

**CLASSIFICATION OF STRESS LEVEL BASED ON SPEECH  
FEATURES**

**ARSHED AHMED JASIM**

**UNIVERSITI UTARA MALAYSIA**

**2014**

# **CLASSIFICATION OF STRESS LEVEL BASED ON SPEECH FEATURES**

A dissertation submitted to Dean of Research and Postgraduate Studies  
Office

In partial Fulfilment of the requirement for the degree  
Master of Science (Information Technology)  
Universiti Utara Malaysia

By  
Arshed Ahmed Jasim

## **Permission to Use**

In presenting this dissertation in fulfilment of the requirements for a Master of Science in Information Technology (MSc. IT) from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this dissertation in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my dissertation.

Requests for permission to copy or to make other use of materials in this dissertation, in whole or in part should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences  
UUM College of Arts and Sciences  
Universiti Utara Malaysia  
06010 UUM Sintok  
Kedah Darul Aman

## Abstrak

Kehidupan kontemporari adalah penuh dengan cabaran, gangguan, tarikh akhir, kekecewaan dan permintaan yang tidak berkesudahan. Ini boleh mengakibatkan seseorang itu stres. Stres telah menjadi satu fenomena global yang dialami dalam kehidupan moden harian kita. Stres mungkin memainkan peranan penting dalam gangguan psikologi dan tingkah laku seperti kebimbangan atau kemurungan. Oleh itu, pengesanan awal tanda-tanda dan gejala stres merupakan penawar ke arah mengurangkan kesan buruk dan kos yang tinggi dalam pengurusan stres. Usaha penyelidikan yang dibentangkan ini merangkumi teknik Pengenalan Percakapan Automatik (ASR) untuk mengesan stres sebagai alternatif yang lebih baik berbanding pendekatan yang lain seperti analisis kimia, kekonduksian kulit, elektrokardiogram yang mahal dan mempunyai kesan halangan dan gangguan. Dua set data suara direkodkan daripada sepuluh orang pelajar Arab di Universiti Utara Malaysia (UUM) iaitu dalam mod rehat dan stress. Ciri-ciri percakapan seperti frekuensi asas ( $f_0$ ); formants (F1, F2, dan F3), tenaga dan Pekali Frekuensi Cepstral Mel (MFCC) ini diekstrak dan dikelaskan menggunakan jiran K-terdekat, Analisisa Diskriminan Linear dan Rangkaian Neural Buatan. Keputusan dari nilai purata frekuensi asas mendedahkan bahawa peningkatan stres adalah berkait rapat dengan pertambahan nilai frekuensi asas. Daripada tiga metod pengelasan, prestasi jiran K-terdekat (KNN) adalah terbaik diikuti oleh analisisa diskriminan linear (LDA) manakala rangkaian neural buatan (ANN) menunjukkan prestasi yg paling rendah. Klasifikasi tahap stres rendah, sederhana dan tinggi telah dilakukan berdasarkan keputusan klasifikasi daripada KNN. Kajian ini menunjukkan kebolegunaan maju ASR sebagai cara yang lebih baik pengesanan stres dan pengelasan.

## Abstract

Contemporary life is filled with challenges, hassles, deadlines, disappointments, and endless demands. The consequent of which might be stress. Stress has become a global phenomenon that is been experienced in our modern daily lives. Stress might play a significant role in psychological and/or behavioural disorders like anxiety or depression. Hence early detection of the signs and symptoms of stress is an antidote towards reducing its harmful effects and high cost of stress management efforts. This research work thereby presented Automatic Speech Recognition (ASR) technique to stress detection as a better alternative to other approaches such as chemical analysis, skin conductance, electrocardiograms that are obtrusive, intrusive, and also costly. Two set of voice data was recorded from ten Arabs students at Universiti Utara Malaysia (UUM) in neural and stressed mode. Speech features of fundamental, frequency ( $f_0$ ); formants (F1, F2, and F3), energy and Mel-Frequency Cepstral Coefficients (MFCC) were extracted and classified by K-nearest neighbour, Linear Discriminant Analysis and Artificial Neural Network. Result from average value of fundamental frequency reveals that stress is highly correlated with increase in fundamental frequency value. Of the three classifiers, *K*-nearest neighbor (KNN) performance is best followed by linear discriminant analysis (LDA) while artificial neural network (ANN) shows the least performance. Stress level classification into low, medium and high was done based of the classification result of KNN. This research shows the viability of ASR as better means of stress detection and classification.

## **Acknowledgement**

### **By the Name of Allah, the Most Gracious and the Most Merciful**

First and foremost, I thank to Allah for blessing me with good health to be able to complete this project. This dissertation is accomplished by the student but the knowledge and experiences gathered, evolved during the process through the efforts of many people. Without their cooperation, encouragement and suggestion; this study would not have been possible.

I heartily thank my supervisors Dr. Shahrul Azmi Mohd.Yusof and Ms. Aniza Mohamed Din They have supported me throughout my research process. They gave support, insight, guidance and encouragement throughout to fulfil my study.

My acknowledgements would not be complete until I thank my great father and my lovely mother for their prayers, love, and support in the duration of my study in Malaysia. Who sacrificed much and supported my efforts with understanding and constant encouragement. Without them, it is almost impossible for me to complete this master's degree. May Allah bless them.

I would like to express my gratitude to my brother and my sister for their love and moral care. They are supporting me spiritually throughout my life. Without them, I'm nothing here.

Finally, many thanks go to all of Malaysian people, especially UUM lecturers and staffs, for their very good dealing with all of international students. As well as, Thanks to College of Arts and Science with its community whom has made this possible by organizing this course which provided the opportunity for me to share and learn their information in applying the essential apparatus to the mission.

**“Thank you UUM”**

# Table of Contents

Permission to Use.....	i
Abstrak.....	ii
Abstract.....	iii
Acknowledgement .....	iv
Table of Contents.....	v
List of Tables .....	viii
List of Figures .....	ix
List of Acronyms .....	xi
<b>CHAPTER ONE INTRODUCTION.....</b>	<b>1</b>
1.0 Introduction.....	1
1.1 Background .....	1
1.2 Problem Statement .....	4
1.4 Research Objectives .....	6
1.5 Motivation.....	6
1.6 Contributions.....	7
1.7 Scope.....	8
<b>CHAPTER TWO LITERATURE REVIEW .....</b>	<b>9</b>
2.0 Introduction.....	9
2.1 ASR performance and variability.....	9
2.2 Types of variability .....	10
2.2.1 Intrinsic variability in speech .....	10
2.2.2 Environmental variability.....	10
2.2.3 Long term variability.....	11
2.2.4 Short term variability .....	11
2.3 Stress as a source of variability.....	11
2.4 Types of stress and stressor.....	13
2.5 Measuring stress.....	14

2.6 Stressed speech features .....	15
2.7 Speech corpus .....	17
2.8 Speech recognition .....	18
2.8.1 Feature extraction.....	19
2.8.2 Pattern recognition .....	21
2.8.3 Artificial neural network as vowel classifier.....	23
2.8.4 Linear Discriminant Analysis.....	24
2.8.5 K-Nearest Neighbors.....	25
2.9 Stress Classification Related Works .....	26
<b>CHAPTER THREE METHODOLOGY .....</b>	<b>33</b>
3.0 Introduction .....	33
3.1 Research Design.....	33
3.2 Speakers' selection.....	35
3.3 Experimental setup.....	35
3.4 Data collection .....	36
3.5 Speech corpus .....	37
3.6 Pre-Processing.....	38
3.6.1 Segmentation/ Endpoint Detection.....	38
3.6.2 Normalization.....	38
3.6.3 Pre-Emphasizing .....	39
3.6.4 Windowing .....	40
3.7 Feature extraction.....	40
3.8 Pattern classification .....	40
3.8.1 Training .....	41
3.8.2 Testing.....	42
3.9 Stressor.....	42
3.10 Classifier Settings .....	43
<b>CHAPTER FOUR PARAMETER ANALYSIS .....</b>	<b>44</b>
4.0 Introduction .....	44



4.1 Heart rate analysis .....	44
4.2 Fundamental frequency ( $f_0$ ) Analysis .....	47
4.3 Formants Analysis.....	49
4.4 Energy .....	52
4.5 Mel Frequency Cepstral Coefficients (MFCC) .....	54
4.5.1 Classification result by Gender .....	55
4.5.2 Classification result by words .....	58
4.5.2.1 Classification result by words with 13 MFCC .....	58
4.5.2.2 Classification result by words with 4 MFCC .....	61
4.6 Stress classification .....	65
<b>CHAPTER FIVE CONCLUSION AND FUTURE RESEARCH .....</b>	<b>67</b>
5.0 Introduction .....	67
5.1 Conclusion .....	67
5.3 Future research .....	71
REFERENCES.....	72

## List of Tables

<b>Table</b>	<b>Page</b>
Table 2.1: Classification of Stressor in Speech Production System.....	14
Table 2.2: Summary of related research studies.....	31
Table 3.1: Data collection procedure.....	36
Table 4.1: Summary of speaker's statistics.....	44
Table 4.2: Statistical measures of fundamental frequency of normal and stressed speech.....	47
Table 4.3: Statistical measures of first formant (F1) of normal and stressed speech.....	50
Table 4.4: Statistical measures of second formant (F2) of normal and stressed speech.....	51
Table 4.5: Statistical measures of third formant (F3) of normal and stressed speech.....	52
Table 4.6: Statistical measures of energy (dB) of normal and stressed speech.....	53
Table 4.7: Classification result of 13-MFCC for normal and stressed speech.....	55
Table 4.8: Classification result of 4-MFCC for normal and stressed speech.....	57
Table 4.9: Word based ANN classification result of 13-MFCC normal and stressed speech....	58
Table 4.10: Word based KNN classification result of 13-MFCC normal and stressed speech...59	
Table 4.11: Word based LDA classification result of 13-MFCC normal and stressed speech...60	
Table 4.12: Word based ANN classification result of 4-MFCC normal and stressed speech.....62	
Table 4.13: Word based KNN classification result of 4-MFCC normal and stressed speech....	63
Table 4.14: Word based LDA classification result of 4-MFCC normal and stressed speech....	64
Table 4.15: Stress Level and percentage of stressed speech.....	66

## List of Figures

<b>Figure</b>	<b>Page</b>
Figure 2.1: Modelling of changes caused by Emotion in Speech.....	12
Figure 2.2: MFCC Feature Extraction process.....	21
Figure 2.3: Block Diagram of Pattern Recognition Speech Recognizer.....	22
Figure 2.4: Architecture of a typical artificial neural network.....	24
Figure 3.1: Research Framework.....	34
Figure 3.3: Data Acquisition Process.....	37
Figure 3.2: Recording setup for the physical task stress of the UUM-CycleStress corpus.....	42
Figure 4.1: Average heart rate for neutral and stress speech for male and female speakers.....	46
Figure 4.2: Fundamental frequency against time for neutral and stress speech for male and female speakers .....	49
Figure 4.3: First formant (F1) against time for neutral and stress speech for male and female speakers.....	50
Figure 4.4: Second formant (F2) against time for neutral and stress speech for male and female speakers.....	51
Figure 4.5: Third formant (F3) against time for neutral and stress speech for male and female speakers.....	52
Figure 4.6: Energy against time for neutral and stress speech for male and female speaker.....	54
Figure 4.7: Classification Rate (CR %) chart of 13-MFCC for neutral and stress speech for male and female speakers.....	56
Figure 4.8: Classification Rate (CR %) chart of 4-MFCC for neutral and stress speech for male and female speakers.....	57
Figure 4.9: ANN Word based Classification Rate (CR %) chart of 13-MFCC for neutral and stress speech.....	59

Figure 4.10: KNN Word based Classification Rate (CR %) chart of 13-MFCC for neutral and stress speech.....	60
Figure 4.11: LDA Word based Classification Rate (CR %) chart of 13-MFCC for neutral and stress speech.....	61
Figure 4.12: ANN Word based Classification Rate (CR %) chart of 4-MFCC for neutral and stress speech.....	63
Figure 4.13: KNN Word based Classification Rate (CR %) chart of 4-MFCC for neutral and stress speech.....	64
Figure 4.14: LDA Word based Classification Rate (CR %) chart of 4-MFCC for neutral and stress speech.....	65
Figure 4.15: Percentage classification of stress level of stressed speech.....	66

## List of Acronyms

<b>A</b>	Age of the speakers
<b>ANN</b>	Artificial Neural Network
<b>ANS</b>	Automatic Nervous System
<b>ASR</b>	Automatic Speech Recognition
<b>BN</b>	Bayesian Network
<b>BPNN</b>	Back-propagation Neural Network
<b>BVP</b>	Blood Volume Pressure
<b>CB</b>	Critical Band
<b>CRs</b>	Classification Rates
<b>CV</b>	Consonant-Vowel
<b>DFT</b>	Discrete Fourier Transform
<b>DTW</b>	Discrete Wavelet Transform
<i>f<sub>0</sub></i>	Fundamental Frequency
<b>F1, F2, F3</b>	First, Second and Third Formants
<b>F, M</b>	Female, Male
<b>FFT</b>	Fast Fourier Transform
<b>FM</b>	Frequency Modulation
<b>FS, FN</b>	Female Stress, Female Neutral
<b>GMM</b>	Gaussian Mixture Model
<b>GSR</b>	Galvanic Skin Response
<b>HCI</b>	Human- Computer Interaction
<b>HCNN</b>	Hidden Control Neural Network
<b>HMM</b>	Hidden Markov Model
<b>HR</b>	Current Heart Rate

<b>KNN</b>	K-Nearest Neighbor
<b><i>l</i></b>	Exertion level
<b>LDA</b>	Linear Discriminant Analysis
<b>LM</b>	Levenberg-Marquardt back propagation
<b>LPC</b>	Linear Predictive Coding
<b>MAP</b>	Maximum A Posteriori
<b>MFCC</b>	Mel-Frequency Cepstral Coefficient
<b>MHR</b>	Maximum Heart Rate
<b>MLP</b>	Multi-Layer Perceptron
<b>MS, MN</b>	Male Stress, Male Neutral
<b>NN</b>	Neural Network
<b>PAD</b>	Pitch, Amplitude, Duration
<b>PLP</b>	Perceptual Linear Prediction
<b>RHR</b>	Normal Heart Rate
<b>RNN</b>	Recurrent Neural Network
<b>ROS</b>	Rate of Speech
<b>SLM</b>	Sound Level Meter
<b>ST</b>	Skin Temperature
<b>STT</b>	Speech to Text
<b>SUSAS</b>	Speech under Simulated and Actual Stress
<b>SVM</b>	Support Vector Machine
<b>TEO</b>	Teager Energy Operator
<b>VQ</b>	Vector Quantization
<b>WER</b>	World Error Rate

# CHAPTER ONE

## INTRODUCTION

### 1.0 Introduction

This section serves as a broad introduction to the study. It contains the background of the study, motivation, and problem statement. In addition, it also presents research questions and the objectives of the research, the scope and significance of the study.

### 1.1 Background

Contemporary life is filled with challenges, hassles, deadlines, frustrations, disappointments, and endless demands. The consequent of which might be stress. Stress has become a global phenomenon that is been experienced in our modern daily lives (Lu et al., 2012). For many people, stress is so commonplace – in traffic, markets, schools, or at work that it has become a way of life so much that ability to cope with stress is seen as a plus quality. While to some stress is a nightmare. Stress is not always bad, in small doses, it can help propel and motivate an individual under pressure to do better (Dhole & Gurjar, 2013). But being constantly running in emergency mode (stressed), the body and mind might pay the price. Affirming this is the studies report that stress might play a significant role in psychological and/or behavioural disorders like anxiety or depression (Dhole & Gurjar, 2013; Lu et al., 2012). Early detection of the signs and symptoms of stress is an antidote towards reducing its harmful effects and high cost of stress management efforts. Ability to detect stress and the level can be of use vital in applications that are stress sensitive such as

The contents of  
the thesis is for  
internal user  
only



## REFERENCES

- Abu Shariah, M., Ainon, R. N., Zainuddin, R., & Khalifa, O. O. (2007). *Human computer interaction using isolated-words speech recognition technology*. Paper presented at the Intelligent and Advanced Systems, 2007. ICIAS 2007. International Conference on.
- Acerro, A. (1990). *Acoustical and environmental robustness in automatic speech recognition* (Doctoral dissertation, Carnegie Mellon University).
- Adami, A. G., Lazzarotto, G. B., Foppa, E. F., & Couto Barone, D. (1999). *A comparison between features for a residential security prototype based on speaker identification with a model of artificial neural network*. Paper presented at the Computational Intelligence and Multimedia Applications, 1999. ICCIMA'99. Proceedings. Third International Conference on.
- Amuda, S., Boril, H., Sangwan, A., & Hansen, J. H. (2010). *Limited resource speech recognition for Nigerian English*. Paper presented at the Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on.
- Anusuya, M., & Katti, S. (2010). Speech recognition by machine, A review. *arXiv preprint arXiv:1001.2267*.
- Anusuya, M., & Katti, S. (2011). Classification Techniques used in Speech Recognition Applications: A Review. *Int. J. Comp. Tech. Appl*, 2(4), 910-954.
- Azmi, M. M., & Tolba, H. (2008). *Noise robustness using different acoustic units*. Paper presented at the Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on.
- Azmi, M. Y. S., Idayu, M. N., Roshidi, D., Yaakob, A., & Yaacob, S. (2012). Noise Robustness of Spectrum Delta (SpD) Features in Malay Vowel Recognition. In *Computer Applications for Communication, Networking, and Digital Contents* (pp. 270-277): Springer.

- Bakker, J., Holenderski, L., Kocielnik, R., Pechenizkiy, M., & Sidorova, N. (2012). *Stess@work: From measuring stress to its understanding, prediction and handling with personalized coaching*. Paper presented at the Proceedings of the 2nd ACM SIGHT International Health Informatics Symposium.
- Benzeghiba, M., De Mori, R., Deroo, O., Dupont, S., Erbes, T., Jouviet, D., et al. (2007). Automatic speech recognition and speech variability: A review. *Speech Communication, 49*, 763-786.
- Bou-Ghazale, S. E., & Hansen, J. H. (2000). A comparative study of traditional and newly proposed features for recognition of speech under stress. *Speech and Audio Processing, IEEE Transactions on, 8*(4), 429-442.
- Buchanan, C. R. (2005). Informatics Research Proposal-Modelling the Semantics of Sound. *School of Informatics, University of Edinburgh, United Kingdom*.
- Casale, S., Russo, A., Scebba, G., & Serrano, S. (2008). *Speech emotion classification using machine learning algorithms*. Paper presented at the Semantic Computing, 2008 IEEE International Conference on.
- Chee, L. S., Ai, O. C., Hariharan, M., & Yaacob, S. (2009, November). MFCC based recognition of repetitions and prolongations in stuttered speech using k-NN and LDA. In *Research and Development (SCOReD), 2009 IEEE Student Conference on* (pp. 146-149). IEEE.
- Cohen, S. E., Kessler, R. C., & Gordon, L. U. E. (1995). *Measuring stress: A guide for health and social scientists*: Oxford University Press.
- Costello, A., Abbas, M., Allen, A., Ball, S., Bell, S., Bellamy, R., et al. (2009). Managing the health effects of climate change: lancet and University College London Institute for Global Health Commission. *The Lancet, 373*(9676), 1693-1733.

- Daniel, J., & Martin, J. (2009). *Speech and Language Processing-Edition: 2. Prentice-Hall Inc, ISBN, 131873210, 2-16,489.*
- Deng, L., & Li, X. (2013). Machine learning paradigms for speech recognition: An overview. *Audio,Speech And Language Processing, IEEE Transactions on; 21(5), 1-30.*
- Dhole, N. P., & Gurjar, A. A. (2013). Detection of Speech under Stress: A Review. *International Journal of Engineering and Innovative Technology (IJEIT) on (Vol. 2, issue 10, pp. 36-38).*
- Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification. New York: John Wiley, Section, 10, 1.*
- Entwistle, M. S., & Adviser-Granaas, M. (2005). *Training methods and enrollment techniques to improve the performance of automated speech recognition systems under conditions of human exertion: University of South Dakota.*
- Godin, K. W., Hansen, J. H., Busso, C., & Katz, W. F. (2009). Classification based analysis of speech under physical task stress. *Master's thesis, University of Texas at Dallas, Richardson, TX.*
- Gray, S. S. (2006). *Speech Science Modeling for Automatic Accent and Dialect Classification. University of Colorado.*
- Han, Z., Lung, S., & Wang, J. (2012, March). A study on speech emotion recognition based on ccbc and neural network. In *Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on (Vol. 2, pp. 144-147). IEEE.*
- Hansen, J. H., Sangwan, A., & Kim, W. (2012). Speech under stress and Lombard effect: impact and solutions for forensic speaker recognition. In *Forensic Speaker Recognition (pp. 103-123): Springer.*
- Haykin, S. S. (2009). *Neural networks and learning machines (Vol. 3): Prentice Hall New York.*

- He, L., Lech, M., Maddage, M. C., & Allen, N. (2009, August). Stress detection using speech spectrograms and sigma-pi neuron units. In *Natural Computation, 2009. ICNC'09. Fifth International Conference on* (Vol. 2, pp. 260-264). IEEE.
- Hong, J.-H., Ramos, J., & Dey, A. K. (2012). *Understanding physiological responses to stressors during physical activity*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.
- Huang, X., Acero, A., & Hon, H. W. (2001). *Spoken language processing* (Vol. 18). Englewood Cliffs: Prentice Hall.
- Ibiyemi, T., & Akintola, A. (2012). Automatic Speech Recognition for Telephone Voice Dialling in Yorùbá. *International Journal of Engineering, 1*.
- Khalifa, O. O., El-Darymli, K. K., Abdullah, A.-H., & Daoud, J. I. (2013). Statistical Modeling for Speech Recognition.
- Khanna, P., & Kumar, M. S. (2011). Application of Vector Quantization in Emotion Recognition from Human Speech. In *Information Intelligence, Systems, Technology and Management* (pp. 118-125). Springer Berlin Heidelberg.
- Koolagudi, S. G., Kumar, N., & Rao, K. S. (2011, February). Speech emotion recognition using segmental level prosodic analysis. In *Devices and Communications (ICDeCom), 2011 International Conference on* (pp. 1-5). IEEE.
- Koolagudi, S. G., & Rao, K. S. (2012). Emotion recognition from speech using source, system, and prosodic features. *International Journal of Speech Technology, 15*(2), 265-289.
- Kurniawan, H., Maslov, A. V., & Pechenizkiy, M. (2013). *Stress detection from speech and Galvanic Skin Response signals*. Paper presented at the Computer-Based Medical Systems (CBMS), 2013 IEEE 26th International Symposium on.

- Lai, M., Chen, Y., Chu, M., Zhao, Y., & Hu, F. (2006). *A hierarchical approach to automatic stress detection in English sentences*. Paper presented at the Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on.
- Lemos, M. C. D., Valim, V., Zandonade, E., & Natour, J. (2010). Intensity level for exercise training in fibromyalgia by using mathematical models. *BMC musculoskeletal disorders*, *11*(1), 54.
- Levit, M., Huber, R., Batliner, A., & Noeth, E. (2001). *Use of prosodic speech characteristics for automated detection of alcohol intoxication*. Paper presented at the ISCA Tutorial and Research Workshop (ITRW) on Prosody in Speech Recognition and Understanding.
- Lu, H., Frauendorfer, D., Rabbi, M., Mast, M. S., Chittaranjan, G. T., Campbell, A. T., et al. (2012). *StressSense: Detecting stress in unconstrained acoustic environments using smartphones*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.
- Mao, X., Chen, L., & Fu, L. (2009, March). Multi-level speech emotion recognition based on HMM and ANN. In *Computer Science and Information Engineering, 2009 WRI World Congress on* (Vol. 7, pp. 225-229). IEEE.
- Martin, R. (2005). Statistical methods for the enhancement of noisy speech. In *Speech Enhancement* (pp. 43-65): Springer.
- Mohd Yusof, S., & Yaacob, S. (2008). Classification of Malaysian vowels using formant based features. *Journal of ICT*, *7*, 27-40.
- Murray, I. R., Baber, C., & South, A. (1996). Towards a definition and working model of stress and its effects on speech. *Speech Communication*, *20*(1), 3-12.

- Nakatani, T., Juang, B.-H., Kinoshita, K., & Miyoshi, M. (2005). *Harmonicity based dereverberation with maximum a posteriori estimation*. Paper presented at the Applications of Signal Processing to Audio and Acoustics, 2005. IEEE Workshop on.
- Narayana, M., & Kopparapu, S. (2009). *On the use of stress information in speech for speaker recognition*. Paper presented at the TENCON 2009-2009 IEEE Region 10 Conference.
- Neely, S. T., & Allen, J. B. (1979). Invertibility of a room impulse response. *The Journal of the Acoustical Society of America*, 66, 165.
- Patil, S. A., & Hansen, J. H. (2008). Detection of speech under physical stress: Model development, sensor selection, and feature fusion.
- Rabiner, L. R., & Schafer, R. W. (1978). *Digital processing of speech signals*(Vol. 100, p. 17). Englewood Cliffs: Prentice-hall.
- Rao, K. S., Koolagudi, S. G., & Vempada, R. R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, 16(2), 143-160.
- Razak, Z., Ibrahim, N. J., & Idna Idris, M. (2008). Quranic Verse recitation recognition module for support in J-QAF learning: A Review. *International Journal of Computer Science and Network Security (IJCSNS)*, 8(8), 207-216.
- Scherer, S., Hofmann, H., Lampmann, M., Pfeil, M., Rhinow, S., Schwenker, F., et al. (2008). *Emotion Recognition from Speech: Stress Experiment*. Paper presented at the LREC.
- Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., & Wendemuth, A. (2009). *Acoustic emotion recognition: A benchmark comparison of performances*. Paper presented at the Automatic Speech Recognition & Understanding, 2009. ASRU 2009. IEEE Workshop on.

- Shnayderman, I., & Katz-Leurer, M. (2013). An aerobic walking programme versus muscle strengthening programme for chronic low back pain: a randomized controlled trial. *Clinical rehabilitation*, 27(3), 207-214.
- Sigmund, M. (2010). *Changes in frequency spectrum of vowels due to psychological stress*. Paper presented at the Radioelektronika (RADIOELEKTRONIKA), 2010 20th International Conference.
- Sigmund, M., & Dostal, T. (2004). Analysis of emotional stress in speech. *Proc. IASTED AIA 2004*, 317-322.
- Siraj, F., Shahrul Azmi, M., Paulraj, M., & Yaacob, S. (2009). *Malaysian Vowel Recognition Based on Spectral Envelope Using Bandwidth Approach*. Paper presented at the Modelling & Simulation, 2009. AMS'09. Third Asia International Conference on.
- Sjölander, K., & Beskow, J. (2000). *Wavesurfer-an open source speech tool*. Paper presented at the INTERSPEECH.
- Sun, Z., Yuan, X., Bebis, G., & Louis, S. J. (2002). *Neural-network-based gender classification using genetic search for eigen-feature selection*. Paper presented at the Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on.
- Tanaka, H., Monahan, K. D., & Seals, D. R. (2001). Age-predicted maximal heart rate revisited. *Journal of the American College of Cardiology*, 37(1), 153-156.
- Torabi, S., AlmasGanj, F., & Mohammadian, A. (2008, December). Semi-Supervised Classification of Speaker's Psychological Stress. In *Biomedical Engineering Conference, 2008. CIBEC 2008. Cairo International* (pp. 1-4). IEEE.
- Wang, Y. (2009). *Speech recognition under stress*. Southern Illinois University Carbondale.

- Wu, M., & Wang, D. (2006). A two-stage algorithm for one-microphone reverberant speech enhancement. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14, 774-784.
- Yan, Q., Vaseghi, S., Rentzos, D., & Ho, C.-H. (2007). Analysis and synthesis of formant spaces of British, Australian, and American accents. *Audio, Speech, and Language Processing, IEEE Transactions on*, 15, 676-689.
- Yegnanarayana, B., & Murthy, P. S. (2000). Enhancement of reverberant speech using LP residual signal. *Speech and Audio Processing, IEEE Transactions on*, 8, 267-281.
- You, H., & Adviser-Alwan, A. (2009). *Robust automatic speech recognition algorithms for dealing with noise and accent*: University of California at Los Angeles.
- Zhai, J., & Barreto, A. (2008). Stress detection in computer users through non-invasive monitoring of physiological signals. *Blood*, 5, 0.
- Zhai, J., Barreto, A. B., Chin, C., & Li, C. (2005). *Realization of stress detection using psychophysiological signals for improvement of human-computer interactions*. Paper presented at the SoutheastCon, 2005. Proceedings. IEEE.
- Zhang, H. (2012). Emotional Speech Recognition Based on Syllable Distribution Feature Extraction. In *Foundations of Intelligent Systems* (pp. 415-420). Springer Berlin Heidelberg.
- Zhou, G., Hansen, J. H., & Kaiser, J. F. (2001). Nonlinear feature based classification of speech under stress. *Speech and Audio Processing, IEEE Transactions on*, 9(3), 201-216.