

Contents

1	Introduction	1
1.1	Background	1
1.2	Motivation	2
1.3	Chronological Survey	4
1.3.1	Infant Period (~1970)	4
1.3.2	Foundation Period (1970s)	5
1.3.3	Exploring Period (1980s)	6
1.3.4	Transcending Period (1990–2005)	7
1.3.5	Regenerative Period (Since 2006)	9
1.4	Evaluation Metrics	11
1.4.1	Character Segmentation Rates	11
1.4.2	Recognition Rates	12
1.5	Organization	12
	References	14
2	HIT-MW Database	23
2.1	Introduction	23
2.2	Sampling Strategy	26
2.2.1	Writer Sampling	26
2.2.2	Text Sampling	27
2.3	Database Acquisition	27
2.3.1	Text Partition	27
2.3.2	Layout Design	28
2.3.3	Form Collection	30
2.4	Database Processing	30
2.4.1	Handwriting Digitalization	30
2.4.2	Image Preprocessing	30
2.4.3	Database Labeling	31
2.4.4	Text Line Extraction	32

2.5	Database Statistics	37
2.5.1	Basic Statistics	39
2.5.2	Erasing Statistics.	41
2.5.3	Miswriting Statistics	43
2.6	Application of HIT-MW Database.	44
2.7	Discussions and Conclusions.	45
	References	46
3	Integrated Segmentation-Recognition Strategy	49
3.1	Introduction	49
3.2	Reference Systems.	50
3.2.1	Segmentation-Based Baseline Systems.	50
3.2.2	Integrated Segmentation-Recognition Baseline System	53
3.3	Perceptron Learning of MQDF	54
3.3.1	Perceptron Learning	55
3.3.2	Dynamic Margin Regularization	57
3.3.3	Interpretation	58
3.4	Active Set Technique.	59
3.5	String Sample Synthesizing	60
3.5.1	String Sample Distortion	61
3.5.2	Verisimilitude Consideration	63
3.6	String-Level Training.	63
3.6.1	MCE Criterion	64
3.6.2	String-Level Training	64
3.7	Linguistic Constraints	65
3.8	Experiments	66
3.8.1	Databases.	66
3.8.2	Evaluation on Character Databases	67
3.8.3	Evaluation on HIT-MW Database.	73
3.9	Conclusions	76
	References	77
4	Segmentation-Free Strategy: Basic Algorithms	79
4.1	Introduction	79
4.2	System Overview	80
4.3	Sliding Windowing	80
4.4	Feature Representation.	82
4.5	HMM Training and Decoding.	83
4.6	Experiments	86
4.6.1	Evaluation on Segmentation-Free Systems	87
4.6.2	Comparison Between Segmentation-Based and Segmentation-Free Systems	93
4.6.3	Discussions	95

4.7	Conclusions	96
	References	97
5	Segmentation-Free Strategy: Advanced Algorithms	99
5.1	Introduction	99
5.2	Ensemble	100
5.2.1	Curve of Character Matching Rate	100
5.2.2	Serial Structure	101
5.2.3	Parallel Structure	102
5.3	Feature Enhancement	104
5.3.1	Enhanced FPF (enFPF)	104
5.3.2	Delta Features	106
5.4	Text Line Normalization	106
5.5	Synthesizing Samples	106
5.6	Distributed N-gram Training	109
5.7	MPE/MWE Training	110
5.8	Experiments	111
5.8.1	HIT-SW Database	111
5.8.2	Evaluation of enFPF	112
5.8.3	Evaluation of Ensemble Classifiers	114
5.8.4	Evaluation of Text Line Normalization	117
5.8.5	Evaluation of Delta Features	118
5.8.6	Evaluation of Synthetic Samples	118
5.8.7	Evaluation of MPE/MWE Training	120
5.8.8	Evaluation of Distributed N-gram Training	122
5.9	Conclusions	122
	References	123