

ASL Number Recognition using Open-finger Distance Feature Measurement Technique

Asha Thalange
Asst. Prof., E and TC Department,
Walchand Institute of Technology,
Solapur, Maharashtra, India

Shantanu Dixit, Ph.D.
E and TC Department,
Walchand Institute of Technology,
Solapur, Maharashtra, India

ABSTRACT

In recent years, lot of research is done in regard to the use of computers to recognize sign language. Computer recognition of sign language is an important research problem for enabling communication with hearing impaired people without the help of interpreter. In this article we propose a method to detect the static image based number of American Sign Language (ASL). This method is based on counting the open fingers in the static images and extracting the feature vector based on the successive distance between the adjacent open fingers. Further neural network is used for the classification of these numbers. This method is qualified to provide an average recognition rate of 92 percent.

General Terms

Pattern Recognition, Gesture Recognition.

Keywords

ASL Number, Neural Network, Static Hand Gesture Recognition, Open-finger Distance, Thinning.

1. INTRODUCTION

Computer is used by many people either at their work or in their spare-time. Special input and output devices have been designed over the years with the purpose of easing the communication between computers and humans, the two most known are the keyboard and mouse [1].

The idea is to make computers understand human language and develop a user friendly human computer interfaces (HCI). Making a computer understand speech, facial expressions and human gestures are some steps towards it. Gestures are the non-verbally exchanged information. A person can perform innumerable gestures at a time. Since human gestures are perceived through vision, it is a subject of great interest for computer vision researchers. Coding of these gestures into machine language demands a complex programming algorithm.

A gesture is scientifically categorized into two distinctive categories: dynamic and static [1]. A dynamic gesture is intended to change over a period of time whereas a static gesture is observed at the spurt of time. Dynamic gestures are considered as temporally consecutive sequences of hand or head or body postures in sequence of time frames. A waving hand means goodbye is an example of dynamic gesture and the stop sign is an example of static gesture. To understand a full message, it is necessary to interpret all the static and dynamic gestures over a period of time. This complex process is called gesture recognition. Gesture recognition is the process of recognizing and interpreting a stream continuous sequential gesture from the given set of input data. Dynamic gestures recognition is accomplished using Hidden Markov Models (HMMs), Dynamic Time Warping, Bayesian networks or other patterns recognition methods that can

recognize sequences over time steps. Static gesture (or pose gesture) recognition can be accomplished by using template matching, eigen spaces or PCA (Principal Component Analysis), Elastic Graph Matching, neural network or other standard pattern recognition techniques. Template matching techniques are actually the pattern matching approaches. It is possible to find out the most likely hand postures from an image by computing the correlation coefficient or minimum distance metrics with template images.

Gesture recognition has significant application in sign language recognition. Sign languages are the most raw and natural form of languages that could be dated back to as early as the advent of the human civilization, when the first theories of sign languages appeared in history. 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.

Hearing impaired people have over the years developed a gestural language where all defined gestures have an assigned meaning. The language allows them to communicate with each other and the world they live in. Fig. 1 shows the different gestures in American Sign Language for the numbers 0 to 9.



Fig. 1: Gestures in American Sign Language

2. RELATED WORK

In sign language recognition, it is desirable to use a shape representation technique that will sufficiently describe the shape of the hand while also being capable of fast computations, enabling the recognition to be done in real-time. It is also desirable for the technique to be invariant to translation, rotation, and scaling. In addition, a method that will allow for easy matching would be beneficial.

Gesture recognition was first proposed by Myron W. Krueger as a new form of interaction between human and computer in the middle of seventies [2]. Currently, there are several available techniques that are applicable for hand gesture recognition. Zimmerman [3] developed a VPL data glove that is linked to the computer to recognize signs. The glove can measure the bending of fingers, the position and orientation of the hand in 3-D space. In vision-based gesture recognition, hand shape segmentation is one of the toughest problem under

a dynamic environment. It can be simplified by using visual markings on the hands. Some researchers have implemented sign language and pointing gesture recognition based on different marking modes [4]. Y. Cui [5] represented a system which is able to recognize different hand gestures under a complex background. Klimis Symeonidis [6] used an orientation histogram of the image to develop a simple and a fast algorithm to extract features from the static image for comparison and recognize a static ASL using neural network. Becky Sue Parton [7] discusses the various projects involving sign language and the potential impact these endeavours will have on deaf education and communities at large. He also discusses the use of artificial intelligence in the field robotics, virtual reality, computer vision, neural networks, Virtual Reality Modelling Language (VRML), three-dimensional (3D) animation, natural language processing (NLP), and intelligent computer aided instruction (ICAI), for sign language manipulation. Jonathan C. Rupe [8] developed the system to identify hand shapes commonly found in American Sign Language using the region-based Generic Fourier Descriptor. Other approaches such as; local linear embedding, Neural Network shape fitting, object based key

frame selection, and Haar wavelet representations have been presented in [9] [10-12]. Rokade et al [13] used thinning method on one to ten numbers of American Sign Language (ASL). The feature vector consists of angle between the lines, which join center point to each endpoint, and center point to the vertical extended line. Ravikiran J et al [14] introduced an efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet of the American Sign Language. Asanterabi Malima, [15] et al proposes an algorithm invariant to translation, rotation, and scale of the hand for counting the open fingers to control a robot.

3. SYSTEM DESCRIPTION

Some of the signs of ASL numbers have same number of fingers open. Such as sign of number three, six, seven, eight and nine, all have total of 3 fingers open. It becomes difficult to recognize the exact number just by counting only the number of fingers open. This article deals with an algorithm that not only counts the number of fingers open in ASL number but also recognizes the exact number from 0 to 9. Following fig. 2 shows the system block diagram

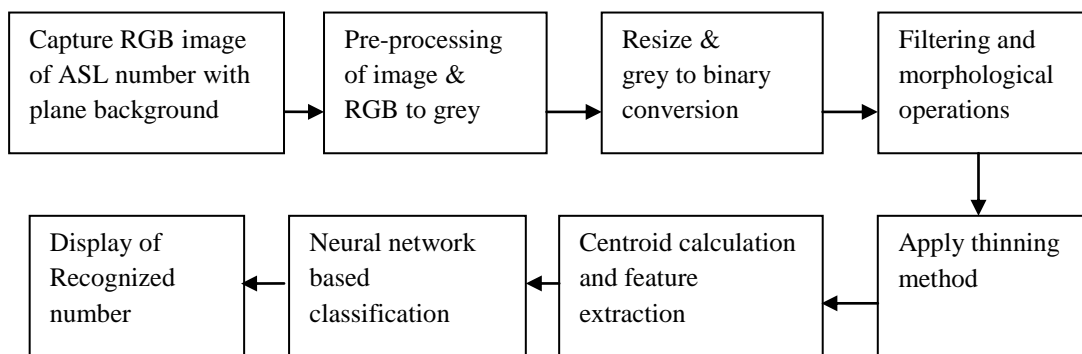


Fig. 2: System Block Diagram

The whole system functioning is divided into four main modules namely:

- Image capture and Pre-processing
- Binary image conversion and Morphological operations
- Image thinning and Feature extraction
- Classification

3.1 Image Capture and Pre-Processing

Colour image of sign of ASL number with a plane black background is captured by a 5 megapixel web-camera concentrating on the palm of the hand. The plane background is used for simplicity of processing at the initial stage. The extracted image of ASL gesture from the signer is first pre-processed to enhance the image's quality. For this low pass filtering and median filtering is applied to the input image. Further the colour image is converted to grey scale image and resized to 256 x 420 image. Initially the dataset is created by us. Total of 300 images are captured, 30 images for each number sign by a single signer along with a total of 376 images of all numbers for testing. The database is made robust such as the images are scale independent and contains rotated samples between +45 and -45 degree in order to make system robust.

3.2 Binary Image Conversion and Morphological Operations

The resized grey image is converted to binary image with an appropriate threshold that converts the background to black and foreground i.e. palm to white. Also further filtering and morphological operations are applied to remove noise.

3.3 Image Thinning and Feature Extraction

Image thinning is applied to some extent to binary image in order to get clear separation among the fingers. The extent of thinning is limited so that the image is not totally converted to skeleton. The centroid of the hand is obtained and with this centroid as centre, the circle is plotted so that it cuts all the open fingers of the hand. Total fingers open is the count-1, where count is number of times the circle traces the continuous '1' along the circumference. Further features are extracted. Here the equivalent distance i.e. the number of black pixels traced by the circle between the two adjacent open fingers are used as the feature vector for classification. Here the circle is traced anti-clock wise starting from bottom of image i.e. from 2700. The number of black pixels are counted back to back and stored in a vector form. The amount of separation between the respective fingers for different signs of ASL numbers is different thus giving different type of vector sequence. This vector obtained itself is a feature vector which is used for further recognition.

3.4 Classification

Multi-layered feed forward back-propagation Neural Network based classification engine is used here. Initially the network is trained with feature vectors obtained from training set consisting of 30 images of each ASL number obtained from a

single user. The feature vector obtained from the actual test image is applied to this trained network for classification or recognition of ASL number. Further the detected number is displayed in text form.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

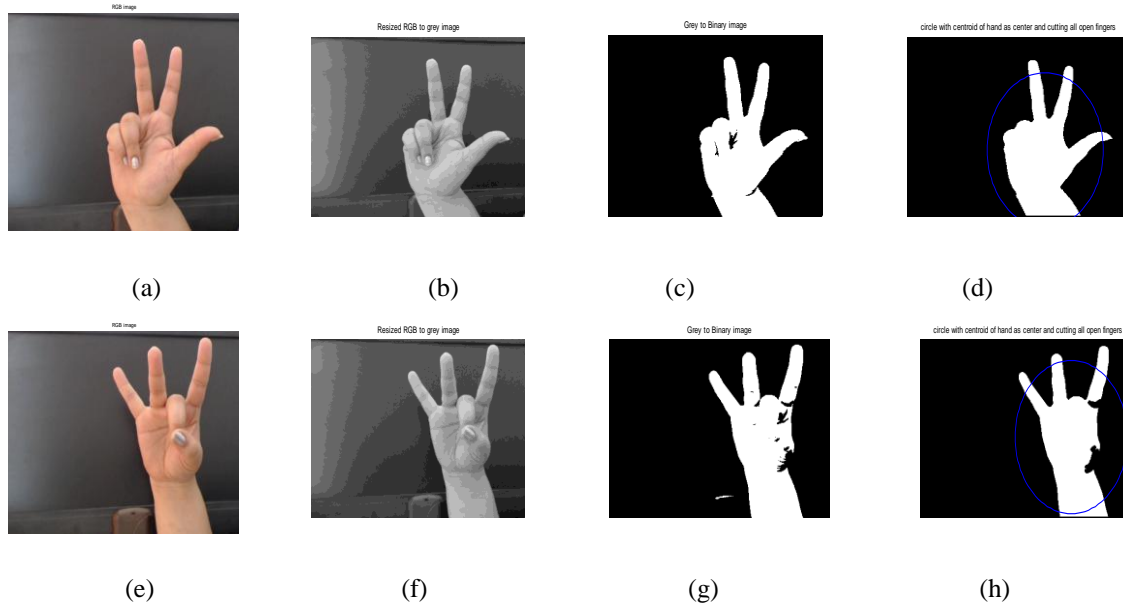


Fig. 3: Processing Stages of Number Recognition for ASL number 3 (a-d) and ASL number 8 (e-h)

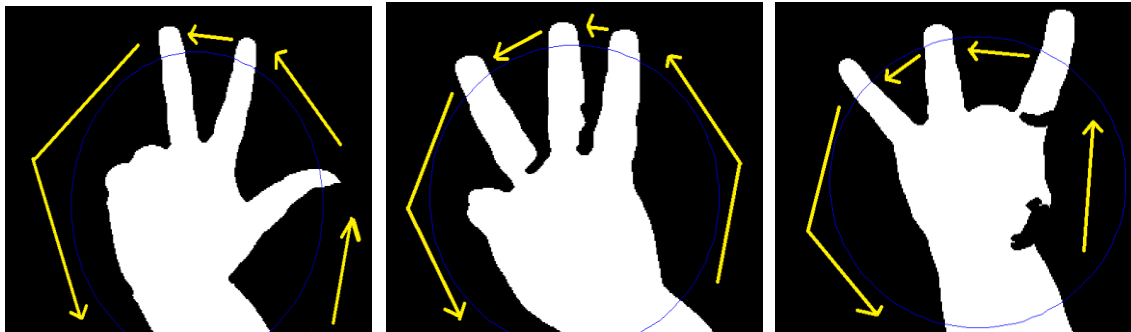


Fig. 4: Images of ASL number 3, 6 and 8 with arrows indicating the distance between adjacent open fingers.

Fig. 3 shows the different processing stages of signs of ASL numbers 3 and 8. Fig. 3.a and 3.e shows the captured RGB image. Fig. 3.b and 3.f shows the resized RGB to grey scale image. Fig. 3.c and 3.g shows the grey to binary converted image. Fig. 3.d and 3.h shows the plot of circle, cutting open fingers on the filtered and morphologically processed binary image.

Fig. 4 shows the images of ASL number 3, 6, 8 respectively with arrows indicating the direction of trace and the

successive distance between adjacent open fingers. As shown all the three number signs have exactly three fingers open but the successive distance between the open fingers is different.

Table 1 shows the summary of the total number of signs of ASL numbers 0 to 9 applied to the system and recognition rate. ASL number 0 to 5 are recognised 100% correctly where as other numbers from 6 to 9 are recognized with less percentage with overall average recognition rate of 92.31%.

Table 1: Classification Result of Test Images for ASL Numbers 0 To 9

O/P Detected Sign	I/P Test Images →									
	0	1	2	3	4	5	6	7	8	9
0	38	0	0	0	0	0	0	0	0	0
1	0	33	0	0	0	0	0	0	0	0
2	0	0	36	0	0	0	0	0	0	0
3	0	0	0	33	0	0	1	0	0	0
4	0	0	0	0	44	0	0	0	0	0
5	0	0	0	0	0	42	0	0	0	0
6	0	0	0	0	0	0	37	10	0	0
7	0	0	0	0	0	0	4	23	3	4
8	0	0	0	0	0	0	0	6	32	2
9	0	0	0	0	0	0	0	0	0	28
Total detected correctly	38	33	36	33	44	42	37	23	32	28
Total Images	38	33	36	33	44	42	42	39	35	34
% of Recognition	100	100	100	100	100	100	88.10	58.97	91.43	82.35
Average Recognition Rate	92.09%									

5. CONCLUSION

This paper proposes a simple feature of successive distances between the adjacent open fingers of the signs of ASL numbers for recognition. Here neural network is used for classification. As the method works perfect for the numbers 0 to 5, further by adding some more features along with this distance feature, the recognition rate for the numbers 6 to 9 can be improved. Also some other classification algorithms can be applied to same feature vector in order to increase the recognition rate.

6. REFERENCES

- [1] Henrik Birk and Thomas Baltzer Moeslund, "Recognizing Gestures From the Hand Alphabet Using Principal Component Analysis", Master's Thesis, Laboratory of Image Analysis, Aalborg University, Denmark, 1996.
- [2] Myron W. Krueger, *Artificial Reality II*, Addison-Wesley, Reading, 1991.
- [3] Thomas G. Zimmerman and Jaron Lanier, "A Hand Gesture Interface Device", ACM SIGCHI/GI, pages 189-192, 1987.
- [4] James Davis, and Mubarak Shah, "Recognizing hand gestures", ECCV, pages 331-340, Stockholm, Sweden, May 1994.
- [5] Y. Cui, D. Swets and J. Weng, "Learning-based Hand Sign Recognition using SHOSLIF-M", Proceedings of 5th International Conference on Computer Vision, pp 631–636, Boston, 1995.
- [6] Klimis Symeonidis, "Hand Gesture Recognition Using Neural Networks", Master's Thesis, School of Electronic and Electrical Engineering On August 23, 2000.
- [7] M. Lamar and M. Bhuiyant. "Hand alphabet recognition using morphological PCA and neural networks". International Joint Conference on Neural Networks, pages 2839–2844, Washington, USA, 1999.
- [8] Jonathan C. Rupe, "Vision-Based Hand Shape Identification for Sign Language Recognition", Master's thesis, Department of Computer Engineering Kate Gleason College of Engineering Rochester Institute of Technology Rochester, NY April 2005
- [9] X. Teng, B. Wu, W. Yu, and C. Liu, "A hand gesture recognition system based on local linear embedding", Journal of Visual Languages and Computing 16 (2005) 442–454.
- [10] E. Stergiopoulou, N. Papamarkos, "Hand gesture recognition using a neural network shape fitting technique", Engineering Applications of Artificial Intelligence Journal (2009).
- [11] U. Rokade, D. Doye, and M. Kokare, "Hand Gesture Recognition Using Object Based Key Frame Selection", International Conference on Digital Image Processing (2009).
- [12] W. Chung, X. Wu, and Y. Xu, "A Real-time Hand Gesture Recognition based on Haar Wavelet Representation", International Conference on Robotics and Biomimetics Bangkok, Thailand, February 21 - 26, 2009.
- [13] R. Rokade, D. Doye, and M. Kokare, "Hand Gesture Recognition by Thinning Method", International Conference on Digital Image Processing (2009).
- [14] Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari, "Finger Detection for Sign Language Recognition", Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.
- [15] Asanterabi Malima, Erol Özgür, and Müjdat Çetin, "A Fast Algorithm For Vision-based Hand Gesture Recognition For Robot Control", 2004.