Contents

Foreword xi
List of Figures xiii
List of Tables xvii
List of Contributors xix

1 Introduction 1
Raphaël Troncy, Benoit Huet and Simon Schenk

2 Use Case Scenarios 7
Werner Bailer, Susanne Boll, Oscar Celma,
Michael Hausenblas and Yves Raimond

2.1 Photo Use Case 8
2.1.1 Motivating Examples 8
2.1.2 Semantic Description of Photos Today 9
2.1.3 Services We Need for Photo Collections 10

2.2 Music Use Case 10
2.2.1 Semantic Description of Music Assets 11
2.2.2 Music Recommendation and Discovery 12
2.2.3 Management of Personal Music Collections 13

2.3 Annotation in Professional Media Production and Archiving 14
2.3.1 Motivating Examples 15
2.3.2 Requirements for Content Annotation 17

2.4 Discussion 18
Acknowledgements 19

3 Canonical Processes of Semantically Annotated Media Production 21
Lynda Hardman, Željko Obrenović and Frank Nack

3.1 Canonical Processes 22
3.1.1 Premeditate 23
3.1.2 Create Media Asset 23
3.1.3 Annotate 23
3.1.4 Package 24
3.1.5 Query
3.1.6 Construct Message
3.1.7 Organize
3.1.8 Publish
3.1.9 Distribute

3.2 Example Systems
3.2.1 CeWe Color Photo Book
3.2.2 SenseCam

3.3 Conclusion and Future Work

4 Feature Extraction for Multimedia Analysis
Rachid Benmokhtar, Benoit Huet,
Gaël Richard and Slim Essid

4.1 Low-Level Feature Extraction
4.1.1 What Are Relevant Low-Level Features?
4.1.2 Visual Descriptors
4.1.3 Audio Descriptors

4.2 Feature Fusion and Multi-modality
4.2.1 Feature Normalization
4.2.2 Homogeneous Fusion
4.2.3 Cross-modal Fusion

4.3 Conclusion

5 Machine Learning Techniques for Multimedia Analysis
Slim Essid, Marine Campedel, Gaël Richard, Tomas Piatrik,
Rachid Benmokhtar and Benoit Huet

5.1 Feature Selection
5.1.1 Selection Criteria
5.1.2 Subset Search
5.1.3 Feature Ranking
5.1.4 A Supervised Algorithm Example

5.2 Classification
5.2.1 Historical Classification Algorithms
5.2.2 Kernel Methods
5.2.3 Classifying Sequences
5.2.4 Biologically Inspired Machine Learning Techniques

5.3 Classifier Fusion
5.3.1 Introduction
5.3.2 Non-trainable Combiners
5.3.3 Trainable Combiners
5.3.4 Combination of Weak Classifiers
5.3.5 Evidence Theory
5.3.6 Consensual Clustering
5.3.7 Classifier Fusion Properties

5.4 Conclusion
6 Semantic Web Basics
Eyal Oren and Simon Schenk
6.1 The Semantic Web
6.2 RDF
6.2.1 RDF Graphs
6.2.2 Named Graphs
6.2.3 RDF Semantics
6.3 RDF Schema
6.4 Data Models
6.5 Linked Data Principles
6.5.1 Dereferencing Using Basic Web Look-up
6.5.2 Dereferencing Using HTTP 303 Redirects
6.6 Development Practicalities
6.6.1 Data Stores
6.6.2 Toolkits

7 Semantic Web Languages
Antoine Isaac, Simon Schenk and Ansgar Scherp
7.1 The Need for Ontologies on the Semantic Web
7.2 Representing Ontological Knowledge Using OWL
7.2.1 OWL Constructs and OWL Syntax
7.2.2 The Formal Semantics of OWL and its Different Layers
7.2.3 Reasoning Tasks
7.2.4 OWL Flavors
7.2.5 Beyond OWL
7.3 A Language to Represent Simple Conceptual Vocabularies: SKOS
7.3.1 Ontologies versus Knowledge Organization Systems
7.3.2 Representing Concept Schemes Using SKOS
7.3.3 Characterizing Concepts beyond SKOS
7.3.4 Using SKOS Concept Schemes on the Semantic Web
7.4 Querying on the Semantic Web
7.4.1 Syntax
7.4.2 Semantics
7.4.3 Default Negation in SPARQL
7.4.4 Well-Formed Queries
7.4.5 Querying for Multimedia Metadata
7.4.6 Partitioning Datasets
7.4.7 Related Work

8 Multimedia Metadata Standards
Peter Schallauer, Werner Bailer, Raphaël Troncy and Florian Kaiser
8.1 Selected Standards
8.1.1 MPEG-7
8.1.2 EBU P_Meta
8.1.3 SMPTE Metadata Standards
8.1.4 Dublin Core 133
8.1.5 TV-Anytime 134
8.1.6 METS and VRA 134
8.1.7 MPEG-21 135
8.1.8 XMP, IPTC in XMP 135
8.1.9 EXIF 136
8.1.10 DIG35 137
8.1.11 ID3/MP3 137
8.1.12 NewsML G2 and rNews 138
8.1.13 W3C Ontology for Media Resources 138
8.1.14 EBUCore 139

8.2 Comparison 140
8.3 Conclusion 143

9 The Core Ontology for Multimedia 145

Thomas Franz, Raphaël Troncy and Miroslav Vacura

9.1 Introduction 145
9.2 A Multimedia Presentation for Granddad 146
9.3 Related Work 149
9.4 Requirements for Designing a Multimedia Ontology 150
9.5 A Formal Representation for MPEG-7 150
9.5.1 DOLCE as Modeling Basis 151
9.5.2 Multimedia Patterns 151
9.5.3 Basic Patterns 155
9.5.4 Comparison with Requirements 157
9.6 Granddad’s Presentation Explained by COMM 157
9.7 Lessons Learned 159
9.8 Conclusion 160

10 Knowledge-Driven Segmentation and Classification 163

Thanos Athanasiadis, Phivos Mylonas, Georgios Th. Papadopoulos, Vasileios Mezaris, Yannis Avrithis, Ioannis Kompatsiaris and Michael G. Strintzis

10.1 Related Work 164
10.2 Semantic Image Segmentation 165
10.2.1 Graph Representation of an Image 165
10.2.2 Image Graph Initialization 165
10.2.3 Semantic Region Growing 167
10.3 Using Contextual Knowledge to Aid Visual Analysis 170
10.3.1 Contextual Knowledge Formulation 170
10.3.2 Contextual Relevance 173
10.4 Spatial Context and Optimization 177
10.4.1 Introduction 177
10.4.2 Low-Level Visual Information Processing 177
10.4.3 Initial Region-Concept Association 178
10.4.4 Final Region-Concept Association 179
10.5 Conclusions 181
## 11 Reasoning for Multimedia Analysis 183

*Nikolaos Simou, Giorgos Stoilos, Carsten Saathoff, Jan Nemrava, Vojtěch Svátek, Petr Berka and Vassilis Tzouvaras*

11.1 Fuzzy DL Reasoning 184

*11.1.1 The Fuzzy DL \textit{f-SHIN} 184*

*11.1.2 The Tableaux Algorithm 185*

*11.1.3 The FiRE Fuzzy Reasoning Engine 187*

11.2 Spatial Features for Image Region Labeling 192

*11.2.1 Fuzzy Constraint Satisfaction Problems 192*

*11.2.2 Exploiting Spatial Features Using Fuzzy Constraint Reasoning 193*

11.3 Fuzzy Rule Based Reasoning Engine 196

11.4 Reasoning over Resources Complementary to Audiovisual Streams 201

## 12 Multi-Modal Analysis for Content Structuring and Event Detection 205

*Noel E. O’Connor, David A. Sadlier, Bart Lehane, Andrew Salway, Jan Nemrava and Paul Buitelaar*

12.1 Moving Beyond Shots for Extracting Semantics 206

12.2 A Multi-Modal Approach 207

12.3 Case Studies 207

12.4 Case Study 1: Field Sports 208

*12.4.1 Content Structuring 208*

*12.4.2 Concept Detection Leveraging Complementary Text Sources 213*

12.5 Case Study 2: Fictional Content 214

*12.5.1 Content Structuring 215*

*12.5.2 Concept Detection Leveraging Audio Description 219*

12.6 Conclusions and Future Work 221

## 13 Multimedia Annotation Tools 223

*Carsten Saathoff, Krishna Chandramouli, Werner Bailer, Peter Schallauer and Raphaël Troncy*

13.1 State of the Art 224

13.2 SVAT: Professional Video Annotation 225

*13.2.1 User Interface 225*

*13.2.2 Semantic Annotation 228*

13.3 KAT: Semi-automatic, Semantic Annotation of Multimedia Content 229

*13.3.1 History 231*

*13.3.2 Architecture 232*

*13.3.3 Default Plugins 234*

*13.3.4 Using COMM as an Underlying Model: Issues and Solutions 234*

*13.3.5 Semi-automatic Annotation: An Example 237*

13.4 Conclusions 239
14  Information Organization Issues in Multimedia Retrieval Using Low-Level Features 241
   Frank Hopfgartner, Reede Ren, Thierry Urruty and Joemon M. Jose

14.1  Efficient Multimedia Indexing Structures 242
   14.1.1  An Efficient Access Structure for Multimedia Data 243
   14.1.2  Experimental Results 245
   14.1.3  Conclusion 249

14.2  Feature Term Based Index 249
   14.2.1  Feature Terms 250
   14.2.2  Feature Term Distribution 251
   14.2.3  Feature Term Extraction 252
   14.2.4  Feature Dimension Selection 253
   14.2.5  Collection Representation and Retrieval System 254
   14.2.6  Experiment 256
   14.2.7  Conclusion 258

14.3  Conclusion and Future Trends 259

Acknowledgement 259

15  The Role of Explicit Semantics in Search and Browsing 261
   Michiel Hildebrand, Jacco van Ossenbruggen and Lynda Hardman

15.1  Basic Search Terminology 261

15.2  Analysis of Semantic Search 262
   15.2.1  Query Construction 263
   15.2.2  Search Algorithm 265
   15.2.3  Presentation of Results 267
   15.2.4  Survey Summary 269

15.3  Use Case A: Keyword Search in ClioPatria 270
   15.3.1  Query Construction 270
   15.3.2  Search Algorithm 270
   15.3.3  Result Visualization and Organization 273

15.4  Use Case B: Faceted Browsing in ClioPatria 274
   15.4.1  Query Construction 274
   15.4.2  Search Algorithm 276
   15.4.3  Result Visualization and Organization 276

15.5  Conclusions 277

16  Conclusion 279
   Raphael Troncy, Benoit Huet and Simon Schenk

References 281

Author Index 301

Subject Index 303