# Chapter 54
# Emotion Recognition From Speech Using Perceptual Filter and Neural Network

**Revathi A.**
*SASTRA University, India*

**Sasikaladevi N.**
*SASTRA University, India*

## ABSTRACT

*This chapter on multi speaker independent emotion recognition encompasses the use of perceptual features with filters spaced in Equivalent rectangular bandwidth (ERB) and BARK scale and vector quantization (VQ) classifier for classifying groups and artificial neural network with back propagation algorithm for emotion classification in a group. Performance can be improved by using the large amount of data in a pertinent emotion to adequately train the system. With the limited set of data, this proposed system has provided consistently better accuracy for the perceptual feature with critical band analysis done in ERB scale.*

## INTRODUCTION

Speech signal is considered as the acoustic signal obtained by exciting the vocal tract by quasi periodic pulses of air for voiced sounds and noise like excitation for unvoiced sounds. Speech utterances reveal the linguistic content, accent, slang and emotional state of a speaker. It is really cumbersome to recognize the emotions from speech with limited set of data. Emotion recognition from speech has found applications in call centers and unmanned control of risky processes. This system would be useful for treating the mentally retarded patients and patients with depression and anxiety. Web related services, retrieval of information and synthesis of data would use this automated emotion recognition system. These systems will find place in operating robots for the speech commands given by the emotional operator. Modulation spectral feature is used as a new feature by Siging Wu et.al (Wu, 2011) for emo-

tion recognition. Chi-Chun Lee et.al (Lee, 2011) have used hierarchical binary classifier and acoustic & statistical feature for emotion recognition. K. Sreenivasa Rao et.al (Rao, 2012) have used MFCC and GMM for recognizing emotions. Ankur Sapra et.al (Sapra, 2013) has used modified MFCC feature and NN classifier for emotion recognition. Shashidar G. Koolakudi et.al (Koolagudi, 2012) have used MFCC and GMM for speaker recognition in emotional environment. In this chapter on speaker independent emotion recognition, SVM is used to create templates for all emotions and system is evaluated with the speeches of a speaker not considered for training. Training speeches are converted into set of features and SVM models are developed as representative of emotions. During testing, group classification is done using minimum distance classifier and subsequently individual emotion classification is done in a group containing pertinent emotion models using linear binary classifier. Perceptual linear predictive cepstrum with critical band analysis done in BARK and ERB scale are used as features in this work and they provide complimentary evidence in assessing the performance of the system based on ANN modeling technique. ANN modeling technique is based on the selection of hidden layers and number of neurons in hidden layer. Weights between the layers are optimized using iterative procedure and output layer with two neurons to choose one of the two emotions in a group. This chapter also deals with the comparative analysis between the features and analysis is done comprehensively to assess the performance of the speaker independent and dependent emotion recognition system.

Affective computing has played a pivotal role in acting as an interface between humans and machines. Speech based emotion recognition system is difficult to be implemented because of the dataset which is containing limited set of speech utterances spoken limited set of speakers. Emotion recognition from speech is performed by using various databases. This chapter on multi speaker independent emotion recognition encompasses the use of perceptual features with filters spaced in Equivalent rectangular bandwidth (ERB) and BARK scale and vector quantization (VQ) classifier for classifying groups and artificial neural network with back propagation algorithm for emotion classification in a group. Performance can be improved by using the large amount of data in a pertinent emotion to adequately train the system. With the limited set of data, this proposed system has provided consistently better accuracy for the perceptual feature with critical band analysis done in ERB scale with overall accuracy as 76% and decision level fusion classification yielded 100% as accuracy for all emotions except FEAR and BOREDOM. Overall accuracy of the decision level fusion classifier is 78%. Speaker dependent emotion recognition system has provided 100% as accuracy for all the emotions for perceptual feature with critical band analysis done in ERB scale and perceptual linear predictive cepstrum has given 100% as accuracy for all emotions except anger and fear emotions.

## MATERIALS AND METHODS

## PLPC and ERBPLPC Extraction

The extraction of perceptual linear predictive cepstrum deals with computation of spectrum using FFT technique, wrap the spectrum along the BARK and ERB frequency scales, performing convolution between the warped spectrum and power spectrum of the simulated critical band masking curve, performing pre-emphasis by simulated loudness equalization, mapping between the intensity and the perceived loudness done by cube root compression, generation linear predictive (LP) coefficients and conversion into LP derived cepstrum. Probability is computed by counting the number of samples whose spectral

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/emotion-recognition-from-speech-using-perceptual-filter-and-neural-network/289004

# Related Content

### E-Government Avatar-Based Modeling and Development
Anara Kizabekovaand Vsevolod Chernyshenko (2020). *Avatar-Based Control, Estimation, Communications, and Development of Neuron Multi-Functional Technology Platforms (pp. 19-34).*
www.irma-international.org/chapter/e-government-avatar-based-modeling-and-development/244785

### VLSI Implementation of Neural Systems
Ashok Kumar Nagarajan, Kavitha Thandapani, Neelima K., Bharathi M., Dhamodharan Srinivasanand SathishKumar Selvaperumal (2023). *Neuromorphic Computing Systems for Industry 4.0 (pp. 94-116).*
www.irma-international.org/chapter/vlsi-implementation-of-neural-systems/326835

### Simultaneous Evolution of Network Architectures and Connection Weights in Artificial Neural Networks
Ruhul A. Sarkerand Hussein A. Abbass (2006). *Artificial Neural Networks in Finance and Manufacturing (pp. 28-42).*
www.irma-international.org/chapter/simultaneous-evolution-network-architectures-connection/5347

### Neural Network-Based Stock Market Return Forecasting Using Data Mining for Variable Reduction
David Encke (2006). *Artificial Neural Networks in Finance and Manufacturing (pp. 43-63).*
www.irma-international.org/chapter/neural-network-based-stock-market/5348

### Influential Nodes in Social Networks: Centrality Measures
Kousik Das, Rupkumar Mahapatra, Sovan Samantaand Anita Pal (2020). *Handbook of Research on Advanced Applications of Graph Theory in Modern Society (pp. 371-385).*
www.irma-international.org/chapter/influential-nodes-in-social-networks/235544