

Report ITU-R BT.2556-0

(09/2025)

BT Series: Broadcasting service (television)

Experiences of ultra-high definition television and high dynamic range television

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from <https://www.itu.int/ITU-R/go/patents/en> where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Reports

(Also available online at <https://www.itu.int/publ/R-REP/en>)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	Broadcasting service (television)
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
TF	Time signals and frequency standards emissions

Note: This ITU-R Report was approved in English by the Study Group under the procedure detailed in Resolution ITU-R 1.

Electronic Publication
Geneva, 2025

© ITU 2025

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without written permission of ITU.

REPORT ITU-R BT.2556-0

Experiences of ultra-high definition television and high dynamic range television

(2025)

TABLE OF CONTENTS

Page

1	London Olympics SUPER Hi-VISION public viewing operations	4
2	Implementation of UHD TV broadcasting	5
3	Initiatives for 8K/120 Hz broadcasting	5
4	Practice of 8K UHD programmes broadcasting on large screens in China.....	5
5	Deployment of NextGen TV (ATSC 3.0) in the United States of America.....	5
5.1	NextGen TV High Dynamic Range with Standard Dynamic Range compatibility using single-layer High Dynamic Range (SL-HDR1)	6
5.2	NextGen TV with native HDR or SDR Video Essence, Advanced Audio and an Interactive Application Layer	6
6	Examples of single-master HDR productions	7
6.1	SDR-focused workflows.....	7
6.2	Dual-focused workflows.....	7
7	Use cases of 8K UHD TV for HDTV applications	8
Annex 1 London Olympics SUPER Hi-VISION Public Viewing Operations Report		8
1	Introduction	8
2	Overview	8
3	Venue production	9
4	Transmission and editing.....	14
5	Distribution and screening.....	17
6	Conclusion.....	20
Annex 2 Implementation of UHD TV broadcasting in Japan.....		21
1	UHD TV satellite broadcasting services	21
2	Specifications for UHD TV satellite broadcasting.....	21
3	List of ARIB Standards	23
References.....		23
Annex 3 Initiatives for 8K/120 Hz broadcasting		24
1	Development of high frame frequency 8K equipment	26
1.1	High speed camera.....	26

1.2	Studio Monitor.....	27
1.3	Video/audio encoder and decoder.....	28
1.4	Sheet-type flexible OLED display	28
2	Live production and transmission.....	29
Annex 4	Practice of 8K UHD programmes broadcasting on large screens in China	30
1	Overview	30
2	8K UHD video and 3D audio broadcasting and distribution system workflow	30
3	8K programme production and broadcasting in CMG	32
3.1	Basic technical specifications for 8K UHD programme video and audio	32
3.2	8K UHD programme video and audio production requirements and specifications	32
3.3	8K UHD signal routing system.....	35
3.4	8K UHD broadcast system	35
4	8K distribution and transmission.....	36
4.1	Distribution and transmission of video and audio over private network.....	36
4.2	Distribution and transmission of audio over Internet	36
4.3	Video and audio synchronization in heterogeneous network (HetNet).....	36
5	8K signal reception and display.....	37
5.1	Signal reception and display of “BaiChengQianPing”	37
5.2	System architecture.....	37
5.3	Reception and display technical specifications	38
6	Main technology	42
6.1	AVS3	42
6.2	HDR VIVID.....	42
6.3	Audio Vivid	44
	References.....	45
Annex 5	Deployment in the United States of America of NextGen TV HDR with SDR compatibility using Single-Layer HDR (SL-HDR1).....	46
	References.....	47
Annex 6	Deployment of NextGen TV in the USA using PQ HDR (with SMPTE ST 2094-10, SMPTE ST 2094-40, HDR10 metadata), Immersive Audio and a new interactive HTML5 application layer and regionalized data.....	48
Annex 7	Examples of single-master HDR production with SDR-focused workflow	50
Annex 8	Examples of single-master dual-focused HDR productions	53
Annex 9	Use cases of 8K UHD TV for HDTV applications	56
1	Multi-shot HDTV production system using a single 8K camera	56
1.1	Overview.....	56

1.2	System control software	57
1.3	Programme production	57
2	Interactive viewing of 8K UHD TV on HDTV display with zooming and panning.....	58
2.1	Overview.....	58
2.2	System architecture.....	58
2.3	Procedure for zooming and panning.....	60

Abbreviations/Glossary

BBC	British Broadcasting Corporation
CCU	Camera control unit
CMG	China Media Group
HDR	High dynamic range
HDR-TV	High dynamic range television
HEVC	High-efficiency video coding
HLG	Hybrid log-gamma
HOA	Higher-order ambisonics
IBC	International Broadcasting Centre
MMT	MPEG Media Transport
NGA	Next Generation Audio
NHK	Nippon Hoso Kyokai
NHK STRL	NHK Science and Technology Research Laboratories
NTP	Network Time Protocol
OB	Outside broadcasting
OBS	Olympic Broadcasting Service
PCM	Pulse code modulation
PQ	Perceptual quantizer
PV	Public viewings
SDR	Standard dynamic range
SHV	SUPER Hi-VISION
SL HDR1	Single Layer HDR
UHDTV	Ultra-high definition television
UWA	UHD World Association
WCG	Wide colour gamut

This Report details the experiences of those involved in producing, distributing and presenting ultra-high definition television (UHDTV) and high dynamic range television (HDR-TV) programmes.

1 London Olympics SUPER Hi-VISION¹ public viewing operations

Nippon Hoso Kyokai (NHK) and the British Broadcasting Corporation (BBC), in cooperation with the Olympic Broadcasting Service (OBS), held public viewing of SUPER Hi-VISION at venues in Japan, the United Kingdom and the United States of America for the 2012 London Olympics. This

¹ SUPER Hi-VISION complies with 7 680 × 4 320 UHDTV system described in Recommendation ITU-R BT.2020 – Parameter values for ultra-high definition television systems for production and international program exchange, and a 22.2 multichannel sound system described in Recommendation ITU-R BS.2051 – Advanced sound system for programme production.

operation consists of programme production at Olympic Games venues, post-production at a BBC studio, IP distribution of live and non-live programmes through global and domestic IP networks, and regular presentations during the Olympic Games at nine presentation venues in three countries. A summarized report is shown in Annex 1.

2 Implementation of UHD TV broadcasting

4K and 8K UHD TV satellite broadcasting has been in operation in Japan. The detailed information and specifications are described in Annex 2.

3 Initiatives for 8K/120 Hz broadcasting

8K/120 Hz broadcasting is yet to be implemented and effort has been made to develop various pieces of equipment for the 8K/120 Hz ecosystem including programme production, transmission and display. Live production and transmission of 8K/120 Hz programmes were also conducted to demonstrate the feasibility of 8K/120 Hz broadcasting. More details are provided in Annex 3.

4 Practice of 8K UHD programmes broadcasting on large screens in China

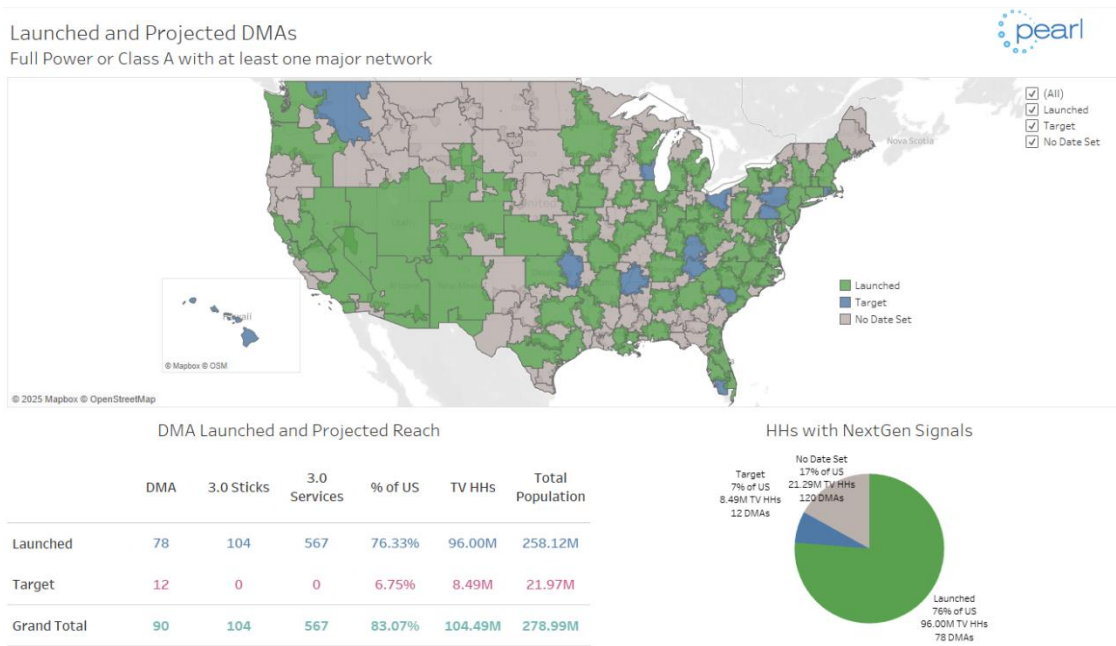
In October 2021, China Media Group (CMG) began to launch a three-year “Bai Cheng Qian Ping” (100 cities, 1 000 screens) UHD video and 3D audio promotion activities, cooperating with IT, radio and television, culture, tourism, transportation and other industries. In January 2022, CMG launched a specific 8K UHD channel to provide 8K programmes for the event. The campaign mainly showcased the UHD video and 3D audio-visual effects through the outdoor and indoor UHD large screens in public, using multicast technology to distribute CMG 8K programmes over IP to more than 70 cities, realizing the display of more than 310 UHD large screens for the Beijing Winter Olympics. More details are provided in Annex 4.

5 Deployment of NextGen TV (ATSC 3.0) in the United States of America

ATSC 3.0 (deployed on consumer devices under the logo certification program known as “NextGen TV”) offers:

- Over-the-air broadcasts of UHD video quality through advanced video codecs which support HDR and SDR. NextGen TV provides a huge advancement from the current ATSC 1.0 over-the-air DTV broadcasts that utilize MPEG-2 Video. It effectively lowers data rate requirements for high quality video by 4 times.
- Immersive audio thru the advanced AC-4 codec. It lowers data rate requirements for high quality audio.
- A modern IP-based streaming system for playback using MPEG-DASH.
- An interactive application-layer that uses a hybrid over-the-air and IP-based model which presents an HTML5 rendered “canvas” with composited video, images and audio for televisions.
- Advanced emergency alerting and informing, including hyper localization.
- Additional innovative features to enhance and expand the broadcast viewing experience.
- See Figure 1 for a regional breakdown of NextGen TV deployment for the United States of America.

FIGURE 1
Current NextGen TV deployment in the USA (January 2025)



Through a single shared encoding and transmission system per region (“lighthouse”), multiple broadcasters have deployed NextGen TV in cities around the United States of America. These stations take advantage of the new capabilities in NextGen TV using multiple new standards defined in ATSC, MPEG, SMPTE and W3C.

5.1 NextGen TV High Dynamic Range with Standard Dynamic Range compatibility using single-layer High Dynamic Range (SL-HDR1)

ATSC 3.0, also known as “NextGen TV” under a certification programme for consumer TVs, STBs, etc. offers 4K ultra high-definition video quality, theatre-like sound, mobile reception (handheld and vehicular reception of broadcast television outside of the home) and innovative new features to enhance and expand the broadcast viewing experience. One of the most prominent features of ATSC 3.0 broadcasting is the use of High Dynamic Range (HDR) to provide consumers with better picture quality. One of the adopted solutions for HDR is the single layer HDR (SL-HDR1) technology, providing both an SDR picture and an HDR picture from the same video stream (Standard Dynamic Range (SDR) plus metadata to reconstruct an HDR image). From a distribution perspective, the single-layer format thus provides a cost-efficient workflow that cuts bandwidth in half, to deliver all available programming in both HDR and SDR formats using a single video stream. As of March 2025, SL-HDR1 is being transmitted in more than 50 markets in the USA, where more than 170 channels deliver content in HDR, 24 hours a day, seven days a week, with SL-HDR1 technology². More details are provided in Annex 5.

5.2 NextGen TV with native HDR or SDR Video Essence, Advanced Audio and an Interactive Application Layer

Using the very same “lighthouses” described in § 5, additional broadcasters are presenting native HDR and SDR channels utilizing improved video and audio codecs. Currently there are 300 SDR HDTV, 86 UHD TV (Perceptual Quantizer (PQ)), and 16 SDTV channels (see Recommendations

² A set-top-box or TV with an ATSC 3.0 SL-HDR1 decoder is required to view HDR.

ITU-R BT.709, ITU-R BT.2100 and ITU-R BT.601, respectively) broadcasting signals in their native formats³. Hybrid Log-Gamma (HLG) is supported in ATSC 3.0 but not yet used for current broadcast channels; one broadcaster tentatively intends to use it. In all use-cases, the native HDR or SDR video essence is delivered to the display device without conversion.

NextGen TV tuners typically support native HDR displays. If the TV does not have a NextGen TV tuner, an outboard set-top box is required for HDR rendering.

In earlier systems like ATSC 1.0, and conventional cable, SDR essence will continue to be used during the transition period from HDTV to UHDTV and from SDR to HDR. Newer encoders and playout systems are available to convert HDR to SDR (or the reverse) from incoming video.

Audio is delivered via AC-4 which supports lower bit rate, higher quality contribution or distribution rate delivery of immersive audio with support for many other advanced features like audio objects for multiple languages, dialog adjustments/isolation, or audio description.

Finally, in some cases, an application layer is utilized to present regional interactive data pulled from the internet through a hybrid model supported in NextGen TV. Examples of real-world deployments with the methods described in this section including the application layer are provided in Annex 6.

6 Examples of single-master HDR productions

Single-master workflows for live HDR production are documented in Report ITU-R BT.2408 – Suggested guidance for operational practices in HDR television production. It enables simultaneous production of HDR and SDR signals that conform to standards for international programme exchange.

6.1 SDR-focused workflows

An SDR-focused workflow has been widely used by the BBC and Sky UK for their UHD-HDR live productions. It has been found to facilitate easy migration to HDR production and to be efficient in terms of staff education, hardware, and complexity of infrastructure, whilst protecting the SDR pictures. It uses HLG as the core HDR format, with a fixed down-mapper to generate an SDR output, so no metadata is required. Annex 7 provides examples of single-master HDR productions with an SDR-focused workflow that uses a gamma-adapted tone mapping.

6.2 Dual-focused workflows

A Dual-focused workflow has been widely used in the USA by NBCU Sports Group, NBCU Saturday Night Live 50th Anniversary Radio City Music Special, Fox Sports, Amazon Prime Video Live Sports productions. It allows the content creators to balance the capabilities of each format with limited sacrifices given that both HDR and SDR formats can be monitored simultaneously by video shaders, vision supervisors and with control room multiviews so efficient evaluation of both images is possible. These workflows currently use HLG as the core HDR format, with a fixed down-mapper to generate the SDR-essence for output. The most common transmission method uses the HLG HDR video format for production and then converts HLG to perceptual quantizer (PQ) for final transmission so that an absolute-luminance mapping is provided to a target display (television) for viewers. PQ supports sophisticated tone mapping capabilities through the use of static or dynamic metadata that attempts to reduce clipping and preserve more detail given the vastly different luminance capabilities in the television marketplace. See Annex 8 for examples of productions using a single-master UHD-HDR dual-focused workflow.

³ A set-top-box or TV with an ATSC 3.0 tuner is required. Source for station count: PearlTV “lighthouse” database.

7 Use cases of 8K UHDTV for HDTV applications

8K UHDTV that provide a wider field of view with higher resolution can provide use cases for HDTV production and viewing. Annex 9 provides examples of a multi-shots HDTV production system using a single 8K camera and interactive viewing of 8K UHDTV on HDTV display with zooming and panning.

Annex 1

London Olympics SUPER Hi-VISION Public Viewing Operations Report

1 Introduction

For the London Olympics, the British Broadcasting Corporation (BBC) and Nippon Hoso Kyokai (NHK) held SUPER Hi-VISION (SHV) public viewings (hereinafter PV) at venues in Japan, the United Kingdom (Fig. 2) and the United States of America, in cooperation with the Olympic Broadcasting Service (OBS).

SHV is a next-generation broadcasting system being developed by NHK, which reproduces strong sensations of presence, as though the viewer was actually there, through the use of ultra-high-resolution, 33-megapixel video (7 680 pixels horizontally by 4 320 pixels vertically) and 22.2 multichannel audio (22.2ch audio).

FIGURE 2
PV venue (Bradford, UK)

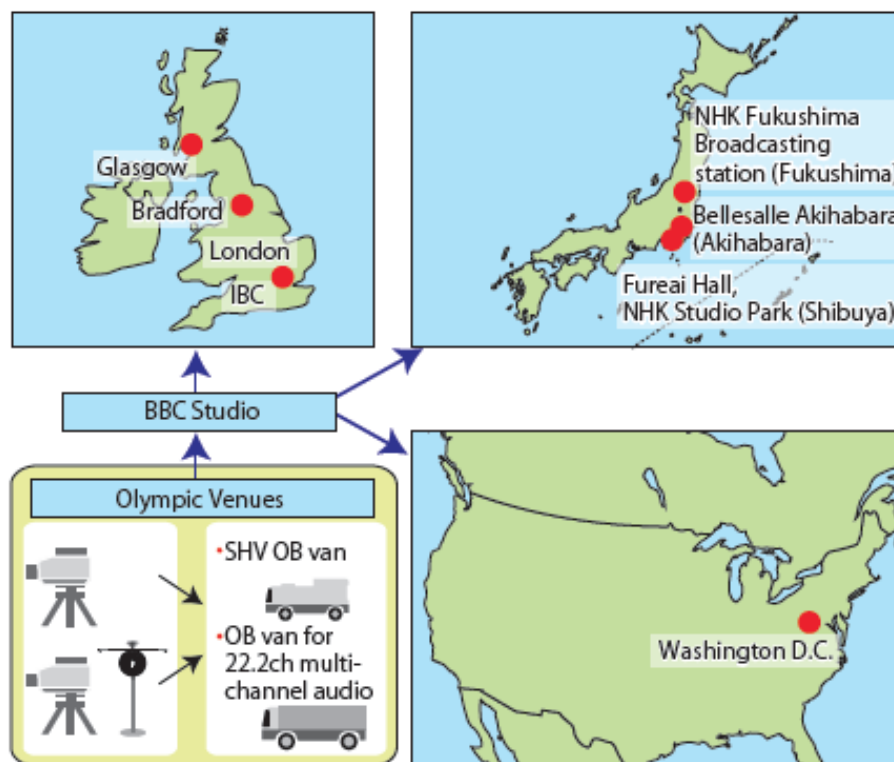


2 Overview

Content from seven events at the Olympic venues was produced. These were the opening ceremony, swimming, basketball, athletics, cycling, synchronized swimming and the closing ceremony. Content created at the venues was sent by optical fibre to a temporary production and transmission base established in TC0, a studio at the BBC Television Centre London. Live content was output as is from Television Centre while other content was edited and packaged. Both live and non-live programmes were compressed into transmission streams (TS) with a rate of approximately 300 Mbit/s, then

distributed by IP lines to PV venues within the United Kingdom except for the International Broadcasting Centre (IBC), in the United States of America, and in Japan for presentation (Fig. 3).

FIGURE 3
System overview



An uncompressed signal was sent from the BBC Television Centre to a PV theatre in the IBC within the Olympic Park to hold PVs for broadcasting-related attendees gathered from around the world at the Olympic venues.

The opening and closing ceremonies and swimming competitions were presented live in the United Kingdom and the United States of America. Due to time differences, only the morning (local time) part of the July 30 session of the swimming competition was shown live in Japan.

3 Venue production

1) Outside broadcasting-van (OB-van) system

Two trucks were used at each location: one for video (Fig. 4), and the other for audio (Fig. 5). They were staffed with a single crew and after each event they had to de-rig the equipment before moving to the next venue.

FIGURE 4
OB-van for Video



FIGURE 5
OB-van for Audio



The video OB-van was a rental truck that was furnished with empty equipment racks before being shipped to London from Japan. Once there, it was rigged with the SHV production equipment to create the video OB-van. Production equipment included an eight-input switch system with resources including two SHV cameras (three at the opening and closing ceremonies), two SSD live-slow-motion devices, a host HD signal up-converter, and graphics equipment.

2) *Shooting*

Three cameras equipped with 1.25-inch four CMOS image sensors were used at the opening and closing ceremonies, this was reduced to two cover the competitions. Two sets of lenses were used, each included a 5×12 mm to 60 mm lens (Fig. 6) and a 10×18 mm to 180 mm lens. These were selected by venue.

FIGURE 6
1.25-inch 4-CMOS camera and 5x lens

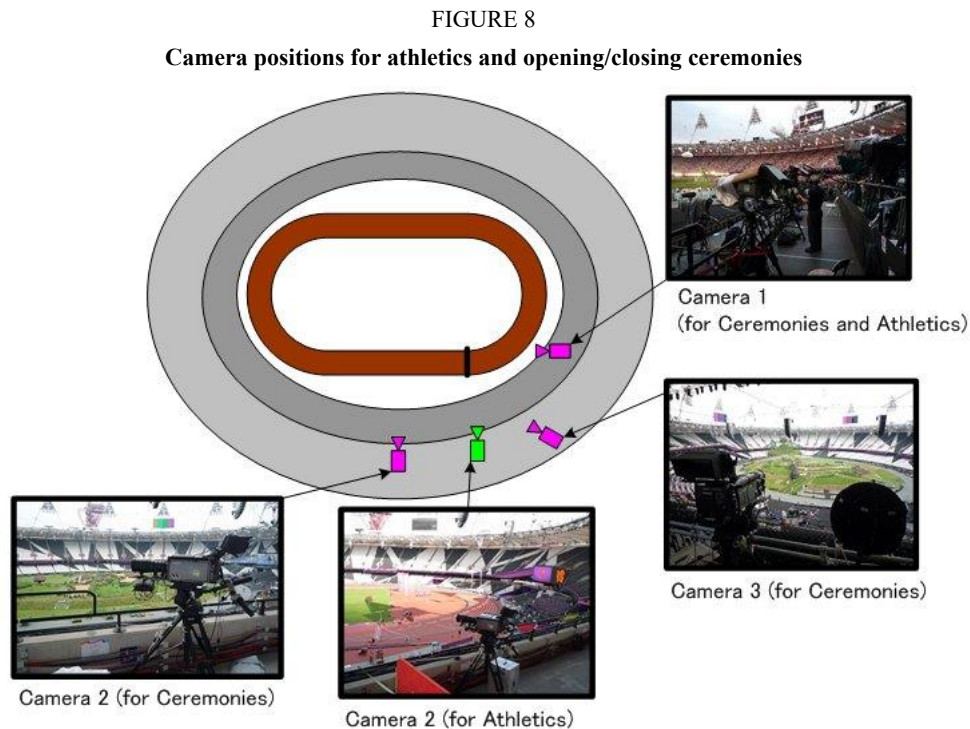


One feature of SHV is the different viewing distance. The optimal viewing distance for ordinary Hi-Vision is three-times the screen height, but for SHV it is 0.75-times the screen height. This gives it a horizontal field of view of 100 degrees. SHV covers a wide viewing angle with high resolution images and produces an immersive sensation when viewed at close distances. With large screens greater than 400 inches, viewers get a sensation of being wrapped in the field of view, producing a sense of presence, even though they cannot recognize any information in the peripheral areas without consciously looking there. This meant shots had to be composed to place the most important action near the centre of the screen. The camera viewfinder had “action area” markings to help the camera operators maintain a good shot framing (Fig. 7).

FIGURE 7
Viewfinder markers showing large screen



It was decided that each venue would have a master or base camera in a high position while the other camera would give a different angle from a low position and was primarily used for close-ups. Depending on the venue, the low-position camera gave more depth and was able to convey a greater sense of presence, so it was switched in and out with the base camera shot during each event. The camera positions for athletics and opening/closing ceremonies are described in Fig. 8.



Care was also taken to make sure there was a link between the sense of presence given by the images and the sense of presence provided by the audio. An audio monitoring environment with 5.1-channel audio down-mixed from the 22.2 ch audio was provided in the OB-van production room, so that video switching could complement the sounds from the venue and from the cheering crowds.

Due to space limitations in the video monitoring environment, work was done with 28-inch 4K monitors, so technicians had to struggle with a perceptual gap between the monitoring in the OB-van and the actual viewing environments at the large PV screens. In order to check the content in a real viewing environment, the OB crew had to view the 145-inch PDP display in the IBC theatre, or go to the PV venue in London to view the 300-inch screen there.

3) *Audio*

A temporary, locally rented, audio OB-van (Fig. 5) was used. It was installed with a 22.2 ch live mixing board and 22.2 ch speakers (Fig. 9). The mixing desk was developed at NHK Science and Technology Research Laboratories (NHK STRL) with operations optimized for live production and incorporating 3D sound-image positioning (3D panning) functions.

A 22.2 ch-audio, one-point microphone (Fig. 10) was used. It was a fixed, monolithic microphone holder consisting of a 45-cm sphere divided into upper, middle and lower layers. Each layer was partitioned into eight directions using sound baffles and a compact microphone installed in each partition. The microphone was positioned near the base camera, and by panning to arbitrary locations on the sound desk (using the outputs from each microphone) a sound field similar to that heard by the event audience could be reproduced. It was not permitted to place microphones freely for SHV production in the venues, so audio was produced using the one-point microphone as a base, and mixing it with the microphone feeds distributed from the international feed. Use of 3D reverb equipment (22.2 ch audio reverb) also helped in reproducing the expansive feeling in the venues faithfully.

The 22.2 ch output feeds were recorded using a multi-track recorder in case the signals from live PVs are used for other purposes later.

FIGURE 9
Inside audio OB-van



FIGURE 10
22.2 ch one-point microphone



4) Other *issues*

Presentations at the IBC on-site theatre also gathered large audience, suggesting interest in SHV increased daily amongst overseas broadcasters. Within the venue compounds, people from the media visited in groups to see the video and audio OB-vans. As promoting SHV was one of the objectives of this SHV trial, in addition to holding the PVs, visitors were welcomed and given explanations of SHV video and 22.2 ch audio. Visiting broadcaster staff (Fig. 11) came to see the high resolution video and explanations of SHV in the OB-van were done using 4K monitors. In the audio OB-van, visitors were able to sit in the main-mixer seat and experience 22.2 ch sound.

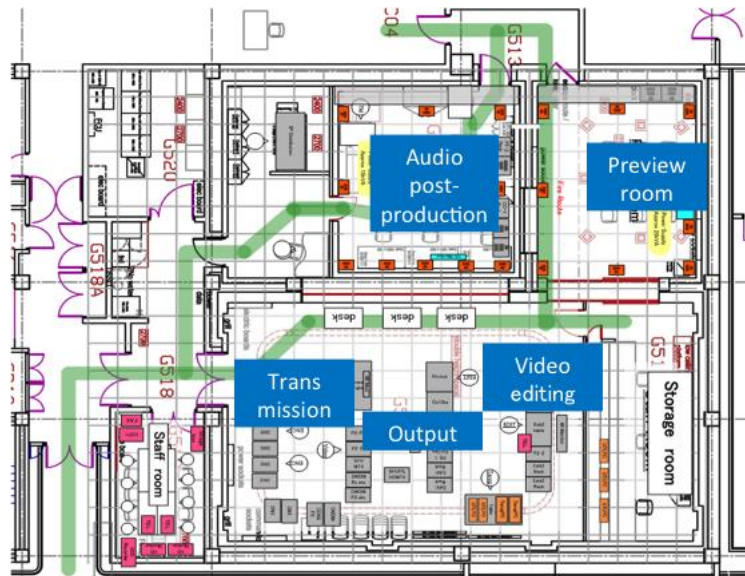
FIGURE 11
Broadcasting staff viewing the monitor beside the OB-van



4 Transmission and editing

For production of the SHV PVs at the London Olympics, transmission, play-out, video editing, audio post-production, and a monitor room were combined in a compact layout in the BBC TC0 studio (Fig. 12). Much consideration was given to the layout of cabling and provision of quality monitoring.

FIGURE 12
BBC TC0 studio layout



1) *Line and transmission systems (Fig. 13)*

An SHV 8×8 routing switcher system (an HD 128×128 router configured to switch 16 channels at a time) was used to access the various resources in TC0 effectively, including the live signals (main and backup) brought in from each venue by optical lines, and the P2 recorders (main and backup) for playing back content. An additional monitoring system was also built separately from the main system to allow the signals from the relay site to be checked even during a PV (Fig. 14).

FIGURE 13
Line and transmission equipment



FIGURE 14

8K-4K down-converter and 28-inch 4K screen for monitoring (left)



2) *Video post-production*

For competitions that were not covered live, the main and backup feeds were recorded using the P2 recorders in the TC0 studio. After the competition, two SHV editors were used to produce content for the next day. Two versions were edited simultaneously, a World Feed (WF) for the UK and USA, and a Japan Feed (JF) for Japan. Segments of approximately ten minutes were edited according to medal standings or popularity, and several events were combined into a package of approximately 45 minutes duration. Editing work was done overnight from midnight to 10 am, for advance distribution to Japan and presentation in the UK and USA starting at noon. The SHV recorder used 17 P2 cards for recording. One of these was used to record HD proxy data.

This recording was then used with an off-line HD proxy editor for editing (Fig. 15). The SHV video was later matched to the HD proxy using this editing data (Fig. 16). Captions were created in 8K (7 680 × 4 320) TIFF files using graphics overlay equipment and imported to the SHV editor to create video, and a 24-channel WAV file of the completed audio master was imported and merged into the SHV editor time line to complete the content.

FIGURE 15

HD proxy editor



FIGURE 16
SHV editor



3) *Audio post-production*

For live transmissions, the on-site audio mix was transmitted to the PV venues, but for edited productions, a 22.2 ch multichannel audio post-production process had to be undertaken. Two simultaneous audio edits were produced (to match the two video edits) by using the HD proxy editing data. Two audio post-production systems were built that included a digital audio workstation (DAW) with a mixing desk that supports 22.2 ch audio and a backup DAW.

These were set up in the audio mixing room (Fig. 17). The audio editing systems also used the 22.2 ch sound speaker system (Fig. 18) in the SHV preview room (see next section). Audio post-production made it possible to manage the overall audio balance for multiple events, improving the quality of the content.

FIGURE 17
Main mixing room



FIGURE 18

Sound equipment in the preview room



4) *SHV preview room*

A preview room was created using an 85-inch SHV LCD monitor and a 22.2 multichannel audio system to enable video and audio signals for live transmissions and the edited content to be previewed reliably (Fig. 19). This was primarily to manage the quality of the signals, but it also allowed viewing on a large monitor at the ideal distance and with the ideal speaker layout. This was used to give advice to on-site teams on aspects such as the camera work and switch timing in order to increase the sense of presence and immersion given by SHV.

FIGURE 19

SHV preview room



5 Distribution and screening

1) *Distribution and screening overview*

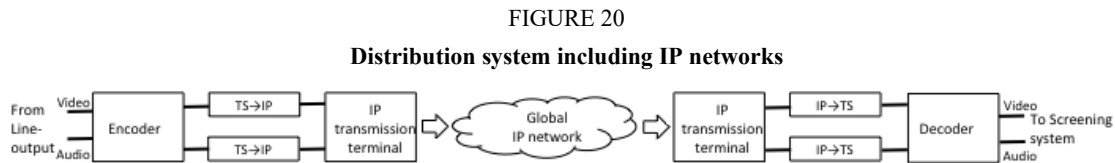
Distribution and screening refers to delivery of the packaged content and live signals from the BBC studio to each of the PV venues and the presentation at the venues. Within the overall system shown in Fig. 3, this section corresponds to the part after the BBC studio and includes encoding of the baseband signals from the BBC studio equipment, transmission over IP, the transport system that decodes and restores the baseband signals and the screening system that reproduces the video and audio signals from the baseband signals.

IP networks within the United Kingdom were used to transmit signals to the three locations (London, Bradford, and Glasgow). Dark fibre was used to transmit uncompressed signals to the IBC within the Olympic complex. Global IP networks were used to transmit to Washington in the USA, and to Tokyo in Japan. The domestic network in the USA was used to transmit from the terminus of the global network to the venue within the USA. From the Tokyo terminus, signals were transmitted to NHK

STRL by domestic lines and then distributed by IP multicast over dedicated domestic lines to three locations in Japan (Shibuya, Akihabara, and Fukushima).

2) *Distribution system*

A system diagram of the distribution system [1][2], consisting of encoder, transport equipment and IP network, is shown in Fig. 20.



The encoder [3] used AVC/H.264 to encode the video and MPEG-2 AAC-LC to encode the audio. The interfaces between SHV equipment converted the so-called dual-green format equivalent to the RGB Bayer matrix signals to ordinary YUV signals that can be compressed using AVC/H.264. In this process, the green signal composed with diagonally-pixel-offset two-channel signals with $3\,840 \times 2\,160$ pixels is converted to the $7\,680 \times 2\,160$ format by shifting the one of the two signals by half pixel. Then Y signal with $7\,680 \times 2\,160$ pixel is calculated from this new green signal and up-converted red and blue signals. The YUV signals were partitioned into eight $1\,920 \times 1\,080$ pixel signals (1 080-60P signals) for handling. An encoder for each of these elementary signals compressed the 1 080-60P signal using AVC/H.264, and up to four audio channel signals using MPEG-2 AAC-LC. Finally, the MPEG-2 TS signals output by the encoder units were multiplexed into two MPEG2-TS signals by TS multiplexers. After the MPEG-2 TS signals output by the encoder equipment were converted to IP signals, they were transmitted over IP networks.

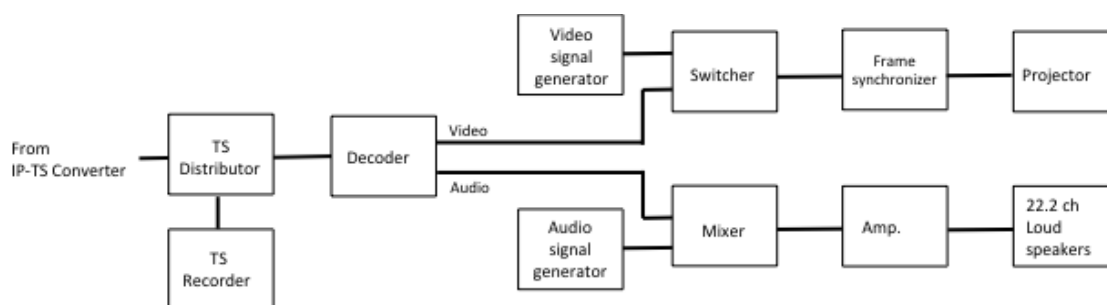
In order to transmit the 300 Mbit/s compressed MPEG-2 TS signal SHV video and audio over global IP networks, the required bandwidth had to be secured, fluctuations that occurs on the network had to be compensated for, and a mechanism for maintaining security was needed. Specifically, functions to control jitter for synchronous transmission, for real-time encoding and decoding, to handle packet loss, and for advanced error correction [4] were needed. These functions were built into the IP transmission terminal equipment.

The decoding equipment performed the same processing as the encoding equipment, in reverse, to produce the baseband video and audio signals.

3) *Screening system*

A system diagram of the screening system is shown in Fig. 21. Content was distributed from NHK STRL to the PV venues in Japan, and from the BBC Studio TC0 to the PV venues in the UK and the USA. The video and audio signals were decoded then presented using SHV projectors with 22.2 ch audio equipment at each venue. Each venue was also equipped with TS recording and playback equipment. This was to be used as a backup system if the lines between NHK STRL and each venue experienced difficulty or, if for any reason any of the venues needed to make their own playback schedule. A baseband signal generator is also built into the system to enable adjustments to the audio and video systems.

FIGURE 21
System diagram for screening system



For video display two types of projector were used as well as an 85-inch LCD, a 145-inch PDP, and a 360-inch multi-screen LCD. At five of the theatre venues the SHV projectors [5] used had 8-megapixel display devices and a pixel-shifting technology called e-shift to increase the resolution, resulting in compact size and low power consumption. The projector in the NHK Minna no Hiroba Furea Hall used 33 megapixel display devices for RGB [6], making it a so-called full-resolution projector. This projector has a high output power so was used as here as the venue has a 520-inch large screen.

The 85-inch LCD [7] was used for direct viewing in the Washington venue, and also in the Akihabara and Fukushima venues. The 85-inch display was combined with a sound system to demonstrate a possible home system. The 145-inch PDP [8] was used to set up a theatre in a room at the IBC and the 360-inch multi-screen LCD was installed at the entrance to the NHK Studio Park tour for NHK visitors.

A new 22.2 multichannel audio system was developed for direct viewing displays and combined with the 85-inch LCDs and the 145-inch PDP. In theatres using a projector, the theatre sound system was used.

4) Screening venues and reactions

TABLE 1
List of screening venues

Country	City	Venue	Screen size (inches)	Display device
Japan	Tokyo	NHK Minna no Hiroba Fureai Hall	520	Projector
		NHK Studio Park	360	Multi-screen LCD
		Belle Salle Akihabara	300	Projector
	Fukushima	NHK Fukushima broadcasting station	350	Projector
United Kingdom	London	BBC Broadcasting House	300	Projector
		IBC	145	Plasma
	Bradford	National Media Museum	250	Projector
	Glasgow	BBC Pacific Quay	350	Projector
United States of America	Washington DC	Comcast (NBC)	85	LCD

The display equipment and screen size used in each PV venue are shown in Table 1.

PVs were held in Japan from July 28 to August 12 at the NHK Minna no Hiroba Fureai Hall in Shibuya (Fig. 22), NHK Studio Park, in the Belle Salle Akihabara event space near the JR Akihabara station, and at the NHK Fukushima broadcasting station. At each venue several related events were held at the same time as the screening of Olympic content in order to attract many more people and to ensure that those who came really enjoyed it. The Shibuya venue catered to families during the summer holidays, the Akihabara venue to young people and those interested in technology, and the Fukushima venue appealed to supporters of local athletes. A total of over 200 000 people attended over all three events. At the venues, there were exclamations of wonder as the screening began, and many comments about the strong sense of presence and intensity of the presentation.

FIGURE 22
PV in the Fureai Hall



In the United Kingdom, the BBC took the lead with NHK cooperating, and screenings were held from July 23 to August 12 at the BBC Broadcasting House in London, at the National Media Museum in Bradford, and at the BBC Scotland building in Glasgow. At the IBC, the OBS took the lead providing demonstrations mainly for broadcasting-related attendees. Before the opening ceremony, a short item introducing SHV was screened along with a NHK and BBC co production featuring sights and sounds around London and the Olympic venues in the run-up to the games. Many dignitaries and people related to broadcasting visited during the screenings, and there were many comments about how wonderful they were, and inquiries of when broadcasting would begin.

In the United States of America, NBC took the lead with cooperation from NHK, and screenings were held from July 27 to August 12 in meeting rooms in the Comcast building in Washington DC, mainly by invitation to people from government and the content, communications and electronics industries.

6 Conclusion

Advanced development of SHV with “presence” is its strongest feature. These events have once again shown the extremely strong sense of presence delivered by SHV video and audio, and the unprecedented levels of emotion can be imparted on viewers, giving them a sense they are actually at the Olympic venue. It was also showed that SHV can operate much like ordinary broadcasting, by producing and transmitting programmes continuously, every day during the Olympic Games using live coverage and recorded and edited content.

A completely different style was also used in production of the content, without using voice-overs (announcing or comments), and using mainly wide camera angles and long (slow) cut ratios. These were met with many comments of surprise, admiration, and of the new possibilities presented for broadcasting businesses.

Annex 2

Implementation of UHDTV broadcasting in Japan

1 UHDTV satellite broadcasting services

UHDTV satellite broadcasting has been operating in Japan on the basis of the domestic standards developed in 2014 and additional specifications for HDR-TV issued in 2016. There are two platforms for UHDTV satellite broadcasting in Japan, which are called BS and CS. In March 2015, 4K broadcasting via CS was commenced by SKY Perfect JSAT Corporation. 4K and 8K test broadcasting via BS was launched by NHK (Japan Broadcasting Corporation) in August 2016 and was taken over by the joint operation by NHK and A-PAB (Association for Promotion of Advanced Broadcasting Services) in December 2016. 4K and 8K regular broadcasting services via BS started on December 1, 2018. As of October 2020, nine channels of 4K and one channel of 8K broadcasting services via BS, and eight channels of 4K services via CS are in operation.

2 Specifications for UHDTV satellite broadcasting

Japanese domestic standards for broadcasting consist of Ministerial Ordinances and Notices and ARIB (Association of Radio Industries and Business) Standards and Technical Reports. They have been developed on the basis of international standards including ITU-R Recommendations, ITU-T Recommendations, and ISO/IEC International Standards.

Figure 23 illustrates the protocol stack of the ISDB-S3 system and Table 2 shows the specifications for UHDTV satellite broadcasting.

The UHDTV video signal and 22.2-channel audio signal are encoded by ITU-T H.265|MPEG-H HEVC and MPEG-4 AAC, respectively. As for closed captions and multimedia content, Timed Text Markup Language (TTML) and HTML 5 are used for the coding, respectively.

These coded signals are encapsulated into MPEG Media Transport (MMT), which is a media transport protocol over IP. IP packets constructed like this are multiplexed over the broadcast channel with the TLV multiplexing scheme.

FIGURE 23

Protocol stack of ISDB-S3 system

4K·8K HLG video	22.2-ch audio	Closed caption	Multimedia content
HEVC	AAC, ALS	TTML	HTML5
MMT/IP			
TLV multiplexing scheme			
ISDB-S3 channel coding & modulation			

TABLE 2
Specifications for UHDTV satellite broadcasting in Japan

		Wide-band bandwidth: 34.5 MHz	Narrow-band bandwidth: 27 MHz
Video signal	Format	8K: 7 680 × 4 320/59.94/P, 4K: 3 840 × 2 160/59.94/P, (2K: 1 920 × 1 080/59.94/P)	4K: 3 840 × 2 160/59.94/P, (2K: 1 920 × 1 080/59.94/P)
		ARIB STD-B32 ^{*1} , Rec. ITU-R BT.2020, Rec. ITU-R BT.2100	
	Colorimetry	Wide colour gamut	
		ARIB STD-B32, Rec. ITU-R BT.2020, Rec. ITU-R BT.2100	
	Transfer function	HLG, (PQ), SDR	
		ARIB STD-B32, Rec. ITU-R BT.2100	
	Coding	H.265 HEVC Main 10 profile	
		ARIB STD-B32, Rec. ITU-T H.265 ISO/IEC 23008-2, Rec. ITU-R BT.1870, Rec. ITU-R BT.2073	
Audio signal	Speaker configuration	22.2, 7.1, 5.1, 2.0	5.1, 2.0
		ARIB STD-B32, Rec. ITU-R BS.2051	
	Coding	MPEG-4 AAC LC, MPEG-4 ALS	MPEG-2 AAC
		ARIB STD-B32, Rec. ITU-R BS.1196	
Multimedia		HTML5	None
		ARIB STD-B62 ^{*2} , Rec. ITU-R BT.2075, Rep. ITU-R BT.2342	
Transport		MMT/TLV, (MPEG-2 TS)	MPEG-2 TS
		ARIB STD-B32, ARIB STD-B60 ^{*3} , Rec. ITU-R BT.2074, Rec. ITU-R BT.1869	ARIB STD-B10 ^{*4} , ARIB STD-B32, Rec. ITU-T H.222.0 ISO/IEC 13818-1
CAS		AES, (Camellia)	Multi 2
		ARIB STD-B61 ^{*5} , Rec. ITU-R BT.1852	Private
Transmission		ISDB-S3	DVB-S2
		ARIB STD-B44 ^{*6} , Rec. ITU-R BO. 2098	Rec. ITU-R BO.1784
Receiver configuration		ARIB STD-B63 ^{*7}	ARIB STD-B1 ^{*8} , ARIB STD-B16 ^{*9}
Operational guidelines		ARIB TR-B39 ^{*10}	Private

NOTE 1: Parameter values in parentheses are neither specified in the operational guidelines nor currently used in the broadcasting.

NOTE 2: Superscripts indicate the relevant standards shown in § 3.

3 List of ARIB Standards⁴

- 1) ARIB STD-B32, “Video coding, audio coding and multiplexing specifications for digital broadcasting”,
http://www.arib.or.jp/english/html/overview/doc/6-STD-B32v2_1-E1.pdf
- 2) ARIB STD-B62, “Multimedia coding specification for digital broadcasting (2nd generation)”,
http://www.arib.or.jp/english/html/overview/doc/6-STD-B62v1_2-1p2-E1.pdf
- 3) ARIB STD-B60, “MMT-based media transport scheme in digital broadcasting systems”
- 4) ARIB STD-B10, “Service information for digital broadcasting system”,
http://www.arib.or.jp/english/html/overview/doc/6-STD-B10v4_6-E2.pdf
- 5) ARIB STD-B61, “Conditional access system (second generation) and CAS program download system specifications for digital broadcasting”
- 6) ARIB STD-B44, “Transmission system for advanced wide band digital satellite broadcasting”,
http://www.arib.or.jp/english/html/overview/doc/6-STD-B44v2_0-E1.pdf
- 7) ARIB STD-B63, “Receiver for advanced wide band digital satellite broadcasting”
- 8) ARIB STD-B1, “Digital receiver for digital satellite broadcasting services using communication satellite”
- 9) ARIB STD-B16, “Digital receiver commonly used for digital satellite broadcasting services using communication satellites”
- 10) ARIB TR-B39, “Operational guidelines for advanced digital satellite broadcasting”

References

- [1] Nojiri, Iguchi, Noguchi, Fujii, Ogasawara: “National Super Hi-Vision Transmission Test using IP Networks for Global Research and Education”, Broadcast Technology (Hoso Gijutsu), Vol. 64, No. 6, pp. 135-141 (2011) (Japanese).
- [2] Sakaida, Iguchi, Kimura, Ogasawara, Fujii: “International Super Hi-Vision Transmission Test and Exhibition of Related Equipment at IBC2011”, Broadcast Technology (Hoso Gijutsu), Vol. 65, No. 1, pp.151-156 (2012) (Japanese).
- [3] Y. Shishikui, K. Iguchi, S. Sakaida, K. Kazui, A. Nakagawa: “Development of High Performance Video Codec for Super Hi-Vision”, 65th NAB Broadcast Engineering Conference, pp. 234-239 (2011).
- [4] Y. Tonomura, D. Shirai, T. Nakachi, T. Fujii, H. Kiya: “Layered Low-Density Generator Matrix Codecs for Super High Definition Scalable Video Coding System”, IEICE Trans. on Fundamentals, Vol. E92-A, No. 3, pp. 798-807 (2009).
- [5] F. Okano, M. Kanazawa, Y. Kusakabe, M. Furuya, Y. Uchiyama: “Complementary Field Offset Sampled-Scanning for GRB Video Elements”, IEEE Trans. Broadcasting, Vol. 58, No. 2, pp. 291-295 (2012).
- [6] T. Nagoya, T. Kozakai, T. Suzuki, M. Furuya, and K. Iwase, “The D-ILA device for the world’s highest definition (8K4K) projection system”, Proc. Int. Display Workshop (IDW) 2008, Vol. 15, pp. 203-206 (2008).

⁴ Some of the ARIB Standards are available in English and free of charge.

- [7] T. Kumakura, M. Shiomi, S. Horino, Y. Yoshida, S. Mizushima: “Development of Super Hi-Vision 8Kx4K Direct-View LCD for Next Generation TV”, SID 2012 Digest, pp. 780-783 (2012).
- [8] K. Ishii, T. Usui, Y. Murakami, Y. Motoyama, M. Seki, Y. Noguchi, T. Furutani, T. Nakakita, T. Yamashita: “Developments of a 145-inch Diagonal Super Hi-Vision Plasma Display Panel”, SID 2012 Digest, pp. 71-74 (2012).

Annex 3

Initiatives for 8K/120 Hz broadcasting

Various pieces of production equipment as well as transmission and display technologies for high frame frequency 8K programmes have been developed to achieve much higher quality with clearer and smoother subject movement (Fig. 24). A complete system ranging from live on-site production to encoding, satellite transmission, and display reproduction at a frame frequency of 120 Hz (120/1.001 Hz) was exhibited at the NHK STRL Open House 2019 in cooperation with Broadcasting Satellite System Corporation (Fig. 25).

FIGURE 24
8K/120 Hz equipment

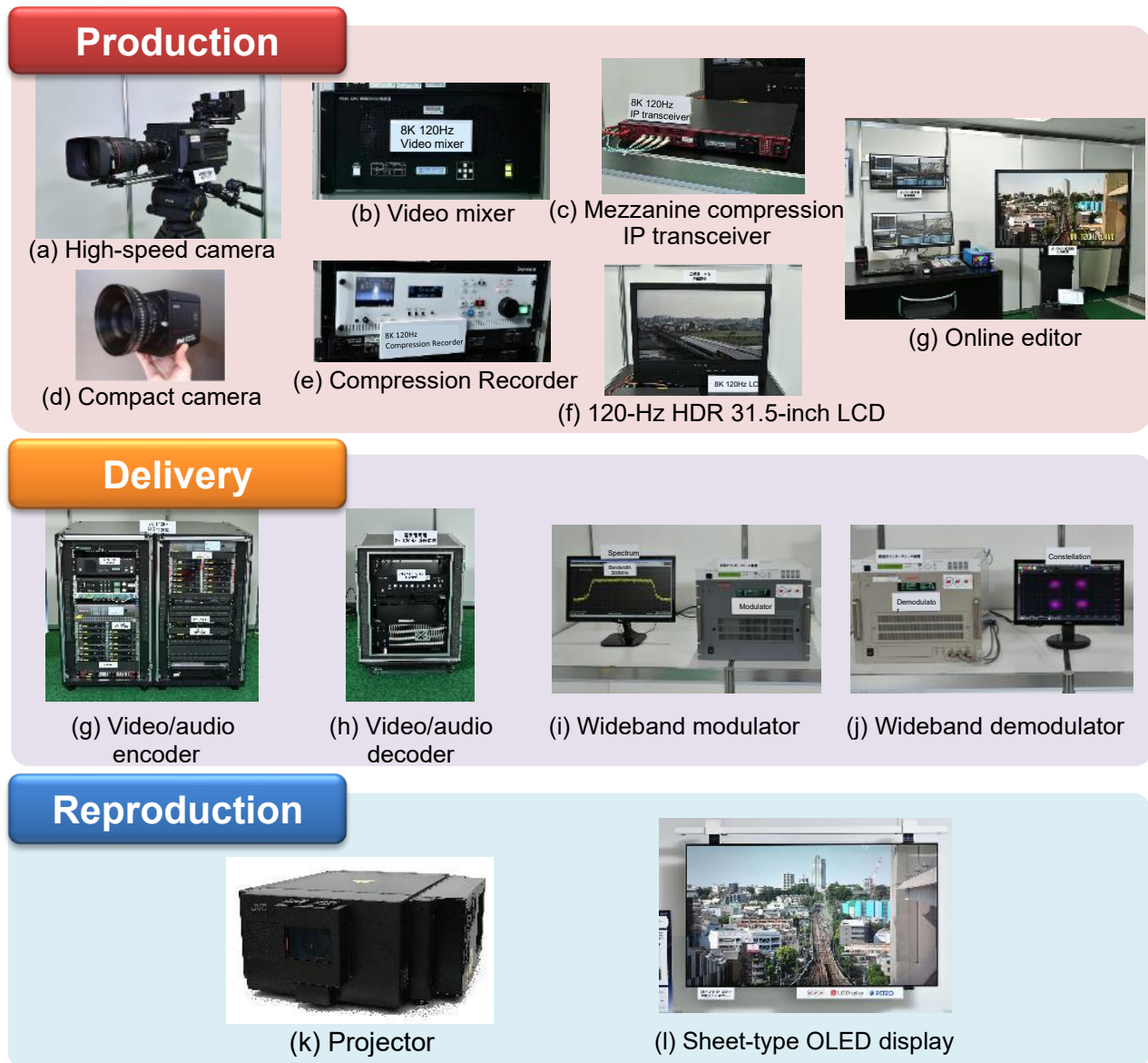
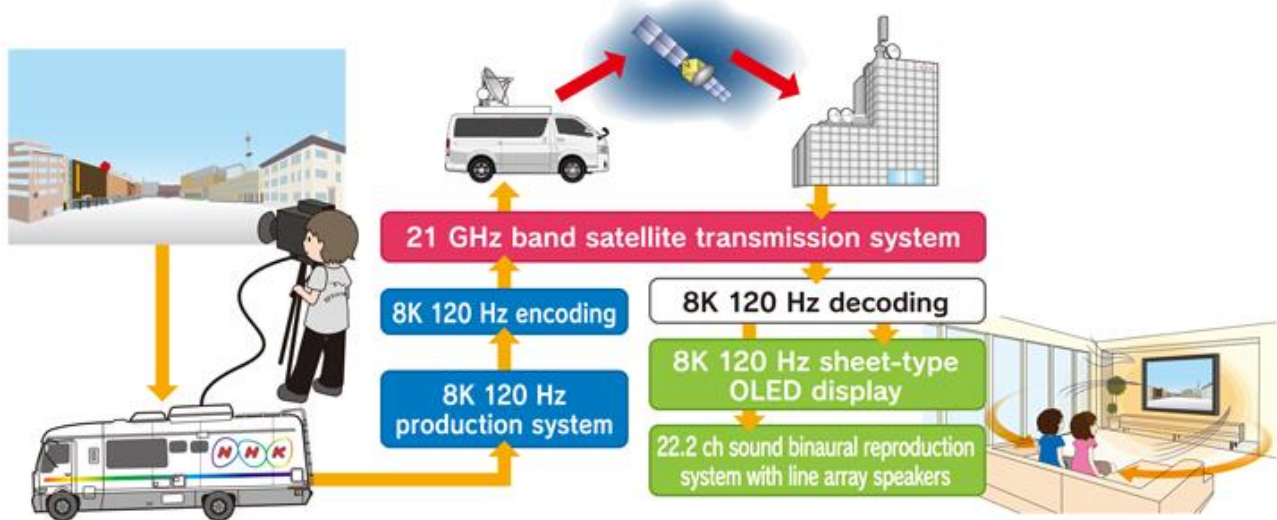


FIGURE 25
8K/120 Hz live broadcast experimental system



1 Development of high frame frequency 8K equipment

1.1 High speed camera

An 8K/240 Hz full-resolution high-speed camera (Fig. 24(a)) and slow-motion replay server system for the creation of sports content have been developed. The high-speed camera acquires 8K/240 Hz video via three 1.25-in 33-Mpixel 240-Hz CMOS image sensors. Signals are transmitted between a camera head and a camera control unit (CCU) via a four-core optical camera cable, and the CCU then processes the sensor signal and outputs 8K/240 Hz RGB 4:4:4 video signals. The 8K/240 Hz slow-motion replay server system can record up to 240 min of 8K/240 Hz video while simultaneously playing back 8K/60 Hz slow-motion video. The input 8K/240 Hz video is coded for bit-rate reduction by a factor of 12 and stored on solid-state drives. The replay server offers loop-recording, live editing, and slow-motion replay functions with a dedicated remote controller. The specifications of the high-speed camera and slow-motion replay server are shown in Tables 3 and 4, respectively.

TABLE 3
Specification of high speed 8K camera

Optical size	1.25-in
Image sensor	33-Mpixel CMOS
Active pixel count	7 680(H) × 4320(V)
Frame frequency	239.76 (240/1.001), 119.88 (120/1.001), 59.94 (60/1.001) Hz
Colour gamut	Rec. ITU-R BT.2020
HDR system	Hybrid Log-Gamma as per Rec. ITU-R BT.2100
Transmission between head and CCU	Head to CCU: 448 Gbit/s (1310 nm) CCU to Head: 40 Gbit/s (1550 nm)
Signal output	U-SDI × 2ch (RGB 4:4:4) as per Rec. ITU-R BT.2077 Part 2
Sensitivity	F2.8 / 2 000 lux
S/N	50 dB

TABLE 3 (*end*)

Dynamic range	450%
MTF	12% at 3 200 TVL
Size	Head: 162 (W) × 206 (H) × 353 (D) mm CCU: 8RU (19-in rack size)
Weight	Head: 18 kg, CCU: 37 kg
Power consumption	Head: 300 W, CCU: 750 W

TABLE 4

Specification of slow-motion replay server

Video format	Input: 8K/240 Hz RGB 4:4:4 Output: 8K/60 Hz YC _B C _R 4:2:2
Signal input	U-SDI × 2
Signal output	12G-SDI × 4
Compression ratio	1/12 (intraframe compression)
Recording media	12 TB SSD
Recording time	240 min
Replay latency	Less than 10 frames
Size	5RU (19-in. rack size)
Functions	Instant replay, loop recording, variable replay speed, live editing
Power consumption	800 W

1.2 Studio Monitor

A 31.5-inch, 8K/120Hz monitor was developed for the studio production. The appearance is shown in Fig. 24(f) and the specifications are given in Table 5. The monitor has a dynamic range of more than 100,000:1 and maximum brightness of more than 1000 cd/m² by controlling the backlight with local dimming.

TABLE 5

Specification of studio monitor

Screen size	31.5-in
Number of pixels	7 680 × 4 320
Frame frequency	119.88 (120/1.001) Hz
Bit depth	12-bit equivalent
EOTF	HLG, PQ, SDR
Maximum brightness	1 000 cd/m ²
Dynamic range	100,000:1
Coverage of colour gamut	85% (Rec. ITU-R BT.2020)

1.3 Video/audio encoder and decoder

An 8K/120 Hz real-time video codec using high-efficiency video coding (HEVC) has been developed that complies with 8K broadcasting standards in Japan (see Annex 4). The encoder processes 8K/120 Hz video by 12 4K/60 Hz sub-encoders (four and three partitions in the spatial and temporal directions, respectively). Video and audio decoders were implemented in a PC using highly parallelised software. The codec has achieved high image quality with a bit rate of 85 Mbit/s, which is equivalent to the bit rate of 8K/60 Hz broadcasting, and can run at higher bit rates of up to 400 Mbit/s. The appearance of the 8K/120 Hz codec is shown in Fig. 24(g) and (h), and the specifications of the encoder and the decoder are presented in Tables 6 and 7, respectively.

TABLE 6
Specification of 8K/120 Hz video/audio encoder

Video	Coding scheme	MPEG-H HEVC ITU-T H.265 Main 10 profile (4:2:0 / 10 bits), Level 6.2
	Format	7 680 × 4 320/119.88 Hz (120/1.001 Hz)
Audio	Coding scheme	MPEG-4 AAC Low complexity profile
	Format	22.2 ch / 24 bits / 48 kHz
Multiplexing	Multiplexing scheme	MPEG-2 TS SMPTE ST 2022-2

TABLE 7
Specification of decoder

CPU	2 × Intel Xeon E5-2669 v4 (2.20 GHz, total 44 cores)
Main memory	DDR4-2400 ECC Registered 128 GB
Graphics board	2 × NVIDIA Quadro P6000

1.4 Sheet-type flexible OLED display

An 8K/120 Hz sheet-type flexible OLED display was developed for the easy enjoyment of ultra-high-definition, highly realistic images at home. The appearance of the display is shown in Fig. 24(l) and the specifications are shown in Table 8.

TABLE 8
Specification of 8K/120 Hz sheet-type OLED display

Screen size	88-in
Number of pixels	7 680 × 4 280
Frame frequency	119.88 (120/1.001) Hz

2 Live production and transmission

In the live production and transmission of 8K/120 Hz programmes, low-latency, mezzanine compression IP transmission was used to transmit audio and video from the live venue to the production centre, where the transmitted video was edited in real time by online equipment and 22.2-channel audio mixing was performed. The appearance and specifications of the IP transmission are demonstrated in Fig. 24(c) and Table 9, respectively. The completed 8K/120 Hz content was coded using the HEVC codec at 250 Mbit/s for transmission via 21-GHz band satellite. The wideband modulator and demodulator (Fig. 24(i) and (j)) have been developed for 21-GHz band satellite transmission. The BSAT-4a satellite, which is equipped with a 21-GHz-broadband, high-capacity transponder was used. The specifications of the satellite transmission are shown in Table 10. The live video transmitted from the venue was presented on the 88-inch sheet-type OLED display and the 22.2-channel sound was reproduced by a binaural reproduction system using line-array speakers.

TABLE 9
Specification of IP transmission

Video	Coding scheme	TICO (Tiny Codec)
	Format	$7\,680 \times 4\,320/119.88$ (120/1.001) Hz
	Compression ratio	1/6
	Transmission bandwidth	24 Gbit/s
	Error correction	FEC (Exclusive OR)
	Interface	4×10 GbE
Audio	Format	$2 \times 64\text{ch} / 24 \text{ bits} / 48 \text{ kHz}$ (bidirectional)
	Transmission bandwidth	250 Mbit/s
	IP protocol	AES67
	Interface	Gigabit Ethernet

TABLE 10
Specification of 21-GHz satellite transmission

Satellite	BSAT-4a
Uplink frequency	18.24375 GHz
Downlink frequency	21.84375 GHz
Diameter of uplink and downlink antennas	1.5 m
Bandwidth of satellite transponder	300 MHz (21.7-22.0 GHz)
Symbol rate	250 Mbaud
Roll-off factor	0.1
Modulation scheme	QPSK
Error correction (code rate)	LDPC (2/3)
Transmission capacity	327.726 Mbit/s
Interface	Gigabit Ethernet

Annex 4

Practice of 8K UHD programmes broadcasting on large screens in China

1 Overview

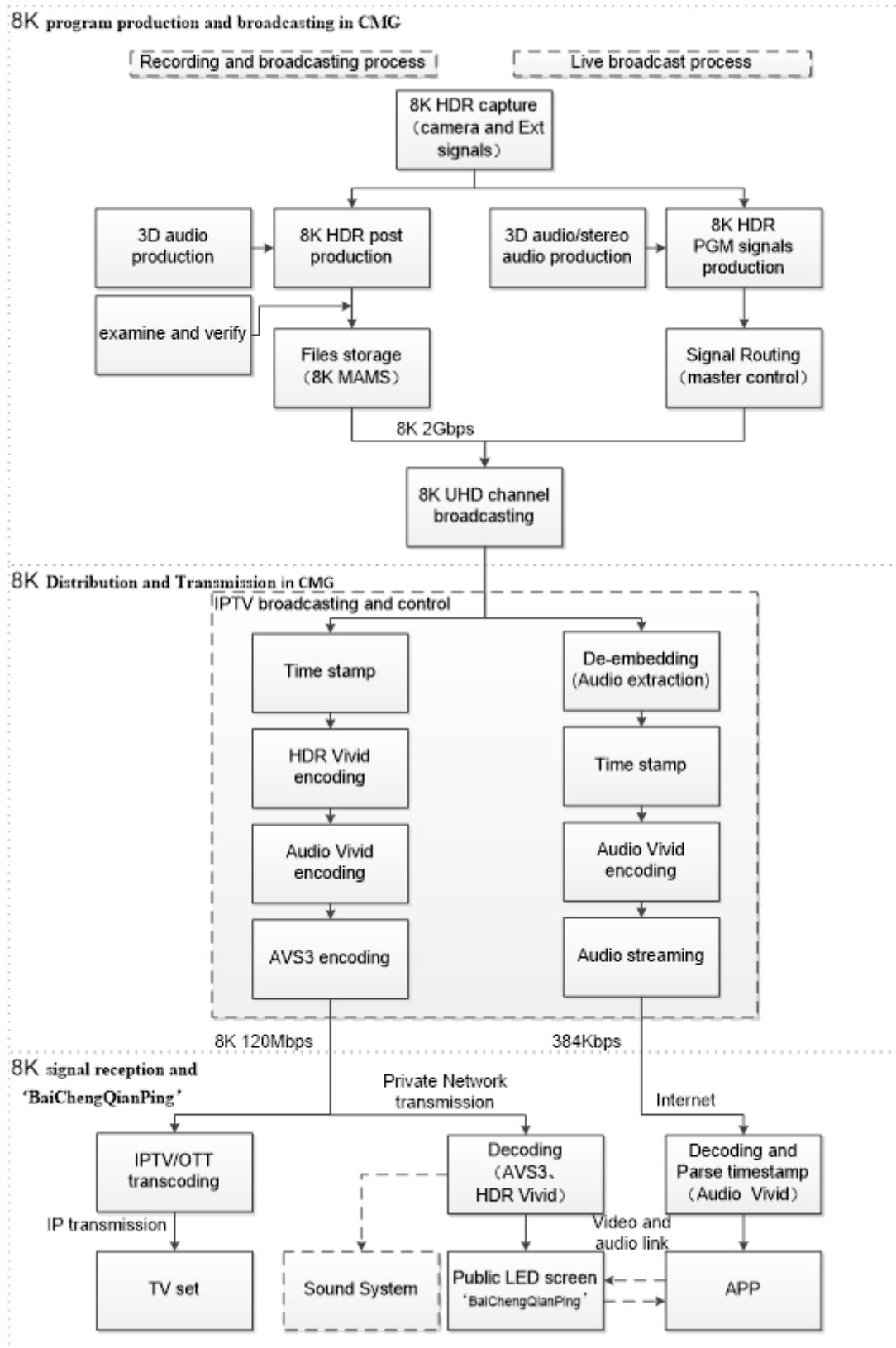
In October 2021, China Media Group (CMG) began to launch a three-year “Bai Cheng Qian Ping” (100 cities, 1,000 screens) UHD video and 3D audio promotion activities, cooperating with IT, radio and television, culture, tourism, transportation and other industries. In January 2022, CMG launched a specific 8K UHD channel to provide 8K programmes for the event. The campaign mainly showcased the UHD video and 3D audio-visual effects through the outdoor and indoor UHD large screens in public, using multicast technology to distribute CMG 8K programmes over IP to more than 70 cities, realizing the display of more than 310 UHD large screens for the Beijing Winter Olympics.

2 8K UHD video and 3D audio broadcasting and distribution system workflow

By promoting industrialization in 2022, China has now fully built the end-to-end all-elements 8K production and broadcast chain, and realized the technical support of the whole workflow of 8K content from capture, recording, production, signal routing, broadcasting, encoding and transmission to terminal displays. The 8K UHD signal broadcasting-distribution workflow is shown in Fig. 26.

FIGURE 26

8K UHD signal broadcasting-distribution workflow



3 8K programme production and broadcasting in CMG

3.1 Basic technical specifications for 8K UHD programme video and audio

Basic technical specifications for 8K UHD programme video and audio refers to Table 11.

TABLE 11

Basic technical specifications for 8K UHD programme video and audio

No.	Item	Technical specification
1	Picture aspect ratio	16:9
2	Pixel count Horizontal × vertical	7 680 × 4 320
3	Sampling lattice	Orthogonal
4	Pixel aspect ratio	1:1 (square pixels)
5	Pixel addressing	Pixel ordering in each row is from left to right, and rows are ordered from top to bottom.
6	Frame frequency	50 Hz
7	Scan mode	Progressive
8	Quantization levels	10 bits
9	Color gamut	Rec. ITU-R BT.2020
10	HDR	HLG 1 000 cd/m ²
11	Sampling	4:2:2
12	Audio encoding	PCM 24 bit
13	Audio sampling frequency	48 kHz
14	Channel	Supports 16-channel PCM audio
15	File format	MXF OP-1a
18	8K video coding format	JPEG XS High 4:4:4 12 Profile as defined by ISO/IEC 21122-2
19	8K video encoding bitrate	2 073.6 Mbit/s

3.2 8K UHD programme video and audio production requirements and specifications

3.2.1 8K UHD programme video and audio production requirements

The 8K UHD studio/ outside broadcasting van generally uses IP system to produce 8K UHD signals, realizing the whole workflow of 8K UHD production functions from signal source, graphics, subtitle, real-time monitoring, to program recording and time-delayed broadcasting. The external signal in the studio comes from the 8K HDR signal provided by the master control, or the 4K signal provided by the master control and other 4K signal sources are used after up-conversion in the studio.

Post-production includes functions such as 8K signal recording and storage, editing and production, colour grading and rendering. In CMG, 8K editing production integrated network system has the ability to co-product 8K HDR programme, with automated colour management, 8K multi-format mixing, high-performance 8K real-time editing, 3D audio and surround sound production, etc. realizing 8K multi-format high-code mixing and real-time output with different 8K encoding and different dynamic range. 8K colour grading and rendering has the function of colour reproduction

correction for the characteristics of different types of 8K materials. The 8K UHD programme files from the post-production system are reviewed and stored in the MAMS for programme broadcasting, or directly delivered to 8K UHD TV broadcasting system.

The audio production uses surround sound/3D audio and achieves PCM audio signal transmission of surround sound and 3D audio over IP in the live broadcast workflow. Surround sound/3D audio production picks up signals such as special effects, sound effects, commentary and hosting, and mix them together with playback sounds and other signals. With the live broadcast mixer, adjust the location, distance and other information of sound elements to create PGM signals, and transmit them to the master control with 8K video.

After surround sound/3D audio are recorded in the studio or outfield, they are pushed to the video and audio post-production platform in the form of files for post-production. The platform is based on the digital audio workstation, and is used to adjust the frequency, dynamics, location, distance and other information of sound elements, create the surround sound/3D audio files, which are pushed to the MAMS or broadcasting system with the video files.

3.2.2 8K UHD programme production basic specifications

3.2.2.1 Basic technical specifications of 8K video signals and files

Basic technical specifications of 8K UHD signal and programme files refers to Table 12.

TABLE 12

Basic technical specifications of 8K video signals and files

No.	Item	Channel
1	Video format	7 680 × 4 320/50/P
2	Sampling	4:2:2
3	Bit depth	10 bit
4	Coding format	JPEG XS High 444 12 Profile
5	Horizontal wavelet transform layers	5
6	Vertical wavelet transform layers	2
7	Dequantization method	Uniform quantizer, Qp _{ih} = 1
8	Level	8k-2
9	Sub Level	Sublev3 bpp
10	Coding method	Intra-frame encoding with a fixed frame length
11	Compression ratio	16:1
12	Bitrate	2 073.6 Mbit/s

3.2.2.2 Audio production system basic requirements

3.2.2.2.1 Surround sound production

Surround sound audio production technical specifications, audio level and loudness specifications and channel allocation of surround sound refer to Table 13.

TABLE 13

Technical specifications of surround sound production

No.	Item	Technical specifications
1	Format	PCM
2	Bit depth	24 bits
3	Sampling frequency	48 kHz

Audio level and loudness specifications

No.	Item	Technical specifications
1	Maximum true peak level	−2 dBTP
2	Integrated loudness	−24 LKFS
3	Variations of loudness	±2 LU

Channel distribution of surround sound

No.	Track	Channel
1	Track 1	Left (L)
2	Track 2	Right (R)
3	Track 3	Centre (C)
4	Track 4	Low frequency effects (Lfe)
5	Track 5	Left surround (Ls)
6	Track 6	Right surround (Rs)
7	Track 7	Stereo left (L)
8	Track 8	Stereo right (R)

3.2.2.2.2 3D audio production

In 3D audio production, the audio level and loudness requirements can be referred to surround sound-related specifications, the 3D audio channel allocation reference Table 14.

TABLE 14

Channel distribution of 3D audio

No.	Track	Channel
1	Track 1	Left (L)
2	Track 2	Right (R)
3	Track 3	Centre (C)
4	Track 4	Low frequency effects (Lfe)
5	Track 5	Left surround (Ls)
6	Track 6	Right surround (Rs)
7	Track 7	Stereo left (L)
8	Track 8	Stereo right (R)
9	Track 9	Left top front (Ltf)

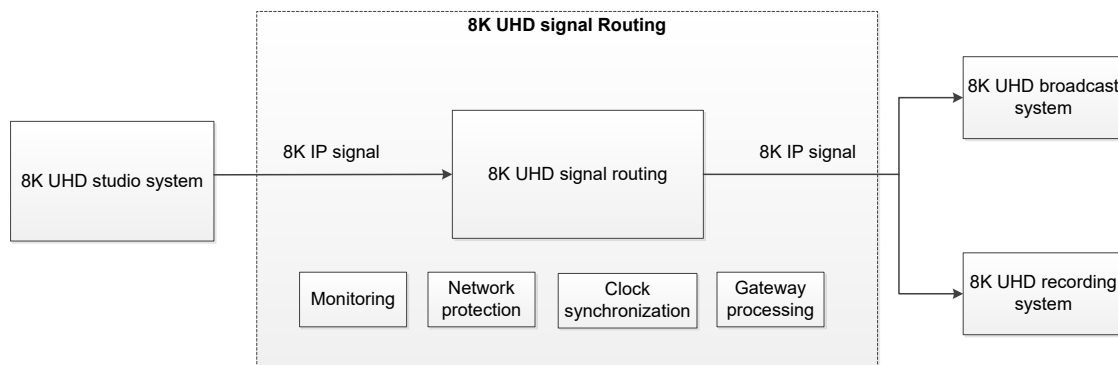
TABLE 14 (*end*)

No.	Track	Channel
10	Track 10	Right top front (Rtf)
11	Track 11	Left top rear (Ltr)
12	Track 12	Right top rear (Rtr)

3.3 8K UHD signal routing system

The 8K UHD signal routing system is mainly responsible for the routing of 8K UHD signals exchange among the studio, the recording system and the broadcast system. The system adopts an all-IP system architecture to realize the routing, recording, broadcasting, transmission and distribution of 8K UHD signals, as shown in Fig. 27.

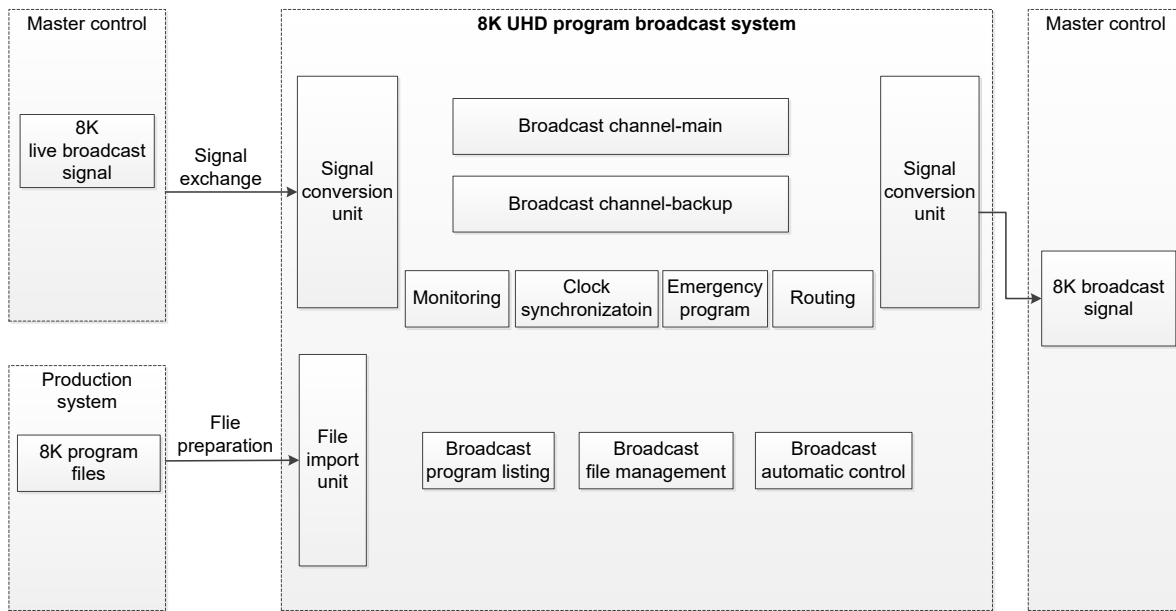
FIGURE 27
8K UHD signal routing system



3.4 8K UHD broadcast system

The 8K UHD broadcast system takes the broadcast video server as the core, realizing the functions of live broadcast signal routing, broadcast file decoding and playback, signal switching, monitoring, and emergency programme. The broadcast system is logically divided into several processing units, among which the 8K UHD video and audio signals conversion unit adopts IP technology. The file import unit migrates the 8K programme files from the production system to the broadcast video server for storage, and completes the decoding and broadcasting of the files according to the pre-programmed programme playlist, as shown in Fig. 28.

FIGURE 28

8K UHD programme broadcast system**4 8K distribution and transmission**

The 8K UHD broadcast signal from the 8K UHD TV broadcast system is sent to the IPTV broadcasting control platform, which carries out HDR Vivid encoding, Audio Vivid encoding and AVS3 encoding, then distributes and transmits the encoded 8K 120 Mbit/s IP stream.

4.1 Distribution and transmission of video and audio over private network

The video and audio encoding system in the IPTV broadcasting control platform provides timestamp embedding and HDR Vivid and Audio Vivid metadata are extracted by the AVS3 encoder. After inserting the timestamp through the encoder, the 8K signal uses the PES packet of TS stream to carry and transmit the timestamp data for extraction with decoder system. Meanwhile, the encoder extracts the HDR Vivid metadata and Audio Vivid metadata from the uncompressed live streams, and encodes by AVS3 and outputs the live streams. 8K Streaming Service routes and distributes the received IP streams and pushes them to the private network for transmission.

4.2 Distribution and transmission of audio over Internet

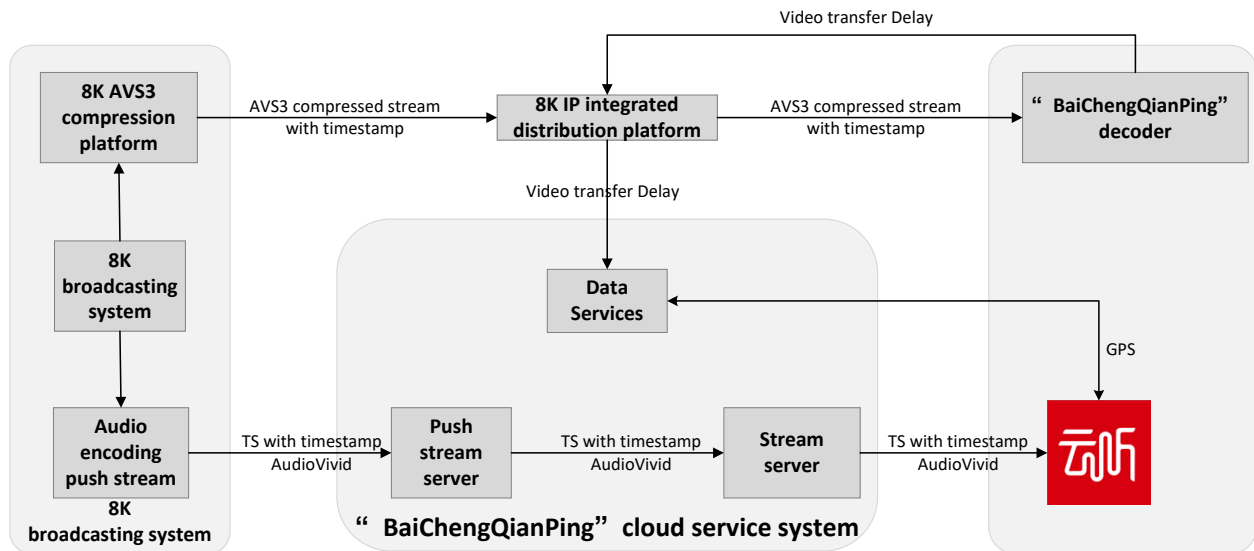
The Audio Vivid encoder system in IPTV broadcast control platform provides metadata extraction, timestamp embedding and encoding distribution functions for Audio Vivid distributed over Internet. The 8K signal is de-embedded by the Audio Vivid encoder, which then extracts the pulse code modulation (PCM) audio signal for Audio Vivid encoding. Audio Vivid encoder system use the bitrate of 384 kbit/s and insert a timestamp, which is carried and transmitted by the PES packet of TS stream, and the Audio Vivid stream is packaged and pushed to the internet and transmitted to the "EtherealSound" mobile APP. The video and audio are synchronized using the timestamp extracted by the mobile receiver.

4.3 Video and audio synchronization in heterogeneous network (HetNet)

Since the 8K TV accompaniment cannot be played through high-power speakers in the public square, a HetNet AV synchronization technology was developed. The audience are able to hear the TV accompaniment sound while watching 8K video through the public large screen. Automatic end-to-

end delay detection is achieved by inserting timestamp into the stream information in the chain, and Audio Vivid with a timestamp is sent to the “EtherealSound” mobile APP Through the delay time measurement between large screen and mobile terminal, audio distribution delay control, and mobile decoding and rendering delay dynamic compensation technology, the audio and video delay of different network transmission is controlled within $-90\text{ ms} \sim +120\text{ ms}$, which meets the requirements of the Chinese Standards “Relative timing of sound and vision for broadcasting” (GB/T 22150-2008). The video and audio synchronization workflow in HetNet is shown in Fig. 29.

FIGURE 29
Video and audio synchronization workflow in HetNet



5 8K signal reception and display

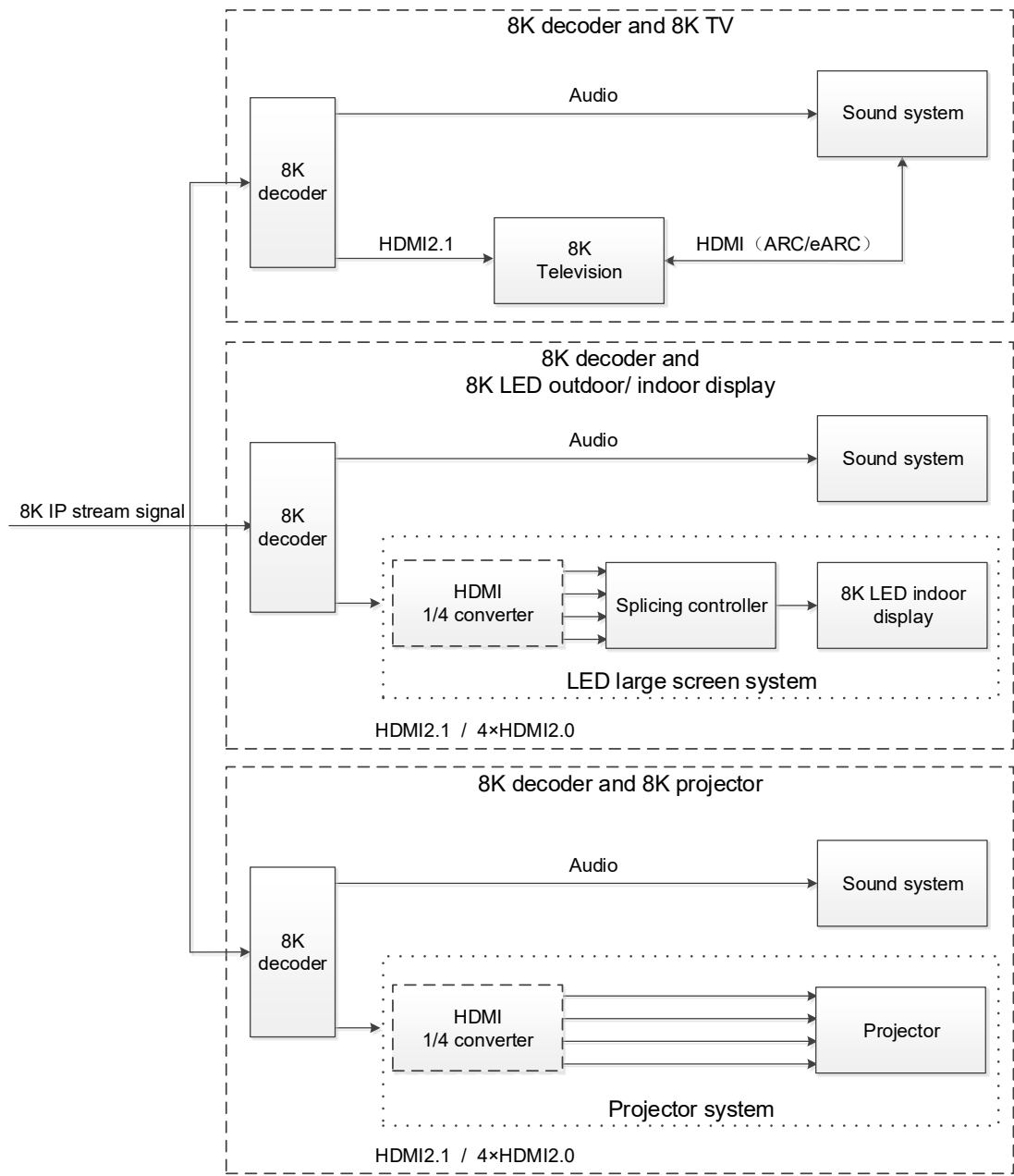
5.1 Signal reception and display of “BaiChengQianPing”

Through the construction or upgrading of 4K/8K UHD public large screens, the built of operation management platforms, routing centres, we realize real-time routing and monitoring and broadcast control of national large screens.

5.2 System architecture

After AVS3 decoding in terminal, the video and audio can be presented through 8K TV, 8K LED outdoor/indoor screens, 8K projector and sound system. The system architecture is shown in Fig. 30.

FIGURE 30
System architecture



5.3 Reception and display technical specifications

5.3.1 Decoder specifications

8K professional decoder should have dual-link access and Network Time Protocol (NTP) client, support NTP protocol. It has the functions of remote control, operation data feedback, security authentication, remote upgrade of systems and applications, and log management. The 8K decoder specifications refer to Table 15.

TABLE 15
8K decoder specifications

No.	Item		Technical specifications
	Type of input transport stream		Transport streams conformed by GB/T 17975.1-2010
	Video streaming protocol		UDP and RTP transmission methods specified in T/UWA012.2-2022
	Input transport stream packet length		Input 188-byte packet length and 204-byte packet length
	File format		TS file
	Single-channel video decoding	Profile and level	AVS3 basic profile 10bit (profile_id value 0 × 22), level 10.0.60
		Frame frequency	50 Hz (Progressive)
			100 Hz (Progressive)
		Picture aspect ratio	16:9
		Sampling	4:2:0
		Quantization levels	10 bits
			12 bits
		Pixel count Horizontal × vertical	7 680 × 4 320 maximum
		Colour gamut	Rec. ITU-R BT.2020
		HDR	Rec. ITU-R BT.2100
HDR Vivid			
	Audio output		Stereo audio
			5.1 Surround sound encoding signal passthrough and output
	Audio interface		HDMI, fiber S/PDIF; L/R RCA
	Video interface		HDMI2.1
			HDMI2.0 × 4, SQD

5.3.2 LED Screen specifications

At present, there are about 40 cities in China which have UHD large LED screens larger than 200 m². Take Shenzhen COCO Park as an example, the screen resolution is 7 680 × 4 416 pixels, the dot pitch is P2.5, the maximum brightness is 5 500 cd/m², and the display supports HDR and WCG. The present of UHD large LED screen in Shenzhen COCO Park are shown in Figs 31 and 32, the multi-city present effect is shown in Fig. 33.

FIGURE 31

Shenzhen COCO Park in the day



FIGURE 32

Shenzhen COCO Park at night



FIGURE 33
Multi-city present effect



The LED screen should follow the technical specifications in Table 16.

TABLE 16
LED display specifications

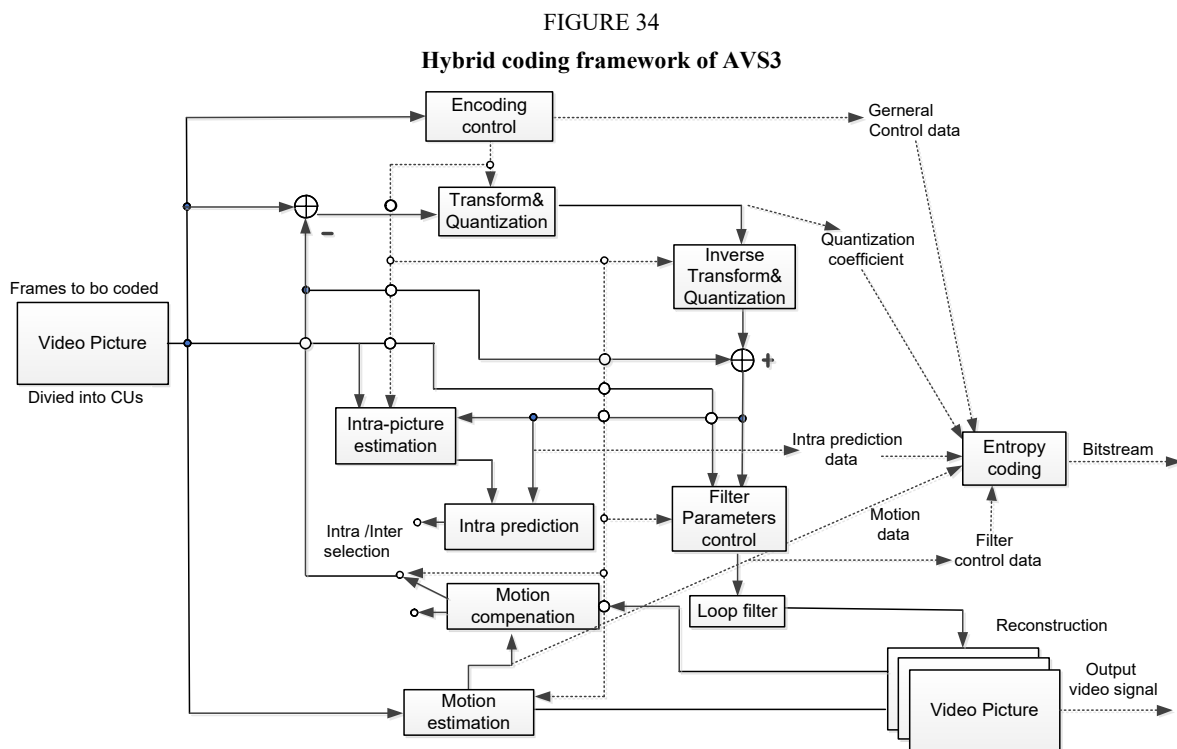
No.	Item		Technical specifications
	Physical resolution		$\geq 7\,680 \times 4\,320(8K)$
	Maximal luminance	white lamp	$\geq 5\,000\text{ cd/m}^2$
		black lamp	$\geq 3\,500\text{ cd/m}^2$
	Viewing angle of luminance	horizontal	$\geq 140^\circ$
		vertical	$\geq 120^\circ$
	Viewing angle of chromaticity	horizontal	$\geq 150^\circ$
		vertical	
	Contrast		$\geq 5\,000:1$
	Large area uniformity		$\geq 95\%$
	Area coverage of the Rec. ITU-R BT.2020 gamut		$\geq 78\%$ (CIE 1931)
	Display frame rate		$\geq 50\text{ Hz}$
	Single lamp refresh rate		$\geq 3\,840\text{ Hz}$
	Bit depth supported by input		10 bits
	Colour temperature		adjustable from 3 200 K~9 300 K

6 Main technology

6.1 AVS3

AVS3 (The 3rd generation of the Audio Video coding Standard, T/AI 109.2-2021: “Intelligent Media Coding – Part 2: Video”) is the third-generation video coding standard for 8K and 5G industrial applications in China. It included base profile finished in March 2019 and advanced profile finished in June 2021. In 2022, A001 (r20) published by DVB and TS 101 154 (v2.8.1) published by ETSI (European Telecommunications Standardization Association) incorporated AVS3 codec into the coding specification part.

AVS3 uses hybrid coding framework of block-based prediction coding and transform coding, which includes block division, intra prediction, inter prediction, transform quantization, entropy coding, loop filtering, etc. Figure 34 shows the hybrid coding framework of AVS3.



6.2 HDR VIVID

HDR Vivid is an HDR video technology standard recommended by UHD World Association (UWA) (<http://theuwa.com/>), which released its core part: High Dynamic Range Video Technology Part 1: Metadata and Tone mapping in September 2020, and has released seven companion standards in the past two years.

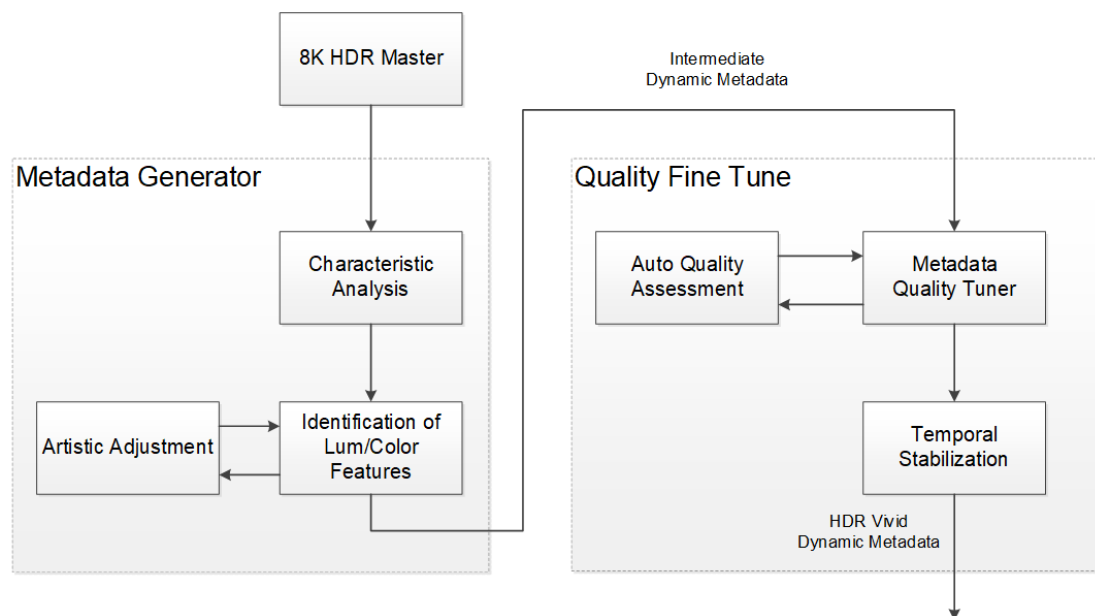
HDR Vivid technology is an end-to-end system solution. Based on human vision perception, it defines dynamic metadata in HDR video content, thus be adaptable to various applications.

6.2.1 Introduction to HDR Vivid encoding

The dynamic metadata assures the consistency across different terminal displays, and allows creators to incorporate their preferred artistic style. Metadata generator module first analyses the characteristics of the HDR master content, including image brightness features, histogram distribution, etc. Based on the characteristics of different scenes, multiple luminance colour feature

points are identified. At this point, the content creator can adjust these feature points according to their preference. After the feature points mentioned above are determined, the metadata quality tuning module can evaluate and adjust the quality of the dynamic metadata generated by the algorithm without human intervention. Finally, the temporal stabilizer module prevents video flickering that might be introduced by dynamic metadata. The HDR Vivid encoding workflow is shown in Fig. 35.

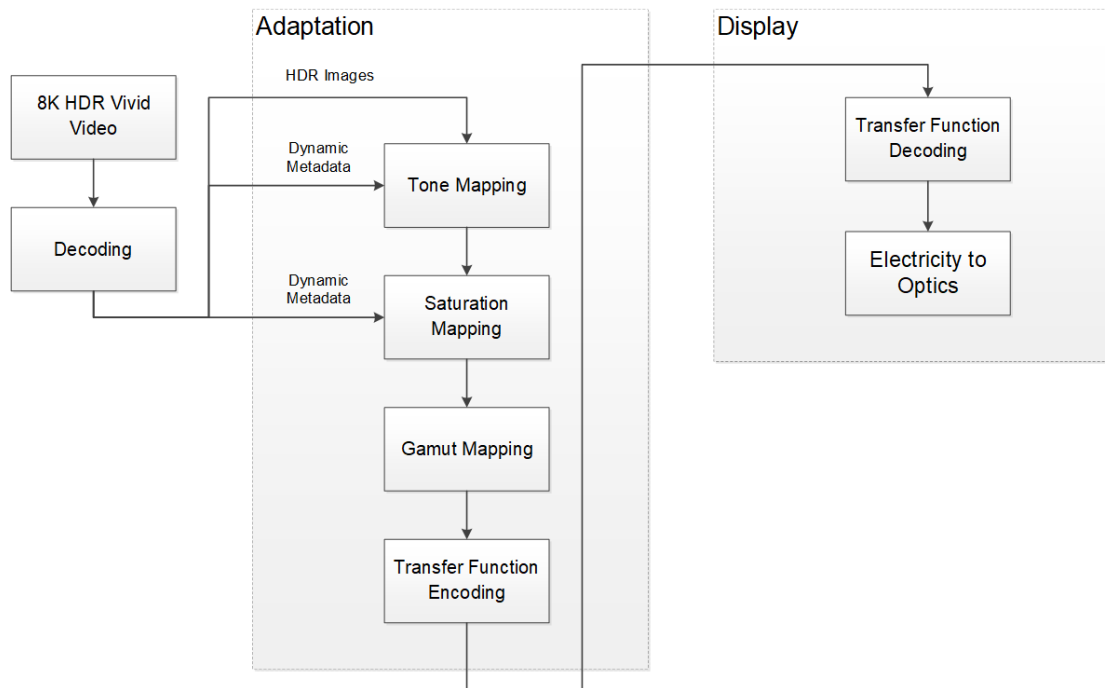
FIGURE 35

HDR Vivid encoding workflow**6.2.2 HDR Vivid decoding**

The HDR Vivid terminal first decodes the AVS3 video stream. Then, the HDR Vivid image is rendered with luminance and colour gamut for display according to dynamic metadata. This adaptation process is carried out through Tone Mapping and Saturation Mapping to maximize the original artistic intention of HDR Vivid content on various terminals. The final display process is achieved through transfer function decoding and electro-optical conversion to render HDR Vivid content. The HDR Vivid decoding workflow is shown in Fig. 36.

FIGURE 36

HDR Vivid decoding workflow



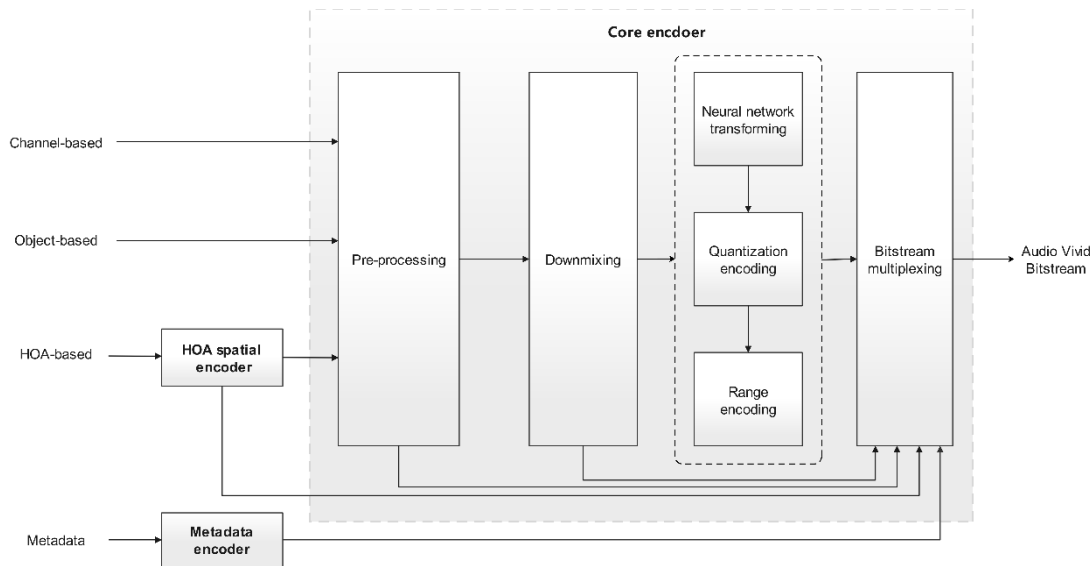
6.3 Audio Vivid

Audio Vivid is a surround/3D audio Codec and rendering technology developed by China for the “BaiChengQianPing” platform, and is a high quality, end-to-end audio production and broadcast core processing technology solution based on AI technology. It was standardized by a joint effort of UHD World Association (UWA) and Audio Video Coding Standard Workgroup of China (AVS), and was first published in April 2022. Audio Vivid adopts a novel hybrid AI Codec architecture that extracts abstract features and increases entropy coding efficiency through deep learning. It also incorporates Higher-Order Ambisonics (HOA) spatial coding technology, makes novel 3D audio metadata structure and definitions, supports multi-speaker rendering and binaural rendering. Audio Vivid brings enhanced sense of sound spatialization and localization, fulfils customer’s demand on high fidelity, high immersive sound experience while enabling personalization, interaction ability unique to Next Generation Audio (NGA) system. Audio Vivid encoder and distribution/transmission server can support package of various protocols and multiple source formats.

6.3.1 Introduction to Audio Vivid

The Audio Vivid encoder (see Fig. 37) comprises a lossy audio encoding tool, a lossless audio encoding tool and a metadata encoding tool. It supports encoding of channel signal, object signal, HOA signal and metadata. Particularly, the lossy audio encoding tool consists of a HOA spatial encoder and a core encoder. The core encoder consists of a pre-processing module, a downmixing module and neural network-based modules for feature transform, quantization and range coding, and is able to encode channel/object signals into bitstream. The core encoder together with the HOA spatial encoder can encode HOA signals. The Audio Vivid metadata consists of two parts – basic metadata and extended metadata, where the basic metadata references to the audio definition model defined in Recommendation ITU-R BS.2076-2 (ADM) and the extended metadata are newly defined to serve for enhanced binaural rendering experience. Audio Vivid metadata can be either forward compatible and backward extendable which offers sufficient flexibility and extendibility while fulfilling the requirement of metadata global interworking.

FIGURE 37
Audio Vivid encoding workflow



6.3.2 Audio Vivid decoding and rendering

The decoding workflow of Audio Vivid is basically a reverse of encoding workflow. Audio Vivid comprises two default renderers – a multi-speaker renderer and a binaural renderer and also supports interface to 3rd party renderers. Both the default renderers support rendering/playback in content production and device and can receive channel signals, object signals, HOA signals and their combinations so to offer enhanced experience, more creative ideas and allow interaction and personalization. The default binaural renderer also supports 3DoF immersive experience.

References

- [1] GB/T 22150-2008, Relative timing of sound and vision for broadcasting.
- [2] GY/T 358-2022, Technical requirements for display adaptation metadata of high dynamic range television systems.
- [3] UWA official website: <http://theuwa.com/>.
- [4] TUWA 012.1-2022, Technical requirements of program broadcasting of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.
- [5] TUWA 012.2-2022, Coding of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”: system.
- [6] TUWA 012.3-2022, Coding of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”: video.
- [7] TUWA 012.4-2022, Technical requirements of Professional Decoder of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.
- [8] TUWA 012.5-2022, Technical requirements of the public display system (outdoor) of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.
- [9] TUWA 012.6-2022, Technical requirements of network transmission of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.
- [10] TUWA 012.7-2022, Measuring method of UHD encoder for “Bai Cheng Qian Ping”.
- [11] TUWA 012.8-2022, Measuring method of the public display system (outdoor) of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.

- [12] TUWA 012.9-2022, Technical requirements of the public display system (indoor LCD) of UHD video and audio broadcasting system for “Bai Cheng Qian Ping”.

Annex 5

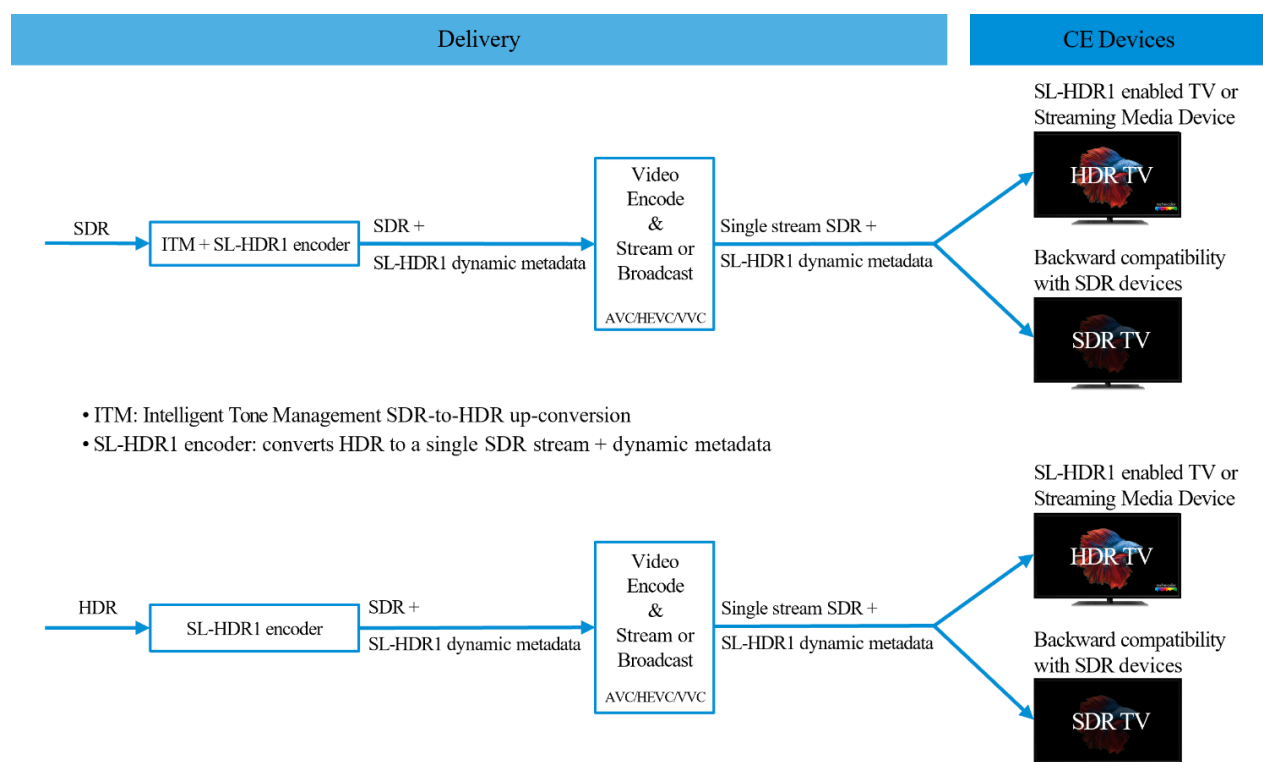
Deployment in the United States of America of NextGen TV HDR with SDR compatibility using Single-Layer HDR (SL-HDR1)

ATSC 3.0 [1] – also known as “NextGen TV” under a certification programme for consumer TVs, STBs, etc.– offers 4K ultra high-definition video quality, theatre-like sound, mobile reception (handheld and vehicular reception of broadcast television outside of the home), and innovative new features to enhance and expand the broadcast viewing experience. ATSC 3.0 enables local TV broadcasters to better personalize their programming with relevant information and interactive features, creating a more compelling and immersive viewer experience, whether it is breaking news, live sports, a drama or reality show. This broadcast technology also supports hyper-localized weather alerts and other emergency information, with targeted public announcements that are interactive and mobile. ATSC 3.0 is based on the Internet Protocol (IP), just like online video services, and uses web languages for interactivity, enabling more innovation and new services, while combining online and broadcast television.

One of the most prominent features of ATSC 3.0 broadcasting is the use of High Dynamic Range (HDR) to provide consumers with better picture quality. One of the adopted solutions for HDR is the Single Layer HDR (SL-HDR1) technology [2] [3], providing both an SDR picture and an HDR picture from the same video stream (SDR plus metadata). From a distribution perspective, the single-layer format thus provides a cost-efficient workflow that cuts bandwidth in half, to deliver all available programming in both HDR and SDR formats using a single video stream.

As of March 2025, SL-HDR1 is being transmitted in more than 50 markets in the United States of America, where more than 170 channels deliver content in HDR, 24 hours a day, seven days a week, with SL-HDR1 technology [4]. Additionally, broadcast-enabled virtual channels are deployed using SL-HDR1, where viewers with a connected TV can see and select the channel from the over-the-air service guide and the programme essence is delivered over the Internet [5]. An SL-HDR1 pre-processor converts the HDR signal into an SDR signal with metadata. The SDR with metadata signal is used for distribution and provides a high-quality SDR picture to receivers that do not have an SL-HDR1 decoder. TV receivers that do have an SL-HDR1 decoder can reconstruct an HDR signal that is perceptually equal to the HDR source [6]. When the source content is SDR, conversion –from SDR to HDR and back to SDR– can preserve the integrity of the original SDR source while also providing an excellent HDR signal, through the use of tone mapping tools that are dynamic and adjustable [7]. When the source content is HDR, and specifically in a live single master HDR production, the SL-HDR pre-processor produces a high-quality SDR down-conversion, that can automatically adapt to capture changes in lighting conditions to retain some of the details and contrast present in the original HDR image. The SL-HDR pre-processor further enables stable graphics management in the down-conversion process and a degree of compatibility with static (LUT based) conversion methods [8] [9]. Figure 38 shows a typical SL-HDR1 signal distribution chain [10].

FIGURE 38
SL-HDR1 signal distribution chain



References

- [1] Spotlight ATSC 3.0 – ATSC : NextGen TV, <https://www.atsc.org/nextgen-tv/>
- [2] ATSC A/341, “Video – HEVC”.
- [3] ETSI TS 103 433-1 V1.4.1 (2021-08), High-Performance Single Layer High Dynamic Range (HDR) System for use in Consumer Electronics devices; Part 1: Directly Standard Dynamic Range (SDR) Compatible HDR System (SL-HDR1).
- [4] Advanced HDR By Technicolor, <https://advancedhdrbytechnicolor.com/>
- [5] Sinclair Announces Virtual Carriage – Tennis Channel’s T2 | Business Wire, <https://www.businesswire.com/news/home/20240320459650/en/>
- [6] D. Touze and L. van de Kerkhof, Single-Layer HDR Video Coding with SDR Backward Compatibility, SCTE/ISBE 2017.
- [7] D. Touze et al., “HDR Production–Tone Mapping Techniques and Roundtrip Conversion Performance for Mastering With SDR and HDR Sources,” in SMPTE Motion Imaging Journal, vol. 132, no. 6, pp. 31-44, July 2023, doi: 10.5594/JMI.2023.3277092.
- [8] D. Touze et al., “HDR-SDR conversion: Live HDR Single Master Production Conversion Interoperability Challenges,” in Proceedings of the 2024 NAB Broadcast Engineering and Information Technology Conference.
- [9] D. Touze et al., “Enhancing Live Event Production With SDR/HDR Conversion Compatibility and Stable Graphics Management: A Metadata-Driven Approach,” in SMPTE Motion Imaging Journal, vol. 133, no. 6, pp. 22-34, Oct. 2024, doi: 10.5594/JMI.2024/DDLP8310.
- [10] B. Redmann et al. “HDR Challenges and Solutions,” in SMPTE Motion Imaging Journal, vol. 132, no. 10, pp. 27-42, Nov.-Dec. 2023, doi: 10.5594/JMI.2023.3325513.

Annex 6

Deployment of NextGen TV in the USA using PQ HDR (with SMPTE ST 2094-10, SMPTE ST 2094-40, HDR10 metadata), Immersive Audio and a new interactive HTML5 application layer and regionalized data

NBCUniversal currently uses dynamic and static metadata tone mapping systems (SMPTE ST 2094-10, SMPTE ST 2094-40 and HDR10) in order to support the standards provided by individual television manufacturers TVs.

The AC-4 audio codec provides advanced capabilities for immersive audio, which can be object or channel-based with some additional level of interactive control over audio parameters like dialog muting, selection of alternate dialog (language, audio description, etc.), and level.

NextGen TV's application layer allowed NBCUniversal to create an interactive "app" which provide users the ability to access regionalized information(weather) or to select additional streaming sources from the internet (NBCU's Peacock streaming service). Application layer rendered graphics can be composited over a video channels content using NextGen TV's HTML5 "canvas" which is hoped will ultimately natively support HDR.

Work is under way to analyze and improve picture quality to make NextGen TV a differentiator over current SDR-BT.709 HDTV systems. The components of this effort include the improved use of statmux and evaluation of new station manager systems that automate the process of changing encoder modes dynamically (per a broadcast schedule).

Efforts in the improvement of accessibility include the enhancement of audio description (AD) to leverage the already available immersive surround and place the describer on top of the main soundtrack. This method includes system generated automatic "ducking" and is potentially much more efficient than the use of an additional complete main soundtrack to support the AD.

Example screenshots from the current use of the NextGen TV (ATSC 3.0) application layer by NBCUniversal where regionalised, contextual information is shown in Figs 39, 40, 41 and 42.

FIGURE 39

NextGen TV UHD broadcast HTML5 Rendered Interactive Menu System

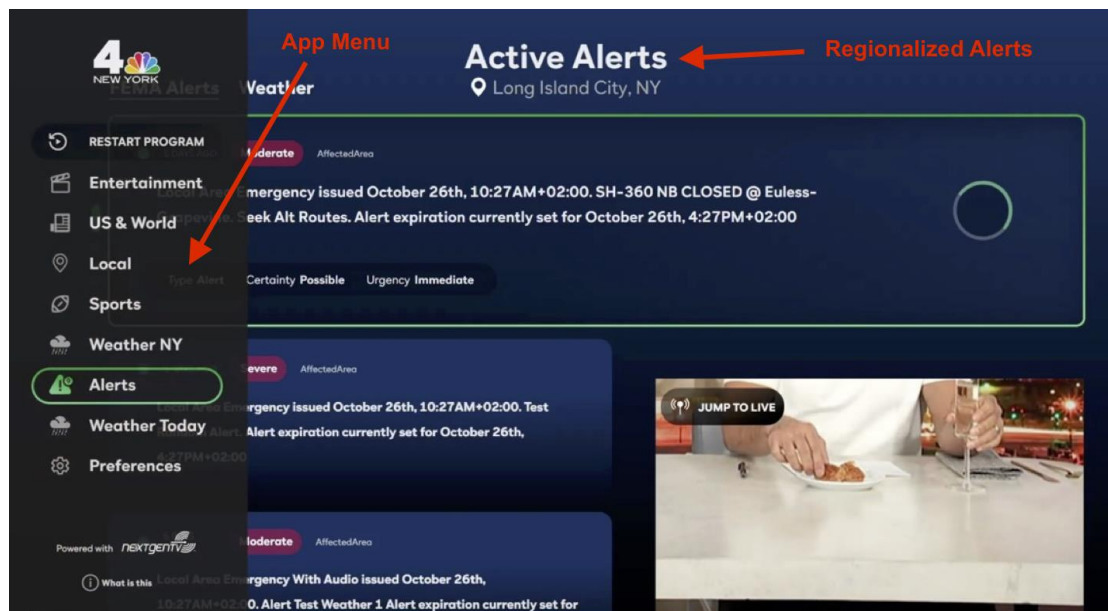


FIGURE 40

NextGen TV UHD broadcast HTML5 Rendered Regionalized Weather (via Internet)



FIGURE 41

NextGen TV UHD broadcast NextGen HTML5 Rendered Localized Weather Display (Detailed)

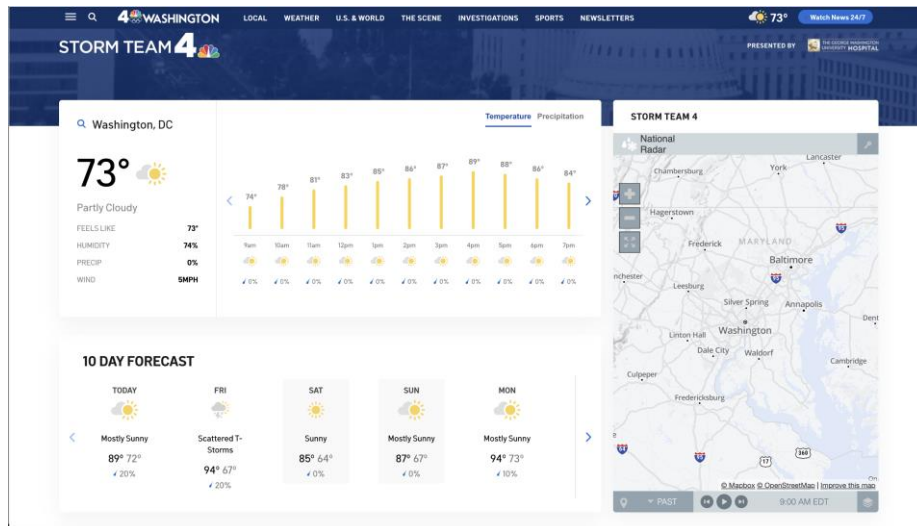
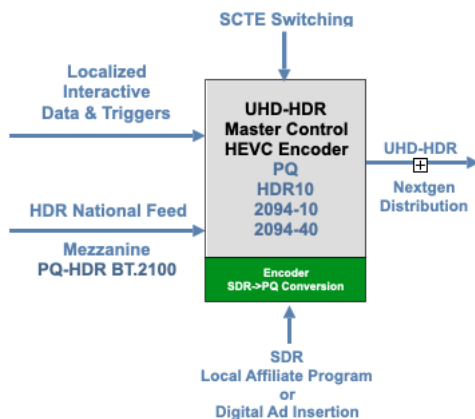


FIGURE 42
NBCUniversal NextGen TV broadcast encoder video and data paths



Annex 7

Examples of single-master HDR production with SDR-focused workflow

Single-master workflows for live HDR production are documented in Report ITU-R BT.2408 – Suggested Guidance for operational practices in HDR television production. The SDR-focused workflow has been widely used by the BBC and Sky UK for their UHD-HDR live productions. It enables simultaneous production of HDR and SDR signals that conform to ITU-R Recommendations for international programme exchange. It has been found to facilitate easy migration to HDR production and to be efficient in terms of staff education, hardware, and complexity of infrastructure, whilst preserving SDR images. It uses HLG as the core HDR format, with a fixed down-mapper to generate an SDR output, therefore no metadata is required.

For major events, a host broadcaster will typically provide a “world feed” for international programme exchange. Downstream rights holding broadcasters incorporate the world feed into their own local production to produce a complete programme tailored to their audience, for example by adding studio discussions, competitor interviews and branding graphics. The host production and downstream production may both be single-master HDR productions. Where the host production is a single-master HDR production, it is essential for the downstream broadcasters to have knowledge of the video processing and conversions applied by the host broadcaster, to ensure that an intended “look” is maintained during possible round-trip conversions from SDR to HDR to SDR.

Table 17 lists examples of single-master HDR production with SDR-focused workflow being used by a host broadcaster. Table 18 lists examples of a single-master HDR production with SDR-focused workflow being used in a downstream production.

In all of these examples, critical SDR monitors that were part of the SDR-focused production were set to have a peak luminance of 100 cd/m².

TABLE 17

Examples of single-master HDR production with SDR-focused workflow used by a host broadcaster

Host Broadcaster	Genre	Details
Sky Sports	Sport: Football	English Premier League English and Welsh Internationals English Football League English Football League Cup English Football League Trophy English Women's Super League Scottish Premier League Scottish Women's Premier League
	Sport: Cricket (Men's and Women's)	The Ashes 2023 England vs India Test Cricket Series 2022 The Hundred The T20 Blast Twenty20 Internationals County Championship
	Sport: Boxing	Boxer Fight Nights
	Sport: Rugby League	Super League Grand Final 2022 and 2023
	Sport: Darts	World Darts Championships 2022
	Rolling news channel	Sky Sports News
Sky	Music Festival	Isle of Wight Festival
BBC	Sport: Football	FA Cup 2019
	Music Festivals	Glastonbury Festival 2022, 2023, 2024
	State Events	Queen Elizabeth II's Platinum Jubilee 2022 Coronation of King Charles III and Queen Camilla 2023

TABLE 18

Examples of Single-master HDR production with SDR-focused workflow used by a downstream broadcaster

Downstream Broadcaster	Genre	Details
Sky Sports	Sport: Formula 1	F1 race weekends and supporting events
	Sport: American Football	NFL Superbowl 2023
Sky News (Presentation)	State Event	Coronation of King Charles III and Queen Camilla 2023
BBC	Sport: Football	Euros 2020, 2024 FIFA World Cup 2022
	Sport: Tennis	Wimbledon 2023, 2024
	Sport/Entertainment	Olympics Opening and Closing Ceremonies 2024

Case Study: The Coronation of King Charles III and Queen Camilla

The Coronation of King Charles III and Queen Camilla in May 2023 provides one example of a single-master HDR production with SDR-focused workflow used by a host broadcaster which fed an independent single-master HDR production with SDR-focused workflow used by a downstream broadcaster.

In the host production, over 100 UHD HDR cameras were deployed, covering the procession between Buckingham Palace and Westminster Abbey, Wellington Barracks and inside Westminster Abbey itself. The Coronation was produced in UHD (2160p) HDR. Production in UHD HDR adds considerable complexity to such a large-scale production, but it was considered important to produce and archive such a historic event in the highest possible quality.

The coverage utilised seven outside broadcast (OB) trucks from four different providers – one covering the abbey, five covering the procession route and Wellington Barracks, and a final ‘presentation’ (Pres) truck covering the BBC studio at Canada Gate and the main BBC and international programme feeds. With different equipment being used in each of the seven outside broadcast trucks, it was essential to ensure consistent results from each production unit.

The entire production used an SDR-focused single-master UHD HDR production workflow, based on the HLG HDR format. As part of the single-master HDR production workflow, BBC R&D’s HDR to SDR format conversion ‘3D-LUTs’⁵ were used. They use a non-linear, gamma-adjusted conversion algorithm, ensuring the HDR image and the SDR image, when shown on a reference 100 cd/m² nominal peak luminance SDR display, look subjectively similar.

For such a large TV production, it was important to keep signal routing as simple as possible. Each OB truck had its own director and provided a finished UHD HDR programme for their section of the parade or abbey coverage. The Pres truck then mixed between those feeds and the BBC’s studio cameras at Canada Gate, to produce the final UHD programme output. This was then down-mapped in the Pres truck from HDR to SDR using a BBC display-light down-mapping LUT, and converted from UHD 2160p to HD 1080i, to provide the BBC HD SDR programme output.

As a confidence check, the 1080i HD SDR programme output created by the Pres truck was fed back to the other OB trucks to allow them to compare against their locally generated HD SDR signal. This

⁵ <https://www.bbc.co.uk/rd/blog/2020-06-lut-format-conversion-hdr-video-production>

also allowed the vision supervisors in each truck to ensure their pictures were consistent with those of the other OB trucks. The 1080i HD SDR ‘return’ programme feed allowed any format converter configuration errors to be quickly spotted.

Sky used the same feeds intercutting with their independent SDR-focused production to enable seamless addition of Sky branding and commentators. This understanding of the workflows and LUTs used by the host production delivered both HDR and SDR to both the BBC and Sky’s outputs without issues or compromise. Use of the HLG HDR format for the HDR distribution meant that no metadata was needed.

Annex 8

Examples of single-master dual-focused HDR productions

Single-master HDR production using a dual-focused workflow is described in clause 7.4 of Report ITU-R BT.2408 – Suggested guidance for operational practices in HDR television production. This workflow has been widely adopted in the United States by NBCUniversal, Fox Sports and Amazon Prime Video Live Sports for UHD-HDR live production and transmission, allowing simultaneous HDR and SDR signal generation that meets international program exchange standards. The dual-focused monitoring approach standardizes diffuse white levels across HDR and SDR displays, enabling side-by-side monitoring with minimal eye adaptation issues. This method familiarizes production teams and shaders with native HDR visuals, technical levels, and adjustments, facilitating a smoother transition to full HDR production.

While dual-focused shading introduces slightly more complexity in monitoring, it streamlines staffing, allowing each vision supervisor and shader to monitor HDR and SDR signals simultaneously as viewed by the primary audience. Both HDR and SDR outputs are produced with HLG as the core HDR format and a fixed down-mapper for SDR signal generation.

In these examples, the final signal to the consumer is converted from HLG to the PQ HDR format which utilizes an absolute display-referred mapping of luminance, where the goal is to deliver a consistent viewing experience for the primary audience. Additionally, PQ can support advanced static or dynamic tone mapping on streaming devices and displays to prevent simple clipping of video and instead to optimize the video to a target displays specific peak luminance capability.

Table 19 provides examples of Single-Master HDR productions using dual-focused workflows, highlighting instances where SDR monitoring was adjusted to match HDR reference white levels (203 cd/m²), in order to closely match the primary audience experience.

TABLE 19

Some Examples: United States Dual-Focused Single Master HDR/SDR Broadcasts

NBCU Sports Group	Fox Sports	Amazon Prime Video
Saturday Night Live 50 th Anniversary Radio City Music Special (2024)	Superbowl LIV, LVII, LIX	Thursday Night Football
Sunday Night Football (2024/2025)		
Paris Olympics (2024) ⁽¹⁾	NFL Football	Monday Night Hockey
Tokyo Olympics (2020) ⁽¹⁾	Major League Baseball	National Women's Soccer League
Beijing Olympics (2021) ⁽¹⁾	MLB World Series (2024)	
PGA Tour Golf	FIFA World Cup (2022)	
Notre Dame Football (2020-2024)	MLB All Star Games	
College Football	MLB Post Season Divisional, League Championship and World Series	
Kentucky Derby (2024)	Over 100 other events throughout each year	

⁽¹⁾ NBCUniversal studio events used dual-focused shading. HDR-to-SDR conversions occur via NBCU LUT3 which feeds legacy SDR infrastructure and streaming encoders.

Case Study: Sunday Night Football (NBCUniversal, LLC)

Sunday Night Football employs a dual-focused workflow where shaders, vision supervisors, and control rooms view HDR and SDR displays simultaneously. Using diffuse white to anchor the image close to the knee point, a fixed LUT is used to create dual HDR/SDR outputs from the core HDR Production.

- *Vision supervisors* can set color and initial exposure for both formats in simultaneously, effectively guiding shaders to achieve the desired artistic intent of the production.
- *Shaders* adjust HDR cameras efficiently while monitoring HDR and down-mapped SDR simultaneously so adjustments create optimal output on both paths.
- *Control room multiviews* allow for side-by-side comparison across cameras, graphics, and sources, normalized within a single display format.

Adjacent displays are adjusted to unify diffuse white on the HDR and SDR displays as described in Report ITU-R BT.2408, clause 7.4. In this example, SDR displays are set at approximately 203 cd/m².

Figures 43 and 44 describe Dual-Focused Monitoring and by extension an overview of the single-master workflow which uses HLG for production. Final transmission using PQ is also shown. This approach is used by several broadcasters in the United States of America.

FIGURE 43
Dual-focused monitoring

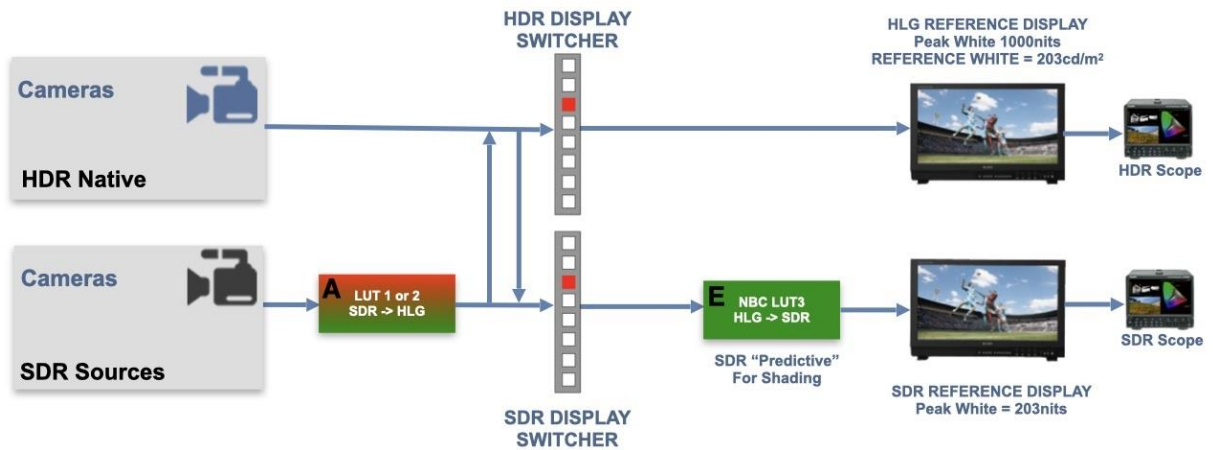
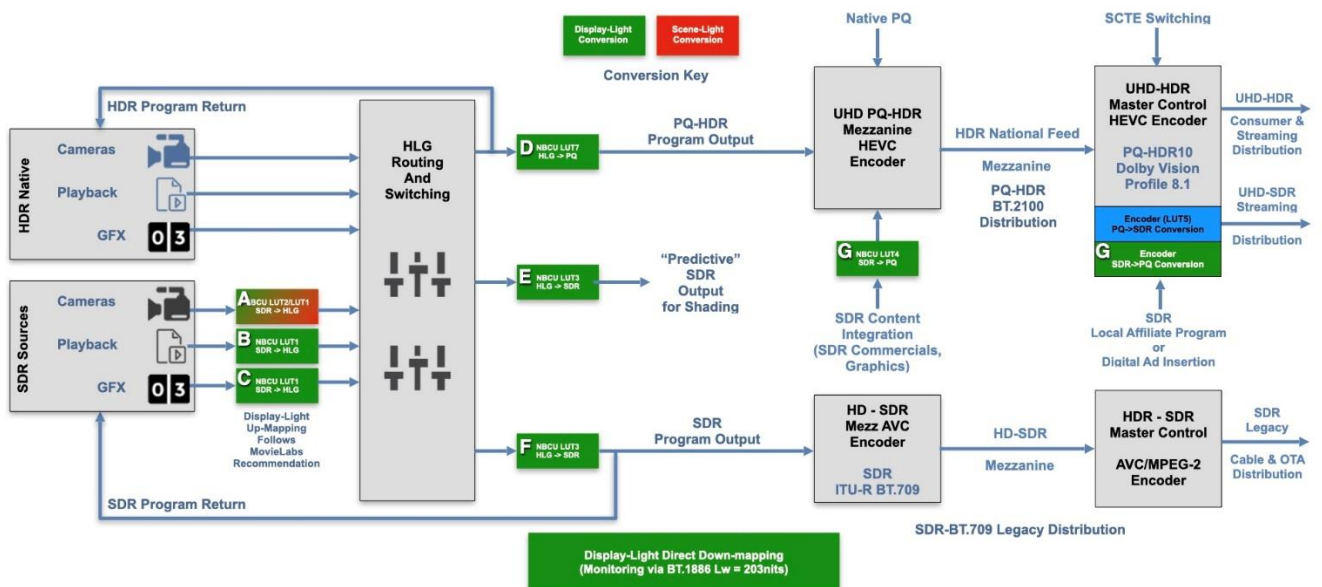


FIGURE 44
Single-Master HDR production and transmission workflow example



Annex 9

Use cases of 8K UHDTV for HDTV applications

1 Multi-shot HDTV production system using a single 8K camera

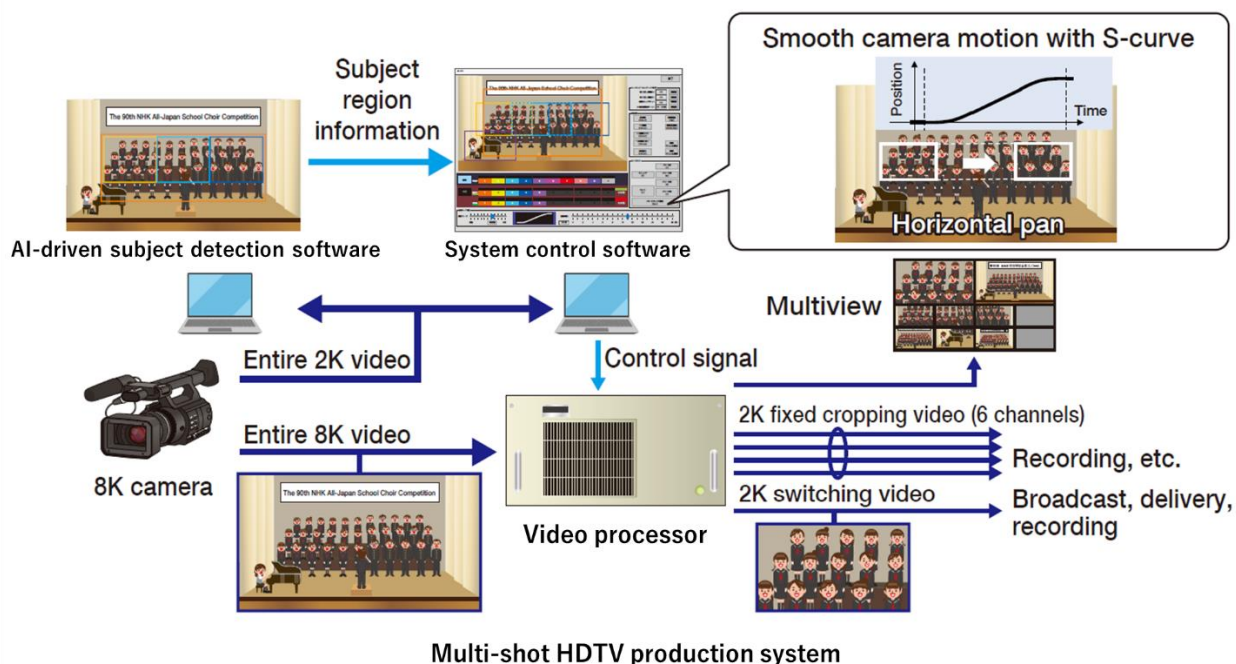
1.1 Overview

8K UHDTV images can be used for HDTV production by cropping at arbitrary position and then applying down-conversion. To take advantage of these features, a system for multi-shot HDTV production using a single 8K camera was developed and is now in practical use. With AI-driven subject detection, virtual camerawork and multi-shot switching functions, HDTV programmes can be produced with one 8K camera and one operator.

An overview of the HDTV multi-shot production system using a single 8K camera is shown in Fig. 45. The entire 8K video captured by an 8K camera is fed to a video processor. Simultaneously, the entire HDTV video down-converted from the entire 8K video is fed to an AI-driven subject detection software and system control software. The AI-driven subject detection software detects the region of the images containing the subject, such as performers, from the input video and supplies the subject region information to the system control software. The system control software then adjusts the cropping ranges based on the subject region information, creating multiple virtual cameras. These cropping ranges can also be manually determined by the operator. The system control software additionally provides virtual camerawork and multi-shot switching. The system control software supplies the control signal, including the cropping ranges, virtual camerawork and multi-shot switching, to the video processor. The video processor then converts 8K video to 2K video according to this control signal. The video processor has a main and next outputs, as well as six virtual camera outputs for recording and/or monitoring.

FIGURE 45

Overview of HDTV multi-shot production system using a single 8K camera

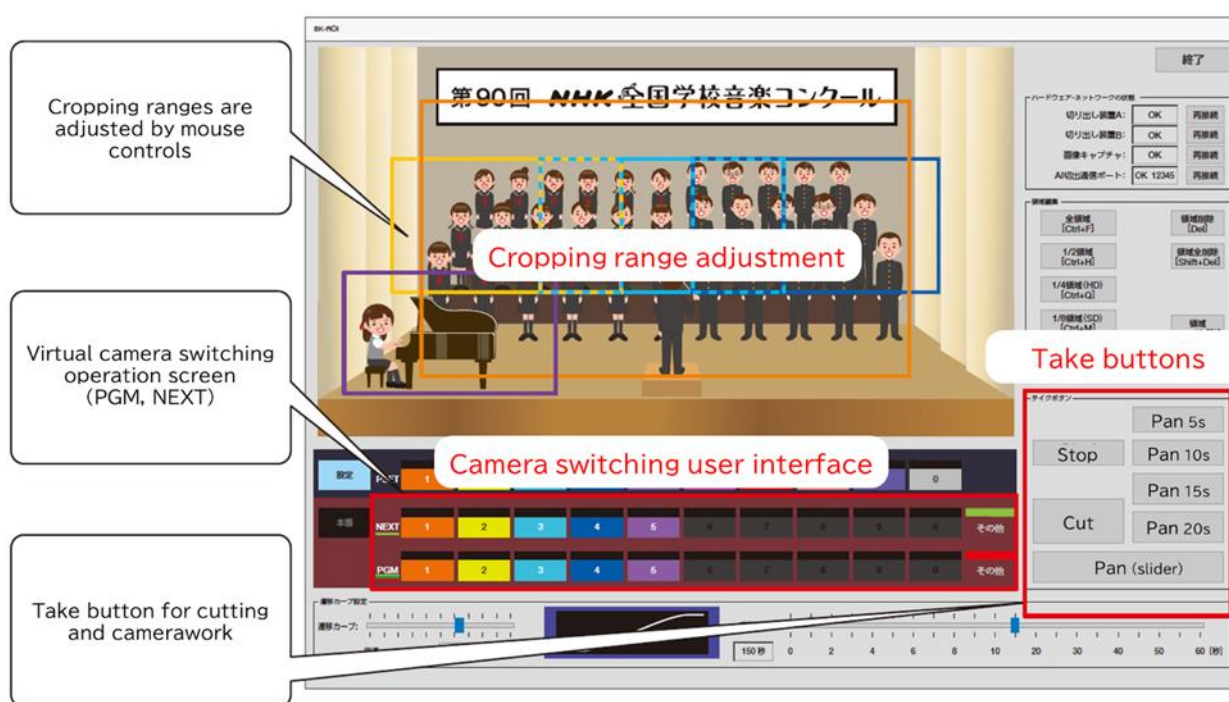


1.2 System control software

The graphical user interface of the system control software is shown in Fig. 46. This software enables cropping range adjustment, multi-shot switching and virtual camerawork. Some cropping ranges can be automatically adjusted by the AI-driven subject detection technology, while others can be manually selected by dragging the mouse cursor over the entire 2K image displayed on the software. The cropping ranges can also be finely readjusted using the mouse and the keyboard.

The system control software creates a maximum of ten virtual cameras and enables multi-shot switching. Panning is possible, where the cropping range smoothly transitions between two cropping ranges for a selected period with a control using an S-curve with a constant jerk, allowing for natural camerawork movement similar to professional camera operation.

FIGURE 46
System control software



1.3 Programme production

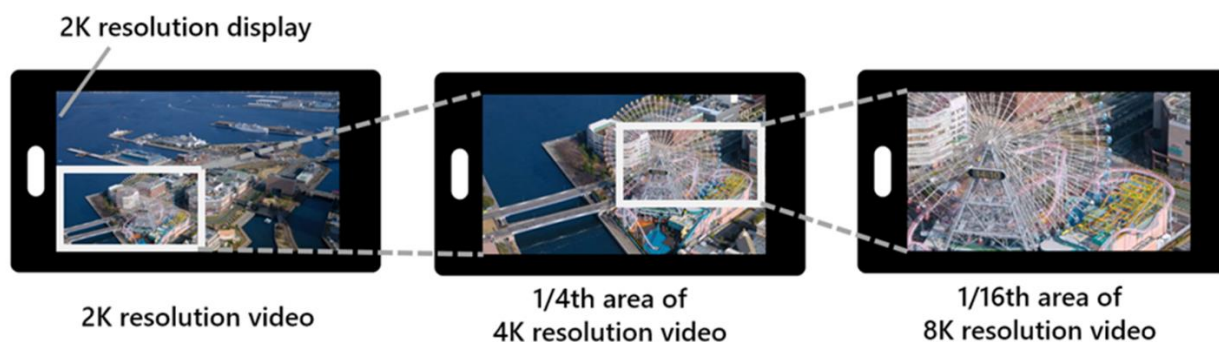
Using this production system, HDTV programmes featuring live music performance were produced (Fig. 47). Owing to the smooth virtual camerawork control and the simple and intuitive cropping range adjustment, as well as a straightforward camera switching user interface, the system was highly appreciated for efficient programme production.

FIGURE 47

Operation in programme production**2 Interactive viewing of 8K UHDTV on HDTV display with zooming and panning****2.1 Overview**

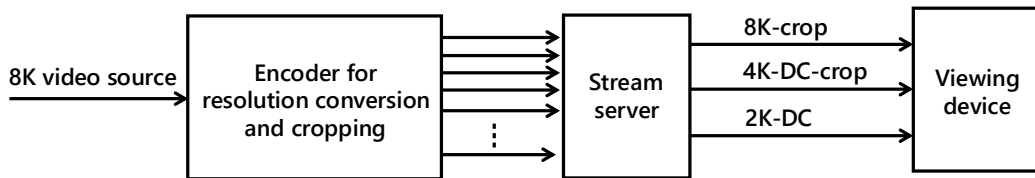
A system has been developed to enable interactive viewing of 8K UHDTV on an HDTV display with zooming and panning. Figure 48 shows the conceptual diagram of the system. Initially, on a viewing device with 2K ($1\,920 \times 1\,080$ pixels) resolution, the entire 8K video frame, after being down-converted to 2K resolution, is displayed. When a user wants to closely examine a specific area, instead of enlarging the 2K video with a zoom factor of 200%, 1/4th area of the corresponding 4K ($3\,840 \times 2\,160$ pixels) video with a zoom factor of 100%, or even 1/16th area of the corresponding 8K ($7\,680 \times 4\,320$ pixels) video with a zoom factor of 100% is displayed. In each case, the onscreen area of the video maintains 2K resolution even while zooming operation. This technology allows users to freely enlarge or move the viewing area on handheld/mobile devices such as smartphones and tablets, which typically have challenges in decoding and displaying 8K video.

FIGURE 48

Viewing of 8K video on HDTV display with zooming and panning**2.2 System architecture**

The system architecture is shown in Fig. 49. The system comprises the following three components: (a) a video encoder that produces video files with different resolutions and cropping areas through resolution conversion and cropping, (b) a stream server that stores the video files produced by the encoder and distributes them to viewing devices, and (c) a viewing device that displays the appropriate videos based on the user zooming/panning operations.

FIGURE 49
System architecture



An overview of the encoding for resolution conversion and cropping is shown in Fig. 50. “2K-DC” represents the 2K video down-converted from the original 8K video, “4K-DC-crop” represents an element of the 25 video files down-converted from the original 8K video to 4K resolution and cropped to a 2K frame size with a stride of 480 pixels (horizontal) and 270 pixels (vertical), and “8K-crop” represents an element of the 49 video files cropped to a 2K frame size with a stride of 960 pixels (horizontal) and 540 pixels (vertical).

The viewing device chooses the optimal files from a total of 75 video files (2K-DC, 4K-DC-crop, and 8K-crop) based on the zoom factor and playback position, and displays them by stacking in descending order of the resolution. A set of example video encoding parameters for the video files of 2K-DC, 4K-DC-crop, and 8K-crop is listed in Table 20.

FIGURE 50
Video encoding for resolution conversion and cropping

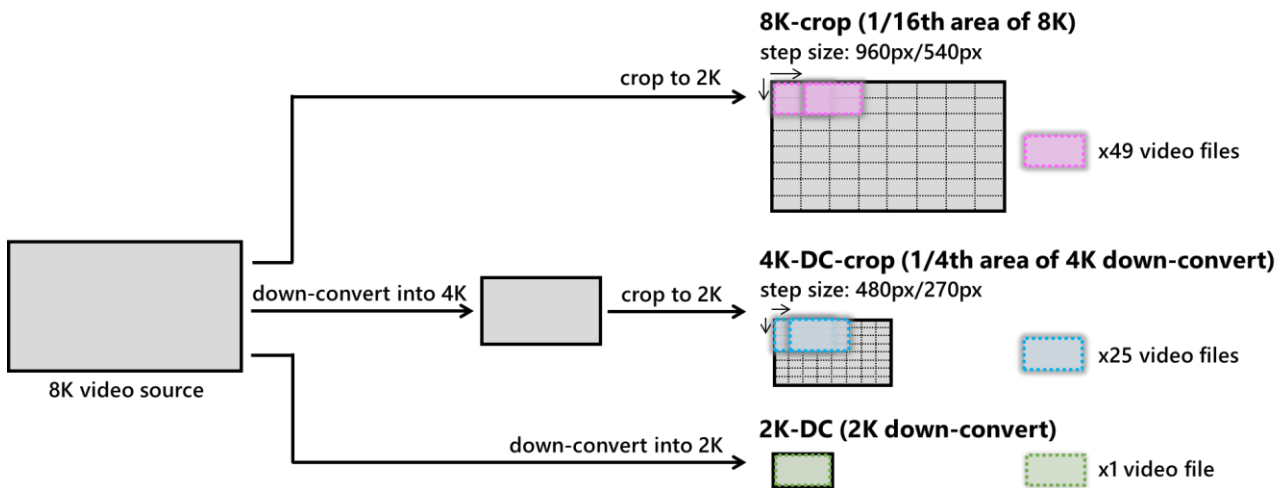


TABLE 20

Example video encoding parameters for video files

Pixel count	1 920(H) × 1 080(V)
Frame frequency	30 Hz
Video coding	H.264/AVC
Bitrate	6 Mbit/s
Streaming protocol	MPEG-DASH
Segment length	0.4 s
Total number of video files	75

2.3 Procedure for zooming and panning

When the zoom factor is changed from 100% to 200% by a pinch-in operation while playing 2K-DC, the viewing device chooses the optimal element of 4K-DC-crop closest to the touchpoint. While the viewing device requests the 4K-DC-crop from the server, it continues playing the 2K-DC with a zoom factor of 200% until it receives the new segments of 4K-DC-crop. Once ready, the viewing device seamlessly starts to play the 4K-DC-crop over the 2K-DC. A similar procedure is used when changing the zoom factor from 200% to 400%.

When the video onscreen area is changed by a swipe operation while watching 8K-crop, the viewing device chooses the optimal element of 8K-crop closest to the destination area determined by the swipe. While the viewing device requests the element of 8K-crop from the server, it continues playing the 4K-DC-crop and 2K-DC until it receives new segments of 8K-crop. Once ready, the viewing device seamlessly starts to play 8K-crop over the 4K-DC-crop and 2K-DC, and simultaneously requests the 4K-DC-crop from the server. The video layered behind the 8K-crop is seamlessly switched to the 4K-DC-crop when ready for playing.
