

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction	1
1.2	Emotion from Psychological Perspective	2
1.3	Emotion from Speech Signal Perspective	3
1.3.1	Speech Production Mechanism	4
1.3.2	Source Features	4
1.3.3	System Features	5
1.3.4	Prosodic Features	5
1.4	Emotional Speech Databases	6
1.5	Applications of Speech Emotion Recognition	8
1.6	Issues in Speech Emotion Recognition	8
1.7	Objectives and Scope of the Book	9
1.8	Main Highlights of Research Investigations	10
1.9	Brief Overview of Contributions in this Book	10
1.9.1	Emotion Recognition using Spectral Features Extracted from Sub-syllabic Regions and Pitch Synchronous Analysis	10
1.9.2	Emotion Recognition using Global and Local Prosodic Features Extracted from Words and Syllables	11
1.9.3	Emotion Recognition using Combination of Features	11
1.9.4	Emotion Recognition on Real Life Emotional Speech Database	11
1.10	Organization of the Book	11
	References	14

<b>2</b>	<b>Robust Emotion Recognition using Pitch Synchronous and Sub-syllabic Spectral Features . . . . .</b>	<b>17</b>
2.1	Introduction. . . . .	17
2.2	Emotional Speech Corpora . . . . .	19
2.2.1	Indian Institute of Technology Kharagpur-Simulated Emotional Speech Corpus: IITKGP-SESC . . . . .	20
2.2.2	Berlin Emotional Speech Database: Emo-DB . . . . .	21
2.3	Feature Extraction . . . . .	22
2.3.1	Linear Prediction Cepstral Coefficients (LPCCs). . . . .	23
2.3.2	Mel Frequency Cepstral Coefficients . . . . .	24
2.3.3	Formant Features . . . . .	25
2.3.4	Extraction of Sub-syllabic Spectral Features . . . . .	26
2.3.5	Pitch Synchronous Analysis . . . . .	29
2.4	Classifiers . . . . .	31
2.4.1	Gaussian Mixture Models (GMM). . . . .	32
2.4.2	Auto-Associative Neural Networks . . . . .	32
2.5	Results and Discussion . . . . .	34
2.6	Summary . . . . .	44
	References . . . . .	45
<b>3</b>	<b>Robust Emotion Recognition using Sentence, Word and Syllable Level Prosodic Features . . . . .</b>	<b>47</b>
3.1	Introduction. . . . .	47
3.2	Prosodic Features: Importance in Emotion Recognition . . . . .	48
3.3	Motivation . . . . .	51
3.4	Extraction of Global and Local Prosodic Features . . . . .	52
3.4.1	Sentence Level Features . . . . .	53
3.4.2	Word and Syllable Level Features . . . . .	54
3.5	Results and Discussion . . . . .	56
3.5.1	Emotion Recognition Systems using Sentence Level Prosodic Features . . . . .	56
3.5.2	Emotion Recognition Systems using Word Level Prosodic Features . . . . .	60
3.5.3	Emotion Recognition Systems using Syllable Level Prosodic Features . . . . .	64
3.6	Summary . . . . .	68
	References . . . . .	69
<b>4</b>	<b>Robust Emotion Recognition using Combination of Excitation Source, Spectral and Prosodic Features . . . . .</b>	<b>71</b>
4.1	Introduction. . . . .	71
4.2	Feature Combination: A Study. . . . .	72
4.3	Emotion Recognition using Combination of Excitation Source and Vocal Tract System Features . . . . .	74

4.4	Emotion Recognition using Combination of Vocal Tract System and Prosodic Features . . . . .	75
4.5	Emotion Recognition using Combination of Excitation Source and Prosodic Features . . . . .	78
4.6	Emotion Recognition using Combination of Excitation Source, Vocal Tract System and Prosodic Features . . . . .	79
4.7	Summary . . . . .	83
	References . . . . .	84
<b>5</b>	<b>Robust Emotion Recognition using Speaking Rate Features. . . . .</b>	<b>85</b>
5.1	Introduction. . . . .	85
5.2	Motivation . . . . .	87
5.3	Two Stage Emotion Recognition System . . . . .	89
5.4	Gross Level Emotion Recognition . . . . .	90
5.5	Finer Level Emotion Recognition . . . . .	92
5.6	Summary . . . . .	94
	References . . . . .	94
<b>6</b>	<b>Emotion Recognition on Real Life Emotions . . . . .</b>	<b>95</b>
6.1	Introduction. . . . .	95
6.2	Real Life Emotion Speech Corpus . . . . .	96
6.3	Recognition Performance on Real Life Emotions. . . . .	97
6.4	Summary . . . . .	100
	References . . . . .	100
<b>7</b>	<b>Summary and Conclusions . . . . .</b>	<b>101</b>
7.1	Summary of the Present Work. . . . .	101
7.2	Contributions of the Present Work . . . . .	103
7.3	Conclusions from the Present Work . . . . .	104
7.4	Scope for Future Work . . . . .	105
	References . . . . .	107
	<b>Appendix A: MFCC Features . . . . .</b>	<b>109</b>
	<b>Appendix B: Gaussian Mixture Model (GMM). . . . .</b>	<b>113</b>

**Robust Emotion Recognition using Spectral and  
Prosodic Features**

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

2013, XII, 118 p. 37 illus., 15 illus. in color., Softcover

ISBN: 978-1-4614-6359-7