Scientific
Research

Real-Time Static Hand Gesture Recognition for American Sign Language (ASL) in Complex Background

Jayashree R. Pansare[1], Shravan H. Gawande[2], Maya Ingle[3]

[1]Department of Computer Engineering, M. E. Society's College of Engineering, Pune, India; [2]Department of Mechanical Engineering, M. E. Society's College of Engineering, Pune, India; [3]Department of Computer Engineering, School of Computer Engineering, Devi Ahilya University, Indore, India.
Email: jayshreerahul@gmail.com

ABSTRACT

Hand gestures are powerful means of communication among humans and sign language is the most natural and expressive way of communication for dump and deaf people. In this work, real-time hand gesture system is proposed. Experimental setup of the system uses fixed position low-cost web camera with 10 mega pixel resolution mounted on the top of monitor of computer which captures snapshot using Red Green Blue [RGB] color space from fixed distance. This work is divided into four stages such as image preprocessing, region extraction, feature extraction, feature matching. First stage converts captured RGB image into binary image using gray threshold method with noise removed using median filter [medfilt2] and Guassian filter, followed by morphological operations. Second stage extracts hand region using blob and crop is applied for getting region of interest and then "Sobel" edge detection is applied on extracted region. Third stage produces feature vector as centroid and area of edge, which will be compared with feature vectors of a training dataset of gestures using Euclidian distance in the fourth stage. Least Euclidian distance gives recognition of perfect matching gesture for display of ASL alphabet, meaningful words using file handling. This paper includes experiments for 26 static hand gestures related to A-Z alphabets. Training dataset consists of 100 samples of each ASL symbol in different lightning conditions, different sizes and shapes of hand. This gesture recognition system can reliably recognize single-hand gestures in real time and can achieve a 90.19% recognition rate in complex background with a "minimum-possible constraints" approach.

Keywords: Image Preprocessing; Region Extraction; Feature Extraction; Median Filter; ASL

1. Introduction

Recognition of sign language is one of the major concerns for dump and deaf people. Sign language recognition is a research area involving pattern recognition, computer vision, natural language processing. Sign language recognition is a comprehensive problem because of the complexity of the visual analysis of hand gesture and the highly structured nature of sign language. As well as it is considered as a very important function in many practical communication applications, such as sign language understanding, entertainment, and human computer interaction (HCI). Among natural human gestures occurring during non-verbal communication, pointing gesture can be easily recognized and included in more natural new human computer interfaces. The video streams of backgrounds are frequently influenced by the background changes such as illumination changes and changes due to adding or removing parts of the background. Therefore, the quality of the foreground and the segmented image of hand gesture severely drop. The novel method proposed in [1] is based on difference background image between consecutive video frames, of using the "3σ-principle" of normal distribution for hand gesture detection to cope with the problem. The adaptive method of automatic threshold selection based on the method of maximal between-class variance is proposed for hand gesture segmentation to select optimal threshold. Lee [2] proposed a method to recognize hand gestures extracted from images with complex background for more natural interface. This method is based on obtaining the image through subtract one image from another sequential image, by measuring the entropy, separating hand region from images, tracking the hand region and recognizing hand gestures. The limitation for YCbCr segmentation method is that background should be plain and uniform. Rokade [3] proposed RGB segmentation which is more sensitive to light conditions and the threshold value for conversion of output image to binary

image that value is different for different lighting conditions. Fang [4] proposed a robust real-time hand gesture recognition method which is based on hand detection followed by hand tracking and then hand is segmented using motion and color cues. Messer [5] proposed static hand gesture recognition, which mainly consists of the recognition of well defined signs based on a posture of the hand. Hand gesture recognition system is proposed in [6] which is glove-free, fast, accurate and an effective sign recognition system. Rokade [7] uses Euclidean distance for shape similarity for hand gesture recognition. Ren [8] proposed Contour retrieve to get the object's contour. Feng [9] proposed research on Features Extraction from Frame Image Sequences using hand gesture contour algorithm (HGCA). Howe [10] proposed comparison of hand segmentation methodologies for hand gesture Recognition.

2. Flow of Hand Gesture Recognition System

In this section, the flow of hand gesture recognition system algorithm is presented as shown in **Figure 1**. The hand image with a resolution of 160×120 is first captured using Fron-Tech E-cam (**Figure 2**). A hand region is then extracted from image using skin detector. Feature extracted from filtered binary image and feature matching compares running and training images. Hand recognition uses closer matching hand gesture. **Figure 3** shows ASL symbols, but as this system supports static gestures only so J and Z symbols taken as different sign. And E, H, M, T signs are changed to avoid ambiguity while calculating features. Changed ASL symbols are shown in **Figure 4**.

3. Proposed Hand Gesture Recognition System

System Architecture shown in **Figure 5** consists of six stages of Hand gesture recognition system.

3.1. Image Capturing

First, confirm that you have the correct template for your paper size. This template has been tailored for output on the US-letter paper. Color is a robust feature. Image capturing is done by different color space methods. It is vulnerable to changing lighting conditions and it differs among people. In this proposed method, first of all, the snapshot of RGB image is captured by the 10 megapixel web camera. In this database, the image size is 160×120. If this size of image is used for the processing, the computation time will be very high.

3.2. Image Preprocessing

For skin detection purpose minimum and maximum skin probabilities of input RGB image are calculated. Data
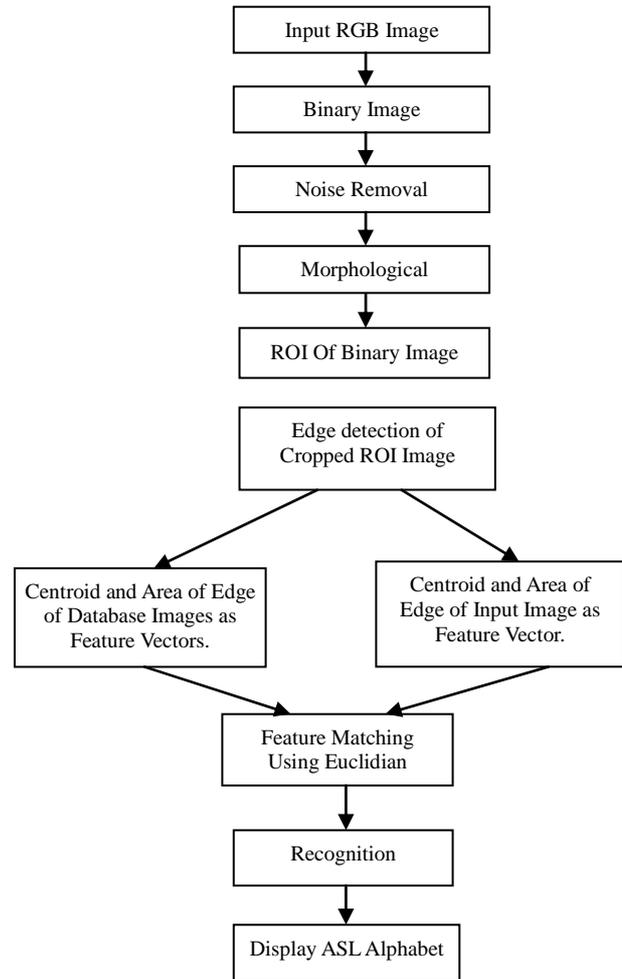

Figure 1. Flow of hand gesture recognition system.

Figure 2. Fron-Tech E-cam.

using normalized skin probability used for producing gray threshold at zero which gives binary image. Median filtering is applied to red channel of image to reduce "salt and pepper" noise and it is more effective than convolution because the goal is to simultaneously reduce noise and preserve edges. Then filtered binary image passed
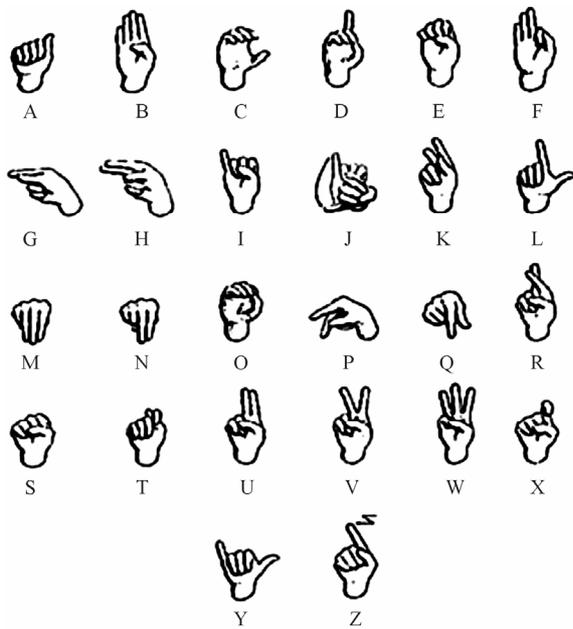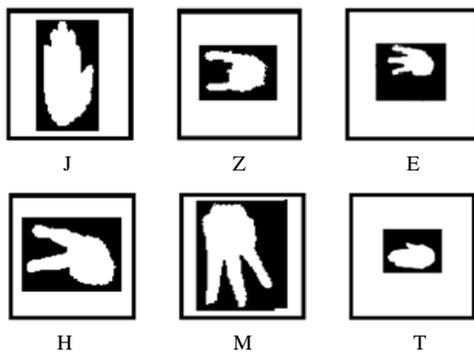
Figure 3. ASL symbols.
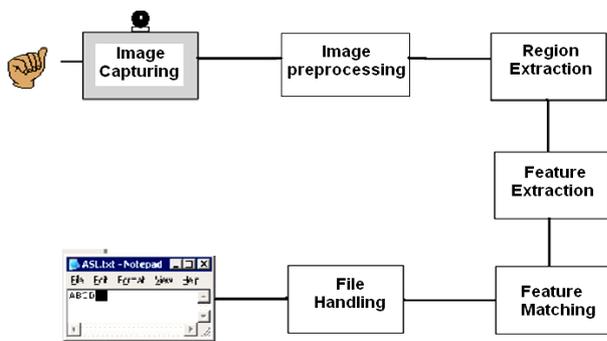

Figure 4. Modified ASL symbols.

Figure 5. System architecture for hand gesture recognition.

through Guassian filter of size 5×5 for smoothing of image. Removal of zeros are applied on normalized image. Morphological operation imclose applied on filtered image. Which performs closing the binary image and close operation is a dilation followed by erosion, using structuring element.

3.3. Region Extraction

8-Component connectivity of binary image is calculated using bwlebel and regionprops is applied to find bounding box for white pixels of hand image and imcrop is applied to extract region of interest (ROI). Edge detection of extracted image using "Sobel" method finds edges using the Sobel approximation to the derivative edges at those points where the gradient of input binary image I is maximum.

3.4. Feature Extraction

Feature vector is formed using centroid given by (C_x, C_y). And area of hand edge region given by bwarea.

$$\left(C_x, C_y\right) = \frac{1}{n}\left(\sum_{t=1}^{n} x_t, \sum_{t=1}^{n} y_t\right)$$

Feature vectors of training images are stored in mat files of MATLAB and feature vector of input hand gesture image are calculated at run time.

3.5. Feature Matching

Feature matching is done using Euclidian distance. Euclidian distance between feature vector of input real-time image and feature vector of each training image is calculated using following formula.

$$Lsc_t = \sqrt{\left(x_{t+1} - x_1\right)^2 + \left(y_{t+1} - y_1\right)^2}$$

Least Euclidian distance is used for recognition of perfect matching hand gesture.

3.6. File Handling

The recognized hand gestures are used to display concerned A-Z alphabet or word formed from combination alphabets.

4. Experimental Results

In this section, the performance of real time static hand gesture recognition system in complex background is evaluated for each of the 26 hand gestures 100 samples are stored in database. And running input image is used for the performance evaluation.

Based on the proposed algorithm the recognition results are reported in **Table 1** and graph of recognition results of 26 hand gestures is shown in **Figure 6**. The overall success rate is 90.19% with the recognition time about 0.5 seconds using 10 megapixel web camera and MATLAB 7.5.

5. Conclusion

The single hand gesture recognition system works successfully for real-time static hand gesture recognition

Table 1. Recognition results of 26 hand gestures.

Sign	Correct	Sign	Correct
A	90	N	95
B	86	O	86
C	90	P	90
D	85	Q	92
E	100	R	90
F	85	S	93
G	90	T	90
H	92	U	85
I	85	V	100
J	86	W	96
K	90	X	85
L	92	Y	92
M	90	Z	90
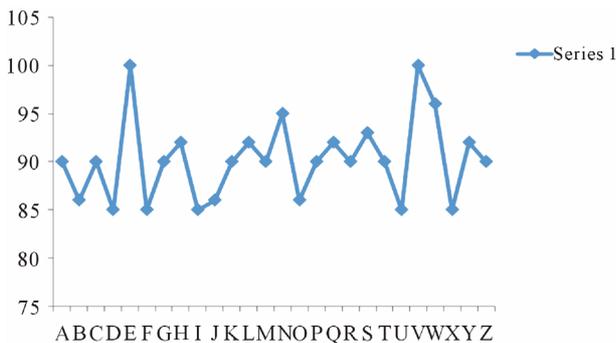
Figure 6. Graph of recognition results of 26 hand gestures.

system in different lightning conditions. From experiments it is observed that the results are more accurate with low-cost 10 mega pixel Fron-Tech E-cam as compared to 1.3 megapixel USB web camera and 1.3 megapixel Lifecam. Fixed position of Web camera mounted on the top of monitor or attached to laptop produces more results for RGB color space images captured from 1 meter distance from camera in complex background. Low-level feature extraction from extracted region reduces computation and works efficiently while matching feature vectors of real-time hand gesture images with feature vectors of training dataset. Total 2600 samples of hand gestures are collected having 100 samples per gestures of 26 gestures in different lightning conditions with different hand shapes and sizes produces more accuracy by using Euclidian distance for feature matching. The work is implemented in MATLAB 7.5. The experiments shows that the least Euclidian distance identifies perfect matching hand gesture and system achieves 90% hand

gesture recognition rate and is suitable for ASL alphabets and meaningful words and sentences.

6. Acknowledgements

REFERENCES

[1] Q. Y. Zhang, F. Chen and X. W. Liu, "Hand Gesture Detection and Segmentation Based on Difference Background Image with Complex Background," *Proceedings of the* 2008 *International Conference on Embedded Software and Systems*, Sichuan, 29-31 July 2008, pp. 338-343.

[2] J.-S. Lee, Y.-J. Lee, E.-H. Lee and S.-H. Hong, "Hand Region Extraction and Gesture Recognition from Video Stream with Complex Background through Entropy Analysis," *Proceedings of the* 26*th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Francisco, 1-5 September 2004, pp. 1513-1516.

[3] R. Rokade, D. Doye and M. Kokare, "Hand Gesture Recognition by Thinning Method," *International Conference on Digital Image Processing*, Bangkok, 7-9 March 2009, pp. 284-287. doi:10.1109/ICDIP.2009.73

[4] Y. K. Fang, K. Q. Wang, J. Cheng and H. Q. Lu, "A Real-Time Hand Gesture Recognition Method," *Proceedings of IEEE International Conference on Multimedia and Expo*, Beijing, 2-5 July 2007, pp. 995-998.

[5] T. Messer, "Static Hand Gesture Recognition," University of Fribourg, Switzerland, 2009.

[6] W.-K. Chung, X. Y. Wu and Y. S. Xu, "A Realtime Hand Gesture Recognition Based on Haar Wavelet Representation," *Proceedings of the* 2008 *IEEE International Conference on Robotics and Biometrics*, Bangkok, 22-25 February 2009, pp. 336-341. doi:10.1109/ROBIO.2009.4913026

[7] U. S. Rokade, D. L. Doye and M. Kokare, "Hand Gesture Recognition Using Object Based Key Frame Selection," *Proceedings of International Conference on Digital Image Processing*, Bangkok, 7-9 March 2009, pp. 288-291. doi:10.1109/ICDIP.2009.74

[8] Y. Ren and F. M. Zhang, "Hand Gesture Recognition Based on MEB-SVM," *Proceedings of the* 2009 *International Conference on Embedded Software and Systems*, Zhejiang, 25-27 May 2009, pp. 344-349. doi:10.1109/ICESS.2009.21

[9] Z. Q. Feng, P. L. Du, X. N. Song, Z. X. Chen, T. Xu and D. L. Zhu, "Research on Features Extraction From Frame Image Sequences," *Proceedings of* 2008 *International Symposium on Computer Science and Computational Technology*, Shanghai, 20-22 December 2008, pp.762-766.

[10] L. W. Howe, F. Wong and A. Chekima, "Comparison of Hand Segmentation Methodologies for Hand Gesture Recognition," *Proceedings of International Symposium on Information Technology*, Kuala Lumpur, 26-28 August 2008, pp. 1-7. doi:10.1109/ITSIM.2008.4631669