Vision Based Hand Gesture Recognition for Indian Sign Languages

Gobhinath.S
Assistant Professor,
SKCET, Coimbatore.

Dhivya.R
UG Scholar,
SKCET, Coimbatore.

Karnila.M
UG Scholar,
SKCET, Coimbatore.

Abstract—Sign Language is a language which uses visually transmitted sign patterns to convey meaning by simultaneously combining hand shapes, orientation and movement of the hands to fluently express one's thoughts / communicate with others and is commonly used by the physically impaired people who cannot speak or hear. Automatic Sign Language system requires fast and accurate techniques for identifying static signs or a sequence of produced signs to help interpret their correct meaning. Major component of a Sign Language is Hand Gesture Recognition. Gestures are powerful means of communication among humans. It may be based on hand, body, lip movements (speech expression), eye movements, facial expression (mood/emotions) of the human. Among all, hand gesture is the easiest and most natural way of communication between people. With the latest advances in the field of computer vision, image processing and pattern recognition, real -time vision-based hand gesture classification is belonging more and more feasible for human-computer interaction in virtual environments. Hand gestures are an intuitive yet powerful communication modality which has not been fully explored for Human Computer Interaction. Sign language study shows that among various gesture communication modalities, hand gesture plays significant role. This research aim towards working on hand gesture recognition for sign language interpretation as a HCI application.

Index Terms—Indian Sign Language (ISL), Hand Gesture Recognition (HGR), Human Computer Interaction (HCL), Support Vector Machine (SVM) *and* Local Binary Pattern (LBP)

I. *INTRODUCTION*

Sign language (SL) is a natural language used for communication by hearing and / or speech impaired persons. It is a language which uses manual communication and body language to convey meaning, as opposed to acoustically conveyed sound patterns. They share many similarities with spoken languages (sometimes called "oral languages", which depend primarily on sound), which is why linguists consider both to be natural languages, but there are also some significant differences between signed and spoken languages [1]. Though sign language is spread and used all over the world, it is not universal. Wherever deaf community exist, sign languages develop. As is the case with spoken languages, sign languages also vary from region to region. Hundreds of sign languages are in use around the world and are at the core of local deaf cultures. Some sign languages have obtained some form of legal recognition, while others have no status at all. Regionally American Sign Language (ASL), French Sign Language (LSF), German Sign Language (GSL), Quebec Sign Language (LSQ), Indian Sign Language (ISL), British Sign Language (BSL) etc. have been evolved.

Indian Sign Language is one of the first known sign language systems and is considered extremely important in the history of sign languages, but it is rarely used today. In linguistic terms, sign languages are as rich and complex as any spoken language, despite the common misconception that they are not "real languages". Professional linguists have studied many sign languages and found that they exhibit the fundamental properties that exist in all languages [2] [3]. The elements of a sign are Hand shape (or Hand form), Orientation (or Palm Orientation), Location (or Place of Articulation), Movement, and Facial Expression summarized in the acronym HOLME [4].

Sign languages, like spoken languages, organize elementary, meaningless units (phonemes) into meaningful semantic units. Like in spoken languages, these meaningless units are represented as features, although often crude distinctions are also made in terms of manual and non-manual parameters. The manual parameters include hand shape, hand orientation, location and motion whereas the non-manual parameters include gaze, facial expression, mouth parameters, position and motion of the trunk and head [5]. Sign languages are independent of spoken languages and follow their own paths of development. The grammars of sign languages do not usually resemble that of spoken languages used in the same geographical area in fact, in terms of syntax, ASL shares more with spoken Japanese than it does with English. Actually, sign languages can convey meaning more than spoken languages by simultaneous means, e.g. by the use of space, two manual articulators, and the signer's face and body.

II. LITERATURE SURVEY

The spoken and written language of a country is different from other country. Although the same language has been used by a number of countries, however, the syntax and semantics of a language is dependent on a country/region. For example, English is the official language of the UK, USA and many other nations. The usage of English differs at country level. Similarly, the sign language of a country is not similar than that other country. The focus of this study is on the development of sign languages at international level [2].

As discussed in the introduction, the development of sign language is for each country is at varied with time. The sign languages listed in table 1 presents some of the important international sign languages. The focus of this chapter is also on literature survey in the field of sign language recognition. Therefore, only brief discussions on linguistics characteristics of BSL (British Sign Language), ASL (American Sign Language), JSL (Japanese Sign Language), CSL (Chinese Sign Language) and ISL (Indian Sign Language) are given in this table[1, 4].

Table 1 : International Sign Languages

S. No.	Country	Sign Language	Abbn.
1	United Kingdom	British Sign Language	BSL
2	United States of America	American Sign Language	ASL
3	Japan	Japanese Sign Language	JSL
4	Republic of China	Chinese Sign Language	CSL
5	**Republic of India**	**Indian Sign Language**	**ISL**
6	Ukraine	Ukrainian Sign Language	UKL
7	Republic of Sri Lanka	Sri Lankan Sign Language	SLTSL
8	Federative Republic of Brazil	Brazilian Sign Language	Libras

III. HAND GESTURE RECOGNITION

A sign in a Sign Language (SL) as discussed earlier consists of three main parts: Manual features, non-manual features and finger spelling [1]. To interpret the meaning of a sign, all these parameters are to be analyzed simultaneously. Sign language poses a main challenge of being multichannel. Each channel in this system is separately built and analyzed and the output of each channel is combined at the final stage to draw conclusion.

The research in Sign Language Interpretation (SLI) began with Hand Gesture Recognition (HGR). Hand gestures are extensively used in human non-verbal communication by hearing impaired and speech impaired people. Even normal people sometimes use sign language for communication. Though sign language is spread and used all over the world, it is not universal. Wherever hearing impaired community exists, sign languages develop. To make communication between the hearing impaired and normal people simple and efficient, it is necessary that this process be automated. Number of techniques have been developed for automatic HGR. The overall process of Hand Gesture Recognition (HGR) system block diagram is as shown in figure 1. There are three similar steps in HGR:

1. Hand acquisition which deals with hand extraction from the given static image and hand extraction and tracking from a video.
2. Feature extraction which basically deals with compressed representation of the data which will facilitate the recognition of the hand gesture.
3. Classification/ recognition of the hand gesture following some rule.

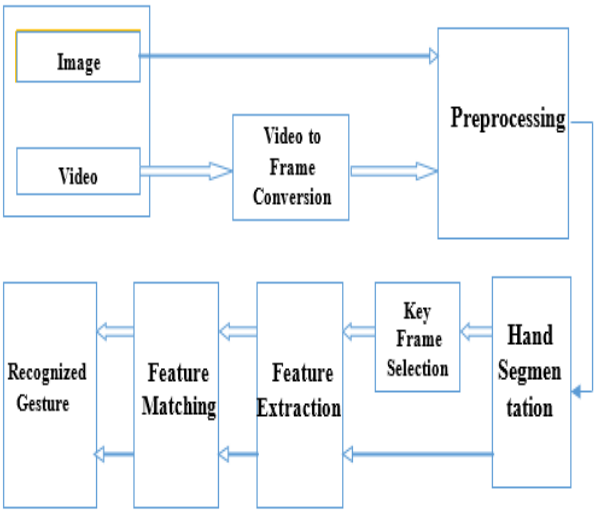

Figure 1. Block Diagram for Process of Hand Gesture Recognition

IV. DATA SETS ACQUISTION

In this research, two different data sets are created for the development of an ISL recognition system. The data sets are ISL digits (0-9) and single handed ISL alphabets (A-Z). For data set acquisition, dark background for uniformity and easy in manipulation of images for feature extraction and classification is initially chosen. A Sony digital camera, Cyber shot H70, is used for capturing the images. All images are captured with flash light in an intelligent auto mode. The common file format JPEG is used to capture the images as it is a common image standard now a days. Each original image is 4608×3456 pixel and required approximately 5.5 MB storage space. To create an efficient data set with a reasonable size, the images are cropped to 200×300 RGB pixels and barely 25 KB memory space is required per image. The data set is collected from 100 signers. Out of these signers, 69 are male and 31 are female with average age group of 27. The average height of a signer is about 66 inches. The data set contains isolated ISL numerical signs (0-9). Five images per ISL digit sign is captured from each signer.

Therefore, a total of 5,000 images are collected in the data set. After the creation of the data set, it is divided into training set and testing set. The sample images of the data set are shown in figure 2.

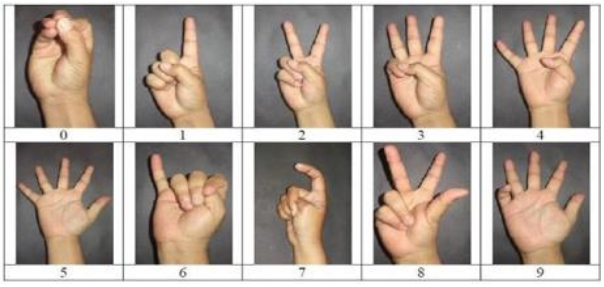

Figure 2. The ISL Digit Signs Data set

In the first data set, 100 signs of each single handed characters of ISL are captured using a digital camera, as described above. So, in this data set, a total of 2600 images cropped to 200×300 RGB pixel sizes are collected. The images are collected from four males and six females. Each signer contributed 10 samples of ISL alphabet sign. The average age and height of signers in this data set is 27 years and 66 inch respectively. The backgrounds of sign images are dark, as only hand orientations are required for the feature extraction process. The images are stored in JPEG format because it can be easily exported and manipulated in different hardware and software environments. Each pre-processed ISL sign image required nearly 25 KB of storage space with 72 dpi. The size of the images is 200×300 pixels. The skin colors of these images are neither very dark complexion nor very white complexion. This is due to the proposed application can be useful in India only. The colors corresponding to human skins are mainly used in capturing the sign images. A sample data set is shown in figure 3.

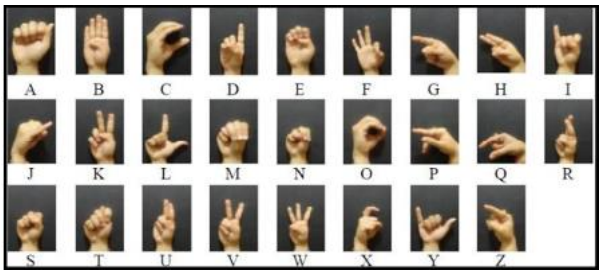

Figure 3. The ISL Single Handed Alphabet Signs Data Sets

V. HAND SEGMENTATION

Segmentation is used to detect hand from background [5]. The experimentation in this work is carried out using two datasets representing hand gestures performed with one hand for alphabets A to Z using Indian Sign Language. The images of this dataset before and after preprocessing stage are shown in figure 4.

(a)

(b)

Figure 4. (a) Original images of alphabet 'A', (b) Images after RGB to Gray conversion and resizing.

VI. FEATURE EXTRACTION

The feature extraction [6] approaches in image processing, extracts valuable information present in an image. This deals with conversion of a high dimensional data space into lower dimensional data space. The lower dimensional data extracted from images should contain precise information which is the representative of the actual image. The image can be reconstructed from the lower dimensional data space. The lower dimensional data is required as input to any classification technique as it is not feasible to process higher dimensional data with speed and accuracy. The

inputs to an automatic sign language recognition system are either static signs (images) or dynamic signs (video frames) [2]. In order to classify input signs in an automatic sign language recognition system, extraction of valuable features from signs is required. All the algorithms that are used for facial feature extraction are used for Hand feature extraction as well. Each one of the feature extraction technique used is described in short below.

VI.1 LOCAL BINARY PATTERN

Local Binary Pattern (LBPs) have proven to be very effective for image representation and have been applied in various analysis. The LBPs are tolerant against monotonic illumination changes and are able to detect various texture primitives like spot, line end, edge, corner etc. The most popular and efficient version of LBP i.e. Block LBP (figure 5) with uniform / no uniform patterns is used as the first technique for extracting hand features [6].

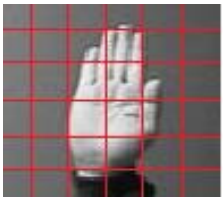

Figure 5. Block LBP Operator

VI.2 .SUPPORT VECTOR MACHINE CLASSIFIER

Classification is a technique part of machine learning. The technique is used to classify each item in a data set into one of a predefined set of groups. Classification methods use mathematical models including linear programming, decision trees, statistics and neural networks for pattern classification. In classification, a software module is created that could learn the art of classifying the data items into different groups. With initial experimentation using multiclass SVM and decision trees, a huge amount of misfits have been identified in the process of classification. Hence these classifiers are not further used for final experimentation towards recognition.

Figure 6. Output of ISL Digits Produced by the System

Although 26 classes are present in ISL single handed alphabet, the system is able to predict single handed characters with more than 95% accuracy. This is possible with LBP and SVM feature extraction technique. In figure 6, a sample output is shown for single handed ISL sign 'B'. The input sign image is processed through the system and a prediction is shown in the right hand side of the output screen. The sign interpreted as single handed 'B' which is the correct prediction. The ISL alphabets in English domain are A, B, C,..., Z. There is no distinction between lower case and upper case letters in ISL alphabet [5, 6].

V. CONCLUSION

In this work a vision-based automatic sign language recognition was presented which is able to recognize sentences in American Sign Language. Several features and different methods to combine them were investigated and experimentally evaluated. Tracking algorithms with applications to hand and head tracking were presented and experiments were carried out to determine the parameters of these algorithms. An emphasis was put on appearance-based features that use the images itself to represent signs. Other systems for automatic sign language recognition usually require a segmentation of input images to calculate features for the segmented image parts. The results presented in this work show that the usage of appearance-based features yields a promising recognition performance.

REFERENCES

[1] Sahoo, A. K. and K. K. Ravulakollu. Indian Sign Language Recognition using Skin Color Detection. International Journal of Applied Engineering Research (IJAER). vol. 9, no. 20, 2014, pp. 7347-7360.

[2] Joyeeta Singha, Karen Das " Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique " (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 2, 2013.

[3] Noor Adnan Ibraheem and Rafiqul Zaman Khan,"Survey on Various Gesture Recognition Technologies and Techniques" International Journal of Computer Applications, Volume 50 – No.7, July 2012.

[4] Matt Huenerfauth, Elaine Gale, Brian Penly, Mackenzie Willard, Dhananjai Hariharan, "Comparing methods of displaying language feedback for student videos of american sign language", Proceedings of the 17th InternationalACM SIGACCESS Conference on Computers & Accessibility. ACM, pp. 139-146, 2015.

[5] Chenyang Zhang, Xiaodong Yang, YingLi Tian, "Histogram of 3d facets: A characteristic descriptor for hand gesture recognition", Automatic Face and Gesture Recognition, pp. 1-8, 2013.

[6] Raviteja Vemulapalli, Felipe Arrate, Rama Chellappa, "Human action recognition by representing 3d skeletons as points in a lie group", CVPR. IEEE, pp. 588-595, 2014.