IJCSE International Journal of Computer Sciences and Engineering **Open Access**

Research Paper **Volume-4, Issue-9** **E-ISSN: 2347-2693**

Gradient Based Key Frame Extraction for Continuous Indian Sign Language Gesture Recognition and Sentence Formation in Kannada Language: A Comparative Study of Classifiers

Ramesh M. Kagalkar[1*] and S.V Gumaste[2]

[1*]Research scholar, Visvesvaraya Technological University Regional Resource Centre (VTU-RRC), Visvesvaraya Technological University, Belgaum, Karnataka, India.

[2]Professor & Head, Computer Engineering, R. H. Sapat College of Engineering, Nashik, Maharashtra, India

Available online at: www.ijcseonline.org

Received: 22/Aug/2016 Revised: 02/Sept/2016 Accepted: 20/Sept/2016 Published: 30/Sep/2016

***Abstract*— Human hands are delicate instruments. Hand gestures and finger gestures are excellent ways of emphasizing what we say, but on the other hand they can also reveal our true intentions. In this paper introduced a continuous Indian sign language recognition system, wherever each the hands are used for playacting any gesture. Recognizing a sign language gestures from continuous gestures could be a terribly difficult analysis issue. This paper solve the problem using gradient based key frame extraction technique. These key frames are useful for splitting continuous language gestures into sequence of signs further as for removing uninformative frames. After splitting of gestures every sign has been treated as associate degree isolated gesture. Then features of pre-processed gestures are extracted using orientation histogram (OH) with principal component analysis (PCA) is applied for reducing dimension of features obtained after OH. Experiments are performed on our own continuous ISL dataset which is created using EOS camera in PG Research Laboratory (SPPU, Pune). Probes are tested exploitation varied forms of classifiers like, Manhattan distance, Correlation, Manhattan distance, City block distance, Euclidian distance etc. Comparative analysis of our projected theme is performed with varied forms of distance classifiers. From this analysis we tend to found that the results obtained from Correlation and Euclidian distance offers higher accuracy then alternative classifiers.**

Keywords-* *Gesture Recognition, Orientation histogram (OH); Correlation; Indian sign language (ISL); Principal component analysis (PCA);

I. INTRODUCTION

Humans know each other by conveying their ideas, thoughts, and experiences to the people around them. There are numerous ways to achieve this and the best one among the rest is the gift of Speech. Through speech everyone can very convincingly transfer their thoughts and understand each other. It will be injustice if we ignore those who are deprived of this invaluable gift. The only means of communication available to the vocally disabled is the use of sign language. Using sign language they are limited to their own world. This limitation prevents them from interacting with the outer world to share their feelings, creative ideas and potentials. Very few people who are not themselves deaf ever learn to sign language. These limitation increases the isolation of deaf and dumb people from the common society. Technology is one way to remove this hindrance and benefit these people. Several researchers have explored these possibilities and have successfully an achieved finger spelling recognition with high levels of accuracy. But progress in the recognition of sign language, as a whole has various limitations in today's applications.

The problem of automated sign language recognition can be put across as, given a video of a sign language sentence, can We identify the signs in the sentence and reconstruct the sentence? The solution to the problem of sign language recognition has many practical implications. Firstly, advances in automated sign language recognition are necessary to improve the quality of life of deaf persons by facilitating their interaction with hearing populace in public situations. For instance, the use of innovative computer technologies can provide a solution to the dilemma a security screener faces in attempting to communicate with deaf passengers during the course of daily business activities. Also, it can be helpful in other places like courtroom, conventions or even a grocery store. On the other note, human computer interaction (HCI) is gradually moving towards a modality where speech recognition will play a major role. While speech recognition has made rapid advances, gesture recognition is lagging behind. With this gradual shift to speech based Input output devices, there is a great danger that persons who rely solely on sign languages for communication will be deprived access to state-of-the-art technology unless there are significant advances in automated recognition of sign languages. Secondly, the problem of automated sign language recognition is also worthwhile from a scientific and technological point of interest, since advances in this problem would definitely impact the general problem of automated gesture recognition, which is at the core of designing next generation man-machine interface.

Most of the work in continuous sign language recognition has used Hidden Markov Models (HMM) for recognition and have not used in any way facial or non-manual information. Although facial information is considered to be fairly important, no prior work on continuous ASL recognition has made use of it. There has been some work on detecting `head shakes' and `nods' only, but these works do not show results on continuous sign language. Hence, there arises the need to develop some assistive tools which would provide an interface between a sign language and its spoken counterpart so that the gap between the two segments of the society can be minimized. Various past work were observed for this objective to solve the interaction problem as outline.

However the research community has paid less attention for the translation of sign language into text information in Kannada language. For the generating text description in Kannada language is challenging work for researcher. The Kannada is a Dravidian language spoken primarily in Karnataka State in South India, and has a literature that dates from the ninth century. It has a population of 35,346,000 speakers, and is spoken not only in Karnataka, but to some extent in the neighbouring states of Andhra Pradesh, Tamil Nadu, and Maharashtra. The literacy rate in Kannada is estimated to be about 60%. Kannada is written with its own script, which is similar to the script used also for Telugu. The Kannada script is also used for writing Tulu. Kannada is one of the most well-known Dravidian languages of India. It is as old as Tamil, the truest language of the Dravidian family. It is spoken predominantly in the state of Karnataka in India. Though a significant number of Kannada speaking people can also be found in USA, UAE, Singapore, Australia and UK, all of which have migrated from India. On an average, there are about 35 million Kannadigas i.e. the Kannada speaking people in the world, making it the27th most spoken language in the world. It is one of the official languages of India and the official and administrative language of the state of Karnataka in the country.

The script of Kannada language is syllabic. The language uses forty nine phonemic letters which are segregated into three groups- Swaragalu, the vowels, the Vyanjanagalu, the consonants and Yogavaahakagalu, the two characters which are not vowel or consonant. The Character set is very similar to that of other Indian languages. The script is fairly complex as like other complex scripts it has also been derived from Brahmi script. As far as Kannada Grammar is concerned, it is a highly inflected language with three genders- the masculine, feminine and the neutral, there are two numbers-singular and plural. Kannada is inflecting for gender, number and tense, among other things. A primary objective of this work is to recognize sign and generate equivalent text description in Kannada language. The significant contribution made in this paper helps in translation of sign language to text I Kannada language for single signer. The present work attempts to recognize the sign based on shape and features. The extracted features are input to ANN Classifier independently. The study leves scope for further research on translating different signer of sign language into text information in Kannada language with grammatically correct meaning. The literature on related woks is reviewed to know the state of the art and survey is organized.

Sign language continues to be the preferred method of communication among the deaf and the hearing-impaired. Advances in information technology have prompted the development of systems that can facilitate automatic translation between sign language and spoken language. More recently, systems translating between Arabic sign and spoken language have become popular. This paper reviews systems and methods for the automatic recognition of Arabic sign language. Additionally, this paper highlights the main challenges characterizing Arabic sign language as well as potential future research directions [1]. In this paper a Sign Language Recognition system has been proposed. The first step of this system is to create a database of Indian Sign Language. This is done by acquiring the videos from the signers while they are performing the hand gestures. Next step is Hand tracking and Segmentation. This is performed in order to extract features from a particular gesture. A three step algorithm has been used in the proposed system to get better quality hand tracking and segmentation. This algorithm works on motion tracking, edge detection and skin colour detection. The system is implemented successfully and results are presented in this paper. The results demonstrate working of motion tracking, edge detection and skin colour detection individually as well as their combined effect [2].

It is difficult for most of us to imagine, but many who are Deaf-mute rely on sign language as their primary means of communication. They, in essence, hear and talk through their hands. Sign languages are visual languages. They are natural languages which are used by many deaf mute people all over the world. In sign language the hands convey most of the information. Hence, vision -based automatic sign language recognition systems have to extract relevant hand features from real life image sequences to allow correct and stable gesture classification. In our proposed system, we intend to recognize some very basic elements of sign language and to translate them to text. Firstly, the video shall be captured frame by frame, the captured video will be processed and the appropriate image will be extracted, this retrieved image will be further processed using BLOB analysis and will be sent to the statistical database; here the captured image shall compared with the one saved in the database and the matched image will be used to determine the performed alphabet sign in the language. Here, we will be implementing only

American Sign Language Finger spellings, and we will construct words and sentences with them [3]. In the recent years many approaches have been made that uses computer vision algorithms to interpret sign language. This endeavour is yet another approach to accomplish interpretation of human hand gestures. The first step of this work is background subtraction which achieved by the Euclidean distance threshold method. Thinning algorithm is then applied to obtain a thinned image of the human hand for further analysis. The different feature points which include terminating points and curved edges are extracted for the recognition of the different signs. The input for the project is taken from video data of a human hand gesturing all the signs of the American Sign Language [4].

In recent years, enormous research is progressing in the field of computer vision and human computer interaction where hand gestures play a vital role. Hand gestures are more powerful means of communication for hearing impaired when they communicate to the normal people everywhere in day to day life. As the normal people find little difficulty in recognizing and interpreting the meaning of sign language expressed by the hearing impaired, it is inevitable to have an interpreter for translation of sign language. To overcome this difficulty, an automatic hand gesture recognition system which translates the sign language into text needs to be developed. In this paper, a static hand gesture recognition system for American Sign Language using Edge Oriented Histogram (EOH) features and multiclass SVM is proposed. The edge histogram count of input sign language alphabets is extracted as the features and applied to a multiclass SVM for classification. The average accuracy of the system is compared with different number of features and the experimental findings demonstrate that the proposed method gives a success rate of 93.75% [5].

Visual Interpretation of gestures can be useful in accomplishing natural Human Computer Interactions (HCI). In this paper we proposed a method for recognizing hand gestures. We have designed a system which can identify specific hand gestures and use them to convey information. At any time, a user can exhibit his/her hand doing a specific gesture in front of a web camera linked to a computer. Firstly, we captured the hand gesture of a user and stored it on disk. Then we read those videos captured one by one, converted them to binary images and created 3D Euclidian Space of binary values. We have used supervised feed-forward neural net based training and back propagation algorithm for classifying hand gestures into ten categories: hand pointing up, pointing down, pointing left, pointing right and pointing front and number of fingers user was showing. We could achieve up to 89% correct results on a typical test set [6].

This paper presents an automatic translation system of gestures of the manual alphabets in the Arabic sign language. The proposed Arabic Sign Language Alphabets Translator (ArSLAT) system does not rely on using any gloves or visual markings to accomplish the recognition job. As an alternative, it deals with images of bare hands, which allows the user to interact with the system in a natural way. The proposed system consists of five main phases; pre-processing, best-frame detection, category detection, feature extraction and classification. The extracted features used are translation, scale, and rotation invariant, which make the system more flexible. Experiments revealed that the system was able to recognize the 30 Arabic alphabets with an accuracy of 91.3% [7].

This paper presents a novel technique for hand gesture recognition through human–computer interaction based on shape analysis. The main objective of this effort is to explore the utility of a neural network-based approach to the recognition of the hand gestures. A unique multi-layer perception of neural network is built for classification by using back-propagation learning algorithm. The goal of static hand gesture recognition is to classify the given hand gesture data represented by some features into some predefined finite number of gesture classes. The proposed system presents a recognition algorithm to recognize a set of six specific static hand gestures, namely: Open, Close, Cut, Paste, Maximize, and Minimize. The hand gesture image is passed through three stages, pre-processing, feature extraction, and classification. In pre-processing stage some operations are applied to extract the hand gesture from its background and prepare the hand gesture image for the feature extraction stage. In the first method, the hand contour is used as a feature which treats scaling and translation of problems (in some cases). The complex moment algorithm is, however, used to describe the hand gesture and treat the rotation problem in addition to the scaling and translation. The algorithm used in a multi-layer neural network classifier which uses back-propagation learning algorithm. The results show that the first method has a performance of 70.83% recognition, while the second method, proposed in this article, has a better performance of 86.38% recognition rate [8].

Organization of paper as follows: section I tell us about the introduction of sign gestures overviews and also we explains about analysis of previous research where we explains what are the works already done and what are the drawbacks. A brief of system overview about detail overviews acquisition of dataset, Steps of algorithm, extraction of key frame, feature extraction and different classifiers is explained in section II. In section III we have discussed the experimental results and what are the findings from those experimental results. In Section IV incorporates conclusion and future work of the paper. End of the paper includes acknowledgement and references.

II. System Overviews

This paper focuses on the planned continuous ISL gesture recognition system. The system has two phases, Training and testing phase. In training phase set of trained gesture signed are stored and trained to extract feature and compute values from each sign image frame from the videos[9,10,11]. In testing phase takes the testing sample and extract the features from sample sign image frame from video sample and compute the values and compare with data base previously stored features values and then classifies based on similar matched features where matched feature are extracted using orientation histogram (OH) with principal component analysis (PCA) is applied for reducing dimension of features obtained after OH and then perform mapping of gesture sign classified to equivalent characters, words and sentences. Finally generate the meaning full equivalent Kannada text description. The system dataset consisting of a set of signs wherever single hand or both the hands has been used for playacting continuous ISL gestures. 75 sentences database has been created[12,13,13,15].

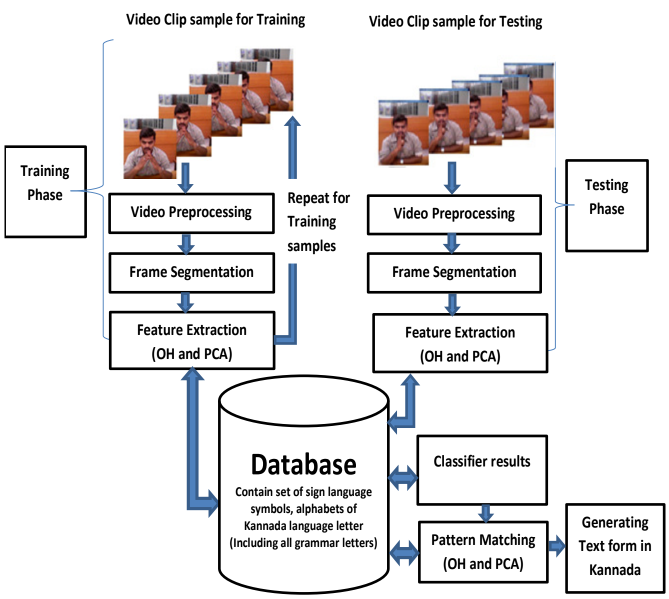

Figure 1. Shows block diagram of the proposed method.

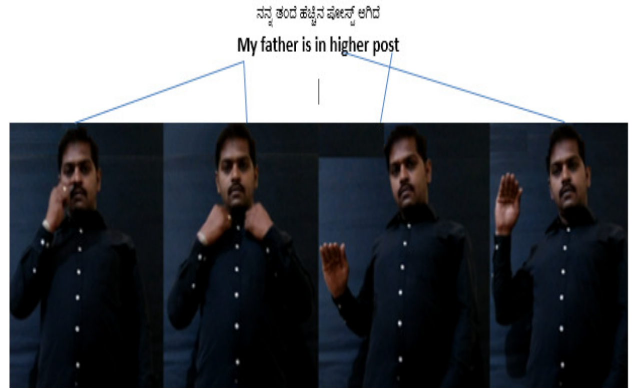

Figure 2. Shows sentence in English "My father is in higher post" and equivalent text in Kannada form is "ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ".

Every sentence consists of 3, 4 and 5 forms of gestures that is shown in figure 2 and. Each sentence may be a combination of static and dynamic gestures. Extracting begin frame and finish frame of every gesture is that the main downside in continuous language gesture recognition system as a result of it consists of a set of substantive gestures and additionally an imprecise gestures having no that means. We tend to deal this problem using gradient based key frame extraction method[16,17]. Here major modification within the gradient shows finish of the one gesture and begin of another gesture. Key frame helps to interrupt every sentence into sequence of words (isolated gestures) and additionally obliging for extracting frames of substantive gestures. Orientation histogram, DWT and PCA is employed for extracting features options of these frames that includes of substantive gestures. The block diagram of proposed framework is shown in figure 1.

1.1 Acquisition of Dataset

Dataset has been created using an associate external camera with the configuration of Canon EOS with 18–55mm lens, 18 mega pixels, 29 frames per second and resolution is 3920*320 bits/sec. Here we tend to used single camera for creation of gesture dataset. Black background is employed for data base creation of ISL gestures. Here we tend to consider the upper body part only[18,19,20]. Movement of the upper body part is appropriate. Position of camera is extremely vital (camera calibration), for clarity of the dataset and for removing several backgrounds connected issues like background noise, body motion etc.

Here we have taken 75 Indian sign sentences of 5 totally different hearing impaired individuals from deaf and dumb school at Laxmeshware Taluk, Gadag District, Karnataka, India, wherever every sentence has been recorded 10 times, 6 for coaching and 4 for testing. Each video is split into sequence of frames of size 640*480. Continuous gesture “My father is in higher post” (“ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ”) is created of three gestures (“My father” (“ನನ್ನ ತಂದೆ “), “post” (“ಪೋಸ್ಟ್ ಆಗಿದೆ“) and “higher”(“ಹೆಚ್ಚಿನ“)) shown in figure 2.

2.2 Pre-processing

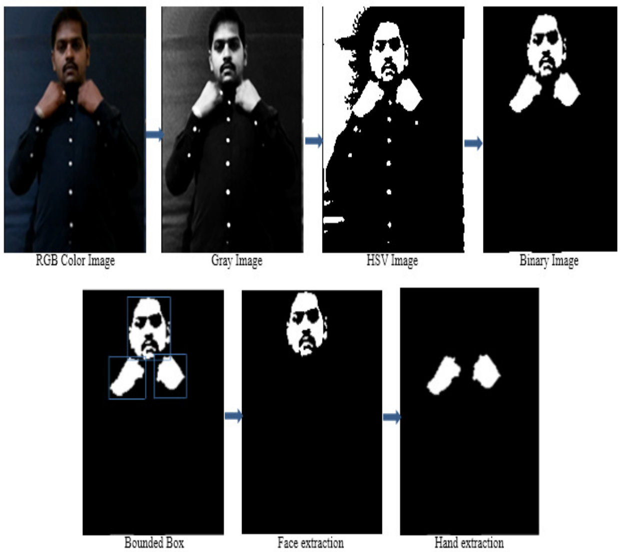

Figure 3.Shows pre-processing steps of each RGB frames from continuous gesture for the sentence “My father is in higher post”. (“ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ”)

In this step silhouette pictures of each hand gestures square measure created. Here we have a tendency to extracts foreground image from complete image means that it removes background of a picture and acquire the skeleton of upper body part. Then hand region square measure ablated from these foreground pictures by eliminating largest connected region that is face. Finally we have a tendency to get the hand portion from upper body[21,22,23].

In this proposed system we initial convert every video into sequence of RGB frames. Every frame having dimension 640*480. Skin colour segmentation is applied for extraction of skin region that is split into range of chunks. For locating skin region, every frame is regenerate into HSV (Hue, saturation, value) plane wherever solely H and S value having threshold (H > 0.55 or S <= 0.20 or S > 0.95) is employed for locating non skin region of an image. Then this region mark as zero for extracting skin region. Median filter is applied for preserving outer boundary (edges) of segmented region. It mainly removes salt pepper noise and impulsive noise for edge preservation[24]. Images obtained once median filtering are regenerate into binary form. At the end of pre-processing compute largest connected region that is face. Eliminate face region from upper half the body and can get hand gestures[25,26]. Each step of pre-processing is shown in figure 4 and explained below.

Algorithm Steps

1. Every video is converted-again into RGB frames (I). Then RGB frames are born-again into HSV plane. Here we have a tendency to contemplate solely H and S half as a result of H and S shows non-skin region. The edge worth of H and S for non-skin region is (H > 0.55 or S <= 0.20 or S >0.95). Create 0 for all non-skin region in image I and obtained colour region in image (I) denoted as I1.
2. Convert image I into grey type by compute I1 image from grey scale of I.
3. Then apply median filter of window size [3, 3] to get rid of noise from I ‘m’ image.
4. Calculative the world of every binary region, initial we've to label them then calculate space of every region.
5. Type spaces of all region for extracting largest area. Here we have a tendency to contemplate face has the biggest space. From this extract face region from the full binary image. Currently compute binary image (bw) from face region to induce hands in binary image that's the region of interest[27].

2.3 Extraction of Key Frame

A finite sequence of frames kind every video. During which frames can be any gesture frame or non-gesture frame[28]. Therefore we want to extract those frames that belong to any significant gesture and take away insignificant frames as a result of these frame creates an additional affliction of process. During this paper we tend to used gradient methodology for extracting key frames of every video sequence. We take frames to try and do segmentation and calculate gradient of every frame. Key frame extraction graph of “My father is in higher post” (“ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ“) gesture is shown in figure 4. Graph shown in figure 4 shows that gesture1 starts in between frame 0th to frame 4th and finish at in between frame 254th to frame 258th. From 0th to frame 4th, there is a constant gradient value and similarly from 254th to frame 258th we tend to obtained constant gradient values that shows end of one gesture or begin of another gesture[29]. Suppose if gesture 1 ends at frame 256th than this can be the beginning of next gesture. From this we will calculate the whole variety of frames in every gesture. 15 frames each gesture are considered from the middle of the each gesture data set because in between of frames gesture are clearly visible and identifiable. This methodology has been used for creating database for dividing every continuous gesture into isolated gestures[30].

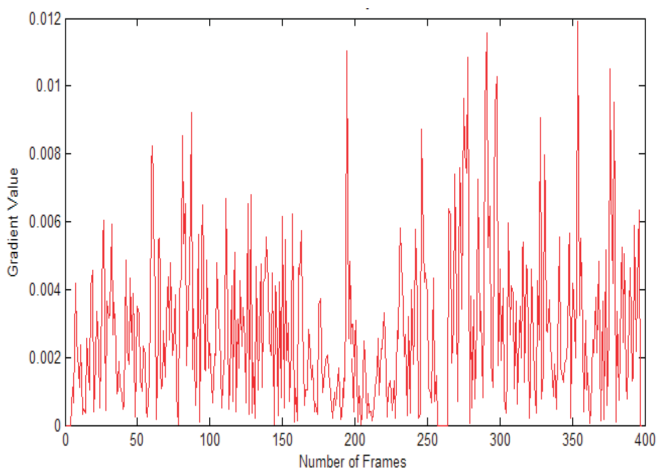

Figure 4. Key frame extraction of “ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ” gesture.

2.4 Feature Extraction

We applied orientation histogram as a feature extraction technique for extracting most applicable features of every gesture. It offers convenience to even light weight condition changes and scene illumination changes. The edge of the image of the scene is still same. All the continual ISL gestures are obtained in normal lighting mode wherever element intensities may be urged to vary the scene lighting. Another advantage that is utilized on orientation histogram is translation invariant property. It determines that same feature vectors are made by constant frames at totally different position of gestures. It is achieved to measure the local orientation histogram for all the frames of the dynamic gestures. Local orientation histogram doesn’t modification by the interpretation of the enclose the gesture. Then dimension of those options are reduced using Principal component analysis (PCA). The steps of orientation histogram algorithm are:

i) Subsample the 640*480 image into 60*40 size that reduces space complexity and makes processing time fast.
ii) Finding the edges of an image using 3-tab derivative filter a = [0 − 1 1] b = [01− 1]. It helps us for finding the image gradient in a-direction and b-direction.
iii) The gradient in a-direction furthermore as in b-direction.

$$da = \frac{\delta p(a,b)}{\delta a} = \frac{p(a+1,b)-p(a-1,b)}{2} \quad (1)$$

$$db = \frac{\delta p(a,b)}{\delta b} = \frac{p(a,b+1)-p(a,b-1)}{2} \quad (2)$$

Where p(x, y) represents intensity function at (x, y) pixels position.

iv) Find out the gradient direction using arctan2 function which is expressed as,

$$X(a,b) = \mathrm{atan2}\left(\frac{\delta p}{\delta a}\delta p\delta b\right) = \mathrm{atan2}\left(\frac{\delta b}{\delta a}\right) \quad (3)$$

The value of X lies between $\left[-\frac{\pi}{2} \quad \frac{\pi}{2}\right]$

v) Magnitude

$$mg\,(a,b) = \sqrt{da^2 + db^2} \quad (4)$$

vi) Convert these values into a column vector so the radian values will be converted into the degrees. Here 180^0 degree is split into 18 and 36 bins, in 18 bins every bin is of 10^0 and in 36 bins every bin is of 5^0. The polar plot for 18 bins and 36 bins of “My father” (“ನನ್ನ ತಂದೆ”) gesture are shown in Figure 5. This polar plot shows angle of variation within the hand at the time of performing any gesture.

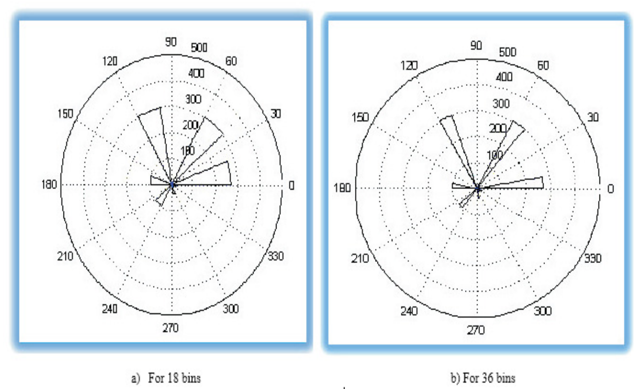

Figure 5. Shows polar plot of orientation histogram of “My father” (“ನನ್ನ ತಂದೆ”) gesture.

Principal Component Analysis

It is essentially used for locating patterns in input data for light similarities and variations between them and for reducing dimension of data set. Once extracting these patterns kind data then we have to compress the data using PCA. Finding patterns from the large data set is very difficult. Thus for analyzing data set PCA may be a robust tool. PCA may be a non-parametric, easy technique for extracting important information from confused data. Eigen value generated from PCA gives projection direction of confused data set.

2.5 Different Classifiers

Here gesture recognition is done using different distance Metrics like Correlation Distance, Chess Board Distance, Cosine Distance, Mahalanobis Distance, City Block Distance and Euclidean Distance. When pre-processing every probe sentences are divided into sequence of isolated gestures. Then extract features of every gesture and matched the number of frames have been successfully classified. After classification process a text formation has been performed.

1. Correlation Distance

Statistical dependency between random variables like projected trained image (r) and projected test image (s) is measured by using correlation distance.

$$CO(r,s) = 1 - \frac{(r-mean\ (r)).(s-mean\ (s))}{|(r-mean\ (r))|\ .\ |(s-mean\ (s))|} \quad (5)$$

2. Chess Board Distance

This distance comes into focus from game of chess wherever kings take minimum range of moves to travel from one square to another in chessboard. The chessboard distance between two vectors projected trained image (r) and projected test image (s) is represented by,

$$CB\ (r,\ s) = Max_i\,(|ri| - |si|) \qquad (6)$$

It gives some times better result than Euclidean distance.

3. Cosine Distance

It is a complement of the cosine similarity in positive space. It provides distance in terms of angular cosine in between two vectors r and s. Cosine similarity is a measure of similarity between two vectors projected trained image (r) and projected test image (s) by calculating the cosine of the angle between them. It can be represented as,

$$C\,S\ (r,\ s) = cos(Q) = \frac{r\ .\ s}{|r|\ |s|} \qquad (7)$$

Cosine distance can be represented as,

$$C\,D\ (r,\ s) = 1 - \frac{r\ .\ s}{|r|\ |s|} \qquad (8)$$

4. Mahalanobis Distance

When we have two vectors projected trained image (r) and projected test image (s) of the same distribution with covariance matrix S then the Mahalanobis Distance will be,

$$M\ (r,\ s) = \sqrt{(\bar{r} - \bar{s})'\ Co^{-1}\ (\bar{r} - \bar{s})} \qquad (9)$$

Covariance is a measure of the random variables from two ordered pair of sets either data move in the same direction.

$$Co\ (r,\ s) = 1/r\ \left(\textstyle\sum_{i=1}(ri - \bar{s})\ \ (ri - \bar{s})\right) \quad (10)$$

5. City Block Distance

The shortest distance between two points is along the hypotenuse that is calculated by using Euclidean distance. Rather than this area distance is measured because the length in x direction another with length in y direction. The space measured by area is often larger than or equal to zero. If points are same than distance is zero otherwise larger than zero. Most of the times city block distance gives similar result with Euclidean geometrician distance.

$$CT\ (r,\ s) = \sum |ri - si| \qquad (11)$$

6. Euclidean Distance

Here Euclidean distance is employed to search out the gap between projected trained image (r) and projected test image (s). Keep minimum distance value from the all trained image and discard the rest. Distance between two points p and q in Euclidean space is named Euclidean distance. Euclidean space is of n-dimension if r = (r1, r2, . . . , r_n) and s = (s1, s2, . . . , s_n), here r and s are called as Euclidean Vectors.

$$E(r,\ s) = \textstyle\sum_{i=0}^{n}(|ri - si|^2)^{1/2} \qquad (12)$$

III. Experimental Results and Analysis

Experiments are performed on 75 different form of sentences. Every sentence having 3, 4 or 5 gestures. Here every continuous gesture is formed of static as well dynamic gesture for performing experiments. The system considered training data base sets of 75 samples, in the paper only 5 data set is shown in table1. Every sentence are going to be recorded 10 times, 6 times for training and 4 times sentence is employed for testing. In every sentence we have a tendency to contemplate 20 frames of every gesture out of n range of frames (n is vary from sentence to sentence) for training and 10 frames of every gesture for testing. Here only those frames are considered which are present at the middle because it consists of most informative frames. Here experiments are performed on 18 bins additionally as on 36 bins which suggest 180^0 is split into 8 elements every half is of 10^0. Equally for 36 bins [30]. From experimental results we have a tendency to found that the results obtained from Euclidian distance and correlation have higher recognition rate than alternative distance primarily based classifiers like city-block distance, checkerboard distance etc. Table 2 and table 3 shows that the orientation histogram with 36 bins provide higher accuracy then 18 bins as a result of it measures angle of change of hand abundant acceptable. In 36 bins histogram there is a 5 degree resolution and in 18 bins. Figure 6 and figure 7 shows various classifier Vs Different sentences of OH (18 bins) and OH (36 bins). There is a 10^0 degree resolution. From above table we also seen that the recognition rate of some gesture is much higher than some other gestures. This will happen because some gestures are of similar type like “ನೀವು” and “ನಾನು” etc. Those gestures are of same type it gets misclassified and provides wrong results.

Table 1.Shows training data base sets.

No	Description In English	Description In Hindi	Sign Video Samples	Description In Kannada
1.	Hello, What is your name	नमस्कार, आपका नाम क्या है	Hello,What is your name_.3gp	ನಮಸ್ಕಾರ, ನಿಮ್ಮ ಹೆಸರೇನು
2.	Please wait for sometime	कुछ देर के लिए कृपया प्रतीक्षा करें	Please wait for sometime.3gp	ದಯವಿಟ್ಟು ಸ್ವಲ್ಪ ಸಮಯದ ನಿರೀಕ್ಷಿಸಿ
3.	Do you want something to drink	आपको कुछ चाहिए पीने के लिए	Do you want something to drink_.3gp	ನೀವು ಕುಡಿಯಲು ಏನಾದರೂ ಬಯಸುವಿರಾ
4.	I had to say something but I forgot	मैं कुछ कहना चाहता था लेकिन मैं भूल गया	I had to say something but i forgot.3gp	ನಾನು ಏನೋ ಹೇಳಲು ಹೊಂದಿತ್ತು ಆದರೆ ನಾನು ಮರೆತಿದ್ದಾರೆ
5.	My father is in higher post	मेरे पिता ने उच्च पद में है	My father is in higher post.3gp	ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ

Table 2. Classification results of OH (18 bins) with PCA at various distance based classifier.

Sr. No	Sentences	Correlation distance	Chessboard distance	Cosine distance	Mahalanobis distance	City block distance	Euclidean distance
1.	ನಮಸ್ಕಾರ, ನಿಮ್ಮ ಹೆಸರೇನು ?	92%	85%	93%	83%	83%	92%
2.	ದಯವಿಟ್ಟು ಸ್ವಲ್ಪ ಸಮಯದ ನಿರೀಕ್ಷಿಸಿ	91%	83%	89%	83%	80%	91%
3.	ನೀವು ಕುಡಿಯಲು ಏನಾದರೂ ಬಯಸುವಿರಾ	89%	81%	91%	79%	82%	94%
4.	ನಾನು ಏನೋ ಹೇಳಲು ಹೊಂದಿತ್ತು ಆದರೆ ನಾನು ಮರೆತಿದ್ದಾರೆ	91%	83%	90%	80%	86%	91%
5.	ನಿಮ್ಮ ಕುಟುಂಬದಲ್ಲಿ ಎಷ್ಟು ಜನರಿರುತ್ತಾರೆ	92%	84%	87%	84%	84%	90%
6.	ನನ್ನ ಕುಟುಂಬದಲ್ಲಿ ನಾಲ್ಕು ಜನರಿದ್ದಾರೆ	91%	81%	90%	79%	82%	91%
7.	ನಿಮ್ಮ ತಂದೆ ಮಾಡುತ್ತಿರುವುದಾದರೂ ಏನು ?	93%	80%	91%	80%	79%	92%
8.	ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ	81%	81%	81%	68%	72%	83%
9.	ದಯವಿಟ್ಟು ಕಸ ಎಸೆದು, ಬುಟ್ಟಿ ಬಳಸಿ	91%	79%	88%	83%	80%	94%
10.	ನಾವು ನಾಳೆ ಭೇಟಿ ಮಾಡಬಹುದು	89%	79%	88%	81%	77%	92%

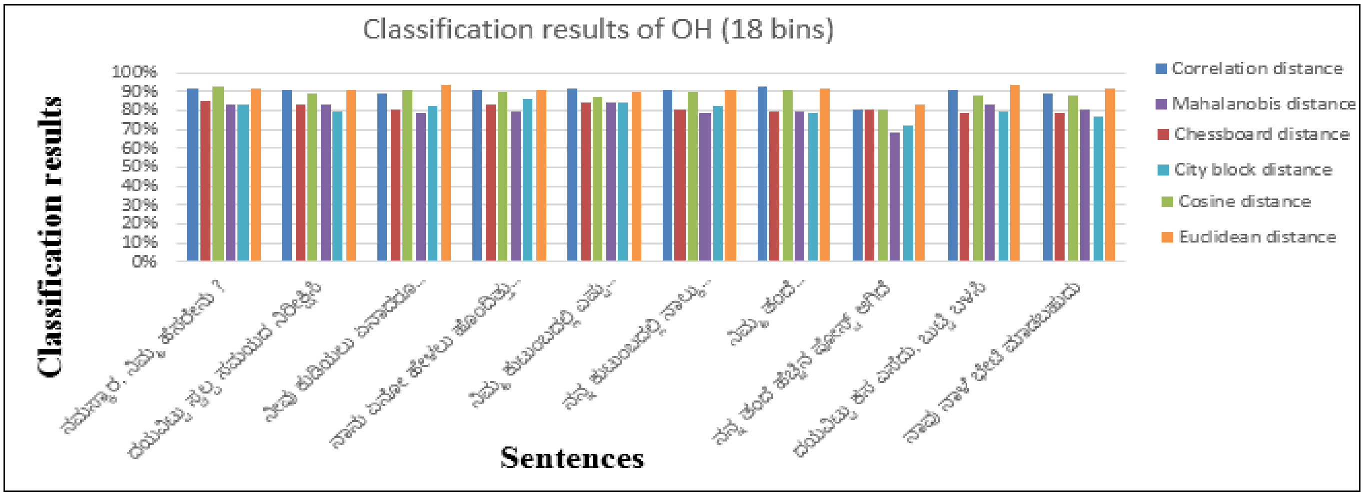

Figure 6.Shows various classifier vs Different sentences of OH (18 bins)

Table 3. Classification results of OH (36 bins) with PCA at various distance based classifier.

Sr. No	Sentences	Correlation distance	Chessboard distance	Cosine distance	Mahalanobis distance	City block distance	Euclidean distance
1.	ನಮಸ್ಕಾರ, ನಿಮ್ಮ ಹೆಸರೇನು ?	94%	85%	93%	83%	83%	94%
2.	ದಯವಿಟ್ಟು ಸ್ವಲ್ಪ ಸಮಯದ ನಿರೀಕ್ಷಿಸಿ	91%	83%	89%	83%	80%	91%
3.	ನೀವು ಕುಡಿಯಲು ಏನಾದರೂ ಬಯಸುವಿರಾ	92%	90%	91%	91%	88%	94%
4.	ನಾನು ಏನೋ ಹೇಳಲು ಹೊಂದಿತ್ತು ಆದರೆ ನಾನು ಮರೆತಿದ್ದಾರೆ	91%	91%	90%	83%	86%	93%
5.	ನಿಮ್ಮ ಕುಟುಂಬದಲ್ಲಿ ಎಷ್ಟು ಜನರಿರುತ್ತಾರೆ	92%	84%	87%	85%	84%	94%
6.	ನನ್ನ ಕುಟುಂಬದಲ್ಲಿ ನಾಲ್ಕು ಜನರಿದ್ದಾರೆ	91%	81%	90%	79%	82%	91%
7.	ನಿಮ್ಮ ತಂದೆ ಮಾಡುತ್ತಿರುವುದಾದರೂ ಏನು ?	93%	80%	91%	80%	79%	92%
8.	ನನ್ನ ತಂದೆ ಹೆಚ್ಚಿನ ಪೋಸ್ಟ್ ಆಗಿದೆ	83%	81%	81%	72%	76%	86%
9.	ದಯವಿಟ್ಟು ಕಸ ಎಸೆದು, ಬುಟ್ಟಿ ಬಳಸಿ	92%	82%	88%	83%	80%	94%
10.	ನಾವು ನಾಳೆ ಭೇಟಿ ಮಾಡಬಹುದು	89%	79%	88%	81%	77%	94%

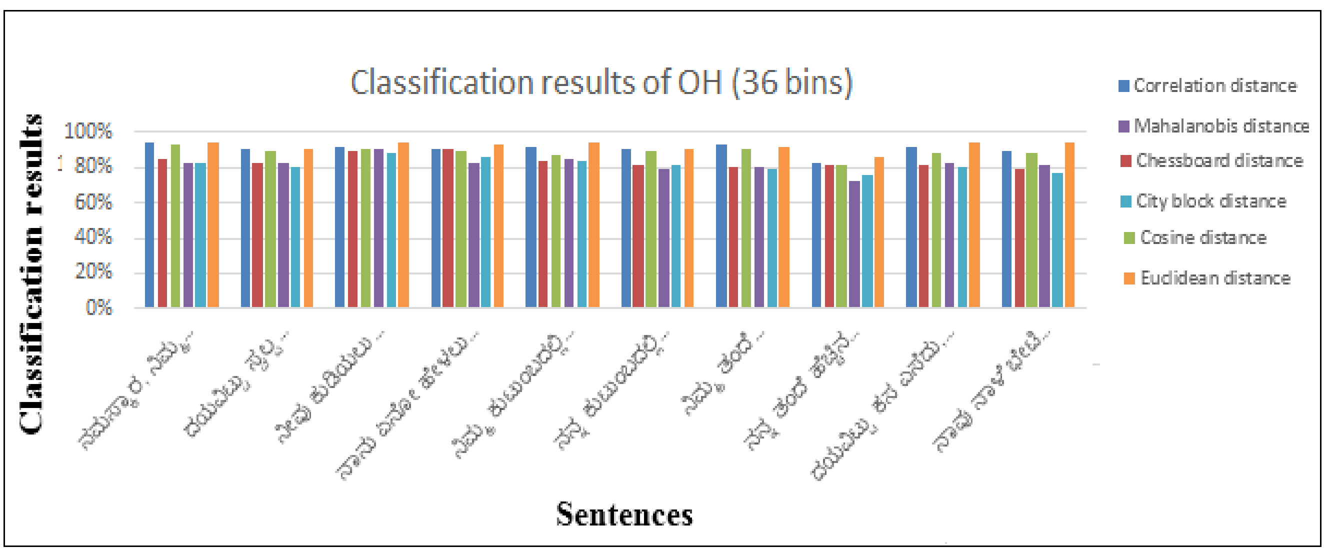

Figure 7.Shows various classifier vs Different sentences of OH (36 bins)

CONCLUSION

The proposed system for continuous ISL gesture provides satisfactory performance together with feature obtained using orientation histogram with Principal component analysis, wherever each the hands are used for performing any gesture. In continuous ISL recognition system key frame extraction is the foremost step. It helps for extracting the continuous gesture into isolated gestures also it show how many number of frames an isolated gesture can have. Once hand segmentation we have a tendency to applied Orientation histogram for extracting features of hands for training dataset and similarly for testing. Minimum distance shows maximum classification rate from probe dataset to training dataset. Here classification accuracy is measured with the maximum number of matched frames. Experimental result shows that the designed method provides satisfactory results with Euclidean distance correlation. Results also are tested using normal webcam and get appropriate results.

This work has been enhanced by creating dataset with different background and different illumination conditions. Here we have a tendency to apply a lot of acceptable features that incorporate form of hand within the time of acting gestures, speed of performing every gesture etc. There are varied different classifiers like Support vector machine (SVM) and Hidden Markov Model (HMM) has been applied for classification.

ACKNOWLEDGMENT

We would like to thank Chairman Dr. Ajeenkya D Y Patil, Principal Dr. Ashok Kasnale, Colleague of the Department and Colleagues of the varies Department the Dr. D. Y. Patil School of Engineering and Technology, Lohegaon, Pune Maharashtra, India, for their support, suggestions and encouragement.

REFERENCES

[1] M. Mohandes, M. Deriche, and J. Liu, "Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition", In Proc. IEEE Transaction on Human Machine System, ,pp.2168-2291, 2014.

[2] P. Nanivadekar and V. Kulkarni, "Indian Sign Language Recognition: Database creation, Hand Tracking and Segmentation " , in IEEE Conference On Circuit System, Communication And Information Technology Applications, 2014,

[3] K.Modi and A More, "Translation of Sign Language Finger-Spelling to Text using Image Processing", In conference (IJCA), Volume 77, No.11, September, 2013.

[4] S .N . Omkar and M. Monisha," Sign Language Recognition Using Thinning Algorithm", ICTACT Journal On Image And Video Processing, Volume 02, Issue: 01¸ August 2011.

[5] S. Nagarajan and T. S. Subashini, "Static Hand Gesture Recognition for Sign Language Alphabets using Edge Oriented Histogram and Multi Class SVM", International Journal of Computer Applications (0975 –8887),Volume 82, No4, November 2013.

[6] G. R. S. Murthy and R. S. Jadon , "Hand gesture recognition using neural networks", Advance Computing Conference (IACC), IEEE 2nd International Conference, Patiala, pp 134 – 138, Feb2010.

[7] Nashwa El-Bendary , Hossam M. Zawbaa , Mahmoud S. Daoud and Aboul Ella Hassanien, "ArSLAT: Arabic Sign Language Alphabets Translator", Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on Date of Conference:8-10 Oct. 2010 Page(s):590-595 Conference Location:Krackow Publisher:IEEE.

[8] Haitham Hasan and S. Abdul-Kareem, "Static hand gesture recognition using neural networks", Spinger link: Artificial Intelligence Review February 2014, Volume 41, Issue 2, pp 147–181.

[9] Ramesh Kagalkar and Dr. Nagaraja H.N, "New Methodology for Translation of Static Sign Symbol to Words in Kannada Language"," International Journal of Computer Applications(IJCA), Volume 121, No.20,PageNo. 25-30, July-2015.

[10] Ramesh M. Kagalkar, Dr. Nagaraja H.N and Dr. S.V Gumaste",A Novel Technical Approach for Implementing Static Hand Gesture Recognition", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Volume 4, Issue 7, July 2015.

[11] Ramesh M. Kagalkar and Dr. S.V Gumaste, "Automatic Graph Based Clustering for Image Searching and Retrieval from Database", CiiT International Journal of Software Engineering and Technology, Volume 8, No 2, 2016.

[12] Ramesh M. Kagalkar and Dr. S.V Gumaste, "Review Paper: Detail Study for Sign Language Recognition Techniques", CiiT International Journal of Digital Image Processing, Volume 8, No 3, 2016.

[13] Rashmi B. Hiremath and Ramesh Kagalkar, "A Methodology for Sign Language Video Analysis and Translation into Text in Hindi Language",CiiT International Journal of Fuzzy System, Volume 8, No 5, 2016.

[14] Rashmi B. Hiremath and Ramesh Kagalkar, "Methodology for Sign Language Video Interpretation in Hindi Text Language", International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE), Volume. 4, Issue 5, May 2016.

[15] Amitkumar and Ramesh Kagalkar, "Sign Language Recognition for Deaf User", Internal Journal for Research in Applied Science and Engineering Technology (IJRASET), Volume 2, Issue 12, December 2014.

[16] Amit kumar and Ramesh Kagalkar, "Advanced Marathi Sign Language Recognition using Computer Vision", International Journal of Computer Applications (IJCA), Volume 118, No. 13, May 2015.

[17] Amit kumar and Ramesh Kagalkar, "Methodology for Translation of Sign Language into Textual Version in Marathi", CiiT, International Journal of Digital Image Processing, Volume 07, No.08, Aug 2015.

[18] Mrunmayee and Ramesh Kagalkar, "A Review On Conversion

of Image To Text as Well as Speech using Edge Detection and Image Segmentation", International Journal of Science and Research (IJSR), Volume 3, Issue 11, November 2014.

[19] Mrunmayee Patil and Ramesh Kagalkar, "An Automatic Approach for Translating Simple Images into Text Descriptions and Speech for Visually Impaired People", International Journal of Computer Applications (IJCA), Volume 118, No. 3, May 2015.

[20] Kaveri Kamble and Ramesh Kagalkar, "A Review: Translation of Text to Speech Conversion for Hindi Language", International Journal of Science and Research (IJSR), Volume 3, Issue 11, November 2014.

[21] Kaveri Kamble and Ramesh Kagalkar, "Audio Visual Speech Synthesis and Speech Recognition for Hindi Language" ,International Journal of Computer Science and Information Technologies (IJCSIT), CiiT International Journal of Data Mining Knowledge Engineering, Volume 6, Issue 2, April 2015.

[22] Kaveri Kamble and Ramesh Kagalkar " A Novel Approach for Hindi Text Description to Speech and Expressive Speech Synthesis" International Journal of Applied Information Systems (IJAIS), Volume 8, No.7, May 2015.

[23] Shivaji Chaudhari and Ramesh Kagalkar "A Review of Automatic Speaker Recognition and Identifying Speaker Emotion Using Voice Signal", International Journal of Science and Research (IJSR), Volume 3, Issue 11, November 2014.

[24] Shivaji Chaudhari and Ramesh Kagalkar, "Automatic Speaker Age Estimation and Gender Dependent Emotion Recognition", International Journal of Computer Applications (IJCA), Volume 117, No. 17, May 2015.

[25] Shivaji J. Chaudhari and Ramesh M Kagalkar, "A Methodology for Efficient Gender Dependent Speaker Age and Emotion Identification System", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), Volume 4, Issue 7, July 2015.

[26] Ajay R. Kadam and Ramesh Kagalkar, "Audio Scenarios Detection Technique", International Journal of Computer Applications (IJCA), Volume 120, No. 16, June 2015.

[27] Ajay R. Kadam and Ramesh Kagalkar, "Predictive Sound Recognition System", International Journal of Advance Research in Computer Science and Management Studies (IJARCSMS), Volume 2, Issue 11.

[28] Swati Sargule and Ramesh M Kagalkar, "Hindi Language Document Summarization using Context Based Indexing Model", CiiT International Journal of Data Mining Knowledge Engineering, Volume 08, No. 01, Jan 2016.

[29] Swati Sargule and Ramesh M Kagalkar, "Methodology of Context Centered Term Indexing Style Intended For Hindi Language Document Summarization", CiiT International Journal of Software Engineering, Volume 8, No 5, 2016.

[30] Vandana D. Edke and Ramesh M. Kagalkar, "Video Object Description of Short Videos in Hindi Text Language", International Journal of Computational Intelligence Research, Volume 12, Number 2 (2016), pp. 103-116 © Research India Publications.

Authors Profile

Ramesh. M. Kagalkar

He was born on Jun 1st, 1979 in Karnataka, India and presently working as an Assistant. Professor, Department of Computer Engineering, Dr. D Y Patil School of Engineering and Technology, Charoli, B.K.Via Lohegaon, Pune, Maharashtra, India. He has 14.5 years of teaching experience at various institutions. He is a Research scholar in Visveswaraiah Technological University, Belgaum, He had obtained M.Tech (CSE) Degree in 2006 from VTU Belgaum and He received BE (CSE) Degree in 2001 from Gulbarga University, Gulbarga Karnataka, India. He is the author of two text book; 1.Advance Computer Architecture, 2. The Swift Practical Approach of Learning C-Programming in LAP LAMBERT Academic Publishing, Germany (Available in online) and One of his research article A Novel Approach for Privacy Pre- serving has been consider as text in LAP LAMBERT Academic Publishing, Germany (Available in online). He is waiting for submission of two research articles for patent right. He has published more than 35 research papers in International Journals and presented few of there in international conferences. His main research interest includes Image processing, Gesture recognition, Speech processing, Voice to sign language and CBIR. Under his guidance Ten ME students awarded degree in SPPU, Pune, three students at the edge of completion their ME final dissertation reports and four students started new research work and they have publish their research papers on International Journals and International conference. He can be contacted by email rameshvtu10@gmail.com.

Dr. S.V Gumaste

Dr. S.V Gumaste was born on 02 April 1968 in Karnataka, India and presently working as a Professor & Head, Computer Engineering, Sharadchandra Pawar College of Engineering, Dumberwadi, Otur Dist: Pune Maharashtra, India. He has 22.5 years of teaching experience at various institutions. He received Ph.D degree from Sant Gadge Baba Amravati University, Amravati in 2015. He obtained M.E (CSE) Degree in 2007 from Prof. Ram Meghe Institute of Technology & Research, Badnera, Amravati. (Sant Gadge Baba Amravati University, Amravati, Maharashtra) and He received BE (CSE) Degree in 1992 from BLDE College of Engineering & Technology, Bijapur. (Karnataka University, Dharwar). He is the author of 03 Books such as fundamentals of C Programming, Exel Publication, Amravati. (In August, 2005), C Programming with 150 solved programs, Exel Publication, Amravati. (In January, 2005) and Basics of C Programming: A Practical Approach, Jai Publication, Pusad. (In August, 2004). He has published more than 12 research papers in International Journals and presented 12 papers in international conferences. His main research interest includes networking, image processing and speech processing. He is the life Membership with Professional Bodies of LM-20507, Indian Society for Technical Education, New Delhi, Institutional Member (N0046703), Computer Society of India, Mumbai (2008-09) and (MH/222/1645), All India Federation of University & College Teachers' Organisations, Hyderabad (AIFUCTO). He can be contacted by email: svgumaste@gmail.com

