

Contents

- 1 Introduction** 1
 - 1.1 *Emotion: Psychological Perspective* 2
 - 1.2 *Emotion: Speech Signal Perspective*..... 3
 - 1.2.1 *Speech Production Mechanism* 4
 - 1.2.2 *Source Features* 5
 - 1.2.3 *System Features*..... 5
 - 1.2.4 *Prosodic Features* 7
 - 1.3 *Emotional Speech Databases* 8
 - 1.4 *Applications of Speech Emotion Recognition* 10
 - 1.5 *Issues in Speech Emotion Recognition* 10
 - 1.6 *Objectives and Scope of the Work* 11
 - 1.7 *Main Highlights of Research Investigations* 12
 - 1.8 *Brief Overview of Contributions to This Book* 12
 - 1.8.1 *Emotion Recognition Using Excitation Source Information* 12
 - 1.8.2 *Emotion Recognition Using Vocal Tract Information* 12
 - 1.8.3 *Emotion Recognition Using Prosodic Information* 13
 - 1.9 *Organization of the Book* 13
- 2 Speech Emotion Recognition: A Review** 15
 - 2.1 *Introduction* 15
 - 2.2 *Emotional Speech Corpora: A Review* 16
 - 2.3 *Excitation Source Features: A Review* 23
 - 2.4 *Vocal Tract System Features: A Review* 25
 - 2.5 *Prosodic Features: A Review* 27
 - 2.6 *Classification Models* 28
 - 2.7 *Motivation for the Present Work*..... 32
 - 2.8 *Summary of the Literature and Scope for the Present Work* 33
- 3 Emotion Recognition Using Excitation Source Information** 35
 - 3.1 *Introduction* 35
 - 3.2 *Motivation* 36

3.3	Emotional Speech Corpora.....	39
3.3.1	Indian Institute of Technology Kharagpur- Simulated Emotional Speech Corpus: IITKGP-SESC	40
3.3.2	Berlin Emotional Speech Database: Emo-DB.....	42
3.4	Excitation Source Features for Emotion Recognition	43
3.4.1	Higher-Order Relations Among LP Residual Samples	43
3.4.2	Phase of LP Residual Signal	44
3.4.3	Parameters of the Instants of Glottal Closure (Epoch Parameters)	46
3.4.4	Dynamics of Epoch Parameters at Syllable Level	50
3.4.5	Dynamics of Epoch Parameters at Utterance Level.....	51
3.4.6	Glottal Pulse Parameters	52
3.5	Classification Models	53
3.5.1	Auto-associative Neural Networks.....	53
3.5.2	Support Vector Machines.....	54
3.6	Results and Discussion	56
3.7	Summary	66
4	Emotion Recognition Using Vocal Tract Information.....	67
4.1	Introduction	67
4.2	Feature Extraction	68
4.2.1	Linear Prediction Cepstral Coefficients (LPCCs)	69
4.2.2	Mel Frequency Cepstral Coefficients (MFCCs)	70
4.2.3	Formant Features	71
4.3	Classifiers	72
4.3.1	Gaussian Mixture Models (GMM).....	72
4.4	Results and Discussion	73
4.5	Summary	77
5	Emotion Recognition Using Prosodic Information	79
5.1	Introduction	79
5.2	Prosodic Features: Importance in Emotion Recognition	80
5.3	Motivation	82
5.4	Extraction of Global and Local Prosodic Features.....	85
5.5	Results and Discussion	87
5.6	Summary	91
6	Summary and Conclusions	93
6.1	Summary of the Present Work	93
6.2	Contributions of the Present Work.....	95
6.3	Conclusions from the Present Work	95
6.4	Scope for Future Work	95

- A Linear Prediction Analysis of Speech** 99
 - A.1 The Prediction Error Signal 101
 - A.2 Estimation of Linear Prediction Coefficients 102
- B MFCC Features** 105
- C Gaussian Mixture Model (GMM)** 109
 - C.1 Training the GMMs 110
 - C.1.1 Expectation Maximization (EM) Algorithm 110
 - C.1.2 Maximum A Posteriori (MAP) Adaptation..... 111
 - C.2 Testing 113
- References**..... 115

Emotion Recognition using Speech Features

Rao, K.S.; Koolagudi, S.G.

2013, XII, 124 p. 30 illus., 6 illus. in color.,

ISBN: 978-1-4614-5143-3