# Sign Energy Images for Recognition of Sign Language at Sentence Level

Chethana Kumara B.M. Department of Studies in Computer Science University of Mysore Mysore, India

## ABSTRACT

In this paper, the task of sign language recognition at sentence level is addressed. The idea of Sign Energy Image (SEI) and a method of extracting Fuzzy-Gaussian Local Binary Pattern (FzGLBP) features from SEI to characterize the sign are explored. The suitability of interval valued type symbolic data for efficient representation of signs in the knowledgebase is studied. A Chi-square proximity measure is used to establish matching between reference and test signs. A simple nearest neighbor classification technique is used for recognizing signs. Extensive experiments are conducted to study the efficacy of the proposed system. A data base of signs called UoM-ISL is created for experimental analysis.

# **General Terms**

Pattern Recognition; Image processing, Sign language Recognition, Symbolic data analysis.

## **Keywords**

Fuzzy Gaussian LBP, Interval valued features, Sign Energy Image, Sign language, Video sequence

# **1. INTRODUCTION**

Communication is an essential activity for human beings to express their needs and feelings in their day-to-day life. Natural languages have been evolved over several decades and are used by normal people for communication. However, differently abled (hearing impaired) people who are not able to hear and speak cannot use natural languages as a medium of communication. In order to overcome this limitation, the community of hearing impaired people has practiced a medium of communication using hand movements and facial expression. This medium of communication is evolved overtime, standardized to some extent and is called sign language.

The sign language used by the hearing impaired people is not fully understood by the community of people who use natural language. Thus, there is a communication gap between the normal and hearing impaired society. Several researchers have made an attempt to minimize this gap through computer vision technology. Pattern recognition and image processing techniques have been explored to translate the sign into natural language text. However, from the literature survey it is evident that the work done in this direction is very less and it is still at infant stage. Lot of challenges need to be addressed to develop a complete automated system of sign language recognition / interpretation. Most of the research work done on sign language recognition is limited to fingerspelling / word level recognition, which seems to be not feasible and cumbersome. Since the signs made by the signers are very abstract, it is difficult to extract them at fingerspelling or word level. Thus, there is a need to map the signs to natural language sentences. To the best of our knowledge, no Nagendraswamy H.S. Department of Studies in Computer Science University of Mysore Mysore, India

concrete work has reported in this direction. Also, there are lots of challenges to be addressed in accomplishing the task.

# 2. RELATED WORKS

Several researchers have made an attempt to propose automatic sign language recognition (SLR) system from the past two decades. Many systems have been developed for American [18, 41], Australian [31], Korean [34], Arabic [1, 2, 3], British [9], Chinese [19], French [4], and German [8, 16] sign languages. But no concrete work has been reported for Indian Sign Language [ISL].

Several methods such as Elastic graph matching [44], modified census transformation [30], weighted Eigenspace size function [33], Hidden Markov Models (HMM) [3], weighted combination of geometric and appearance based features such as the area, gravity center [8], moment based size function [26], outline hand length and downscale intensity directly extracted from images in video frames [45] have been studied for user-independent hand posture recognition.

Moments were also explored for static gesture sign recognition. Zernike [22] and Pseudo-Zernike moments features introduced in [11] were used for static gesture sign recognition [24]. More recently, geometric and orthogonal moments for view independent gesture classification were introduced in [39]. The Tchebichef moments computed from internal and external hand edges for hand shape representation and hand posture recognition was reported in [15]. Kernel principal component analysis (KPCA) and non-parametric discriminant analysis (NDA) are applied to recognize Australian Sign Language and use KPCA to detect discriminated features from the KPCA feature space can be found in [36].

In [32], segment and merge approach for subject independent continuous sign language recognition was reported. Bayesian network was used to segment the given continuous sign into sub segments and a two layer CRF was used to merge and recognize the sub segments. They used CyberGlove and Polhemus FASTRACK database for their experiments. Author's dataset comprises of eight signers, in which seven are deaf persons and one person is hearing person. A detailed survey on automatic sign language analysis can be found in [38].

A graph matching technique for recognizing continuous gesture of Arabic Sign Language is proposed in [43]. In this approach, gestures have been divided into elementary static postures. The gestures, which represent alphabets or words, are stored in the database as model graphs. An Algorithm based on tree paradigm was used for graph and sub graph isomorphism detection. They used 158 static postures comprising of 50 dynamic gestures and achieved 80 percentage of recognition accuracy. A framework for sign

language sentence recognition by commonsense context is proposed in [29]. They considered a set of signs for Italian Sign Language (ISL) for domestic applications. The proposed framework is composed of three main steps: 1) sensorial input is processed in order to obtain a feature vector by standard image processing algorithms; 2) a Self-Organizing Map (SOM) neural network is used to classify the single sign and to provide a list of probable meanings; and 3) a commonsense engine to choose the right meaning depending on the context of the whole sentence. They conducted experiments on two sets of signs consisting of 20 and 40 signs called experiment-A and experiment-B respectively. They reported 83.30 percentage and 82.50 percentage of correct translating sentences for experiment-A and experiment-B respectively.

From the literature survey, it is very much clear that several researchers have proposed different models to address the problem of automatic sign language recognition. Most of the works reported in the literature have addressed the problem at fingerspelling level [5, 12, 13, 17, 24, 35, 42] and few attempts were made at word level [10, 14, 19, 27]. It can be observed from the sign language that it is very abstract and it is difficult for a signer to express sentences using fingerspelling alphabets or words. Therefore, the problem must be addressed at sentence level. Only few attempts [6, 7, 29, 32, 43] were made in this direction. However, the works have not treated the entire sentence as a single unit. Also, a practical continuous sign language recognition system needs to be signer independent. In order to address the problem of signer independent recognition system, it is necessary to device a robust representation system, which is tolerant to signer variations and yield good generalization; however, this has not received much attention in the literature except the work reported in [28, 45].

In view of this, in this research, an attempt has been made to recognize the signs at sentence level. The proposed method captures the variations in sign due to different signers or due to different instances by exploring the concept of symbolic data analysis [21, 25, 37] particularly the interval type data. Also, in order to reduce the complexity of processing every frame for feature extraction and keyframe selection for representation, we propose to exploit the idea of Sign Energy Image (SEI), which aggregates the entire video (sequence of frames) into a single frame producing the texture image that captures the signer variations in the form of varying intensity.

Rest of the paper is organized as follows: Section 3 presents the proposed methodology including preprocessing, segmentation of region of interest, generating SEI, feature extraction and representation. Experimental results are presented in section 4 followed by conclusion in section 5

#### 3. PROPOSED METHODOLOGY

The Proposed sign language recognition system tries to capture the hand movements of a signer by aggregating the frames of the videos of sign into single frame called Sign Energy Image (SEI). The pixel intensity of SEI varies depending on the hand movements of a signer. The static and dynamic information of a signer during signing is captured through aggregated pixel intensities in SEI.

The pixel intensity captured through SEI is characterized by exploring fuzzy Gaussian Membership function [40] with LBP [23] texture model called Fuzzy Gaussian LBP texture model. Interval-valued type symbolic feature vector have been used to effectively represent the sign in the knowledgebase to make the system more robust. A Chi-square symbolic proximity measure is used to compute the dissimilarity between test sign feature vector and reference sign feature vector. A simple *K*-NN classification technique is applied to recognize a given test sign as one among the reference sign in the knowledgebase. An overview of the various stages of the proposed sign language recognition system can be schematically depicted as shown in Figure 1 and the following subsections give a detailed description about the proposed methodology.



Figure 1 Overview of the proposed sign language recognition system.

#### 3.1 Preprocessing

In any computer vision / image processing application, preprocessing is an essential stage and plays a vital role. A continuous sign made by a signer is captured through a camera and is preprocessed to remove noise if any. All the frames in a sign video are extracted by preserving their sequence. Further, these frames are subjected to color transformation from RGB to HSV. Skin region segmentation is carried out as in [20] based on HSV color space. Series of morphological operations are applied on the segmented frame to produce more clear and noiseless hand and face regions of a signer. Due to hand movements, horizontal alignment may vary from frame to frame. In order to achieve the standard size, a size normalization factor is introduced by super imposing the resultant frame with actual frame size. It includes proportionally resizing each binary image into actual frame having the same height and width. Figure 2 shows an example preprocessing and segmentation of region of interest.



Figure 2 Preprocessing and segmentation of region of interest.

#### **3.2 Sign Energy Images**

A sign captured by the video contain several frames and the number of frames varies from sign to sign. In many video processing applications, it is not necessary to consider all the frames for analysis / processing as the consecutive frames do not differ significantly. In order to address this problem, idea of key frame selection from videos is suggested by the researchers. In some real-time video processing applications like sign language recognition, several videos corresponds to several signs need to be processed. Even if key frames are considered as representatives for videos, the size and processing complexity will become difficult to manage when the database size is considerably large. This problem has motivated us to think of capturing the entire sign video information in a single frame and study the performance of the sign recognition system. Thus, the idea of Sign Energy Image (SEI) is explored in this work to reduce the storage and processing complexity of the sign recognition system. SEI for a given sign can be computed as follows.

Let  $F_1, F_2, F_3...F_u$  be the *u* number of frames extracted from the given sign video. Each frame is processed has discussed in section 3.1 to produce a binary frame with segmented hand and face regions. The Sign Energy Image (SEI) is obtained by aggregating the pixel intensities of the frames as follows

$$SEI_{i}(x, y) = \frac{1}{u} \sum_{j=1}^{u} F_{k}(x, y)$$
(1)

Where  $F_k$  (x, y) denote the intensity of a pixel at the location (x, y) of  $k^{\text{th}}$  frame of a sign. Figure 3 shows an example SEI computed for a sign.



Figure 3 Sequence of binary images aggregated to produce SEL.

SEI reflects the intensity changes over hand movements in a sequence of frames. Higher pixel intensity value at a position reflects the frequent hand movements at that position.

#### **3.3 Fuzzy Gaussian LBP Histogram**

Once a SEI is obtained for a sign, descriptive features are extracted from SEI for its representation. Fuzzy-Gaussian membership function is used to fuzzify the intensity values of SEI and then basic LBP texture descriptor is applied to characterize the images.

Let  $\mu$  and  $\sigma$  respectively denote the mean and standard deviation of the intensity of pixels in a 3X3 neighbourhood of pixels in an image. Let  $G_P$  denote the grey value of a pixel in a topology of circle with radius *R* and *N* number of neighbors and  $G_c$  denote the centre pixel. The values  $\mu$  and  $\sigma$  are computed as follows:

$$\mu = \frac{\sum_{i=1}^{P} G_i + G_c}{(P+1)}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{P} (G_i - \mu)^2 + (G_c - \mu)^2}{(P+1)}}$$

The degree of belongingness of each pixel's intensity value within 3X3 neighbourhood of pixels to a fuzzy set is computed using the fuzzy Gaussian membership function as follows:

$$\mu_{\tilde{I}}^{-}(Gp) = e^{-\frac{1}{2} \left(\frac{Gp - \mu}{\sigma}\right)^2}$$
(2)

Let  $\mu_{\tilde{I}}(G_P)$  (for  $P = 1, 2, 3, \ldots, 8$ ) and  $\mu_{\tilde{I}}(G_c)$  respectively denote the computed fuzzy membership values of intensity of boundary pixels and the centre pixel within the 3X3 neighbourhood as described in the above paragraph. The Fuzzy Gaussian LBP number (FzGLBP) can be computed as follows:

$$F_{z}GLBP = \sum_{i=1}^{P} S * 2^{i-1}$$

$$i=1$$

$$re \qquad S = \begin{cases} 1 & \text{if } \mu_{\tilde{I}}(G_{\boldsymbol{P}}) \ge \mu_{\tilde{I}}(G_{\boldsymbol{C}}) \\ 0 & \text{otherwise} \end{cases}$$
(3)

Figure 4 shows an example  $3\times3$  neighborhood pixel and the computation of Fuzzy Gaussian LBP code. Figure 5 shows an example instances of two SEIs and the corresponding Fuzzy Gaussian LBP Histograms computed for a sign "*I Want Coffee*". Figure 6 to Figure 8 shows example instances of SEIs for different signs.

whe



Figure 4 Computation of Fuzzy Gaussian LBP code.



Figure 5 (a) (and (b) shows Fuzzy Gaussian LBP Histogram for SEI of two different instances of sign "*I want Coffee*".



Figure 6 SEI of Ten different instances of sign "I want Coffee".



Figure 7 SEI of Ten different instances of sign "I want note book".



Figure 8 SEI of Ten different instances of sign "I want pen".

#### 3.4 Symbolic Representation of Signs

Further, in order to capture intra-class variations of a particular sign and to choose multiple representatives, all the instances of a sign are clustered using hierarchical clustering technique. In order to obtain natural clusters, the idea of inconsistency coefficient is used during clustering. Inconsistency coefficient values obtained depend on how the samples in a class are clustered at various levels. In order to find the adaptive threshold to cut the dendrogram to obtain natural clusters, the maximum inconsistent coefficient value where all samples are clustered into one cluster is considered and the standard deviation ( $\gamma$ ) of all the non-zero inconsistency coefficient values is obtained. The threshold can be computed by subtracting the standard deviation ( $\gamma$ ) multiplied by a small constant value ( $\delta$ ) from the maximum inconsistency coefficient values as follows:

Th = max (Inconsistency coefficients) – 
$$(\delta * \gamma)$$
 (4)

The values for  $\delta$  are empirically chosen through several experiments.

Let  $S = \{S_{l_{1}}, S_{2}, S_{3,...}, S_{n}\}$  be the *n* number of signs considered to train the system for recognition. Let  $S_{k} = \{S_{k}^{(1)}, S_{k}^{(2)}, S_{k}^{(3)}, ..., S_{k}^{(m)}\}$  be the *m* number of instances of a sign  $S_{k}$  captured from different signers or same signers at different instances of time and  $SEI_{k} = \{SEI_{k}^{(1)}, SEI_{k}^{(2)}, SEI_{k}^{(3)}, ..., SEI_{k}^{(m)}\}$  be the corresponding *m* Sign Energy Images obtained respectively for *m* instances of a sign  $S_{k}$ .

Let  $F_k^{p} = \{F_{k1}^{p}, F_{k2}^{p}, F_{k3}^{p}, ..., F_{ku}^{p}\}$  be the feature vector of u - dimension describing  $p^{\text{th}}$  instance of  $k^{th}$  SEI. The m instances of SEIs of a sign  $S_k$  are clustered to capture intraclass variations of a sign  $S_k$ .

Let  $C_k = \{C_{k1}, C_{k2}, C_{k3}, ..., C_{kd}\}$  be the *d* number of clusters obtained for a particular sign and let  $T_l^{(ij)} = \{TF_{l1}^{(ij)}, TF_{l2}^{(ij)}, TF_{l3}^{(ij)}, ..., TF_{lu}^{(ij)}\}$  be the feature vector of u-dimension describing the texture of  $l^{th}$  sign sample belonging to  $j^{th}$  cluster of  $i^{th}$  class, in general.

$$RS^{(ij)} = \left\{ \left[ TF_1^{(ij)-}, TF_1^{(ij)+} \right], \left[ TF_2^{(ij)-}, TF_2^{(ij)+} \right], \left[ TF_3^{(ij)-}, TF_3^{(ij)+} \right], \dots, \left[ TF_u^{(ij)-}, TF_u^{(ij)+} \right] \right\}$$
(8)

Likewise, symbolic feature vectors are computed for all the clusters of all the signs. The collection of all the symbolic feature vectors effectively represents signs of all classes of a dataset in the knowledgebase.

#### **3.5 Matching and Recognition**

In order to recognize the given test sign, the sign is preprocessed and its SEI is computed.

Assuming that there is r number of samples in a cluster, in general, the texture feature vectors describing all the r SEI samples belonging to  $j^{th}$  cluster of  $i^{th}$  class as shown in eq (5)

$$T_{1}^{(ij)} = \left\{ TF_{11}^{(ij)}, TF_{12}^{(ij)}, TF_{13}^{(ij)}, \dots, TF_{1u}^{(ij)} \right\}$$
$$T_{2}^{(ij)} = \left\{ TF_{21}^{(ij)}, TF_{22}^{(ij)}, TF_{23}^{(ij)}, \dots, TF_{2u}^{(ij)} \right\}$$
$$T_{3}^{(ij)} = \left\{ TF_{31}^{(ij)}, TF_{32}^{(ij)}, TF_{33}^{(ij)}, \dots, TF_{3u}^{(ij)} \right\}$$

$$T_r^{(ij)} = \left\{ TF_{r1}^{(ij)}, TF_{r2}^{(ij)}, TF_{r3}^{(ij)}, \dots, TF_{ru}^{(ij)} \right\}$$
(5)  
The variation among the  $k^{th}$  feature is captured by

The variation among the  $k^{in}$  feature is captured by consolidating the feature values to an interval as follows:

Let 
$$\mu T_k^{(ij)} = \frac{1}{r} \sum_{l=1}^r TF_{lk}^{(ij)}$$
 and  $\sigma T_k^{(ij)} = \left[\frac{1}{r} \sum_{l=1}^r \left(TF_{lk}^{(ij)} - \mu T_k^{(ij)}\right)^2\right]^{1/2}$  be the mean and

standard deviation of the  $k^{th}$  texture feature values of all the *r* SEI samples belonging to  $j^{th}$  cluster of  $i^{th}$  class respectively.

Now, the variations with respect to the  $k^{th}$  feature due to texture among all samples in a cluster is captured as follows:

$$TF_k^{(ij)+} = \mu T_k^{(ij)} + \sigma T_k^{(ij)}$$
<sup>(6)</sup>

$$TF_{k}^{(ij)-} = \mu T_{k}^{(ij)} - \sigma T_{k}^{(ij)}$$
<sup>(7)</sup>

Thus,  $\left[TF_k^{(ij)-}, TF_k^{(ij)+}\right]$  is the interval representing the  $k^{th}$  feature value due to texture of sign samples belonging to  $j^{th}$  cluster of  $i^{th}$  class.

Hence, the feature vectors representing the reference sign samples in terms of texture belonging to  $j^{th}$  cluster of  $i^{th}$  class is of interval valued type as follows:

Let 
$$TS = \{TF_1, TF_2, TF_3, ..., TF_u\}$$
 be the *u*-dimension texture feature vector representing the test sign sample to be recognized. Chi-square distance measure is explored for the purpose of comparing reference sign feature vector with the test sign feature vector as follows

Dissimilarity between the test and reference sign with respect to  $k^{th}$  feature represented interms of crisp and interval type data is computed as

$$DS(TS_k, RS_k]) = \begin{cases} 0 & if(TF_k^{(ij)-} \le TF_k \le TF_k^{(ij)+}) \\ min\left[\frac{\left(TF_k^{(ij)+} - TF_k\right)^2}{\left(TF_k^{(ij)+} + TF_k\right)}, \frac{\left(TF_k^{(ij)-} - TF_k\right)^2}{\left(TF_k^{(ij)-} + TF_k\right)}\right] otherwise \end{cases}$$
(9)

The total similarity value between test and reference sign due to all the features is given by

$$DS(TS,RS) = \frac{1}{u} \sum_{k=1}^{u} DS(TS_k,RS_k])$$
(10)

where u = 256.

First Nearest Neighbour classification technique is used to classify the given test sign sample as one among the known sign class.

## 4. EXPERIMENTATION

In order to study the efficacy of the proposed sign language recognition system, an extensive experimentation was conducted on a considerable more number of videos of signs. A sign language database which consisting of videos of signs describing the sentences used by the hearing impaired people in their day-to-day life is created. The signs made by the students from different schools of Mysore zone are captured in a controlled environment.

A single digital camera is used to capture signs at the rate of 30 frames / second in VGS model. The Digital camera is placed in front of the signer to capture the signs made by the signer. Once the sign videos are captured, a pre-processing phase is followed as explained in section 3.1. A particular sing made by the different signers at different instances possess variations with respect to time takes to complete a sign. In order to study these variations among the signs made by the same signer at different instances of time and also by the different signers, a sign is captured with four different signers and each signer performing ten repetitions of signs. Hence, the UoM-ISL dataset contains 1040 videos of 26 different signs. Thus, there are 1040 videos of signs (17.3 Hours) used for the experimental analysis. Several experiments are conducted for different percentages of training and testing. The experiments are repeated for 50 random sets of training and testing samples.

The values for  $\delta$  in Eqn. (4) can be in the range 0. 1 to 1. However, the number of clusters is significantly changed for only some values of  $\delta$ . In our experiments, it is found that the number of clusters obtained for  $\delta = (0.1, 0.5, 1)$  are significantly different and hence, these three values are chosen for three different experiments on dataset of signs considered.

Performance of any classification or recognition system is measured in terms of its accuracy, precision, recall and Fmeasure and are defined as follows: These measures are

defined on the basis of correct positive (CP), correct negative (CN), false positive (FP) and false negative (FN) for the overall test samples (NS).

$$Accuracy = \frac{CP + CN}{NS} \tag{11}$$

$$Precision = \frac{CP}{CP + FP} \tag{12}$$

$$Recall = \frac{CP}{CP + FN}$$
(13)

$$F - measure = \frac{(2*Precision*Recall)}{(Precision+Recall)}$$
(14)

Table 1 presents the recognition rate in terms of average F-measure of 50 random trails for different percentages of training and testing samples.

The confusion matrices have shown in Figure 7and 8 presents the class wise performance of the proposed method interms of F-measure. From the above results, we have observed that the proposed method has shown good recognition rate for most of the signs. Proposed model achieved the overall average recognition accuracy of 79.13 percentage for 60:40 ratios of training and testing with 324 numbers of representatives. Experiments have also revealed the fact that higher recognition rate is achieved for more number of representatives with more number of training samples. With this observation, we understand that more representatives are needed for a class when there are more intra-class variations in that class and hence the total sign representatives also increase for the entire sign database.

In order to demonstrate the efficacy of the proposed Fuzzy Gaussian LBP feature extraction technique, with interval valued type symbolic representation scheme, we compared the results obtained by the proposed methodology on UoM-ISL data set with that of LBPV feature extraction technique with symbolic representation scheme. The reason for considering LBPV feature extraction technique is that it is accepted as the best texture descriptor among the variants of LBP. Table 2 presents the results obtained for LBPV feature extraction technique with symbolic representation. It can be observed from the results presented in Table 1 and Table 2 that the proposed Fuzzy Gaussian LBP technique has yielded better results compared to LBPV technique.

Also, the results of the proposed methodology are compared with different combinations of feature extraction and representation techniques. Table 3 presents the overall average recognition accuracy in terms of F-measure for different combinations. It can be observed from the results that the proposed feature extraction and representation technique has yielded better results compared to other combinations. First nearest neighbor classification technique is used in all the experiments.

Sign Language Dataset	Ratio of Training and	Number of	Recognition rate in terms of
	Testing samples	representatives	average F-measure
UoM – ISL Dataset	60:40	211	$70.43 \pm 1.70$
		291	$75.92 \pm 0.83$
		324	$79.13 \pm 1.22$
	50:50	211	$68.38 \pm 1.38$
		267	$74.76 \pm 1.36$
		324	$75.76 \pm 1.45$
	40:60	211	$68.38 \pm 0.38$
		291	$74.28 \pm 1.36$
		324	$76.81 \pm 0.76$

Table 1: Overall classification performance of the proposed method

Table 2: Overall classification	performance of LBPV features	with symbolic representation
---------------------------------	------------------------------	------------------------------

Sign Language Data	Ratio of Training and	Number of	Recognition Rate
Set	Testing	representatives	Average for overall Classes
UoM –ISL Sign Language Data	60:40	188	$66.48 \pm 0.70$
		264	$69.19 \pm 0.66$
		314	$77.08 \pm 0.60$
	50:50	188	$64.86 \pm 1.12$
		264	$67.08 \pm 0.75$
		314	$73.99 \pm 0.55$
	40:60	188	$60.92 \pm 0.59$
		264	$61.42 \pm 1.10$
		314	$70.50 \pm 0.60$

In order to demonstrate the superiority of the proposed model (FzGLBP + Symbolic representation) with other feature extraction and representation techniques, experiments were also conducted using LBPV with crisp and symbolic representation and the results are tabulated in Table 3.

In our experiments, the number of training samples used for LBPV with crisp representation is 624 for 60:40 ratios and an average recognition accuracy of 30.40 percentage is obtained, whereas for the proposed method has yielded 79.13 percentage of recognition accuracy of 79.13 for 60:40 ratios with only 324 representatives. The results clearly show that the proposed feature extraction and representation techniques performs better compared to other combinations.

# 5. CONCLUSION

In this paper, an attempt has been made to address the problem of sign language recognition at sentence level. The proposed method exploits the idea of Sign Energy Image (SEI), texture description and symbolic data analysis concepts to characterize and effectively represent a sign. The idea of SEI reduces the number of frames needed to be processed and also the problem of uneven number of frames captured for a sign. Since, SEI aggregates the entire video information into a single frame, it reduces the overall time complexity and storage space requirement. In addition, as the method explored the concept of symbolic data analysis, it provides effective representation capturing intra-class variation in a sign. To study the efficacy of the proposed approach, extensive experiments were carried out on a considerably large database of signs created by us. The experimental results illustrate that the proposed method shows good performance in terms of precision, recall and F-measure rates. Also, in order to demonstrate the superiority of the proposed fuzzy Gaussian LBP feature extraction technique with that of LBPV with crisp and symbolic representation techniques, several experiments were carried out for different combination of feature extraction and representation

techniques. The results are compared and found that the proposed combination has yielded better results.

Table 3: Results for different combinations of feature		
extraction and representation		

Different combinations	Recognition rate (Average F- measure)
LBPV with crisp representation (624 training samples)	30.40
LBPV with symbolic representation (314 training samples)	77.08
FzGLBP with crisp representation (624 training samples)	61.02
FzGLBP with symbolic representation (Proposed Method) (324 training samples)	79.13

# 6. ACKNOWLEDGEMENT

**Authors** would like to thank the students and the teaching staff of Sai Ranga Residential Boy's School for Hearing Impaired, Mysore, and N K Ganpaiah Rotary School for physically challenged, Sakaleshpura, Hassan, Karnataka, INDIA, their immense support in the process of UoM - ISL Sign language data set creation.



Figure 7 Confusion matrix for (a) (60:40), (b) (50:50) and c) (40:60) training and testing samples respectively ( $\delta = 1$ ).



Figure 8 Confusion matrix for (a) (60:40), (b) (50:50) and c) (40:60) training and testing samples respectively ( $\delta = 0.5$ ).

#### 7. REFERENCES

- O. Al-Jarrah, A. Halawani, Recognition of gestures in Arabic sign language using neuro-fuzzy systems, Artif. Intell. 133 (1-2) 117–138, 2001.
- [2] M. Al-Roussan, M. Hussain, Automatic recognition of Arabic sign language finger spelling, Int. J. Comput. Appl. (IJCA) 8 (2) 80–88 (Special issue on Fuzzy Systems), 2001.
- [3] M. Al-Roussan, K. Assaleh, A. Talaa, 2009: Videobased Signer independent Arabic sign language recognition

using hidden-Markov models, Appl. Softw. Comput. 9 990–999,2009.

- [4] O. Aran, T. Burger, A. Caplier, L. Akarun, A beliefbased sequential fusion approach for fusing manual signs and non-manual signals, Pattern Recognit. 42 812–822, 2009.
- [5] Aryanie, D. and Heryadi, Y. American sign languagebased finger-spelling recognition using k-Nearest Neighbors classifier. 3rd International Conference on Information and Communication Technology (ICoICT), 2015.
- [6] K. Assaleh, T. Shanableh, M. Fanaswala, F. Amin and H. Bajaj. Continuous Arabic Sign Language Recognition in User Dependent Mode. Journal of Intelligent Learning Systems and Applications, Vol. 2 No. 1, pp. 19-27, 2010.
- [7] B. Bauer, H. Hienz, Relevant features for video-based continuous sign language recognition, in: FG00 Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 440-445, 2000.
- [8] B. Bauer, K.F. Kraiss, Video-based sign recognition using self-organizing subunits. Proceedings of the 16th International Conference on Pattern Recognition, pp. 434–437, 2002.
- [9] R. Bowden, D. Windridge, T. Kadir, A. Zisserman, M. Brady, A linguistic feature vector for the visual interpretation of sign language, in: Proceedings of the Eighth European Conference on Computer Vision, pp. 391–401,2004.
- [10] Brashear, H., Starner, T., Lukowicz, P., & Junker, H. Using multiple sensors for mobile sign language recognition. Georgia Institute of Technology,2003.
- [11] C-C. Chang, J.J. Chen, W.K. Tai, C.C. Han, New approach for static gesture recognition, J. Inf. Sci. Eng. 22,1047–1057,2006.
- [12] H. Cooper, B. Holt, R. Bowden, Sign language recognition, Chapter in Visual Analysis of Humans: Looking at People, pp. 539–562, 2011.
- [13] Daniel Kelly, John McDonald., A person independent system for recognition of hand postures used in sign language. Charles Markham., Pattern Recognition Letters 31 (2010) pp1359–1368, 2010.
- [14] César Roberto de Souza and Ednaldo Brigante Pizzolato. Sign language recognition with support vector machines and hidden conditional random fields: going from fingerspelling to natural articulated words. MLDM'13 Proceedings of the 9th international conference on Machine Learning and Data Mining in Pattern Recognition. Pages 84-98,2013.
- [15] Djamila ahmani and Slimane Larabi.,User-independent system for sign language finger spelling recognition Journal of Visual Communication and Image Representation Vol 25, Iss 5,pp1240–1250, July 2014.
- [16] P. Dreuw, D. Stein, T. Desealers, D. Rybach, M. Zahedi, H. Ney, Spoken language processing techniques for sign language recognition and translation, Technol. Disability 20,2008.
- [17] Ebling, S., Wolfe, R., Schnepp, J., Baowidan, S., McDonald, J., Moncrief, R., ... & Tissi, K. Synthesizing

the finger alphabet of Swiss German Sign Language and evaluating the comprehensibility of the resulting animations. In 6th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT) (p. 10), September, 2015.

- [18] The American sign language hand shape dictionary, Gallaudet University,1998.
- [19] W. Gao, G.L. Fang, D.B. Zhao, Y.Q.A. Chen, A Chinese sign language recognition system based on SOFM/SRN/HMM, Pattern Recognit. 37,2389–2402, 2004
- [20] Gasparini, Francesca, and Raimondo Schettini. "Skin segmentation using multiple thresholding." Electronic Imaging 2006. International Society for Optics and Photonics, 2006.
- [21] Gowda, K. Chidananda, and Edwin Diday.: Symbolic clustering using a new dissimilarity measure. Pattern Recognition, 24.6: 567-578, 1991.
- [22] L. Gu, J. Su, Natural hand posture classification based on Zernike moments and hierarchical classifier, IEEE Int. Conf. Robotics Autom. 3088–3093, 2008.
- [23] Zhenhua Guo, Lei Zhang, David Zhang. Rotation invariant texture classification using LBP variance (LBPV) with global matching". Pattern Recognition 43, pp.706–719, 2010.
- [24] Guru, D. S., and Suraj, M. G. Recognition of postal codes from fingerspelling video sequence. International Journal of Image and Graphics. 2009.
- [25] Guru, D. S., and H. S. Nagendraswamy.: Clustering of interval-valued symbolic patterns based on mutual similarity value and the concept of k-mutual nearest neighborhood. Computer Vision–ACCV 2006. Springer Berlin Heidelberg, 234-243, 2006.
- [26] Handouyahia, M., Ziou, D., Wang, S., Sign language recognition using moment-based size functions. In: Proc. Intl. Conf. on Vision Interface, pp.210–216, 1999.
- [27] Eun-Jung Holden, Gareth Lee, and Robyn Owens. Australian sign language recognition. Mach. Vision Appl. 16, 5, 312-320, December 2005.
- [28] Nagendraswamy, H. S., Chethana Kumara B M., Guru, D. S., & Naresh, Y. G. Symbolic Representation of Sign Language at Sentence Level. IJIGSP, 9, 49-60. DOI: 10.5815/ijigsp.2015.09.07 2015.
- [29] Ignazio Infantino, Riccardo Rizzo, and Salvatore Gaglio. A framework for sign language sentence recognition by commonsense context". Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 37. No 5 1034-1039, 2007.
- [30] A. Just, Y. Rodriguez, S. Marcel, Hand posture classification and recognition using the modified census transform, in: 7th Internat. Conf. on Automatic Face and Gesture Recognition, FGR. Pp. 351–356, 2006.
- [31] M.W. Kadous, Machine recognition of Australian signs using Powergloves: Towards large-lexicon recognition of sign languages, in: Workshop on the Integration of Gestures in Language and Speech, Wilmington Delaware, 1996.

- [32] Kong W W., Surendra Ranganath. Towards subject independent continuous sign language recognition: A segment and merge approach. Pattern Recognition 47 (2014) 1294–1308, 2014.
- [33] D. Kelly, J. McDonald, C. Markham, A person independent system for recognition of hand postures used in sign language, Pattern Recognit. Lett.31, 1359– 1368, 2010.
- [34] J.S. Kim, W. Jang, Z. Bien, A dynamic gesture recognition system for the Korean sign language (KSL), IEEE Trans. Syst. Man Cybern. 26 (2) 354–359, 1996.
- [35] Kindiroglu, A. A., Yalcin, H., Aran, O., Hrúz, M., Campr, P., Akarun, L., & Karpov, A. Automatic recognition fingerspelling gestures in multiple languages for a communication interface for the disabled. Pattern Recognition and Image Analysis, 22(4), 527-536, 2012.
- [36] Wun-Guang Liou and Chung-Yang Hsieh and Wei-Yang Lin. Trajectory-based sign language recognition using Discriminant Analysis in higher-dimensional feature space. IEEE International Conference on Multimedia and Expo (ICME), pp1-4,2011.
- [37] Mohan kumar H P and H S Nagendraswamy. Change Energy Image for Gait Recognition: An approach based on symbolic representation. Int. J. Image Graphics Signal Proc. (IJIGSP) 6 (4) 1-8, 2014.
- [38] Sylvie C.W. Ong and Surendra Ranganath. Automatic Sign Language Analysis: A Survey and the Future beyond Lexical Meaning. IEEE transactions on pattern analysis and machine intelligence, vol. 27, no. 6, june 2005.
- [39] S.P. Priyal, P.K. Bora, A study on static hand gesture recognition using moments, in: IEEE International Conference on Signal Processing and Communications (SPCOM), pp. 1–5, 2010.
- [40] Ross, T. J.: Fuzzy logic with engineering applications. John Wiley & Sons. 2009.
- [41] T. Starner, A. Pentland, Real-time merican sign language recognition from video using hidden Markov models, in: AAAI Fall Symposium on Disabilities, Cambridge, MA, 1996.
- [42] Suraj and Guru D S. Secondary diagonal FLD for fingerspelling recognition. International Conference on Computing: Theory and Applications, ICCTA'07, 2007.
- [43] Tolba, M. F., Ahmed Samir, and Magdy Aboul-Ela. Arabic sign language continuous sentences recognition using PCNN and graph matching". Neural Computing and Applications 23.3-4: 999-1010, 2013.
- [44] J. Triesch, C. Von der malsuburg. Classification of hand postures against complex backgrounds using elastic graph matching, Image Vision Comput. 20 (13-14) 937– 943. 2002.
- [45] Nagendraswamy, H. S., BM Chethana Kumara, and R. Lekha Chinmayi. "GIST Descriptors for Sign Language Recognition: An Approach Based on Symbolic Representation." Mining Intelligence and Knowledge Exploration. Springer International Publishing, 2015. 103-114.