1

Introduction

Oliver Schreer¹, Jean-François Macq², Omar Aziz Niamut³, Javier Ruiz-Hidalgo⁴, Ben Shirley⁵, Georg Thallinger⁶ and Graham Thomas⁷

¹Fraunhofer Heinrich Hertz Institute, Berlin, Germany
²Alcatel-Lucent Bell Labs, Antwerp, Belgium
³TNO, Delft, The Netherlands
⁴Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
⁵University of Salford, Manchester, United Kingdom
⁶Joanneum Research, Graz, Austria
⁷BBC Research & Development, London, United Kingdom

The consumption of audio-visual media has changed rapidly in the past decade. Content is now viewed on a variety of screens ranging from cinema to mobile devices. Even on mobile devices, today’s user expects to be able to watch a personal view of a live event, for example, with a level of interactivity similar to that of typical web applications. On the other hand, current video and media production technology has not kept up with these significant changes. If we consider the complete media processing chain, the production of media, the delivery of audio-visual information via different kinds of distribution channels and the display and interaction at the end user’s terminal, many challenges have to be addressed. The major challenges are the following.

Due to reuse of video content for different distribution channels, there is a need for conversion and post-production of the content in order to cope with different screen sizes. It is widely accepted that a movie production for cinema is recorded in a significantly different way to that intended for smaller screens. However, production budgets are limited; hence complex and costly re-purposing must be avoided. A good example is the production of 3D movies, where the aim is to develop camera technologies that allow 2D and 3D capture at the same time. Approaches to multiformat production that require parallel shooting or significant manual re-editing are no longer financially viable.
The convergence of broadcast and Internet requires future media production approaches to embrace the changes brought by web-based media. The habits of media consumption have changed drastically, partially due to the availability of user interaction with users freely navigating around web pages and interactively exploring maps and views of the street for example. Hence, future media production and delivery must support interactivity.

Although the overall bandwidth available for media delivery is continuing to increase, future media services will still face limitations, particularly if the end user at home or on-the-go is considered. Hence, new distribution formats are required to allow for the provision of audio-visual media beyond current HDTV formats, to support interactivity by the end user and to support intelligent proxies in the network that are capable of performing processing, which cannot be offered by low capacity devices. First developments towards resolution beyond HD are already appearing commercially, such as 4K camera and display technologies.

In addition, the user wants to decide when, where and on which device to watch audio-visual media as nowadays a variety of devices are available (including mobiles, TV at home and immersive large projection systems in cinemas). All of these devices must be supported by media delivery and rendering. Therefore, a large variety of audio-visual formats must be provided for the full spectrum of terminals and devices taking their special capabilities and limitations into account.

Even in live events, a lot of human operators such as directors or cameramen are involved in content creation and capturing the event from different viewpoints. Due to the increasing number of productions, automated viewpoint selection may be able to make a significant contribution to limiting production costs.

A new concept appearing on the horizon that could provide answers to these issues and challenges is referred to as format-agnostic media production. The basic idea is to define a new approach to media production that supports the necessary flexibility across the whole production, delivery and rendering chain. A key aspect of this approach is to acquire a representation of the whole audio-visual scene at a much higher fidelity than traditional production systems, and to shift closer to the user-end the decision of how the content is experienced. This idea allows end users to experience new forms of immersive and interactive media by giving them access to audio-visual content with the highest fidelity and flexibility possible. This book discusses current challenges, trends and developments along the whole chain of technologies supporting the format-agnostic approach. This approach could lead to a gradual evolution of today’s media production, delivery and consumption patterns towards fully interactive and immersive media.

In Chapter 2 “State-of-the-art and Challenges in Media Production, Broadcast and Delivery”, we give an overview on the current situation in audio-visual acquisition, coding and delivery and the evolution of terminal devices at the end-user side in current media production and delivery. Based on the review of the state-of-the-art and a summary of current and upcoming challenges, the format-agnostic concept is explained. This concept offers the capability to deal successfully with the new requirements of current and future media production.

The acquisition and processing of audio-visual media following a format-agnostic approach is discussed in two separate chapters, Chapter 3 and Chapter 4. In Chapter 3 “Video Acquisition”, the three major video format parameters, spatial resolution, temporal resolution and colour depth (i.e., the dynamic range) are investigated with respect to the benefits they offer for future immersive media production. Due to the large variety of future video formats moving towards higher resolution, frame rate and dynamic range, the need for a format-agnostic
Introduction

The concept is particularly helpful in supporting media production and rendering independent of the specific format. The composition and merging of visual information from different sensors will lead to more appealing and higher quality images. In Chapter 4 “Platform-Independent Audio”, the current challenges faced in audio broadcast using a channel-based approach and sound scene reproduction techniques such as wave field synthesis are reviewed. The problem of having many competing audio formats is addressed at both the production and reproduction (user) ends. The concept of object-based audio representation is introduced and several example implementations are presented in order to demonstrate how this can be realised.

In Chapter 5 “Semi-automatic Content Annotation”, both manual and automatic content annotation technologies that support format-agnostic media production are discussed. The specific requirements on those tools, in particular under real-time constraints of live scenarios are investigated. Relevant video processing approaches such as detection and tracking of persons as well as action detection are presented. Finally, user interfaces in media production are discussed, which help the production team to perform semi-automatic content annotation.

One of the advanced concepts of media production currently under discussion and development is presented in Chapter 6 “Virtual Director”. This concept builds on various audio-visual processing techniques that allow for automatic shot framing and selection to be used at the production side or by the end user. Approaches are discussed for addressing the semantic gap between data from low-level content analysis and higher-level concepts – a process called Semantic Lifting, finally leading to content and view selection that fulfils the desires of the user.

Chapter 7 “Scalable Delivery of Navigable and Ultra-High Resolution Video” deals with the main challenges in delivering a format-agnostic representation of media. As the final decision on how content will be presented is moved closer to the end user, two factors have a significant impact on delivery: higher data rate at the production side and higher levels of interactivity at the end-user side. The chapter focuses on coding and delivery techniques, which support spatial navigation based on the capture of higher resolution content at the production side. Methods for content representation and coding optimisation are discussed in detail. Finally, architectures for adaptive delivery are presented, showing how ultra-high resolution video can be efficiently distributed to interactive end users.

Chapter 8 “Interactive Rendering” starts with a list of challenges for end user devices resulting from increased interaction with the content supported by the format-agnostic media production and delivery concept. Gesture-based interaction is one of the recent trends in interactive access to media, and this is discussed in detail. A number of technologies already on the market and currently under development are presented. This chapter concludes with user studies of gesture interfaces showing that technology development must coincide with continuous evaluation in order to meet user requirements.

Finally, Chapter 9 “Application Scenarios and Deployment Domains” discusses the format-agnostic concept from an application point of view. Based on the technologies described in the previous chapters, various application scenarios are derived. An analysis is presented of the impact of the format-agnostic concept and related new technologies in the production, network, device and end user domains. Based on this future outlook, this chapter concludes the book.

This book offers a comprehensive overview of current trends, developments and future directions in media production, delivery and rendering. The format-agnostic concept can be considered as a paradigm shift in media production, moving the focus from image to scene
representation and from professionally-produced programmes to interactive live composition driven by the end user. Therefore, this will influence how media is produced, delivered and presented leading to more efficient, economic and user-friendly ways for media to be produced, delivered and consumed. Offering new services, better accessibility to content and putting the user in control are the main aims.

The idea for this book was born in the European FP7 research project FascinatE (Grant agreement no.: FP7 248138, http://www.fascinate-project.eu), which was proposing and investigating the format-agnostic concept for the first time. Beside the editors and the co-authors, which contributed to this book, there are several other colleagues to be mentioned. Without their expertise, their ideas and the fruitful discussion over more than 5 years, this book would not have been possible. Therefore we gratefully thank the following colleagues from several institutions and companies in Europe: R. Schäfer, P. Kauff, Ch. Weissig, A. Finn, N. Atzpadin and W. Waizenegger (Fraunhofer Heinrich Hertz Institute, Berlin Germany); G. Kienast, F. Lee, M. Thaler and W. Weiss (Joanneum Research, Graz, Austria); U. Riemann (Deutsche Thomson OHG, Hannover, Germany); A. Gibb and H. Fraser (BBC R&D, London, United Kingdom); I. Van de Voorde, E. Six, P. Justen, F. Vandeputte, S. Custers and V. Namboodiri (Alcatel-Lucent Bell Labs, Antwerp, Belgium); J.R. Casas, F. Marqués and X. Suau (University Politecnica Catalunya, Barcelona, Spain); O. Juhlin, L. Barkhuus and E. Önnevall (Interactive Institute, Stockholm, Sweden); I. Drumm (University of Salford, Manchester, United Kingdom); and F. Klok, S. Limonard, T. Bachet, A. Veenhuizen and E. Thomas (TNO, Delft, The Netherlands).

The editorial team, August 2013