Vision based multi-feature hand gesture recognition for indian sign language manual signs

Article *in* International Journal on Smart Sensing and Intelligent Systems · March 2016

DOI: 10.21307/ijssis-2017-863

2 authors:

Gajanan K. Kharate
Matoshri College of Engineering & Research Centre

Archana Ghotkar
Pune Institute Of Computer Technology

www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

VISION BASED MULTI-FEATURE HAND GESTURE RECOGNITION FOR INDIAN SIGN LANGUAGE MANUAL SIGNS

Gajanan K. Kharate [1] and Archana S. Ghotkar [2]

[1]Matoshri College of Engineering and Research Centre, Department of Electronics and Telecommunication Engineering, Savitribai Phule Pune University, Nashik,

[2]Pune Institute of Computer Technology, Department of Computer Engineering, (Research Scholar in I2IT, Hinjawadi, Pune) Savitribai Phule Pune University, India.

gkkharate@rediffmail.com, archana.ghotkar@gmail.com

Submitted: Nov. 4, 2015   Accepted: Jan. 4, 2016   Published: Mar. 1, 2016

Abstract- Indian sign language (ISL) is the main communication medium among deaf Indians. An ISL vocabulary show that the hand plays a significant role in ISL. ISL includes static and dynamic hand gesture recognition. The main aim of this paper is to present multi-feature static hand gesture recognition for alphabets and numbers. Here, comparative analysis of various feature descriptors such as chain code, shape matrix, Fourier descriptor, 7 Hu moments, and boundary moments is done. Multi-feature fusion descriptor is designed using contour (Boundary moments, Fourier descriptor) and region based (7Hu moments) descriptors. Analysis of this new multi-feature descriptor is done in comparison with other individual descriptors and it showed noteworthy results over other descriptors. Three classification methods such as, Nearest Mean Classifier (NMC), k-Nearest Neighborhood (k-NN) and Naive Bayes classifier are used for classification and comparison. New Multi-feature fusion descriptor shows high recognition rate of 99.61% among all with k-NN. Real time recognition for number signs 0-9, of fusion descriptor with NMC gave 100% accuracy.

Index terms: Indian sign language interpretation, k-nearest neighborhood classifier, nearest mean Classifier, Boundary moments, Fourier descriptor, 7 Hu moments, Chain code, Shape matrix, Naïve Bayes.

I. INTRODUCTION

Hand gesture recognition (HGR) plays a significant role in any sign language recognition (SLR) and it is one of the requirements of any HCI application. Number of hearing impaired and deaf people is very large in India as compared to other countries [1]. Since 2010, Indian researchers have started working on ISL interpretation which will definitely help Indian deaf people as a communication mode in near future. Various researchers are working on various sign language recognition (SLR) such as American sign language [2, 3], Polish SL [4], Chinese SL [5] and other SLs. In India, sign language varies from state to state like spoken language so, researchers are also working on their native sign language. Pansare et al.[6] worked on the Devnagri sign language translation using histogram and achieved an average accuracy of 87.82% for manual alphabets. Subha and Balkrishnan [7] considered South Indian sign language and achieved 98.44% accuracy using feature point extraction. Geetha and Manjusha [8] worked on ISL alphabets and numerals using B-spline approximation method and achieved up to 90% accuracy. Singha and Das [9] worked on the ISL interpretation for 24 alphabets and achieved 96.25% success rate whereas, Tewari and Kumar [10] also worked on static ISL alphabets using a Kohonen self-organizing map and achieved up to 80% recognition rate. For ISL manual alphabet and numeral recognition static hand signs are needed. Shape is a very important feature of hand. Various shape descriptor methods such as contour or region based [11] are available as a feature extraction technique for static hand gesture recognition. SLR problem is tackled by two approaches viz. data glove based and vision based [12]. Many early systems used data glove based approach for acquisition purpose [2]. But the data glove based approach has limitation such as restriction of freedom of hand, need to wear cumbersome devices and high cost of hardware unit. Vision based system has become more popular due to the features of flexibility and low cost. Researchers working on vision based hand gesture recognition for HCI application face the problem of variable lighting condition, dynamic background and skin color detection [29]. ISL vocabulary shows that, majority of signs are manual (hand gesture) signs compared to non-manual (face, lips and other body gestures) signs. Most of the non-manual signs are used with manual signs [12]. Hand gestures are classified into static and dynamic hand gestures. Researchers who are working on static hand gesture recognition applied to sign language used various region based [13], contour based [14] and texture based [16] features [32]. Some researchers worked on local feature extraction methods for numerals [7, 15, 16]. For classification and recognition of hand gesture recognition, work has been done on various classification methods [19] such as nearest neighborhood classification with Euclidean distance [18] and other similarity measures [17], Bayes classifier [20], neural network [21, 22], hybrid recognizer [23], HMM [5]. ISL static sign consists of alphabets and numerals signs. Figure 1 shows that few signs such as (E, F, X), (M, N, R), (G, S), (2, V, U) and (5, 9) are ambiguous in nature. ISL manual signs are one as well as two handed signs for which efficient feature descriptor (invariant to translation, rotation and scale) is required. Available individual shape descriptor is not sufficient to deal with HGR for complex ISL manual signs. So, there is a need for designing efficient multi-feature shape descriptor which will be robust, less noise sensitive and having better computational complexity. The scope of the paper is restricted to static manual signs for better explanation. In this paper, new vision based multi-feature descriptor for hand gesture recognition has been presented which can be used in any object recognition application in general. Analysis of new descriptor has been done with other descriptors. Performance analysis has also been checked with real- time HGR. Recognized signs are displayed in text as well as in voice so that blind people can also get enabled to the system. Person independent skin color segmentation algorithm with dynamic skin color feature extraction is developed. Boundary moments (BM), Fourier descriptor (FD), and chain code are contour based, while 7Hu moments (Hu) and shape matrix are region based methods selected for feature extraction. An attempt has been made to develop new multi-feature fusion descriptor by combining contour and region based descriptors such as, fusion of (BM, Hu and FD). Multi-feature fusion descriptor gave noteworthy recognition rate compared to individual descriptor and other spatial domain descriptor such as shape matrix and chain code. For performance analysis, recognition is done using three classifiers viz. Nearest mean classifier (NMC), Nearest neighborhood (NN) and Naive Bayes classifier. Performance analysis of each descriptor with NMC, NN, k-NN and Naive Bayes has been given. Testing has been done on real-time HGR for number recognition. The rest of the paper is organized as given. Section 2 introduces manual data set of ISL. Section 3 shows methodology. Section 4 presents experimental result and analysis for hand gesture recognition. Section 5 presents comparative analysis with existing ISL work. Conclusion and Future work covers in section 6.

II. DATASET FOR ISL

ISL is a visual-spatial language. The present work is focused on static hand gesture so, possible signs based on static hand gestures such as, alphabets and numbers are considered for testing the proposed algorithm. Figure 1 shows ISL manual alphabets and numbers [24,36]. For ISL interpretation, there is no source of image/video standard dataset. So, in this research, around 3600 samples images are captured on 40 persons for training and testing with different lighting condition, angle and distance. Two data sets are formed out of collected samples. Dataset 1 consist of ideal sign images with little change in sign posture, and illumination condition, whereas dataset 2 consist of major change in illumination, skin color and variation in sign postures.

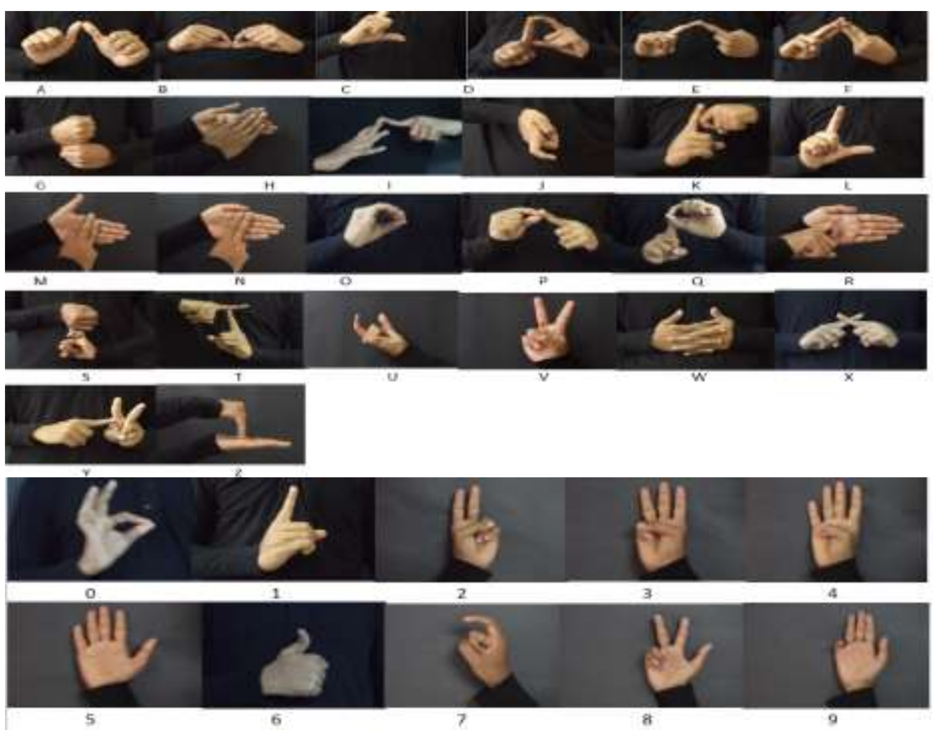


Figure 1. ISL Alphabet and number [32]

III. METHODOLOGY- HAND GESTURE RECOGNITION

Hand gesture recognition includes four major modules viz. image acquisition, pre-processing and hand segmentation, feature extraction, and classification. In this work, bare hand segmentation is explored. Efficient segmentation results are achieved for single as well as two handed complex static hand gestures for alphabets and numbers. Analysis of individual descriptor viz. BM, 7Hu moments and FD, chain code and shape matrix has been done and compared with new multi-feature fusion descriptor (BM+7Hu+FD). Nearest mean classifier (NMC) and k-nearest neighborhood (k-NN) classifier using Euclidean distance are used for classification. Naive Bayes classifier has also been experimented with other contour and region based descriptor viz. chain code and shape matrix which belongs to spatial domain.

a. Pre-processing and Hand Segmentation using decision rule based approach

Image acquisition is carried out with the help of web camera of 2M pixel. Captured images are resized into 320X240 resolutions for fast processing. In this approach, supervised learning approach is used with non-parametric model. The specialty of the algorithm is person independent hand segmentation. Due to the vast diversity of individual skin tone, feature extraction for skin color is done by selecting skin color dynamically with the help of mouse click event. HSV colour space is used for skin colour detection as it gave better result than L*a*b, YCbCr for skin colour segmentation under complex background [25, 26, 27, 28]. Selected feature color values (HSV) get stored into feature array and used as a feature for segmentation purpose. If there is extreme variation of skin tone, then and only then features need to be extracted with mouse click event and previous feature array gets refresh. The algorithm is adapted skin values of new user as a feature for segmentation. So, algorithm is working for almost all variety of skin colors. These feature values are used for converting H, S, V plane into binary image for defined decision rule with experimented threshold value.

Decision rules used in the algorithm 1 are included in InRange function which returns true or false value. These decision rules are:

Boolean InRange (feature[ ], I)

{

$if\ (|F_h - h| \leq 40\ or \left|(F_h - (h - 255)\right) \leq 40\ ) return\ true\ else\ false$

$if\ (\ s > F_s)\ return\ true\ else\ false$

$if\ (\ v > F_v)\ return\ true\ else\ false$

}

Here h, s, and v are the image pixel value and F_h, F_s and F_v are the feature values. Feature array gets refreshed by selecting new skin tone color in case of new user.

This is a very efficient algorithm, which works for person independent skin color with different illumination condition just by selecting skin color values as a feature.

The algorithm 1 for Hand segmentation using HSV color space by decision rules is discussed below

Algorithm 1: Person Independent Hand Segmentation using HSV Color Space by Decision Rules

Procedure SegHand (Image I)

```
Color=0; Image I,S,O,D,Seg;
F[3]; \\  H,S,V dynamically selected color values from RGB image with
        mouse click
I= RGB_Image;
I=CV_GUASSIAN(I);
O=HSV(I);
Calculate  height(h) and width(w) of the image
for i=0 to h
    for j=0 to w
         Boolean InRange (F, O)   \\Compare these value with
                                      Feature H,S and V
              if( InRange (current value && F)) \\ apply decision rule
               Sij= white;
               else
               Sij= black;
               end if
          end for
      end for
      D=dilation (S);  \\ Perform dilation on segmented image
      Seg=Median_Filter(D);  \\ Perform median filtering on dilated image
      Display(Seg);   \\ Display segmented image after preprocessing
```

EndProcedure

Figure 2 shows the detail steps for hand segmentation. Figure 2 shows (a)input RGB image to HSV conversion and (b)the selection of color feature values by mouse click and segmentation result for sign A. Figure 3 and 4 shows the segmentation result for ISL manual dataset of proposed hand segmentation algorithm.

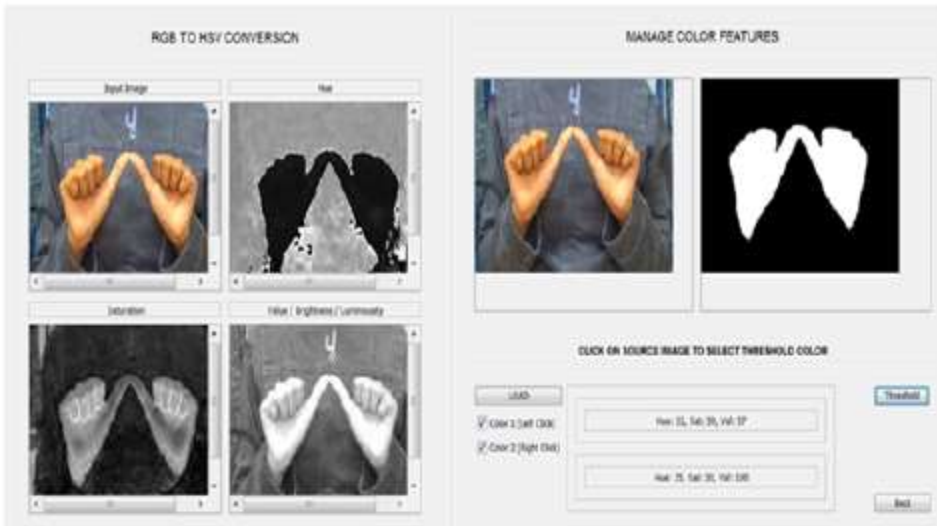


Figure 2. Segmentation result of sign A (a) RGB to HSV conversion (b) Feature selection

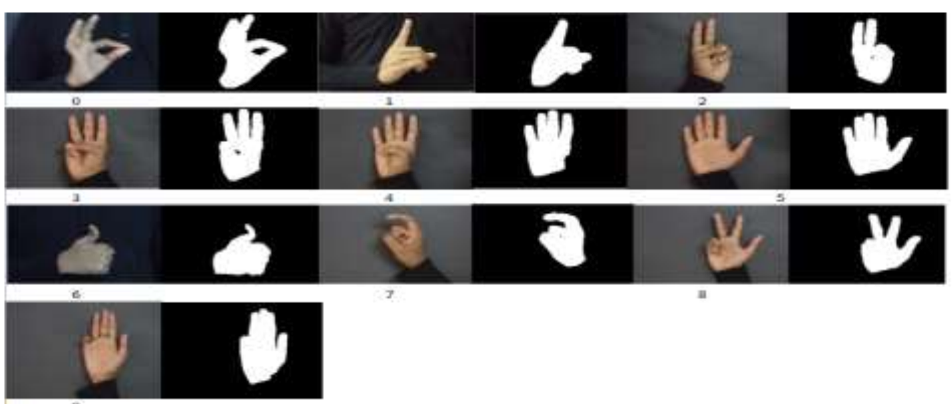


Figure 3. Segmentation result for digit 0-9

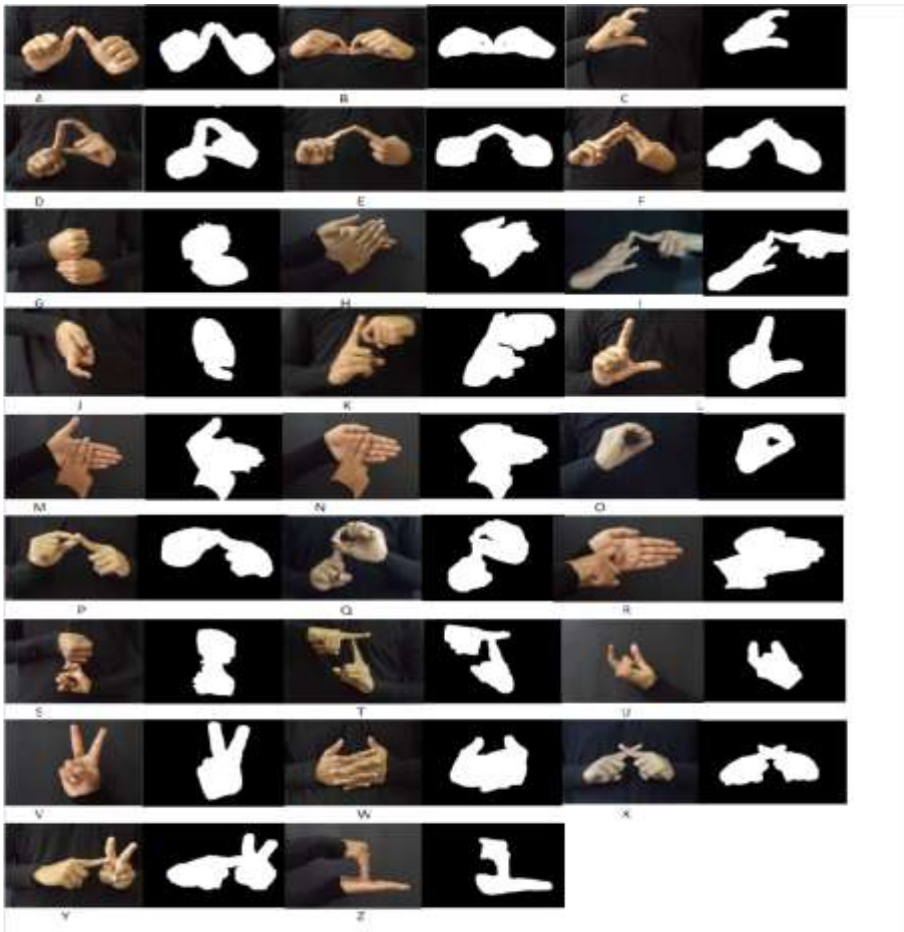


Figure 4. Segmentation result for alphabet A-Z

b. Hand Feature Extraction

ISL manual alphabets and numbers consist of different combinations of hand shape. Some signs are complex and ambiguous. Extraction of features of such complex and ambiguous signs for high recognition accuracy is a challenge. Shape is a visual feature of hand. Study of literature shows that, shape descriptors are classified into contour and region based descriptors. Contour based descriptors possess only boundary information neglecting inner shape region whereas, region based descriptors contain inner region information [11]. Here, BM , chain code and FD of contour based and 7 Hu moments and shape matrix of region based methods are identified which are invariant to translation, rotation and scale. Figure 5 shows these methods further classified into transform, statistical and spatial domain. The main aim here is to develop new multi-feature descriptor, which will consider the strength of each descriptor from different domains and will yield high recognition rate. BM and 7 Hu moments are the statistical, contour and region based descriptor respectively. They possess the moment based statistical information of the hand shape. Fourier descriptor is the contour based descriptor that belongs to transform domain which is advantageous for fast computation, robustness and high information preserving over other techniques. Achieving scale, translation and rotation invariance is easier using Fourier descriptor than other spatial domain descriptors. Analysis has been done with individual descriptor with new multi-feature descriptor using three classifiers.

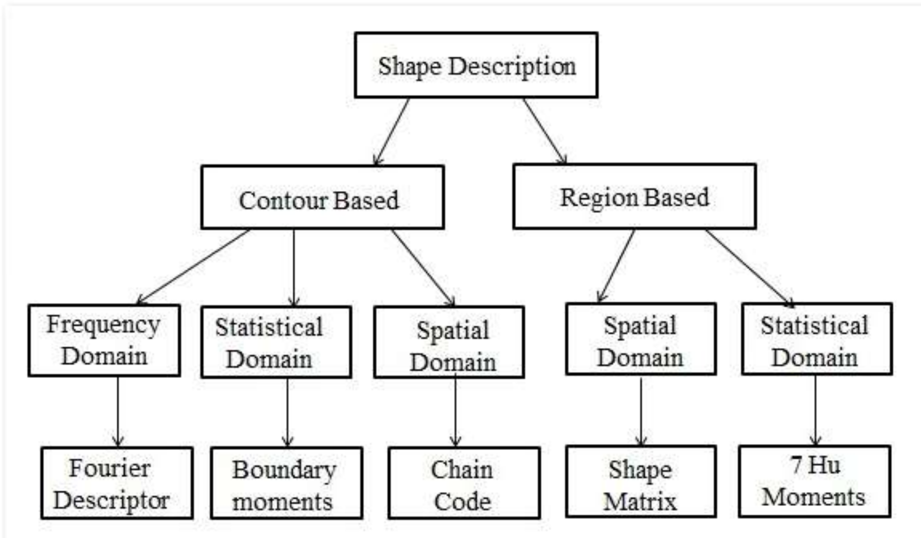


Figure 5. Classification of shape description used for empirical analysis.

b.i New Multi-feature shape descriptor using fusion of boundary moments, 7Hu moments and Fourier descriptor

Recognition rate, speed and robustness are the important aspects for any sign language recognition. The aim to design multi-feature shape descriptor for HGR, is to increase recognition rate by fusing contour (BM and FD) and region (7 Hu) based descriptors.

Boundary moment [11]: It gives the statistical information of the shape. Calculations of features are done on 1-D representation of boundary using centroid distance shape signature. Equal angle sampling method is used to calculate 256 boundary points (z_1-z_{256}). r^{th} moments m_r and central moments μ_r are calculated for 256 boundary points using equation 1.

$$m_r = \frac{1}{256}\sum_{i=1}^{256}[z(i)]^r \quad , \quad \mu_r = \frac{1}{256}\sum_{i=1}^{256}[z(i) - m_1]^r \qquad (1)$$

Boundary points are sensitive to noise so, from r_{th} moments m_r and central moment μ_r, less noise sensitive shape descriptors are calculated which is given below in equation 2. [11]

$$F_1 = \frac{(\mu_2)^{1/2}}{m_1}, \quad F_2 = \frac{\mu_3}{(\mu_2)^{3/2}}, \quad F_3 = \frac{\mu_4}{(\mu_2)^2} \qquad (2)$$

Table 1 shows the boundary moments feature values (F_1 - F_3) for sign alphabet A with different transformation. It is observed that the values are varying in certain range for rigid transformation for image 4 and 5 in table 1 but, for ideal images, it is preserving the same information.

Table 1: Boundary moments for sign 'A' of different transformation

Sr.No.	Image	F1	F2	F3
1		0.7218	0.1242	1.6637
2		0.7275	0.1134	1.6505
3		0.7336	0.1344	1.6366
4		0.6994	0.2569	1.6264
5		0.6339	0.24805	1.9285

7 Hu Moments : 7 Hu is a region based statistical shape description technique. The first six descriptors are invariant under change of size, translation and rotation. The seventh descriptor ensures skew invariant which enables to distinguish between mirror image [13]. In this work, most of the static hand gestures are two handed so, all seven descriptors are

considered for recognition purpose. Table 2 shows Hu moments preserve almost same information in different transformation.

Table 2. 7 Hu moments for sign ‘A’ of different rotation angle

Sign ‘A’ image	**M_1**	**M_2**	**M_3**	**M_4**	**M_5**	**M_6**	**M_7**
	-0.4295	-1.0888	-2.4278	-3.3365	6.2873	3.9314	6.5026
	-0.4292	-1.0882	-2.4375	-3.3543	6.2878	3.9104	6.6496
	-0.4292	-1.0884	-2.5198	-3.4775	6.5767	4.0925	6.6918
	-0.42066	-1.1803	-2.1755	-3.7840	6.7908	5.8045	7.2293
	-0.4124	-1.2002	-2.7675	-2.8007	6.7124	3.4479	5.58605

Fourier descriptor: It is a transform based descriptor [14] which is calculated on 1-D centroid distance shape signature and which is translation invariance. Normalization and sampling are done using equal angle sampling and 256 boundary points are calculated. Discrete Fourier transform (DFT) is applied on s(t), t = 0, 1, ...256 which gives Fourier descriptors and represented the hand shape in frequency domain. Here, lower frequency descriptors are neglected and considered fd_{10}- fd_{256} descriptors for classification. Rotation invariance is achieved by ignoring phase coefficient and retaining magnitude efficient of FD. Scale invariance is achieved by dividing magnitude values of the FDs by DC component. The properties of Fourier descriptors are robustness, ability to capture perceptual property of shape boundary and less sensitivity to noise which makes 255
FD is efficient feature extraction in the field of object recognition. Here, due to complex and ambiguous nature of ISL sign 246 descriptor are considered for recognition purpose

but theoretically 60-80 FDs are sufficient. Observation and empirical study show that, BM, 7 Hu and FD are the suitable descriptors for shape representation but do not give high recognition result individually.

Multi-feature Descriptor: New multi-feature descriptor is formed by fusing BM, Hu and FD. As all these feature descriptors have different numerical range and dimension So, fusion of three descriptors was a challenge. Here, similarity difference (error values) of each feature descriptor BM, Hu and FD with unknown sample has been calculated with all training samples. Two weight factors are devised and used in Fusion descriptor (T). T is a total error value calculated by adding $\mathrm{b}'_{\mathrm{error}}, \mathrm{f}'_{\mathrm{error}}$ and $\mathrm{h}'_{\mathrm{error}}$ given in equation 3 which is a fusion descriptor.

$$(\mathrm{T}) = \mathrm{f}'_{\mathrm{error}} + (\mathrm{h}'_{\mathrm{error}} * \eta_1) + (\mathrm{b}'_{\mathrm{error}} * \eta_2) \;\; \forall\, 1..\mathrm{N} \text{ samples} \qquad (3)$$

$$\because \eta_1, \eta_2 \text{ are weight factors}$$

These three different descriptors have different numerical ranges. Here, η_1 and η_2 weight factors are identified and used to form fusion vector for classification purpose. Error (distance) is calculated for fusion descriptor (Fd+Hu+BM) which is given below for all training samples along with testing sample.

$\mathrm{h}'_{\mathrm{error}} = \sqrt{\sum_{i=0}^{i=7}(u_i - v_i)^2}$ // error difference between trained and test symbol

$\mathrm{f}'_{\mathrm{error}} = \sqrt{\sum_{i=10}^{i=256}(u_i - v_i)^2}$ // error difference between trained and test symbol

$\mathrm{b}'_{\mathrm{error}} = \sqrt{\sum_{i=0}^{i=3}(u_i - v_i)^2}$ // error difference between trained and test symbol

So, Fusion descriptor (T) is a new multi-feature descriptor having 256 dimension and possesses both strengths of contour and region based descriptors. Figure 6 shows sample result of recognition signs 'K' and 'L'.

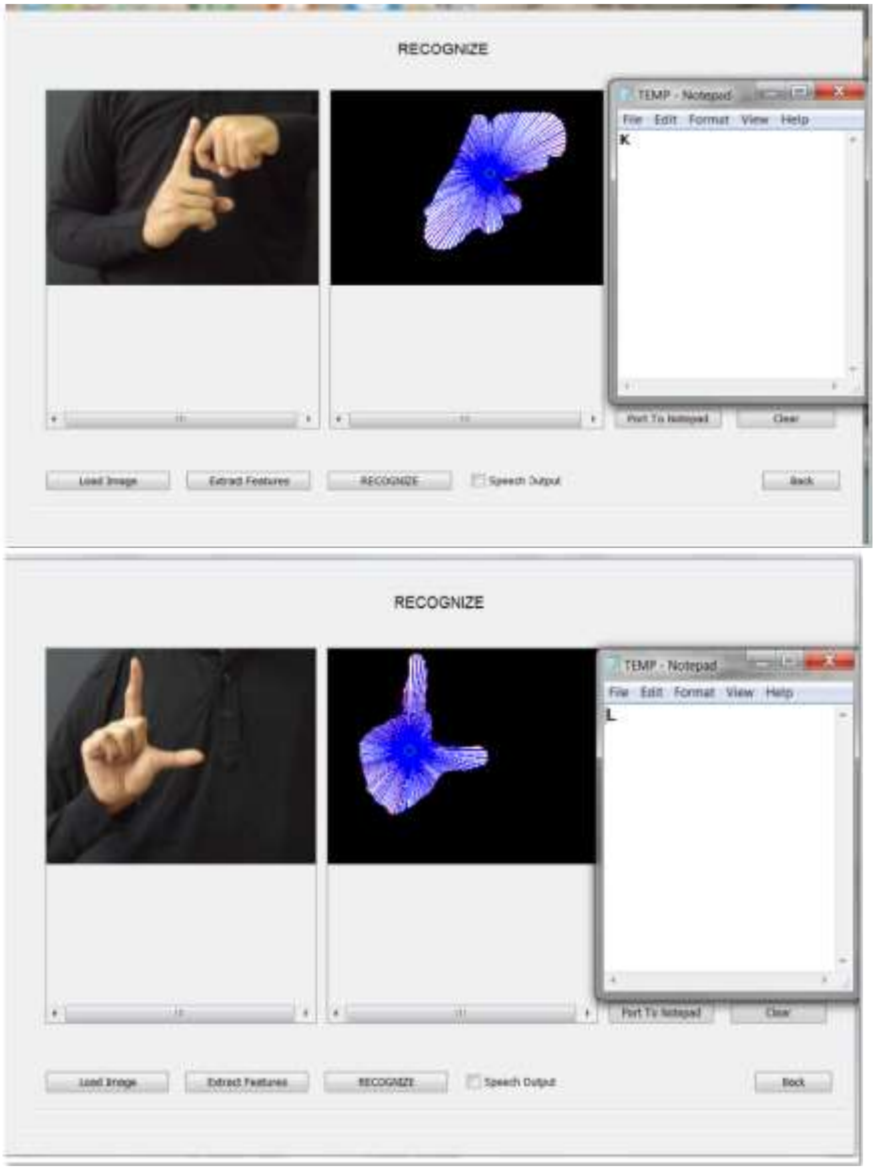


Figure 6. Recognition result for few sample signs such as 'K','L'.

b.ii HGR using shape matrix and chain code descriptor

Shape matrix [11]: It is a region based descriptor which is calculated in spatial domain. It is a descriptor of MXN size which represents the region shape. Polar model of shape matrix is used. Centre of gravity of shape and maximum radius of shape region are calculated. From centre(Rc), maximum radius (Rm) of hand shape is calculated and divided into ten equally spaced radius (Rm,Rm-1,…R0) and ten circles are drawn and considered as rows of matrix M. Fifty equally spaced arcs are drawn on these circles and considered as columns of matrix N. The value of M and N are experiment based values for trained sample images.

$M_{ij} = 1$ If intersection of circle and arc fall into shape region else it is defined as 0

Principal component analysis has been applied to reduce the dimension of matrix and used for recognition. Figure 7 shows shape matrix representation for sign A. The red point (intersection of circle and arc) in the given image makes as 1 value in the corresponding row and column in the matrix.

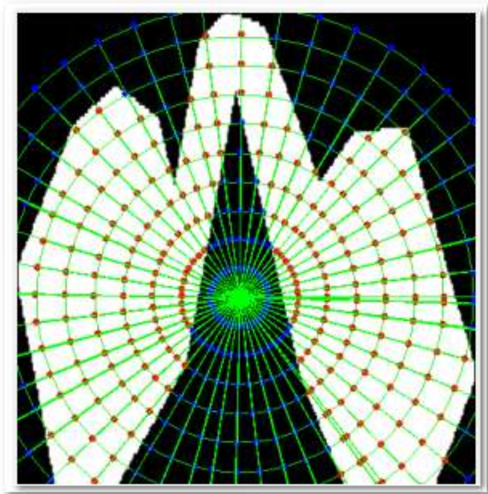


Figure 7. Shape matrix representation for sign 'A'

Chain code [11]: It is a contour based shape description technique that falls under spatial domain. Free man eight directional differential chain code is used. It represents an object boundary by a connected sequence of straight line segments of specified length and direction. Shape number as a descriptor is formed after normalization and used as a descriptor for matching purpose. $\vec{c}$ is sequence of integer between 0-7 following the eight directions. Dimension reduction is done using PCA on $\vec{c} = \{ c_1 \ldots . c_n\}$ and used as descriptor for recognition purpose.

c. Hand Gesture classification and results

In present work, three classifiers such as, nearest mean classifier (NMC), non-parametric nearest neighbourhood (NN) and parametric Naive Bayes are used. These classifiers are simple in nature and suitable for multi-dimensional feature vector. Performance analysis of each algorithm is done with these classifiers using hold-out validation technique wherein 50% sample data are used for training and 50% sample data are used for testing purpose. Shape matrix and chain code performance are checked with nonparametric NN and parametric Naive Bayes classifier. Analysis of fusion descriptor has been done with other descriptor such as chain code and shape matrix which belong to spatial domain. Real time HGR performance is checked with NMC for number sign.

c. i Nearest Mean Classifier

ISL static hand gesture was recognized using nearest mean classifier [18, 19], and output is displayed in the text (displayed to notepad) as well as in voice. Here, mean prototype is used to calculate typical sample to represent a particular class from a large number of samples. Total 36 mean values corresponding to each class are calculated from training data samples. Theses mean values represent each class and used to calculate minimum distance with the unknown gesture and minimum distance class is declared as recognized class. Euclidean distance measure [18] is used for distance calculation with all proposed feature descriptors which is a direct measure of the similarity between two multi-dimensional vectors. Figure 8 shows that recognition accuracy of fusion (BM+Hu+FD) vector, is higher than that of individual descriptors. It is observed that, combining contour and region based descriptors with the combination of statistical and frequency domain descriptor, the recognition rate is found to increase. Performance of real-time HGR is checked for signs 0-9 for all descriptors and gave good true positive rate (TPR) of 100 % for fusion descriptor. Fusion descriptor considers all strong features of each descriptor and gives high recognition rate. It is observed that, due to the large number of FDs, FD gives a higher recognition rate than the other two descriptors BM and 7Hu. In fusion descriptor also, FD dominates other two descriptors. But in some cases where FD fails, BM and 7Hu contribute to recognize the correct ISL symbol in fusion such cases is given in Table 3. In few cases where FD and 7Hu fail, BM contributes to recognize the correct ISL symbol in fusion and such cases are given in Table 4. In some cases where all the three descriptors individually fail, the fusion gives the correct recognition of ISL symbol such few sample cases are given in table 5. Observations given in table 3, 4 and 5 are the results of NMC and they showed that fusion descriptor perform better in giving high recognition accuracy than individual descriptor which is essential in any object recognition application. Table 6 presents performance analysis for A-Z and 0-9 using NMC based on confusion matrix. Accuracy (ACC), true positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR) are the parameters used for performance analysis [35].

Table 3. Few cases where, contribution of BM and 7Hu in fusion

Test sign	**7Hu**	**FD**	**BM**	**Fusion (7Hu+FD+BM)**
S	S	K	Z	S
J	J	9	C	J
P	P	X	P	P
T	T	E	T	T
X	X	D	X	X
5	5	9	R	5

Table 4. Few cases where, contribution of BM in fusion

Test sign	**7Hu**	**FD**	**BM**	**Fusion (7Hu+FD+BM)**
Y	X	S	Y	Y
3	5	J	3	3
W	O	C	W	W
K	7	2	K	K

Table 5. Few cases where fusion work when 7Hu, BM and FD fail

Test sign	**7Hu**	**FD**	**BM**	**Fusion (7Hu+FD+BM)**
A	B	F	Y	A
S	V	K	F	S
R	6	J	8	R

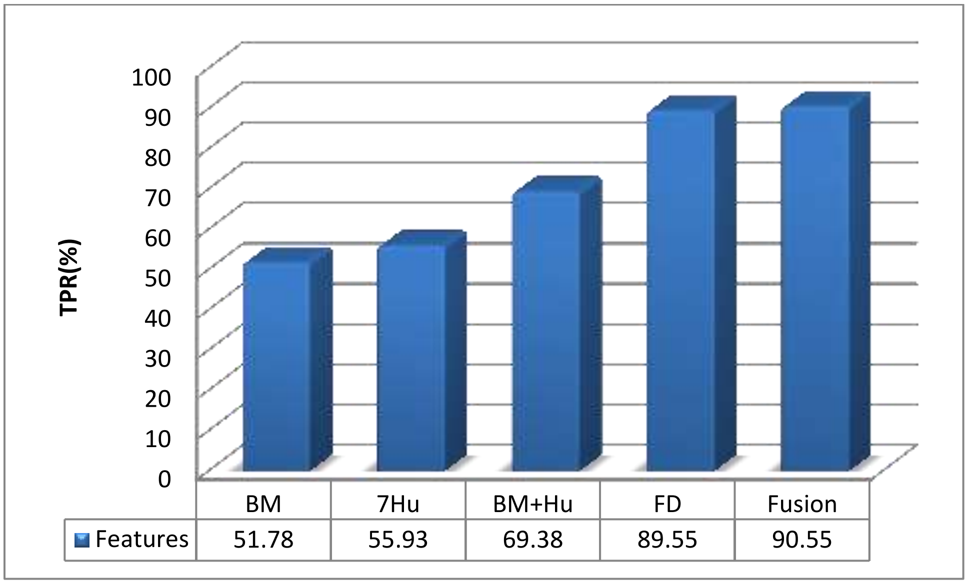


Figure 8. Recognition accuracy by NMC

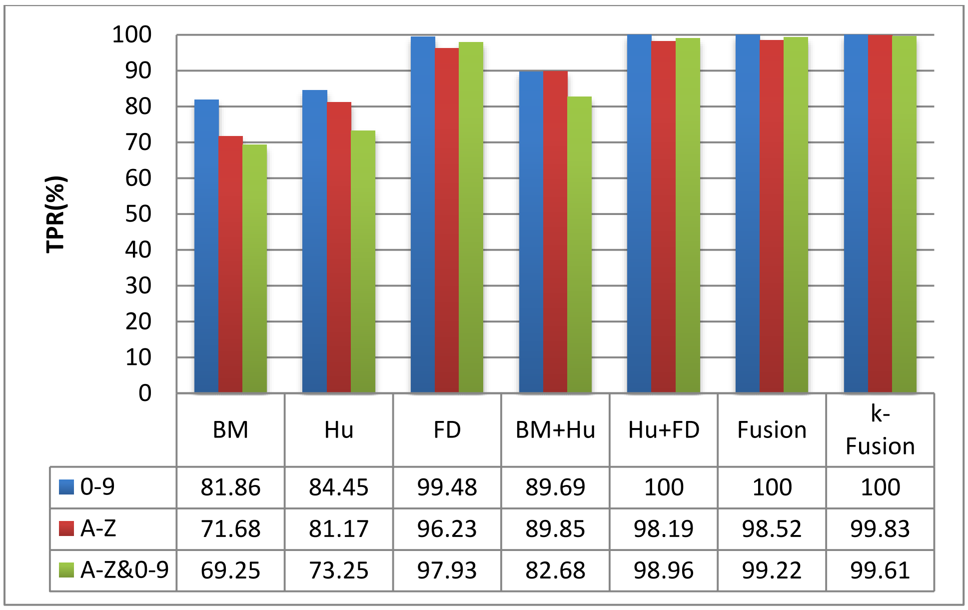
Table 6. Performance analysis for A-Z and 0-9 (36 classes) with NMC

Performance Parameter	**BM**	**7Hu**	**FD**	**BM+Hu**	**Fusion**
Accuracy	97.32	97.55	99.41	98.69	99.47
True positive rate	51.78	55.93	89.55	69.38	90.55
True negative rate	98.62	98.74	99.70	99.12	99.73
False positive rate	1.37	1.25	0.29	0.87	0.26
False negative rate	48.21	44.06	10.44	30.61	9.4

c.ii Nearest Neighbourhood Classifier (NN)

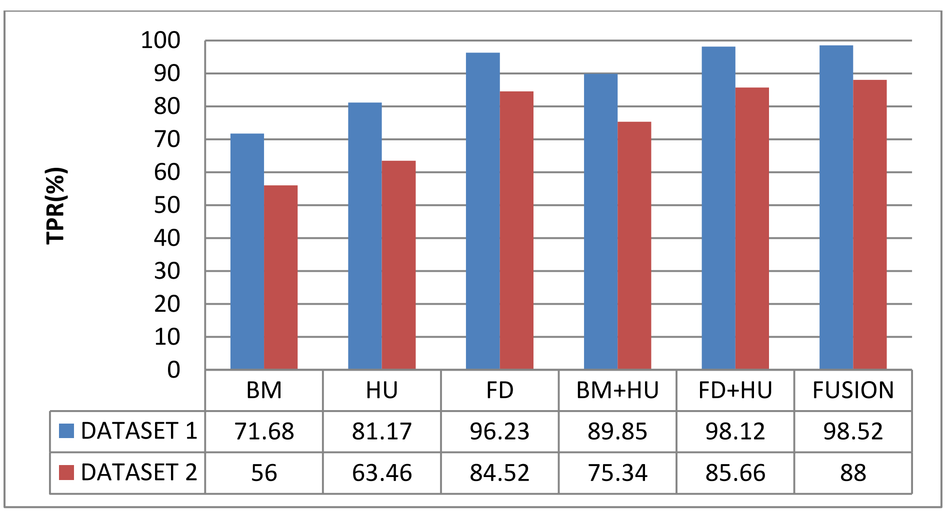
NMC classifier shows that, fusion descriptor gives better recognition rate than other single descriptors. So, to compare the performance with other classifiers, analysis has been carried out with other non-parametric classifiers such as a nearest neighbourhood with Euclidean distance measures. The results are compared with k-NN considering experimented k value as a 5. Analysis of NN and 5-NN is presented here. The alphabet from training database, which corresponds with the minimum distance with the unknown gesture, is the recognized gesture. Figure 9 shows detailed result analysis of all methods for three manual sets and k-fusion shows highest recognition rate of 100% for number, 99.83%

for alphabets and 99.61% for alphabet and numbers. In figure 10 comparison analysis of two different dataset has been presented with NN classifier. It shows that in any dataset fusion descriptor performance is better than traditional single method. It is observed that though the sample images are not ideal and distorted by transformation, variable distance and variation in illumination, performance of fusion descriptor is better. In all cases, fusion of contour and region based descriptor gives increased recognition rate as it considers strength of each feature descriptor. Table 7 shows performance analysis based on confusion matrix using NN and k-NN classifier.



	BM	Hu	FD	BM+Hu	Hu+FD	Fusion	k-Fusion
0-9	81.86	84.45	99.48	89.69	100	100	100
A-Z	71.68	81.17	96.23	89.85	98.19	98.52	99.83
A-Z&0-9	69.25	73.25	97.93	82.68	98.96	99.22	99.61

Figure 9. Recognition accuracy by NN and k-NN method



	BM	HU	FD	BM+HU	FD+HU	FUSION
DATASET 1	71.68	81.17	96.23	89.85	98.12	98.52
DATASET 2	56	63.46	84.52	75.34	85.66	88

Figure 10. Comparative analysis with dataset 1 and dataset 2

Table 7. Performance analysis for NN, k-NN classifier

Performance analysis for A-Z and 0-9 (36 classes)							
Performance Parameter	**BM**	**Hu**	**FD**	**BM+Hu**	**FD+Hu**	**Fusion**	**k-Fusion**
Accuracy	98.29	98.51	99.88	99.03	99.94	99.95	99.97
True positive rate	69.25	73.25	97.93	82.68	98.96	99.22	99.61
True negative rate	99.12	99.23	99.94	99.5	99.97	99.97	99.98
False positive rate	0.87	0.76	0.05	0.49	0.02	0.02	0.01
False negative rate	30.74	26.74	2.06	17.31	1.03	0.7	0.38

c.iii. Performance Analysis with k-NN and Naive Bayes Classifier for chain code and shape matrix

Naive Bayes classifier is one of the classifiers used for the classification. Figure 11 shows the performance analysis of k-NN and Naive Bayes on chain code and shape matrix. It is observed that TPR for Naive Bayes on shape matrix is better than chain code. Fusion descriptor gave good and steady performance amongst all the features with all classifier.

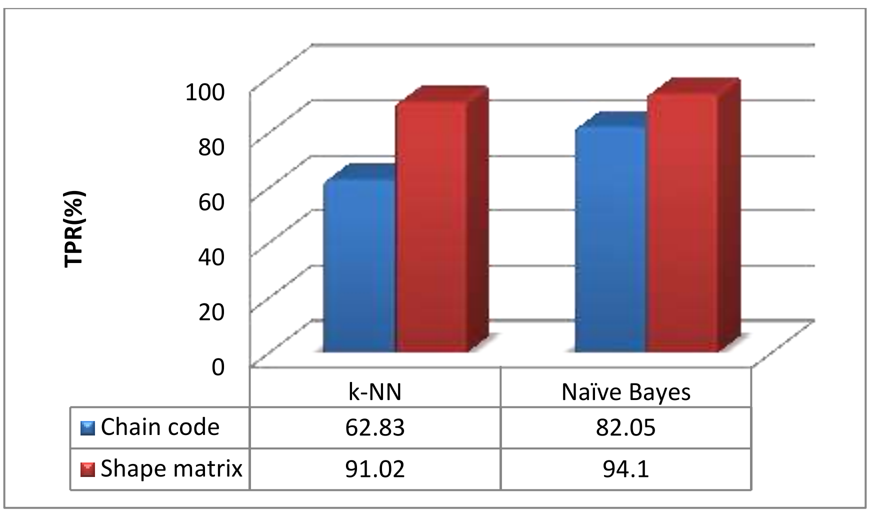


Figure 11. Performance analysis of chain code and shape matrix

IV. COMPARATIVE ANALYSIS WITH EXISTING ISL WORK

The research work is compared with related ISL work. There is a lot of research being done on various foreign sign languages. Comparing the proposed work with other SL is not feasible and appropriate due to difference in morphology, phonology and grammar of each SL like spoken language. Comparison of proposed work also cannot be directly done with other ISL work, due to the non-availability of standard image/video data set. Currently all researchers are working on self-created own image/video data set. Result may vary with camera model, resolution and image size. Some of them are using regional sign languages such as Tamil and Marathi where signs are language dependent. Therefore it remains subjective analysis peculiar to individual researcher. However, the comparisons with other researchers on standard ISL are mentioned in table 8.

Table 8. Comparison Analysis of proposed with exiting ISL Static HGR work

Method	**Dataset**	**Accuracy**	**Author-year**
Kohenon self-organizing map	24 alphabets	96.25%	[*Singha and Das* 2013] [9]
B-spline approximation	ISL alphabets and numerals	Up to 90%	[*Geetha and Manjusha* 2012] [8]
Kohenon self-organizing map	24 alphabets	Up to 80%	[*Tiwari and Kumar* 2012] [10]
Histogram of Edge Frequency	26 Alphabets	98.10%	[*Lilha and shivmurthy 2013*] [33]
SIFT	9 alphabets	95%	[S. Goyal et al., 2013][34]
New fusion descriptor	**36 alphabets**	**99.61%**	**Proposed approach**

IV. CONCLUSION AND FUTURE WORK

New multi-feature hand gesture recognition algorithm with person independent hand segmentation is explored. Recognition of static manual signs is one of the steps of ISL interpretation. Key focus of this paper is to give multi-feature hand gesture recognition algorithm, which can be applied to any HCI application or object recognition problem in general. Fusion descriptor is designed by combining contour and region based descriptors. Further experimentation has been carried out with chain code and shape matrix which falls under spatial domain feature and tested with both non-parametric NN and parametric Naive

Bayes classifier. There are certain limitations of spatial feature descriptor over other descriptors such as noise sensitivity, difficulty in achieving invariance, and high computational complexity. So, fusion descriptor was formed with statistical and transforms based shape descriptors and analysis has been done with spatial feature descriptors. Result revealed that, fusion descriptor gave high recognition rate amongst all because shape information of contour and region based descriptor get added in fusion descriptor moreover that showed statistical and transform based combination preserves shape information of both domains. So, this attempt can give a perspective to other researchers that combining contour and region based descriptors outperform over individual descriptors. Future work is in progress for ISL dynamic sign words and sentence interpretation.

REFERENCES

[1] U. Zeshan, M. Vasistha and Sethna, "Implementation of Indian Sign Language in Educational Setting," Asia pacific Disability Rehabilitation Journal ,vol. 16, no. 1, 2001, pp. 16-39.

[2] C. Vogler, D. Metaxas, "A Framework for Recognizing the Simultaneous Aspect of American Sign Language", Computer Vision and Image Understanding, 81, 1998, pp. 358-384.

[3] T. Starner, J. Weaver, A. Pentland, "Real-time American Sign Language recognition using desk and wearable computer based video," *IEEE Transaction* on Pattern Analysis and Machine Intelligent, 20, 2010, pp. 1371-1375.

[4] M. Flasinski, S. Myslinski , "On the Use of Graph parsing for recognition Of isolated hand posture of Polish Sign Language", Pattern Recognition, Elsevier, 43, 2004, pp. 2249-2264.

[5] W. Gao, G. Fang, D. Zhao, "A Chinese sign language recognition system based on SOFM/SRN/HMM", Pattern Recognition. Elsevier, vol. 37, 2013, pp. 2389-2402.

[6] J. Pansare et al., "Real-Time Static Devnagri sign Language Translation using Histogram," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),vol. 2, no. 4, 2011, pp. 1455-1459.

[7] P. Subha, G. Balakrishnan , "Recognition of Tamil Sign Language Alphabet using Image Processing to aid Deaf-Dumb People", International Conference on Communication Technology and System Design, vol. 1, 2012, pp. 861-868.

[8] M. Geetha, U. Manjusha , "A Vision based Recognition of Indian Sign Language Alphabets and Numerals using B-Spline Approximation," International Journal on Computer Science and Engineering, vol. 4, no. 3, 2013, pp. 406-415.

[9] J. Singha, K. Das, "Recognition of Indian Sign Language in Live Video", International Journal of Computer Applications, vol. 70, no. 19, 2012, pp. 17-22.

[10] D. Tewari, S. Kumar, "A Visual Recognition of Static Hand Gesture in Indian Sign Language based on Kohonen Self organizing Map Algorithm", International Journal of Engineering and Advanced Technology , vol. 2, no. 2, 2004, pp. 165-170.
[11] D. Zhang, C. Lu, "The Journal of The Pattern Recognition Society, Elsevier, vol. 37, no.1,2014,1-19.
[12] H. Cooper, B. Holt, R. Browden, Sign language recognition. http://epubs.surrey.ac.uk/531441/1/SLR-LAP.pdf, accessed in sept-2014.
[13] Z. Huang, J. Leng, "Analysis of Hu Moment Invariants on Image Scaling and Rotation", Second International Conference on Computer Engineering and Technology ,vol. 1, 2007, pp.476-480.
[14] S. Conseil, S. Bourennane, L. Martin, "Comparison Of Fourier Descriptors And Hu Moments for Hand Posture Recognition", European Signal Processing Conference, (EUSIPCO),vol.1, 2009, pp. 1-6.
[15] J. Ravikiran et al., "Finger detection for sign language detection", Proceedings of the International Multi-Conference of Engineers and Computer Scientists , Hong Kong ,vol. 1, 2011, pp. 1-5.
[16] J. Rekha, J. Bhattacharya and S. Majumder , "Hand Gesture Recognition for Sign Language: A New Hybrid Approach", 15th International Conference on Image Processing, Computer Vision & Pattern Recognition IPCV'11, CSREA Press, Las Vegas, Nevada, USA, vol.1, 2000, pp. 30-35.
[17] A. Jain , R. Duin, "Statistical Pattern Recognition: A Review", IEEE Transactions On Pattern Analysis And Machine Intelligence, vol. 22, no. 1, 2009, pp. 4-37.
[18] J. Li, B. Lu, "An adaptive image Euclidean Distance", Elsevier, Vol. 42, no. 3, 2009, pp.349-357.
[19] S. Cha, "Comprehensive Survey on Distance/Similarity Measures between Probability Density Function", International Journal of Mathematical Models and Methods in Applied Science, vol. 1, no. 4, 2007, pp. 300-307.
[20] K. Wong,R. Cipollas, "Continuous gesture recognition using a sparse Bayesian classifier", International conference on pattern recognition, 2009, pp. 1084-87.
[21] T. Maung, "Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks," Proceedings of World Academy of Science, Engineering and Technology, vol. 3, 2007, pp. 470-478.
[22] Q. Munib, Moussa, Bayean, Hiba, " American Sign Language(ASL) recognition based on Hough transform and neural network", Expert Systems with Applications, Elsevier,2002, pp. 24-37.
[23] A. Corradini, "Real-Time Gesture Recognition by means of Hybrid Recognizers," Gesture Workshop, LNAI 2298, Springer-Verlag Berlin Heidelberg, 2014, pp. 34-47.
[24] ISL Dictionary: http://indiansignlanguage.org/isl-dictionary, accessed May 2014.
[25] S. Phung, A. Bouzerdoum, D. Chai, "Skin Segmentation Using Colour Pixel Classification: Analysis and Comparison", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 27, no. 10, 2005, pp. 148-154.
[26] A. Elgammal, C. Muang, D. Hu, D. Chai , "Skin Detection a Short Tutorial" , Encyclopedia of Biometrics, Springer-Verlag Berlin Heidelberg, 2002.
[27] M. Jones, J. Rehg, "Statistical Color Models with Application to Skin Detection", International Journal of Computer Vision, vol. 46, no. 1, 2002, pp. 81-96.
[28] A. S. Ghotkar , G. K. Kharate, "Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction", International

Journal of Human Computer Interaction(IJHCI),Computer Science Journal, Malaysia, vol. 3, no. 1, 2012, pp. 15-25.
[29] Archana S. Ghotkar, Gajanan K. Kharate, "Vision based Real time Hand Gesture Recognition Techniques for HCI ", *International Journal of computer Application*, Volume. 70, No. 16, Foundation of Computer Science, New York, USA, 2013, pp.1-6.
[30] T. Dasgupta, S. Shulka, S. Kumar, S. Diwakar, A. Basu, "A Multilingual Multimedia Indian Sign Language Dictionary Tool," Proceedings 6th Workshop on Asian Language Resources, vol.1, 2008, pp. 57-64.
[31] C. Harshith et al., "Survey on Various Gesture Recognition Techniques for Interfacing Machine Based on Ambient Intelligence", International Journal of Computer Science and Engineering Survey, 2011, vol. 1, no. 2, pp. 31-33.
[32] Archana S. Ghotkar and Gajanan K. Kharate, "Study of Hand Gesture Recognition for Indian Sign Language", International Journal of Smart Sensing and Intelligent Systems, Vol. 7, No.1, 2014, pp. 96-115.
[33] H. Lilha, D. Shivmurthy, "Evaluation of Features for Automated Transcription of Dual-Handed Sign Language Alphabet", Proceedings of International Conference on Image Information Processing, 2013, pp. 1-5.
[34] S. Goyal, I. Sharma, S. Sharma, "Sign Language Recognition System for Deaf and Dumb People", International Journal on Engineering and Research Technology, vol. 2, no. 4, 2013, pp. 382-387.
[35] Tom Fawcett, "An introduction to ROC analysis", Pattern recognition letters, Elsevier,vol.27,2006,pp.861-874.