IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

Gradient Feature based Static Sign Language Recognition

M. Mahadeva Prasad

Department of Studies in Electronics, Post Graduate Centre, University of Mysore, Hassan, India

Corresponding Author: prasada9@gmail.com

Available online at: www.ijcseonline.org

Accepted: 26/Dec/2018, Published: 31/Dec/2018

Abstract**—** In this paper, the work carried out to design the gradient feature based static sign language is presented. Sign languages are the gestures used by the hearing and speaking impaired people for communication. The sign languages are classified into static or dynamic or both static and dynamic sign languages. In static sign languages, still hand postures are used to convey information. In the dynamic sign languages, sequence of hand postures is used to convey information. In the present work, efforts have been made to design the computer vision based static sign language recognition system for the American Sign Language alphabets. The images that represent the static sign language alphabets are grouped into training and test images. The training sign language images are subjected to preprocessing. From the preprocessed images, magnitude and direction gradient features are extracted. These features are used to train the recognition system. The test images are subjected to preprocessing and feature extraction. The extracted features from the test sign language images are used to test the designed sign language recognition system. To classify the static sign language hand gestures, nearest neighbor classifier has been used. Independent experiments are carried out to evaluate the performance of the gradient magnitude and the gradient direction features. The average recognition accuracy of 95.4% for magnitude gradient feature and 80.3% for direction gradient feature are obtained.

Keywords**—** Sign Language Recognition System; American Sign Language; Static Sign Language; Gradient Features.

I. Introduction

Machine based sign language recognition system is one among the many applications developed under the category of human computer interaction (HCI) system. The sign languages (SL) are the gestures used by hearing and speaking impaired people to convey information. In SL, different hand shapes, orientation and movement of hands, facial expressions and lip-patterns are used for message communication. The hand gestures used in the SL communication are the spatio-temporal patterns. These spatio-temporal hand patterns may be static or dynamic or both static and dynamic [1]. While the still hand shapes or postures are used to represent the message in the static SL, a series of hand postures are used to convey the message in dynamic SL. The SL is classified into three categories [2]. They are; finger spelling, word level, and non-manual features. In finger spelling SL, different hand gestures are used to represent the spellings of a language. To form the word of a SL, sequence of finger spelling gestures is used. In case of non-manual features, facial expressions, movement of tongue, mouth and body positions are carried out to convey the message.

The task of identifying and interpreting the gestures of a SL is called the SL recognition. The knowledge of the defined gestures of a SL is essential to understand the signs or gestures made by the people. Many efforts have been made to design and develop the SL recognition system for HCI. The developed SL system can be classified into three main categories. They are, data glove method, marker glove method, and computer vision based method [3]. In data glove method, a hand glove mounted with accelerometers, RGB depth sensors have been used [4-6]. The data generated from these sensors due to the movement of the hands to show different SL gestures are used to design the classification system. The drawbacks of this system are; the system requires the frequent calibrations of sensors and the designed recognition system is costly. Since the sensors are very sensitive even for a small hand movement, misclassification of gestures occurs. In case of marker glove method, various colour markers are used on the hand glove to identify finger tips and wrist of a hand [7, 8]. This method demands the complex colour image processing to identify the colour coding used on the glove. Both data glove method and marker glove method are not suitable for multiple users. The reasons for this are; variation in hand shapes of people and the problem of hygiene associated with the use of common glove by different people. These drawbacks can be eliminated by the use of computer vision based SL recognition system.

The computer vision based SL recognition system can be classified into either signer dependent or signer independent systems. In signer dependent system, SL gestures generated by the single signer are used for both training and testing the recognition system. In signer independent system, SL gestures used to test the system are different from that of used to train the system.

The hand gestures used to represent the SL differs from region to region and language to language. There is a well-defined signs to represent the alphabets and other components of a language. In this paper, the work carried out in designing the computer vision based SL recognition system for the American Sign Language (ASL) alphabets is presented. In ASL alphabets, there are twenty-six hand sign gestures corresponding to twenty-six English alphabets. Figure 1 shows these twenty-six alphabets [9]. Out of these twenty-six alphabets, letters J and Z require the movement of hand for representation [1, 10]. Therefore, these letters are termed as dynamic hand gestures. Since the present work is

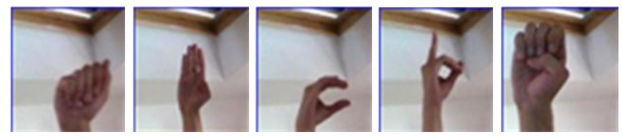
Figure 1. Twenty-six hand sign gestures corresponding to twenty six American sign language alphabets.

focused on the static hand gestures, the recognition system is designed to recognize letters from A to I and K to Y.

The remainder of this paper is organized as follows: In Section 2, information on the ASL alphabet hand gesture database used in the present work and the preprocessing of the hand gesture figures are discussed. The magnitude and gradient feature extraction and their salient features are presented in Section 3. The results of the experiments carried out to validate the proposed work are given in Section 4. Finally, in Section 5, the concluding remarks and the scope for future work are highlighted.

II. DATA COLLECTION AND PREPROCESSING

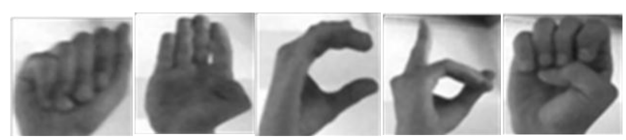
In the present work, the open source ASL alphabet hand gesture figure database has been used [11]. The hand gesture figures corresponding to twenty-four ASL alphabets which are used in this work are collected from this database. From this database, twenty figures corresponding to each of the twenty-four ASL alphabets are selected. The hand gesture figures are selected such that all the figures have the uniform illumination. The sample images of ASL alphabets A to E are shown in Figure 2.



A B C D E

Figure 2. Image data samples of ASL alphabets A to E.

Since the collected ASL colour images are having different background, position variations, and different sizes, they are subjected to preprocessing. The preprocessing methods used in the present work are; image cropping, image resizing, converting colour image to gray-scale image, and median filtering of gray-scale image. To begin with, the images are cropped to obtain the region of interest image part. The cropped images are resized to a size of $100\text{x}100$ so that all figures have the uniform size. The resized images are converted to gray-scale images and they are subjected to median filtering. The preprocessed figures of ASL alphabets A to E are shown in Figure 3.



A B C D E

Figure 3. Preprocessed image data samples of ASL alphabets A to E.

III. FEATURE EXTRACTION

In the present work, the magnitude and the direction values computed from the image gradient vector are used as features. In practice, the magnitude gradient and the direction gradient of an image are computed to find the strength and direction of the edges in the image.

In an image f, the edge strength and the direction at location (x, y) is found by computing the image gradient vector ∇f. The image gradient vector ∇f at a point (x, y) is a two-dimensional function having the first order derivatives in horizontal and vertical directions as given in (1).

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (1)$$

Where,

$g_x = \frac{\partial f}{\partial x}$ is the gradient in x-direction

$g_y = \frac{\partial f}{\partial y}$ is the gradient in y-direction.

The geometrical property of the gradient vector ∇f is that it points in the direction of the greatest change of f at location (x, y) [12]. The magnitude $Gmag(x, y)$ and the direction $Gdir(x, y)$ of vector ∇f are computed by using the expressions given in (2) and (3) respectively.

$$Gmag(x, y) = \sqrt{g_x^2 + g_y^2} \quad (2)$$

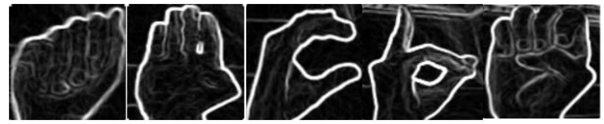
$$Gdir(x, y) = \tan^{-1}\left[\frac{g_x}{g_y}\right] \quad (3)$$

The common method used to compute g_x and g_y is by spatially filtering the image f by using the odd-sized Prewitt or Sobel masks. Due to the better noise-suppression property of Sobel masks, they are preferred over the Prewitt masks. The 3x3 Sobel masks used in this work for computing the x-direction gradient g_x and y-direction gradient g_y are given in (4) and (5) respectively.

$$g_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (4)$$

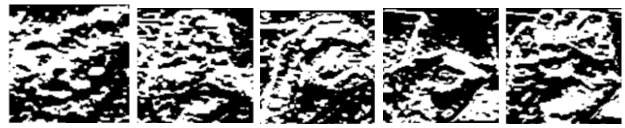
$$g_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad (5)$$

The static hand gesture figures of ASL alphabets are subjected to gradient feature extraction. The image gradient vector is computed from the preprocessed images using the Sobel masks. From the image gradient vector, magnitude gradient image and direction gradient image are computed. The magnitude Gmag and the direction Gdir images corresponding to the preprocessed ASL images of Figure 3 are shown in Figure 4 and Figure 5 respectively.



A B C D E

Figure 4. Gradient magnitude images of ASL alphabets A to E.



A B C D E

Figure 5. Gradient direction images of ASL alphabets A to E.

IV. EXPERIMENTS AND REULTS

The ASL alphabet figures in the collected database are segmented into the training set and the test set. Out of twenty image data samples corresponding to twenty-four ASL alphabets, ten image samples of each letter are reserved for training and the remaining ten image data samples are used for testing the designed recognition system. The image data samples are subjected to preprocessing and feature extraction as explained in sections II and III.

The ASL recognition experiments are carried out by using the nearest neighbor classifier with Euclidian distance as the distance measure metric. The average recognition accuracies of 95.4% and 80.3% are obtained for image magnitude gradient and image direction gradient features respectively.

V. CONCLUSION AND FUTURE SCOPE

The recognition system for the ASL alphabets is designed and the experiments are carried out with the twenty-four static hand gesture images corresponding to twenty-four ASL alphabets. The average recognition accuracies of 95.4% and 80.3% are obtained for image magnitude gradient and image direction gradient features. As part of the future work, experiments are carried out to design the signer-independent hand gesture recognition system. In addition, efforts will be made to design the recognition system that can handle the gesture images that are captured under varied illuminations and image depths.

REFERENCES

[1] Mahmoud Elmezain, Ayoub Al-Hamadi, Omer Rashid, Bernd Michaelis, "Posture and Gesture Recognition for Human-Computer Interaction", Advanced Technologies, Kankesu Jayanthakumaran (Ed.), InTech Publisher, pp. 415-440, 2009.

[2] Richard Bowden, Andrew Zisserman, Timor Kadir, Mike Brady, "Vision based Interpretation of Natural Sign Languages", In the Proceedings of the 2003 Int. Conf. on Computer Vision System, pp. 391-401, 2003

[3] Prashan Premaratne, "Human Computer Interaction using Hand Gestures," Springer, 2014.

[4] D. Karthikeyan, G. Muthulakshmi, "English Letters Finger Spelling Sign Language Recognition System", Int. Jl. of

Engineering Trends and Technology, Vol. 10, No. 7, pp. 334-339, 2014.
[5] Twinkel Verma, S.M. Kataria, "Hand Gesture Recognition Techniques", Int. Research Jl. of Engineering and Technology, Vol. 03, Issue 4, pp. 2316-2319, 2016.
[6] R.M. McGuire, J. Hernandez-Rebollar, T. Starner, V. Henderson, H. Brashear, D.S. Ross, "Towards a One-way American Sign Language Translator", In the Proceedings of IEEE Int. Conf. on Automatic Face and Gesture Recognition, pp. 620-625, 2004.
[7] Qi Wang, Xilin Chen, Liang-Guo Zhang, Chunli Wang, Wen Gao, "Viewpoint Invariant Sign Language Recognition", Computer Vision and Image Understanding, Vol. 1081, pp. 87–97, 2007.
[8] Nancy, Gianetan Singh Selhan, "An Analysis of Hand Gesture Technique using Finger Movement Detection based on Color Marker", Int. Jl. of Computer Science and Communication, Vol. 3, No. 1, pp. 129-133, 2012.
[9] https://www.nidcd.nih.gov/sites/default/files/Content%20Images/NIDCD-ASL-hands-2014.jpg
[10] https://en.wikipedia.org/w/index.php?title=American_manual_alphabet&oldid=873597841
[11] https://www.kaggle.com/grassknoted/asl-alphabet/data
[12] Gonzalez R.C., Woods R.E., "Digital Image Processing", 3rd Edn., Pearson Education, Inc., 2013.

Author Profile

Dr. M. Mahadeva Prasad is working as Assistant Professor in the Department of Studies in Electronics, Post Gradueate Cetnre, University of Mysore, Hassan, India. His research interests are signal processing, patterrn recogntion, audio and image processing, handwriting recognition. He has about 20 years of teaching and research experience. He is the the member of BoS and BoE of University of Mysore and other Universities of Karnataka. He is the life member of IETE, India.