

Contents

Preface	v
1 Video Search.....	1
1.1 Introduction	1
1.2 Addressing the Opportunity.....	2
1.3 Classification of Web Video Sites.....	5
1.3.1 Content Originators and Traditional Broadcasters	5
1.3.2 Aggregators	6
1.3.3 Download	6
1.3.4 Sharing.....	6
1.3.5 Application Specific	7
1.3.6 Other Video Systems	7
1.4 Classification of Video Sources.....	8
1.4.1 Webcams / Security	9
1.4.2 Video Telephony / Teleconferencing	9
1.4.3 Industrial / Academic / Medical	9
1.4.4 User Generated Content.....	10
1.4.5 Public Access and Government (PEG) Content	10
1.4.6 Enterprise Content	10
1.4.7 Rushes, Raw Footage	11
1.4.8 News	11
1.4.9 Advertising	11
1.4.10 Episodic TV Programming	11
1.4.11 Feature Films	12
1.4.12 Content Value	12
1.5 Challenges of Video Search.....	13
1.5.1 Acquisition	14
1.5.2 Media File Formats.....	15
1.5.3 Data Transport	16
1.5.4 Browsing.....	16
1.5.5 Duplication	17
1.5.6 Ranking and Indexing.....	17
1.6 Advantages of Video Search over Text.....	18

1.6.1 Applications.....	18
1.6.2 Metadata	19
1.7 Metadata vs. Content	19
1.7.1 Content-based retrieval.....	19
1.8 Conclusion	20
References	21
2 Video Data Sources and Applications.....	23
2.1 Introduction	23
2.1.1 Evolution of Digital Media Metadata.....	23
2.1.2 Consumer Video Metadata	24
2.1.3 Metadata Loss.....	24
2.1.4 Metadata Standards	25
2.1.5 Dublin Core	26
2.1.6 MPEG-7.....	27
2.1.7 MPEG-21.....	27
2.2 Essential Media Metadata.....	29
2.2.1 Embed Global Metadata	29
2.2.2 Elementary Metadata.....	29
2.3 Metadata for Personal Media Collections.....	31
2.3.1 Consumer Media Libraries	31
2.3.2 UPnP Forum	33
2.3.3 MP3 ID3	33
2.3.4 3GP / QuickTime / MP4.....	34
2.3.5 Metadata Services.....	34
2.3.6 Content Identification.....	36
2.3.7 Recorded Television.....	37
2.4 Media Syndication: RSS Content Description	39
2.4.1 Content Syndication	39
2.4.2 Media Enclosures	39
2.4.3 Podcasts	41
2.4.4 RSS for Content Ingest.....	42
2.4.5 MediaRSS.....	43
2.5 Metadata for Broadcast Television.....	43
2.5.1 Electronic Programming Guide (EPG).....	44
2.5.2 Extended Data Service (XDS).....	46
2.5.3 Program and System Identifier Protocol (PSIP).....	47
2.6 Metadata for Video on Demand	47
2.6.1 Introduction	47
2.6.2 Cable Labs	49
2.7 Production Metadata.....	50
2.8 Timed Text Formats	51

2.8.1 Introduction	51
2.8.2 Synchronization Precision and Resolution	52
2.8.3 Transcripts	53
2.8.4 Closed Captions.....	54
2.8.5 Synchronized Accessible Media Interchange	55
2.8.6 Metadata from Social Sources	55
2.8.7 Metadata Issues.....	55
2.9 Conclusion	56
References	56
3 Internet Video	59
3.1 Introduction	59
3.2 Digital Video	59
3.2.1 Aspect Ratio	59
3.2.2 Luminance and Chrominance Resolution.....	61
3.2.3 Video Compression	62
3.3 Internet Protocol Media Systems.....	66
3.3.1 Transport.....	66
3.3.2 Searching VoD vs. Live.....	67
3.3.3 IPTV	68
3.3.4 Rights Management.....	70
3.3.5 Redirector Files	70
3.3.6 Layered Encoding.....	73
3.3.7 Illustrated Audio	73
3.4 Media Captioning	74
3.5 Conclusion	75
References	76
4 Video Search Engine Systems.....	77
4.1 Introduction	77
4.2 Content Acquisition	78
4.2.1 Metadata Normalization	78
4.2.2 User Contributed.....	79
4.2.3 Syndicated Contribution.....	80
4.2.4 Broadcast Acquisition.....	81
4.3 Content Processing	82
4.3.1 Asset Management	82
4.4 Retrieval.....	84
4.5 User Perspectives.....	85
4.5.1 Interaction States	85
4.5.2 Granularity of Search Results Representation	87
4.6 Factors Concerning Scalability.....	88

- 4.6.1 Introduction 88
- 4.6.2 Acquisition 89
- 4.6.3 Processing..... 89
- 4.6.4 Storage..... 90
- 4.6.5 Retrieval 91
- 4.7 Retrieval Interfaces..... 92
- 4.8 Typical System Features..... 93
- 4.9 Conclusion..... 94
- References 94

- 5 Media Processing 97**
- 5.1 Introduction 97
- 5.2 Feature Extraction..... 99
- 5.3 Media Segmentation 100
- 5.4 Clustering, Structure Generation 101
- 5.5 Real-Time Processing..... 103
- 5.6 Systems Issues and Architectures..... 103
- 5.7 Conclusion 104
- References 105

- 6 Video Processing 107**
- 6.1 Introduction 107
- 6.2 Shot Boundary Determination 108
 - 6.2.1 Feature Extraction 110
 - 6.2.2 Shot Boundary Detectors..... 111
 - 6.2.3 Fusion of Detector Results 117
 - 6.2.4 Evaluation Results 117
- 6.3 Representative Image Selection..... 118
- 6.4 Face Detection 121
- 6.5 Face Recognition 126
- 6.6 Video Optical Character Recognition..... 129
- 6.7 Concept Detection 131
 - 6.7.1 Color Feature 133
 - 6.7.2 Texture Feature..... 133
 - 6.7.3 Edge Feature 135
- 6.8 Video Browsing..... 135
- 6.9 Conclusion 140
- References 141

- 7 Audio Processing..... 145**
- 7.1 Introduction 145
- 7.2 Audio Signal and Its Representation 146

7.3 Audio Features.....	148
7.3.1 Frame-Level Features	148
7.3.2 Clip-Level Features	154
7.4 Audio Segmentation	156
7.4.1 Speaker Segmentation	157
7.4.2 Audio Scene Segmentation.....	158
7.5 Audio Content Categorization	160
7.5.1 Speaker Recognition.....	160
7.5.2 Audio Scene Detection	162
7.5.3 Music Genre Classification	163
7.6 Speech Recognition	164
7.7 Audio Query and Browsing Techniques.....	166
7.7.1 SpeechLogger	167
7.7.2 Query by Example	171
7.8 Conclusion	172
References	173
8 Text Processing	177
8.1 Introduction	177
8.2 Story Segmentation.....	178
8.2.1 Cue Phrases	178
8.2.2 Cosine Similarity	179
8.2.3 Dynamic Programming.....	181
8.2.4 Topic Classification.....	183
8.3 Named Entity Extraction	183
8.3.1 Rule Based NEE	184
8.3.2 Data Driven NEE.....	185
8.3.3 NEE Tools	186
8.4 Part-of-Speech Tagging.....	187
8.5 Capitalization.....	189
8.5.1 Linguistic Processing Architecture.....	191
8.5.2 Web Document Collection	191
8.5.3 Text Capitalization Algorithm.....	192
8.6 Information Retrieval.....	194
8.6.1 Stemming.....	194
8.6.2 Term Weighting.....	195
8.6.3 Ranking.....	196
8.7 Text Summarization	197
8.7.1 Keyword Extraction.....	199
8.8 Conclusion	201
References	201

9 Multimodal Processing	203
9.1 Introduction	203
9.2 Case Studies.....	205
9.2.1 Closed Caption Alignment	205
9.2.2 Multimodal News Story Segmentation.....	209
9.2.3 Major Cast Detection.....	214
9.3 Conclusion	217
References	217
10 Research Systems	221
10.1 Introduction	221
10.2 Academic and Industrial Research	222
10.3 Early Internet Deployments.....	226
10.3.1 SpeechBot.....	226
10.3.2 StreamSage	227
10.3.3 SingingFish.....	227
10.4 Selected Commercial Systems.....	228
10.4.1 Virage and Convera	228
10.4.2 Nexidia (FastTalk).....	228
10.5 Resources: Datasets, Evaluations, Conferences	229
10.6 Media Monitoring Deployments.....	231
10.7 Case Study: AT&T MIRACLE	232
10.7.1 Introduction	232
10.7.2 System Architecture	232
10.7.3 Collections.....	233
10.7.4 Data Organization.....	235
10.7.5 Acquisition / Ingest.....	236
10.7.6 Content Processing	238
10.7.7 Real-time processing	239
10.7.8 Query Engine.....	239
10.7.9 Applications.....	240
10.7.10 Performance.....	240
10.8 Conclusion	242
References	242
11 Current Trends in Video Search	247
11.1 Introduction	247
11.2 Video Production.....	248
11.2.1 Metadata Retention.....	248
11.2.2 Multiple Distribution Channels	248
11.2.3 Mobisodes and Webisodes	249
11.3 Video Distribution	249

11.3.1 Streaming Protocols.....	250
11.3.2 Electronic Sell Through.....	250
11.3.3 Peer-to-peer Delivery	251
11.3.4 Managed Download.....	251
11.3.5 Syndication	252
11.4 The Video Web and User Interaction	252
11.4.1 Web-Based Editing.....	252
11.4.2 Media Browsing	252
11.4.3 Social Tagging.....	253
11.4.4 Dynamic Interfaces.....	253
11.4.5 Video Blogs (vlogs).....	254
11.4.6 Integrated Collections.....	254
11.5 Television Technology and Consumption	254
11.5.1 Proliferation of Channels.....	255
11.5.2 Live to Time Shifted.....	255
11.5.3 Mobile Consumption	255
11.6 Trends in Media Devices	256
11.6.1 Increased Media Capabilities.....	256
11.6.2 Increasing Accessibility.....	257
11.6.3 DRM.....	257
11.6.4 Home Media Systems.....	257
11.7 Media Processing Research	257
11.8 Deployments	260
11.9 Conclusion	261
References	261
Glossary	265
Index.....	271