Preface

Riding on the success of 3D cinema blockbusters and the advancements in stereoscopic display technology, 3D video applications have been gathering momentum in recent years, which further enhance visual experience by vividly extending the conventional "flat" video into a third dimension. Several 3D video prototypes have been developed based on distinct techniques in 3D visualization, representation, and content production. Among them, stereoscopic 3D video systems evoke 3D perception by binocular parallax, in which scene is presented in a fixed perspective defined by two transmitted views, while further manipulations on depth perception require expensive computation with current technologies.

Depth-image-based rendering (DIBR) is being considered to significantly enhance the 3D visual experience relative to the conventional stereoscopic systems. With DIBR techniques, it becomes possible to generate additional viewpoints using 3D warping techniques to adjust the perceived depth of stereoscopic videos or to provide the necessary input for auto-stereoscopic displays that do not require glasses to view the 3D scene. This functionality is also useful for free-viewpoint video (FVV), where the viewer has the freedom to move about in front of the display, and is able to perceive natural perspective changes as if looking through a window. In recognition of progress being made in this area and a strong interest from the industry to provide equipment and services supporting such applications, MPEG is also embarking on a new phase of 3D video standardization based on DIBR techniques.

The technologies surrounding DIBR-oriented 3D video systems, however, are not mature enough at this stage to fully fulfill the above targets. Depth maps, which are central to the synthesis of virtual views, need to be either captured with specialized apparatus or estimated from scene textures using stereo matching. Existing solutions are either costly or not sufficiently robust. Besides, there is a strong need to achieve efficient storage and robust transmission of this additional information. Knowing that the depth maps and scene textures are different in nature, and that synthesized views are the ultimate information for display, DIBR-oriented depth, and texture coding may employ different distortion measures for rate-distortion or rate-quality optimization, and possibly different coding principles to make better use of available bandwidth. Since view synthesis, coupled with errors introduced by depth generation and compression, may introduce new types of artifacts that are different from those of conventional video acquisition and compression systems, it is also necessary to understand the visual quality of the views produced by DIBR techniques, which is critical to ensure a comfortable, realistic, and immersive 3D experience.

This book focuses on this depth-based 3D-TV system which is expected to be put into applications in the near future as a more attractive alternative to the current stereoscopic 3D-TV system. Following an open call for chapters and a few rounds of extensive peer reviews, 15 chapters of good quality have been finally accepted, ranging from a technical review and literature survey on the whole system or a particular topic, solutions to some technical issues, to implementation of some prototypes. According to the scope of these chapters, this book is organized into four sections, namely system overview, content generation, data compression and transmission, and 3D visualization and quality assessment, with the chapters in each section summarized below.

- Part I (Chap. 1) provides an overview of the depth-based 3D-TV system.

Chapter 1 entitled "*An overview of 3D-TV system using depth-image-based rendering*" covers key technologies involved in this depth-based 3D-TV system using the DIBR technique, including content generation, data compression and transmission, 3D visualization, and quality evaluation. It also compares the conventional stereoscopic 3D with the new depth-based 3D systems, and reviews some standardization efforts for 3D-TV systems.

- Part II (Chaps. 2–7) focuses on 3D video content creation, which specifically targets at depth map generation and view synthesis technologies.

As the leading chapter in the section, Chap. 2, entitled "*Generic content creation for 3D displays*", discusses future 3D video applications and presents a generic display-agnostic production workflow that supports the wide range of all existing and anticipated 3D displays.

Chapter 3 entitled "*Stereo matching and viewpoint synthesis FPGA implementation*" introduces real-time implementation of stereo matching and view synthesis algorithms, and describes "Stereo-In to Multiple-Viewpoint-Out" functionality on a general FPGA-based system, demonstrating a real-time high quality depth extraction and viewpoint synthesizer, as a prototype toward a future chipset for 3D-HDTV.

Chapter 4 entitled "*DIBR-based conversion from monoscopic to stereoscopic and multi-view video*" provides an overview on 2D-to-3D video conversion that exploits depth-image based rendering (DIBR) techniques. The basic principles and various methods for the conversion, including depth extraction strategies and DIBR-based view synthesis approaches, are reviewed. Furthermore, evaluation of conversion quality and conversion artifacts are discussed in this chapter.

Chapter 5 entitled "*Virtual view synthesis and artifact reduction techniques*" presents a tutorial on basic view synthesis framework using DIBR and various quality enhancement approaches to suppressing synthesis artifacts. The chapter also discusses the requirements of and solutions to real-time implementation of view synthesis.

Chapter 6 entitled "*Hole filling for view synthesis*" addresses the inherent disocclusion problem in the DIBR-based system of the newly exposed areas appearing in novel synthesized views. The problem is solved in two manners: the preprocessing of the depth data, and the image inpainting of the synthesizing view.

Chapter 7 entitled "*LDV generation from multi-view hybrid image and depth video*" presents a complete production chain for 2-layer LDV format, based on a hybrid camera system of five color cameras and two time-of-flight cameras. It includes real-time preview capabilities for quality control in the shooting and post-production algorithms to generate high quality LDV content consisting of foreground and occlusion layers.

- Part III (Chaps. 8–11) deals with the compression and transmission of 3D video data.

Chapter 8 entitled "*3D video compression*" first explains the basic coding principles of 2D video compression, followed by the coding methods for multi-view video. Next, 3D video is described with video and depth formats, special requirements, coding, and synthesis methods for supporting multi-view 3D displays. Finally, the chapter introduces the 3D video evaluation framework.

Chapter 9 entitled "*Depth map compression for depth-image-based rendering*" focuses on depth map coding. It discusses unique characteristics of depth maps, reviews recent depth map coding techniques, and describes how texture and depth map compression can be jointly optimized.

Chapter 10 entitled "*Effects of wavelet-based depth video compression*" also concentrates on the compression of depth data. This chapter investigates the wavelet-based compression of the depth video and the coding impact on the quality of the view synthesis.

Chapter 11 entitled "*Transmission of 3D video over broadcasting*" gives a comprehensive survey on various standards for transmitting 3D data over different kinds of broadcasting networks, including terrestrial, cable, and satellite networks. The chapter also addresses the important factors in the deployment stages of 3D-TV services over broadcast networks with special emphasis on the depth-based 3D-TV system.

- Part IV (Chaps. 12–15) addresses 3D perception, visualization, and quality assessment.

Chapter 12 entitled "*The psychophysics of binocular vision*" reviews psychophysical research on human stereoscopic processes and their relationship to DIBR. Topics include basic physiology, binocular correspondence and the horopter, stereo-acuity and fusion limits, non-corresponding inputs and rivalry, dynamic cues to depth and their interactions with disparity, and development and adaptability of the binocular system.

Chapter 13 entitled "*Stereoscopic and autostereoscopic displays*" first explains the fundamentals of stereoscopic perception and some of the artifacts associated with 3D displays. Then, a description of the basic 3D displays is given. A brief history is followed by a state of the art covering glasses displays through volumetric, light field, multi-view, head tracked to holographic displays.

Chapter 14 entitled "*Subjective and objective visual quality assessment in the context of stereoscopic 3D-TV*" discusses current challenges in relation to subjective and objective visual quality assessment for stereo-based 3D-TV (S-3DTV). Two case studies are presented to illustrate the current state of the art and some of the remaining challenges.

Chapter 15 entitled "*Visual quality assessment of synthesized views in the context of 3D-TV*" addresses the challenges on evaluating synthesized content, and proposes two experiments, one on the assessment of still images and the other on video sequence assessment. The two experiments question the reliability of the usual subjective and objective tools when assessing the visual quality of synthesized views in a 2D context.

As can be seen from the above introductions, this book spans systematically a number of important and emerging topics in the depth-based 3D-TV system. In conclusion, we aim to acquaint the scholars and practitioners involved in the research and development of depth-based 3D-TV system with such a most updated reference on a wide range of related topics. The target audience of this book would be those interested in various aspects of 3D-TV using DIBR, such as data capture, depth map generation, 3D video coding, transmission, human factors, 3D visualization, and quality assessment. This book is meant to be accessible to audiences including researchers, developers, engineers, and innovators working in the relevant areas. It can also serve as a solid advanced-level course supplement to 3D-TV technologies for senior undergraduates and postgraduates.

On the occasion of the completion of this edited book, we would like to thank all the authors for contributing their high quality works. Without their expertise and contribution, this book would never have come to fruition. We would also like to thank all the reviewers for their insightful and constructive comments, which helped to improve the quality of this book. Our special thanks go to the editorial assistants of this book, Elizabeth Dougherty and Brett Kurzman, for their tremendous guidance and patience throughout the whole publication process. This project is supported in part by the National Basic Research Program of China (973) under Grant No.2009CB320903 and Singapore Ministry of Education Academic Research Fund Tier 1 (AcRF Tier 1 RG7/09).

Springer
http://www.springer.com/978-1-4419-9963-4

3D-TV System with Depth-Image-Based Rendering
Architectures, Techniques and Challenges
Zhu, C.; Zhao, Y.; Yu, L.; Tanimoto, M. (Eds.)
2013, XII, 480 p., Hardcover
ISBN: 978-1-4419-9963-4