For private use. Final DOI: https://doi.org/10.1007/978-3-319-77842-6_8

Chapter 8
Transmission of 3D Video Content

Emil Dumic, Anamaria Bjelopera, Khaled Boussetta, Luis A. da Silva Cruz, Yuansong Qiao, A. Murat Tekalp and Yuhang Ye

Abstract This chapter describes different video transport technologies that support the existing 3D video formats, such as frame-compatible side-by-side and multi-view video plus depth. Particular emphasis is given to the DVB systems (terrestrial, satellite, and cable) and IP transport, focusing HTTP/TCP streaming, adaptive HTTP streaming, RTP/UDP streaming, P2P Networks, and Information-Centric Networking-ICN. Hybrid transport technologies, combining broadcast and broadband networks for video delivery are also addressed. The chapter highlights important aspects of 3D video transmission over wireless

E. Dumic (✉)
Department of Electrical Engineering, University North, 42000 Varaždin, Croatia
e-mail: emil.dumic@gmail.com

A. Bjelopera
Department of Electrical Engineering and Computing, University of Dubrovnik, 20000 Dubrovnik, Croatia
e-mail: anamaria.bjelopera@unidu.hr

K. Boussetta
L2TI, University of Paris, 13, 99, Avenue J-B Clement, 93430 Villetaneuse, France
e-mail: Khaled.Boussetta@univ-paris13.fr

L. A. da Silva Cruz
Department of Electrical and Computer Engineering, Instituto de Telecomunicações, University of Coimbra, Pólo II, 3030-290 Coimbra, Portugal
e-mail: lcruz@deec.uc.pt

Y. Qiao · Y. Ye
Software Research Institute, Athlone Institute of Technology, Athlone, Ireland
e-mail: ysqiao@research.ait.ie

Y. Ye
e-mail: yye@research.ait.ie

A. Murat Tekalp
Koç University, Istanbul, Turkey
e-mail: mtekalp@ku.edu.tr

P. A. Assunção and A. Gotchev (eds.), *3D Visual Content Creation, Coding and Delivery*, Signals and Communication Technology,
https://doi.org/10.1007/978-3-319-77842-6_8

networks, together with their benefits and limitations in the delivery of this type of content. Recent research results are summarized for different delivery systems and transport technologies.

8.1 Introduction

With the rapid development of different 3D video content, its transmission and delivery to the end user become a challenging problem to the existing network systems due to the ever-increasing capacity needs, different errors that may arise in the transmission chain, used format of 3D content, etc. This chapter describes different video transport technologies that support most of the existing 3D video formats (e.g., side-by-side, multi-view + depth). The DVB systems (terrestrial, satellite, and cable) are addressed and also IP transport (HTTP/TCP streaming, adaptive HTTP streaming, RTP/UDP streaming, P2P Networks, and Information-Centric Networking-ICN), hybrid transport technologies (a combination of broadcast and broadband video delivery), and 3D video transmission over wireless networks, together with their benefits and limitations in 3D video delivery content.

First, in Sect. 8.2 the main processing blocks of DVB systems (DVB-T/T2, DVB-S/S2, and DVB-C/C2) are presented, along with transport of 3D video (side-by-side, multi-view + depth), from compressed video, audio, and ancillary information that form elementary streams, packetized elementary streams. The most important specifications of Program Stream (PS) or Transport Stream (TS), as defined in MPEG-2 Systems are also addressed. Hybrid transport technologies to pool existing DVB broadcast of stereo 3D TV and IP streaming for multi-view 3DTV broadcast applications are highlighted in Sect. 8.3. In Sect. 8.4, IP transport technologies will be discussed in more detail, like multi-casting, content-distribution networks (CDN), peer-to-peer (P2P) streaming. In this section, HTTP/TCP streaming, adaptive HTTP streaming, and RTP/UDP streaming are also explained. Section 8.5 describes Information-Centric Networking (ICN) concept. Section 8.6 describes support of 3D stereo and multi-views video over wireless networks and finally, Sect. 8.7 gives conclusions.

8.2 DVB-T/T2, C/C2, and S/S2 Systems

Terrestrial, cable or satellite broadcast have been the most commonly used delivery method for bringing 3D TV content to home and in the technological context the Digital Video Broadcasting–Terrestrial (DVB-T) standard is of utmost importance, as defined by the European Telecommunications Standards Institute (ETSI) standard EN 300 744 [1], DVB-T2 in ETSI EN 302 755 [2], Digital Video Broadcasting-Satellite (DVB-S) in ETSI EN 300 421 [3], DVB-S2 in ETSI EN 302

307 [4], Digital Video Broadcasting-Cable (DVB-C) in ETSI EN 300 429 [5], and DVB-C2 in ETSI EN 302 769 [6]. DVB systems today usually carry stereo video in one of the frame-compatible formats that combines the left and right video sequences in one high-definition (HD) stream.

- SbS (side-by-side)—left and right images are one next to the other in an HD image;
- TaB (top-and-bottom)—put left and right images one above the other in a HD image.

A frame-compatible stereoscopic video format combines the left-eye and right eye images in a spatial multiplex arrangement which results in a composite image that can be treated like a conventional high-definition television (HDTV) image by the receiver demodulator and compression decoder. SbS and TaB formats are compatible with actual HD systems and can be transmitted using current DVB standards. Of course, at the receiver side, one needs to have 3D monitor to properly show 3D video. Regarding how a decoded signal is sent to a stereo display, current stereoscopic systems usually use a frame-sequential 3D signal. Left and right frames are alternately sent to the display and by diverse systems like active shuttered glasses or polarized glasses are then shown to each eye. This means that the real frame frequency is half the video frame frequency. New standard for DVB-3DTV that is currently being developed, ETSI TS 101 547, in Part 2 describes frame-compatible stereoscopic 3DTV formats (SbS, TaB), ETSI TS 101 547-2 [7].

8.2.1 DVB-T

DVB-T is the DVB European-based consortium standard for the broadcast transmission of digital terrestrial television [1]. This system is used for transport of compressed digital audio, video, and other data, multiplexed into a Moving Picture Experts Group (MPEG) TS [8, 9] using Coded Orthogonal Frequency Division Multiplexing (COFDM) modulation [10]. Rather than carrying the data on a single radio frequency (RF) carrier, Orthogonal Frequency Division Multiplexing (OFDM) works by splitting the digital data stream into a large number of slower digital streams, each of which digitally modulates a set of closely spaced adjacent carrier frequencies. In the case of DVB-T, there are two choices for the number of carriers known as 2K-mode or 8K-mode (4K-mode is rarely used) [1].

A DVB-T transmitter shown in Fig. 8.1 (taken from ETSI EN 300 744 [1]), consists of the following signal processing blocks, explained in more detail in [11].

- *Source coding and MPEG-2 multiplexing*: Compressed video, audio, and private data streams form elementary streams (different video and audio compression). Elementary streams are first cut into packetized elementary streams (PES) and afterward multiplexed into MPEG-2 transport stream (MPEG-2 TS) [12]. Each TS packet is 188 bytes long and can contain data from only one PES

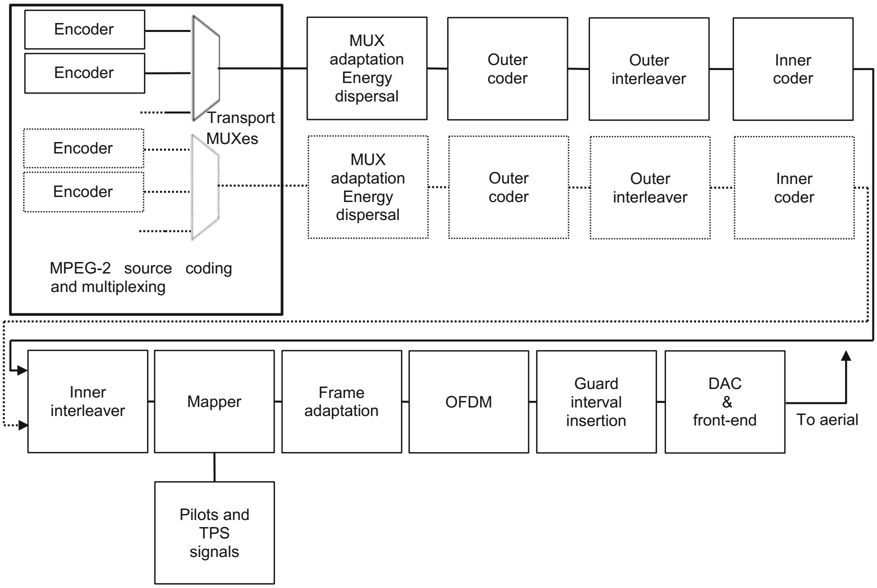

Fig. 8.1 DVB-T transmitter block diagram

packet. The system also allows two-level hierarchical channel coding and modulation, including uniform and multi-resolution constellation. In this case, the functional block diagram of the system shall be expanded to include the modules shown dashed in Fig. 8.1.

- *External encoder (RS encoder)*: Reed–Solomon coding or RS (204, 188, T = 8) code is used which is a shortened version of the code RS (255, 239, T = 8). Reed–Solomon code RS (204,188) uses 16 parity bytes and it can correct up to eight erroneous bytes per packet.
- *External interleaver (convolutional interleaver, I = 12)*: Convolutional interleaver rearranges the transmitted packets with the purpose to increase the efficiency of the Reed–Solomon decoding by spreading the burst errors introduced by the channel over a longer time.
- *Internal encoder (Punctured Convolutional Code)*: Uses convolutional coding and is an efficient complement to the Reed–Solomon coder and external interleaver. Possible Forward Error Correction (FEC) codes: 1/2, 2/3, 3/4, 5/6, and 7/8.
- *Internal interleaver*: Two separate interleaving processes are used to reduce the influence of burst errors, one operating on bits (bit interleaver) and another on groups of bits (symbol interleaver).
- *Mapper (+ pilots and Transmission Parameter Signaling (TPS) carriers)*: All data carriers in one OFDM frame are modulated using Quadrature Phase-Shift Keying (QPSK), 16QAM (Quadrature Amplitude Modulation), 64QAM, nonuniform 16QAM or nonuniform 64QAM constellations.

- *OFDM Transmitter and Guard Interval Insertion*: In DVB-T, OFDM usually uses 2048 or 8192 carriers (2K and 8K mode). Insertion of the guard interval extends symbol duration by 1/4, 1/8, 1/16 or 1/32 to give the total symbol duration. Cyclic prefix serves as a guard interval and eliminates the intersymbol interference from the previous symbol.
- *DAC (digital to analogue converter) and front-end*: Digital signal is transformed into an analogue signal with a DAC and then modulated to radio frequency (very high frequency (VHF), ultra high frequency (UHF)) by the RF front-end. The occupied bandwidth is designed to accommodate DVB-T signal into 5, 6, 7, or 8 MHz channels.

DVB-T receiver consists of below mentioned signal processing blocks, which are as follows:

- Front-end and ADC (analogue to digital converter);
- Time and frequency synchronization;
- Guard interval disposal and OFDM Receiver;
- Channel Estimator and Channel Compensation;
- Demapper;
- Inner Deinterleaver;
- Internal Decoding (Viterbi Decoder);
- External Deinterleaving (Convolutional Deinterleaver, $I = 12$);
- External decoding (RS Decoder);
- MUX (multiplexer) adaptation;
- MPEG-2 demultiplexing and source decoding.

The receiving STB (Set-Top Box) adopts techniques which are dual to those used in the transmission. Its practical performance depends on hardware construction (it is not standardized like encoder). Details of an example simulation of DVB-T transmitter and receiver can be found in [11], while the simulation itself can be downloaded from [13].

8.2.2 DVB-T2

Because of the inefficient frequency capacity usage in the terrestrial television platform, defined in DVB-T standard [1], a more efficient transmission system was developed to fulfill the market demands and allow launching new services, such as 3D and multi-view video delivery. To maximize spectrum efficiency, the DVB Project developed the second-generation digital terrestrial television standard, the DVB-T2 standard [2]. This new specification includes a newer coding scheme, interleaving and modulation techniques which provide increased capacity and robustness in the terrestrial transmission environment, compared to DVB-T. Transmission can be adapted to the characteristics of the actual channel conditions, thanks to all configurable parameters of the new standard.

Similarly to the DVB-T, the DVB-T2 uses COFDM, but new modulation and coding techniques are introduced. The possibility of using the 256QAM mode allows higher number of bits to be carried per data cell, which increases the spectral efficiency and bitrate. This increase is possible due to the better coding scheme Low-density parity-check + Bose–Chaudhuri–Hocquenghem (LDPC + BCH). The support for the 16K and 32K transmission modes allowed increase of the guard interval length without decreasing the spectral efficiency of the system. It is possible to choose between normal or extended carrier modes. The extended carrier mode gives the possibility to use more carriers per symbol, resulting in increased data capacity. Comparison of available modes in DVB-T and DVB-T2 specifications is shown in Table 8.1 [14]. Bolded values in DVB-T2 are newly added in standard.

The DVB-T2 transmitter, shown in Fig. 8.2, consists of several signal processing blocks. First novelty in the DVB-T2 standard is LDPC code [15] in combination with BCH, used as a protection against interference and noise. LDPC and BCH codes offer excellent performance resulting in a robust signal reception in different signal transmission condition. An important new feature is also bit, cell, time, and frequency interleaver. Additionally, a new technique called rotated constellations [16] resulted in improved robustness against loss of data cells. Similarly to the DVB-T, the DVB-T2 uses COFDM, but new modulation and coding techniques are introduced. 256QAM mode allows higher number of bits to be carried per data cell, which increases the spectral efficiency and bitrate. Higher number of bits per symbol, compared with DVB-T, is possible due to the better protection coding scheme in DVB-T2 (LDPC + BCH). The support for the 16K and 32K transmission modes increases the guard interval length without decreasing the spectral efficiency of the system. It is also possible to choose between normal or extended carrier modes. Extended carrier modes achieve even better spectral efficiency, comparing to normal carrier modes. DVB-T2 standard uses eight scattered pilot pattern modes (depending on Fast Fourier Transform (FFT) size and guard

Table 8.1 Comparison of parameters in DVB-T and DVB-T2 standard

	DVB-T	DVB-T2
FEC	Convolutional coding (1/2, 2/3, 3/4, 5/6, 7/8) + Reed–Solomon	LPDC (1/2, **3/5**, 2/3, 3/4, **4/5**, 5/6) + BCH
Modes	QPSK, 16QAM, 64QAM	QPSK, 16QAM, 64QAM, **256QAM**
Guard interval	1/4, 1/8, 1/16, 1/32	1/4, **19/256**, 1/8, **19/128**, 1/16, 1/32, **1/128**
FFT size	2K, 8K	**1K**, 2K, **4K**, 8K, **16K**, **32K**
Scattered Pilots	8% of total	**1%**, **2%**, **4%**, 8% of total
Continual Pilots	2.6% of total	**0.35%** of total
Spectrum	5, 6, 7, or 8 MHz	**1.7**, 5, 6, 7, 8, **10** MHz

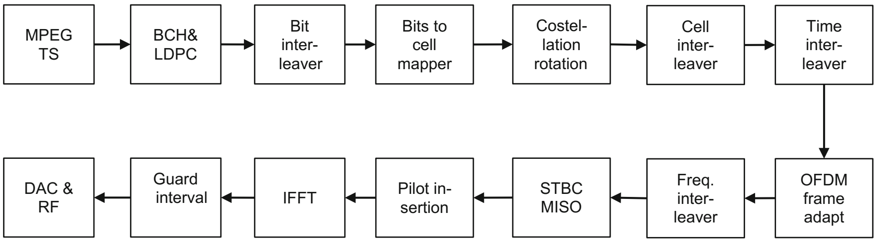

Fig. 8.2 DVB-T2 transmitter block diagram

interval, only some combinations are allowed), in order to maximize the data payload. The DVB-T2 standard also allows the transmission of single or multiple PLPs (Physical Layer Pipes) simultaneously. Each PLP carries one or more logical data streams (DVB-T2 services and signaling data) and can have different physical parameters, like coding rate or constellation. Optionally, the DVB-T2 standard supports Multiple Input Single Output (MISO) systems. Two techniques for Peak to Average Power Ratio (PAPR) reduction can be used: Active Constellation Extension (ACE) for lower order constellations and Tone Reservation method for higher order constellations.

An example of net bitrate calculation application in DVB-T2 network can be downloaded from [13] (written in MATLAB). Net bitrate for specific symbols/frame can be also calculated (if exists). Supported modes are given as follows (but only with allowed combinations of parameters):

- scattered pilots: PP1–PP8;
- bandwidth: 1.7, 5, 6, 7, 8, 10 MHz;
- FFT: 1k, 2k, 4k, 8k, 16k, 16k-ext, 32k, 32k-ext;
- guard interval: 1/4, 19/128, 1/8, 19/256, 1/16, 1/32, 1/128;
- modulation: QPSK, 16QAM, 64QAM, 256QAM;
- L1 post-modulation: BPSK (Binary Phase-shift keying), QPSK, 16QAM, 64QAM;
- FEC: 1/2, 3/5, 2/3, 3/4, 4/5, 5/6;
- PAPR (Peak to Average Power Reduction): no tr-PAPR, tr-PAPR;
- efficiency mode: High-efficiency mode—HEM (no Deletion of Null Packets), normal (no Input Stream Synchronization, no Deletion of Null Packets).

8.2.3 DVB-S/S2

DVB-S standard is defined in ETSI EN 300 421 [3] for digital video broadcasting over satellite. Functional block diagram is shown in Fig. 8.3. Basically, its structure is similar with DVB-T up to Inner encoder: Outer coder is Reed–Solomon (204,188); Outer interleaver is convolutional interleaver; Inner coder can have FEC

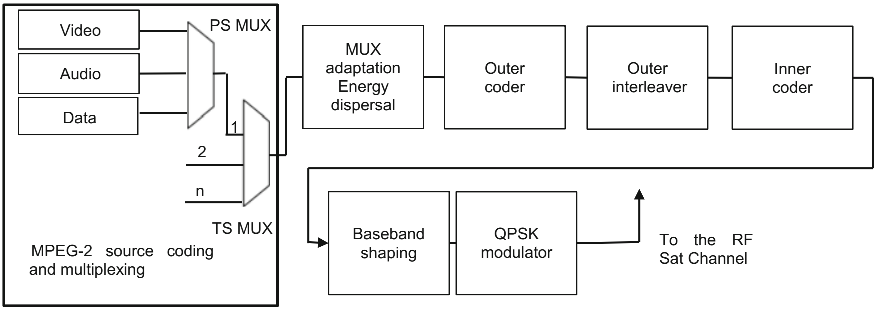

Fig. 8.3 Functional block diagram of DVB-S system

Table 8.2 Comparison of parameters in DVB-S and DVB-S2 standard

	DVB-S	DVB-S2
Input interface	Single transport stream (TS)	Multiple transport stream and generic stream encapsulation
Modes	Constant coding and modulation	Variable coding and modulation and adaptive coding and modulation
FEC	Reed–Solomon (RS) 1/2, 2/3, 3/4, 5/6, 7/8	LDPC + BCH 1/4, 1/3, 2/5, 1/2, 3/5, 2/3, 3/4, 4/5, 5/6, 8/9, 9/10
Modulation	QPSK	QPSK, 8PSK, 16APSK, 32APSK
Roll-off	0.35	0.2, 0.25, 0.35
Pilots	N/A	Pilot symbols

1/2, 2/3, 3/4, 5/6, 7/8. After it, signal is shaped in baseband (square root cosine filter with roll-off factor 0.35) and modulated using QPSK modulation. QPSK modulation, although with lower spectral efficiency, is chosen because satellite services are particularly affected by power limitations in transmission channel.

The second-generation digital satellite television standard, DVB-S2, is explained in detail in ETSI EN 302 307 [4]. Main differences between DVB-S and DVB-S2 are shown in Table 8.2.

In March 2014, an optional extension of DVB-S2 standard was proposed as DVB-S2X [16].

8.2.4 DVB-C/C2

The DVB-C standard, described in detail in ETSI EN 300 429 [5], is used for cable delivery systems A DVB-C transmitter, Fig. 8.4, consists of several signal processing blocks, which are given as follows:

- MPEG-2 source coding and multiplexing;
- MUX adaptation and energy dispersal;

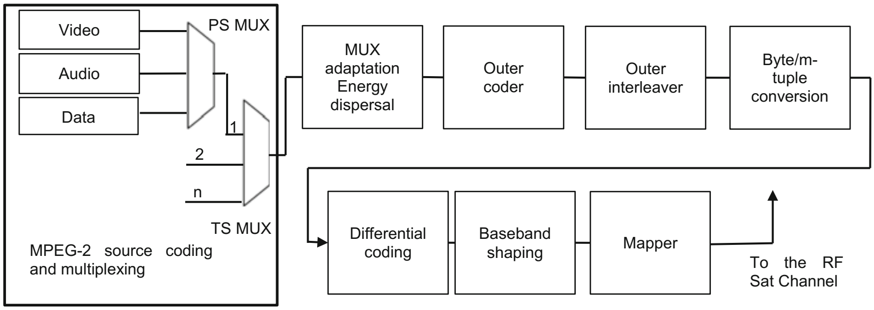

Fig. 8.4 Functional block diagram of DVB-C system

- *Outer encoder*: Reed–Solomon RS (204, 188) code;
- Outer interleaver;
- *Byte/m-tuple conversion*: This unit shall perform a conversion of the bytes generated by the interleaver into QAM symbols (m = 4, 5, 6, 7 or 8).
- *Differential coding*: In order to get a rotation-invariant constellation, this unit shall apply a differential encoding of the two Most Significant Bits (MSBs) of each symbol.
- Baseband shaping
- *QAM Mapper and RF front-end*: Five allowed modulation modes are 16QAM, 32QAM, 64QAM, 128QAM, and 256QAM.

DVB-C STB has techniques which are dual to those used in the transmission. DVB-C2, second-generation digital cable television standard is explained in detail in ETSI EN 302 769 [6]. Main differences between DVB-C and DVB-C2 are shown in Table 8.3.

Table 8.3 Comparison of parameters in DVB-C and DVB-C2 standard

	DVB-C	DVB-C2
Input interface	Single transport stream (TS)	Multiple transport stream and generic stream encapsulation
Modes	Constant coding and modulation	Variable coding and modulation and adaptive coding and modulation
FEC	Reed–Solomon (RS)	LDPC + BCH 1/2, 2/3, 3/4, 4/5, 5/6, 8/9, 9/10
OFDM	–	4K IFFT
Modulation schemes	16QAM to 256QAM	16QAM to 4096QAM
Guard interval	–	1/64 or 1/128
Interleaving	Bit-interleaving	Bit, time, and frequency interleaving
Pilots	–	Scattered and continual pilots

8.2.5 Transport of 3D Video in DVB Systems

Compressed video, audio, and ancillary information can be transported over DVB physical layers after multiplexing and packetization using the formats defined in MPEG-2 Systems specification, as laid out in ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 138181-1 [8] and ETSI technical standard 101 154 [9]. MPEG-2 Systems describes two packetizing formats, one at program stream (PS) level adequate for use in storage or transmission over error-free communication systems, and another one at transport layer level (TS—Transport Stream) with provisions for multiplexing of multiple program streams and better suited for transmission over error-prone channels such as those typical of the DVB systems described in Sects. 8.2.1–8.2.4. In both cases, the source encoded media information is first formatted as elementary video, audio, and other data streams that are packetized and prepended with headers containing timing information to form elementary stream packets, as illustrated in Fig. 8.5. The figure shows packets with variable lengths, reflecting the fact that the media information can be sourced at different data rates. The PES associated with a given program usually share a common time base and are multiplexed, defining a program stream. The packet headers carry important timing information such as presentation time stamps (PTS) and decode time stamps (DTS) which are used by the audio and video decoders to decide in which order the packet contents should be decoded and rendered and also allows synchronization of the audio and video streams at play time.

The program stream packets can be stored or transmitted directly over error-free transmission channels. Alternatively, PES packets belonging to different programs possibly with different time bases can segment into smaller packets, to form TS packets which can be passed on to the lower layers of the DVB system for transmission after addition of FEC bits. The segmentation of PES packets into TS packets is shown in Fig. 8.6 and the multiplexing of TS packets from different programs into a final transport stream is sketched in Fig. 8.7.

The transport layer also fulfills very important functions of synchronization, both end-to-end and between elementary streams, relying on a reference signal provided

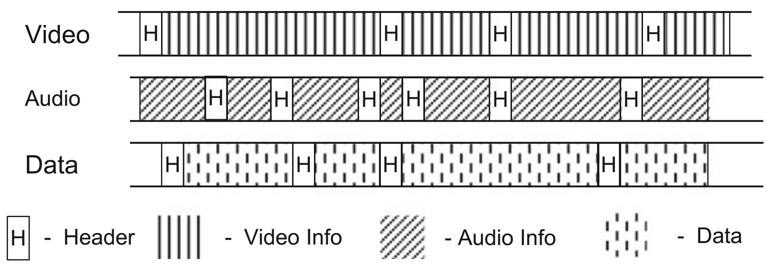

Fig. 8.5 Packetized elementary streams

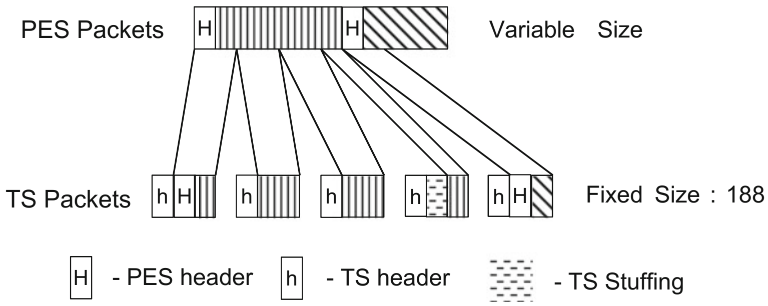

Fig. 8.6 PES packetization into TS packets

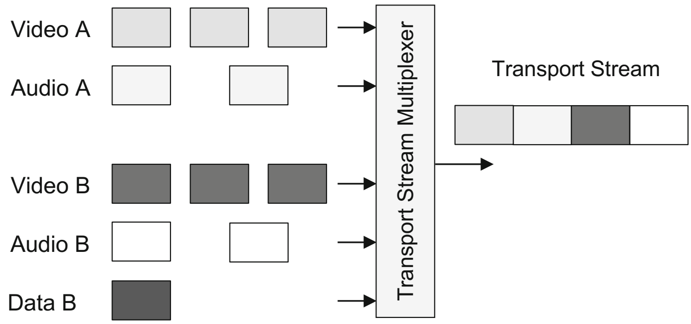

Fig. 8.7 Multiplexing of TS packets

by the system time clock (STC) whose value is periodically inserted in the TS packets.

Originally, MPEG-2 Systems was designed for use in the transmission of 2D video but it has since suffered amendments to support the transmission of 3D video in stereo, multi-view video, and in texture plus depth formats. Encoded 3D stereo video can be transported in several different ways over MPEG-2 TS systems, some choices depending on how the left and right views are encoded. The simplest solution is applicable to video which has been encoded in "simulcast", i.e., both views have been encoded separately. In this case, the two streams can be transported as two elementary streams of a given program stream or two program streams of a transport stream and the decoder receives information about the 3D nature of the content through specific signaling. A slightly more efficient method involves using layered video coding, such as in MPEG-2 with the Multi-view Profile, to encode one of the views as a base view and encode the other with reference to the base view. The two streams can then be multiplexed into a transport stream. These two transport methods are very convenient as they are easily compatible with 2D decoders which just have to ignore one of the elementary video streams and decode the other.

Instead of encoding both views and separate video streams, frame-compatible encoders assemble a composite frame which is the encoded using a regular 2D video encoder. The composite frames can be constructed either by joining the left and right view frames side-by-side, or in a top–bottom arrangement or using some other spatial multiplexing scheme possibly with some previous spatial sub-sampling operations. The encoded video is then delivered using a normal MPEG-2 TS complemented with information to signal the decoder that the stream contains a frame-compatible-encoded stereo video. Several variants of these schemes are documented in ETSI technical standard series 101 547 [17–20].

Both H.264/AVC (Advanced Video Coding) and HEVC (High-Efficiency Video Coding) have extensions for encoding multi-view video, according to which view is independently encoded (base view) while the others are encoded using the frames of the base view as reference for inter-view prediction. Several streams result from the encoding procedure, which are interleaved on a frame basis to produce a multi-view elementary stream that can be delivered using an MPEG-2 transport stream together with signaling information.

Video plus depth emerged as 3D video format with several advantages over stereo video, as it supports compatibility with 2D decoders in a natural way and results in a relatively small additional amount of bits when compared to 2D video. Due to these advantages and the foreseeable wide adoption of the format, the MPEG group standardized a method to encode auxiliary video information such as the depth component of this 3D video format (the standard also supports representing parallax information). The standard is informally known as MPEG-C Part 3, or ISO/IEC 23002-3 and also specifies how to signal the receiver that a stream contains 3D video in video plus depth format. Both the video part as well as the depth component can be transported over MPEG-2 TS systems, together with supplemental information which can specify the encoding standard used to encode the depth information. Further details can be found in [21].

In [22], the authors describe how to deliver layered media (such as MVC and SVC) over DVB-T2 using multiple PLPs. The key is to use the common PLP to deliver one representation of the layered media stream and transmit the second representations of the layered media stream in a data PLP. The combination of layered media codecs with multiple PLPs in DVB-T2 can enable flexible and cost-efficient delivery of high-quality HDTV and 3DTV services. Figure 8.8 shows flexible implementation of 3D services using MVC and 2D services using SVC.

A review of state of the art in 3D video formats, coding methods for different transport options and video formats, IP streaming protocols, and streaming architectures is presented in [23]. The authors also describe asymmetric stereoscopic video streaming, adaptive, and peer-to-peer (P2P) streaming of multi-view video, view-selective streaming, and future directions in broadcast of 3D media over IP and jointly over DVB and IP. In [24], the authors present a complete framework of an end-to-end error resilient transmission of 3D video over DVB-H and provides an analysis of transmission parameters. Figures 8.9 and 8.10 show basic principle of proposed 3D video transmitter and receiver over DVB-H networks.

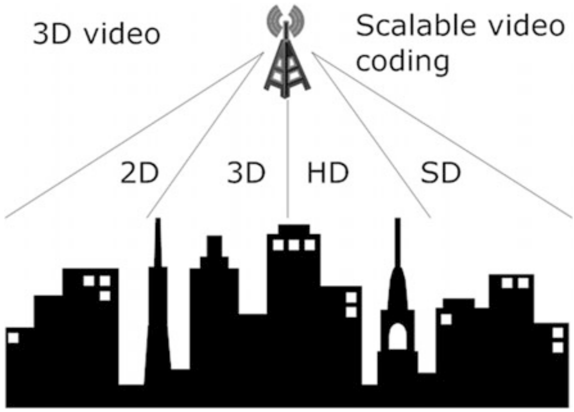

Fig. 8.8 3D services with MVC and 2D services with SVC

The effect of different slice modes and protection methods on the error performance of video + depth-based 3D video broadcast over DVB-H was studied in [25]. In [26], the authors propose a complete framework for end-to-end transmission of stereo video for regular services using a digital video broadcasting-terrestrial version 2 (DVB-T2) system. The proposed system incorporates existing services (such as fixed and mobile) to deliver stereoscopic 3D content in order to maintain backward compatibility without requiring any additional bandwidth. Figure 8.11 shows transmitter side of Hybrid DVB-T2 3DTV and Fig. 8.12 shows receiver side of the proposed system.

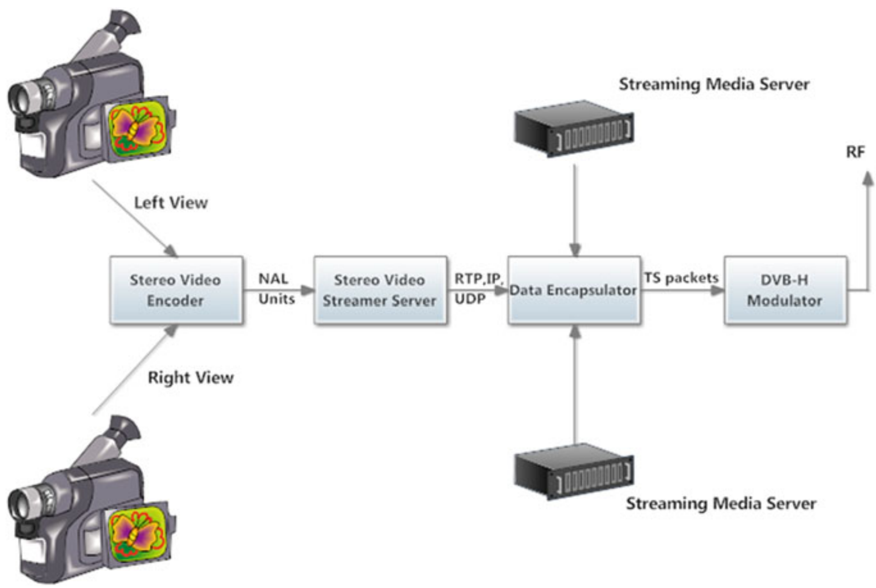

Fig. 8.9 3D video transmitter over DVB-H

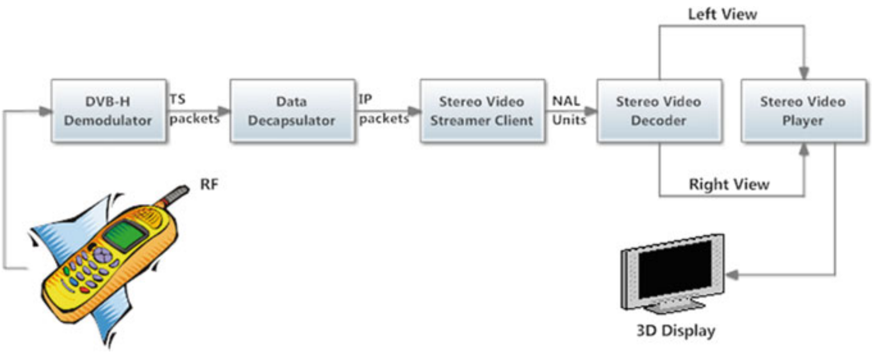

Fig. 8.10 3D video receiver over DVB-H

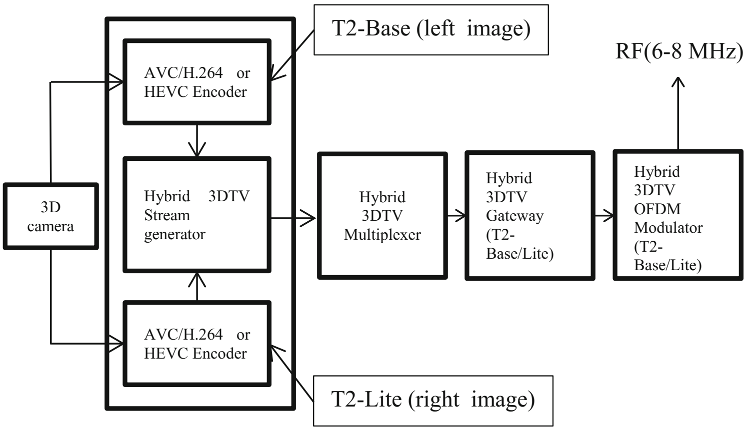

Fig. 8.11 Hybrid DVB-T2 3DTV system overview (transmitter side)

The work described in [27] presents DVB-T2 and T2- Lite/NGH hybrid 3DTV system design and implementation for DVB-T2 terrestrial 3DTV broadcasting services. The proposed system uses the video streams from two different PLP as left and right eye videos to provide high-quality 3D services, while it is backward compatible with DVB-T2/T2-Lite and NGH standard.

8.3 Hybrid Broadcast/Broadband 3DTV

The broadcast transport and existing IP networks do not sufficiently scale up to carry multi-view video with depth maps. Hence, it may be of interest to pool existing DVB broadcast of stereo 3D TV and IP streaming for multi-view 3DTV broadcast applications. A particular system, where frame-compatible stereo

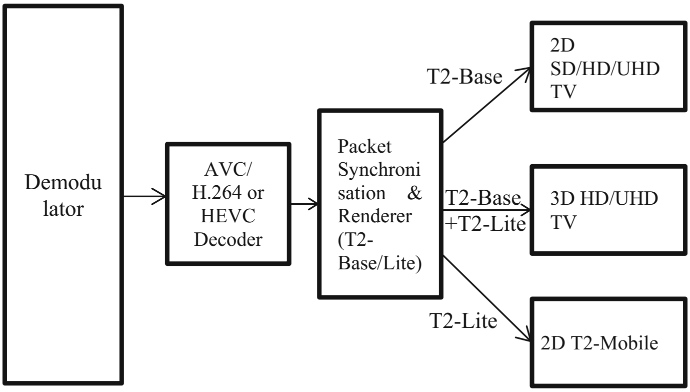

Fig. 8.12 Hybrid DVB-T2 3DTV system overview (receiver side)

broadcast is supplemented by several synchronized auxiliary views and associated depth maps over IP has been proposed in [28]. A hybrid system is constructed in the proposed work, where the DVB network and a P2P overlay network transport multi-view media together. The DVB network has been used to deliver part of the 3D service owing to its robustness and wide availability, as a mechanism to guarantee the minimum 3D Quality of Experience (QoE). P2P transport technology is used for real-time (overlay) multimedia delivery, together with its benefits that include a user preference-aware adaptation mechanism, adaptive redundant chunk scheduling for robustness, incentives to decrease the load on the content server for improved system scalability, and resynchronization capability with the DVB transmission.

A converged broadcast and broadband platform in order to deliver 3D media to both mobile and fixed users with guaranteed minimum quality of experience (QoE) is described in [29]. It offers an ideal business model for operators having both digital video broadcast and Internet Protocol (IP)-based media services. In [30], the authors propose a new synchronization and transport system target decoder (T-STD) model of 3D video distribution based on heterogeneous transmission protocol in a hybrid network environment, where a broadcasting network and broadband network are combined. Proposed technology has been proved to be successfully used as a core element for synchronization and T-STD model in a hybrid 3D broadcasting.

In [31], a general reference model was devised to allow the convergence of 3DTV and 3D Web by defining a general architecture and some extensions of current Smart TV concepts as well as the related standards, like HbbTv (Hybrid Broadcast Broadband TV) [32] and HTML-5 [33]. The authors propose two scenarios: the first scenario is designed for TV sets which include 3D rendering resources and where the apps that are executed on the device have access both to Web and broadcast content. The second scenario is less restrictive and only needs capability to display the remotely rendered 3D content overlaid on top of the broadcasting signal.

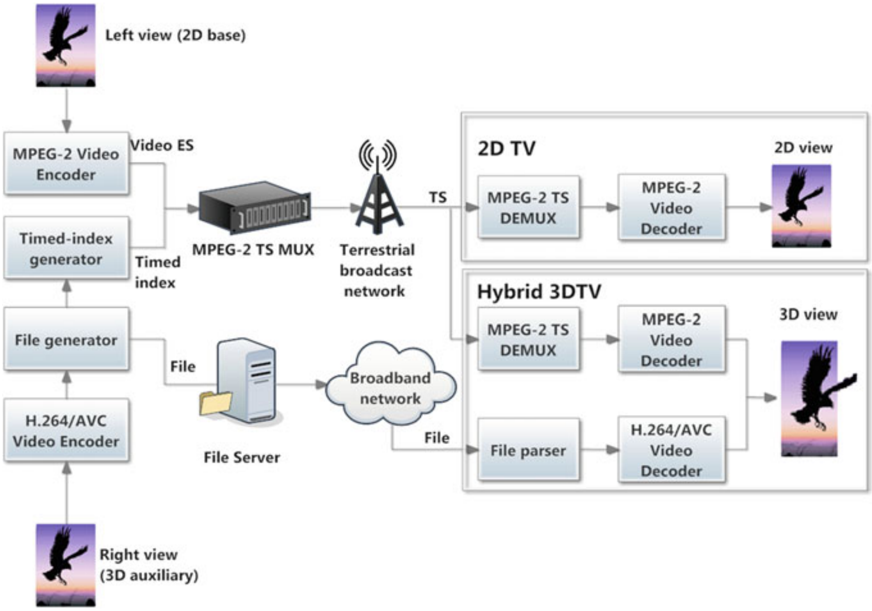

Fig. 8.13 Hybrid broadcast/broadband system overview for hybrid 3DTV service

In [34], the authors propose a hybrid 3DTV broadcasting system, which utilizes both a terrestrial broadcast network and a broadband network. In the proposed system, two elementary streams of left and right views for a stereoscopic video service are transmitted over a terrestrial broadcasting network and a broadband network, respectively. In addition, the proposed system suggests a new mechanism for synchronization between these two elementary streams. Basic diagram of the proposed hybrid 3DTV broadcasting service is shown in Fig. 8.13.

8.4 3D Video Delivery Over IP

In this section, different IP transport technologies that can be used to deliver 3D video content are discussed: HTTP/TCP streaming, adaptive HTTP streaming, RTP/UDP streaming, multi-casting, content-distribution networks (CDN), peer-to-peer streaming (P2P).

8.4.1 HTTP and RTP-Based 3D/Multi-view Streaming

Delivery of media over the Internet Protocol (IP), where the client player can start playback before the entire file has been sent is called streaming. Streaming over IP is a flexible transport system for multi-view video since the transmission bitrate can be configured according to requirements of video format and user equipment.

Streaming systems can be classified as server-client or peer-to-peer (P2P) systems. A server-client streaming system consists of a streaming server and a client that communicate using a set of standard protocols. The client may be a stand-alone player or a Web browser. In the server-client model, typically a different stream is sent to each client. This model is not scalable since server traffic increases linearly with the number of stream requests. Different solutions have been proposed to solve this problem including multi-casting, content-distribution networks (CDN) and P2P streaming. Multi-casting is a one-to-many delivery system, where the server sends each packet only once and the nodes in the network replicate packets only when necessary to reach multiple clients. It can be implemented at the network (IP) or application level. In P2P streaming, clients (peers) forward packets to other peers (as opposed to network nodes) to minimize the load on the source server. Some P2P technologies employ the multi-cast concept when distributing content to multiple recipients, known as peer-casting.

Streaming applications can be classified as one-way video-on-demand, broadcast, or live streaming sessions and two-way real-time communication (RTC) sessions. In a video-on-demand session, the server streams from a pre-encoded and stored file. Live streaming refers to live content delivered in real-time over the Internet, which requires a live camera and a real-time encoder on the server side [35].

A streaming service can be a pull-application (client-driven stream request) or a push-application (server driven packet transmission). In order to avoid the burden of processing individual client status information on the server side, one-way streaming applications are often configured as pull-applications using HTTP/TCP (Hypertext Transfer Protocol/Transmission Control Protocol), where the server is a simple HTTP server and all intelligence reside on the client side. Streaming over HTTP works by breaking a stream into a sequence of small HTTP-based file downloads, each download loading one short *chunk* of the whole stream. Delay critical applications, such as real-time communication sessions, are set up as push applications using Real-Time Streaming Protocol (RTSP). RTSP is an open standard published by the Internet Engineering Task Force (IETF) in 1998. RTSP servers use the Real-time Transport protocol (RTP) over User Datagram Protocol (UDP) for media stream delivery, which supports a range of media formats. Smartphone platforms also include support for RTSP as part of the Third-Generation Partnership Project (3GPP) standard. The main problem with UDP-based streaming is that streams are frequently blocked by firewalls since they are not being sent over HTTP (port 80). In order to alleviate this problem, protocols have been extended to allow for stream encapsulation within HTTP requests, called tunneling, as a fallback solution. Streaming protocols also have secure variants that use encryption to protect the stream.

Since the Internet is a best-effort channel, packets may be delayed or dropped by the routers and the effective end-to-end bitrates fluctuate in time. Adaptive streaming aims to adapt the video encoding rate according to estimated available end-to-end network rate. One possible solution is stream switching, where the server encodes video at multiple preselected bitrates and the client requests switching to the stream encoded at the rate that is closest to its network access rate. An alternative solution is based on scalable video coding, where one or more enhancement layers of video may be dropped to reduce the bitrate as needed.

HTTP streaming solutions include support for adaptive streaming (bitrate switching) to allow clients dynamically switch between streams of varying quality and chunk size during playback, in order to adapt to changing network conditions and available central processing unit (CPU) resources. HTTP streaming allows chunks to be cached within Internet service providers (ISP) or corporations, which would reduce the bandwidth required to deliver HTTP streams. Different vendors have implemented different HTTP-based streaming solutions, such as HTTP Live Streaming (HLS) by Apple, Smooth Streaming by Microsoft, HTTP Dynamic Streaming (HDS) by Adobe, which use similar mechanisms but are incompatible. MPEG-DASH (Dynamic Adaptive Streaming over HTTP) is an international standard, published in April 2012, for adaptive bitrate HTTP-based streaming that is audio/video codec agnostic in order to address this fragmentation problem.

The work done in [36] addresses multi-view video streaming over HTTP for emerging multi-view displays. In this research, the effect of various adaptations of decision strategies is evaluated and, as a result, a new quality-aware adaptation method is designed. The proposed method is benefiting from layer-based video coding in such a way that high Quality of Experience (QoE) is maintained in a cost-effective manner. A different approach was followed in [37], where the authors proposed a method of efficient 3D adaptive streaming service based on the DASH, which covers service-compatible stereoscopic content in a single segment sequence. The 3D adaptive HTTP streaming method introduced in this chapter is able to provide seamless streaming service for various kinds of stereoscopic contents.

In [38], the authors propose a dynamic adaptive rate control system and its associated rate-distortion model for multi-view 3D video transmission, which improves the user's quality of experience in the face of varying network bandwidth. Rate control system has been built on top of two state-of-the-art key technologies: HEVC and MPEG-DASH.

In [39], the authors investigated the possibility of greatly improving the average quality and also attenuating the quality variations for a better user experience, by leveraging multipath communication requiring no extra bandwidth. Algorithm exploits concurrently multiple paths using a Content Delivery Network (CDN) compliant distributed streaming infrastructure and standard HTTP range requests.

8.4.2 3D Video Distribution Over P2P Networks

Due to the large bandwidth consumptions to deliver 3D video content (e.g., stereoscopic and multi-view video), it is a challenge to achieve service scalability and satisfy increasing numbers of recipients [40]. Therefore, the throughput provided by the traditional server-client model may not fulfill the requirements of high-quality 3D video delivery. On the contrary, the peer-to-peer model enables video streaming clients to share content with others, which enhances the utilization of spare bandwidth of the users, such that the overall throughput can be improved and the burden of service providers can be mitigated.

The P2P protocols were originally designed to distribute larger files over the Internet via an overlay solution [41]. In the early stage, the P2P streaming protocols choose a tree-based topology [42] and hierarchical connections among peers. As this approach is difficult to cope with network dynamics, e.g., peer churns, and cannot utilize the outbound bandwidth of the peers at the leaves of the tree, mesh-based topology has become the widely adopted structure in practice [43]. Usually, a tracker server [44] is deployed to provide a list of relevant peers for a content object, which allows any peer to download the missing content chunks from other peers. Later, due to the creation of the distributed hash table (DHT) [45] for discovering peers without centralized storage and inquire, the tracker becomes redundant and is deprecated in some protocols.

A popular P2P file sharing protocol—BitTorrent [46] operates according to a divide-and-conquer principle and divides content into equal size chunks to facilitate delivery. Different successors are proposed for optimizing the video streaming quality via modifying chunk scheduling policies [47, 48] and interleaving chunks [49]. In the original BitTorrent protocol, two scheduling policies of downloading chunks are ***rarest-first*** and ***tit-for-tat***. The rarest-first policy downloads the chunk that is the least distributed within the swarm, which is inferred as the least distributed chunk are most likely to be required by neighbors. The tit-for-tat policy determines the qualities (contributions) of the neighboring peers and rejects the requests from the low-rank peers, which prevents leeches from downloading much more than their uploading [50].

For video streaming applications, the rarest-first policy may encounter some serious issues. For instance, when a peer has started playback of the video, the rarest-first policy may prefer to download the chunks that are far away from the playback time and consequently stall the playback. To solve this problem, a novel method [47] that combines the sequential downloading and rarest-first is proposed. The peer has a priority to retrieve the chunks sequentially within a ***ready-to-play*** buffer that guarantees the least playback time without stalling. If more bandwidth is available, the peer downloads content chunks with the rarest-first policy to improve the content richness of the swarm.

Streaming the 3D video can be either supported by a push-based P2P [45, 51] or a pull-based P2P [52]. In push-based P2P, peers are arranged in a tree structure and the video chunks are pushed from a parent node to its child nodes. It is an efficient solution if connections between peers are stable. However, tree regeneration is required when a parent node leaves. Thus, some solutions make use of redundant parents to prevent tree loss (disconnection) caused by leaving nodes. In pull-based P2P, peers can be arranged in a mesh-based structure and peers pull (request) the needed content from neighbors. In order to achieve so, each peer maintains a bitmap of chunks that are currently available by it or its neighbors and downloads the required video chunks from the neighbor with the highest bandwidth.

In 2D video streaming, video quality adaptation is the main design concern. For example, in delivering of scalable videos (e.g., H.264/SVC (Scalable Video Coding) [53]), the base layer is indispensable and the enhancement layer can be abandoned according to the adaptation logic when bandwidth is insufficient.

In contrast, 3D video content delivery is more complicated than 2D video delivery because of the complexity of the video layer dependency. The scheduling and adaptation logics in 3D video delivery require considering more factors besides the quality of a view, e.g., current viewpoint, viewpoint switching, and the possibility of missing viewing synthesis. Taking multi-view streaming, for example, the viewpoint of the viewer (peer) determines the importance of different views, the farther the less important it is. The adaptation logic should guarantee the chunks of the view in the current field of interest. The loss of the enhancement layer in this view may affect the QoE (Quality of Experience) much greater than the loss of the layers that is outside the field of interest. Furthermore, as the viewer can switch the viewpoint during playback, the views outside the field of interest also need to be pre-fetched. Otherwise, playback stalling or image distortion caused by missing chunks will largely affect QoE. The prediction of viewer's behavior becomes critical, which allows viewers to achieve adequate playback quality by spending less bandwidth. Finally, the coding techniques affect the video delivery design and performance significantly. In the early stage, the different views are coding independently (simulcasting) and symmetrically [54]. The delivery and adaptation algorithms proposed for 2D video can be utilized here without lots of modifications. With the development of asymmetric coding [54] for different views, peers need to distinguish the coding scheme for each view and perform optimal adaptation. Moreover, if the coding scheme supports to recover the lost view from the depth map and the reference views [55], peers can proactively abandon some views by determining the importance of each one, which enhances QoE with a limited bandwidth.

8.5 3D Video Distribution in ICN

The Information-Centric Networking (ICN) [56] concept is an important approach to the future Internet research activities. It is proposed to shift the Internet infrastructure away from the host-centric paradigm to a network architecture based on named data objects. In ICN, content is named and routed directly in the network and it can be retrieved through in-network caching and adaptive routing, which is a promising solution to improve efficiency, scalability, and robustness of the network.

An obvious benefit of ICN is that it enables caching of content in intermediate nodes to reduce congestion and improve delivery speed. For instance, if a large group of users is accessing the same data source (e.g., website or video) on the Internet simultaneously, the bottleneck link can be overloaded in the current TCP/IP architecture. By caching content in network [56] or at edge [57], the requests from multiple clients can retrieve content from intermediate nodes along the forwarding path, which can shorten the delay and offload the burden of bottlenecks. This applies to scenarios like large content delivery such as 3D video delivery. Currently, there are basically no studies on 3D video streaming on ICN. In this section, the challenges and possible solutions are discussed.

For the ICN implementations (e.g., Content-Centric Networking [58]) that support adaptive forwarding, the traffic control is still undergoing work which may affect the deployment of 3D video streaming. As interest packets (requests) of a video can be forwarded to multiple content producers concurrently, the consumer (i.e., a node is requesting the content) can hardly predict the congestion of the network based on its local information (e.g., timeout or duplicate acknowledgment) as used in TCP/IP. To this end, recent works [59, 60] prefer to utilize the explicit congestion notification signals (e.g., ECN—Explicit Congestion Notification or NACK—Negative Acknowledgement) to notify the congestion explicitly and another type of promising solutions is to use path-labeling [61] to detect congestion for each path.

Although it is similar to P2P, that ICN enables a consumer to download different content chunks from multiple producers simultaneously, there are some fundamental differences between ICN and P2P. For instance, P2P allows peers to communicate with other peers directly and routers are not responsible for discovering peers or balancing the load. By removing the host information in ICN, consumers are not designed to communicate to producers directly. Instead, routers are responsible for balancing the traffic to peers or servers to maximize the throughput of peers. Thus, the peer selection in traditional P2P is mapped to content publishing and adaptive forwarding. For the time being, the design of forwarding strategy becomes critical to maximize the downloading performance of peers.

As ICN introduces in-network caching, how and where to cache 3D video chunks becomes a novel topic that must be explored. In layered video streaming, the base layer is more important than the enhancement layers from the decoding point of view, thus differentiated service can be applied to cache chunks with different priority if the cache is limited.

Another possibility is to apply differentiated service to selective forwarding. When the network cannot provide sufficient bandwidth to deliver the content, some ICN solution enables routers to drop interest packets proactively. If the priority of content is considered, routers are able to discard the interest packets that do not have large effects on QoE.

8.6 3D Stereo and Multi-view Video in Wireless Networks

The support of 3D stereo and multi-view video over wireless networks is still an emerging research topic and very few research works have investigated the related challenges. First categories of works have focused on assessing the quality of experience obtained when transmitting 3D contents over wireless networks. In [62], the authors have conducted several experiments with subjective QoE evaluation of a 3D stereoscopic video files transported through TCP connections and over 802.11n WiFi access network. Several metrics were ranked, such as video continuity, 3D visual quality or video/voice synchronization while varying some network QoS (Quality of Service) metrics, such as WiFi channels and available bandwidth. The study showed the strong correlation between QoS degradation, in particular, bandwidth capacity, and the perceived QoE.

In [63], the authors address the support of 3D video streaming over LTE wireless networks and propose a framework that aims to dynamically control the system parameters in order to optimize the perceived 3D QoE. The key idea relies on a proxy that enables, through a Machine learning approach, real-time control and monitoring of the parameters impacting the 3D QoE. In [64], the authors extend this idea of context-aware 3D rendering and streaming over heterogeneous wireless networks to the cloud paradigm. The real-time control and monitoring of the network resources are facilitated thanks to the SDN (Software-defined networking) controller. The dynamic adjustment of the 3D coding accordingly with the available cloud resources and the 3D video visual quality is then enabled thanks to a second module.

Actually, the critical issue associated with 3D/multi-view video streaming in nowadays wireless networks is related to the limited and variable wireless bandwidth capacity and their inability to support the huge data rates associated with 3D contents. Moreover, packet losses and disconnections are quite frequent, either due to user's mobility or are as a consequence of interferences. In addition, expected use case scenarios are those were wireless resources are shared among several users. Dedicated or reserved wireless resources to support 3D streaming is a technically possible approach, but it is not economically viable for large public. To cope with the wireless bandwidth limitations, a first direction is to dynamically adjust the encoding and transmission bitrate depending on the available wireless bandwidth. In [65], authors proposed an encoding and transmission technique over WiFi networks of 3D objects of an MPEG-4 video depending on their perceived importance. The objects are segmented thanks to the depth information of 3D content.

Another recent proposal is the collaborative rendering of 3D contents. This approach can be used for instance to stream 3D video games contents in cloud-based services. In such services, the conventional approach is to stream the video games 3D contents to thin clients (tablets, smartphones, etc.). While this solution offers several advantages, such as the independence between the games and the terminals, it requires the satisfaction of stringent QoS constraints, including short delays and large bandwidth. To accommodate the users with the limited bandwidth of wireless networks, the authors of [66] propose to offload part of the GPU computation. Two strategies are analyzed. The first one assumes that the client renders each frame with reduced details. Therefore, the cloud has to stream only the per-frame differences between the high detail version and the client frame. In the second strategy, the client renders with high details, a subset of frames while the server streams only the remaining ones.

In [67], a rate adaptation method that uses the packet buffer size as an indicator of network load was proposed. During congestion, transmitter proactively drops packets belonging to layers with less impact. On the receiver side, data from the dropped layers are approximated using error concealment strategy, based on synthesizing the missing texture and depth frames when possible. A different strategy was followed in [68], where the authors proposed a multi-view video-aware transmission over MIMO wireless systems. The basic idea is to exploit the channel diversity of multiple antennas and the source coding characteristics so to achieve unequal error protection against channel errors. To achieve this goal, authors developed a nonlinear mixed integer programming framework to perform antenna selection and power allocation and proposed low-complexity algorithms to assign these resources.

A multiple description coding approach for multipath transmission of free-viewpoint video, with joint interview and temporal description recovery capability was investigated in [69]. Even frames of the left view and the odd frames of the right view are separately encoded and transmitted as one description on one path. Remaining frames in the two views are encoded and transmitted over a second path. If the receiver receives only one description due to burst loss in the other path, the missing frames in the other description are partially reconstructed using newly proposed frame recovery procedure from the same authors.

Another strategy was followed in [70], in mobile delivery of 3D content, where the authors propose to leverage depth-image-based rendering (DIBR) in multi-view 3D, which allows each mobile client to synthesize the desired view from nearby left and right views, in order to effectively reduce the bandwidth consumption in wireless networks. The authors developed the Multi-View Group Management Protocol (MVGMP) for multi-view 3D multicast. When a user joins the video multicast group, it can receive the most suitable right and left views, so that the view failure probability is guaranteed to stay below a threshold. When a user leaves the video multicast group, MVGMP carefully selects and withdraws a set of delivered views to reduce the network load, so that the video failure probability for other users will not exceed the threshold. Bandwidth consumption can be effectively reduced since it is not necessary to deliver all the views subscribed by the clients.

Another scheme called the Temporal Synchronization Scheme (TSS) for live 3D video streaming over 802.11 wireless networks was developed in [71]. The TSS scheme delivers video frames for the left and right views in the same frame order with the same transmission priority and compensates for frame damage and loss during the decoding phase. A new metric called the Stereoscopic Temporal Variation Index (STVI) is also proposed to measure the degree of temporal asynchrony in 3D video.

While the main idea in previous approaches is to act on the 3D contents in order to fit with the wireless network capacity, another promising direction is to the increase the wireless bandwidth in the next generations of wireless access technologies. Hence, regarding WLAN, new amendments to the IEEE 802.11 standard are already under development. Precisely, the 802.11ax is the next WiFi access technology that is actually under development with the targeted maximum bitrate around 10 Gbit/s. In complementary, 802.11ay is expected to replace cabled Ethernet LAN. This technology is predicted to provide up to 176 Gbit/s.

Regarding cellular network technologies, the next fifth-generation wireless systems (5G) is also under active development. The standard is expected at the horizon of 2020. Several promising features are announced such as tens of Mb/s per connection for thousands of users and very short latencies.

8.7 Conclusion

This chapter presented different network systems that can be used to deliver different 3D video content. A short description was provided about the DVB broadcast systems (DVB-T/T2, DVB-S/S2 and DVB-C/C2), IP transport technologies (multi-casting, content-distribution networks, peer-to-peer streaming, Information-Centric Networking concept, HTTP/TCP streaming, adaptive HTTP streaming, RTP/UDP streaming), hybrid transport technologies (a combination of broadcast and broadband), and 3D video over wireless networks that can be used to efficiently transmit 3D media content.

References

1. ETSI EN 300 744: Digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television. V.1.6.1, Sept 2008
2. ETSI EN 302 755: Digital video broadcasting (DVB); frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2). V1.1.1, Sept 2009
3. ETSI EN 300 421: Digital video broadcasting (DVB); framing structure, channel coding and modulation for 11/12 GHz satellite services. V.1.1.2, Aug 1997
4. ETSI EN 302 307: Digital video broadcasting (DVB); second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications (DVB-S2). V.1.2.1, Aug 2009
5. EN 300 429: Digital video broadcasting (DVB); framing structure, channel coding and modulation for cable systems. V.1.2.1, Apr 1998
6. ETSI EN 302 769: Digital video broadcasting (DVB); frame structure channel coding and modulation for a second generation digital transmission system for cable systems (DVB-C2). V.1.1.1, Oct 2010
7. ETSI TS 101 547-2: DVB plano-stereoscopic 3DTV; Part 2: frame compatible plano-stereoscopic 3DTV. V1.2.1, Nov 2012
8. ISO/IEC 13818-1: Information technology—generic coding of moving pictures and associated audio—Part 1: Systems, July 2015
9. ETSI TS 101 154 (V1.8.1): Digital video broadcasting (DVB); specification for the use of video and audio coding in broadcasting applications based on the MPEG-2 transport stream, July 2007
10. Stott, J.H.: The how and why of COFDM. EBU Techn. Rev. **278** (1998)
11. Dumic, E., Sisul, G., Grgic, S.: Evaluation of transmission channel models based on simulations and measurements in real channels. Frequenz **66**(1–2), 41–54 (2012)

12. Reimers, U: Digital Video Broadcasting. Springer, Berlin (2001)
13. http://beam.to/datasets
14. DVB fact sheet: DVB-T2—2nd generation terrestrial broadcasting. DVB Project Office, Apr 2012
15. Richardson, T., Urbanke, R.: The renaissance of Gallager's low-density parity-check codes. IEEE Commun. Mag. **41**, 126–131 (2003)
16. ETSI EN 302 307-2: Digital video broadcasting (DVB); second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications: Part 2: DVB-S2 extensions (DVB-S2X). V1.1.1, Mar 2014
17. ETSI TS 101 547-1: DVB plano-stereoscopic 3DTV; Part 1: Overview of the multipart, V1.2.1, Dec 2015
18. ETSI TS 101 547-2: DVB plano-stereoscopic 3DTV; Part 2: frame compatible plano-stereoscopic 3DTV, V1.2.1, Nov 2012
19. ETSI TS 101 547-3: DVB plano-stereoscopic 3DTV; Part 3: HDTV service compatible plano-stereoscopic 3DTV, V1.1.1, Nov 2012
20. ETSI TS 101 547-4: DVB plano-stereoscopic 3DTV; Part 4: service frame compatible planostereoscopic 3DTV for HEVC coded services. V1.1.1, June 2016
21. ISO/IEC 23002-3: Information technology—MPEG video technologies—Part 3: Representation of auxiliary video and supplemental information (2007)
22. Hellge, C., Wiegand, T., Torre, E.G., Gomez-Barquero, D., Schierl, T.: Efficient HDTV and 3DTV services over DVB-T2 using multiple PLPs with layered media. IEEE Commun. Mag. **51**(10), 76–82 (2013). https://doi.org/10.1109/MCOM.2013.6619569
23. Gurler, C.G., Gorkemli, B., Saygili, G., Tekalp, A.M.: Flexible transport of 3-D video over networks. Proc. IEEE **99**(4), 694–707 (2011). https://doi.org/10.1109/JPROC.2010.2100010
24. Bugdayci, D., Akar, G.B., Gotchev, A.: Optimized transmission of 3D video over DVB-H channel. In: IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, pp. 20–24 (2012). https://doi.org/10.1109/ccnc.2012.6181056
25. Buğdaya, D., Bid, M.O., Aksay, A., Demirtaş, M., Akar, G.B.: Video + depth based 3D video broadcast over DVB-H. In: 2010 IEEE 18th Signal Processing and Communications Applications Conference, Diyarbakir, 2010, pp. 902–905. https://doi.org/10.1109/siu.2010.5652310
26. Hossen, M.S., Kim, S.H., Kim, K.D.: Stereoscopic video transmission over DVB-T2 system using future extension frame. IEEE Trans. Broadcast. **62**(4), 817–825 (2016). https://doi.org/10.1109/TBC.2016.2590831
27. Kim, S.-H., Lee, J., Jeong, S., Choi, J., Kim, J.: Development of fixed and mobile hybrid 3DTV for next generation terrestrial DTV. In: 2013 3DTV Vision Beyond Depth (3DTV-CON), Aberdeen, pp. 1–2 (2013)
28. Ekmekcioglu, E., Gurler, G., Kondoz, A., Tekalp, A.M.: Adaptive multi-view video delivery using hybrid networking. IEEE Trans. Circ. Syst. Video Technol. **27**(6), 1313–1325 (2017)
29. Lykourgiotis, A., et al.: Hybrid broadcast and broadband networks convergence for immersive TV applications. IEEE Wirel. Commun. **21**(3), 62–69 (2014). https://doi.org/10.1109/MWC.2014.6845050
30. Yun, K., Cheong, W., Kim, K.: A synchronization and T-STD model for 3D video distribution and consumption over hybrid network. IEICE Trans. Inf. Syst. **98**(10), 1884–1887 (2015), Oct 2015. https://doi.org/10.1587/transinf.2014edl8242
31. Olaizola, I.G., Pérez, J., Zorrilla, M., Martín, Á., Laka, M.: Reference model for hybrid broadcast web 3D TV. In: Proceedings of the 18th International Conference on 3D Web Technology (Web3D '13). ACM, New York, NY, USA, pp. 177–180, June 2013. https://doi.org/10.1145/2466533.2466560
32. ETSI TS 102 796 v1.4.1: Hybrid broadcast broadband TV, Aug 2016
33. https://www.w3.org/TR/2016/REC-html51–20161101/. Accessed 30 Sept 2017

34. Lee, J., Yun, K., Kim, K.: A 3DTV broadcasting scheme for high-quality stereoscopic content over a hybrid network. IEEE Trans. Broadcast. **59**(2), 281–289 (2013). https://doi.org/10.1109/TBC.2013.2256678
35. Tekalp, A.M.: Digital Video Processing. Prentice Hall (2015)
36. Ozcinar, C., Ekmekcioglu, E., Kondoz, A.: Quality-aware adaptive delivery of multi-view video. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1–6, Mar 2016
37. Park, G., Lee, J., Lee, G., Kim, K.: Efficient 3D adaptive HTTP streaming scheme over internet TV. In: IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB), pp. 1–6, June 2012
38. Su, T., Javadtalab, A., Yassine, A., Shirmohammadi, S.: A DASH-based 3D multi-view video rate control system. In: 2014 8th International Conference on Signal Processing and Communication Systems (ICSPCS), pp. 1–6, Dec 2014
39. Gouache, S., Bichot, G., Bsila, A., Howson, C.: Distributed & adaptive HTTP streaming. In: IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6, July 2011
40. Gürler, C.G., Tekalp, M.: Peer-to-peer system design for adaptive 3D video streaming. IEEE Commun. Mag. **51**(5), 108–114 (2013)
41. Lua, E.K., Crowcroft, J., Pias, M., Sharma, R., Lim, S.: A survey and comparison of peer-to-peer overlay network schemes. IEEE Commun. Surv. Tutor. **7**(2), 72–93 (2005)
42. Hudzia, B., Kechadi, M.-T., Ottewill, A.: Treep: a tree based p2p network architecture. In: Cluster Computing, 2005. IEEE International, pp. 1–15 (2005)
43. Magharei, N., Rejaie, R., Guo, Y.: Mesh or multiple-tree: a comparative study of live p2p streaming approaches. In: INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE, pp. 1424–1432 (2007)
44. Sen, S., Spatscheck, O., Wang, D.: Accurate, scalable in-network identification of p2p traffic using application signatures. In: Proceedings of the 13th International Conference on World Wide Web, pp. 512–521 (2004)
45. Galuba, W., Girdzijauskas, S.: Distributed hash table. In: Encyclopedia of Database Systems. Springer, Berlin, pp. 903–904 (2009)
46. B. Inc; BitTorrent. Available: http://www.bittorrent.com/. Accessed 11 June 2017
47. Vlavianos, A., Iliofotou, M., Faloutsos, M.: BiToS: enhancing BitTorrent for supporting streaming applications. In: INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings, pp. 1–6 (2006)
48. Paris, J.-F., Shah, P.: Peer-to-peer multimedia streaming using BitTorrent. In: Performance, Computing, and Communications Conference, 2007. IPCCC 2007. IEEE International, pp. 340–347 (2007)
49. Liu, Z., et al.: H. 264/MVC interleaving for real-time multiview video streaming. J. Real-Time Image Process **10**(3), 501–511 (2015)
50. Wang, P., Wu, L., Aslam, B., Zou, C.C.: A systematic study on peer-to-peer botnets. In: Proceedings of 18th International Conference on Computer Communications and Networks, 2009. ICCCN 2009, pp. 1–8 (2009)
51. Bracciale, L., Piccolo, F.L., Luzzi, D., Salsano, S., Bianchi, G., Blefari-Melazzi, N.: A push-based scheduling algorithm for large scale P2P live streaming. In: Telecommunication Networking Workshop on QoS in Multiservice IP Networks, 2008. IT-NEWS 2008. 4th International, pp. 1–7 (2008)
52. Zhang, M., Zhang, Q., Sun, L., Yang, S.: Understanding the power of pull-based streaming protocol: can we do better? IEEE J. Sel. Areas Commun. **25**(9) (2007)
53. SVC extension of H.264/AVC—Fraunhofer Heinrich Hertz Institute. Available: https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/image-video-coding/research-topics/svc-extension-of-h264avc.html. Accessed 11 June 2017
54. Gurler, C.G., Savas, S.S., Tekalp, A.M.: Quality of experience aware adaptation strategies for multi-view video over p2p networks. In: 19th IEEE International Conference on Image Processing (ICIP), pp. 2289–2292 (2012)

55. Fehn, C.: Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. In: Electronic Imaging 2004, pp. 93–104 (2004)
56. Ahlgren A, Dannewitz C, Imbrenda C, Kutscher D, Ohlman B, "A survey of information-centric networking", *IEEE Commun. Mag.*, vol. 50, no. 7, 2012
57. Fayazbakhsh, S.K., et al.: Less pain, most of the gain: incrementally deployable icn. ACM SIGCOMM Comput. Commun. Rev. **43**, 147–158 (2013)
58. Jacobson, V., Smetters, D.K., Thornton, J.D., Plass, M.F., Briggs, N.H, Braynard, R.L.: Networking named content. In: Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies pp. 1–12 (2009)
59. Zhang, F., Zhang, Y., Reznik, A., Liu, H., Qian, C., Xu, C.: A transport protocol for content-centric networking with explicit congestion control. In: 2014 23rd International Conference on Computer Communication and Networks (ICCCN), pp. 1–8 (2014)
60. Zhou, J., Wu, Q., Li, Z., Kaafar, M.A., Xie, G.: A proactive transport mechanism with explicit congestion notification for NDN. In: IEEE International Conference on Communications (ICC), pp. 5242–5247 (2015)
61. Carofiglio, G., Gallo, M., Muscariello, L., Papalini, M., Wang, S.: Optimal multipath congestion control and request forwarding in information-centric networks. In: 21st IEEE International Conference on Network Protocols (ICNP), pp. 1–10 (2013)
62. Kulik, I., Trinh, T.: Investigation of quality of experience for 3D video in wireless network environment. In: 18th European Conference on Information and Communications Technologies (EUNICE), Aug 2012, Budapest, Hungary. Lecture Notes in Computer Science, LNCS, vol. 7479, pp. 340–349, Information and Communication Technologies. Springer, Berlin (2012)
63. Politis, I., Lykourgiotis, A., Dagiuklas, T.: A framework for QoE-aware 3D video streaming optimisation over wireless networks. Mobile Inf. Syst. **2016**, Article ID 4913216, 18 p (2016
64. Ho, D., Kim, H., Kim, W., Park, Y., Chang, K., Lee, H., Song, H.: Mobile cloud-based interactive 3D rendering and streaming system over heterogeneous wireless networks. IEEE Trans. Cir. Sys. Video Technol. **27**(1), 95–109 (2017)
65. Nasir, S., Hewage, C.T.E.R., Mrak, M., Worrall, S., Kondoz, A.M.: Depth based object prioritisation for 3D video communication over Wireless LAN. In: 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, pp. 4269–4272 (2009)
66. Cuervo, E., Wolman, A., Cox, L.P., Lebeck, K., Razeen, A., Saroiu, S., Musuvathi, M.: Kahawai: high-quality mobile gaming using GPU offload. In: Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys '15). ACM, New York, NY, USA, pp. 121–135 (2015)
67. Oztas, B., Pourazad, M.T., Nasiopoulos, P., Leung, V.C.M.: Adaptive 3D-HEVC video streaming over congested networks through layer prioritization. In: 24th International Conference on Telecommunications (ICT), Limassol, pp. 1–5 (2017). https://doi.org/10.1109/ict.2017.7998258
68. Chen, Z., Zhang, X., Xu, Y., Xiong, J., Zhu, Y., Wang, X.: MuVi: multi-view video aware transmission over MIMO wireless systems. IEEE Trans. Multim **99**, 1–1. https://doi.org/10.1109/tmm.2017.2713414
69. Liu, Z., Cheung, G., Chakareski, J., Ji, Y.: Multiple description coding and recovery of free viewpoint video for wireless multi-path streaming. IEEE J. Sel. Top. Signal Process. **9**(1), 151–164 (2015). https://doi.org/10.1109/JSTSP.2014.2330332
70. Lin, C.H., Yang, D.N., Lee, J.T., Liao, W.: Efficient error-resilient multicasting for multi-view 3D videos in wireless network. In: IEEE Global Communications Conference (GLOBECOM), Washington, DC, pp. 1–7 (2016). https://doi.org/10.1109/glocom.2016.7841779
71. Yoon, Y., Kim, M., Lee, B., Go, K.: Temporal synchronization scheme in live 3D video streaming over IEEE 802.11 wireless networks. In: Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014, Sydney, NSW, pp. 1–7 (2014). https://doi.org/10.1109/wowmom.2014.6918944