1 Introduction ................................................................. 1
  1.1 Objectives ............................................................. 1
  1.2 Multimedia Database Retrieval ...................................... 2
    1.2.1 Background ................................................. 2
    1.2.2 Challenges .................................................. 2
    1.2.3 The Development of Multimedia Database Retrieval Technology ....................................... 3
  1.3 Technology Perspective .............................................. 3
    1.3.1 Human Centered Search and Retrieval .................... 3
    1.3.2 Internet Scale Multimedia Analysis and Retrieval .......... 5
    1.3.3 Mobile Visual Search ....................................... 7
    1.3.4 Multimedia Retrieval in a Cloud Datacenter .............. 8
    1.3.5 Technologies of 2-D Video and 3-D Motion Database Retrieval ........................................ 10
  1.4 Application Perspective .............................................. 14
  1.5 Organization of the Book ............................................ 15

2 Kernel-Based Adaptive Image Retrieval Methods ...................... 17
  2.1 Introduction ........................................................... 17
  2.2 Kernel Methods in Adaptive Image Retrieval ...................... 18
    2.2.1 Adaptive Retrieval Framework ................................ 18
    2.2.2 Query Adaptation Method .................................. 19
    2.2.3 Metric Adaptation Method .................................. 20
    2.2.4 Query and Metric Adaptive Method ......................... 21
    2.2.5 Nonlinear Model-Based Adaptive Method .................. 23
  2.3 Single-Class Radial Basis Function Based Relevance Feedback . 24
    2.3.1 Center Selection ............................................ 24
    2.3.2 Width Selection ............................................. 26
    2.3.3 Experimental Result ........................................ 27
  2.4 Multi-Class Radial Basis Function Method ......................... 30
    2.4.1 Local Model Network ........................................ 34
## Contents

2.4.2 Learning Methods for the RBF Network .................. 35  
2.4.3 Adaptive Radial-Basis Function Network .................. 37  
2.4.4 Gradient-Descent Procedure ............................... 40  
2.4.5 Fuzzy RBF Network with Soft Constraint ................ 43  
2.4.6 Experimental Result ........................................ 44  

2.5 Bayesian Method for Fusion of Content and Context in Adaptive Retrieval ................................................. 47
2.5.1 Fusion of Content and Context ............................. 47  
2.5.2 Content-Based Likelihood Evaluation in Short-Term Learning ......................................................... 51  
2.5.3 Context Model in Long-Term Learning ................... 52  
2.5.4 Experimental Result ........................................ 54  

2.6 Summary .................................................................. 58  

3 Self-adaptation in Image and Video Retrieval ........................... 59
3.1 Introduction ........................................................... 59  
3.2 Pseudo Relevance Feedback Methods ............................... 60  
3.2.1 Re-ranking Domain ......................................... 60  
3.2.2 Self-organizing Tree Map .................................. 62  
3.2.3 Pseudo Labeling ............................................ 65  
3.2.4 Experimental Result ........................................ 67  

3.3 Re-ranking in Compressed Domains ................................ 69  
3.3.1 Descriptor in Discrete Cosine Transformation ............ 69  
3.3.2 Descriptor in Wavelet Based Coders ....................... 70  
3.3.3 Experimental Result ........................................ 74  

3.4 Region-Based Re-ranking Method .................................. 80  
3.4.1 Segmentation of the Region of Interest .................... 82  
3.4.2 Edge Flow Method ......................................... 82  
3.4.3 Knowledge-Based Automatic Region of Interest ........ 83  
3.4.4 Pseudo-Relevance Feedback with Region of Interest ..... 84  
3.4.5 Experimental Result ........................................ 84  

3.5 Video Re-ranking ..................................................... 87  
3.5.1 Template Frequency Model Implementing Bag-of-Words Model .................................................. 87  
3.5.2 Adaptive Cosine Network .................................. 89  
3.5.3 Experimental Result ........................................ 94  

3.6 Summary .................................................................. 99  

4 Interactive Mobile Visual Search and Recommendation at Internet Scale ............................................................ 101
4.1 Introduction ........................................................... 101  
4.2 BoW-Based Mobile Visual Search Using Various Context Information ................................................... 103  
4.2.1 The Bag-of-Word (BoW) Model ................................ 104  
4.2.2 Mobile Visual Search ........................................ 106
7.2 Video Parsing in Compressed Domain ......................................... 171
  7.2.1 Conventional Method .................................................. 171
  7.2.2 Twin Window Amplification Method ............................... 172
  7.2.3 Demonstration ....................................................... 174

7.3 News Video Retrieval ...................................................... 175
  7.3.1 Characterization of News Video Units ............................ 175
  7.3.2 Indexing and Retrieval of News Video ........................... 178
  7.3.3 Demonstration ....................................................... 180

7.4 Segmentation of Video Objects ......................................... 182
  7.4.1 Graph Cut Video Segmentation ................................... 182
  7.4.2 Object Segmentation ............................................... 187
  7.4.3 Histogram of Oriented Gradients ................................ 188

7.5 Segmentation of Face Object Under Illumination Variations .... 191
  7.5.1 Automatic Face Detection using Optimal Adaptive
        Correlation Method with Local
        Normalization .................................................. 193
  7.5.2 Experimental Result ............................................... 197

7.6 Play Event NFL Video Classification Using
MPEG-7 and MFCC Features .................................................. 200
  7.6.1 Localization of Play Events .................................... 201
  7.6.2 Classification of American Football Events ................. 205
  7.6.3 Experimental Results ............................................ 209

7.7 Summary ........................................................................... 210

8 Adaptive Retrieval in a P2P Cloud Datacenter ................. 213
  8.1 Introduction ............................................................. 213
  8.2 Distributed Database System ........................................ 214
    8.2.1 Cloud Datacenter .............................................. 214
    8.2.2 Application of a Multimedia Retrieval System
    in a P2P Datacenter ............................................... 215

  8.3 Adaptive Image Retrieval in a Self-organizing
Chord P2P Network .......................................................... 217
    8.3.1 System Architecture .......................................... 217
    8.3.2 Indexing of Nodes and Data Items on the
    Distributed Hash Table .......................................... 218
    8.3.3 Query Processing on the P2P Network ......................... 221

  8.4 Social Network Image Retrieval Using
Pseudo-Relevance Feedback .................................................. 227
    8.4.1 Social Network Discovery ..................................... 227
    8.4.2 Query Within the Social Network ............................ 229
    8.4.3 Pseudo Relevance Feedback in the Distributed
    Database System .................................................. 229
    8.4.4 Experimental Result ............................................ 233

  8.5 Video Re-ranking on the Social P2P Network ................. 237
    8.5.1 System Architecture ........................................... 238
8.5.3 Re-ranking Approach to P2P Video Retrieval ............. 239
8.5.4 Experimental Result .......................................... 243
8.6 Summary .................................................................... 246

9 Scalable Video Genre Classification and Event Detection ............ 247
9.1 Introduction .................................................................. 247
9.1.1 Overview ............................................................ 249
9.2 Video Representation and Genre Categorization .................. 252
9.2.1 Related Work .................................................... 252
9.2.2 Bottom-Up Codebook Generation .......................... 254
9.2.3 Low-Level Genre Categorization ................................ 256
9.3 High-Level Event Detection Using Middle-Level View as Agent ............................................................... 256
9.3.1 Related Work .................................................... 257
9.3.2 Middle-Level Unsupervised View Classification .......... 259
9.3.3 High-Level Event Detection .................................... 264
9.4 Experimental Result .................................................. 268
9.4.1 Genre Categorization Using K-Nearest Neighbor Classifier ................................................................. 270
9.4.2 Middle-Level View Classification Using Supervised SVM and Unsupervised PLSA .................... 273
9.4.3 Event Detection Using Coarse-to-Fine Scheme .......... 275
9.5 Summary .................................................................... 278

10 Audio-Visual Fusion for Film Database Retrieval and Classification ............................................................. 279
10.1 Introduction .................................................................. 279
10.2 Audio Content Characterization .................................... 280
10.2.1 Finite Mixture Model ......................................... 281
10.2.2 Laplacian Mixture Model and Parameter Estimation .... 282
10.2.3 Comparison of Gaussian Mixture Model and Laplacian Mixture Model ............................................ 284
10.2.4 Feature Extraction from Audio Signal ..................... 286
10.2.5 Performance of Video Retrieval Using Audio Indexing .. 287
10.3 Visual Content Characterization ..................................... 289
10.3.1 Visual Indexing Algorithm .................................... 289
10.3.2 Performance Comparison for Retrievals from Movie Database ............................................................ 290
10.4 Audio-Visual Fusion ................................................. 294
10.4.1 Decision Fusion Model ......................................... 295
10.4.2 Support Vector Machine Learning .......................... 296
10.4.3 Implementation of Support Vector Machine .......... 298
10.4.4 Results of Movie Clip Classification ...................... 300
10.5 Summary .................................................................... 303

11 Motion Database Retrieval with Application to Gesture Recognition in a Virtual Reality Dance Training System ................. 305
11.1 Introduction ........................................................... 305
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.2</td>
<td>Dance Training System</td>
<td>306</td>
</tr>
<tr>
<td>11.3</td>
<td>Spherical Self-organizing Map (SSOM)</td>
<td>309</td>
</tr>
<tr>
<td>11.4</td>
<td>Characterization of Dance Gesture Using Spherical Self-organizing Map</td>
<td>311</td>
</tr>
<tr>
<td>11.5</td>
<td>Trajectory Analysis</td>
<td>312</td>
</tr>
<tr>
<td>11.5.1</td>
<td>Sparse Code of Spherical Self-organizing Map</td>
<td>314</td>
</tr>
<tr>
<td>11.5.2</td>
<td>Posture Occurrence</td>
<td>315</td>
</tr>
<tr>
<td>11.5.3</td>
<td>Posture Transition and Posture Transition Sparse Code</td>
<td>316</td>
</tr>
<tr>
<td>11.5.4</td>
<td>Performance Comparison</td>
<td>317</td>
</tr>
<tr>
<td>11.6</td>
<td>Online Gesture Recognition and Segmentation</td>
<td>319</td>
</tr>
<tr>
<td>11.7</td>
<td>Trajectory Analysis on the Multicodebook SSOM Using Hidden Markov Model</td>
<td>321</td>
</tr>
<tr>
<td>11.7.1</td>
<td>The Self-organizing Map Distortion Measurement</td>
<td>322</td>
</tr>
<tr>
<td>11.7.2</td>
<td>The Hidden Markov Models of Gesture</td>
<td>325</td>
</tr>
<tr>
<td>11.7.3</td>
<td>Obtaining Learning Parameters</td>
<td>328</td>
</tr>
<tr>
<td>11.7.4</td>
<td>Experimental Result</td>
<td>329</td>
</tr>
<tr>
<td>11.8</td>
<td>Summary</td>
<td>333</td>
</tr>
</tbody>
</table>

References                                                                                      335
Multimedia Database Retrieval
Technology and Applications
Muneesawang, P.; Zhang, N.; Guan, L.
2014, XII, 350 p. 142 illus., 111 illus. in color., Hardcover
ISBN: 978-3-319-11781-2