International Journal of Latest Trends in Engineering and Technology
Vol.(10)Issue(3), pp.057-060
DOI: http://dx.doi.org/10.21172/1.103.10
e-ISSN:2278-621X

A STRESS IDENTIFICATION USING COMBINED METHOD OF NEURAL NETWORK AND EMD COMPONENTS IN SPEECH

Mrs. N.P. Dhole[1], Dr. S.N. Kale[2]

Abstract-Stress is a burden which is acting on individuals body in formof tremor, anxiety, pressure, greater heart beats, sweating,shakiness of hands or legs and movements of eyes. Everybodyagrees to avoid the stress but when the actual time comes noone can escape from it. Speech is a non-linear and non-stationarysignal and termed as a physical quantity which canbe measured. Analysis of speech is measuring the stress inducedin speech in the situations like fatigue, emotions and medicalproblems [1]. This paper presents a combined method of Neural Network and EMD Components to analyze Stress in Speech.
Keywords: EMD, NN, DB8 Transform, MFCC.

1. INTRODUCTION

Today's world is filled with fast and growing technology.Day to day the working environment is having ease ofComfort with this technology. Due to this there has beenless physical movements and has induced mental stressLevels which introduced new field of research as analysisof speech. The speech signal is a way of communicationwhich happens between the two or more persons and againit is cheapest medium of stress analysis. Stress analysis isnot only the field of medical but has diverted attentioninto engineering field too. Stress recognition from speechsignal is the study of speech as normal or stressed. Againstress speech will be further considered as high level stress or low level stress depending on its amplitude level. These levels can be measured by the changes or fluctuations inmuscles of vocal chords named as microtremor frequency in speech [2]. The definition of microtremor is a low amplitude oscillation in the range of 8-12 Hz. When stresslevel is increased it increases this microtremor [3]. In thispaper we have used method called Adaptive EmpiricalMode Decomposition for decomposing speech intoIntrinsic Mode Functions (IMF) [4] along-with Neural Network to classify Stress Types.

2. AUDACITY SOFTWARE

Audacity is a free and Open Source Software, it's an easy-to-use audio editor and recorder for Windows, Mac OS X, GNU/Linux, and other operating systems. Audacity is free software, developed by a group of volunteers and distributed under the GNU General Public License (GPL) [5]. We can use Audacity to Record live audio, Convert tapes and records into digital recordings or CDs Edit Ogg Vorbis, MP3, and WAV sound files to Cut, copy, splice, and mix sounds together to Change the speed or pitch of a recording. Audacity can record live audio through a microphone or mixer, or digitize recordings from cassette tapes, vinyl records, or minidiscs. In this research work we have recorded the speech using audacity with different frequencies 8 kHz, 16 kHz and 44.1 kHz.

3. FLOWCHART OF THE ALGORITHM

3.1 EMD Components-

Adaptive Empirical Mode Decomposition Algorithm has been implemented in the papers [6][7]. We got the components i.e EMD components [8][9]. The EMD components are again segmented and filtered out.

[1] Assistant Professor, Department of Electronics and Telecommunication Engineering, PRMIT&R Badnera, Amravati (Maharashtra), India
[2] Associate Professor, Department of Applied Electronics, Sant Gadge Baba Amravati University, Amravati, (Maharashtra), India

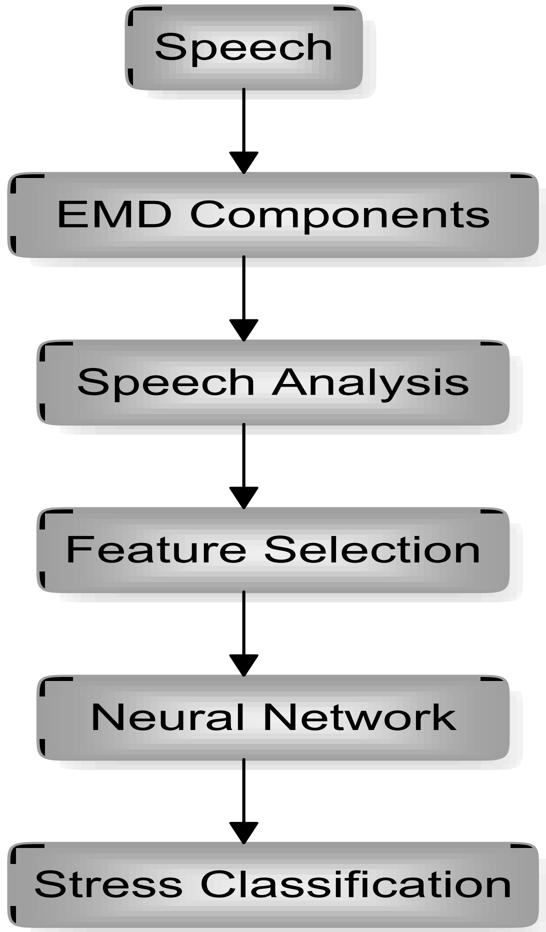

Figure 1. AEMD and NN Algorithm

3.2. Feature Selection-

Feature Selection Extracts MFCC. Mel frequency cepstral co-efficient are mostly used features for any speech recognition system. We are using MFCC for stress speech feature extraction [8]. Feature extraction undergoes raw speech transformation into useful parameters without changing speech information. It consists of Pre-emphasis, Framing, windowing, spectral estimation, Mel Filtering DCT etc. as procedures for this features extraction. In stress speech extraction we convert into useful data to classify and train the neural network. The daubichies 8 (DB8) variant of the wavelet transform is the most suitable for feature reduction and pattern preservation.

3.3. Neural Network Approach-

This approach is designed for complicated tasks but it is not as efficient as HMM in the case of large vocabularies. Phoneme recognition is the general approach of neural networks. In this approach the technique of intelligence, analyzing and visualizing of speech signal is done to measure phonetic features [9].The network includes a huge number of neurons. Each neuron computers nonlinear weight of inputs and broadcast result to the outgoing units, training sets are used for assigning pattern of values to input and output neurons, training set determines the weight of strength of each pattern.

3.4 Stress Classification

We tested our stress detection systems under 5 different categories, namely,
Stress Type 1
Stress Type 2
Stress Type 3
Stress Type 4
No Stress
Stress type 1 arises from problems like workload and anxiety. Stress Type 2 induces from noise and speech quality. Stress type 3 corresponds to effects causing due to medicines, illness and narcotics. Stress Type 4 refers to problem arises from vibration and acceleration. Finally No stress means persons is in normal condition.

4. RESULT

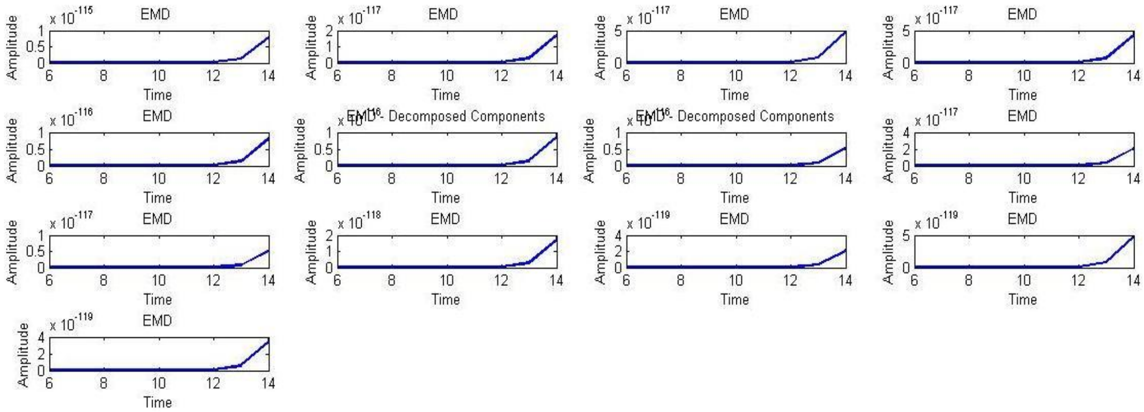

Figure 2. Speech Signal EMD

Above figure 2.shows the Empirical ModeDecomposition Components from AEMD Algorithm whereX-label is amplitude and Y-label is time.

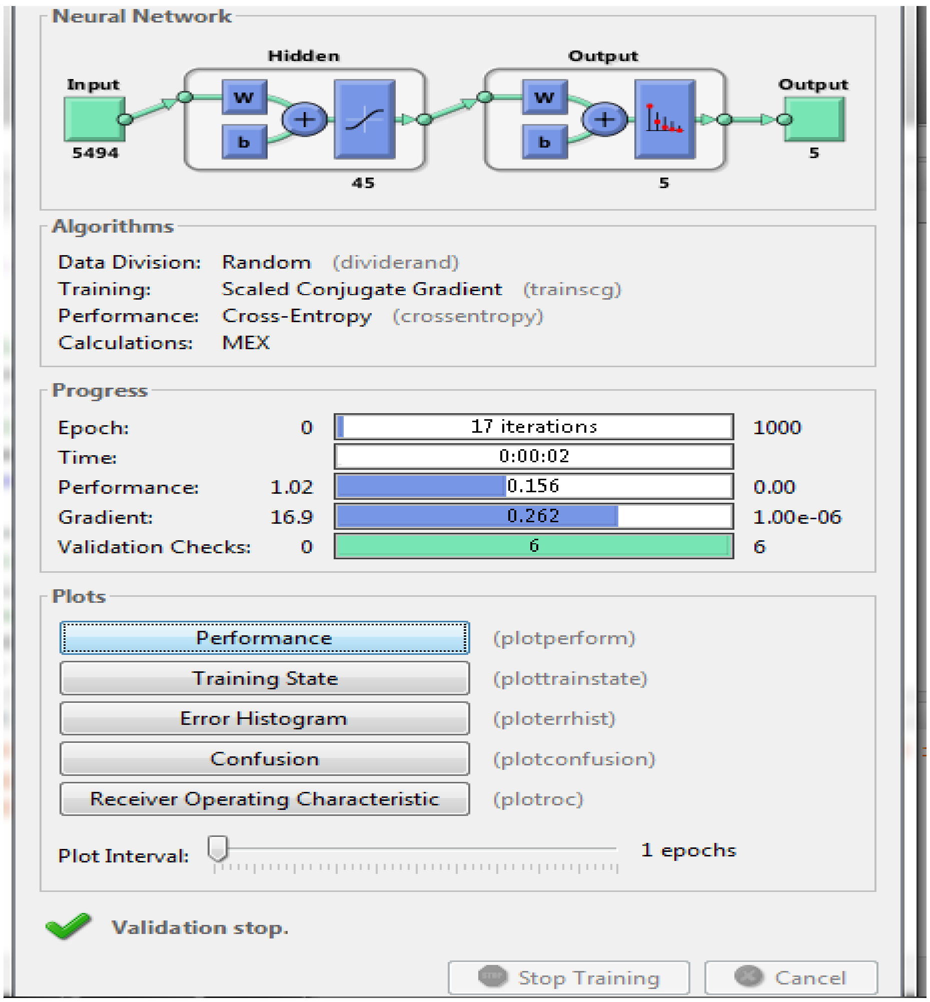

Figure.3. Neural Network Training

Above figure 3 depicts the Neural Network Performance with hidden neurons 17 iterations.

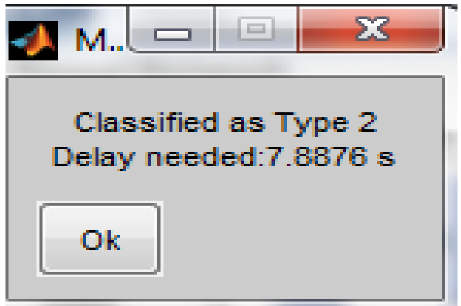

Figure 4.Delay Needed

From fig.4describes the Delay needed for NN classifier for a particular Wave file and classified as Stress Type 2. Delay arises due to echo and reverberations in speech.

5. CONCLUSION

From the above results we have combined two methods AEMD with NN to form a new algorithm which contributes in identification of stress in speech signal more prominently in near future we will try this algorithms on data sets or healthy speech identification.

6. REFERENCES

[1] James Z. Zhang, NyagaMbitiru, Peter C. Tay, RobertD Adams, “Analysis of Stress in Speech UsingAdaptive Empirical Mode Decomposition”, IEEEAsilomar , pp. 361-365, 2009.
[2] Roberto Cabrera Cosetl, J. M. David Baez Lopez,“Voice Stress Detection: A Method for Stress AnalysisDetecting Fluctuations on Lippold micrometerspectrum using FFT”, IEEE, pp. 928-932, 2008.
[3] C.Shahnaz ,S.Sultana, S.A.Fattah,et.al, “Emotionrecognition based on EMD-Wavelet analysis of speechsignals,” 2015 IEEE Conference on Digital SignalProcessing, pp. 307-310, DOI10.11.09/ICDSP.2015.725188
[4] A. Ahrabian and D. P. Mandic, "A class of multivariatedenoising algorithms based on synchrosqueezing,"IEEE Transactions on Signal Processing,vol. 63, no. 9, pp. 2196-2208, 2015.
[5] http://audacity.sourceforge.net/download.
[6] HindraKurniawan, Alexandr V. Maslov, MykolaPechenizkiy, “Stress Detection from Speech andGalvanic Skin Response Signals”, IEEE Press 2013, pp.209-214.
[7] Hamid “Empirical Mode Decomposition for advancedspeech signal processing”, Journal of Sig. Processing,vol. 17, no. 6, pp. 215-229, Nov 2013.
[8] James Z. Zhang, NyagaMbitiru, Peter C. Tay, RobertD Adams, “Analysis of Stress in Speech UsingAdaptive Empirical Mode Decomposition”, IEEEAsilomar, pp. 361-365, 2009.
[9] Mrs. N.P. Dhole, Dr. S. N. Kale, “Adapive Empirical Mode Decomposition of Speech Signal”,International Journal of Emerging Trends & Technology in Computer Science, Vol.6, issue 3,pp.422-425,May-June 2017.
[10] Mrs. N.P. Dhole, Dr. S. N. Kale, “Identification of Stress Speech in Marathi Language using AEMD Algorithm,” International Journal of Scientific Research & Review, Vol 7, issue 3, pp. 471-474, 2018
[11] BOGERT, B. P.; HEALY, M. J. R.; TURKEY, J. W. “The QuefrencyAlanysis of Time Series for Echoes: Cepstrum, Pseudo Autocovariance, Cross-Cepstrum and Saphe Cracking”, Proceedings of the Symposium on Time Series Analysis, (M. Rosenblatt, Ed) Chapter 15, New York: Wiley, pp.209-243, 1963.
[12] Weintraub, M., Murveit, H., Cohen, M., Price, P., Bernstein, J., Baldwin, G., & Bell, D. (1989, May). Linguistic constraints in hidden Markov model based speech recognition. In Acoustics, Speech, and Signal Processing, 1989.ICASSP-89, International Conference on pp. 699-702, 1989.