Contents

Preface **V**

I Introduction 1

1 Fundamental Concept **3**
1.1 Motivation 5
1.2 Active and Assisted Living 6
1.3 Digital Medicine 9
1.4 Outline and Contribution 11
References 13

II Visual Scene Analysis 17

2 Large-Scale Multimedia Retrieval **19**
2.1 Hierarchical Organisation of Semantic Meanings 19
2.2 Concept Detection 24
2.2.1 Global versus Local Features 24
2.2.2 Feature Learning 28
2.3 Event Retrieval 31
2.3.1 Event Retrieval within Images/Shots 32
2.3.2 Event Retrieval over Shot Sequences 33
2.4 Conclusion and Future Trends 34
2.4.1 Reasoning 35
2.4.2 Uncertainties in Concept Detection 35
2.4.3 Adaptive Learning 36
References 39

3 Shape-Based Object Recognition **53**
3.1 Problem Statement and Motivation 53
3.2 Shape Representation 54
3.2.1 Survey of Related Methods 54
3.2.2 Coarse-grained Shape Representation 57
3.2.3 Fine-grained Shape Representation 58
3.3 Shape Matching 61
3.3.1 Survey of Related Methods 61
3.3.2 Shape Matching using Coarse-grained Features . . 63
3.3.3 Shape Matching using Fine-grained Features . . . 64
3.4 Experiments and Results 65
3.4.1 Shape Retrieval using Coarse-grained Features . . 65
3.4.2 Shape Retrieval using Fine-grained Features . . . 67
3.5 Conclusion and Future Trends 68
References 71

4 Moving Object Analysis for Video Interpretation **81**
4.1 Object Tracking in 2D Video 81
4.1.1 Survey of Related Approaches 82
4.1.2 Tracking-Learning-Detection 89
4.1.3 Tracking in Omnidirectional Video 90
4.1.4 Experiments and Results 92
4.2 3D Trajectory Extraction from 2D Video 93
4.2.1 RJ-MCMC Particle Filtering 94
4.2.2 Convoy Detection in Crowded Video 99
4.2.3 Experiments and Results 102
4.3 Conclusion and Future Trends 104
References 106

III Human Data Interpretation **111**

5 Physical Activity Recognition **113**
5.1 Atomic Activity Recognition 113
5.1.1 Survey of Related Approaches 115
5.1.2 Codebook Approach for Classification 118
5.1.3 Experiments and Results 124
5.2 Gait Recognition 129
5.2.1 Survey or Related Approaches 130

5.2.2 Spatiotemporal Representation of Gait 132
5.2.3 Experiments and Results 135
5.3 Conclusion and Future Trends 141
References 144

6 Cognitive Activity Recognition **157**
6.1 Definition, Taxonomy, Impact on Health 157
6.2 Sensing the Brain Activity 158
6.2.1 Electroencephalography 158
6.2.2 Electrooculography 159
6.2.3 Functional Magnetic Resonance Imaging 159
6.2.4 Functional Near-InfraRed Spectroscopy 159
6.3 Survey of Related Methods 159
6.4 Electrooculography-Based Approach 161
6.4.1 Cognitive Activity Recognition Method 161
6.4.2 Investigating Codewords 162
6.5 Application and Validation 163
6.5.1 Collecting a Dataset 163
6.5.2 Implementation Details 165
6.5.3 Results for Cognitive Activity Recognition 165
6.5.4 Results for Codewords Investigation 166
6.6 Conclusion and Future Trends 167
References 169

7 Emotion Recognition **173**
7.1 Automatic Recognition of Emotions 173
7.1.1 Definition and Taxonomy of Emotions 174
7.1.2 Existing Techniques for Emotion Recognition . . . 180
7.1.3 Emotion Recognition Challenges 182
7.2 Multimodal Emotion Recognition 186
7.2.1 Arousal/Valence Estimation 186
7.2.2 Basic Emotion Recognition 189
7.3 Approaches Based on Physiological Data 192
7.3.1 Stress Detection Using Hand-crafted Features . . 194
7.3.2 Codebook Approach for Feature Generation 196
7.3.3 Deep Neural Networks for Feature Generation . . 199
7.4 Conclusion and Future Trends 203
References 204

IV Conclusion **211**

8 Summary and Future Vision **213**
8.1 Visual Scene Analysis 213
8.2 Human Data Interpretation 215
8.3 Data-Driven Society 217
References 219

List of Figures **223**

List of Tables **227**