

Contents

	Page
Abstract	(3)
Acknowledgement	(4)
Contents	(5)
List of Tables	(7)
List of Figures	(8)
Glossary of Abbreviations	(9)
Chapter	
1. Introduction	1
1.1 Motivations	2
1.2 Goals and Expected Outcome	3
1.3 Outline	3
2. Pitch Detection Algorithm	5
2.1 Backgrounds	5
2.1.1 Pre-processing of Pitch Tracking	5
2.1.2 Pitch Tracking Algorithm	6
2.1.3 Smoothing	9
2.1.4 Feature Extraction	10
2.2 Implementations	11
2.2.1 AMDF	11
2.2.2 Auto-correlation	12
2.2.3 Classification Framework	13
2.3 Experiments and Discussions	14
2.3.1 Experiments Setting	14
2.3.2 Experiments Results and Discussions	15
2.3.2.1 Voiced/Unvoiced Decision and Smoothing	15
2.3.2.2 Effects of Pre-processing	16
2.3.2.3 Feature Extraction	17
2.3.2.4 Classification	18
2.4 Summary	19
3. Configuration of Tone Feature	20
3.1 Backgrounds	20
3.1.1 Scaling	20
3.1.2 Normalization	22
3.2 Implementations	23

3.2.1	Framework	23
3.2.2	Feature Adaptation	23
3.3	Experiments and Discussions	24
3.3.1	Experiments Subject	24
3.3.2	Experiments Setting	24
3.3.3	Experiments Results and Discussions	25
3.3.3.1	Tone-critical Segment	25
3.3.3.2	Feature Setting	26
3.3.3.3	Frequency Scaling	27
3.3.3.4	Normalization	28
3.3.3.5	Confusion Analysis	30
3.4	Summary	30
4.	Final Consonants and Tone Classification	31
4.1	Final Consonant of Thai Language	31
4.2	Final consonants and Tone Classification	32
4.3	Experiments and Discussions	33
4.4	Summary	35
5.	2-stage NN Approaches for Tone Classification	36
5.1	The observation of biased and unbiased NN	36
5.2	Proposed Method	38
5.3	Experiments and Discussions	38
5.4	Summary	44
6.	Conclusions and Future Works	45
6.1	Summary	45
6.2	Future works	48
	Bibliography	49
	Appendix	53
	Vitae	56

List of Tables

Table	Page
2-1. Confusion-matrix of Tone Classification for Thai Digit	18
3-1. Tone-critical segment classification results	25
3-2. Feature Setting Classification Results	26
3-3. Classification Results of Scaling	28
3-4. Classification Results of Normalization	30
3-5. Confusion-matrix	30
4-1. Distribution of final consonants in speech database	32
4-2. Tone Classification with Two sub-classifier based on voiced and voiceless of final consonants	34
4-3. Tone Classification with Three sub-classifier based on distributions of final consonants	35
4-4. Tone Classification with four-sub-classifier based on final consonants	35
5-1. Tone classification using same number of training data for each tone	37
5-2. Confusion-matrix for biased NN	38
5-3. Distribution of NN Training Data	38
5-1. Classification confusion-matrix of tone0-tone1,4-tone2,3	41
5-2. Classification results for NN2	41
5-3. Classification results for NN3	42
5-4. Final Classification Results and Original Results	42
5-5. Confusion-matrix for NN0,1,3-2-4	43
5-6. Classification Results for tone0-1-3	43
5-7. Comparison of Final Classification Results	43
5-8. Classification Results of Tone0 output	44
5-9. Classification Results of Tone1 Output	44
5-10. Classification Results of Tone3 Output	45
5-11. Comparison of Final Classification Results	45
A1-1. Tone Classification Results of HTK Experiments	54
A3-1. Distribution of Tone Data	55

List of Figures

Figure	Page
2-1. $x(n)$, $\text{clc}[x(n)]$, $\text{sgn}[x(n)]$ and the auto-correlation	6
2-2. Block Diagram of the Coarse Pitch Detection using AMDF	12
2-3. Block Diagram of Pitch Detection Algorithm using Modified autocorrelation	13
2-4. Classification Framework for Pitch Detection Algorithms	13
2-5. Waveform of Thai Digit("07229")	14
2-6. Waveform of Mandarin Speech "hao" with 3 rd tone	14
2-7. Pitch Track Using Autocorrelation Method and AMDF	15
2-8. the Voice/Unvoiced Detection in Autocorrelation Method	16
2-9. Smoothing of "hao(3)" Pitch Contour	16
2-10. Effects of LPF in Auto-correlation method	17
2-11. The pitch pattern extracted	17
2-12. Four Discrete Legendre bases(Point: base 1 * : base 2 Δ : base 3 o: base 4)	18
2-13. Average F0 contours of the five Thai tones produced in isolation	19
3-1. Experiment's Framework	23
3-2. Feature Adaptation Framework	24
3-3. Pitch contour with final consonant and the vowel only	25
3-4. Pitch Contour and Tone Feature	26
3-5. Histogram of Scaling	27
3-6. Histogram of Two Speaker's Pitch	28
3-7. Histogram after Normalization	29
4-1. Framework of Tone classification using alphabet information	33
5-1. Framework of 2-layer NN	39
5-2. NN framework based on confusion information	40
5-3. Grouping of tone according the initial level	41
5-4. Tone Groups based on the shape of them	42

Glossary of Abbreviations

PDA	Pitch Detection Algorithm
AMDF	Average Magnitude Difference Function
ACF	Auto-correlation Function
V/UV	Voiced/Unvoiced
C_L	Clipping Threshold
clc	Center Clipping
NN	Neural Network
LPF	Low Pass Filter
LMS	Least Mean Square
JND	Just Noticeable Difference
HTK	Hidden Markov Model Toolkit
F0	Fundamental Frequency
CB	Critical Band
ERB	Equivalent Rectangle Bandwidth
MFCC	Mel-frequency Cepstrum Coefficients
Tone 0	Middle Tone in Thai Language
Tone 1	Low Tone in Thai Language
Tone 2	Falling Tone in Thai Language
Tone 3	High Tone in Thai Language
Tone 4	Rising Tone in Thai Language