

Contents

- 1 General Framework 1**
 - 1.1 Introduction 2
 - 1.2 Classical Pattern Recognition Paradigm 3
 - 1.2.1 Decision Theory and Pattern Recognition 7
 - 1.3 Interactive Pattern Recognition and Multimodal Interaction 9
 - 1.3.1 Using the Human Feedback Directly 11
 - 1.3.2 Explicitly Taking Interaction History into Account 12
 - 1.3.3 Interaction with Deterministic Feedback 12
 - 1.3.4 Interactive Pattern Recognition and Decision Theory 15
 - 1.3.5 Multimodal Interaction 16
 - 1.3.6 Feedback Decoding and Adaptive Learning 20
 - 1.4 Interaction Protocols and Assessment 21
 - 1.4.1 General Types of Interaction Protocols 22
 - 1.4.2 Left-to-Right Interactive–Predictive Processing 24
 - 1.4.3 Active Interaction 24
 - 1.4.4 Interaction with Weaker Feedback 25
 - 1.4.5 Interaction Without Input Data 25
 - 1.4.6 Assessing IPR Systems 26
 - 1.4.7 User Effort Estimation 26
 - 1.5 IPR Search and Confidence Estimation 27
 - 1.5.1 “Word” Graphs 28
 - 1.5.2 Confidence Estimation 33
 - 1.6 Machine Learning Paradigms for IPR 35
 - 1.6.1 Online Learning 36
 - 1.6.2 Active Learning 40
 - 1.6.3 Semi-Supervised Learning 41
 - 1.6.4 Reinforcement Learning 41
 - References 43
- 2 Computer Assisted Transcription: General Framework 47**
 - 2.1 Introduction 47
 - 2.2 Common Statistical Framework for HTR and ASR 48

2.3	Common Statistical Framework for CATTI and CATS	50
2.4	Adapting the Language Model	52
2.5	Search and Decoding Methods	52
2.5.1	Viterbi-Based Implementation	53
2.5.2	Word-Graph Based Implementation	54
2.6	Assessment Measures	58
	References	58
3	Computer Assisted Transcription of Text Images	61
3.1	Computer Assisted Transcription of Text Images: CATTI	62
3.2	CATTI Search Problem	63
3.2.1	Word-Graph-Based Search Approach	64
3.2.2	Word Graph Error-Correcting Parsing	64
3.3	Increasing Interaction Ergonomics in CATTI: PA-CATTI	66
3.3.1	Language Model and Search	68
3.4	Multimodal Computer Assisted Transcription of Text Images: MM-CATTI	70
3.4.1	Language Model and Search for MM-CATTI	73
3.5	Non-interactive HTR Systems	75
3.5.1	Main Off-Line HTR System Overview	75
3.5.2	On-Line HTR Subsystem Overview	79
3.6	Tasks, Experiments and Results	81
3.6.1	HTR Corpora	82
3.6.2	Results	88
3.7	Conclusions	94
	References	96
4	Computer Assisted Transcription of Speech Signals	99
4.1	Computer Assisted Transcription of Audio Streams	100
4.2	Foundations of CATS	100
4.3	Introduction to Automatic Speech Recognition	101
4.3.1	Speech Acquisition	101
4.3.2	Pre-process and Feature Extraction	102
4.3.3	Statistical Speech Recognition	102
4.4	Search in CATS	103
4.5	Word-Graph-Based CATS	103
4.5.1	Error Correcting Prefix Parsing	104
4.5.2	A General Model for Probabilistic Prefix Parsing	105
4.6	Experimental Results	107
4.6.1	Corpora	108
4.6.2	Error Measures	109
4.6.3	Experiments	109
4.6.4	Results	110
4.7	Multimodality in CATS	113
4.8	Experimental Results	115
4.8.1	Corpora	115

4.8.2	Experiments	116
4.9	Conclusions	116
	References	117
5	Active Interaction and Learning in Handwritten Text Transcription	119
5.1	Introduction	119
5.2	Confidence Measures	121
5.3	Adaptation from Partially Supervised Transcriptions	122
5.4	Active Interaction and Active Learning	122
5.5	Balancing Error and Supervision Effort	124
5.6	Experiments	126
5.6.1	User Interaction Model	126
5.6.2	Sequential Transcription Tasks	127
5.6.3	Adaptation from Partially Supervised Transcriptions	128
5.6.4	Active Interaction and Learning	129
5.6.5	Balancing User Effort and Recognition Error	130
5.7	Conclusions	132
	References	132
6	Interactive Machine Translation	135
6.1	Introduction	136
6.1.1	Statistical Machine Translation	136
6.2	Interactive Machine Translation	138
6.2.1	Interactive Machine Translation with Confidence Estimation	140
6.3	Search in Interactive Machine Translation	141
6.3.1	Word-Graph Generation	141
6.3.2	Error-Correcting Parsing	142
6.3.3	Search for n -Best Completions	143
6.4	Tasks, Experiments and Results	144
6.4.1	Pre- and Post-processing	145
6.4.2	Tasks	145
6.4.3	Evaluation Measures	145
6.4.4	Results	146
6.4.5	Results Using Confidence Information	148
6.5	Conclusions	149
	References	150
7	Multi-Modality for Interactive Machine Translation	153
7.1	Introduction	153
7.2	Making Use of Weaker Feedback	154
7.2.1	Non-explicit Positioning Pointer Actions	154
7.2.2	Interaction-Explicit Pointer Actions	156
7.3	Correcting Errors with Speech Recognition	157
7.3.1	Unconstrained Speech Decoding (DEC)	158
7.3.2	Prefix-Conditioned Speech Decoding (DEC-PREF)	159
7.3.3	Prefix-Conditioned Speech Decoding (IMT-PREF)	159
7.3.4	Prefix Selection (IMT-SEL)	160

7.4	Correcting Errors with Handwritten Text Recognition	160
7.5	Tasks, Experiments and Results	162
7.5.1	Results when Incorporating Weaker Feedback	162
7.5.2	Results for Speech as Input Feedback	163
7.5.3	Results for Handwritten Text as Input Feedback	165
7.6	Conclusions	166
	References	167
8	Incremental and Adaptive Learning for Interactive Machine Translation	169
8.1	Introduction	169
8.2	On-Line Learning	170
8.2.1	Concept of On-Line Learning	170
8.2.2	Basic IMT System	171
8.2.3	Online IMT System	172
8.3	Related Topics	174
8.3.1	Active Learning on IMT via Confidence Measures	174
8.3.2	Bayesian Adaptation	174
8.4	Results	175
8.5	Conclusions	176
	References	176
9	Interactive Parsing	179
9.1	Introduction	180
9.2	Interactive Parsing Framework	182
9.3	Confidence Measures in IP	184
9.4	IP in Left-to-Right Depth-First Order	186
9.4.1	Efficient Calculation of the Next Best Tree	187
9.5	IP Experimentation	188
9.5.1	User Simulation Subsystem	188
9.5.2	Evaluation Metrics	189
9.5.3	Experimental Results	190
9.6	Conclusions	191
	References	192
10	Interactive Text Generation	195
10.1	Introduction	195
10.1.1	Interactive Text Generation and Interactive Pattern Recognition	196
10.2	Interactive Text Generation at the Word Level	197
10.2.1	N -Gram Language Modeling	198
10.2.2	Searching for a Suffix	199
10.2.3	Optimal Greedy Prediction of Suffixes	199
10.2.4	Dealing with Sentence Length	203
10.2.5	Word-Level Experiments	204
10.3	Predicting at Character Level	205
10.3.1	Character-Level Experiments	205

10.4	Conclusions	207
	References	207
11	Interactive Image Retrieval	209
11.1	Introduction	209
11.2	Relevance Feedback for Image Retrieval	210
11.2.1	Probabilistic Interaction Model	210
11.2.2	Greedy Approximation Relevance Feedback Algorithm	213
11.2.3	A Simplified Version of GARF	214
11.2.4	Experiments	214
11.2.5	Image Feature Extraction	215
11.2.6	Baseline Methods	216
11.2.7	Discussion	218
11.3	Multimodal Relevance Feedback	218
11.3.1	Fusion by Refining	219
11.3.2	Early Fusion	219
11.3.3	Late Fusion	220
11.3.4	Proposed Approach: Dynamic Linear Fusion	222
11.3.5	Experiments	223
11.3.6	Discussion	225
	References	225
12	Prototypes and Demonstrators	227
12.1	Introduction	228
12.1.1	Passive, Left-to-Right Protocol	228
12.1.2	Passive, Desultory Protocol	230
12.1.3	Active Protocol	231
12.1.4	Prototype Evaluation	231
12.2	MM-IHT: Multimodal Interactive Handwritten Transcription	231
12.2.1	Prototype Description	232
12.2.2	Technology	233
12.2.3	Evaluation	235
12.3	IST: Interactive Speech Transcription	239
12.3.1	Prototype Description	240
12.3.2	Technology	241
12.3.3	Evaluation	242
12.4	IMT: Interactive Machine Translation	242
12.4.1	Prototype Description	243
12.4.2	Technology	244
12.4.3	Evaluation	246
12.5	ITG: Interactive Text Generation	246
12.5.1	Prototype Description	247
12.5.2	Technology	249
12.5.3	Evaluation	250
12.6	MM-IP: Multimodal Interactive Parsing	251
12.6.1	Prototype Description	251

12.6.2 Technology	254
12.6.3 Evaluation	255
12.7 GIDOC: GIMP-Based Interactive Document Transcription	255
12.7.1 Prototype Description	255
12.7.2 Technology	260
12.7.3 Evaluation	260
12.8 RISE: Relevant Image Search Engine	261
12.8.1 Prototype Description	261
12.8.2 Technology	262
12.8.3 Evaluation	264
12.9 Conclusions	264
References	265
Glossary	267
Index	271

Multimodal Interactive Pattern Recognition and
Applications

Toselli, A.H.; Vidal, E.; Casacuberta, F.

2011, XVI, 274 p., Hardcover

ISBN: 978-0-85729-478-4