Real Time Hand Gesture Recognition Using Different Algorithms Based on American Sign Language

Md. Mohiminul Islam,[1,*] Sarah Siddiqua,[2] and Jawata Afnan[3]

Department of Electrical and Electronic Engineering, Shahjalal University of Science and Technology

Sylhet - 3114, Bangladesh

[1]misakib1994@gmail.com, [2]sarah.siddiqua@ymail.com, [3]jawata24@gmail.com

Abstract— **Human Computer Interaction (HCI) is a broad research field based on human interaction with computers or machines. Basically, Hand Gesture Recognition (HGR) is a subfield of HCI. Today, many researchers are working on different HGR applications like game controlling, robot control, smart home system, medical services etc. The purpose of this paper is to represent a real time HGR system based on American Sign Language (ASL) recognition with greater accuracy. This system acquires gesture images of ASL with black background from mobile video camera for feature extraction. In the processing phase, the system extracts five features such as fingertip finder, eccentricity, elongatedness, pixel segmentation and rotation. For feature extraction, a new algorithm is proposed which basically combines K curvature and convex hull algorithms. It can be called "K convex hull" method which can detect fingertip with high accuracy. In our system, Artificial Neural Network (ANN) is used with feed forward, back propagation algorithm for training a network using 30 feature vectors to recognize 37 signs of American alphabets and numbers properly which is helpful for HCI system. The total gesture recognition rate of this system is 94.32% in real time environment.**

Keywords- American sign language; hand gesture recognition; fingertip finder algorithm; k curvature; convex hull; pixel segmentation; eccentricity; elongatedness; artificial neural network;

I. INTRODUCTION

Sign language is the primary language of the people who are deaf or hard of hearing and also used by them who can hear but cannot physically speak. It is a complex but complete language which involves movement of hands, facial expressions and postures of the body. Sign language is not universal. Every country has its own native sign language. Each sign language has its own rule of grammar, word orders and pronunciation. The problem arises when deaf and dumb people try to communicate using this language with the people who are unaware of this language grammar. So it becomes necessary to develop an automatic and interactive interpreter to understand them.

Research for sign language recognition was started in the '90s. Hand gesture related research can be divided into two categories. One is based on electromagnetic gloves and sensors which determines hand shape, movements and orientation of the hand. But it is costly and not suitable for practical use. People want something more natural. Another one is based on computer vision based gesture recognition, which involves image processing techniques. Consequently, this category faces more complexity.

Many researchers are working on hand gesture recognition using visual analysis. An automated vision based American Sign Language (ASL) recognition system was presented in [1] using the HSV color model to detect skin color and edge detection to detect hand shape. Another noteworthy development was the HCI system for recognizing faces and hand gesture from a video camera presented in [2]. They combined head position and hand gesture to control equipment. The position of the eyes, mouth and face center were identified to determine head position. Two new methods were introduced in their paper: automatic gesture area segmentation and orientation normalization of the hand gesture. Their recognition rate was 93.6%. The edge detection algorithm and skin detection algorithm were applied together in [3] using MATLAB for a better solution. The Canny edge detection algorithm for the purpose of detecting points at which image brightness changes sharply. They used ANN algorithm for gesture identification for fast computational ability. Static hand gesture recognition analyzing three algorithms named Convexity defect, K curvature and Part based hand gesture recognition was developed using Microsoft Kinect sensors [4]. Microsoft's Kinect camera allows for capturing pseudo-3D image called the depth map which can easily segment the input image and track the image in 3D space. But this camera is very costly. In [5], three techniques were explored: K curvature, Convex Hull, Curvature of Perimeter for fingertip detection. A new approach was suggested called Curvature of Perimeter with its application as a virtual mouse. A static and dynamic hand gesture recognition system was proposed in depth data using dynamic time warping in [6]. A directional search algorithm allowed for entire hand contour, the K curvature algorithm was employed to locate fingertips over that contour. Identification of Bengali Sign Language for 46 hand gestures was presented in [7]. ANN was trained by feature vectors of the fingertip finder algorithm. A database of 2300 images of Bengali signs was constructed. The experiment showed an accuracy of 88.69%.

A real time hand gesture recognition system and a new fingertip finder algorithm are presented in this paper. This proposed algorithm is basically a combination of K curvature and convex hull algorithms. So we can call this proposed

978-1-5090-6004-7/17/$31.00 ©2017 IEEE

algorithm as "K convex hull" method. Other algorithms such as rotation, eccentricity, elongatedness measurement and pixel segmentation are applied for extracting more features. A total of 30 feature vectors is extracted from each image. ANN with feed forward, back propagation algorithm is used to train the network using each sign features previously stored in the database. Then the network is used for real time recognition of a gesture. A graphical user interface (GUI) is created for testing unknown signs and showing text results. In the testing phase, users have to show signs with black background in front of a mobile camera and the meaning of that sign is shown in the GUI display panel. This GUI is shown in Fig. 1.

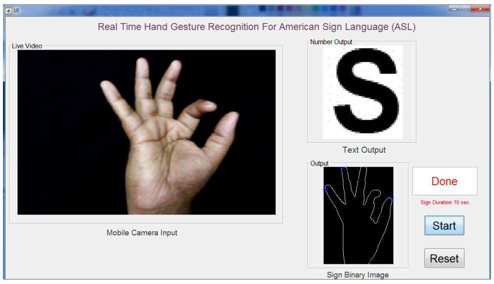


Fig. 1: Graphical User Interface

MATLAB R1016a is used for image processing in this research.

II. Methodology

The system is designed to visually recognize all static gestures of American Sign Language (ASL) with bare hand. Different users have different hand shapes and skin colors, making it more difficult for the system to recognize a gesture. The system combines five feature extraction algorithms for user independent and robust hand gesture recognition. The whole system works in four steps for gesture recognition such as image acquisition, preprocessing, feature extraction and feature recognition. The flow diagram of the proposed system is shown in Fig. 2.

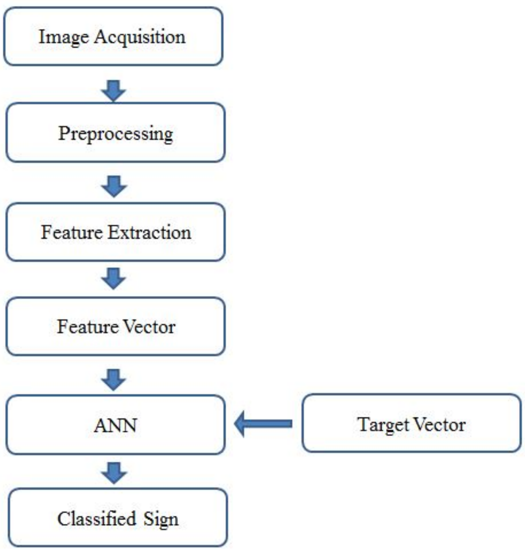


Fig. 2: Block diagram of sign language detection system

A. Image Acquisition

A total of 50 image samples of each sign of ASL is collected from different people. A database of 1850 images of 37 signs is created to extract feature vectors. The signs for all alphabets and numbers of ASL are shown in Fig. 3. A neural network is trained using this database.

Fig. 3: ASL signs from the database

B. Preprocessing

Pre-processing of an image is necessary as the image received from the camera may contain different noises. At first, the images are resized to 260×260 pixels. Then, they are converted from RGB to binary by Global histogram threshold using Otsu's method. Then median filtering is done for removing noise and preserving the edges. Morphological bridge and diag are used for filling holes and smoothing the edges. For sign detection, the portion from wrist to fingers of a hand is needed. So the rest of the part is eliminated from the image by cropping it. Then the fingers of the images are needed to be aligned vertically upward. So the images are rotated from 0 to 360 degrees with respect to hand wrist position. At least 15 consecutive white pixels are searched at the bottom of an image to locate the wrist. If white pixels are found at the bottom of the image, then no rotation is needed. If not, then the image is rotated by 90 degrees clockwise and checked again. In this way, the loop continues until the wrist is found.

C. Feature Extraction

An image can be identified and classified by some points of interest or set of values called the features. In this paper, five distinct features such as fingertip finder, eccentricity, elongatedness, pixel segmentation and rotation are used for feature extraction.

D. Feature Recognition

Multilayer feed forward neural network with back propagation training algorithm is used to identify American numbers and alphabets. ANN mimics biological neural network system. Actually neural networks are trained so that a particular input leads to a specific target output by adjusting the values of the connections between elements. In the training phase of the neural network, input vector is equal to the feature vectors and target vector is equal to the number of signs used for recognition.

Each sign is represented by a vector containing 30 features. ANN has three layers such as input, hidden and output. Hidden layer has a variable number of neurons with radial basis function. Here 20 hidden layers are used with trial and error basis for best performance. Fifty sample images of each sign from different individual are taken to train, test and validate the ANN. The network shows 99.7% accuracy in testing purpose. The ANN structure is shown in Fig. 4.

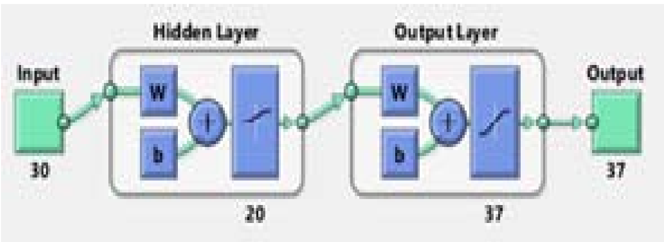


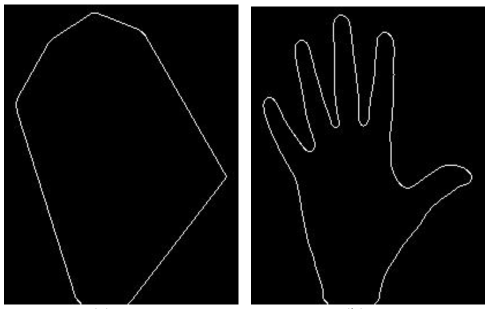
Fig. 4: ANN structure

III. PROPOSED ALGORITHM FOR FEATURE EXTRACTION

In this paper, five distinct features such as fingertip finder, eccentricity, elongatedness, pixel segmentation and rotation are used for feature extraction. By increasing number of features and image samples, the system gets better accuracy and reliability.

A. Fingertip Finder Algorithm

Combination of K curvature and Convex hull: Convex hull is making a polygon around the segmented hand region. It encloses all points of the region. The contour of the polygon is detected by Sobel edge detection algorithm shown in Fig. 5(a). The contour of the original hand gesture is also detected that is shown in Fig. 5(b). Then the AND operation is performed between these two images to find the common borders. This reduces the processing time drastically because there is no need to use the K curvature algorithm to all contour pixels of the hand shape. It also reduces the possibility of detecting false fingertip. So, we need to consider only these common points for fingertip detection. The result is shown in Fig. 6.



(a) (b)

Fig. 5: (a) Sobel edge detection of convex hull and (b) Original hand shape

Fig. 6: Common points after AND operation.

The K curvature algorithm finds an angle between two points. The algorithm measures the angle between two points [*p(i-k), p(i)*] and [*p(i), p(i+k)*], where *p* is the specific point of the contour, *i* is sequence number of the specific point, *k* is a specified pixel distance, here we set the value of *k* at 40 using the trial and error method shown in Fig. 7. A threshold value is set at 50°. If the angle at point *p* is greater than the threshold value, then *p* is considered as a fingertip. This threshold value is also found by trial and error method.

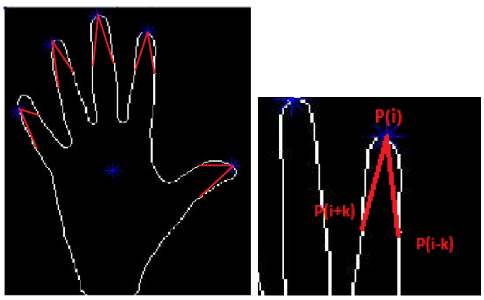


Fig. 7: Application of K curvature

In the proposed method, K curvature is applied only in the common pixels of the images found after AND operation as shown in Fig. 6. When a fingertip is detected, another fingertip is expected to maintain a minimum distance from that fingertip. This minimum distance between two points can be found by measuring the Euclidean distance. It is defined using trial and error method that the Euclidean distance between two closer fingertip should be greater than 22 pixels. This condition helps to get better accuracy by preventing false fingertip detection. This combined algorithm named "K convex hull" calculates fingertip more accurately. It also reduces the processing time as only the common points of the hand contour are listed as boundary points. For feature extraction, distance from the fingertips to the centroid of the hand region, the angle between a line from each fingertip to centroid and horizontal axis crossing through the centroid and total area of the hand region are calculated. So, total 11 features are extracted from each hand. Fig. 8 shows the result of fingertip detection.

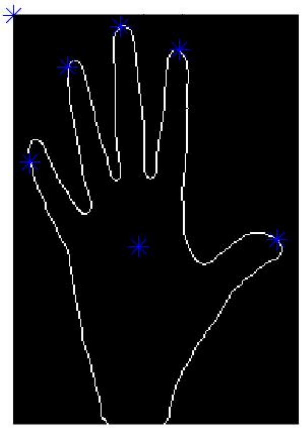


Fig. 8: Detection of the fingertips and the centroid

B. *Automatic Pixel Segmentation*

In this algorithm, an image is segmented into 16 blocks as shown in Fig. 9. Each block contains 256 pixels. Then the number of white pixels in each block is calculated. So, 16 blocks give 16 feature vectors containing the number of white pixels [4]. In this system, this algorithm is used as a feature for HGR.

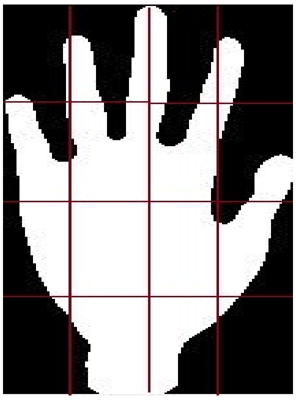


Fig. 9: Pixel segmentation into 16 blocks

C. *Eccentricity*

The major axis of an image is defined as the line segment connecting the two extreme points which creates the maximum Euclidean distance within a shape. The minor axis of the shape is defined as the line perpendicular to the major axis and of such length that if a box passes through the outer four points of the two axes, it completely encloses the shape. The ratio of the major axis to the minor axis is called the eccentricity of the shape. The major and minor axes are shown in Fig. 10.

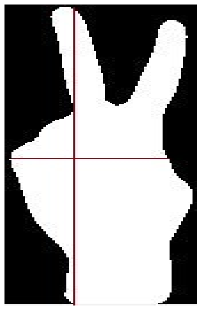


Fig. 10: Eccentricity showing maximum horizontal and vertical chord

D. *Elongatedness*

Elongatedness of an object is defined as the ratio of the smaller side to the larger side of the minimum bounding rectangle along the object. Fig. 11 shows the elongatedness of the hand shape.

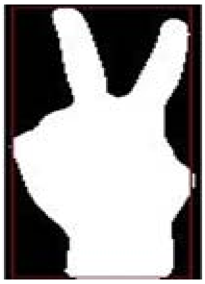


Fig. 11: Minimum bounding rectangle of the hand shape

E. *Rotation*

In ASL, there are some signs which are very similar. The only difference between them is the rotation angle. In such cases, the rotation angle is the only feature that can differentiate those signs. In our research, some complex alphabet signs of ASL are replaced by other rotated ASL signs for better recognition rate. Those signs are shown in Fig. 12.

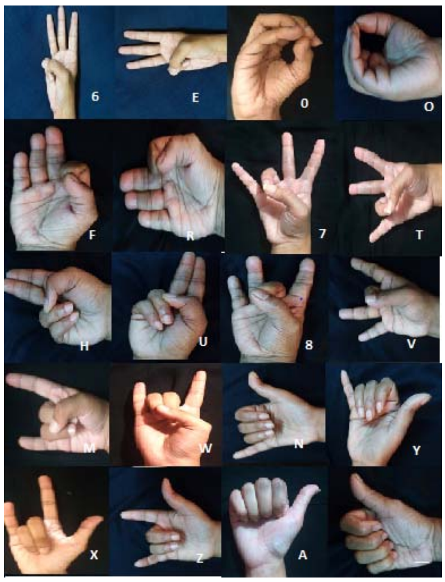


Fig. 12: Modification of some signs from ASL

IV. RESULT ANALYSIS & DISCUSSION

A. Real Time Setup

An Android application called DroidCam is used for real time setup. It allows the mobile camera to take the input images of hand gestures with better resolution. A GUI is created that represents hand gestures by a letter or number on the computer screen as shown in Fig. 13. When the start pushbutton is clicked, response time to take an image frame of the gesture in front of the mobile camera is about 10 seconds. At the end of processing the image, the converted result is shown in text version.

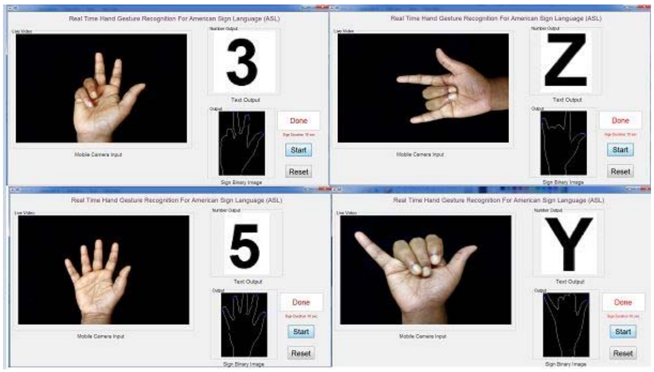


Fig. 13: Real time environment of HGR

B. Comparative Analysis of Different Algorithms:

In this research, several ANN are trained by individual features as well as a combination of all features in order to achieve performance analysis of the neural network.

A comparison analysis of ANN experimental results is shown in Table I. Comparison is done among individual features (i.e. fingertip detection (K convex hull method), automatic pixel segmentation, eccentricity & elongatedness) as well as combination of those features.

TABLE I. COMPARISON ANALYSIS OF DIFFERENT ALGORITHM AND THEIR COMBINATION (IN %)

Feature Extraction Algorithm	Testing Rate (%)	Validation Rate (%)	Average Recognition Rate (%)
Fingertip detection (K convex hull)	85.9	86.1	85.5
Eccentricity & Elongatedness	51.3	51.4	49.5
Pixel segmentation	99.2	98.1	99.02
Combined algorithm	99.5	99.9	99.6

The outcome shows that individual algorithms give the neural network testing accuracy of 85.9%, 51.3% and 99.2% respectively. While combining all features, better accuracy of 99.5% is achieved in the proposed system.

C. Real Time Performance Analysis

To do a real time hand gesture recognition performance analysis, each sign is taken from five different people. Two images are collected from each people while a black background and proper illumination are maintained. Features are extracted from the images and tested in the previously database-trained neural network. The number of correct responses out of 10 times of testing of each sign is shown in Table II.

TABLE II. REAL TIME HAND GESTURE RECOGNITION PERFORMANCE ANALYSIS

Class	Number of Correct Responses (Out of 10)	Recognition Rate (%)
0	9	90
1	10	100
2	8	80
3	10	100
4	10	100
5	10	100
6	7	70
7	10	100
8	10	100
9	10	100
A	10	100
B	10	100
C	8	80
D	10	100
E	9	90
F	8	80
G	10	100
H	8	80
I	9	90
J	9	90
K	10	100
L	10	100
M	9	90
N	9	90
O	10	100
P	9	90
Q	9	90
R	9	90
S	10	100
T	10	100
U	10	100
V	10	100
W	9	90
X	10	100
Y	10	100
Z	10	100
-	10	100
Average Recognition Rate = (3410/3700)×100 = 94.32 %		

The average recognition rate is calculated as follows:

*Average Recognition Rate = (No. of Correct Response) / (No. of Total Samples) * 100 %*

Real time average recognition rate of the proposed system was 94.32% (Table II).

D. Comparative Performance Analysis

A comparative study among different scholarly approaches and our propsed system is shown in Table III. The ANN testing result of the proposed method in [7] shows that it can detect 46 hand gestures with an average accuracy of 88.69%. The real time performance analysis by the method presented in [2] shows the detection of 11 hand gestures with an average accuracy of 93.6%. With the proposed method in this paper, a total number of 37 hand gestures are recognized and the average accuracy rate is 94.32% in real time environment. Our proposed system is superior in both accuracy rate and recognized classes.

TABLE III. COMPARATIVE ANALYSIS OF DIFFERENT HAND GESTURE RECOGNITION METHOD

Author reference	Method used	Hand gesture type	No of class	Average accuracy
Angur M. Jarman [7]	Fingertip finder algorithm with multilayered feed-forward, back propagation training	Bengali alphabet	46	88.69% (ANN testing result)
Yo-Jen Tu, Chung-Chieh Kao, Huei-Yung Lin [2]	Automatic gesture area segmentation and orientation normalization	Facial & Hand Gesture	11	93.6% (Real Time Result)
Proposed Method	K convex hull method, eccentricity, elongatedness, pixel segmentation and rotation	American number and alphabet signs	37	94.32% (Real Time Result)

V. CONCLUSION

We have applied different algorithms for feature extraction of hand gesture recognition system. This includes K convex hull for fingertip detection, pixel segmentation, eccentricity, elongatedness of the object. The experimental results show that K convex hull algorithm gives more accurate fingertip detection. Besides, other algorithms are applied to get greater accuracy in recognition system. Image frames taken by mobile video camera interfaced with the computer are tested by our trained ANN. The ANN is trained with 1850 sample images of our database and it recognizes ASL alphabets and numbers with almost 94.32% accuracy in real time environment. Our future research will be extended for further improvement in recognition accuracy and also for movement detection of hand for word recognition.

REFERENCES

[1] Sharmila Konwar, Sagarika Borah and Dr. T. Tuithung, "An American sign language detection system using HSV color model and edge detection", International Conference on Communication and Signal Processing, IEEE, April 3-5, 2014, India

[2] Yo-Jen Tu, Chung-Chieh Kao, Huei-Yung Lin,"Human Computer Interaction Using Face and Gesture Recognition", Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific, IEEE, Kaohsiung.

[3] Shweta. K. Yewale and Pankaj. K. bharne, "Hand gesture recognition system based on artificial neural network", Emerging Trends in Networks and Computer Communications (ETNCC), IEEE, 22-24 April, 2011.

[4] Marek Vanco, Ivan Minarik and Gregor Rozinaj, "Evaluation of static hand gesture recognition", International Conference on Signal and Image Processing (IWSSIP), IEEE, 12-15 May, 2014.

[5] Javeria Farooq and Muhaddisa Barat Ali, "Real time hand gesture recognition for computer interaction", International Conference on Robotics and Emerging Allied Technologies in Engineering (ICREATE), 22-24 April ,2014.

[6] Guillaume Plouffe and Ana-Maria Cretu, "Static and dynamic hand gesture recognition system in depth data using dynamic time warping" IEEE Transactions on Instrumentation and Measurement (Volume: 65, Issue: 2, Feb. 2016)

[7] Angur M. Jarman, Samiul Arshad, Nashid Alam and Mohammed J.Islam, "An automated Bengali sign language recognition based on finger tip finder Algorithm", International journal of Electronics & Informatics, 2015.