

Bernd Iser • Wolfgang Minker • Gerhard Schmidt

Bandwidth Extension of Speech Signals



Springer

Contents

1	Introduction	1
1.1	Scope of the Book	3
1.2	Nature of Speech Signals	4
1.3	Band Limited Speech Signals	5
1.4	Organization of the Book	7
2	Speech Generation	9
2.1	Human Speech Generation Process	9
2.2	Model for the Speech Generation Process	12
2.2.1	Excitation Signal	12
2.2.2	Vocal Tract Filter	13
2.3	Basic Approach for Bandwidth Extension	14
3	Analysis Techniques for Speech Signals	17
3.1	Linear Predictive Analysis	17
3.1.1	Autocorrelation Method	22
3.1.2	Covariance Method	23
3.2	Parametric Representations of the Spectral Envelope	24
3.2.1	Tube Model	24
3.2.2	AR-Coefficients	27
3.2.3	Cepstral Coefficients	28
3.2.4	MFCCs	31
3.2.5	Line Spectral Frequencies	33
3.3	Scalar Speech Features	35
3.3.1	Zero Crossing Rate	36
3.3.2	Gradient Index	36
3.3.3	Fundamental Frequency	36
3.3.4	Kurtosis	42
3.3.5	Spectral Centroid	42
3.3.6	Energy Based Features	44

3.4	Distance Measures	48
3.4.1	Log Spectral Deviation	48
3.4.2	RMS-Log Spectral Deviation	49
3.4.3	Cepstral Distance	49
3.4.4	Likelihood Ratio Distance	50
3.4.5	Itakura Distance	50
3.4.6	Itakura-Saito Distance	51
3.4.7	Other Spectral Distance Measures	51
4	Excitation Signal Extension	53
4.1	Estimation of the Narrowband Excitation Signal	54
4.2	Extension Using Non-Linear Characteristics	55
4.2.1	Half-Way Rectification	57
4.2.2	Full-Way Rectification	57
4.2.3	Quadratic Characteristic	60
4.2.4	Cubic Characteristic	60
4.2.5	Tanh Characteristic	60
4.2.6	Benesty Characteristic	61
4.2.7	Adaptive Quadratic Characteristic	61
4.3	Extension Using Spectral Shifting/Modulation Techniques	62
4.3.1	Fixed Spectral Shifting	62
4.3.2	Adaptive Spectral Shifting	63
4.4	Extension Using Function Generators	64
4.4.1	Sine Generators	64
4.4.2	White Noise Generators	64
4.5	Power Adjustment	65
4.6	Discussion	66
5	Broadband Spectral Envelope Estimation	67
5.1	Generation and Preparation of a Training Data Set	67
5.2	Estimation of the Narrowband Spectral Envelope	72
5.3	Estimation of the Broadband Spectral Envelope Using Codebooks	77
5.3.1	Codebook Training and Operation Using AR Coefficients	78
5.3.2	Codebook Training and Operation Using Cepstral Coefficients	79
5.3.3	Codebook Training and Operation Using MFCCs	84
5.3.4	Codebook Training and Operation Using LSFs	84
5.4	Estimation of the Broadband Spectral Envelope Using Neural Networks	85
5.4.1	Training and Operation Using Cepstral Coefficients	86
5.5	Estimation of the Broadband Spectral Envelope Using Linear Mapping	89
5.5.1	Training and Operation Using Cepstral Coefficients	89

5.6	Combined Approach Using Codebook Classification and Linear Mapping	90
5.6.1	Training and Operation Using Cepstral Coefficients	92
5.7	Discussion	94
6	Quality Evaluation	97
6.1	Evaluation of the Excitation Signal Extension	97
6.1.1	Objective Quality Criteria	98
6.1.2	Subjective Quality Criteria	99
6.1.3	Rank Correlation Between Objective and Subjective Results	106
6.2	Evaluation of the Envelope Estimation	107
6.2.1	Objective Quality Criteria	109
6.2.2	Subjective Quality Criteria	112
6.3	Discussion	114
7	Summary	115
A	Levinson–Durbin Recursion	119
B	LBG-Algorithm	123
C	Generalized Delta Rule or Standard Backpropagation	125
D	Stability Check	129
E	Feature Distributions	131
F	Evaluation Results	143
G	Abbreviations and Symbols	153
G.1	Abbreviations	153
G.2	Important Symbols	154
	List of Tables	157
	List of Figures	161
	References	167
	Index	179