Emerging Technologies for 3D Video
Creation, Coding, Transmission and Rendering
Editors
Frédéric Dufaux | Béatrice Pesquet-Popescu | Marco Cagnazzo

WILEY
EMERGING TECHNOLOGIES FOR 3D VIDEO
EMERGING TECHNOLOGIES FOR 3D VIDEO CREATION, CODING, TRANSMISSION AND RENDERING

Edited by

Frédéric Dufaux
Télécom Paris Tech, CNRS, France

Béatrice Pesquet-Popescu
Télécom Paris Tech, France

Marco Cagnazzo
Télécom Paris Tech, France
## Contents

| Preface | xvii |
| List of Contributors | xxi |
| Acknowledgements | xxv |

### PART I CONTENT CREATION

#### 1 Consumer Depth Cameras and Applications

*Seungkyu Lee*

1.1 Introduction 3
1.2 Time-of-Flight Depth Camera 3
   1.2.1 Principle 4
   1.2.2 Quality of the Measured Distance 6
1.3 Structured Light Depth Camera 11
   1.3.1 Principle 11
1.4 Specular and Transparent Depth 12
1.5 Depth Camera Applications 15
   1.5.1 Interaction 15
   1.5.2 Three-Dimensional Reconstruction 15

References 16

#### 2 SFTI: Space-from-Time Imaging

*Ahmed Kirmani, Andrea Colaço, and Vivek K. Goyal*

2.1 Introduction 17
2.2 Background and Related Work 18
   2.2.1 Light Fields, Reflectance Distribution Functions, and Optical Image Formation 18
   2.2.2 Time-of-Flight Methods for Estimating Scene Structure 20
   2.2.3 Synthetic Aperture Radar for Estimating Scene Reflectance 20
2.3 Sampled Response of One Source–Sensor Pair 21
   2.3.1 Scene, Illumination, and Sensor Abstractions 21
   2.3.2 Scene Response Derivation 22
   2.3.3 Inversion 24
2.4 Diffuse Imaging: SFTI for Estimating Scene Reflectance 24
3 2D-to-3D Video Conversion: Overview and Perspectives

Carlos Vazquez, Liang Zhang, Filippo Speranza, Nils Plath, and Sebastian Knorr

3.1 Introduction
3.2 The 2D-to-3D Conversion Problem
3.2.1 General Conversion Approach
3.2.2 Depth Cues in Monoscopic Video
3.3 Definition of Depth Structure of the Scene
3.3.1 Depth Creation Methods
3.3.2 Depth Recovery Methods
3.4 Generation of the Second Video Stream
3.4.1 Depth to Disparity Mapping
3.4.2 View Synthesis and Rendering Techniques
3.4.3 Post-Processing for Hole-Filling
3.5 Quality of Experience of 2D-to-3D Conversion
3.6 Conclusions

References

4 Spatial Plasticity: Dual-Camera Configurations and Variable Interaxial Ray Zone

4.1 Stereoscopic Capture
4.2 Dual-Camera Arrangements in the 1950s
4.3 Classic “Beam-Splitter” Technology
4.4 The Dual-Camera Form Factor and Camera Mobility
4.5 Reduced 3D Form Factor of the Digital CCD Sensor
4.6 Handheld Shooting with Variable Interaxial
4.7 Single-Body Camera Solutions for Stereoscopic Cinematography
4.8 A Modular 3D Rig
4.9 Human Factors of Variable Interaxial

References
5.2 Geometrical Models for Stereoscopic Imaging  82
   5.2.1 The Pinhole Camera Model  82
   5.2.2 Stereoscopic Imaging Systems  85
5.3 Stereo Matching Process  88
   5.3.1 Disparity Information  88
   5.3.2 Difficulties in the Stereo Matching Process  88
   5.3.3 Stereo Matching Constraints  89
   5.3.4 Fundamental Steps Involved in Stereo Matching Algorithms  89
5.4 Overview of Disparity Estimation Methods  91
   5.4.1 Local Methods  91
   5.4.2 Global Methods  93
5.5 Conclusion  98
References  98

6 3D Video Representation and Formats  102
Marco Cagnazzo, Béatrice Pesquet-Popescu, and Frédéric Dufaux

6.1 Introduction  102
6.2 Three-Dimensional Video Representation  103
   6.2.1 Stereoscopic 3D (S3D) Video  103
   6.2.2 Multiview Video (MVV)  104
   6.2.3 Video-Plus-Depth  105
   6.2.4 Multiview Video-Plus-Depth (MVD)  107
   6.2.5 Layered Depth Video (LDV)  107
6.3 Three-Dimensional Video Formats  109
   6.3.1 Simulcast  109
   6.3.2 Frame-Compatible Stereo Interleaving  110
   6.3.3 MPEG-4 Multiple Auxiliary Components (MAC)  113
   6.3.4 MPEG-C Part 3  113
   6.3.5 MPEG-2 Multiview Profile (MVP)  113
   6.3.6 Multiview Video Coding (MVC)  114
6.4 Perspectives  118
Acknowledgments  118
References  119

7 Depth Video Coding Technologies  121
Elie Gabriel Mora, Giuseppe Valenzise, Joël Jung, Béatrice Pesquet-Popescu, Marco Cagnazzo, and Frédéric Dufaux

7.1 Introduction  121
7.2 Depth Map Analysis and Characteristics  122
7.3 Depth Map Coding Tools  123
   7.3.1 Tools that Exploit the Inherent Characteristics of Depth Maps  123
   7.3.2 Tools that Exploit the Correlations with the Associated Texture  127
   7.3.3 Tools that Optimize Depth Map Coding for the Quality of the Synthesis  129
7.4 Application Example: Depth Map Coding Using “Don’t Care” Regions  132
   7.4.1 Derivation of “Don’t Care” Regions  133
   7.4.2 Transform Domain Sparsification Using “Don’t Care” Regions  134
8 Depth-Based 3D Video Formats and Coding Technology

Anthony Vetro and Karsten Müller

8.1 Introduction 139
8.1.1 Existing Stereo/Multiview Formats 140
8.1.2 Requirements for Depth-Based Format 140
8.1.3 Chapter Organization 141
8.2 Depth Representation and Rendering 141
8.2.1 Depth Format and Representation 142
8.2.2 Depth-Image-Based Rendering 143
8.3 Coding Architectures 144
8.3.1 AVC-Based Architecture 144
8.3.2 HEVC-Based Architecture 146
8.3.3 Hybrid 146
8.4 Compression Technology 147
8.4.1 Inter-View Prediction 148
8.4.2 View Synthesis Prediction 148
8.4.3 Depth Resampling and Filtering 149
8.4.4 Inter-Component Parameter Prediction 150
8.4.5 Depth Modelling 151
8.4.6 Bit Allocation 152
8.5 Experimental Evaluation 153
8.5.1 Evaluation Framework 153
8.5.2 AVC-Based 3DV Coding Results 155
8.5.3 HEVC-Based 3DV Coding Results 156
8.5.4 General Observations 158
8.6 Concluding Remarks 158
References 159

9 Coding for Interactive Navigation in High-Dimensional Media Data

Ngai-Man Cheung and Gene Cheung

9.1 Introduction 162
9.2 Challenges and Approaches of Interactive Media Streaming 163
9.2.1 Challenges: Coding Efficiency and Navigation Flexibility 163
9.2.2 Approaches to Interactive Media Streaming 165
9.3 Example Solutions 166
9.3.1 Region-of-Interest (RoI) Image Browsing 166
9.3.2 Light-Field Streaming 167
9.3.3 Volumetric Image Random Access 168
9.3.4 Video Browsing 168
9.3.5 Reversible Video Playback 169
9.3.6 Region-of-Interest (RoI) Video Streaming 169
References 169
9.4 Interactive Multiview Video Streaming 172
  9.4.1 Interactive Multiview Video Streaming (IMVS) 172
  9.4.2 IMVS with Free Viewpoint Navigation 179
  9.4.3 IMVS with Fixed Round-Trip Delay 181
9.5 Conclusion 184
References 184

10 Adaptive Streaming of Multiview Video Over P2P Networks 187
C. Göktuğ Gürler and A. Murat Tekalp

10.1 Introduction 187
10.2 P2P Overlay Networks 188
  10.2.1 Overlay Topology 188
  10.2.2 Sender-Driven versus Receiver-Driven P2P Video Streaming 189
  10.2.3 Layered versus Cross-Layer Architecture 190
  10.2.4 When P2P is Useful: Regions of Operation 191
  10.2.5 BitTorrent: A Platform for File Sharing 191
10.3 Monocular Video Streaming Over P2P Networks 192
  10.3.1 Video Coding 193
  10.3.2 Variable-Size Chunk Generation 193
  10.3.3 Time-Sensitive Chunk Scheduling Using Windowing 194
  10.3.4 Buffer-Driven Rate Adaptation 195
  10.3.5 Adaptive Window Size and Scheduling Restrictions 195
  10.3.6 Multiple Requests from Multiple Peers of a Single Chunk 196
10.4 Stereoscopic Video Streaming over P2P Networks 197
  10.4.1 Stereoscopic Video over Digital TV 197
  10.4.2 Rate Adaptation in Stereo Streaming: Asymmetric Coding 197
  10.4.3 Use Cases: Stereoscopic Video Streaming over P2P Network 200
10.5 MVV Streaming over P2P Networks 201
  10.5.1 MVV Streaming over IP 201
  10.5.2 Rate Adaptation for MVV: View Scaling 201
  10.5.3 Use Cases: MVV Streaming over P2P Network 202
References 203

PART III RENDERING AND SYNTHESIS

11 Image Domain Warping for Stereoscopic 3D Applications 207
Oliver Wang, Manuel Lang, Nikolce Stefanoski, Alexander Sorkine-Hornung,
Olga Sorkine-Hornung, Aljoscha Smolic, and Markus Gross

11.1 Introduction 207
11.2 Background 208
11.3 Image Domain Warping 209
11.4 Stereo Mapping 210
  11.4.1 Problems in Stereoscopic Viewing 210
  11.4.2 Disparity Range 210
  11.4.3 Disparity Sensitivity 211
  11.4.4 Disparity Velocity 211

References

Contents ix
11.4.5 Summary 212
11.4.6 Disparity Mapping Operators 212
11.4.7 Linear Operator 212
11.4.8 Nonlinear Operator 212
11.4.9 Temporal Operator 213
11.5 Warp-Based Disparity Mapping 213
  11.5.1 Data Extraction 213
  11.5.2 Warp Calculation 214
  11.5.3 Applications 216
11.6 Automatic Stereo to Multiview Conversion 218
  11.6.1 Automatic Stereo to Multiview Conversion 218
  11.6.2 Position Constraints 219
  11.6.3 Warp Interpolation and Extrapolation 219
  11.6.4 Three-Dimensional Video Transmission Systems for Multiview Displays 220
11.7 IDW for User-Driven 2D–3D Conversion 221
  11.7.1 Technical Challenges of 2D–3D Conversion 222
11.8 Multi-Perspective Stereoscopy from Light Fields 225
11.9 Conclusions and Outlook 228
Acknowledgments 229
References 229

12 Image-Based Rendering and the Sampling of the Plenoptic Function 231
Christopher Gilliam, Mike Brookes, and Pier Luigi Dragotti
12.1 Introduction 231
12.2 Parameterization of the Plenoptic Function 232
  12.2.1 Light Field and Surface Light Field Parameterization 232
  12.2.2 Epipolar Plane Image 234
12.3 Uniform Sampling in a Fourier Framework 235
  12.3.1 Spectral Analysis of the Plenoptic Function 236
  12.3.2 The Plenoptic Spectrum under Realistic Conditions 239
12.4 Adaptive Plenoptic Sampling 242
  12.4.1 Adaptive Sampling Based on Plenoptic Spectral Analysis 244
12.5 Summary 246
  12.5.1 Outlook 246
References 247

13 A Framework for Image-Based Stereoscopic View Synthesis from Asynchronous Multiview Data 249
Felix Klose, Christian Lipski, and Marcus Magnor
13.1 The Virtual Video Camera 249
  13.1.1 Navigation Space Embedding 251
  13.1.2 Space–Time Tetrahedralization 252
  13.1.3 Processing Pipeline 255
  13.1.4 Rendering 256
  13.1.5 Application 257
  13.1.6 Limitations 258
13.2 Estimating Dense Image Correspondences 258
  13.2.1 Belief Propagation for Image Correspondences 259
  13.2.2 A Symmetric Extension 260
  13.2.3 SIFT Descriptor Downsampling 261
  13.2.4 Construction of Message-Passing Graph 261
  13.2.5 Data Term Compression 262
  13.2.6 Occlusion Removal 263
  13.2.7 Upsampling and Refinement 263
  13.2.8 Limitations 263
13.3 High-Quality Correspondence Edit 264
  13.3.1 Editing Operations 264
  13.3.2 Applications 265
13.4 Extending to the Third Dimension 265
  13.4.1 Direct Stereoscopic Virtual View Synthesis 266
  13.4.2 Depth-Image-Based Rendering 267
  13.4.3 Comparison 267
  13.4.4 Concluding with the “Who Cares?” Post-Production Pipeline 268
References 270

PART IV DISPLAY TECHNOLOGIES

14 Signal Processing for 3D Displays 275
  Janusz Konrad
  14.1 Introduction 275
  14.2 3D Content Generation 276
    14.2.1 Automatic 2D-to-3D Image Conversion 276
    14.2.2 Real-Time Intermediate View Interpolation 280
    14.2.3 Brightness and Color Balancing in Stereopairs 286
  14.3 Dealing with 3D Display Hardware 287
    14.3.1 Ghosting Suppression for Polarized and Shuttered
    Stereoscopic 3D Displays 287
    14.3.2 Aliasing Suppression for Multiview Eyewear-Free 3D Displays 289
  14.4 Conclusions 292
Acknowledgments 293
References 293

15 3D Display Technologies 295
  Thierry Borel and Didier Doyen
  15.1 Introduction 295
  15.2 Three-Dimensional Display Technologies in Cinemas 295
    15.2.1 Three-Dimensional Cinema Projectors Based on Light
    Polarization 296
    15.2.2 Three-Dimensional Cinema Projectors Based on Shutters 299
    15.2.3 Three-Dimensional Cinema Projectors Based on
    Interference Filters 300
  15.3 Large 3D Display Technologies in the Home 301
    15.3.1 Based on Anaglyph Glasses 301
15.3.2 Based on Shutter Glasses
15.3.3 Based on Polarized Glasses
15.3.4 Without Glasses
15.4 Mobile 3D Display Technologies
15.4.1 Based on Parallax Barriers
15.4.2 Based on Lighting Switch
15.5 Long-Term Perspectives
15.6 Conclusion

References

16 Integral Imaging
Jun Arai

16.1 Introduction
16.2 Integral Photography
16.2.1 Principle
16.2.2 Integral Photography with a Concave Lens Array
16.2.3 Holocoder Hologram
16.2.4 IP using a Retrodirective Screen
16.2.5 Avoiding Pseudoscopic Images
16.3 Real-Time System
16.3.1 Orthoscopic Conversion Optics
16.3.2 Applications of the Ultra-High-Resolution Video System
16.4 Properties of the Reconstructed Image
16.4.1 Geometrical Relationship of Subject and Spatial Image
16.4.2 Resolution
16.4.3 Viewing Area
16.5 Research and Development Trends
16.5.1 Acquiring and Displaying Spatial Information
16.5.2 Elemental Image Generation from 3D Object Information
16.5.3 Three-Dimensional Measurement
16.5.4 Hologram Conversion
16.6 Conclusion

References

17 3D Light-Field Display Technologies
Péter Tamás Kovács and Tibor Balogh

17.1 Introduction
17.2 Fundamentals of 3D Displaying
17.3 The HoloVizio Light-Field Display System
17.3.1 Design Principles and System Parameters
17.3.2 Image Organization
17.4 HoloVizio Displays and Applications
17.4.1 Desktop Displays
17.4.2 Large-Scale Displays
17.4.3 Cinema Display
17.4.4 Software and Content Creation
17.4.5 Applications
PART V  HUMAN VISUAL SYSTEM AND QUALITY ASSESSMENT

18  3D Media and the Human Visual System  349
Simon J. Watt and Kevin J. MacKenzie

18.1 Overview  349
18.2 Natural Viewing and S3D Viewing  349
18.3 Perceiving 3D Structure
  18.3.1 Perceiving Depth from Binocular Disparity  352
18.4 ‘Technical’ Issues in S3D Viewing
  18.4.1 Cross-Talk  354
  18.4.2 Low Image Luminance and Contrast  355
  18.4.3 Photometric Differences Between Left- and Right-Eye Images  355
  18.4.4 Camera Misalignments and Differences in Camera Optics  356
  18.4.5 Window Violations  356
  18.4.6 Incorrect Specular Highlights  356
18.5 Fundamental Issues in S3D Viewing  357
18.6 Motion Artefacts from Field-Sequential Stereoscopic Presentation
  18.6.1 Perception of Flicker  359
  18.6.2 Perception of Unsmooth or Juddering Motion  359
  18.6.3 Distortions in Perceived Depth from Binocular Disparity  360
  18.6.4 Conclusions  360
18.7 Viewing Stereoscopic Images from the ‘Wrong’ Place
  18.7.1 Capture Parameters  361
  18.7.2 Display Parameters and Viewer Parameters  364
  18.7.3 Are Problems of Incorrect Geometry Unique to S3D?  364
  18.7.4 Conclusions  366
18.8 Fixating and Focusing on Stereoscopic Images
  18.8.1 Accommodation, Vergence and Viewing Distance  367
  18.8.2 Accommodation and Vergence in the Real World and in S3D  367
  18.8.3 Correcting Focus Cues in S3D  368
  18.8.4 The Stereoscopic Zone of Comfort  369
  18.8.5 Specifying the Zone of Comfort for Cinematography  370
  18.8.6 Conclusions  371
18.9 Concluding Remarks  372
Acknowledgments  372
References  372

19  3D Video Quality Assessment  377
Philippe Hanhart, Francesca De Simone, Martin Rerabek, and Touradj Ebrahimi

19.1 Introduction  377
19.2 Stereoscopic Artifacts  378
21 View Selection 416
Fahad Daniyal and Andrea Cavallaro

21.1 Introduction 416
21.2 Content Analysis 417
21.2.1 Pose 417
21.2.2 Occlusions 419
21.2.3 Position 419
21.2.4 Size 421
21.2.5 Events 421
21.3 Content Ranking 421
21.3.1 Object-Centric Quality of View 422
21.3.2 View-Centric Quality of View 423
21.4 View Selection 424
21.4.1 View Selection as a Scheduling Problem 425
21.4.2 View Selection as an Optimization Problem 425
21.5 Comparative Summary and Outlook 426
References 429

22 3D Video on Mobile Devices 432
Arnaud Bourge and Alain Bellon

22.1 Mobile Ecosystem, Architecture, and Requirements 432
22.2 Stereoscopic Applications on Mobile Devices 433
22.2.1 3D Video Camcorder 434
22.2.2 3D Video Player 434
22.2.3 3D Viewing Modalities 434
22.2.4 3D Graphics Applications 435
22.2.5 Interactive Video Applications 435
22.2.6 Monoscopic 3D 435
22.3 Stereoscopic Capture on Mobile Devices 436
22.3.1 Stereo-Camera Design 436
22.3.2 Stereo Imaging 437
22.3.3 Stereo Rectification, Lens Distortion, and Camera Calibration 438
22.3.4 Digital Zoom and Video Stabilization 440
22.3.5 Stereo Codecs 442
22.4 Display Rendering on Mobile Devices 442
22.4.1 Local Auto-Stereoscopic Display 442
22.4.2 Remote HD Display 443
22.4.3 Stereoscopic Rendering 443
22.5 Depth and Disparity 445
22.5.1 View Synthesis 445
22.5.2 Depth Map Representation and Compression Standards 446
22.5.3 Other Usages 447
22.6 Conclusions 448
Acknowledgments 448
References 448
# 23 Graphics Composition for Multiview Displays

_Jean Le Feuvre and Yves Mathieu_

23.1 An Interactive Composition System for 3D Displays
23.2 Multimedia for Multiview Displays
   23.2.1 Media Formats
   23.2.2 Multimedia Languages
   23.2.3 Multiview Displays

23.3 GPU Graphics Synthesis for Multiview Displays
   23.3.1 3D Synthesis
   23.3.2 View Interleaving
   23.3.3 3D Media Rendering

23.4 DIBR Graphics Synthesis for Multiview Displays
   23.4.1 Quick Overview
   23.4.2 DIBR Synthesis
   23.4.3 Hardware Compositor
   23.4.4 DIBR Pre- and Post-Processing
   23.4.5 Hardware Platform

23.5 Conclusion

Acknowledgments
References

---

# 24 Real-Time Disparity Estimation Engine for High-Definition 3DTV Applications

_Yu-Cheng Tseng and Tian-Sheuan Chang_

24.1 Introduction

24.2 Review of Disparity Estimation Algorithms and Implementations
   24.2.1 DP-Based Algorithms and Implementations
   24.2.2 GC-Based Algorithms and Implementations
   24.2.3 BP-Based Algorithms and Implementations

24.3 Proposed Hardware-Efficient Algorithm
   24.3.1 Downsampled Matching Cost for Full Disparity Range
   24.3.2 Hardware-Efficient Cost Diffusion Method
   24.3.3 Upsampling Disparity Maps
   24.3.4 Temporal Consistency Enhancement Methods
   24.3.5 Occlusion Handling

24.4 Proposed Architecture
   24.4.1 Overview of Architecture
   24.4.2 Computational Modules
   24.4.3 External Memory Access

24.5 Experimental Results
   24.5.1 Comparison of Disparity Quality
   24.5.2 Analysis of Sampling Factor
   24.5.3 Implementation Result

24.6 Conclusion

References

Index
Preface

The underlying principles of stereopsis have been known for a long time. Stereoscopes to see photographs in 3D appeared and became popular in the nineteenth century. The first demonstrations of 3D movies took place in the first half of the twentieth century, initially using anaglyph glasses, and then with polarization-based projection. Hollywood experienced a first short-lived golden era of 3D movies in the 1950s. In the last 10 years, 3D has regained significant interests and 3D movies are becoming ubiquitous. Numerous major productions are now released in 3D, culminating with Avatar, the highest grossing film of all time.

In parallel with the recent growth of 3D movies, 3DTV is attracting significant interest from manufacturers and service providers. This is obvious by the multiplication of new 3D product announcements and services. Beyond entertainment, 3D imaging technology is also seen as instrumental in other application areas such as video games, immersive video conferences, medicine, video surveillance, and engineering.

With this growing interest, 3D video is often considered as one of the major upcoming innovations in video technology, with the expectation of greatly enhanced user experience.

This book intends to provide an overview of key technologies for 3D video applications. More specifically, it covers the state of the art and explores new research directions, with the objective to tackle all aspects involved in 3D video systems and services. Topics addressed include content acquisition and creation, data representation and coding, transmission, data representation, compression, and transmission. Relevant standardization efforts are reviewed. Finally, applications and implementation issues are also described.

More specifically, the book is composed of six parts. Part One addresses different aspects of 3D content acquisition and creation. In Chapter 1, Lee presents depth cameras and related applications. The principle of active depth sensing is reviewed, along with depth image processing methods such as noise modelling, upsampling, and removing motion blur. In Chapter 2, Kirmani, Colaço, and Goyal introduce the space-from-time imaging framework, which achieves spatial resolution, in two and three dimensions, by measuring temporal variations of light intensity in response to temporally or spatiotemporally varying illumination. Chapter 3, by Vazquez, Zhang, Speranza, Plath, and Knorr, provides an overview of the process generating a stereoscopic video (S3D) from a monoscopic video source (2D), generally known as 2D-to-3D video conversion, with a focus on selected recent techniques. Finally, in Chapter 4, Zone\(^*\) provides an overview of numerous contemporary strategies for shooting narrow and variable interaxial baseline for stereoscopic cinematography. Artistic implications are also discussed.

A key issue in 3D video, Part Two addresses data representation, compression, and transmission. In Chapter 5, Kaaniche, Gaetano, Cagnazzo, and Pesquet-Popescu address the

\(^*\) It is with great sadness that we learned that Ray Zone passed away on November 13, 2012.
problem of disparity estimation. The geometrical relationship between the 3D scene and the generated stereo images is analyzed and the most important techniques for disparity estimation are reviewed. Cagnazzo, Pesquet-Popescu, and Dufaux give an overview of existing data representation and coding formats for 3D video content in Chapter 6. In turn, in Chapter 7, Mora, Valenzise, Jung, Pesquet-Popescu, Cagnazzo, and Dufaux consider the problem of depth map coding and present an overview of different coding tools. In Chapter 8, Vetro and Müller provide an overview of the current status of research and standardization activity towards defining a new set of depth-based formats that facilitate the generation of intermediate views with a compact binary representation. In Chapter 9, Cheung and Cheung consider interactive media streaming, where the server continuously and reactively sends appropriate subsets of media data in response to a client’s periodic requests. Different associated coding strategies and solutions are reviewed. Finally, Güler and Tekalp propose an adaptive P2P video streaming solution for streaming multiview video over P2P overlays in Chapter 10.

Next, Part Three of the book discusses view synthesis and rendering. In Chapter 11, Wang, Lang, Stefanoski, Sorkine-Hornung, Sorkine-Hornung, Smolic, and Gross present image-domain warping as an alternative to depth-image-based rendering techniques. This technique utilizes simpler, image-based deformations as a means for realizing various stereoscopic post-processing operators. Gilliam, Brookes, and Dragotti, in Chapter 12, examine the state of the art in plenoptic sampling theory. In particular, the chapter presents theoretical results for uniform sampling based on spectral analysis of the plenoptic function and algorithms for adaptive plenoptic sampling. Finally, in Chapter 13, Klose, Lipski, and Magnor present a complete end-to-end framework for stereoscopic free viewpoint video creation, allowing one to viewpoint-navigate through space and time of complex real-world, dynamic scenes.

As a very important component of a 3D video system, Part Four focuses on 3D display technologies. In Chapter 14, Konrad addresses digital signal processing methods for 3D data generation, both stereoscopic and multiview, and for compensation of the deficiencies of today’s 3D displays. Numerous experimental results are presented to demonstrate the usefulness of such methods. Borel and Doyen, in Chapter 15, present in detail the main 3D display technologies available for cinemas, for large-display TV sets, and for mobile terminals. A perspective of evolution for the near and long term is also proposed. In Chapter 16, Arai focuses on integral imaging, a 3D photography technique that is based on integral photography, in which information on 3D space is acquired and represented. This chapter describes the technology for displaying 3D space as a spatial image by integral imaging. Finally, in Chapter 17, Kovács and Balogh present light-field displays, an advanced technique for implementing glasses-free 3D displays.

In most targeted applications, humans are the end-users of 3D video systems. Part Five considers human perception of depth and perceptual quality assessment. More specifically, in Chapter 18, Watt and MacKenzie focus on how the human visual system interacts with stereoscopic 3D media, in view of optimizing effectiveness and viewing comfort. Three main issues are addressed: incorrect spatiotemporal stimuli introduced by field-sequential stereo presentation, inappropriate binocular viewing geometry, and the unnatural relationship between where the eyes fixate and focus in stereoscopic 3D viewing. In turn, in Chapter 19, Hanhart, De Simone, Rerabek, and Ebrahimi consider mechanisms of 3D vision in humans, and their underlying perceptual models, in conjunction with the types of distortions that today’s and tomorrow’s 3D video processing systems produce. This complex puzzle is examined with a focus on how to measure 3D visual quality, as an essential factor in the success of 3D technologies, products, and services.
In order to complete the book, Part Six describes target applications for 3D video, as well as implementation issues. In Chapter 20, Bazin, Saurer, Fraundorfer, and Pollefeys present a semi-automatic method to generate interactive virtual tours from omnidirectional video. It allows a user to virtually navigate through buildings and indoor scenes. Such a system can be applied in various contexts, such as virtual tourism, tele-immersion, tele-presence, and e-heritage. Daniyal and Cavallaro address the question of how to automatically identify which view is more useful when observing a dynamic scene with multiple cameras in Chapter 21. This problem concerns several applications ranging from video production to video surveillance. In particular, an overview of existing approaches for view selection and automated video production is presented. In Chapter 22, Bourge and Bellon present the hardware architecture of a typical mobile platform, and describe major stereoscopic 3D applications. Indeed, smartphones bring new opportunities to stereoscopic 3D, but also specific constraints. Chapter 23, by Le Feuvre and Mathieu, presents an integrated system for displaying interactive applications on multiview screens. Both a simple GPU-based prototype and a low-cost hardware design implemented on a field-programmable gate array are presented. Finally, in Chapter 24, Tseng and Chang propose an optimized disparity estimation algorithm for high-definition 3DTV applications with reduced computational and memory requirements.

By covering general and advanced topics, providing at the same time a broad and deep analysis, the book has the ambition to become a reference for those involved or interested in 3D video systems and services. Assuming fundamental knowledge in image/video processing, as well as a basic understanding in mathematics, this book should be of interest to a broad readership with different backgrounds and expectations, including professors, graduate and undergraduate students, researchers, engineers, practitioners, and managers making technological decisions about 3D video.

Frédéric Dufaux
Béatrice Pesquet-Popescu
Marco Cagnazzo
List of Contributors

Jun Arai, NHK (Japan Broadcasting Corporation), Japan
Tibor Balogh, Holografika, Hungary
Jean-Charles Bazin, Computer Vision and Geometry Group, ETH Zürich, Switzerland
Alain Bellon, STMicroelectronics, France
Thierry Borel, Technicolor, France
Arnaud Bourge, STMicroelectronics, France
Mike Brookes, Department of Electrical and Electronic Engineering, Imperial College London, UK
Marco Cagnazzo, Département Traitement du Signal et des Images, Télécom ParisTech, France
Andrea Cavallaro, Queen Mary University of London, UK
Tian-Sheuan Chang, Department of Electronics Engineering, National Chiao Tung University, Taiwan
Gene Cheung, Digital Content and Media Sciences Research Division, National Institute of Informatics, Japan
Ngai-Man Cheung, Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore
Andrea Colaço, Media Lab, Massachusetts Institute of Technology, USA
Fahad Daniyal, Queen Mary University of London, UK
Francesca De Simone, Multimedia Signal Processing Group (MMSPG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Didier Doyen, Technicolor, France
Pier Luigi Dragotti, Department of Electrical and Electronic Engineering, Imperial College London, UK
Frédéric Dufaux, Département Traitement du Signal et des Images, Télécom ParisTech, France
Touradj Ebrahimi, Multimedia Signal Processing Group (MMSPG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Friedrich Fraundorfer, Computer Vision and Geometry Group, ETH Zürich, Switzerland

Raffaele Gaetano, Département Traitement du Signal et des Images, Télécom ParisTech, France

Christopher Gilliam, Department of Electrical and Electronic Engineering, Imperial College London, UK

Vivek K. Goyal, Research Laboratory of Electronics, Massachusetts Institute of Technology, USA

Markus Gross, Disney Research Zurich, Switzerland

C. Göktüg Gürler, College of Engineering, Koç University, Turkey

Philippe Hanhart, Multimedia Signal Processing Group (MMSPG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Alexander Sorkine-Hornung, Disney Research Zurich, Switzerland

Joël Jung, Orange Labs, France

Mounir Kaaniche, Département Traitement du Signal et des Images, Télécom ParisTech, France

Ahmed Kirmani, Research Laboratory of Electronics, Massachusetts Institute of Technology, USA

Felix Klose, Institut für Computergraphik, TU Braunschweig, Germany

Sebastian Knorr, incube labs GmbH, Technische Universität Berlin, Germany

Janusz Konrad, Department of Electrical and Computer Engineering, Boston University, USA

Péter Tamás Kovács, Holografika, Hungary

Manuel Lang, Disney Research Zurich, Switzerland

Seungkyu Lee, Samsung Advanced Institute of Technology, South Korea

Jean Le Feuvre, Département Traitement du Signal et des Images, Telecom ParisTech, France

Christian Lipski, Institut für Computergraphik, TU Braunschweig, Germany

Kevin J. MacKenzie, Wolfson Centre for Cognitive Neuroscience, School of Psychology, Bangor University, UK

Marcus Magnor, Institut für Computergraphik, TU Braunschweig, Germany

Yves Mathieu, Telecom ParisTech, France

Elie Gabriel Mora, Orange Labs, France; Département Traitement du Signal et des Images, Télécom ParisTech, France
List of Contributors

Karsten Müller, Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Germany

Béatrice Pesquet-Popescu, Département Traitement du Signal et des Images, Télécom ParisTech, France

Nils Plath, imcube labs GmbH, Technische Universität Berlin, Germany

Marc Pollefeys, Computer Vision and Geometry Group, ETH Zürich, Switzerland

Martin Rerabek, Multimedia Signal Processing Group (MMSPG), Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Olivier Saurer, Computer Vision and Geometry Group, ETH Zürich, Switzerland

Aljoscha Smolic, Disney Research Zurich, Switzerland

Olga Sorkine-Hornung, ETH Zurich, Switzerland

Filippo Speranza, Communications Research Centre Canada (CRC), Canada

Nikolce Stefanoski, Disney Research Zurich, Switzerland

A. Murat Tekalp, College of Engineering, Koç University, Turkey

Yu-Cheng Tseng, Department of Electronics Engineering, National Chiao Tung University, Taiwan

Giuseppe Valenzise, Département Traitement du Signal et des Images, Télécom ParisTech, France

Carlos Vazquez, Communications Research Centre Canada (CRC), Canada

Anthony Vetro, Mitsubishi Electric Research Labs (MERL), USA

Simon J. Watt, Wolfson Centre for Cognitive Neuroscience, School of Psychology, Bangor University, UK

Oliver Wang, Disney Research Zurich, Switzerland

Liang Zhang, Communications Research Centre Canada (CRC), Canada

Ray Zone, The 3-D Zone, USA
We would like to express our deepest appreciation to all the authors for their invaluable contributions. Without their commitment and efforts, this book would not have been possible.

Moreover, we would like to gratefully acknowledge the John Wiley & Sons Ltd. staff, Alex King, Liz Wingett, Richard Davies, and Genna Manaog, for their relentless support throughout this endeavour.

Frédéric Dufaux
Béatrice Pesquet-Popescu
Marco Cagnazzo
Part One
Content Creation