An Effective System for Acute Spotting Aberration in the Speech of Abnormal Children Via Artificial Neural Network and Genetic Algorithm

[1]Bharathi. C.R. and [2]V. Shanthi
[1]Research Scholar, Department of Electronics and Communication,
Sathya Bama University, Chennai, India
[2]Department of MCA, St. Josheph's College of Engineering, India

Abstract: Problem statement: In real-world environment, speech signal processing plays a vital role among the research communities. A wide range of researches are carried out in this field for denoising, enhancement and more. Besides the other, stress management is important to identify the spot in which the stress has to be made. **Approach:** In this study, in order to provide proper speech practice for the abnormal person, their speech is analyzed. Initially, the normal and abnormal person's speech are obtained with the same set of words. As an initial process, the Mel Frequency Cepstrum Coefficients (MFCC) is extracted from both words and the Principal Component Analysis (PCA) is applied to reduce the dimensionality of the words. From the dimensionality reduced words, the parameters are obtained and then these parameters are utilized to train the ANN which is used to identify the word that is abnormal. After identifying the abnormal word, the acute word is extracted through the thresholding operation and then FFT is computed for the acute word. From this FFT, the parameters are obtained and then these parameters are used in the genetic algorithm for optimization. GA is used to identify the spot in which the speech practice is required for the abnormal person. **Results:** The proposed system is implemented in the working platform of MATLAB. The performance of the proposed system is tested by generating the dataset for the normal and abnormal female children. **Conclusion:** In this study, an effective system has been proposed to identify the abnormal word and the spot in which the speech has to be improved also identified.

Key words: Principal Component Analysis (PCA), Mel Frequency Cepstrum Coefficients (MFCC), Genetic Algorithm (GA), Speech Recognition System (ASR), Word Accuracy (WA)

INTRODUCTION

Speech is one of the salient forms of communication in daily life (Shirbahadurkar and Bormane, 2009). Speech is formed through the functioning of time-varying vocal tract system. During the production of speech, both excitation and the vocal tract are changed constantly with time (Kinnunen and Li, 2010). In every verbal communication, the quality and precision of speech are given greater importance (Guru *et al.*, 2009). Due to the presence of deleterious properties in the acoustic environment such as multipath distortion (reverberation) and ambient noise, the performance of speech and speaker recognizers are often degraded (Stark and Barkana, 2010). Speech communication applications such as voice-controlled devices, hearing aids and hands-free telephones mostly suffer from poor speech quality because of background noise and room echo (Chomphan, 2011a). Most of the time, particularly during travel, we meet noisy environment. Signal processing techniques removes the noise from the signal-noise mixture and provides an almost noise-free sound for enhanced communication (Guru *et al.*, 2009). Signal processing techniques are exploited in several applications such as speech acquisition, acoustic imaging and communications purposes (Nordholm *et al.*, 2010).

In communication systems, speech signals can be polluted by different environmental noises and so, the communication quality can be degraded making the speech less intelligible. Many attempts have been made for improving the amplitude of speech. If it is feasible to recognize the speech when it is present and offer more gain for that speech than the surrounding environmental sounds, both the accuracy and comfort of speech are enhanced. The enhancement of speech from corrupted noisy observations is mostly based on probabilistic models of speech and noise, e.g., perfect modeling and evaluation of the speech and noise statistics is, therefore, of vast importance (Boubakir and Berkani, 2010). In the modern period, there is a great interest for developing

Corresponding Author: Bharathi. C.R., Reasarch Scholar, Department of Electronics and Communication, Sathya Bama University, Chennai, India

techniques for both speech (and character/word sequences) recognition and synthesis (Chomphan, 2011b). Automatic speech recognition by computer is a process in which the speech signals are mechanically changed into the sequence of words in text (Ghai and Singh, 2012). Generally, the speech recognition has two stages: feature extraction and classification (Bitouk *et al.*, 2010). In speech synthesis, the speeches are automatically generated (acoustic waveforms) from the text (Hazem *et al.*, 2010).

In several studies, the stress and its effects on the acoustic speech signal has given more importance (Rouiha *et al.*, 2008). Since the spoken language expands over time, the auditory information expressed by the stress, intonation and pauses in spoken language must be maintained briefly in memory to make sure that the information is processed and integrated before the stimulus input decays (Lopez-Hurtado and Prieto, 2008). Stress is a psycho-physiological state described by subjective strain, dysfunctional physiological activity and worsening of performance (Hallberg and Bergman, 2011). Stress classification is not only used to enhance the efficacy of speech recognition systems, but also used in telecommunications, military applications, medical applications and law enforcement (Krevelen and Poelman, 2010). The most generally used techniques for inspecting acoustic indicators of stress in speech normally start from pitch, duration, formant frequencies and spectral variation (Hallberg and Bergman, 2011). Stress contributes to the salience of individual words, promoting word identification, integration into the semantic, syntactic and discourse structures, as well as allowing recall of words from memory (Lopez-Hurtado and Prieto, 2008).

MATERIALS AND METHODS

In this study, we concentrate on the abnormal person's speech and to enhance the speech of them by comparing it with the normal persons. This study is very useful for the speech practitioners in the way in which position they have to improve the speech of the abnormal person. Initially, the samples of the normal as well as the abnormal person's speech have been obtained and with the aid of these samples, the further process has to be carried out. Initially, the MFCC of both the speeches are extracted and the PCA is applied to the MFCC to reduce the dimensionality of the speeches. After that the parameters are extracted from this MFCC of both speeches and then these speeches are inputted to generate the ANN.

The abnormal and normal features are used to train the network. Subsequently, the acute word is identified through the thresholding operation and then FFT of the acute word is identified. From the FFT of the acute

word, number of peaks and their amplitude is identified. The amplitude and number of peaks are inputted to the genetic algorithm for the optimization process to identify the position in which the stress has to given. The remaining structure of the study is organized as follows: briefly describes the GA and ANN, then reviews the recent related researches, after that explains the proposed system with adequate equations and neat diagrams along with proper explanations, finally details the implemented results and winds up the study.

Genetic Algorithm (GA): Genetic Algorithm (GA) was developed by Holland in 1970. GA is stochastic search algorithm modeled on the process of natural selection, which emphasize biological evolution (Das and Saha, 2009). The basic concept of GAs is designed to simulate processes in natural system essential for evolution, particularly those that follow the principles first laid down by Charles Darwin of survival of the fittest. As such they represent an intelligent use of an arbitrary search within a defined search space to solve a problem (Ghosh *et al.*, 2010). Genetic algorithms have been utilized effectively to solve several different types of problems, though various factors restrict the success of a GA on a specific function. Problem required are good, but optimal solutions are not perfect for GAs (Hanif *et al.*, 2009).

The GA is an evolutionary process in which a population of solutions obtains over a sequence of generations. During each generation, the fitness of each solution is evaluated and the solutions are selected for reproduction based on their fitness (Kannaiah *et al.*, 2011). The following are the steps involved (Patel *et al.*, 2011):

- Selection: Selection is the process of choosing two parents randomly from the population.
- Crossover (Recombination): In crossover, two parents create an offspring, which enhanced with better individuals
- Mutation: Mutation makes slight random modification in offspring that enables to explore the search space of the algorithm by maintaining diversity in the population. Mutation can be performed either by flipping, interchanging or reversing of bits
- Selection: Replacement is the final stage of reproduction cycle. Crossover and mutation operations produce four populations that is, two parents and two offspring (or children), but not all the four can be used for creation of simulation purpose
- Termination: The convergence criteria decides when to terminate the GA, either by giving the maximum number of generations, deciding elapsed time or ending the algorithm when small enough change in fitness

Artificial Neural Network (ANN): An Artificial Neural Network (ANN) is tailored to imitate natural neural networks using a computing process (El-Shafie and Taha, 2011). Generally, neural networks are utilized to establish a relationship between a set of inputs and a set of outputs. ANNs are non-linear mapping structures based on the function of the human brain. ANNs can detect and learn correlated patterns between input datasets and equivalent target values. Neural Networks are formed by interconnected neurons and each neuron has a certain number of inputs (Sumathi and SanthaKumaran, 2011).

Neurons containing alike characteristics in an ANN are arranged in groups called layers. The neurons in one layer are connected to those in the neighboring layers, but not to those in the same layer. The strength of connection between the two neurons in neighboring layers is represented as a 'connection strength' or 'weight'. An ANN usually has three layers, an input layer, a hidden layer and an output layer (Solaimani, 2009). ANN is a computing system having a number of simple, highly interconnected processing elements called neurons, which processes information in parallel by its dynamic state response to external inputs (Reza *et al.*, 2011).

Related work: A handful of researches are available in literature some of them are listed below in this study.

Anandthirtha and Nagaraj (2009) have integrated the amplitude profile of sampled speech data by utilizing sum of sine functions with a confidence level of more than 90%. Moreover, amplitude correlation technique has been applied between original speech signal samples of normal and pathological subjects and correlation technique has also been applied between the curve fit constant values for normal and pathological subjects. The results of both techniques have been obtained and it has been compared to conclude the varying degrees of speech disability severity.

Caroline Floccia *et al.* (2010) have conducted two tests with 20- 24-month-old English-learning children, to analyze the interaction between segmental and supra-segmental stress-related information in ancient word learning.

After dental rehabilitation with entire dentures, Stelzle *et al.* (2010) have aimed to introduce and authenticate a computer-based Speech Recognition System (ASR) for automatic speech assessment in edentulous patients. The speech result of edentulous patients with and without entire dentures has been compared for inspecting the crash of dentures on speech production. Twenty-eight patients understanding a standardized text have been recorded twice with and without their whole dentures in situ. Under the similar conditions, they have recorded a control set of 40 healthy subjects with natural dentition. By using a polyphone- based ASR, speech quality has been estimated consistent with the percentage of the Word Accuracy (WA). They have proved that, the ASR is very useful and easily applicable tool for automatic speech evaluation in an identical way.

Patel and Rao (2010) have proposed an approach to the identification of speech signal by using frequency spectral information together with Mel frequency, for the development of speech feature representation in a HMM based recognition approach. They have integrated the frequency spectral information to the conventional Mel spectrum base speech recognition approach. For speech signal, the Mel frequency approach has exploited the frequency surveillance in a specified resolution, which results in resolution feature overlapping resulting in recognition limit. A mapping approach like Resolution decomposition with separating frequency has been used for a HMM based speech recognition system. The simulation results have demonstrated that, there is an enhancement in the quality metrics of speech recognition regarding computational time, learning accuracy for a speech recognition system.

Nachamai *et al.* (2010) have discussed that demarcation in human interaction is through two channels: one transmits explicit messages; the other transmits implicit messages about the speakers themselves knowingly or unknowingly. Both linguistics and technology have invested enormous effort in trying to understand the first (explicit) channel, but the second (implicit) is not as well understood. First, building an emotion detection system makes it possible to assess the extent to which theoretical proposals explain people's everyday competence in understanding emotion. Second, model building enforces coherence. It was true that emotions play an important role in the making of speech. The deduction of emotions from speech is of recent origin has been discussed.

Proposed technique for identifying aberration spot: Recently, the speech signal processing plays an important role in the research community and these researches are supportive to the real world environment. In this study, we propose a system to identify the abnormal speech compared with the normal speech and the spot in which the speech has to be improved also identified. The proposed work is comprised of 3 segments Feature extraction, classification and identification of acute spot. Initially, normal human speech and abnormal speech dataset is obtained and then the MFCC is extracted from both the speeches. This speech dataset is developed through in which both the databases are same set of scripts. Subsequently, with the aid of the PCA the dimensionality is reduced

and after that the parameters are obtained from the extracted MFCC feature. Then the speeches are classified through the ANN in order to identify the abnormal one from the normal. After the classification process, the abnormal speech is identified and the FFT is computed for the abnormal speech. The amplitude and its value in the FFT also obtained and these are the parameters utilized for the genetic algorithm to spot the acute location in the speech for improvisation. The step by step processes are explained in detail in this study.

Feature extraction: In this segment, the speech samples are extracted from the normal persons and abnormal persons with the aid of the audio synthesizer. Let D_a and D_b is the abnormal person, normal person speech datasets respectively and from these datasets the MFCC feature is extracted Eq. 1 and 2:

$$D_a = \{w_1, w_2, w_3 ... w_{N_w-1}\} \tag{1}$$

$$D_b = \{\omega_1, \omega_2, \omega_3 ... \omega_{N_w-1}\} \tag{2}$$

Cepstrum and mel-frequency cepstrum differentiated in that in the MFC, frequence bands are equally spaced on the mel scale, which estimates the human auditory system's response more closely than the linearly spaced frequency bands used in the normal cepstrum (Abdalla and Ali, 2010). MFCC's are based on the known variation of the human ear's critical bandwidths with frequency. The MFCC technique makes use of two types of filter, namely, linearly spaced filters and logarithmically spaced filters. To capture the phonetically important characteristics of speech, signal is expressed in the Mel frequency scale. This scale has a linear frequency spacing below 1000 Hz and a logarithmic spacing above 1000 Hz. Normal speech waveform may vary from time to time depending on the physical condition of speakers' vocal cord. Rather than the speech waveforms themselves, MFFCs are less susceptible to the said variations (Sumithra *et al.*, 2011). The following steps are involved in extracting the MFCC feature (Bharathi and Shanthi, 2011):

- Fourier transform is taken for the signal
- With the aid of triangular overlapping windows, the power spectrums are compared with mel scale
- At each of the mel frequency the logs of powers are taken
- Discrete Cosine Transform (DCT) is taken for the mel log powers
- Obtaining the resulting spectrums is the amplitude of MFCCs

With the utilization of above steps, the MFCC features for the above steps the MFCC are obtained from the normal as well as abnormal datasets are obtained which is referred as M_a and M_b Eq. 3-6:

$$N = w - \mu \tag{3}$$

$$cv = \frac{N \times N^T}{n-1} \tag{4}$$

$$Y = N * E^T \tag{5}$$

$$\frac{Y}{E^T} * \mu = w \tag{6}$$

$N -$ Mean Deviation
w-window
μ-mean
cv-covariance vector
$E -$ Eigenvector

The aforesaid eqns are utilized to obtain the PCA of both M_a and M_b here given eqns are the general sets of eqns to generate PCA. In PCA the data are processed by window by window. After PCA the inverse PCA also applied to obtain the dimensionality reduced original information again. After this process completed, the following parameters are obtained from the MFCC featured vectors M_a, M_b. The parameters are mean, standard deviation, maximum amplitude value and its id, minimum amplitude value and its id, MFCC length are extracted for the MFCC featured word and as well as for the original word also extracted and hence for each word we have 14 inputs. For each word the 14 inputs are given to the neural network to classify whether the inputted word is normal or abnormal. For this each word has the parameters as said above and the vector is S inputted to the neural network. The ANN is trained through this inputted vector to classify either normal or abnormal.

Classification through ANN: In the classification, we utilize the ANN (Bharathi and Shanthi, 2011) for the classification purpose in order to separate the normal and for identifying the abnormal dataset. The following steps details the training process of ANN

Step 1: As the first step, set up the input weights to every neuron, apart from the neurons of the input layer

Step 2: This neural network has 14 input layers which are the parameters of a word. N_h hidden layers and one output layer to identify the word inputted is either normal or abnormal. The value at the output layer is either true or false depending on whether the input word is abnormal or normal. In this neural network, 14 input neurons and a bias neuron, N_h hidden neurons and a bias neuron and one output neuron are present

Step 3: The weights are added to the designed N also it is biased. The developed N is shown in Fig. 1

Step 4: The basis function and the activation function which are chosen for the designed N is given below Eq. 7-9:

$$Z_i = \alpha + \sum_{j=1}^{N_h} \left(\begin{array}{l} w_{ij}S_t[1] + w_{ij}S_t[1] + w_{ij}S_t[2] \\ + w_{ij}S_t[3] + w_{ij}S_t[4] + ... + w_{ij}S_t[14] \end{array} \right) \qquad (7)$$

$$h(Z_i) = \frac{1}{1 + e^{-Z_i}} \qquad (8)$$

$$h(Z_i) = Z_i \qquad (9)$$

Equation 7 is the basis function for the input layer, where $S_t[1]$-$S_t[14]$ are the parameters for the word which are mean, standard deviation, maximum amplitude value and its id, minimum amplitude value and its id, MFCC length for MFCC featured word and for the original word we extract the same word for the MFCC length here we extract the word length. Here w_{ij} is the weight of the neuron and α is the bias. The

sigmoid function for the hidden layer is given in Eq. 7 and the activation function for the output layer is given in Eq. 8. The basis function given in Eq. 7 is commonly used in all the remaining layers (hidden and output layer, but with the number of hidden and output neurons, respectively). The output of the ANN is obtained by giving the region vector as its input.

Step 5: The learning error is determined for the network N as follows Eq. 10:

$$Er = \frac{1}{N_h} \sum_{o=0}^{N_h - 1} D_o - Z_o \qquad (10)$$

Here, Er is the error in the FF-ANN, D_o is the desired output and Z_o is the actual output.

Error minimization by BP algorithm: The steps involved in the training of BP algorithm based N is given below:

- Randomly generate weights in the interval [0,1] and assign it to the neurons of the hidden layer and the output layer. But all neurons of the input layer have a constant weight of unity
- Determine the BP error using Eq. 9 and give the training gene data sequence as input to the N Eq. 7-38 and Eq. 9 show the basis function and transfer function
- Adjust the weights of all the neurons when the BP error is determined as follows Eq. 11:
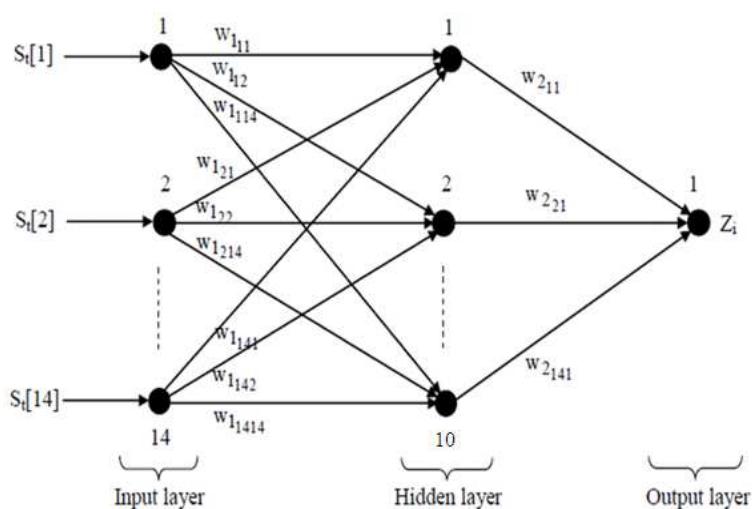

Fig. 1: Network structure of the generated neural network for the classification process

$$w_{ij} = w_{ij} + \delta w_{ij} \tag{11}$$

The change in weight δw_{ij} given in Eq. 7 can be determined as $\delta w_{ij} = \gamma.Z_{ij}.E_r$, where E_r is the BP error and γ is the learning rate and it normally ranges from 0.2-0.5.

(4) After adjusting the weights, repeat steps (7) and (8) until the BP error gets minimized. Normally, it is repeated till the criterion, E<0.1 is satisfied.

Once the error gets minimized to a minimum value it is construed that the designed FF-ANN is well trained for its further testing phase and the BP algorithm is terminated. Thus the neural network is trained using the parameters for each word.

Identifying the acute spot of aberration of abnormal children: In this study, the spot in which the speech practice is required is done here. Prior to this process, the identified abnormal words are stored in the vector A_v With the aid of the thresholding operation the acute word is extracted and stored. In this module, the identified abnormal speech sample is utilized to identify the position to be trained. For the identification process, genetic algorithm is utilized. A Genetic Algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of Evolutionary Algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection and crossover.

Generation of chromosomes: The amplitude and its value of the abnormal word is extracted and as well as for its normal is also extracted. Initially generate N_p number of random chromosomes and the genes are the corresponding indices of the vector abnormal word's amplitude and its value of abnormal word Eq. 12:

$$I^{(k)} = \left\{ I_0^{(k)}, I_2^{(k)}, I_3^{(k)}, \cdots, I_{n_I-1}^{(k)} \right\} 0 \le k \le N_p - 1$$
$$0 \le \eta \le n_I - 1 \tag{12}$$

In Eq. 12, $I_\eta^{(k)}$ represents the ηth gene of the kth chromosome. In this, the abnormal word's amplitude and its corresponding value are located at each of the indices. After the chromosome generation, fitness function is applied on these generated chromosomes.

Fitness function: Here, the fitness of the generated chromosomes is the comparison of the values located at the indices and with their corresponding normal word. This may indicate that the difference between the values of both normal and abnormal identify the acute location of the word. In this fitness process, the generated index's corresponding value is taken into consideration for the comparison process with its corresponding normal word which is already extracted and stored:

For each k, where $0 \le k \le N_p^{-1}$
For each η, where $0 \le \eta \le nI-1$
Compare abnormal word's amplitude value in the index $I_\eta^{(k)}$ with its original word's amplitude value
End For
Compute the mean of chromosome K
Sort chromosomes based on mean
End for

Crossover and mutation: Among different types of crossover we have chosen single point crossover and the crossover point is C. In this crossover the genes (indices here) started from C is interchanged with the genes with another chromosome. From the sorted chromosomes which are the outcome of fitness, the best chromosomes are selected and the crossover is to be occurred. Hence we obtain the children chromosomes which are utilized for the identification of acute spot in which the training has to be required.

Subsequently mutation is performed on the chromosomes obtained after crossover. The mutation process replaces N_M number of genes from every chromosome with the new genes. In this N_M numbers of genes are nothing but the genes which are highly deviated. Then, chromosomes which are selected for crossover operation and the chromosomes which are obtained from the mutation are combined and so the population pool is filled up with the N_p chromosomes. Then, the process is repeated iteratively until it reaches a maximum iteration of I_{max}.

Selection of optimal solution: After the process is repeated I_{max} times, chromosomes that have maximum fitness value are selected from the resultant group of chromosomes as the best chromosomes. Here, the best chromosomes are the chromosomes that have maximum fitness. The indices, which are obtained from the genes of the best chromosomes, represented the acute position of the abnormal word in which the training has to be required for the speech practice.

RESULTS

The proposed system was implemented in the working platform of MATLAB (version 7.11). In this

with the aid of the Free Audio Editor we generate the dataset with the normal and abnormal female children within the age limit 6-10. For normal data we utilized 2 female children and for abnormal data we utilized a female child for our system and their normal frequency range is from 0-300 kHz.

Initially, the words are extracted from the both normal and abnormal children and then the MFCC feature has been extracted from it. Subsequently, the PCA is applied to reduce the dimensionality of the words and then they are inputted to the neural network to identify the abnormal and the normal word. Figure 2 is the generated neural network N structure.

DISCUSSION

In this study, the proposed technique is tested with the database of 100 words with normal children 2 and an abnormal child data. In the genetic algorithm the abnormal occurred spot is identified. Here $N_p = 10$ and each chromosome with gene 20 amplitude has been generated. After generation the fitness comparison with normal data is analyzed. After the crossover process, the fitness is applied on the children chromosomes and hence the highly deviated chromosomes are replaced with the new gens. The iteration is repeated until it reaches the Imax and hence we tested with different words; the output obtained is for example for the words 'Analysis' and 'Dinosaurus' shown in the figures. In Fig. 3 the word ('Analysis') of abnormal child is shown and Fig. 4 and 5 shown the normal children's plotted signal In Fig. 6, the acute spot is identified in the abnormal child's word. Likewise the Fig. 7 shows the abnormal word and Fig. 8 and 9 shows the normal children's word. Figure 10 shows the identified acute spot in the word 'Dinosaurus. Figure 11 show the GUI for the word analysis and the identified acute spot is also shown in it.
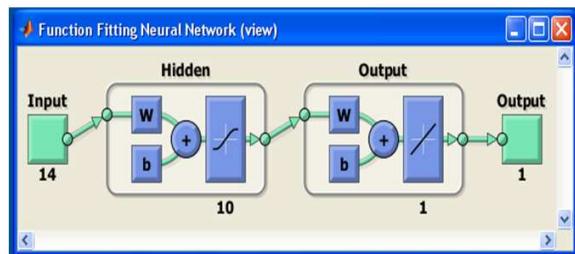

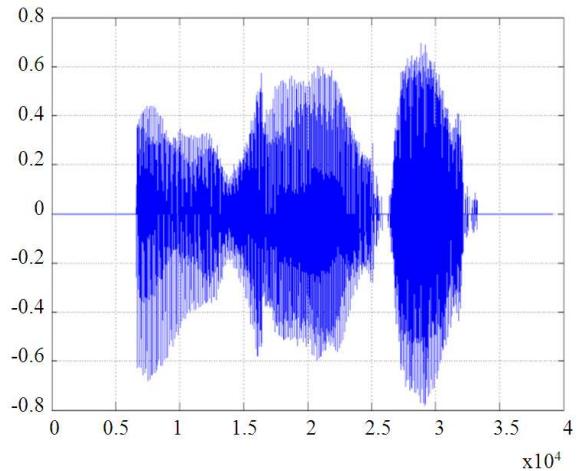Fig. 2: Generated ANN view for classification

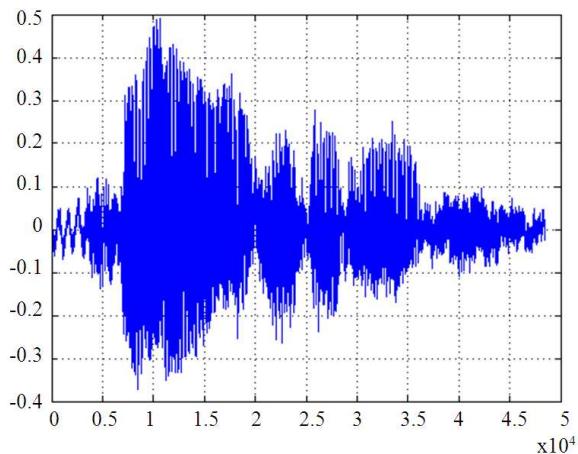Fig. 3: Diagram for the word Analysis [Abnormal child]

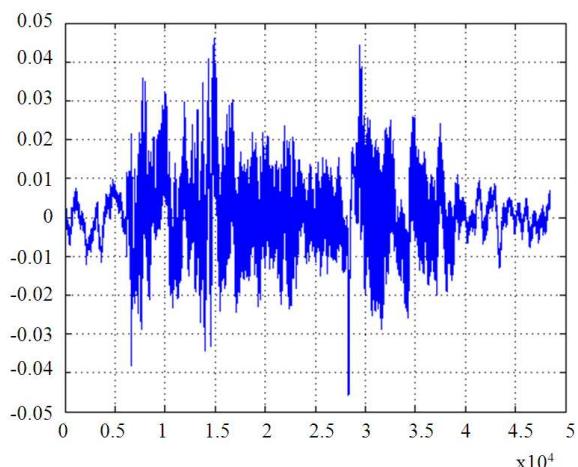Fig. 4: Signal for the word 'Analysis' [Normal child1]

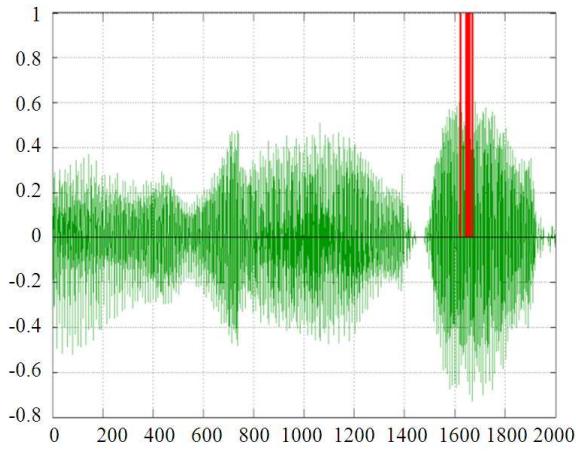Fig. 5: Signal for the word 'Analysis' [Normal child2]

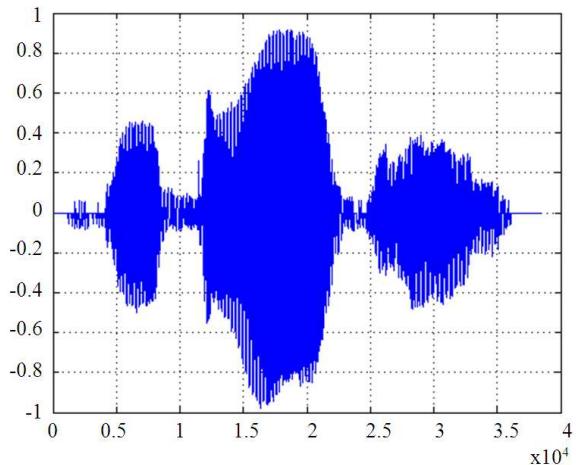Fig. 6: Identified acute spot of aberration for the word 'Analysis' [Abnormal child]

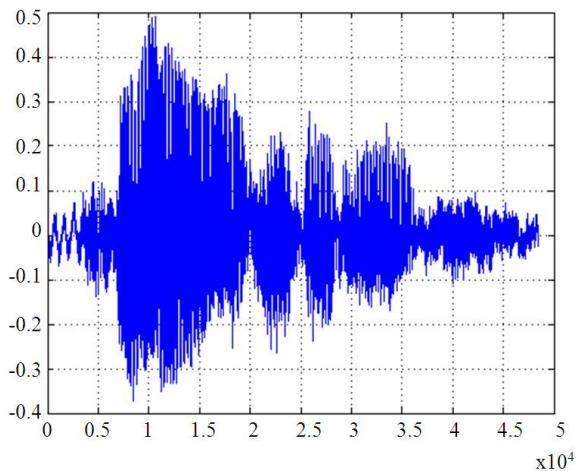
Fig. 9: Signal for the word 'Dinosaurus' [Normal child2]
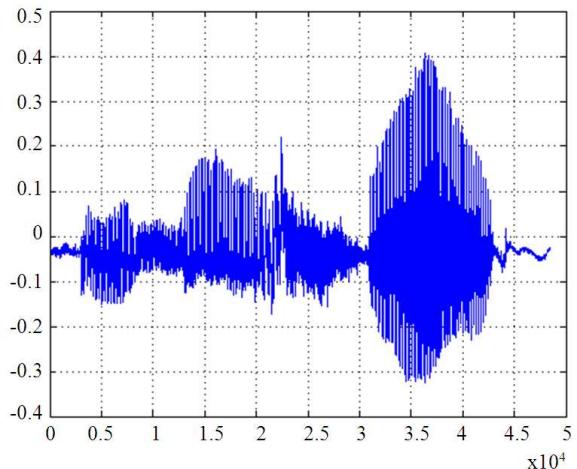

Fig .7: Signal for the word 'Dinosaurus' [Abnormal child]
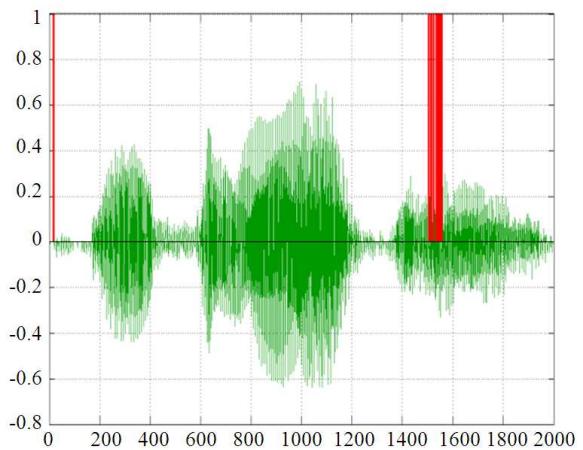

Fig. 10: Identified acute spot of aberration for the word Dinosaurus' [Abnormal child]

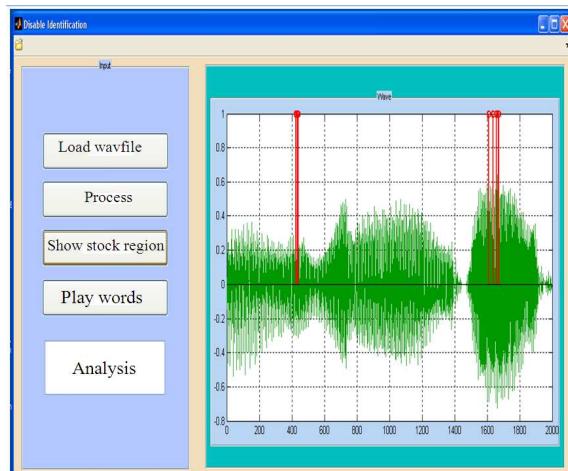Fig. 8: Signal for the word 'Dinosaurus' [Normal child1]

Fig. 11: Output GUI for the abnormal word 'Analysis'

CONCLUSION

In this study, an effective system has been proposed to identify the abnormal word and the spot in which the speech has to be improved also identified. In order to identify this, initially the MFCC is obtained from both the normal and abnormal words and then the parameters are obtained. After that, both the words are inputted to the ANN to identify the abnormal word. Subsequently, FFT has been obtained for the abnormal word and then the number of peaks and the amplitude has been extracted. These parameters are utilized for the genetic algorithm to optimize the spot in which the speech has to be improved. The results have been showed the spot in which the speech has to be improved is identified shown in the Fig. 6 and 10 shows the identified acute spot of the words 'Analysis' and 'Dinosaurus' respectively.

REFERENCES

Abdalla, M.I. and H.S. Ali, 2010. Wavelet-Based mel-frequency cepstral coefficients for speaker identification using hidden markov models. J. Tele Commun., 1: 16-21.

Anandthirtha, B.G. and Nagaraj, 2009. Optimal curve fitting of speech signal for disabled children. Int. J. Comput. Sci. Inform. Technol., 1: 99-107.

Bharathi, C.R. and V. Shanthi, 2011. Classification of speech for clinical data using artificial neural network. Int. J. Comput. Sci., 8: 359-365.

Bitouk, D., R. Verma and A. Nenkova, 2010. Class-level spectral features for emotion recognition. Speech Commun., 52: 613-625. DOI: 10.1016/j.specom.2010.02.010

Boubakir, C. and D. Berkani, 2010. Speech enhancement using minimum mean-square error amplitude estimators under normal and generalized gamma distribution. J. Comput. Sci., 6: 700-705. DOI: 10.3844/jcssp.2010.700.705

Chomphan, S., 2011a. High pitch delay resolution technique for tonal language speech coding based on multi-pulse based code excited linear prediction algorithm. J. Comput. Sci., 7: 65-69. DOI: 10.3844/jcssp.2011.65.69

Chomphan, S., 2011b. Analysis of decision trees in context clustering of hidden markov model based thai speech synthesis. J. Comput. Sci., 7: 359-365. DOI: 10.3844/jcssp.2011.359.365

Das, S. and B. Saha, 2009. Data quality mining using genetic algorithm. Int. J. Comput. Sci. Secu., 3: 105-112.

El-Shafie, M.N. and Taha, 2011. Performance of artificial neural network and regression techniques for rainfall-runoff prediction. Int. J. Physical Sci., 6: 1997-2003.

Floccia, C., T. Nazzi, K. Austin, F. Arreckx and J. Goslin, 2010. Lexical stress and phonetic processing in word learning in 20- to 24-month-old English-learning children. Dev. Sci., 14: 602-613. PMID: 21477198

Ghai, W. and N. Singh, 2012. Analysis of automatic speech recognition systems for indo-aryan languages: punjabi a case study. Int. J. Soft Comput. Eng., 2: 379-385.

Ghosh, S., S. Biswas, D. Sarkar and P.P. Sarkar, 2010. Mining frequent itemsets using genetic algorithm. Int. J. Artificial Intell. Appli., 1: 133-143.

Guru, S., R. Shivaramakrishnan and Srinivasan, 2009. Real time speech enhancement using blackfin processor BF533. J. Instrument Soc. India, 37: 67-79.

Hallberg, L.R. and S. Bergman, 2011. Minimizing the dysfunctional interplay between activity and recovery: A grounded theory on living with fibromyalgia. Int. J. Q. Stud. Health Well-being, 6: 1-10. PMID: 21687553

Hanif, F., K. Nasiruddin, K.S. Inayatullah and S.T. Nizami, 2009. Solving tsp problem by using genetic algorithm. Int. J. Basic Applied Sci., 9: 79-88.

Hazem, R., M. El-Bakry and I.R. Ismail, 2010. Diphone speech synthesis system for arabic using MARY TTS. Int. J. Comput. Sci. Inform. Technol., 2: 18-26.

Kannaiah, S.K., J. Thangavel and D.P. Kothari, 2011. A genetic algorithm based multi objective service restoration in distribution systems. J. Comput. Sci., 7: 448-453. DOI: 10.3844/jcssp.2011.448.453

Kinnunen, T. and H. Li, 2010. An overview of text-independent speaker recognition: From features to supervectors. Speech Commun., 52: 12-40. DOI: 10.1016/j.specom.2009.08.009

Krevelen, D.W.F.V. and R. Poelman, 2010. A survey of augmented reality technologies, applications and limitations. Int. J. Virtual Reality, 9: 1-20.

Lopez-Hurtado, E. and J.J. Prieto, 2008. A microscopic study of language-related cortex in Autism. Am. J. Biochem. Biotechnol., 4: 130-145. DOI: 10.3844/ajbbsp.2008.130.145

Nachamai, M., T. Santhanam and C.P. Sumathi, 2010. A new fangled insinuation for stress affect speech classification. Int. J. Comput. Appli., 1: 19-19.

Nordholm, S., T. Abhayapala, S. Doclo, S. Gannot and P. Naylor *et al.*, 2010. Microphone array speech processing. EURASIP J. Adv. Signal Proc., 2010: 1-3.

Patel, H.K., A. Sircar, S. Sheth and R. Jadvani, 2011. Application of genetic algorithm to hydrocarbon resource estimation. J. Petroleum Gas Eng., 2: 83-92.

Patel, I. and Y.S. Rao, 2010. Speech recognition using HMM with MFCC-An analysis using frequency specral decomposion technique. Signal Image Proce.: Int. J., 1: 101-110.

Reza, M., Z. Mohsen, F.H. Sayyaadi and S.B. Shouraki, 2011. Hysteresis nonlinearity identification using new preisach model-based artificial neural network approach. J. Applied Math. 2011: 1-22. DOI: 10.1155/2011/458768

Rouiha, K., D. Bachir and B. Ali, 2008. Analysis of speech processing strategies in cochlear implants., J. Comput. Sci., 4: 372-374. DOI: 10.3844/jcssp.2008.372.374

Shirbahadurkar, S.D. and S. Bormane, 2009. Speech synthesizer using concatenative synthesis strategy for marathi language (Spoken in Maharashtra, India). Int. J. Recent Trends Eng., 2: 80-82.

Solaimani, K., 2009. Rainfall-runoff prediction based on artificial neural network (A Case Study: Jarahi Watershed). Am. Eurasian J. Agric. Environ. Sci., 5: 856-865.

Stark, B. and B.D. Barkana, 2010. Acoustic echo cancellation: dual architecture implementation. J. Comput. Sci., 6: 101-106. DOI: 10.3844/jcssp.2010.101.106

Stelzle, U., K. Bocklet, N. Schuster, E. Seiss and Nkenke, 2010. Automatic, computer-based speech assessment on edentulous patients with and without complete dentures-preliminary results. J. Oral Rehab., 37: 209-216. PMID: 20085615

Sumathi, B. and A. SanthaKumaran, 2011. Pre-diagnosis of hypertension using artificial neural network. Global J. Comput. Sci. Technol., 11: 43-47.

Sumithra, M.G., K. Thanuskodi and A.A.H. Jenifer, 2011. A new speaker recognition system with combined feature extraction techniques. J. Comput. Sci., 7: 459-465. DOI: 10.3844/jcssp.2011.459.465