International Journal of Engineering & Technology, 7 (2.8) (2018) 271-277

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

A probabilistic feature based SVM model for Hindi/English speech recognition

Priyanka Bansal[1*], Syed Akhtar Imam[2]

[1]*Research Scholar, Department of ECE, Jamia Millia Islamia, New Delhi, India*
[1]*Assistant Professor, Department of ECE, Manav Rachna International University, Faridabad*
[2]*Associate Professor, Department of ECE, Jamia Millia Islamia, New Delhi, India*
**Corresponding author E-mail: priyankabansalchamp@gmail.com*

Abstract

Real time speech recognition has various challenges including noise, turbulence, language and crosstalk problem. In this paper, multi-phase hybridization is applied to cover these challenges and to provide effective speech recognition. The model is explicitly divided into three main stages where each stage is implicitly divided into several sub-stages to provide specific problem solution. The proposed hybrid model resolved the problem of acoustic turbulence, background noise and instrumentation noise problem at the earlier stage. The rectified speech signals are processed using ICA and Fuzzy-HMM approach to generate the structural and statistical features. In this stage, the signal is divided in smaller linear blocks to extract the features. Later on, fuzzy-weighted SVM is implied to recognize the speech signal. The experimentation is applied on Hindi and English characters and sentence datasets. The comparative results are derived against BPNN and PCA models for different sample sets. The comparative results obtained from model signifies that the model has improved the recognition rate effectively.

Keywords: *Speech Recognition, Fuzzy-Weighted, SVM, HMM, Featured.*

1. Introduction

Speech signal carries biometric information, which can be used for authentication, commending and information transfer. Each stage and application of speech capturing suffer from different associated challenges. Some of these applications, technology specific challenges and applications are listed in table 1.

Table 1: Speech Processing Challenges

Technology/Application	Challenges	Impact
Speech Capturing	1. The device specific noise, including instrumentation noise, echo problem 2. Background Noise or the environmental disturbance 3. Voice Mixing or overlapping	1. The quality of speech processing and authentication methods affects 2. Disturbed Speech cannot ensure effective results
Speaker Recognition	1. Speech Variation in two different talks 2. Mood dependent pitch and frequency 3. Words gap and speed 4. For distance, cannot ensure the presence of a person, recorded voice can be used	1. Speaker recognition is used in various voice commanders and authentication system. 2.A highly accurate recognition process is required
Speech Recognition	1. Grammar and sentence formation 2. Language Problem or mispronunciation of some words 3. Specification of phrase or sentence termination for applying the approach 4. Intelligent Mapping required to identify the word sense	1. Used for speech to text conversion and in language transition 2. Improper sense identification can change the meaning of identifying word or sentence
Language Translation	1. The vast dictionary of both languages is required along with relative mapping 2. Intelligent sentence sensing system must map based on exact meaning 3.Mispronunciation can change the meaning and sense	1. The globalization of speech communication to represent in different languages 2. Can be used for distance, regional and language independent talks

In more advanced forms, hybrid methods can rectify the speech signal and to generate the integrated speech features. These two stages are application independent approaches applied to improve the quality of speech and speech acquired information. Later on stages are application specific. This paper is specifically defined for speech recognition method. To accomplish the task, the first requirement is to identify the basic and the descriptive constraints of speech. Some specific constraints are listed in table 2.

Table 2: Constraints for Effective Speech Recognition

Constraint	Description
Language	1. Same method cannot be applied to all languages. Each language has own criticalities. The set of basic components different from language to language. Such as in Hindi, the number of vowels, consonants is different and applied in different form
Type	1. The recognition process can be applied based on the algorithmic specification in characters, words and sentences 2. For some applications such as voice commanders, some specific commands must recognize 3. Single person or multiple speech also affects the recognition process
DB	1. A vast dictionary and training process must improve the robustness and accuracy of recognition process 2. The recording must be performed in the same environment and from similar or same device for which the recognition system is defined. 3. The database must have multiple instances of the same person for a single word or character for robust results 4. Training set must be collected in different male and female voices to improve the probabilistic speech mapping.
Algorithm Specification	1. The feature set formulation is required based on signal level, structural or the statistical observation. 2. Some deterministic, probabilistic and supervised training must generate the recognition rules

A segmented featured transformed and fuzzy-SVM integrated method is defined to perform speech recognition on Hindi and English characters and words. The work is defined by a multi-stage filtration method to resolve different impurities at the earlier stage.

Later on the block segmented method is applied to generate speech features. These all features are applied on fuzzy method to generate the rule framed analysis. Finally, SVM is applied to improve the recognition process.

2. Research Methodology

Speech recognition is adopted by many applications for verification of existence and availability of an individual. But the speech captured in real and complex environment suffers from various kinds of challenges. These challenges are in terms of various noises and impurities that degrade the speech quality. The paper has combined the probabilistic and rule weighted method to recognize the word level speech effectively. The model is defined for dataset of both Hindi and English speech signals separately. To gain the higher recognition rate, the wider and multistage signal rectification is defined in this model. The raw real time signal is taken to form the training and testing datasets. These raw speech signals are processed by DWT, LPC, Band Pass filtration and Gaussian distance scoring methods to remove background noise, instrumentation noise and acoustic turbulence. The rectified signals are divided in linear blocks and processed by Fuzzy-HMM and ICA methods to generate the structural and statistical features. The featured training dataset is processed is processed under fuzzy weighted SVM classifier to generate the classification rules. Finally, to perform the speech signal recognition, the classification model is applied on test set. In this section, algorithmic description of each of the associated work stage is provided.

A. Signal Filteration

The real time speech captured in complex environment can have different kind of noises. These impurities can be recognized as instrumentation noise, background noise or the quality of speech signal. The existence of noise in signal reduces the chances of accurate speech recognition. The multi-phase filteration to handle various kind of signal impurities are defined in this section with algorithmic specification.

1) DWT for Background noise Filteration

The speech capturing in complete silence is almost impossible in real environment. The online speech processing is always disturbed by some background noise or disturbance. This noise can occur throughout the speech in a different form, i.e. in traffic, machine noise, fan sound, etc. A smaller segment driven noise probability identification and rectification using DWT based band pass filtration is applied here to remove the noise. It is a fast and the real time process signal to react against noise coefficient vector appropriately. The integer arithmetic is also applied to identify the noise existence by observing the block distance analysis. The two-level frequency band divisions preserve the information and provided the effective signal exploration. The decomposition modelling for down sampling is shown in figure 1.

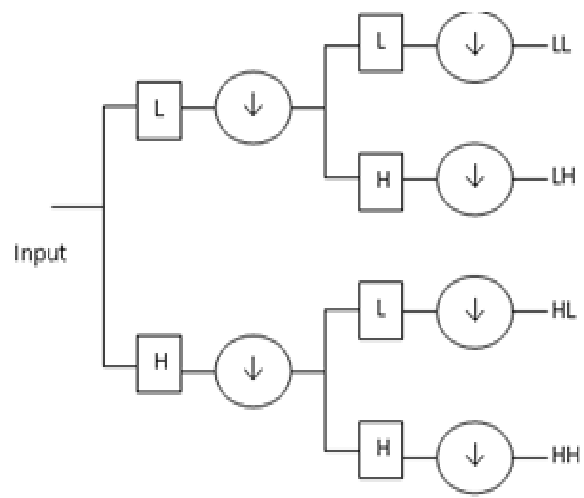

Fig. 1: DWT Decomposition (Down sampling Method)

The down sampling is applied to keep the feature data and to discard the disturbed signal. The frequency band observation can rectify the signal and neglect the noise signal part. The algorithmic specification for noise filtration and background noise removal is explained as:.

Table 3: Background Noise Suppression

```
Algorithm (SpeechSig)
/*SpeechSig is the raw input speech with different kind of inclusive noise element*/
{
Blocks=GetBlocks(SpeechSig,winsize)
[Split the signal into smaller blocks in the specification of time domain based partial analysis]
For i=1 to Blocks. Length
[Process the Speech Signal Blocks for Noise Reduction]
{
SpeechBlk=Blocks(i)
SpeechBlk =LowBandAnalysis(SpeechBlk)
[Apply Low Band Pass Fileration to generate the spectrum signal Block]
SpeechBlk =HighBandPass(SpeechBlk)
[Observe the signal under Base specification frequency drift for signal noise reduction]
[HCoff LCoff]=DWT(SpeechBlk)
[Apply second level decomposition to apply the differentiated frequency signal response]
SpeechBlk=SpeechBlock*LCoff
[Generate the rectified signal Block]
SpeechSignal=[SpeechSignal SpeechBlk]
[Combine the Signal Blocks]
        }
Return SpeechSignal
}
```

The method shows that the segmented signal blocks are processed under algorithm for generating the optimized signal result.

2) Spectral Subtraction for Instrumentation Noise Filteration

Speech capturing can be done using any mike or mobile device. The faulty device can include instrumentation noise within a speech signal. The spectral subtraction approach is applied to analyze the frequency difference and to identify the frequency range of effective signal. After observing these peaks of noise and speech signal in generated segment, the signal subtraction is applied to rectify the signal against instrumentation noise. The noise spectrum is analysed between two consecutive speech periods. The time-domain specific analysis is applied to observe the phase variation and the instantaneous magnitude map to improve the signal. For this observation time domain specific inverse discrete Fourier transformation is applied. This magnitude specific estimation can suppress the noise elements. The difference between the peak signal and the average frequency is obtained to apply the spectral subtraction method. The segment driven analysis is applied to recognize the disorder level. This band level observation is dynamic and change as the frame based frequency fluctuation occurs. The amplitude level analysis with maximum noise specification is observed to identify the probable change required to deduct the signal. The minimum and maximum limits are applied to keep the signal strength and suppress the noise elements in the signal.

3) LPC for Acoustic Turbulence Analysis

Acoustic feedback is the hearing aid based signal leakage occur because of instrumentation problems. As this specialized noise form generates the loop in the signal and become stronger with each trip. In this noise, the magnitude of the particular loop segment is increased up to a high extent. Some predictive an adaptive analysis must remove this noise. Because of this probabilistic LPC (Linear Predictive Coding) is suggested here to improve the speech quality against this noise. A probabilistic modelling is required under LPC method to speech block analysis and synthesis. The person vocal tract is here required to compare with the pitch of loop area with the incremental frequency and duration observation. This time bound magnitude observation must identify the probable change in the speech signal and its estimation of expected turbulence. The linear IIR filter must apply the transfer function specification to improve the signal form. Two pole turbulence observation under LPC method is shown in figure 2.

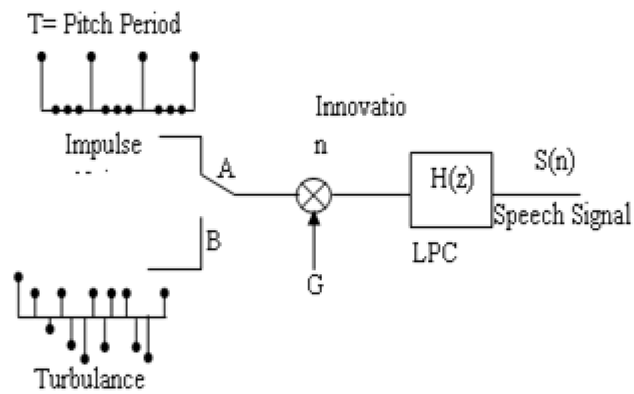

Fig. 2: LPC Method for Turbulance

The conditional cover based probabilistic estimation is induced to perform LPC filteration. The time domain specific dual pole observation is applied to identify the turbulence segments. The fundamental frequency, Particular Pitch Period and the Pitch increase ration under time are observed collectively to identify the problem region. The block specific predictor is applied to differentiate the valid and noisy signals.

4) Gaussian Distance Scored Recifitication

The impurities in the speech content regions or segments are generally more critical. These environment driven or echo disrupts the speech signal. The Gaussian weighted measure is applied to estimate the difference between the normal and the cross captured speech signal. For this Score evaluation, the speech signal is divided in smaller blocks. The time domain specific magnitude difference analysis is done between two signal blocks. The signal blocks with lesser variation in Gaussian score are considered as the effective signal blocks. The blocks with balanced Gaussian score with lesser frequency change and low turbulence are considered as the filtered speech signal. The rectified signal is processed for feature generation and speech recognition.

B. Signal Feature Generation

Once, the speech signal database is fully converted to normalized form, by removing different impurities and unbalancing elements. The next stage of work is to apply the recognition algorithm. Instead of working on signal data directly, a dynamic segmentation based statistical approach is defined at this stage. The dynamic segmentation is applied to extract the signal features. The blocks are generated over the signal between two consecutive peaks. Both the training set and testing set signals are transformed in this form. The segmentation algorithm is shown in the next subsection. After generation of these dynamic blocks, the structured and statistical measures are applied to extract the features from raw input data. The size of this dataset is based on the number of dynamic segments. Because of which, to reduce the dataset size and to obtain most meaningful data elements, a HMM integrated fuzzy rule formulation is applied. In this work, Two level HMM is defined to validate the dataset. To formulate these rules mathematically, the conditional restrictions are applied using fuzzy method. In this section, all the sub stages of generating the signal feature form are described with algorithmic specification.

1) Dynamic Segmentation

To extract the signal features, the primary requirement is to generate the dynamic blocks over the signal. Signal amplitude analysis is done in terms of average, max and min measures. The dynamic limits are applied to identify the frequent peaks. Two consecutive peaks are considered as the lower and upper bounds of dynamically generated blocks. The algorithmic process of this stage is shown in table 4.

Table 4: Dynamic Segmentation

```
FeatureExtract(SpeechSig)
/*SpeechSig is the filtered speech signal*/
{
1.        AbsSignal=SQare(SpeechSig)
ate absolute signal, square the SpeechSig]
2.        [Max Avg Min]=GetStatistics(AbsSignal)
[Generate the Signal Statistics]
3.        DThresh=Avg+(Max-Min)/3
 the Dynamic Threshold for Peak Generation]
4.        For (i =1 to AbsSignal.Length)
[Process the signal instances]
{
6.        If (AbsSignal(i)>DThresh)
[Identify the peak signal based on dynamic threshold value]
{
7.        SignalPeaks(pos)=i
pos++
[Store the signal peaks]
}
}
8.        For i=1 to SignalPeaks.Length
he Identified Peaks to Generate Dynamic Blocks]
{
9.DBlocks=SpeechSig(SignalPeaks(i),SignalPeak(i+1))
[Extract the Dynamic Blocks Based on Dynamically generated peaks]
}
10.       Return DBlocks
}
```

2) Structural Feature Generation

As the classification process is applied on the statistical and structural features generated on dynamically segmented blocks. The featured speech signal not only reduces the processing dataset but also provided the most effective features to improve the accuracy. The structural and statistical features are generated to transform the signal to numerical form and to process the information in discrete form. For structural features ICA is applied independently on each block and the aggregative statistical features are generated. In this section, all kind of aggregative features used to transform the speech signal are provided.

a) ICA: Structural Reformation

ICA (Independent Component Analysis) is used in this work for generation of structural features for each independent block. This is a level based feature generated by distance evaluation. A structural frame generated for weighted distance map is applied on each block for structural feature set generation. Let Blk is the vector represents transformed signal blocks and Frame (F) is the structural frame matrix of size mxm. Where, m is level of features. Then, the equation (1) is showing the generation of structural features

StrucFeatures(i)=Frame*SigBlk(i) (1)

The frame is the distance preserved weighted vector to explore the structural features of particular signal block. The column controlled features are identified by generating the signal distance. The pitch difference is considered as the weighted characterization and represented as ICA feature vector.

b) Block Mean

It is the aggregative statistical feature generated by taking the mean of frequency of magnitude of block features. The mean feature is able to transform the block to relative discrete value. Let SigBlk is the vector representing the transformed signal block. The block mean is applied on raw signal and the structural signal feature. Equation (2) is showing the generation of block mean for particular i^{th} signal block

$$\text{AggMean(i)} = \frac{\sum_{j=1}^{N} Mag(SigBlk(j))}{SigBlk.Length} \quad \text{and}$$

$$\text{AggSMean(i)} = \frac{\sum_{j=1}^{N} Mag(StrucFeatures(j))}{StrucFeatures.Length} \qquad (2)$$

Here, AggMean is the aggregative block mean feature set and AggSMean is the feature set generated on structural features.

c) Block Standard Deviation

Standard Deviation (SD) is another statistical feature generated for each signal and structural signal block. This feature is able to represent the amplitude variation within block. Here, equation (3) is showing the block standard deviation feature.

$$AggSD(i) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(SigBlk(i)_j - \overline{SigBlk(i)_j})^2} \text{and}$$

$$AggSSD(i) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(StrucFeatures(i)_j - \overline{StrucFeatures(i)_j})^2} \qquad (3)$$

Here, N is the length of the blocks.

d) Block Spatial Median

This statistical aggregative feature identifies the middle of probability distribution at block level. The adaptive frequency change within block can be separated using this feature. The frequency distribution can be measured using Aggregative spatial median feature. The feature is generated for each signal block and structural feature block. Let E is the expected distance observation and a is the centralized estimator then equation (4) is showing generated spatial median feature..

AggSM(i)= E‖(SigBlk(i)-a‖ and

AggSSM(i)=E‖($StrucFeatures$ (i)-a‖ (4)

e) Block MSE

The error observation on a signal block can be predicted using MSE (Mean Square Error). MSE is defined as the estimated difference between the actual and predicted signal values. MSE is evaluated for each signal block and structured feature blocks and shown in equation(5).

AggMSE(i)=$\frac{1}{N}\sum_{i=1}^{n}(SigBlk(i) - \widehat{SigBlk}(\imath))^2$ and

AggSMSE(i)=$\frac{1}{N}\sum_{i=1}^{n}(StrucFeatures(i) - \widehat{StrucFeatures}(\imath))^2$ (5)

Here, N is size of block signal

f) RMSD

Another quality specific statistical feature considered in this work is RMSD (Root Mean Square Deviation). This feature identifies the error deviation in aggregative form between the actual and predicted values. Equation (6) is showing the generated RMSD feature for signal and structural signal blocks.

$$AggRMSD(i) = \sqrt{\frac{\sum_{i=1}^{n}(SigBlk(i) - \widehat{SigBlk}(i))^2}{N}} \text{and}$$

$$AggSRMSD(i) = \sqrt{\frac{\sum_{i=1}^{n}(StrucFeatures\ (i) - \widehat{StrucFeatures}(i))^2}{N}} \quad (6)$$

3) Fuzzy Integrated HMM

The generated structural and statistical feature blocks are processed under fuzzy integrated HMM (Hidden Markov Model) method to assign the feature and block weights. The fuzzy is here considered to generate the rules to observe individual blocks and describe them based on content strength. Whereas, HMM is the predictive analysis method in which the mutual analysis between the blocks is done to identify the weight of a block relative to other block. The content level evaluation is performed in this stage for strength evaluation of each feature block. The weight generation and relational observation are collectively able to transform the features in more precise form. The controlled variation and in discrete form represent able to define the feature value in more effective form. The FHMM is applied on both the aggregative statistical featurs and agreegative structural-statistical features.

Speech Recognition: SVM

In the final stage of proposed hybrid model, SVM is applied on weighted fuzzy-predictive features to recognize the speech signal. The SVM is first trained on training feature set to generate the rules to recognize the speech signal effectively. The linear class separator is identified based on featured evaluation. Later, the model is applied on testing set to perform the recognition. The classification process applied on training set is shown in figure 3.

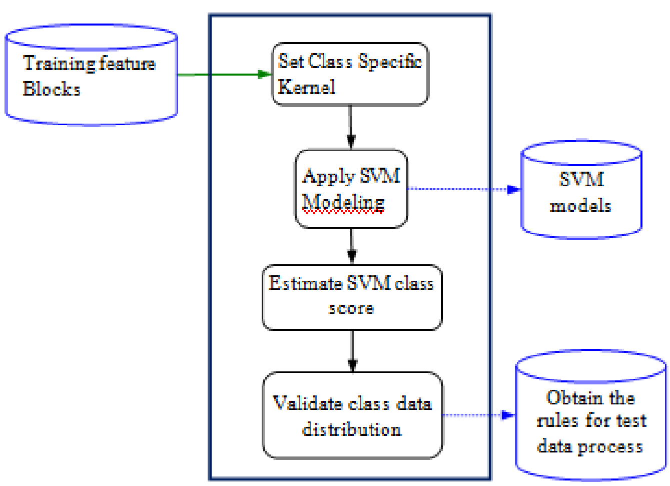

Fig. 3: Training SVM Classifier

The figure demonstrates the classification modelling based on score estimation under SVM kernel specification. The feature modelling is defined to estimate the class score. Once the model is formed, the rules are applied on test data to recognize the speech signal. The test data processing based on defined classifier is shown in figure 4.

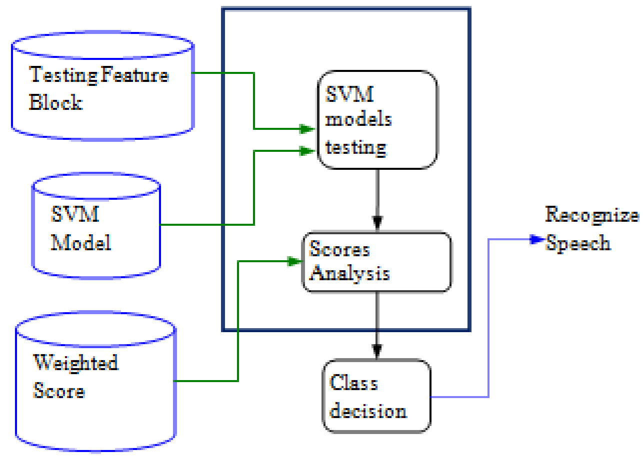

Fig.4 : Testing Speech Feature Data Processing

The figure 4 shows the detailed feature test data processing using designed SVM model. The test data block with SVM model are observed by considering the weight score taken from testing set. The score analysis is applied and the speech class is identified based on score observation.

The complete process model is applied to different Hindi speech sample sets. These sample sets are having words and the sentence data form. The description of these sample sets and obtained results are defined in next experimentation section.

3. Experimentation & Results

The experimentation is applied for different speech forms: character speech dataset, word dataset and the sentence dataset. These datasets are acquired from real time and collected from secondary web sources. Different samplesets are taken for each dataset to obtain a clear and true analysis. The noise element is also included to generate the real time experimentation. The comparative observations are taken against Back propagation neural network and PCA methods. The experimental results with training and testing set specification is defined in this section.

A. English Character Dataset

To verify the model significance, the experimentation is applied on three different sample sets with different number of training and testing instances. Table 5 is having the description of sample set characterizations.

The comparative observations of proposed hybrid model are taken against Back propagation neural network and PCA approach. Table 6 shows the number of instances corrected identified by each of the classifier.

Table 6 also verifies that the proposed model has identified more number of test instances accurately as compared to PCA and BPNN classifiers. The derivation of the results in terms of accuracy ratio is shown in figure 5.

Table 5: English Character Sample sets

Sample Set	Feature	Value
S1	Number of Training Instances	26x5
	Number of Training Speakers	5 Male, 5 Female
	Number of Instances of Each Character (Training)	5
	Number of Test Instances	30
	Number of Testing Speakers	2 Male, 2 Females
	Noise Included (Externally)	No
S2	Number of Training Instances	26x10
	Number of Training Speakers	5 Male, 5 Female
	Number of Instances of Each Character (Training)	10
	Number of Test Instances	50
	Number of Testing Speakers	2 Male, 2 Females
	Noise Included (Externally)	No
S3	Number of Training Instances	26x10
	Number of Training Speakers	5 male, 5 Female
	Number of Instances of Each Character (Training)	10
	Number of Test Instances	50
	Number of Testing Speakers	2 Male, 2 Females
	Noise Included (Externally)	Yes

Table 6: Recognition Rate Analysis

	Test Set Size	BPNN	PCA	Proposed
S1	30	19	20	24
S2	50	35	32	42
S3	50	28	30	39

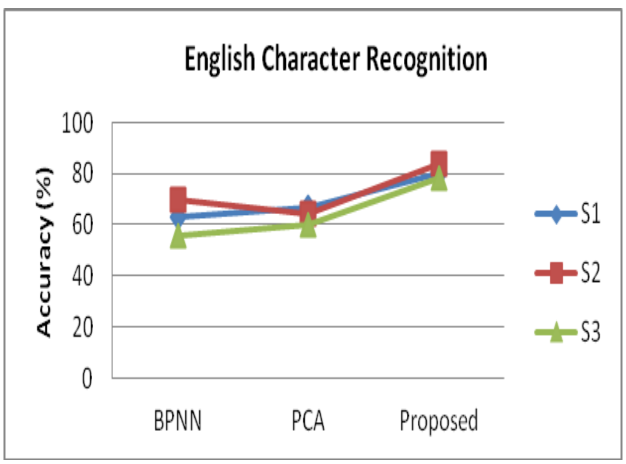

Fig. 5: English Character Accuracy Analysis

B. Hindi Sentence Recognition

To apply the Hindi sentence recognition, the data are collected from external web reference. The database name is Linguistic Data Consortium for Indian Languages, and it is extracted from http://www.ldcil.org/ resourcesSpeechCorpHindi.aspx. The database is having 30 speech signal in WAV format. Three samplesets are generated from this dataset and experimentation is applied against BPNN and PCA approaches. The first sample set is of 30 instances and testing set was of 10 instances. In the second sample set, the training set was of 30 instances and the test set considered with 20 instances. The third sampleset considered here is with same number of instances but included the additional noise in test dataset. The comparative results of correctly identified instances are shown in table 7.

Table 7: Results of Hindi Sentence Sample sets

	Test Set Size	BPNN	PCA	Proposed
S1	10	5	6	7
S2	20	11	12	15
S3	20	8	8	12

The significant improvement in the accuracy rate can be observed in both noisy and noise-free speech signal. The comparative accuracy rate observations can be seen in figure 6

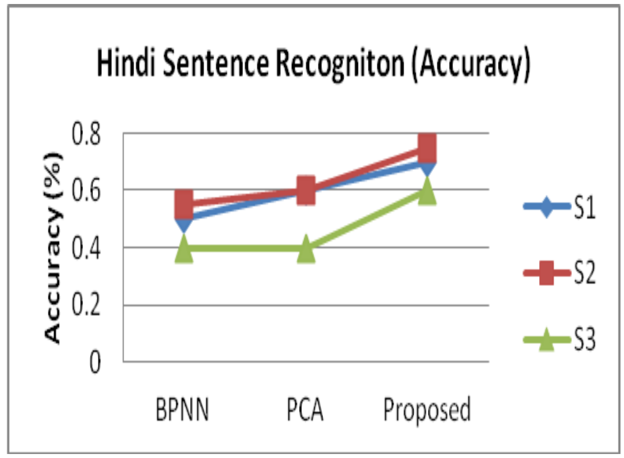

Fig. 6: Hindi Sentence Recognition (Accuracy)

The experimentation applied on Hindi sentence sample set shows that the proposed method has provided the comparatively high recognition rate. In case of the noisy sample set, the accuracy improvement is high. All the results on different sample sets showed that the proposed predictive and probabilistic model has achieved the effectively accurate results on different speech forms.

4. Conclusion

The paper has presented a hybrid hindi/English speech recognition model for real time captured speech. The captured speech is affected from background noise, instrumentation noise and the turbulence. A three stage model is presented in this work to improve the accuracy of speech recognition. In the earlier phase, the decomposition and predictive methods are applied to observe the speech signal and to remove the signal impurities. The improved speech signal was later on processed by structural and statistical features generated on dynamic signal segments. The FHMM method is applied to transform the aggregative features to discrete form. Finally, weighted SVM classifier is applied to recognize the speech signal. The model is applied on multiple samplesets of hindi and English speech instances. The comparative results shows that the proposed model has improved the accuracy of speech recognition effectively.

References

[1] D. Ververidis and C. Kotropoulos (2004), ”Automatic speech classification to five emotional states based on gender information”, Signal Processing Conference, 2004 12th European, Vol, Issue , pp 341-344, 2004

[2] S. Z. Guo; L. J. Yu; G. Y. Kang (2010), ”Band Energy Based GMM Speech with Noise Classification Algorithm”, Pervasive Computing Signal Processing and Applications (PCSPA), 2010 First International Conference on, Vol, Issue , pp 541-544, 2010

[3] C. Shao; M. Bouchard (2003), ”Efficient classification of noisy speech using neural networks”, Signal Processing and Its Applications, 2003. Proceedings. Seventh International Symposium on, Vol1, Issue , pp 357-360 vol.1, 2003

[4] S. Bansal; A. Dev (2015), "Emotional Hindi speech: Feature extraction and classification", Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on, Vol, Issue , pp 1865-1868, 2015

[5] Y. Zeng; Y. Zhang (2007), "Robust Children and Adults Speech Classification", Fuzzy Systems and Knowledge Discovery, 2007. FSKD 2007. Fourth International Conference on, Vol4, Issue , pp 721-725, 2007

[6] M. F. Kaleem; B. Ghoraani; A. Guergachi; S. Krishnan (2011), "Telephone-quality pathological speech classification using empirical mode decomposition", Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE, Vol, Issue , pp 7095-7098, 2011

[7] M. Murugappan; N. Q. I. Baharuddin; S. Jerritta (2012), "DWT and MFCC based human emotional speech classification using LDA", Biomedical Engineering (ICoBE), 2012 International Conference on, Vol, Issue , pp 203-206, 2012

[8] R. Tadeusiewicz; A. Izworski; W. Wszolek; T. Wszolek (1999), "Processing and classification of deformed speech using neural networks", [Engineering in Medicine and Biology, 1999. 21st Annual Conference and the 1999 Annual Fall Meetring of the Biomedical Engineering Society] BMES/EMBS Conference, 1999. Proceedings of the First Joint, Vol2, Issue , pp 927 vol.2-, 1999

[9] T. Ghiselli-Crippa; A. El-Jaroudi (1991), "A fast neural net training algorithm and its application to voiced-unvoiced-silence classification of speech", Acoustics, Speech, and Signal Processing, 1991. ICASSP-91., 1991 International Conference on, Vol, Issue , pp 441-444 vol.1, 1991

[10] D. Ververidis; C. Kotropoulos (2010), "Emotional speech classification using Gaussian mixture models", Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on, Vol, Issue , pp 2871-2874 Vol. 3, 2005

[11] R. Cai (2010), "A Modified Multi-Feature Voiced/Unvoiced Speech Classification Method", Power Electronics and Design (APED), 2010 Asia-Pacific Conference on, Vol, Issue , pp 68-71, 2010

[12] M. Srinivas; D. Roy; C. K. Mohan (2014), "Learning sparse dictionaries for music and speech classification", Digital Signal Processing (DSP), 2014 19th International Conference on, Vol, Issue , pp 673-675, 2014

[13] M. Charfuelan; G. J. Kruijff (2013), "Classification of speech under stress and cognitive load in USAR operations", Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, Vol, Issue , pp 508-512, 2013

[14] Zhongzhe Xiao; E. Dellandrea; Weibei Dou; Liming Chen (2005), "Features extraction and selection for emotional speech classification", Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on, Vol, Issue , pp 411-416, 2005

[15] Guojun Zhou; J. H. L. Hansen; J. F. Kaiser(1998), "Classification of speech under stress based on features derived from the nonlinear Teager energy operator", Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on, Vol1, Issue , pp 549-552 vol.1, 1998

[16] G. Zhou; J. H. L. Hansen; J. F. Kaiser (2001), "Nonlinear feature based classification of speech under stress", IEEE Transactions on Speech and Audio Processing, Vol9, Issue 3, pp 201-216, 2001

[17] T. Ghiselli-Crippa; A. El-Jaroudi (1991), "Voiced-unvoiced-silence classification of speech using neural nets", Neural Networks, 1991., IJCNN-91-Seattle International Joint Conference on, Volii, Issue , pp 851-856 vol.2, 1991

[18] Mahdhaoui; M. Chetouani (2010), "Emotional Speech Classification Based on Multi View Characterization", Pattern Recognition (ICPR), 2010 20th International Conference on, Vol, Issue , pp 4488-4491, 2010

[19] S. Ramamohan; S. Dandapat (2006), "Sinusoidal model-based analysis and classification of stressed speech", IEEE Transactions on Audio, Speech, and Language Processing, Vol14, Issue 3, pp 737-746, 2006

[20] L. He; M. Lech; N. C. Maddage; N. B. Allen (2009), "Time-frequency feature extraction from spectrograms and wavelet packets with application to automatic stress and emotion classification in speech", Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on, Vol, Issue , pp 1-5, 2009.

[21] [7] Priyanka Bansal and Syed Akhtar Imam, "Automated Speaker Recognition Methods: A Critical Review" , International Journal of Advance Research in Science and Engineering, vol. 6 ,issue 10, pp. 519 - 524, Oct 2017.

[22] Priyanka Bansal and Syed Akhtar Imam, "Speaker recognition using MFCC, shifted MFCC with vector quantization and fuzzy", International Conference on Soft Computing Techniques and Implementations, October 2015.

[23] S.V.Manikanthan and T.Padmapriya "Recent Trends In M2m Communications In 4g Networks And Evolution Towards 5g", International Journal of Pure and Applied Mathematics, ISSN NO:1314-3395, Vol-115, Issue -8, Sep 2017.

[24] T. Padmapriya and V. Saminadan, "Inter-cell Load Balancing technique for multi-class traffic in MIMO-LTE-A Networks", International Journal of Electrical, Electronics and Data Communication (IJEEDC), ISSN: 2320- 2084, vol.3, no.8, pp. 22-26, Aug 2015.

[25] S Nazeer Hussain, K Hari Kishore "Computational Optimization of Placement and Routing using Genetic Algorithm" Indian Journal of Science and Technology, ISSN No: 0974-6846, Vol No.9, Issue No.47, page: 1-4, December 2016.