Two-Handed Bangla Sign Language Recognition Using Principal Component Analysis (PCA) And KNN Algorithm

Promila Haque
Computer Scinece & Engineering
East Delta University
Chittagong, Bangladesh
haquepromila29@gmail.com

Badhon Das
Computer Scinece & Engineering
East Delta University
Chittagong, Bangladesh
dasbadhon069@gmail.com

Nazmun Nahar Kaspy
Computer Scinece & Engineering
East Delta University
Chittagong, Bangladesh
nazmun.nahar.kashfy@gmail.com

***Abstract*— Communication is the basic necessity for the existence in the society. Hearing-impaired people connect among themselves utilizing sign language, but typical people find it challenging to recognize it. The purpose of our paper is to decrease the communication gap between ordinary people and people with hearing and speech disabilities. This paper shows the sign language recognition framework prepared for recognizing 26 gestures from the Bangla Sign Language by utilizing PYTHON. In our proposed structure there are three phases: 1. Database formation phase 2. Training phase 3. Classification phase Preprocessing, segmentation, Eigenvector, and Eigenvalues are utilized as a part of the recognition. The Principal Component Analysis (PCA) was being used for recognizing images by extracting their principal component, and K-Nearest Neighbor algorithm is used for the classification phase. This method is very efficient with various backgrounds and effects.**

Keywords— gestures, PCA, eigenvector, SLR, KNN

I. INTRODUCTION

In the current world, the interface with the computing devices has become a necessity, and we cannot live without it. The prime target of the proposed system is to provide a feasible method for the compatible communication between common and hearing and speech disabled people by converting sign language into corresponding text. There are many sign languages including Indian Sign Language (ISL), Bangla Sign Language(BdSL), British Sign Language (BSL), American Sign Language (ASL) and so many. All of these languages are different from each other's. BdSL involves double handed gestures. Because of some difficulties like complex gesture, lack of data, etc., less research work has been developed on BdSL recognition system. The proposed approach introduces an SLR (Sign Language Recognition) method based on PCA for complex background and without any uses of instrumented wired such as "data gloves" and additional hardware and markers. Any user can use the proposed system because it doesn't require any other component and this is the key advantage of our system. It is a gesture dependent system. Here, we present a method of static gesture for two-handed SLR system.

Rest of the paper is organized as the followings: Section 2: Related work and researches on SLR, Section 3. Methodology, Section 4. Database formation phase, Section 5. Training phase, Section 6. Classification phase, Section 7. Implementation and Result, Section 8. Conclusion.

II. RELATED WORK

So many works have been done already on different sign language recognition procedures. In [1], Hidden Markov model is used to implement American Sign Language recognition system. The [2] Korean Sign Language recognition system presents a vision based recognition system which is executed by fuzzy min-max neural network algorithm. Another technique called FSM was used in the reorganization of ASL alphabets for static and dynamic gesture [3]. Four modules described by Ghotkar, A.S., et al., in their work for SLR [4]. They used Generic Fourier Description (GFD) for external features. A final module was the gesture reorganization with Genetic Algorithm. In Brazilian sign language, Andres Jessie Profirio, Kelly La's Wiggers, Luiz E.S Oliveria, Daniel Weingaerter [5] presents a method using 2D and 3D meshes projections of the hand. In this approach framing concept is used for the segmentation of videos. 3D model is further divided into volumetric 3D recognition [6] and skeleton 3D recognition [7]. 3D model is not rendered because of its complexity. In [8], a feature extraction method is used where special features from the images are collected and then compared with the database images. In Bangla sign language system [9], as an input, a video clip of hand gestures are taken and then generates an audio output. All the previous work on Bangla hand sign language recognition is done only with the same background. So the prime target is to derive a computerized sign language recognition system for Bangla Sign languages for various complex environments.

III. METHODOLOGY

The SLR system is primarily based on four phases. The phases are image acquisition, pre-processing, feature extraction and finally the classification. In this system, we have captured images by cameras. After that, we have resized them to reduce the computational efforts. In our proposed framework, at first we have extracted some of the important information from hand sign and then a comparison is made between the test image and images of the database. For the feature extraction phase, a principal component analysis is used, and the classification is done by K-nearest neighbor algorithm. The proposed methodology of the two-handed sign language recognition system presented in this paper shown in Fig. 1.

978-1-5386-9111-3/19/$31.00 ©2019 IEEE

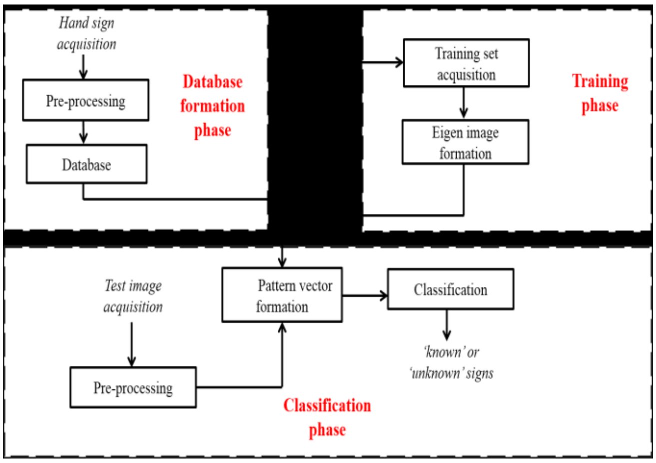

Fig. 1 Proposed Frameworks

IV. Database Formation Phase

This is the initial phase of the recognition process. At first, the acquisition of the hand images is performed. After that, some preprocessing techniques are applied to those images. Then they are added to an empty database. The database we used here is self-created. To achieve high accuracy for sign recognition in SLR system we use 104 images. Some images are given in Fig. 2. We captured the images in the regular camera, and the resolution is 4000×3000 pixel. After acquisition we have applied size normalization in our photos. Then we have converted the images into gray scale format. Lastly to enhance the image's quality histogram equalization was done. These modules automatically reduce every hand sign image into X*Y pixels, can allocate the intensity of hand images (histogram equalization) to improve sign language recognition performance. Then pre-processed images are stored in an initially empty database. Eigenimage formation, different train data processing operation has been performed on this database for testing.

V. Training Phase

In this phase, training set acquisition and Eigenimage formulation have been done. This module composes a feature vector that represents a hand image in a well enough manner. We have used feature extraction to reduce data [10] dimensionality. Its basic approach is to encode related information. PCA is used here for feature extraction purpose or calculating the Eigen images or calculating the Eigen images. PCA mainly reduce data dimension, and it converts high dimension data into lower dimension data. It extracts mostly relevant information from a dataset while it maintains important sample information. PCA algorithm steps are given below:

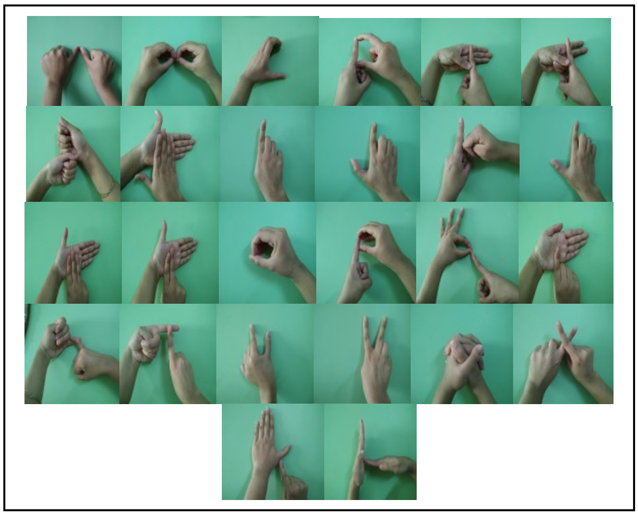

Fig. 2 Sample images in database

Each hand sign image of the database is normalized. Each image can be represented regarding a linear grouping of Eigenimages. These images are estimated using only the best Q Eigen images, having the largest eigenvalues. Let a hand sign image is N x N array i.e., it is basically a two dimensional array. Using PCA we have separated the finest vectors which can best describe the images. Any two dimensional image, for our case the N x N image can be described by a vector of length N^2, and is a linear combination of the original hand sign images. Let the images of the training set be $\Gamma_1, \Gamma_2, \Gamma_3 \ldots \Gamma_Q$. Then the average of all images is defined by

$$\psi = \frac{1}{Q}\sum_{n=1}^{q}\Gamma_n \qquad (1)$$

The average hand sign image Ψ is given in Fig. 3.
Then subtract the mean using the following equation

$$\Phi_i = \Gamma_i - \psi \qquad (2)$$

This differs each hand sign images from the average.
The second step is to calculate a covariance. The covariance is a measure of variations in a pair of variables. In this case, co-variance measured between a 2-dimension. We have 2-dimensional data set (x, y) and we have to measure co-variance between x and y dimensions respectively. The complete covariance matrix as

$$C = \frac{1}{Q}\sum_{n=1}^{Q}\Phi_i{\Phi_i}^T = BB^T \qquad (3)$$

where the matrix $B = [\Phi_1, \Phi_2 \ldots\ldots \Phi_Q]$. The covariance matrix C is a N^2 x N^2 symmetric matrix. It is inflexible to determine N^2 eigenvectors and eigenvalues for usual image sizes. There will be only Q-1 meaningful eigenvectors instead of N^2, if only if the dimension of the space is greater than the number of data points in the image space ($N^2 > Q$). The eigenvectors containing zero eigenvalues are left behind. So at first eigenvectors of a Q x Q matrix are solved, and then the remaining N^2 dimensional eigenvectors. Let Φi is the suitable linear groups of the hand sign images. Consider the eigenvectors u_i of $B^T B$ such that

$$B^T B\, u_i = \lambda u_i \qquad (4)$$

A pre-multiplication is applied on both sides by B, we have

$$BB^T \;\; Bu_i = \lambda \;\; Bu_i \qquad (5)$$

Fig. 3 Mean Image

A pre-multiplication is applied on both sides by B, we have

$$BB^T \; Bu_i = \lambda \; Bu_i \tag{6}$$

So Bu_i are the eigenvectors of

$$C = \; BB^T \tag{7}$$

A Q x Q matrix is constructed then

$$L = BB^T, \tag{8}$$

The next task is to discover the Q eigenvectors, u_i of L which can determine the linear combinations of training set sign images to the formation phase of the eigenimages v_i.

$$v_i = \sum_{k=1}^{Q} u_{lk}\Phi_{k'}, \tag{9}$$

where l=1,2,3,……Q. The calculations are significantly condensed with the above analysis. The eigenvectors with larger eigenvalues are chosen here. It symbolizes the usefulness in symbolizing the difference between the images. The accomplishment of the proposed method depends on the assessment of eigenvalues and eigenvectors of the physical matrix L which is formed from the training set of images. If the eigenvectors u_i are converted to a matrix again, they resemble as a hand like an appearance. That's why they are called Eigenimages. Some Eigenimages are shown in Fig. 4.

VI. CLASSIFICATION PHASE

We have classified each feature extracted image using KNN classifier. Euclidian distance is used to measure the distance between the Eigen images. Among N^2 image space, only Q' subspace of Eigen images are selected. This selection is done by choosing only certain Eigen images associated with the largest Eigenvalues. Practically a smaller Q' is satisfactory for recognition purpose. But for a better result, the number of Eigen images and training set images needed to be identical. In observing for a training set one hundred and four hand sign image, we use twenty-six Eigen images which are well enough. In this phase, a new preprocessed hand sign image (Γ) is transformed into its Eigenimage by using the following equation.

$$J_k = v_k^T(\Gamma - \psi) \tag{10}$$

for k = 1,2,… …Q. The weights J_k formed a feature vector or hand descriptor,

$$\Omega^T = [J_1, J_2, J_3, \ldots\ldots J_{Q'}] \tag{11}$$

Ω^T is known as pattern vector. In the pattern recognition algorithm, this pattern vector finds out whether any new input hand sign falls under any predefined hand sign classes. To classify a test image, a matching is performed between the pattern vectors of the training set and test image, and K-nearest neighbor algorithm is used for this purpose. We used

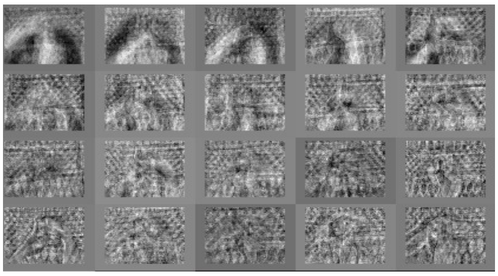

Fig. 4 EigenImage set

Euclidean distance functions for KNN. The distance between the pattern vectors has to be smaller than a threshold value ε_k, which can be determined by the following equation

$$\frac{\|\Omega - \Omega_k\|}{\|\Omega_k\|} \le \varepsilon_k \tag{12}$$

There are three possible output of the proposed method. In the first case, a correct recognition and identification of a hand sign image is being performed. In the second case, it can identify an unknown hand sign. And the last one can identify an image that is not a hand image. For avoiding false classification, the following equation is used

$$\frac{\|\Phi - \Phi_f\|}{\|\Phi_f\|} \le \Phi_k \tag{13}$$

where Φ_k is a user-defined threshold and

$$\Phi_f = \sum_{i=l}^{Q} j_i v_i \tag{14}$$

for the handiness of the input hand images belonging to k-th hand class.

VII. IMPLEMENTATION AND RESULT

The system has been implemented in PYTHON (3.4) version and hardware was core i3 Intel Pentium CPU with 2.10GHz processor machine and 4Gb RAM. And PCA and KNN algorithm is used here. We have created the database of by our own. Database of BdSL (a to z) alphabets is captured by a mobile camera. We have used four images per alphabet, so our database contains (4×26) = 104 images. There are 104 images in training dataset and 26 images in testing data set. Eigenimages are calculated by using the PCA algorithm. Testing images of different persons hand with different backgrounds are taken. Then the classification is performed by matching the pattern vector of the hand sign images of the database with the test image. A sample of testing images is given in Fig. 5

Fig. 5 Sample of test images

The processing results are bellowed:

TABLE I. NUMBER OF SUCCESSFUL RECOGNITION

Test Image	Detected Image	Success Rate(%)
104	81	77.8846

TABLE II. COMPUTATIONAL TIME OF EACH PHASE

Different Phases	Computational Time
Database phase	0.786 s

Different Phases	Computational Time
Training Phase	2.265 s
Classification Phase	0.058 s
Total	3.109 s

TABLE III. PRECISION AND RECALL CHART

Property	Value
True Positive	52
True Negetive	0
False Positive	1
False Negetive	10
Sensivity(Recall)	83.87%
Precision	98.11%
F-Measure	90.46%

Formula for calculating precision and recall.

$$Precision = \frac{True\ Positive}{True\ positive + false\ positive} \quad (15)$$

$$Recall = \frac{True\ Positive}{True\ positive + false\ Negative} \quad (16)$$

$$F - measure = \frac{2(Precision * recall)}{Precision + Recall} \quad (17)$$

VIII. CONCLUSION

This paper presents a mechanism that can be used for the recognition purpose of two-handed Bangla sign languages. The prime target was to build a prototype system using mechanism as mentioned above that will help to communicate with deaf and dumb people. Further, the system is implemented using PYTHON (3.4) version, but any person can use this system as a standalone application. The proposed system successfully recognizes the signs with acceptable running time and accuracy, most importantly with any background. We tried our best to improve the recognition rate as well as recognition success with different background images compared to the previous work. In the future, this work can be extended for input images with different hand orientation. We want to repeat our experiment on more massive databases with more signs.

REFERENCES

[1] C.Vogler,and D.Metaxas, " A Framework for recognizing the simultaneous aspects of American sign language, "in proceedings of the journal of Computer Vision and image Understanding, vol.81, 2001.

[2] S.kim, W.Jang and Z.Bien, "A dynamic gesture recognition system for the Korean Sign Language (KSL)," pp.434-439,vol.26,1996.

[3] Ketki. P.Kshirsagar. Dharmpal doye;"Object Based key Frame Selection for Hand Gesture Recognition "; International Conference of Advances in Recent Technologies in Communication and Computing, 2010.

[4] Ghotkar, A.S.,S.Khatal, S.Khupase, S.Asati and M.hadap. 2012. Hand Gesture Recognition for Indian Sign Language. International Conference on Computer Communication and Informatics (ICCCI-2012), Coimbator. pp. 10-12.

[5] Andress Jess`e Porfirio, Kelly La`is Wiggers, Luiz E.S. Oliveria, Daniel Weingaertner, "LIBRAS Sign Language Hand Configuration Recognition Based on 3D Meshes", 2013. IEEE International Conference on Systems, Man, and Cybernetics.

[6] B. Stenger, P. R. S. Mendonça, and R. Cipolla, "Model-based 3D tracking of an articulated hand", Proc. CVPR, Kauai, HI, USA, pp.310–315, 2001

[7] X. Bai and L. J. Latecki, "Path similarity skeleton graph matching,"IEEE Trans. Pattern Anal.Mach. Intell., vol. 30, no. 7, pp.1282–1292, Jul. 2008.

[8] G. Simion, V. Gui, and M. Otesteanu, "Vision Based Hand Gesture Recognition:A Review" , International Journal Of Circuits, Systems And Signal Processing, Issue 4, Volume 6, 275-282,20

[9] Adnan Eshaque, Tarek Hamid, Shamima Rahman, M Rokunuzzaman, "A Nobel Concept of 3D Animation Based 'Intelligent Assistant' for Deaf People: for Understanding Bengali Expressions." ICCIT, 2002.

[10] Sahoo, A. K., G. S. Mishra and K. K. Ravulakollu. 2014. Sign Language Recognition: State Of The Art. ARPN Journal of Engineering and Applied Sciences. 9(2):