Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44

ORIGINAL ARTICLE

A Probabilistic Feature Based SVM Model for English Speech Recognition

Priyanka Bansal[1] and Syed Akhtar Imam[2]

[1]Research Scholar, Deptt. of Electronics & Communication Engineering, Jamia Millia Islamia, New Delhi, India
[2]Associate Professor, Deptt. of Electronics & Communication Engineering, Jamia Millia Islamia, New Delhi, India

Abstract. In this paper, a multi-phase hybrid model is presented to improve the effectiveness of Speech Recognition. In the first stage of the model, the speech rectification is performed against different real time problems. Then hybridization of four algorithms is adapted to cover different noise and turbulence problems. The decomposition based filtration is then applied to remove background noise and peak observed spectral subtraction is applied to rectify the signal against instrumentation noise. The probabilistic predictor is applied to remove acoustic turbulence and Gaussian weighted map is applied to resolve the crosstalk problem. After recovering the signal from these real time problems, a dynamic block segmentation Fuzzy-HMM based structural-statistical method is defined for feature generation. The stage has combined the ICA, Fuzzy and HMM modeling to generate effective block feature set. Finally, the fuzzy weight based SVM classifier is applied to perform the speech signal classification. The results and observation shows that the proposed hybrid model has improved the maturity level and provided higher accuracy rate.

Keywords: SVM, Structured, Turbulence, Hybrid, Speech Recognition.

1. Introduction

The speech signal is the first communication medium that can provide direct interaction between individuals in verbal form. A distinctive form of communication is used to acquire the knowledge, including question-answering, discussions, lectures, interviews, etc. To provide the human-computer interaction, the role of speech processing exists in different fields. Different speech processing technologies and tools are available nowadays. Major application areas of speech processing include speech recognition, text to speech conversion, speech to text conversion, language transformation, etc. Speech Recognition and classification are having various applications and domains with associated criticalities. It is one of the oldest biometric information-processing methods, still not much popular because of its criticalities. Even though, from the last few decades, researchers have done a lot of work in different speech applications. The accuracy and optimization are the major vectors contributed by the researchers. Some of the associated contributions of earlier researchers are discussed in this section.

Most of the researchers improved the classification and recognition method based on feature extraction. One such feature form includes structural features defined by [1]. Author used the block segmented analysis to generate the feature points from successive frames. Harmonic properties [2] are also utilized to classify the voice and non-voice audios. Vocal features [3] are being used to improve the classification model. The estimation generally applied to cross-sectional area specification with feature aspect including tract length, frequency, mass measure, etc.

Different frameworks and architectures were proposed by the researchers to improve the classification results. One such improvement to neural network was proposed by [4]. Author improved the neural network architecture by using an integrated equalizer. Another classifier applied by [5] was n-gram based Naive Bays

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

method. The grace vector observation was provided by the author. Low level feature encoding method in the form of code book search [6] was applied to improve the classification over rough speech. A weighted features based multi-channel [7] evaluation method is provided to improve the speech classification. Author considered the probabilistic risk analysis using prosodic, gloated and spectral features. A cross validation oriented sequential floating [8] forward selection method under probabilistic measure was defined for classification of isolated words. Author [9] generated the speech atoms using a matrix de-convolution method which later on learned under the least square regression method for improving the classification rate.

2. Research methodology

In this paper, a robust speech recognition method is proposed for English characters, words, sentences. To achieve this robustness and accuracy, the optimization and hybridization are applied at each stage of speech processing. The complete model is divided into three main stages as described in section 1. At the earlier stage, the impurities over the speech signal are rectified and transformation of speech to normalized form is done. To perform this rectification, a hybrid multi-method approach is applied to the specification of associated impurities. This stage can provide the solution against acoustic noise, background noise and the instrumentation noise. To apply the noise reduction and removal a connected segmented method is applied. The hybridization of this stage is done under DWT, LPC, Band Pass Filtration and Gaussian Distance Scoring method. These methods are applied in series by applying the block segmentation over the speech signal. The normalized filtered speech will be obtained as the final outcome of this stage. In the second stage, the normalized speech signal is processed for signal feature generation.

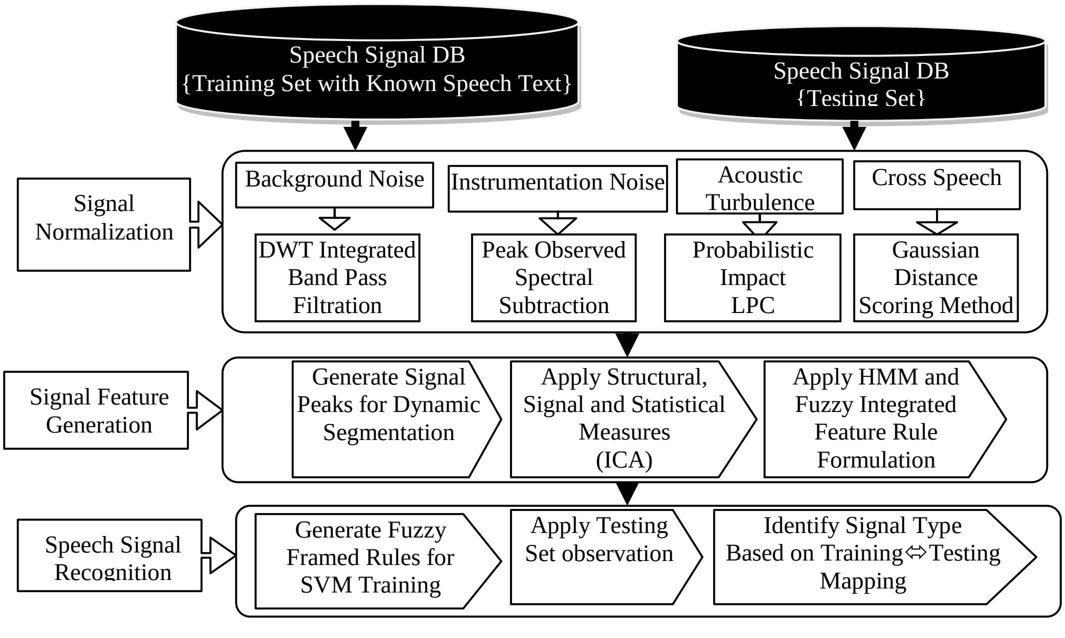

Figure 1. Proposed Speech Recognition Model

In this stage, a block segmented HMM and fuzzy method is applied to generate features. In the final stage of this work model, the rule formulation is obtained based on the prior analysis applied by SVM model and applied the rules on testing set. The experimentation on different training and testing sets are discussed in the next section. The model derived from the work for speech recognition is shown in figure 1. In this section, each of the work stages with algorithm is described.

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

2.1 Signal Normalization

Speech can be extracted through different devices and in different environment. As the source of capturing is not specific, the captured speech can have number of impurities. In this section, the noise rectification from the speech signal and its transition to the standard robust form is formulated.

2.1.1 Background Noise Reduction: Decomposition Approach

At the first level, the signal transfer function is applied using low pass filtration. Difference equation based on the sample period is defined for signal improvement. The filter equation is shown here below

Signal(nS) = 2SignalY(nS -ST) - SignalY(nS – 2S) + SignalX(nS)-2SignalX(nS-6S)+SignalX(nS- 12S) *-(1)*

The filtered signal subtraction is applied by observing the dynamic segmentation window with cutoff frequency to improve the signal effectiveness. The difference equation analysis is applied for signal quality improvement and removing the disturbed aspect. The cutoff frequency based signal improvement is given under high pass filtration shown in equation (2)

Signal(nS)=32SignalX(nS-16S)-{SignalY(nS-S)+SignalX(nS)-SignalX(nS-32S)} - *(2)*

After applying the band pass filtration, the differentiated signal form is obtained. The baseline specific observation is applied to observe the interference noise. The lower limit based secondary deduction is applied to high level inference noise. The equation for next level deduction is shown in equation (3)

Signal(nS) = (1/8 S) [-SignalX(nS - 2 S) – 2SignalX(nS - S)+ 2SignalX(nS + S) +SignalX(nS+ 2S)] - *(3)*

After generating the two levels down-sampled signal, decomposition (DWT) is applied to transform the signal in normalized form.

2.1.2 Instrumentation, Noise Reduction: Spectral Subtraction

In this sub-stage, the spectral subtraction method is applied to observe the spectrum magnitude as the peak value and based on it, valid speech frequency is identified. To apply the spectral subtraction, the frequency driven rectification is applied. Where the speech signal frequency is defined by $\left|\overline{X(e^{j\omega})}\right|$ as average frequency and $\mu(e^{j\omega})$ is the peak signal value. An estimators $\hat{S}(e^{j\omega})$ is observed and applied to rectify the signal and reduce the noise element by $\mu(e^{j\omega})$. When the noise inclusive speech signal lesser than the lower peak, it is the incorrect removal of speech information. The identification of silent speech segment can be obtained by analyzing the difference between N and $\mu(e^{j\theta_n})$. The disorder level narrow band observation can generate the magnitude instances. The spectrum energy analysis is applied to retain the information. The sample spectral observation applied for signal rectification is shown in equation (4)

$$\hat{S}(e^{j\omega}) = \begin{cases} \hat{S}\left(e^{j\omega}\right) \mathrm{T} \geq -12\mathrm{dB} \\ \mathrm{c}\,\mathrm{X}\left(e^{j\omega}\right) \mathrm{T} \leq -12\mathrm{dB} \end{cases} \qquad - (4)$$

2.1.3 Acoustic Turbulence Analysis: Probabilistic LPC

At this stage, the probabilistic modeling is applied to observe the turbulence. The probabilistic estimation identified the requirement of LPC filtration. The analysis of the signal is applied with dual pole inclusion with time domain observation. The exclusive excitation function is applied to identify the valid speech region and the abnormal sound reason. The linear prediction method is here applied as the output filter given by equation (5)

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

$$\hat{S}(n) = -\sum_{k=1}^{p} a_p(k)\, s(n-k) \qquad \text{- (5)}$$

Here s(n) is the block sample of size n, k is the particular signal instance for which the predictor rectifier is applied and p is the number of blocks. The predictive turbulence is represented by us ($\hat{n}$). The error observation can be obtained by getting the differentated value between the actual impulse signal and predicted signal and this error is shown in equation (6)

$$e(n) = s(n) - s(\hat{n}) \qquad \text{- (6)}$$

The pole parameter based observation is applied to differentiate the value under constraint specification. The pole parameters ($a_p(k)$) are observed by differentiating the parameter aspects with linear equation specification. The block aggregative observation using the linear predictive method is given in equation (7)

$$\sum_{k=1}^{p} a_p\,(k) r_{ss}(m-k) = -r_{ss(m)} \qquad \text{- (7)}$$

Here, r_{ss} represents the correlation analysis is applied between consecutive blocks. The equation of this sequence block predictive analysis is given in equation (8)

$$r_{ss(m)} = \sum_{n=0}^{N} s(n)s(n+m) \qquad \text{- (8)}$$

The predictive observations are applied with specification of correlation matrix with model parameters specification. The predictor coefficients are obtained based on the probabilistic estimation at the early stage of this method. Finally, the polynomial map is applied to generate the filtered speech signal. The coefficient vectors defined here are listed as A=[1 A(2) ... A(N+1)], of an Nth order forward linear predictor.
From this observation, the generated signal form with predictive observation is given as equation (9) based on coefficient vectors

$$Xp(n) = -A(2)*X(n-1) - A(3)*X(n-2) - ... - A(N+1)*X(n-N) \qquad \text{- (9)}$$

The difference between the signal and the predictive signal form is -the signal error.

$$err(n) = X(n) - Xp(n) \qquad \text{- (10)}$$

Here, X is matrix for Signal Column
N is order of Polynomial
A is coefficient of polynomials

This whole process is repeated in sequence block pairs recursively till the rectified signal is not obtained from the work.

2.1.4 Cross Speech Impurities: Gaussian Distance Scoring Method

After removing the clear background noise, device and instrumentation noise problems most of the non-voice impurities are removed from the speech signal. However, sometimes the impurities exist in other voice elements. These voice disturbances can occur because of echo problem or accidental capturing of another person talks. To identify these cross voice problems, a level or weight goaded method is required. This weight driven method will distinguish the voice primary user voice element throughout the speech signal under pitch and amplitude analysis and assign it as a higher weight value. Other detected voice elements will be assigned by lower weights. To assign these weights, a Gaussian filter based weight function will be applied. In this work, the cross voice estimation in speech block will be identified as the difference estimation between the primary voice block and the other cross captured voice block. The primary voice captured block is identified as the clear speech block with no turbulence or frequency change. Let i is the primary user voice and the j is the expected cross voice signal based speech block, the weight function is shown in equation (11)

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

$$Wij \ \alpha \ \exp\left(-\frac{||SignalBlk_i - SignalBlk_j||_2^2}{h^2}\right), \ i \neq j \qquad - (11)$$

Here, i and j are two signal blocks
i is the block index of primary user voice
j is the any random block with expected cross talk
h is the spatial trend for block selection method.
And also wii =0 and $\sum_{j=1}^{N} w_{ij} = 1$

The block weight specific observations will be applied throughout the speech signal blocks and the rectification of cross block will be done by deducting the magnitude variation. Based on these collective featured filters, the overall optimized signal form will be obtained. This filtered signal form will be processed under feature generation stage defined in the next subsection.

2.2 Signal Feature Generation

To generate signal set features, at first the signal peaks are identified based on the intensity observation with dynamic threshold mapping. The difference between two peaks is taken as the segmented adaptive feature. The feature generation process is shown in sub-sections of this section.

2.2.1 Dynamic Segmentation

The filtration stage is applied to both the training set and testing set data to rectify the signal against different integrated impurities. Now to generate the structural and statistical features from the signal data, a dynamic segmentation based analysis and rule specific validations are applied. In this subsection, the algorithmic specification is defined to divide the signal in dynamic blocks. To generate these dynamic blocks, the signal amplitude is evaluated under average, min and max functions. Based on these aggregative values, a dynamic threshold is generated to generate the segments peaks. These peaks are the upper and lower bounds of generated segmented blocks. The stage has transformed each of the signals to dynamic splitted blocks.

2.2.2 Structural Feature Generation

This stage is having the significant importance to transform the signal in the numerical data form. Some of the structural and statistical measures are applied to generate these features. Instead of applying the feature extraction measures on complete signal, the transformation of a signal to block form is done. The structural feature is evaluated for each of the dynamic segmented blocks. Number of blocks of the signal defines the size of the numeric database. These numeric features will be evaluated for normal signal blocks and structure specified signal blocks. To highlight the structural aspect, ICA method is applied separately for each block. ICA is a structural matrix based method that can explore the structural characterization up to high extents. The ICA integration for feature generation is described in this section. Once the structural block set is generated, the statistical measures are applied on both normal signal and to ICA based structured feature signals. The structural and statistical measures considered in this work are also described in this section.

2.2.2.1 ICA: Structural Reformation

To generate the structural features ICA (Independent Component Analysis) is applied for each dynamic block in an integrated form. The ICA is a level adaptive method which can apply the distance level observation for individual element to extract effective features. The basic equation for the dynamic block structure formation is given in equation (12)

$$IFeatures(i)=A*DBlocks(i) \qquad - (12)$$

Where, A is Structure framed matrix of size mxm
m is level of features

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

The method is based on the weight matrix specification with the flexibility to explore the structural information within a signal block. These signal features can be controlled with column specification for signal vector. The distance based mapping can identify the signal difference. The structure matrix can identify the difference and provide more significant signal feature exploration. These constraints driven segmented signals represents the eventual information with steady peaks within blocks. The existence of pitch difference can be observed via ICA weight characterization.

2.2.2.2 Block Mean

Each of the dynamic block is estimated for magnitude mean value observation to generate the equalize block data. This mean featured data is able to transform the signal in square waves. The formulation of mean data generation is shown in equation (13). Here, the estimation is defined specifically for ith signal block and structural feature block

MeanData(i) = Sum(Mag(DBlocks(i)))/ Length(DBlocks(i)) and

IMeanData(i) = Sum(Mag(IFeatures (i)))/ Length(IFeatures (i)) - *(13)*

2.2.2.3 Block Standard Deviation

Another statistical measure applied to generate the block amplitude value. The amplitude variation within a block is observed by this feature for both normal signal block and structural signal block. The standard deviation evaluation is shown equation (14)

$$SD(i) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(DBlocks(i)_j - \overline{DBlock(i)_j})^2} \quad \text{and}$$

$$ISD(i) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(IFeatures\ (i)_j - \overline{IFeatures\ (i)_j})^2} \quad - (14)$$

Here, N is the length of the signal block.

2.2.2.4 Block Spatial Median

The median is estimated to observe the middle of probability distribution within the block. It is the breakdown point to identify the frequency change within the block. It can also identify the distribution type within the block. Evaluating block median is shown in equation (15) for both normal and featured signal block. Here a distance driven spatial median is estimated to apply the expected distance observation on normed vector space.

SM(i)= E||(DBlocks(i)-a|| and ISM(i)= E||($IFeatures$ (i)-a|| - *(15)*

Here, a is the centralized estimator for distance level observation

2.2.2.5 Block MSE

MSE (Mean Square Error) is the quality of measure to map the block data with some sample predictor to identify the error observation. The vector specific prediction is applied to evaluate this difference between the actual and predicted values and the average of this difference is MSE. The error estimator for signal block and featured signal block is shown in equation (16)

$$\text{MSE(i)}=\frac{1}{N}\sum_{i=1}^{n}(DBlocks(i) - \widehat{DBlocks}(\imath))^2 \quad \text{and}$$

$$\text{IMSE(i)}=\frac{1}{N}\sum_{i=1}^{n}(IFeatures\ (i) - \widehat{IFeatures}(\imath))^2 \quad - (16)$$

Here, N is size of block signal.

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

2.2.2.6 RMSD

RMSD (Root Mean Square Deviation) is the prediction error method based on the standard deviation modelling to identify the error under distance observation. The aggregative magnitude of predicted and actual data is analysed to predict the error. The RMSD estimation is shown here in equation (17)

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

$$IRMSD(i) = \sqrt{\frac{\sum_{i=1}^{n}(IFeatures\ (i) - \widehat{IFeatures}\ (i))^2}{N}} \quad - \quad (17)$$

2.2.3 Fuzzy Integrated HMM

FHMM (Fuzzy Integrated Hidden Markov Model) is the hybridization of fuzzy system within the Hidden Markov model. The fuzzy where applied to the individual feature parameter obtained from individual speech block analysis and ICA based speech block analysis. This rule formation can identify the strength of each block so that the realistic pruning and filtration will be done on the featured dataset. The number of attributes after this feature generation process is so vast, because of which a predictive rule specific analysis had to generate the effective content features. After generating the weights on individual data values using fuzzy logic, the relational observation is provided by HMM model. HMM model is a finite state machine which observes the discrete sequence and identifies the transition probability based changes. The probabilistic variation from the stage can be observed to identify the transited state. In this work, the normal speech block is considered as the initial stage and the IFeatured (ICA Featured) speech block is the transited speech block on which the state probability estimation will be done. The notation of hmm is given by

λ=(Feature(DBlocks),Features(IFeature),Π)

Where Π is the state probability

At the initial stage of this integrated method, the fuzzy rule formulation is applied on individual attributes. This rule formulation is dynamic, and identify the particular data strength in comparison with complete attribute data. The fuzzy assisted weight assignment is shown in table 5.

Table 1. Fuzzy Assisted Weight Assignment

```
Algorithm(Blocks) /* Blocks is Transformed Featureset*/
{
1. For i=1 to Blocks.Features /*Process Block Features*/
   {2.     For    j=1    to    Blocks(i).Instances/*Block
   Instance Evaluation*/
   {3.     if((FuzzyS(Blocks(j,i))=High             or
   FuzzyS(Blocks(j,i))                      =Medium)And
   FuzzyD(Blocks(j,i))=High)/*Fuzzy analysis*/
   {4.     Blocks(j,i).Weight=High
   }
   Else
   {5.     Blocks(j,i).Weight=Low
   }
6. c=Count(Blocks(:,i),Low)/*Count Low Weight Features*/
7. if    c<..10*Blocks(:,i).Length/*Discard    Low-weight
   Values*/
   {8.     Blocks(:,i).Discard=True    }     }
Return Blocks }
```

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

Here table 1 is showing the algorithmic model for weight assignment using fuzzy logic. The fuzzy rules are here defined with the static data value specified and with dynamic measures. To generate the dynamic fuzzy rules, the max, min and average data values in the particular block data is observed to form the rule. The fuzzy is here based on the amplitude value analysis to represent the existence of non-silence within the block. If the particular data block is not having the minimum number of speech instances, the block will be neglected as the speech feature. The the composite fuzzy rules are defined using static and dynamic measures to improve the reliability in the pruning process. After applying the high level fuzzy rule based filtration, the low level feature rectification is provided by HMM method. The associated featured analysis is applied in this stage. The HMM integrated algorithm for feature selection and pruning is defined in table 2.

Table 2. HMM Based Feature Pruning

```
Algorithm(Blocks) /*Transformed Speech Featured Blocks*/
{
1.  Set the two main feature class : Accepted and
   Rejected Features
2. For i=1 to Blocks.Length /*Process Feature Blocks*/
   {3.     Generate the centroid from the feature data
   analysis for two defined classes
4. Process and map the data from feature class based on
   group level analysis under distance parameter
5. Observe the block data to identify the feature class
6. Analyze the data frequency, If the frequency is
   lower, then eliminate the feature block
7. Predictive Analysis for rule formulation
8. Apply Level 1 HMM to identify associated features
9. if (FeatureClassData.Count<AvgData)
   [Remove low frequency data from feature data]
   {10.    Blocks(:,i).Discard=True
   }
11.  Apply level2 HMM to Analyze data frequency in
   feature block
12.  Apply the associated data analysis and generate
   the frequency count
13.  if (FeatureClassData.Frequency<AvgData)
   [Remove the low frequency block from feature data)
   {14.    Blocks(:,i).Discard=True
   }  }}
```

Here table 2 is showing the algorithm defined to prune the feature attribute with lesser frequency and associated frequency. The two level Markov model is applied to observe the class specific data existence. If the data frequency is lesser than average expected frequency, then remove the data block.

2.3 Speech Recognition: SVM

SVM (Support Vector Machine) is the linear classifier applied over the featured block data to identify the class of data points. The feature data space is observed in high dimensional space with the specified separator rule. At first, the training feature pruned speech data is processed to generate the rule. This rule actually

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

identifies the linear class separator on the dataset. Later, these rules are applied on testing data to identify the class of test speech instances. The data point mapping from the input space is applied through a function to obtain the data class. The SVM based training process is applied to set the data margin under class specification. The complete process model is applied to different English 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 proposed multi-featured SVM trained model is applied on different English speech datasets to perform the speech signal recognition. The first experimentation is here applied on English character dataset collected from real time environment. The speech signals are collected for 10 speakers (5 Male, 5 Females). Each speaker generated 10 instances of the same character. Different training and testing sample considered in this experimentation are listed in table 7.

Table 3. English Character Sample sets

Sample Set	Instances (Training)	Speakers (Training)	Instances Per Speaker (Training)	Instances (Testing)	Speakers (Testing)	Noise
S1	26x5	5 Male, 5 Female	5	30	2 Male, 2 Females	No
S2	26x10	5 Male,5 Female	10	50	2 Male, 2 Females	No
S3	26x10	5 Male,5 Female	10	50	2 Male, 2 Females	Yes

The experimentation of each sample set is performed against proposed model, PCA approach and Back propagation neural network. The number of correctly identified instances in each case is shown in table 4.

Table 4. Recognition Rate Analysis

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

The result in the table shows that the recognition of English alphabets for different sample using the proposed approach provided better results. The accuracy of recognition of speech character accurately is shown in figure 2.

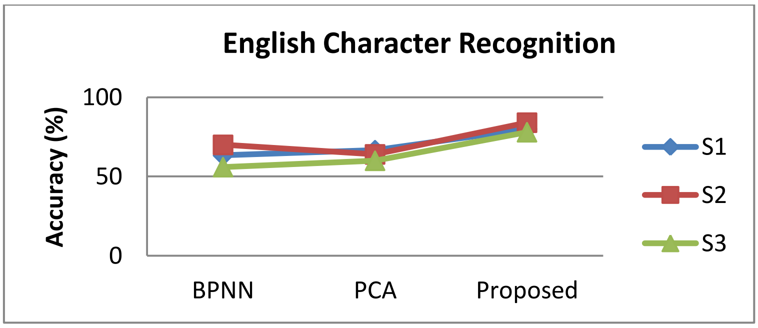

Figure 2. English Character Accuracy Analysis

Journal of Engineering Technology
Volume 6, Special Issue on Technology Innovations and Applications
Oct. 2017, PP. 35-44
ORIGINAL ARTICLE

The figure shows, that the method has provided the absolute improvement for each sample set. The effective improvement obtained in case of noisy signal as the working model has defined a wide noise reduction stage so more accurate results are obtained from the work.

4. Conclusion

In this paper, a multi-level hybrid method is defined to improve the accuracy of the speech recognition model. The working model is defined in three main stages. In first stage, the rectification is provided against different real time speech acquisition problems including noise, turbulence and crosstalk. To improve the recognition, a hybrid feature generation method is applied on normalized speech signal. To process the level based structural analysis, ICA analysis is applied. The normal and ICA featured signal are then processed under different fuzzy-HMM method to assign feature weights and remove silent feature blocks. Finally, the feature data is processed under the SVM classifier to perform speech recognition. The comparative observation shows that the model has improved the speech recognition accuracy between 13 to 25%.

References

[1] Won-Ho Shin, Byoung-Soo Lee, Yun-Keun Lee and Jong-Seok Lee (2000), "Speech/non-speech classification using multiple features for robust endpoint detection", IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, Vol. 3, pp. 1399-1402, 2000.

[2] S. A. Fortune; J. R. Hopgood (2008), ”Speech classification for enhancing single channel blind dereverberation”, *16th European Signal Processing Conference*, Lausanne, pp. 1-5, 2008.

[3] X. Yao; T. Jitsuhiro; C. Miyajima; N. Kitaoka; K. Takeda (2013), ”Estimation of vocal tract parameters for the classification of speech under stress”, *IEEE International Conference on Acoustics, Speech and Signal Processing*, Vancouver, pp. 7532-7536, 2013.

[4] D. Wang; H. Leung; A. P. Kurian; H. J. Kim; H. Yoon (2010), ”A Deconvolutive Neural Network for Speech Classification With Applications to Home Service Robot”, IEEE Transactions on Instrumentation and Measurement, Vol. 59, Issue 12, pp. 3237-3243, 2010.

[5] C. Chelba; M. Mahajan; A. Acero (2003), ”Speech utterance classification”, IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, pp. I-280-I-283, 2003

[6] C. C. Kuo; F. R. Jean; H. C. Wang (1993), ”Speech classification embedded in adaptive codebook search for CELP coding”, IEEE International Conference on Acoustics, Speech, and Signal Processing, Minneapolis, MN, USA, pp. 147-150, 1993

[7] K. E. B. Ooi; M. Lech; N. B. Allen (2013), ”Multichannel Weighted Speech Classification System for Prediction of Major Depression in Adolescents”, IEEE Transactions on Biomedical Engineering, Vol. 60, Issue 2, pp. 497-506, 2013

[8] D. Ververidis; C. Kotropoulos (2005), ”Emotional Speech Classification Using Gaussian Mixture Models and the Sequential Floating Forward Selection Algorithm”, IEEE International Conference on Multimedia and Expo, Amsterdam, pp. 1500-1503, 2005

[9] A. Hurmalainen; T. Virtanen (2013), ”Learning state labels for sparse classification of speech with matrix deconvolution”, IEEE Workshop on Automatic Speech Recognition and Understanding, Olomouc, pp. 168-173, 2013

[10] Roma Bharti; Priyanka Bansal(2015), “ Real Time Speaker recognition system using MFCC and vector quantization technique” , International Journal of Computer Applications, Vol. 117, Issue 1, pp. 25 -31,2015.