Contents

List of Editors and Contributors xiii
List of Abbreviations xvii
Notations xxiii

1 Introduction 1
Oliver Schreer, Jean-François Macq, Omar Aziz Niamut, Javier Ruiz-Hidalgo,
Ben Shirley, Georg Thallinger and Graham Thomas

2 State-of-the-Art and Challenges in Media Production, Broadcast and Delivery 5
Graham Thomas, Arvid Engström, Jean-François Macq, Omar Aziz Niamut,
Ben Shirley and Richard Salmon

2.1 Introduction 5

2.2 Video Fundamentals and Acquisition Technology 7
2.2.1 How Real Should Video Look? 7
2.2.2 Fundamentals of Video 9
2.2.3 Camera Technology 14
2.2.4 Production for Formats with Differing Aspect Ratios 19
2.2.5 Stereoscopic 3D Video 20
2.2.6 Challenges for the Future 21

2.3 Audio Fundamentals and Acquisition Technology 21
2.3.1 Introduction 21
2.3.2 Fundamentals of Audio 21
2.3.3 Non-Live Production 27
2.3.4 Live Production 31
2.3.5 Surround Sound 34
2.3.6 Challenges for the Future 34

2.4 Live Programme Production 34
2.4.1 The Production Area and Roles in Production 35
2.4.2 The Production Workspace 37
2.4.3 Vision Mixing: Techniques for Directing, Selecting and Mixing Camera Feeds 38
2.4.4 Audio Mixing 40
Contents

3.3.2 Photometric Matching of Panoramic and Broadcast Cameras 98
3.3.3 Blending and Switching Between Camera Views 101

3.4 High Frame Rate Video 102
3.4.1 Early Work on HDTV Frame Rates 104
3.4.2 Issues with Conventional Frame Rates 104
3.4.3 Practical Investigations into the Effects of High Frame Rates 107
3.4.4 Future Frame Rates for TV Production and Distribution 111
3.4.5 Consideration of Frame Rates and Motion Portrayal in Synthetic Production 111
3.4.6 Conclusions on Frame Rates 112

3.5 High Dynamic Range Video 112
3.5.1 The Human Visual System in Natural Environments 113
3.5.2 Conventional and HDR Video Cameras 115
3.5.3 Conventional and HDR Displays 117
3.5.4 HDR Video Formats 119
3.5.5 Adaptive Tone-mapping for Format-Agnostic Video 120

3.6 Conclusion 125
Notes 126
References 126

4 Platform Independent Audio 130
Ben Shirley, Rob Oldfield, Frank Melchior and Johann-Markus Batke

4.1 Introduction 130
4.2 Terms and Definitions 132
4.2.1 Auditory Event and Sound Event 132
4.2.2 Basic Room Acoustics Theory 134

4.3 Definition of the Problem Space 135
4.3.1 Reproduction Environment 135
4.3.2 Reproduction Method 137
4.3.3 Audio-visual Coherence 141
4.3.4 User Interaction 143
4.3.5 Example Scenario 143

4.4 Scene Representation 144
4.4.1 Components of a Virtual Sound Scene 144
4.4.2 Representations of Virtual Sound Scenes 146
4.4.3 Implementation Examples 147

4.5 Scene Acquisition 149
4.5.1 Capturing Discrete Audio Objects 150
4.5.2 Capturing the Sound Field Component 152
4.5.3 Capturing the Diffuse Field 153

4.6 Scene Reproduction 153
4.6.1 Scenario: Mobile Consumption Via Headphones 153
4.6.2 Scenario: Interactive Multichannel Reproduction 154
4.6.3 Scenario: Big Screen 154
4.6.4 Scenario: Interactive and Free Viewpoint 3D 155
## Contents

4.7 Existing Systems 156
  4.7.1 Commercial Systems 156
  4.7.2 Research Projects 157
  4.7.3 Perceptive Media 160

4.8 Conclusion 161
  4.8.1 Open Issues 162

References 162

5 Semi-Automatic Content Annotation 166

Werner Bailer, Marco Masetti, Goranka Zorić, Marcus Thaler and Georg Thallinger

5.1 Introduction 166
  5.1.1 Requirements on Semi-automatic Annotation Tools 167
  5.1.2 Requirements on Metadata 168

5.2 Metadata Models and Analysis Architectures 170
  5.2.1 Metadata Models 170
  5.2.2 Architectures for Audio-visual Analysis 171
  5.2.3 Storing MPEG-7 Metadata 172
  5.2.4 Bulk Loading Techniques for Massive MPEG-7 Metadata Storage 175
  5.2.5 An Example Architecture of a Semantic Layer Management System 175

5.3 Domain-independent Saliency 177
  5.3.1 Spatio-temporal Visual Saliency 177
  5.3.2 Estimating Grid-based Saliency 178
  5.3.3 Salient Regions for Controlling Automated Shot Selection 179

5.4 Person Detection and Tracking 180
  5.4.1 Person Detection 181
  5.4.2 Person Tracking 182
  5.4.3 Multicamera and Panoramic Environment 184
  5.4.4 GPU Accelerated Real-time Tracking Beyond HD 187

5.5 Online Detection of Concepts and Actions 189
  5.5.1 Sequence-based Kernels 190
  5.5.2 Kernels for Online Detection 193
  5.5.3 Performance of Online Kernels 194

5.6 Supporting Annotation for Automated Production 195
  5.6.1 User Preferences and Functionality Definition 195
  5.6.2 Design Overview and Principles 196
  5.6.3 Preconfiguration and Predefined Workspaces 198

5.7 Conclusion 204

References 205

6 Virtual Director 209

Rene Kaiser and Wolfgang Weiss

6.1 Introduction 209
  6.1.1 What is a Virtual Director? 210
  6.1.2 Features Enabled by a Virtual Director 211
  6.1.3 Definitions 212
Contents

6.1.4 Requirements for Virtual Director Technology 213
6.1.5 Existing Implementations and Research Activities 215

6.2 Implementation Approaches 219
6.2.1 Behaviour Implementation Approaches 220
6.2.2 Combining Rule Engines and Event Processing Technology 223

6.3 Example Architecture and Workflow 225
6.3.1 Workflow of the Production System 225
6.3.2 Workflow of the Virtual Director 226
6.3.3 Distributed Nature 228
6.3.4 Sources of Knowledge 228

6.4 Virtual Director Subprocesses 230
6.4.1 Semantic Lifting 230
6.4.2 Shot Candidate Identification 231
6.4.3 Shot Framing 233
6.4.4 Shot Prioritisation 234
6.4.5 Decision Making 235

6.5 Behaviour Engineering: Production Grammar 237
6.5.1 Production Knowledge Elicitation Process 237
6.5.2 Cinematic Techniques 238
6.5.3 Audio Scripting 241
6.5.4 Domain Model 241
6.5.5 Limitations in Rule-based Behaviour Engineering 242

6.6 Virtual Director: Example Prototype 243
6.6.1 Architecture and Software Framework 245
6.6.2 Production Scripts 246
6.6.3 Behaviour Implementation 247
6.6.4 Production Grammar Example 248

6.7 Conclusion 251
6.7.1 Summary 251
6.7.2 Limitations 251
6.7.3 Testing and Evaluation 252
6.7.4 Research Roadmap 253
6.7.5 Concluding Thoughts 255

References 256

7 Scalable Delivery of Navigable and Ultra-High Resolution Video 260
Jean-François Macq, Patrice Rondão Alface, Ray van Brandenburg, Omar Aziz Niamut, Martin Prins and Nico Verzijp

7.1 Introduction 260

7.2 Delivery of Format-Agnostic Content: Key Concepts and State-of-the-Art 262
7.2.1 Delivery Agnostic to Content Formats, Device Capabilities and Network Bandwidth 262
7.2.2 Delivery Agnostic to Video Timing – the Temporal Interactivity Case 264
7.2.3 Delivery Agnostic to Video Reframing – the Spatial Interactivity Case 266
7.3 Spatial Random Access in Video Coding

7.3.1 Video Compression and Random Access – A Fundamental Trade-off

7.3.2 Spatial Random Access by Tracking Coding Dependencies

7.3.3 Spatial Random Access by Frame Tiling

7.3.4 Multi-Resolution Tiling

7.3.5 Overlapping Tiling for Low-Powered Devices

7.4 Models for Adaptive Tile-based Representation and Delivery

7.4.1 Saliency-based Adaptive Coding of Tiled Content

7.4.2 Optimisation of Tile Selection Under Delay and Bandwidth Constraints

7.5 Segment-based Adaptive Transport

7.5.1 Video Streaming Over IP

7.5.2 Tiled HTTP Adaptive Streaming Over the Internet

7.5.3 Publish/Subscribe System for Interactive Video Streaming

7.6 Conclusion

References

8 Interactive Rendering

Javier Ruiz-Hidalgo, Malte Borsum, Axel Kochale and Goranka Zorić

8.1 Introduction

8.2 Format-Agnostic Rendering

8.2.1 Available Rendering Solutions in End Terminals

8.2.2 Requirements for Format-Agnostic Video Rendering

8.2.3 Description of a Technical Solution

8.3 Device-less Interaction for Rendering Control

8.3.1 Sensors for Gesture Recognition

8.3.2 Gesture Analysis Techniques

8.3.3 Recognition and Classification Techniques

8.3.4 Remaining Challenges in Gesture Recognition Systems

8.3.5 Description of a Technical Solution

8.3.6 User Study of the Gesture Interface

8.4 Conclusions

References

9 Application Scenarios and Deployment Domains

Omar Aziz Niamut, Arvid Engström, Axel Kochale, Jean-François Macq, Graham Thomas and Goranka Zorić

9.1 Introduction

9.2 Application Scenarios

9.2.1 Digital Cinema: A Totally Immersive Experience

9.2.2 Home Theatre: In Control

9.2.3 Mobile: Navigation and Magnification

9.3 Deployment in the Production Domain

9.3.1 Outlook on New Production Practices