559-572,1901

[16]. L.s.lindsay, a tutorial on principal component analysis february2012.

[17]. Pugeault N.,and Bowden ,R(2011)Spelling it Out Real-Time ASLFingerspelling Recognation In Proceedings of the 1st IEEE Workshop on Consumer Depth Camera for Computer Vision, jointly with ICCV2011.