COPYRIGHT © 2011 IJCIT, ISSN 2078-5828 (PRINT), ISSN 2218-5224 (ONLINE), VOLUME 02, ISSUE 01, MANUSCRIPT CODE: 110749

Recognition Static Hand Gestures of Alphabet in ASL

Md. Atiqur Rahman, Ahsan-Ul-Ambia *and* Md. Aktaruzzaman

Abstract— This paper presents a system for recognizing static gestures of alphabet in American Sign Language (ASL) using artificial neural network (ANN). The required images for the selected alphabet are obtained using a digital camera. The color images are then cropped, resized, and converted to binary images.Then height, area, centroid, and distance of the centroid from the origin (top-left corner) of the image are used as features. Finally, the extracted features are used to train a Backpropagation NN. This recognition system does not use any gloves or visual marking systems. This system only requires the images of the bare hand for the recognition. Experimental results show that this system is able to recognize 26 selected ASL alphabets with an average accuracy of 80.28 %.

Index Terms—American Siign Language, ASL Recognition, Artificial Neural Network, Backpropagation Neural Network, Gestures Recognition.

1 INTRODUCTION

A sign language is a language which uses gestures instead of sound to convey meaning combining hand-shapes, orientation and movement of the hands, arms or body, facial expressions and lip-patterns. It is typically the first language and main means of communication for deaf individuals. The signers, however, still have serious problems of communicating with speaking persons, who are not sign users. The communication difficulty adversely affects the life and interpersonal relationships in the deaf community. Deaf individuals communicate with speaking people usually via interpreters or text writing. Although interpreters can help the communication between deaf and hearing persons, they are often expensive and have negative effect on independency and privacy. There is no universal sign language. Different countries use different sign languages. To help finding a communication aid for deaf people, many researchers have been working on recognition of various sign languages, e.g. Australian, Japanese, Chinese, German, and American sign languages, etc. Research in sign language recognition started to appear in literature at the beginning of 1990s. Stephan Liwicki and Mark Everingham [1] develop a method in which recognition based on hand shape alone is used where motion cues is not required. It also includes scalability to large lexicon recognition with no re-training. But there exists similar discrimination problem like previous work. The letter 'V' and 'I' has very similar representation. Moreover, the dataset contains only a single inexperienced signer and that the imaging conditions are only moderately challenging.Takahashi and Kishino [2] used a range classifier to recognize 46 Japanese Kana manual alphabets with a VPL Data Glove™. The hand gestures were simply encoded with data ranges for joint angles and hand orientations based on experiments. This system could recognize 30 out of 46 hand gestures correctly, but the remaining 16 signs could not be reliably recognized.

Artificial neural networks have been widely used in sign language recognition research. Murakami and Taguchi [3] investigated the use of recurrent neural nets for Japanese Sign Language recognition. Although it achieved a high accuracy of 96%, their system was limited only to 10 distinct signs. Kramer and Leifer [4, 5] developed an ASL finger-spelling system using a Cyberglove, with the use of neural networks for data segmentation, feature classifier, and sign recognition. Using a tree-structured neural classifying vector quantizer, a large neural network with 51 nodes was developed for the recognition of ASL alphabets. They claimed a recognition accuracy of 98.9% for the system.

American Sign Language (ASL) is an efficient technique for communication among most of the 2 million deaf people in United States and Canada. ASL consists of about 6,000 signs for representing the commonly used words [6]. Wilbur [7] stated that most of signs in ASL could be considered as a combination of 36 basic hand shapes. These 36 hand shapes include most of ASL alphabets and their variations. Therefore, the recognition of ASL alphabets is important for spelling a person's name and the words which are not in the ASL vocabulary. Begum, S. and Hasanuzzaman, M. [8] develop a computer vision-based Bangladeshi sign language recognition system. In this system, separate PCA (principal component analysis) is used for Bengali Vowels and Bengali numbers recognition. The system is tested for 6 Bengali vowels and 10 Bengali numbers.

This paper presents an ASL recognizer that has been developed to recognize 26 alphabet of ASL using Backpropagation Neural Network technique. In this recognition system, we have used a new set of features for training and testing of ANN. This system allows fast training and intelli-

- ***Md. Atiqur Rahman*** *is a student of the Department of Computer Science and Engineering, Islamic University, Kushtia-7003, Bangladesh. E-mail: atique0528@gmail.com.*
- ***Ahsan-Ul-Ambia*** *is serving in the Department of Computer Science and Engineering, Islamic University, Kushtia -7003, Bangladesh. E-mail: ambiaiu@yahoo.com*
- ***Md.Akteruzzam*** *is serving in the Department of Computer Science and Engineering, Islamic University, Kushtia-7003, Bangladesh. E-mail: mazaman_iuk@yahoo.com*

gent learning of new gestures. It can recognize 26 letters with average accuracy of 80.26%. The paper is organized as follows. Section 2 describes Used Materials and Method. Section 3 explains Neural Network Training Process. Section 4 shows Results Evaluation. Finally Conclusion is given in section 5.

2 USED MATERIALS AND METHOD

We have used MATLAB tools to implement the algorithm because MATLAB is a high-performance language for education and research [9]. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation and it also has toolboxes for signal processing, neutral network, image processing, database programs etc.

The block diagram of our recognition system is given in the following fig. 1

Take RGB image then resize it 120X110 and convert to gray scale image. Remove noise from gray scale image using Gaussian Filter and convert it to binary image after that clipped it 110X90.

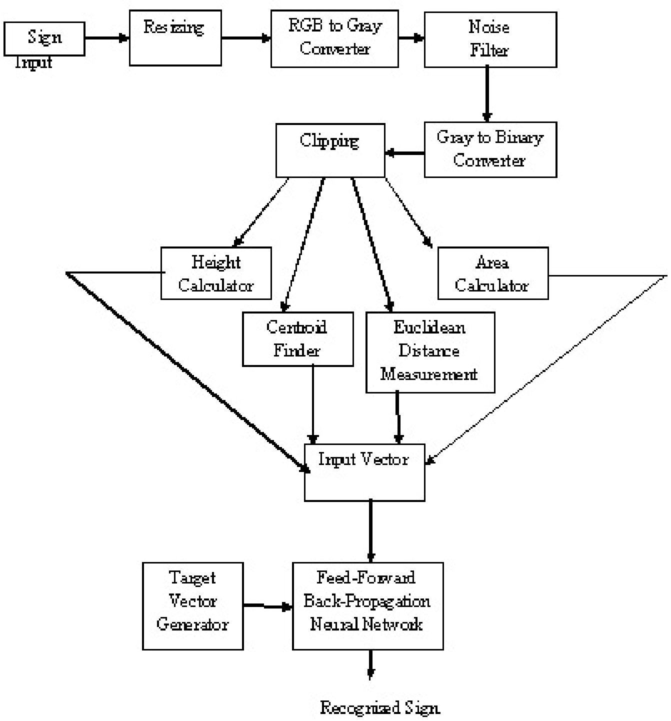

Fig. 1 Block diagram of our developed system.

2.1 Average Height of a Sign

To find the height of a sign from binary image where 0's represent black pixels and 1's represent white pixels, we have used pixel scanning method [10].

The average height of the binary image is calculated as

$Height_{avg}=N / M$.

Where N is the total number of black pixels (0's) of the image and M is the total number of columns containing at least one black pixel (0).

2.2 Centroid of a Sign

The centroid of a polyhedron is simply the average of the respective coordinates of all the vertices of the polyhedron [11]. For example, if the coordinates are (x1, y1) (x2, y2) ... (xN, yN), then the centroid would be,

xC=(x1+x2+...+xN) / N

yC=(y1+y2+...+yN) / N

In our system, the centroid of an image is calculated using the following formula:

$$Xc = \frac{\sum_{i=1}^{N} x_i}{area}$$

Where x_i represents the X-coordinate of each boundary pixel of the image.

$$Yc = \frac{\sum_{i=1}^{N} y_i}{area}$$

Where y_i represents the Y-coordinate of each boundary pixel of the image and N is the total number of boundary points. The centroid of the image is (Xc,Yc).

2.3 Euclidian Distance of the Centroid from the Origin

Given the two points (x1, y1) and (x2, y2), the Euclidian distance between these points is given by the formula [12]:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

The coordinates of the left-top most pixel of binary image is considered as the origin. Hence the Euclidian distance between centroid and the origin is given by

$$d = \sqrt{X_c^2 + Y_c^2}$$

Finally, the features collected from the above sections are combined to form a feature vector in the following order: Feature vector, V= [area, average height, x-centroid, y-centroid, centroid-distance]

3 NEURAL NETWORK TRAINING AND RECOGNITION

In our system, we have used multilayer feed forward neural network. For learning the network, Backpropagation algorithm has been used. The activation function used is "Sigmoid". In this system, 390 samples have been used to train the Network and 208 sample images have been used as the test set. After 4000 epochs the mean square error (MSE) goes to an acceptable level of 0.01. In this work, only 5 input neurons are used to take the input as 1×5 feature vector, 26 output neurons to classify 26 individual signs.

4 RESULTS EVALUATION AND DISCUSSION

In the present work, the signs of ASL are captured by a digital camera of 8 mega pixel resolution. The images are saved as .jpg format. There are 23 samples of each sign of ASL used in the training and recognition. Therefore, a total of 26x23=598 signs are captured. Among 598 images 390 samples are used for training the ANN and the remaining 208 images are used for testing or recognition. The average root mean square error goal is specified for training the network is less than or equal to 0.01. The number of iterations in which the network reached the specified error goal is equal to 10,000. The ASL alphabet recognition rate is presented in table 1.

TABLE 1

THE PERFORMANCE OF RECOGNITION OF ASL ALPHABET.

Letters	Recognized Letters	Unrecognized Letters	Recognition Rate (%)
A	7	1	87.5
B	7	1	87.5
C	6	2	75
D	5	3	62.5
E	7	1	87.5
F	5	3	62.5
G	5	3	62.5
H	8	0	100
I	7	1	87.5
J	6	2	75
K	4	4	50
L	8	0	100
M	8	0	75
N	5	3	62.5
O	6	2	75
P	7	1	87.5
Q	7	1	87.5
R	8	0	100
S	7	1	87.5
T	6	2	75
U	8	0	100
V	8	0	100
W	7	1	87.5
X	4	4	50
Y	7	1	87.5
Z	6	2	75

Recognition Rate= (No of Recognized Letters / No of total samples of that Ltters)*100%

For examples recognition rate of 'V' = (8/8)*100=100%

In the present study, the highest recognition rate is obtained for 'H','L','R','U'and'V', and the lowest recognition rate is found for 'K' and 'X'. Fig. 2 illustrates 3 recognized ASL alphabets 'A', 'U' and 'Y'.

Fig. 2 Some samples of recognized signs of ASL

5 CONCLUSION

In the present paper, a simple model of static gestures of ASL recognition system using ANN with MATLAB tools is discussed. In this paper, the recognition of individual and static signs is discussed. The system can perform static gesture training and recognition of ASL alphabets. The evaluation results show that the proposed method allows fast training and learning of new gestures and reliable recognition of the trained gestures afterwards. The average recognition accuracy of the system is 80.28%. Using more samples for training ANN may improve the performance of the system. Instead of ANN, other statistical recognition methods like HMM with more features may be used in future to improve recognition accuracy. The only limitation of our system is that, for learning NN, the feature vector should have integer values only. Future work will include extending the developed method to recognition of ASL with video based and interactive system.

REFERENCES

[1] Stephan Liwicki and Mark Everingham, "Automatic Recognition of Finger spelled Words in British Sign Language," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 50-57, Miami, FL, USA, June, 2009.

[2] T. Takahashi and F. Kishino, "Gesture coding based in experiments with a hand gesture interface device," ACM SIGCHI Bulletin, Vol. 23, 1991, pp. 67-73.

[3] K. Murakami and H. Taguchi, "Gesture recognition using recurrent neural networks," in Proceedings of the Conference on Human Factors and Computing Systems, 1991, pp. 237-242.

[4] J. Kramer and L. J. Leifer, "A 'talking glove' for nonverbal deaf individuals," Technical Report No. CDR 19900312, Center for Design Research, Stanford University, U.S.A., 1990.

[5] J. Kramer, "The TalkingGlove (RTM): hand-gesture-to-speech using an instrumented glove and a tree-structured neural classifying vector quantizer," Ph.D. Thesis, Department of Mechanical Engineering, Stanford University, U.S.A., 1996.

[6] L. A. Martin and E. D. Sternberg, American Sign Language Dictionary, Revised edition, HarperCollins Publisher, New York, 1994.

[7] R. B. Wilbur, American Sign Language: Linguistic and Applied Dimensions, 2nd ed., College Hill Publication, Boston, 1987.

[8] Begum, S.; Hasanuzzaman, M. "Computer Vision-based Bangladeshi Sign Language Recognition System" Computers and Information Technology, 2009. ICCIT '09. 12th International Conference.

[9] www.mathworks.com/access/helpdesk/help/techdoc/ref/

[10] Md. Aktaruzzaman, Md. Farukuzzaman Khan, Professor Dr. M. Ekin Uddin.(2009) "Recognition of Offline Cursive Begali Handwritten Numerals using ANN" Journal of the Peoples University of Bangladesh,vol. 4, N0.1, July 2009, pp18-28, Bangladesh, ISSN 1812-4747.

[11] http://answers.yahoo.com/question/index?qid=1006041717112

[12] en.wikipedia.org/wiki/Distance

Md. Atiqur Rahman is a student of the Department of Computer Science and Engineering (CSE), Islamic University, Kushtia.He obtained his B.Sc. in Computer Science and Engineering (Honours) degree from Islamic University, Kushtia and currently M.Sc. student of the same department .His research interest includes Image Processing, Speech Processing, Networking, Graph Theory, Wireless Communication.

Dr. Ahshan-Ul-Ambia has been serving as an Associate Professor with the Department of Computer Science and Engineering (CSE), Islamic University, Kushtia. He joined at Islamic University in October 1999 after completion his M.Sc. Degree from the Department of Applied Physics and Electronics, University of Rajshahi, Rajshahi. .His research areas include Biomedical Image and Signal Processing. Dr. Ambia has several papers related to these areas published in national and international journals as well as in referred conference proceedings.

Md. Aktarauzzaman has been serving as an Assistant Professor with the Department of Computer Science and Engineering (CSE), Islamic University, Kushtia. He joined at Islamic University in September 2005 after completion his M.Sc. Degree (CSE) from the Department of Computer Science and Engineering, Islamic University, Kushtia. Right now he teaches courses on Digital System, Automata, Image and Signal Processing, Compiler Design, VLSI Design. His research interest includes Medical Image Processing, Text Mining and Speech Processing.