Ahmet Kondoz · Tasos Dagiuklas *Editors*

3D Future Internet Media

Springer

3D Future Internet Media

Ahmet Kondoz • Tasos Dagiuklas
Editors

3D Future Internet Media

Springer

Editors
Ahmet Kondoz
Faculty Engineering and Physical Sciences
Centre for Vision Speech and Signal Processing (CVSSP)
University of Surrey
Guildford, United Kingdom

Tasos Dagiuklas
Department of Computer Science
Hellenic Open University
Patras, Greece

ISBN 978-1-4614-8372-4 ISBN 978-1-4614-8373-1 (eBook)
DOI 10.1007/978-1-4614-8373-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013950183

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Contributors

Hemantha Kodikara Arachchi I-Lab Multimedia Communication Research, University of Surrey, Guildford, UK

Pedro Assunção Instituto de Telecomunicações/Instituto Politécnico de Leiria, Delegação de Leiria, Portugal

Marcus Barkowsky LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, Nantes, France

Riccardo Bassoli Instituto de Telecomunicações, Campus Universitário Santiago, Aveiro, Portugal

Konstantinos Birkos University of Patras, Patras, Greece

Patrick Le Callet LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, Nantes, France

Tasos Dagiuklas Department of Computer Science, Hellenic Open University, Patras, Greece

Carl James Debono Department of Communications and Computer Engineering, University of Malta, Msida, Malta

Athanasios M. Demiris Micro2gen Ltd., Technology Park NCSR Demokritos, Agia Paraskevi, Greece

Gabriel Dosso VITEC Multimedia, Châtillon, France

Erhan Ekmekcioglu I-Lab Multimedia Communications Research Group, University of Surrey, Guildford, Surrey, UK

Ignacio Digón Escudero Telefónica I+D, Madrid, Spain

Sérgio Faria Instituto de Telecomunicações/Instituto Politécnico de Leiria, Delegação de Leiria, Portugal

Georgios Gardikis NCSR "Demokritos", Institute of Informatics and Telecommunications, Aghia Paraskevi, Attiki, Greece

Michael Grafl Alpen-Adria-Universitaet Klagenfurt, Universitätsstraße, Klagenfurt, Austria

Chaminda T.E.R. Hewage Kingston University, London, UK

Chungeun Kim I-Lab Multimedia and DSP Research Group, Centre for Vision, Speech and Signal Processing (CVSSP), University of Surrey, Guildford, UK

Ahmet Kondoz Faculty Engineering & Physical Sciences, Centre for Vision Speech and Signal, University of Surrey, Guildford, UK

Athanasios Kordelas Department of Electrical & Computer Engineering, University of Patras, Patras, Greece

Jing Li LUNAM Université, Université de Nantes, IRCCyN UMR CNRS 6597, Polytech Nantes, Nantes, France

Evariste Logota Instituto de Telecomunicações, Campus Universitário de Santiago, Aveiro, Portugal

Asimakis Lykourgiotis Department of Electrical and Computer Engineering, University of Patras, Rio, Hellas

Hugo Marques Instituto de Telecomunicações, Campus Universitário de Santiago, Aveiro, Portugal

Instituto Politécnico de Castelo Branco, Castelo Branco, Portugal

Maria G. Martini Kingston University, London, UK

Brian Walter Micallef Department of Communications and Computer Engineering, University of Malta, Msida, Malta

Moustafa M. Nasralla Kingston University, London, UK

Evangelos Pallis TEI of Crete, Department of Applied Informatics and Multimedia, Heraklion, Crete, Greece

Fernando Pascual Blanco Telefónica I+D, Madrid, Spain

Luís Pinto Instituto de Telecomunicações/Instituto Politécnico de Leiria, Delegação de Leiria, Portugal

Ilias Politis Department of Electrical & Computer Engineering, University of Patras, Patras, Greece

Jonathan Rodriguez Instituto de Telecomunicações, Campus Universitário de Santiago, Aveiro, Portugal

Manuel Nuñez Sanz Telefónica I+D, Madrid, Spain

Nicolas Tizon VITEC Multimedia, Châtillon, France

Chapter 1
Introduction

Ahmet Kondoz and Tasos Dagiuklas

There have been many significant advances in 3D media technologies in terms of capturing, representing, coding, transmitting, and visualizing for 3D displays. 3D media have been evolving in various areas, covering diverse market segments such as professional (e.g., scientific, medical, education, and training) and entertainment (3D interactive gaming, broadcasting, social networking, etc.) sectors. These new technologies provide the ability to design and develop applications ranging from virtual collaborative environments (e.g., multimodal interactions, seamless application on generation) to edutainment (e.g., 3D telepresence, combination of the educational content along with entertainment).

Since the Internet has grown beyond its original design objectives due to the increasing demand for performance, availability, scalability, security, and reliability, it progressively reaches a set of fundamental technological (evolution in wireless and mobile networking technologies) and operational limitations (e.g., exhausting the number of possible IP addresses). The Internet was designed for purposes that bear little resemblance to today's usage scenarios and related traffic patterns.

Several organizations have been working towards the development of the Future Internet. Many research projects have been established worldwide. For example, in Europe the FIRE [1], in the USA the FIND [2] and GENI [3], and in Japan and Korea the AKARI [4] programs are the main drivers for the definition of the characteristics and capabilities of the Future Internet. In Europe, the Future Internet Research and Experimentation (FIRE) program, the Future Internet

A. Kondoz
Faculty Engineering & Physical Sciences, Centre for Vision Speech and Signal Processing (CVSSP), University of Surrey, Guildford GU2 7XH, UK
e-mail: a.kondoz@surrey.ac.uk

T. Dagiuklas (✉)
Department of Computer Science, Hellenic Open University, Parodos Aristotelous 18, Patras 26335, Greece
e-mail: dagiuklas@eap.gr

A. Kondoz and T. Dagiuklas (eds.), *3D Future Internet Media*,
DOI 10.1007/978-1-4614-8373-1_1,

Assembly, and the Future Media 3D Internet Task Force are the main endeavors to provide the fundamental foundations for the Future Internet research [5]. These worldwide initiatives take different approaches to Internet evolution as part of their core objectives but are all related to technological and socioeconomic scenarios as envisioned today.

Future Internet is expected to provide the following characteristics:

1. Internet is experiencing a significant shift from PC-based computing to mobile networks. Mobile data is growing at an unprecedented rate well beyond the capacity of today's 3G, 3G+, and 4G networks. Therefore, mobility has become the key driver for the Future Internet, and convergence demands are increasing across heterogeneous networks.
2. Cognitive radio networks aim to increase spectral efficiency and provide radically new and more efficient wireless access methods.
3. Cooperative communications have recently emerged as strong candidate technologies for many future wireless applications in order to improve performance and throughput of wireless networks, respectively. In particular, theory and experiments have shown that they can be extremely useful for wireless networks with disruptive channel and connectivity conditions.
4. In the last decade, Internet underwent a huge evolution, mainly related to services. The concept of Web 2.0, that is, the second generation of Internet, is led by the generation of new services such as user-generated content, applications, and services, social networks, collaborative webs, and Web TV, together with the growing interactive sets of applications. The new environment has a strong impact on the concept of end-to-end services at all levels, from the applications up to the networks and their underlying technologies.
5. New methods of delivery of content towards the end users. This includes the use of content/information-centric networking architecture making the communication network as location independent possible with the use of distributed architecture (e.g., P2P) associated with streaming, data sharing, routing, and forwarding.
6. Multimedia applications and services make strong demands for cloud computing because of the significant amount in terms of computation resources required to serve millions of both Internet and/or mobile users concurrently. In this cloud-based multimedia-computing paradigm, users are able to store, process, and adapt their multimedia application data in the cloud in a distributed manner. The main drivers of the media-aware cloud architecture are simplicity, efficiency, and scalability. Demands for a media cloud are different from the demands for clouds in other industries.

1.1 3D Media Internet

The 3D Media Internet will play a significant role in achieving the key features of Future Internet [6]. For example, the augmented virtual worlds, the collaborative platforms, and the moving holograms created in 3D Internet will originate new

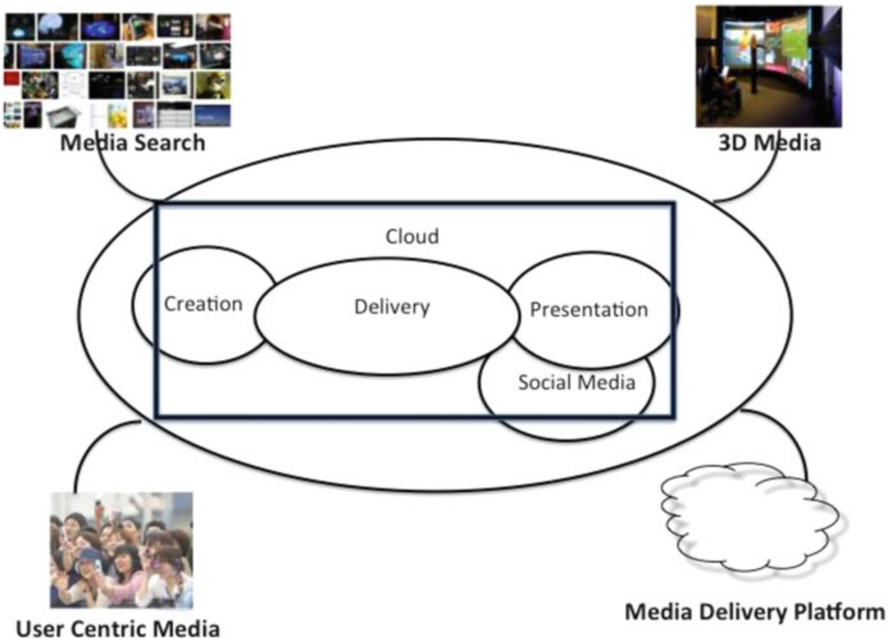

Fig. 1.1 3D media challenges

requirements in terms of information representation and metadata identification. Moreover, the new services and applications will make new demands on ubiquitous user interfaces that will have to support novel inputs (e.g., 3D position sensors), displays (e.g., multiview displays), and presentation (e.g., augmented reality) modalities in different kinds of terminals (e.g., European Commission, 2008).

In these respects, the Future Media Internet will be much more than simply a faster way to go online. It will be designed to overcome current limitations and address emerging trends in areas such as network architecture, content and service mobility, diffusion of heterogeneous nodes and devices, mass digitization, new forms of (3D) user centric/user-generated content provisioning, and emergence of software as a service including interaction with improved security, trustworthiness, and privacy. Figure 1.1 illustrates four 3D Media challenges associated with Media Search, 3D Media, User Centric Media, and Media Delivery Platforms. In order to meet these challenges, the enabling technologies include content creation, delivery technologies, presentation, social media, and cloud computing.

The Future 3D Media Internet has generated a significant amount of recent research goals to overcome current limitations. The resulting innovations could include the following:

1. *3D Content Creation*: In terms of 3D content, the methodological objective is to combine rich interaction-based content production and game-related solutions in order to contribute towards the design and development of novel applications

and services. Instead of the technologically driven approach, content-oriented service development is one of the key requirements for successful Internet applications. A basic understanding and solid background for user behavior can be achieved with the aid of the rich interaction model. Visual elements (look) include texture, light, shadows, reflections, colors, composition, depth perception, shapes, structure, and contrast. Examples of functional concepts (feel) include responsiveness, overall end-user experience, feel of control, how the audiovisual material meets the expectations of the end user, movements, emotional experiences, and affective mechanisms. As opposed to off-line rendering techniques to produce 3D movie effects, in augmented reality applications, the challenge is to perform video processing in real time.

2. *User-Generated Content and Personalization*: Future Internet should provide mechanisms embedded into the network to ease the personalization, adaptation, accessibility, and search aspects but also protect and enforce intellectual property rights related to user-generated content.
3. *Presentation*: In addition it should facilitate a smooth transition from 2D content to 3D content and ease the user participation in 3D content generation and fruition within enhanced 3D collaborative environments. Furthermore, the Future Internet should empower communities to achieve dynamic content creation, provisioning, and sharing (e.g., in social media).
4. *Media Cloud*: Existing cloud computing technologies are not particularly media/video capable. Handling of multiple video flows in terms of encoding, processing, and streaming is a much larger problem that stresses computing infrastructure due to large data and bandwidth requirements. Media cloud infrastructure must be capable of supporting all the functionalities associated with the entire value chain from capturing to processing (encoding, content protection) to content delivery and by greater simplicity to optimize media-related and network infrastructure.
5. *3D Delivery*: P2P technologies have the potential to provide a more cost-effective and flexible delivery solutions for future 3D entertainment services as well as giving users fast interaction with the content and with collaborating partners in their social network. Moreover, recent advances in wireless technologies (e.g., LTE, LTE-A, 5G) will offer the possibility of 3D video to mobile users.
6. *Social Media*: Social networks will play an increasingly important role in the Internet of the future. Internet users will have their own online identity, which will carry them through from one network to the next. Moreover Social Media expects to provide a new type of functionalities and capabilities such as 3D Social Gaming, 3D Life-streaming, and Live-casting.

This book presents recent advances in the area of Future 3D Media Internet. The first four chapters are devoted to the area of 3D Media Coding and presentation. Chapter 2 presents techniques for 3D Media Acquisition and depth map processing. Chapter 3 is a survey of methods for merging the real and the synthetic in augmented 3D worlds. Finally, Chaps. 4 and 5 present techniques to encode 3D Video and Spatial Audio respectively.

The following chapters cover Networking Aspects for 3D Media. Chapter 6 presents current and future developments in the area of transport protocols for 3D video. Chapter 7 presents application-layer filtering techniques through Media Aware Network Element to optimize 3D Video Delivery. Chapter 8 surveys the techniques used for P2P streaming. Finally, Chap. 9 examines the impact of IP Mobility and mobility management protocols on the quality of 3D video.

The third section presents QoE and QoS advances for 3D Media. Chapter 10 discusses the use of QoS/QoE support in P2P overlays. Chapter 11 makes a survey of QoE methodologies for 3DTV applications and services. Finally, Chap. 12 presents error concealment strategies and techniques to improve QoE in multiview video applications.

The last section presents two 3D applications. Chapter 13 describes the use of 3D images and video in medical surgery and training applications to improve diagnosis and surgery operation. Finally Chap. 14 presents innovations in 3DTV capturing, data representation, compression, transmission, and rendering required for a technically efficient and commercially successful 3DTV broadcast system.

References

1. http://www.ict-fire.eu
2. http://www.nets-find.net/
3. http://www.geni.net/
4. http://en.wikipedia.org/wiki/AKARI_Project
5. Tselentis, Georgios (ed) (2009) *Towards the Future Internet: A European Research Perspective*. IOS press
6. Zahariadis, Theodore, Petros Daras, Isidro Laso-Ballesteros (2008) *Towards future 3d media internet*. *NEM Summit*:13–15

Part I
3D Media Coding and Presentation

Chapter 2
3D Media Representation and Coding

Pedro Assunção, Luís Pinto, and Sérgio Faria

Abstract Nowadays, three-dimensional (3D) multimedia provides immersive user experiences in virtual and real 3D environments, based on advanced technology that is rapidly evolving towards different fields of application. An important element in 3D multimedia is undoubtedly the visual information, where 3D video plays a major role. This chapter addresses such element of 3D multimedia, by presenting a comprehensive description of the most common 3D video formats used in various 3D multimedia services and applications. Uncompressed 3D video representation is described, specifically focusing frame compatible formats used for backward compatibility with 2D systems, followed by those formats that explicitly include depth information, either in single-view or multiview representations. Then an overview of standard coding algorithms, currently used for 3D video coding, is presented along with a discussion of their main characteristics in terms of processing methods and performance. Since the response of the human perceptual system to 3D visual content includes specific features different from those already known from 2D perception, the last sections of the chapter describe recent studies dealing with asymmetric representation and coding of stereoscopic video. The use and adaptation of standard coding methods to benefit from asymmetric characteristics of the human visual system is presented and discussed in the light of recent research advances. The chapter concludes by highlighting the most relevant issues in the current context of 3D video representation and coding.

2.1 Introduction

The human perception of the real world is three dimensional (3D) and necessarily includes information about volume and depth, which is not explicitly included in classic representations of natural scenes using digital multimedia signals. The huge

P. Assunção (✉) • L. Pinto • S. Faria
Instituto de Telecomunicações/Instituto Politécnico de Leiria, Delegação de Leiria, Portugal
e-mail: paassunc@ieee.org

A. Kondoz and T. Dagiuklas (eds.), *3D Future Internet Media*,
DOI 10.1007/978-1-4614-8373-1_2, © Springer Science+Business Media New York 2014

amount of structured data required to represent 3D visual scenes leads to the need of standard representation formats in both uncompressed and compressed domains. Since 3D video is the most common type of media content used to provide 3D immersion experience, this is also a driving factor of technology evolution in several domains, ranging from high-resolution 3D cinema and 3D television to small screen applications (e.g., games) using autostereoscopic displays.

Recent evolution of 3D media services along with increasing penetration of 3D-ready equipment in the consumer market leads to the coexistence of emerging 3D systems with legacy ones. Several 3D representation formats are currently used to enable efficient coding for storage and transmission across interoperable systems also enabling operation with equipment in different technological evolution stages either in the segment of professional or consumer market.

The most common 3D video formats are described in the next sections, focusing frame compatible formats used for backward compatibility with 2D systems followed by specific formats explicitly including depth information either in single-view or multiview representations. Then an overview of standard coding algorithms currently used for 3D video is presented along with their main characteristics in terms of processing architectures and coding performance. Since the characteristics of the human perceptual system regarding 3D visual content include specific features, different from those known from 2D perception, the last sections of this chapter describe recent studies dealing with asymmetric representation and coding of stereoscopic video. The use and adaptation of standard systems to benefit from asymmetric characteristics of the human visual system is presented and discussed in the light of recent research advances.

2.2 Raw Formats for 3D Video

There are several different formats used to represent raw 3D video. The common requirement is that all of them must provide means for stereoscopic viewing, since this is the underlying principle behind depth perception from visual information. In the following sections, different formats are described including those based on stereo views, 2D views plus explicit depth information, and multiple views also associated with depth maps.

2.2.1 Frame Compatible Formats

In the context of 3D multimedia services and applications, 3D video representation through frame compatible formats is a key factor to guarantee compatibility with existing 2D video networking technology and equipment. Successful deployment of 3D video delivery services and applications is enabled by making possible transmission of 2D-compatible formats over current networks and legacy decoders with

Table 2.1 Standard frame compatible formats

ID	Compatible format
0	Checkerboard
1	Column-based interleaving
2	Row-based interleaving
3	Side-by-side
4	Top-bottom
5	Temporal interleaving
6	2D
7	Tile format

3D-ready displays already common in the user market. By using frame compatible formats, seamless compression of 3D video content is also accomplished with existing 2D encoders without the need to modify the coding algorithms.

In the case of stereoscopic video, representation in 2D-compatible formats is achieved by multiplexing the two different views into a temporal sequence of 2D signals. This means merging two different views into a classic sequence of single-frame representation. If the full resolution of the two views is maintained, then such representation format has twice the resolution of its equivalent 2D. However, taking into account that good perceived quality of 3D video does not necessarily require two high-quality views, either one or both views may be subsampled in one dimension in order to fit two high-definition (HD) frames into only one HD frame slot [1]. Identification of left and right views is done via specific signaling, used to distinguish the data representing each one. Using H.264/AVC to encode 3D frame compatible formats, the recommended signaling method is the use of supplemental enhancement information (SEI) messages, as shown in Table 2.1, where the *frame_packing_arrangement_type* field of the SEI message is defined according to subclause D.2.25 in the standard [2]. SEI messages are used to convey information about how decoders should handle their output according to the frame-packaging scheme used. There is also an SEI value defining 2D format, which enables switching between stereo and non-stereo content. Additionally to the type of content, the SEI message includes other fields such as the arrangement *id* that can be used to identify which frame is the left or right view.

In the following subsections, these frame compatible formats are described along with their main characteristics.

2.2.1.1 Side-by-Side

The side-by-side format is shown in Fig. 2.1a. In this format, the two stereoscopic views are concatenated side by side, giving rise to a single 2D matrix of pixels with the same resolution in the vertical direction while in the horizontal one there is twice the number of pixels of each single view.

However, since doubling the spatial resolution has strong implications in compressed rates, the horizontal resolution of the original views might be halved

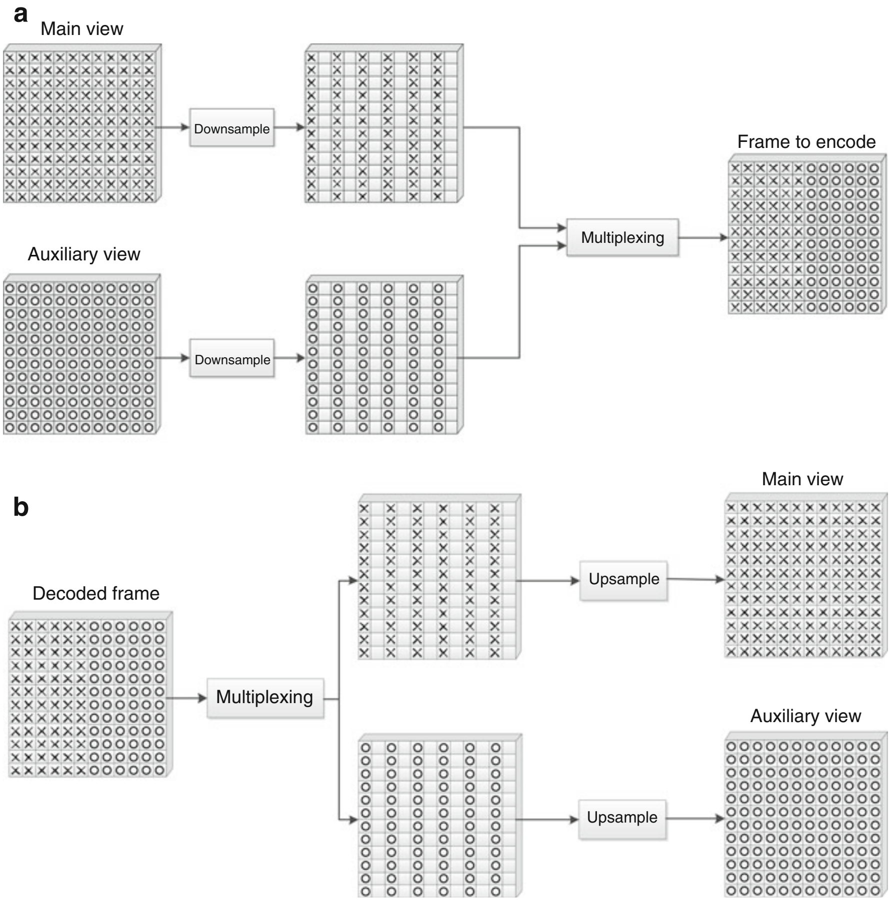

Fig. 2.1 (**a**) Side-by-side packing arrangement. (**b**) Side-by-side up-conversion process

through downsampling before packing into this side-by-side arrangement for subsequent encoding and transmission. Correspondingly, the counterpart up-conversion process must be done after decoding to display the full-resolution stereo images, as shown in Fig. 2.1b.

A different version of the side-by-side arrangement can be accomplished by sampling the stereoscopic views using a quincunx pattern. In this case, even though the horizontal resolution is also reduced to half of the original, still half of each view columns is maintained, as shown in Fig. 2.2. Quincunx sampling relates better with the characteristics of the human visual system (HVS) in terms of frequency domain representation. Thus, this sampling pattern preserves more relevant information from the original signal, which has the potential to result in better perceived quality.

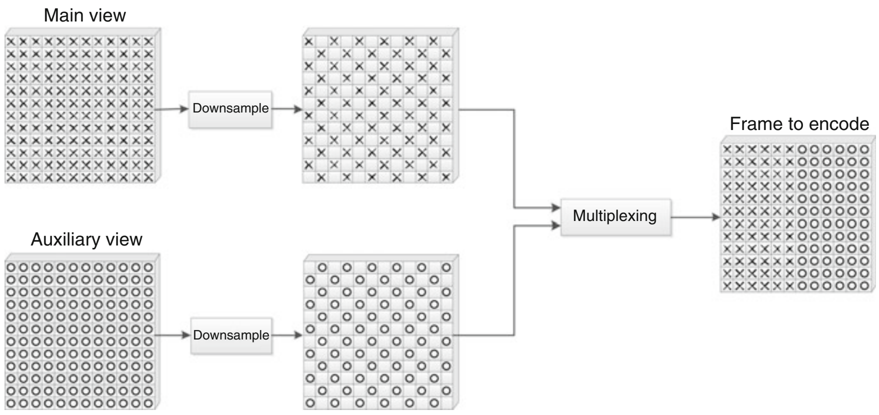

Fig. 2.2 Side-by-side with quincunx arrangement

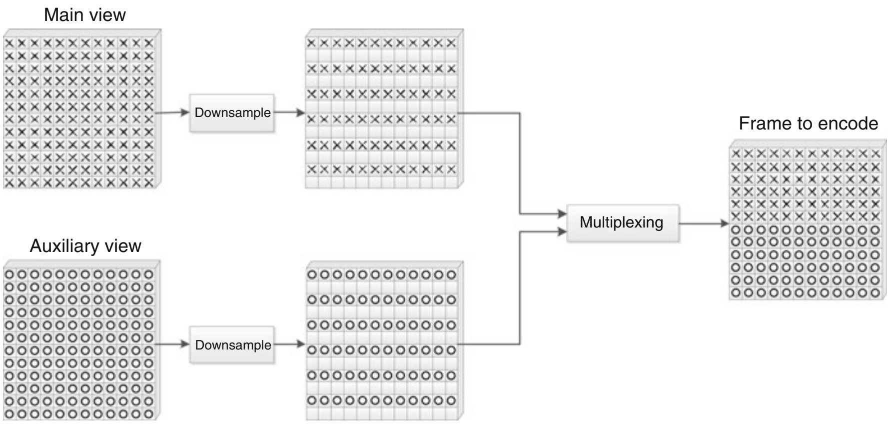

Fig. 2.3 Top-bottom arrangement

2.2.1.2 Top-Bottom

The top-bottom format is based on the same concept as side-by-side, but in this case downsampling is applied to the vertical resolution and the resulting frames concatenated as shown in Fig. 2.3.

Unless otherwise specified by the SEI message, in standard top-bottom format the downsampled left view is concatenated into the first half (top) of a composite frame while the downsampled right view is concatenated into the bottom half. This 3D frame compatible format should not be used with interlaced source material because the vertical resolution of interlaced fields is already half of the full-resolution frame and further downsampling in this dimension could incur in too much loss of spatial detail.

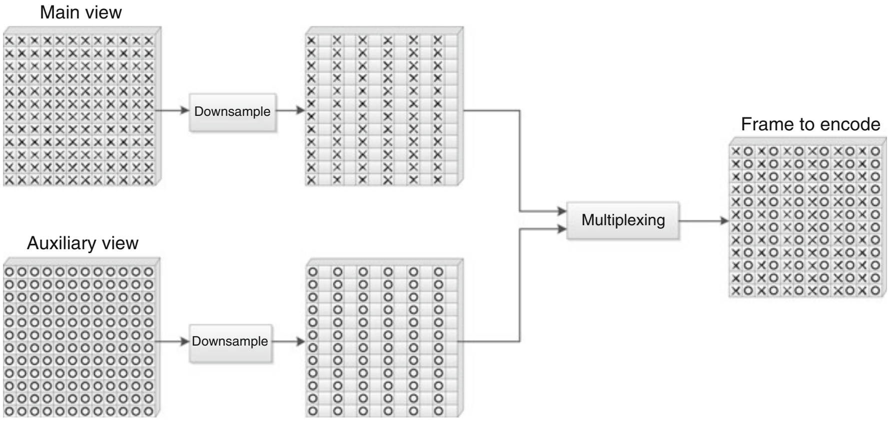

Fig. 2.4 Column interleaving arrangement

Both side-by-side and top-bottom formats are preferred for production and distribution of 3D content in comparison with the ones described next, based on spatial interleaving. This is because interleaving is prone to cross-talk artifacts and color bleeding.

2.2.1.3 Interleaved Formats

Interleaving methods provide higher correlation in the composite frame by multiplexing both downsampled views, either vertically or horizontally according to the downsampled dimension. If both views are half sampled in the horizontal dimension, then a column interleaving arrangement is reached, as shown in Fig. 2.4. Otherwise, if downsampling is performed in the vertical dimension, then multiplexing is done row by row, attaining a row interleaving arrangement for the composite frame.

These interleaving methods can be further combined in order to create a multiplexed frame-like checkerboard such as the one depicted in Fig. 2.5. In this type of format, each view must be downsampled using checkerboard nonmatching patterns. In the case of the left view, this means that in odd rows each other pixel should be kept starting from odd columns, while in even rows each other pixel should kept, starting from even columns. In the case of the right view the complementary pattern must be used.

Other frame compatible arrangement is based on interleaving in the temporal dimension. In this type of interleaving the frame rate of each view is reduced to half of its original rate and then the even frames from the left view are temporally multiplexed with the odd frames from the right view, as shown in Fig. 2.6. In this type of format the spatial resolution of the original views is maintained. It can be particularly suitable to represent low motion 3D content, where frame rate is not a very relevant requirement.

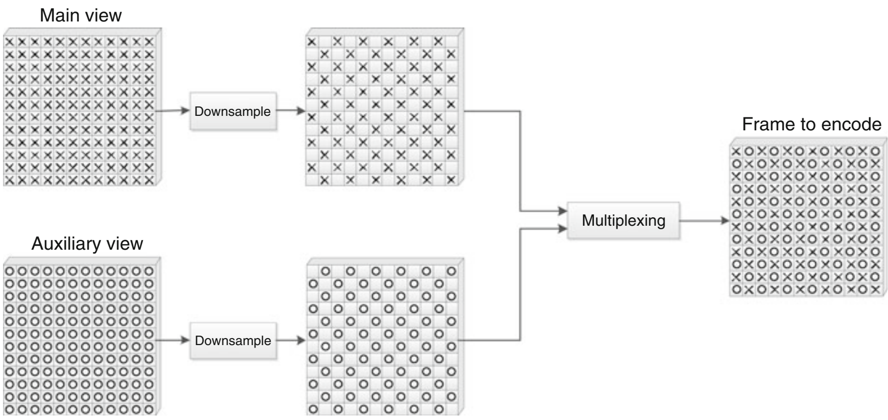

Fig. 2.5 Checkerboard arrangement format

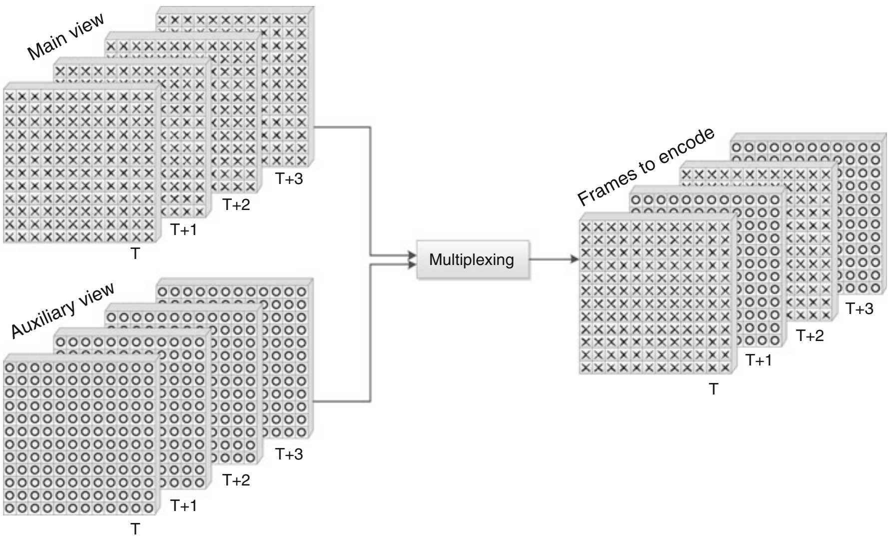

Fig. 2.6 Temporal interleaving frame arrangement

2.2.1.4 Tile Format

The last amendment of H.264/AVC in regard to the use of frame compatible formats introduced the tile format [3, 4]. The arrangement depicted in Fig. 2.7 allows two HD frames (1,280 per 720 pixels) to be packed into a Full HD frame

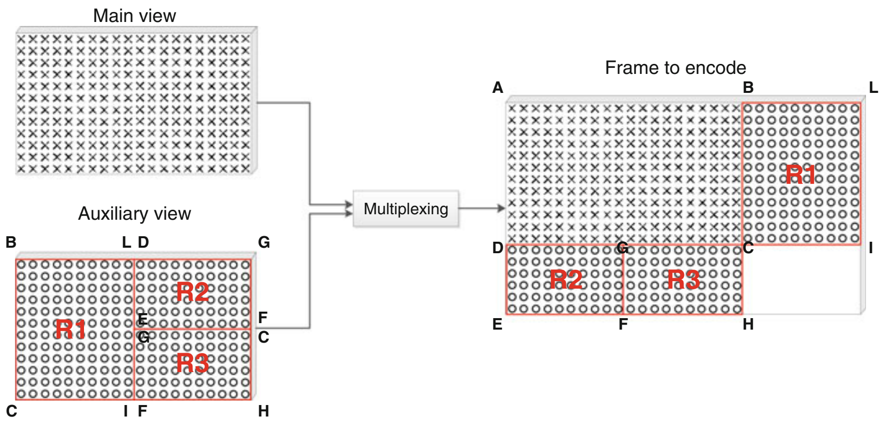

Fig. 2.7 Tile frame arrangement

(1,920 per 1,080 pixels) using a tiling method, where different regions of one view are tiled with the other view.

As seen in the figure, the left view is located at the top left corner of the Full HD frame without any type of preprocessing. The right view is then divided into three regions (R1, R2, and R3), which are placed in specific regions of the resulting Full HD frame.

The great advantage of this method is the backward compatibility with legacy 2D devices, as it requires only a 720p crop to obtain a 2D version of the content in HD resolution. Moreover there are no downsampling operations involved, which means that full resolution of the original frames is maintained in all dimensions.

A potential drawback introduced by this method is lower coding efficiency and annoying artifacts in the coded images due artificial edges created by tiling. However objective tests conducted show that these are not problematic as impairments are not noticeable (comparing with simulcast), mainly above 1 Mbps [4].

2.2.2 Video Plus Depth

An alternative to stereoscopic representation of 3D video consists in two separate 2D signals to convey color image information and the depth associated to each pixel, i.e., the distance to the camera [5]. Such information being available at display side can enable the generation of virtual views through depth-based image rendering techniques. Known as video plus depth ($V + D$), this format has implicit higher complexity than stereo views because it requires either additional computation to obtain the depth values from multiple views of the scene or specific image acquisition hardware to obtain the depth maps directly from the scene, e.g., using hybrid camera systems [6].

Fig. 2.8 Sequence breakdance: video (*left*) plus depth (*right*)

Depth values are usually represented as integers in the range of 0–255, thus using 8 bits per pixel which results in a gray-scale depth map, as shown in Fig. 2.8. These values translate to the maximum and minimum distance of each point. A warping function should be used to reconstruct a stereoscopic sequence by synthesizing the other view of a stereo pair from the color image and the corresponding depth map. Depth image-based rendering (DIBR) is a method commonly used for this purpose [7]. In view synthesis using DIBR there are some problems that may result in image distortions, such as the possibility of occlusions due to the lack of unique texture data that may be needed to render the other stereo view through the depth map.

Separate encoding of each signal (video and depth) is possible by a standard monoscopic codec, such as H.264/AVC. In regard to encoding the depth map, the gray-scale values can be given to the encoder as the luminance component of pseudo video signal where chrominances are set to a constant value. Since the color video is encoded as regular monocular video, this format has inherent backward compatibility with legacy decoders. This format allows extended possibilities at the receiving side compared to the traditional stereo video. For instance, it is possible to adjust the amount of depth perceived by viewers by adjusting view synthesis or to render several different virtual views for multiview displays.

2.2.3 *Multiview Video Plus Depth*

As mentioned before, the $V + D$ format is particularly suited to multiview displays because it enables generation of different virtual views of the same scene. However, if a wide range of views is required, the $V + D$ format is no longer suitable because not many different views can be rendered with enough quality from only one view and corresponding depth map. This is because the original view may become farther away than the one to be synthesized producing visible artifacts due to occlusions, which cannot be properly handled in such cases. This is mainly relevant in wide-range multiview (autostereoscopic) displays or free viewpoint video applications.