

Contents

1	Introduction to ESOLA	1
1.1	Review of Speech Synthesis	1
1.2	Methods and Algorithms of Speech Synthesis	5
1.2.1	Articulatory Synthesis	5
1.2.2	Formant Synthesis	6
1.2.3	Linear Prediction Based Methods	9
1.2.4	Sinusoidal Models	10
1.2.5	Sinusoidal Analysis	11
1.2.6	Sinusoidal Synthesis	11
1.2.7	Concatenative Synthesis	13
1.2.8	PSOLA Methods	15
1.2.9	Other Techniques for Synthesis	16
1.3	Introduction to ESOLA	18
1.4	Organisation of the Book	19
	References	20
2	Epoch Synchronous Overlap Add (Esola) Algorithm	25
2.1	Introduction	25
2.2	Basic Principles of ESOLA	28
2.2.1	Partname: Sub-Phonemic Signal Inventory	29
2.3	Structure of Esola	35
2.3.1	Signal Units Representation	37
2.3.2	Word Number Bus: Word Segmentation	38
2.3.3	Syllable Number Bus: Syllable Breaking Algorithm	38
2.3.4	Special Emphasis Bus	39
2.3.5	Textual Language Processing (TLP) Unit	39
2.4	Speech Engine	40
2.4.1	Epoch Synchronous Overlap Add (ESOLA) Technique	40
2.4.2	Epoch Points for Voiced Speech Signals and Perceptual Pitch Period (PPP)	40

2.4.3	ESOLA Framework	43
2.4.4	Monotonic Properties	49
2.4.5	Properties Related to Peak	51
2.4.6	Properties Related to Valley	51
2.4.7	Pitch Modification Using Extended Bell Function	53
2.5	Preparation of Signal Dictionary	54
2.5.1	Recording	57
2.5.2	Pitch Normalization	57
2.5.3	Amplitude Normalization	58
2.5.4	Complexity Matching: Regeneration of signal	59
2.6	Synthesis Procedures	62
2.6.1	Rules for Token Generation	62
2.6.2	Synthesis Operations	63
2.6.3	Signal Processing Aspects	63
2.7	Esola and Other Concatenative Approaches	65
2.8	Conclusions and Discussion	67
	References	68
3	State Phase Analysis: PDA/VDA Algorithm	71
3.1	Introduction	71
3.2	State Phase Analysis	74
3.2.1	Pseudo Phonemic Labeling	79
3.2.2	Parameter Definitions	81
3.3	Classificatory Analysis	83
3.4	Pitch Extraction	87
3.4.1	Classification Algorithm	87
3.4.2	Experimental Details	87
3.5	Comparative Assessment of Pitch Extraction	90
3.5.1	Comparison of Pitch Data Obtained by State-phase Method	91
3.6	Classification Results	96
3.7	Analysis-Resynthesis Using State Phase Method	99
3.7.1	Extraction of Signal Elements	99
3.7.2	Extraction of Elements in Voiced Region	100
3.7.3	Extraction of Elements in Unvoiced Regions	101
3.7.4	Coding for Data Packet	101
3.7.5	Error Detection and Correction	103
3.7.6	Resynthesis Using Linear Interpolation	103
3.7.7	Decoding and Regeneration	105
3.7.8	Results	107
3.8	Discussion	109
	References	111

4	Phonological Rules for TTS	113
4.1	Introduction	113
4.2	Historical Background of SCB Phonology	115
4.3	Phones and Phonology of SCB	116
4.3.1	Compilation of the Phonological Rules for Bengali	119
4.3.2	Rule for ঐ (E)	121
4.3.3	Rules for ঞ (= J+N1)	121
4.3.4	Rules for ঙ (Y-Ligature)	121
4.3.5	Rules for ঞ (B-Ligature)	122
4.3.6	Rules for ঞ (M-Ligature)	122
4.3.7	Rule for ঞ (R-Ligature)	122
4.3.8	Rule for ঞ (M) and ঞ (N)	122
4.3.9	Rules for ঞ (SH), ঞ (S1) and ঞ (S)	123
4.3.10	Rule for Chandra Bindu (ঃ)	123
4.4	Architecture for G2P Conversion System	123
4.4.1	Structure of RDB Table	125
4.4.2	Generation of Forest from RDB Table	125
4.5	Software Implementation of Phonological Rules	128
4.6	Conclusions and Discussion	128
	Appendix	129
	References	132
5	Intonation Rules for Text Reading	135
5.1	Introduction	135
5.2	Simplification of Pitch Movement	139
5.3	Stylization	141
5.4	Perceptual Evaluation of Syllabic Stylization	144
5.4.1	F ₀ Modification	145
5.5	Perception Test	147
5.5.1	Results	147
5.5.2	Intonation Patterns for SCB	153
5.6	Method of Application in Synthesis	159
5.6.1	Finding of Word Intonation Pattern	160
5.6.2	Finding of Syllabic Intonation Pattern	164
5.6.3	Synthesis	164
5.7	Prosody	169
5.7.1	Duration Rules	172
5.8	Conclusions and Discussion	174
	References	175
6	Shimmer, Jitter and Complexity Perturbation	177
6.1	Introduction	177
6.2	Methodology	181

6.2.1	Glottal Cycle Detection	181
6.2.2	Relative Jitter and Shimmer	181
6.2.3	Complexity Perturbation (CP)	182
6.3	Experimental Procedures	183
6.3.1	Results and Discussion on Obtained Values.	183
6.4	Listening Test	186
6.5	Conclusions and Discussion	190
	References	191
	Appendix	193
	Epilogue	197

Epoch Synchronous Overlap Add (ESOLA)

A Concatenative Synthesis Procedure for Speech

Datta, A.K.

2018, XII, 197 p. 88 illus., 32 illus. in color., Hardcover

ISBN: 978-981-10-7015-0