RESEARCH ARTICLE OPEN ACCESS

Hand Gesture Spotting Using Sign Language through Computer Interfacing

Neha S. Chourasia, Sampa Barman
(Department Of Electronics & Communication Engineering, RTMNU Nagpur)

ABSTRACT
Sign Language is the most natural and expressive way for the hearing impaired. A hybrid feature descriptor, which combines the advantages of SURF & Hu Moment Invariant methods, is used as a combined feature set to achieve a good recognition rate along with a low time complexity. To further increase the recognition rate and make the recognition system resilient to view-point variations, the concept of derived features from the available feature set is introduced. K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) are used for hybrid classification of single signed letter. This paper presents a methodology which recognizes the Indian Sign Language (ISL) and translates into a normal text. The methodology consists of three stages, namely a training phase, a testing phase and a recognition phase. Combinational parameters of Hu invariant moment and structural shape descriptors are created to form a new feature vector to recognize sign. Experimental results demonstrate that the proposed system can successfully recognize hand gesture with 96% recognition rate.
KEYWORDS: Finger Spelled Word Recognition, Hu Moment Invariant, Hidden Markov Model (HMM), Sign Language, Speeded Up Robust Features (SURF), Support Vector Machine (SVM).

I. INTRODUCTION

Sign Language is the means of communication among the deaf and mute community. Sign Language emerges and evolves naturally within hearing impaired community. Sign Language communication involves manual and non-manual signals where manual signs involve fingers, hands, arms and non-manual signs involve face, head, eyes and body. Sign Language is a well-structured language with a phonology, morphology, syntax and grammar. Sign language is a complete natural language that uses different ways of expression for communication in everyday life. Sign Language recognition system transfers the communication from human-human to human-computer interaction. The aim of the sign language recognition system is to present an efficient and accurate mechanism to transcribe text or speech, thus the "dialog communication" between the deaf and hearing person will be smooth. There is no standardized sign language for all deaf people across the world. However, sign languages are not universal, as with spoken languages, these differ from region to region. There are two main approaches used in the sign language recognition that is Glove/Device based and Vision based. In the glove based method the user has to wear a device which carries a load of cables so as to connect the device to a computer. Such devices are expensive and reduce the naturalness of the sign language communication. In contrast, the Vision based method requires only a camera and directly deals with image gestures. It is a two-step process: sign capturing and sign analysis .Vision based methods provide a natural environment to the user and reduces the complications as in the glove based method. Every country has its own sign language with a high level of grammatical variations. The sign language exists in India is commonly known as Indian Sign Language (ISL). It has been argued that the same sign language is used in Nepal, Sri Lanka, Bangladesh, and border regions of Pakistan. Examples of other sign languages are the American Sign Language (ASL), the British Sign Language (BSL), the Korean Sign Language (KSL), and so on.

1.1 BRIEF LITERATURE SURVEY

1. A method developed by Pallavi Gurjal, Real Time Hand Gesture Recognition Paper the algorithm in which video was captured and divided into various frames. The frame with the image was extracted and features like Difference of Gaussian, Scale space Feature Detector etc. extracted though SIFT which helped in gesture recognition [1].
2. A method developed by Archana S Ghotkar, Hand Gesture Recognition for ISL using Cam shift and HSV then recognizing gesture through Genetic Algorithm[2].
3. A method had been developed by P. Subha Rajan, recognising gestures for ISL where they proposed each gesture would be recognized through 7 bit orientation and generation process through RIGHT and LEFT scan [3].

4. A method had been developed by T. Shanableh, for recognizing isolated Arabic sign language gestures in a user independent mode. The signers wore gloves to simplify the process of segmenting out the hands of the signer via colour segmentation [4].
5. A method had been developed by Aleem Khalid Alvi, Pakistan SL Recognition Using Statistical Template Matching for Pakistan people to understand their language[5].
6. A method developed by T. Dasgupta, A Multilingual Multimedia ISL Dictionary Tool, to use the same SL in Nepal, SriLanka, Bangladeshand border regions of Pakistan[6].
7. A method had been developed by N. Otsu, A Threshold Selection method from Gray level Histograms, The segmentation of an input image of a hand gesture is performed using Global thresholding algorithm [7].
8. Byung - woo min, presented the visual recognition of static gesture or dynamic gesture, in which recognized hand gestures obtained from the visual images on a 2D image plane, without any external devices [8].

II. PROBLEM FORMULATION : NEED & SIGNIFICANCE OF PROPOSED RESEARCH WORK & OBJECTIVE

To develop an automatic sign language recognition system with the help of image processing and computer vision techniques. To use natural image sequences, without the signer having to wear data gloves or colored gloves, and to be able to recognize hundreds of signs. The motivation for this work is to provide a real time interface so that signers can easily and quickly communicate with non-signers.
To efficiently and accurately recognize signed words, from Indian Sign Language, using a minimal number of training examples.

III. PRINCIPLE OF THE PROPOSED METHOD

The proposed algorithm consisted of four major steps which are namely Image Acquisition, Feature Extraction, Orientation Detection and Gesture Recognition which is also shown in the below given Fig 1.

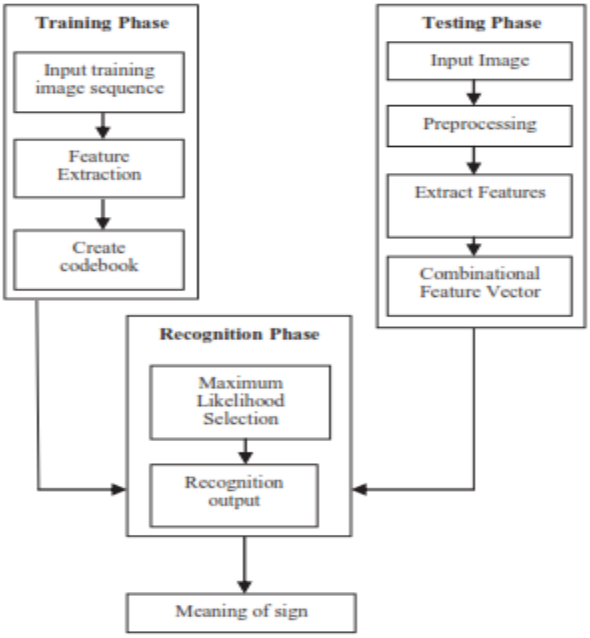

Fig. 1. Block diagram of hand gesture recognition system

All of the steps are explained in details with all the information on how the module is working and what behavior the module is supposedly expected to portray. While deciding on the following algorithm it was observed that pre-processing steps that are to be applied on the images for removal of noise in the background was not at all required.

Software Used: 1. Mat lab.
2.Image processing Toolbox.

Pre-processing: Pre-processing is applied to images before extracting features from hand images. Pre-processing consists of two steps, segmentation and filtering. All the functions are applied on a gray scale image. The segmentation of an input image of a hand gesture is performed using Global thresholding algorithm. The image level is divided into two classes one is hand and the other is background. The assigned value for hand pixel is "1" and the value for background pixel is "0".To eliminate noise from the segmented image a median filter is applied on the image.
The segmentation and filtering is applied below:

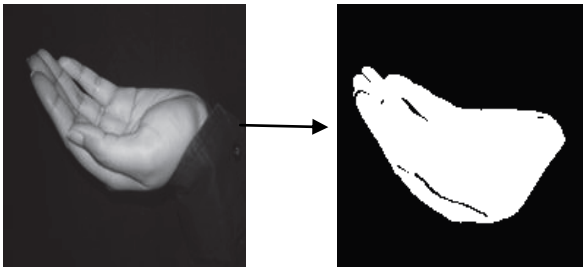

Fig. 2. Segmentation of gray scale image

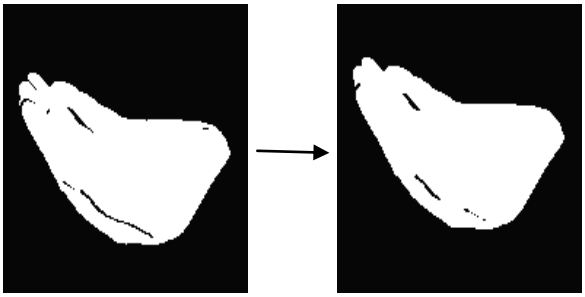

Fig. 3. Filtering of Segmented image

Fig. 4 Sign Database

Feature Extraction through Hu Invariant Moment: A moment known as, Hu invariant moment which is derived from the theory of algebraic invariant. Hu invariant moment is used for scale and position invariant pattern identification. The advantage of using Hu invariant moment is that it can be used for disjoint shapes. In particular, Hu invariant moment set consists of seven values computed by normalizing central moments through order three.

Multi-class Support Vector Machine: In the proposed work, a multi-class support vector machine has been used to appropriately classify the hand gestures among multiple classes. In the proposed approach a binary classifier is converted into multiclass classifier.

Recognition: The images in dataset are trained through MSVM. When an input image is given in the testing phase first the pre-processing is applied to the input image. Then the hu invariant moment and structural shape descriptor features of an input image are calculated and store in a codebook. Then features of the input image are matched with the codebook through MSVM. The most likelihood image is recognized and retrieved with their meaning.

IV. HAND GESTURE RECOGNITION SYSTEM

Hand gesture recognition system consists of the following steps (a) Pre-processing and hand segmentation, (b) Hand detection and tracking, (c) Hand posture recognition and (d) Hand gesture classification as shown in figure 1.

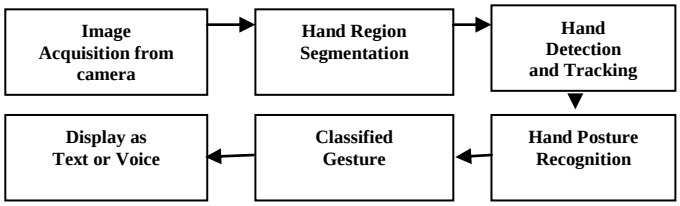

Figure 5. Block diagram of Hand Gesture Recognition System

This section describes about individual letter recognition approach, consisting of hand segmentation, extraction of invariant feature descriptor, recognition of letters and derived feature approach for improving accuracy.

4.1 HAND SEGMENTATION

Skin color segmentation is performed using k-means clustering method. RGB color frames I(m,n,p), (where m, n and p are number of rows , number of columns and number of color planes) are converted into 3 1-dimensional feature vector X of single column and m*n rows. Then I = { X(1), X(2), ….. X(N)} (N=3) is obtained. Through experimental observation classification of colors using 3 clusters (K) under Euclidean distance measure provides better performance. For every pixel in the input, k-means returns an index corresponding to a cluster. Skin pixel region is identified from the different color regions using a thresholding method in RGB color space where the threshold value is selected experimentally. Repeat the cluster for 3 times to avoid local minima. Figure shows example results of the segmentation algorithm.

4.2 RECOGNITION OF LETTERS

The bounding box of the detected hand in each frame is obtained from the previous section. To recognize the posture of detected hand, a combined feature extraction methodology using Speeded Up Robust Features (SURF) and Hu Moment Invariant features is incorporated. Bounding box, BBIm (x,y) is taken as test image. Features are calculated and compared with the database features. Minimum

Euclidean distance between the feature vectors recognizes particular hand posture/letter.

4.3 IMPROVING ACCURACY THROUGH DERIVED FEATURES

To further increase the recognition rate and speed of processing, prominent features are derived from the available data set of features using forward selection algorithm. To improve the performance of classifier on a dataset, it is possible to evaluate each feature's deviation. The deviation is computed per feature xj in the set of N features x={x1,...,xN} by calculating the sum of all differences between the calculated resultant feature vector when feature xj is left out and the actual resultant feature vector oi of sample si in the dataset D=NxP, containing P samples, where each sample si contains N features si={si 1,...si N}. . For clarity, we'll define a new feature set y that excludes xj , such that y={x1,...,xj- 1,xj+1,...,xN}.

V. RESULT AND DISCUSSIONS

For proving the effectiveness and accuracy of the proposed system, we have carried out a number of experiments. A sign image dataset has been created, which contains 720 images. The dataset is classified as a testing set that has 600 images and a training set that has 120 images. 720 images include different signs from ISL of 12 different people. These are divided into 60 classes. The accuracy of the proposed system is evaluated as follows:

$$Accuracy = \left(\frac{correctly\ classified\ gestures}{total\ no\ of\ gestures} \right) * 100\%$$

The resolution of grabbed image is 256*256. Each image is resized by a magnification factor of 0.5 to reduce the computation process. All operations are performed on gray scale image.

VI. CONCLUSION

In this paper, an automatic ISL recognition system has been created which works on real-time basis and employs the use of combinational feature vector with a MSVM classifier. In the combinational feature vector, Hu invariant moment and structural shape descriptors are used collectively for achieving better recognition results. The use of MSVM increases recognition performance. Results demonstrate that the combination of invariant moments and shape descriptors gives better result, as shape descriptors define the boundary of an image while the invariant moments are invariant to change in scale and position of an image. The future directions will focus on signer independent, large vocabulary systems in both isolated and continuous recognition tasks.

REFERENCES

[1] Pallavi Gurjal, Kiran Kunnur, "Real Time Hand Gesture Recognition using SIFT", International Journal for Electronics and Engineering, 2012, pp 19-33.

[2] Ghotkar, Archana S, "Hand Gesture Recognition for ISL", International Conference on Computer Communication and Informatics (ICCCI), 2012, pp 1-4.

[3] Rajam, P. Subha and Dr G Bala krishnan," Real Time ISL Recognition System to aid Deaf and Dumb people", 13th International Conference Vol. 2 Issue 4, April – 2013 ISSN: 2278-0181386 on Communication Technology (ICCT), 2011, pp 737742.

[4] Shanableh Tamer, T. Khaled, "Arabic sign language recognition in user independent mode", IEEE International Conference on Intelligent and Advanced Systems, 2007, pp 597-600.

[5] Aleem Khalid Alvi, M. Yousuf Bin Azhar, "Pakistan Sign Language Recognition Using Statistical Template Matching", International Journal of Information Technology, 2004, pp 1-12.

[6] T. Dasgupta, S. Shukla, S. Kumar, S. Diwakar and A. Basu,"A Multilingual Multimedia ISL Dictionary Tool," The 6th Workshop on Asian Language Resources, 2008, pp 57-64.

[7] N. Otsu, "A Threshold Selection method from Gray level Histograms,"IEEE Transactions on Systems, Man and Cybernetics, vol. 9, no. 1, pp 62–66, 1979.

[8] Byung-woo min, Ho-sub yoon, Jung soh, Takeshiohashi and Toshiaki jima, "Visual Recognition of Static/Dynamic Gesture: Gesture-Driven Editing System", Journal of Visual Languages & Computing Volume10, Issue 3, June 1999, pp 291-309.