

Doc T-22/22-1

Radio, Film & TV Design and Research Institute, NRTA

Immersive Video in China

Ouyang Yue

Abstract

National Radio and Television Administration (NRTA) of China proposed the concept of 5G high-tech video, including four whole new aspects: interactive video, immersive video, VR video and cloud games. Immersive video is an audio and video system, which is viewed with the naked eyes, and the presentation screen covers a field of view of 120° horizontal and 70° vertical, and has 3D sound effects. It has the characteristics of ultra-high image quality, large viewing angle, and a strong sense of immersion, which can bring the audience an immersive experience. As an important part of the high-tech video format, immersive video has enabled a series of new technologies, new products, new applications and new consumption scenarios, and will continue to contribute to the digital development of emerging fields such as tourism, museums and art exhibitions, cultural entertainment. In the future, immersive video technology will be used to build a virtual world that complements the real world, creating a Metaverse entrance that integrates virtual and reality.

1. Introduction

In 2019, National Radio and Television Administration (NRTA) of China proposed the concept of 5G high-tech video, including four whole new aspects: interactive video, immersive video, VR video and cloud games. These new technologies aim to serve the public with videos of "higher technical formats, updated application scenarios, and better audio-visual experience".

Immersive video is an audio and video system, which is viewed with the naked eyes, and the presentation screen covers a field of view of 120° horizontal and 70° vertical, and has 3D sound effects. It has the characteristics of ultra-high image quality, large viewing angle, and a strong sense of immersion, which can bring the audience an immersive experience that cannot be achieved by traditional video and audio technology. Immersive Video focuses on content with clearer and more delicate images, more coherent and smooth pictures, wider and more shocking field of vision, saturated and vivid colors, immersive sounds, and presentation terminals with a strong sense of enveloping.

It is an important development direction of audio-visual technology for the future. To promote the integration of the traditional media with the new information technologies and increase the supply of high-quality cultural contents, the “promotion of immersive video and other applications” is clearly specified in the 14th Five-Year Plan of China.

After three years of development, immersive video has been widely used in China's variety shows, large-scale events, science exhibitions, virtual videography, tourism experience, commercial entertainment, education and training, home terminals and other scenarios, meeting the ever growing need of the Chinese people for better audio-visual experience. The application of immersive video has driven the rise of a huge immersive industry market, and at the same time has raised new requirements for content collection and production, audio/video encoding and decoding, transmission and distribution, as well as diversified presentation terminals.

2. Core Elements of Immersive Video

2.1 Envelopment Feeling of the Image

According to the visual characteristics of human eyes, when image can cover larger area of people's vision, with the horizontal angle of view is greater than 120° and the vertical angle of view is greater than 70°, then people can feel an immersive effect. In order to achieve the above picture envelopment feeling, immersive video can be presented in the form of spherical screen, annular screen, immersive room CAVE and other irregular presentation forms. The display mode can be projection & reflective screen or LED and other self-display screens.

2.2 Picture Quality

Immersive video stands for high-quality pictures. Resolution, contrast, color gamut, frame rate, color depth and other aspects are important indicators of image quality, the index requirements and recommended value are shown in the following table.

2.3 Immersive Sound

The sound system for the immersive video recommends the use of Object based or scenes based 3D sound technology. It is required to achieve a sound system no less than 5.1.4 channels, that is, four top channels are added to the traditional 5.1 surround sound.

The core elements for immersive video production are shown in Table 1.

TABLE 1 Table of Core Elements of Immersive Video

Order	Parameter	Recommended value
1	Horizontal field of view	$\geq 120^\circ$

2	Vertical field of view	$\geq 70^\circ$
3	Presentation	spherical screen, annular screen, immersive room CAVE and other irregular presentation forms
4	Resolution	4K/8K
5	Dynamic Range	HDR (not applicable to the projection of the inner arc spherical screen terminal)
6	Color Gamut	A minimal of DCI-P3, ITU-R BT.2020 is recommended
7	Frame Rate	A minimal of 50/60, 100/120 is recommended
8	Color Depth	A minimal of 10,12 is recommended
9	Sound System	A minimal of 5.1.4 channels

3. Key Technologies

Immersive video involves nine key technologies, which are panoramic video capturing, video image stitching, 3D image mapping, computer image generation, 3D sound, coding and transmission, multimedia control, projection display, LED Screen Display Technology. None of these technologies are entirely new, but for the first time, they are being strung together.

To transmit immersive video, which has bigger picture size and higher picture quality request, the coding and transmission will be a little different. The code stream requirements of immersive video can be divided into three levels that consist of 4K and below, 8K, and above based on the image resolution. The average bit rate of 4K video is 60~100Mbps, and the transmission rate requirement is 80~125Mbps. The average bit rate of 8K video is 150~200Mbps, and the transmission rate required is 180~250Mbps. The average bit rate of video above 8K is 300~500Mbps, and the transmission rate required is 350~600Mbps.

Video coding and decoding technology is the basis for the development of digital TV. Video coding technologies applied to immersive video mainly include H.264, H.265, and H.266 (proposed by ITU-T), as well as the video coding standards AVS/AVS+, AVS2, and AVS3 with independent intellectual property rights developed by the domestic AVS working group. These standards correspond to international standards.

Immersive video can be transmitted by cable TV, 5G wireless network, and Internet according to different application scenarios and actual conditions. These three transmission schemes are suitable to various application requirements of

immersive video.

4. Immersive Video end-to-end Solution

Immersive video service can be mainly divided into live broadcast and on-demand scenarios. The two scenarios correspond to different production processes. The overall framework of its shooting, production, transcoding, transmission, distribution and presentation system is shown in Figure 1.

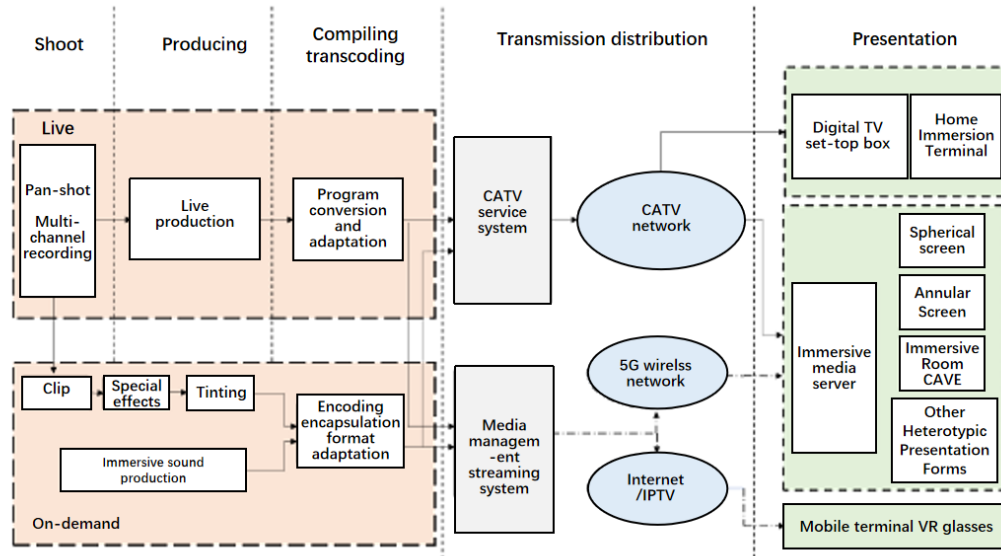


FIGURE 1 Overall Framework of Immersive Video End-to-end System

For the live broadcast immersive video service, panoramic camera is used to capture and shoot on location, and real time seaming is performed to synthesize the panoramic image. At the same time, multi-channel audio recording is completed, and then the video and audio signals are synchronously encoded, compressed, and packaged for adaptive network transmission, and high bit rate panoramic video and audio stream signals are transmitted to the media server at the presentation end through cable TV network, 5G wireless network, and the Internet. Finally, it is rendered on terminal display device after audio and video decoding and rendering.

For on-demand immersive video service, post production links such as editing, special effect, and color grading are required for stitched panorama shot on site or CG produced pictures. The sound need to be mixed to a 3D sound format, and then the produced video and audio are encoded, packaged and output for subsequent playback or on-demand use. The transmission, distribution and presentation links are similar to the live broadcast process.

5. Application Scenarios

Immersive video is widely used in sports events, science and technology venues,

exhibition venues, amusement venues, commercial venues, home terminals, education and training, and virtual filming shooting fields due to its unique visual presentation.

Scenarios	Applications
Live Sports Events / Live Major Events	The content can be presented on a spherical screen terminal by employing panorama camera technology for broadcast television content acquisition and recording, video content creation, video content splicing, coding and decoding, distribution, and transmission. Viewers can experience interaction with the content scene.
Science Museum	The application is a spherical screen theater, and projection-based display technology is used to achieve it. Immersive video typically includes nature scenes such as the sky, ocean, and forest, along with 3D sound and other sensory effects to create an immersive atmosphere.
Exhibition Venue	The majority of the applications consist of L-shaped folding screens, tri-folding screens, immersion room CAVE spaces, and spherical screens. The display technology is based on projection and has been widely used in exhibitions for the Communist Party of China.
Amusement Parks	By deploying flying theater, spherical screen theater, and other equipment configurations, the application achieves entertainment and science teaching purposes. The display system is based on projection or LED and other self-display screens.
Commercial Venues	Typically, the applications take the form of a unique dome structure combined with digital 3D projection technology to create a super stunning immersive experience space or interactive projection art installation. It has been widely used in new product launches, business meetings, promotional activities, shows, parties, concerts, and other commercial activities.
Education and Training	The application method is generally a spherical virtual simulation system in the form of an arc screen or a tri-fold screen. The system is based on multi-channel projection technology and three-dimensional synchronous imaging design and is suited for virtual simulation teaching training, and visual simulation exercises.
Home terminal	Application form can be a small spherical screen of 1.5-3 meters in diameter, with a single projection or multiple projectors stitching to present the image and surround sound system to produce a sense of sound envelopment. Combined with motion effect seat equipment, blowing wind, water mist, and different smells and vibrations, it provides the audience with an immersive audio-visual and physical experience.
Virtual filming	The filming background is displayed in real-time on a projection bowl screen or LED tri-fold screen, providing an immersive space where actors perform in front of a giant screen with a potent visual effect and immersion. It can be utilized in virtual

	studios, films, and television programs. It facilitates filming and provides the audience with the effect of naked-eye 3D.
--	--

6. Typical Cases

6.1 Major Event Applications

During the parade for the 70th anniversary of the founding of the People's Republic of China, Xinhua News Agency built a VR live broadcast matrix, relying on VR + 5G + 8K. 12 cameras are used on site to instantly transmit this unprecedented feast in VR, providing the director to switch perspectives and the freedom to select different camera positions in order to create an immersive sense of realism for the tens of millions of online viewers.

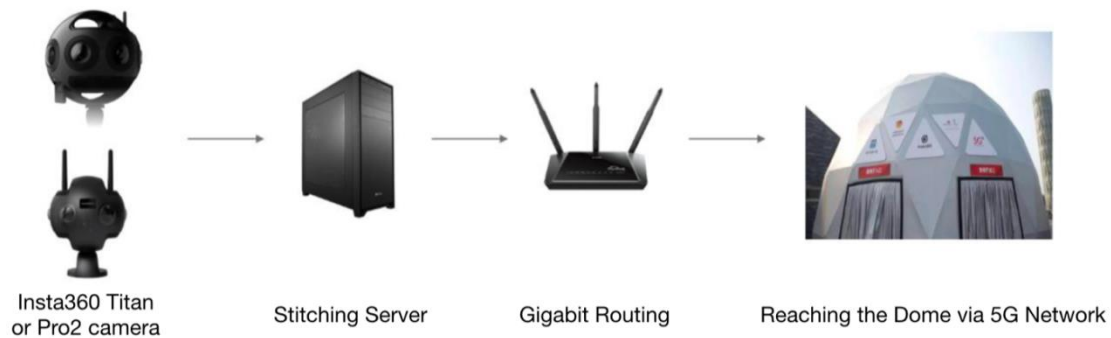


FIGURE 2 Spherical live system diagram for the 70th National Day

6.2 Virtual Production Application

Extended reality (XR) technology is the frontier of the application of immersive video for virtual production. XR technology synchronizes the shooting background on a projection bowl or LED tri-fold screen, so what you see is what you get. A media server with integrated rendering capabilities, a pan-tilt camera for filming, a bowl projection or LED tri-fold screen for presentation, and a switcher constitute the virtual filming system.

2022 Dragon TV's "Bright Moonlight Oriental Moonlight Camping Club" breaks the tradition based on the previous Mid-Autumn Festival brand "Moonlight Oriental", innovatively applies the hecoos xR virtual shooting technology, and conducts the reunion core of "Water Melody Song Head" with a new interpretation. In the program "Hope for a Long Life", the actors used the form of twin dance to present the brotherhood of Su Shi and Su Zhe in the realm of national style and fantasy expanded by xR technology, bringing the audience the best wishes of "a thousand miles together".



FIGURE 3 Immersive LED screen for virtual shooting scenes



FIGURE 4 Virtual production site effect

7. Products

Since 2020, the Radio, Film & TV Design and Research Institute(DRFT) has taken the lead in undertaking the task, assigned by NRTA, of compiling the Immersive Video White Paper and Standard System. DRFT built an immersive video laboratory to continuously carry out technical research, system testing, standard formulation and application demonstration of immersive video. With an open gesture, DRFT has cooperated with around one hundred partners in production, academic, research and application fields to advance the industrial promotion of immersive video. A batch of

good technologies and products with independent intellectual property rights have been hatched, reaching international leading level in panoramic shooting, video playback control, LED screen presentation, and 3D audio.

China has continuously innovated and created the best immersive video shooting tools, using self-developed software and hardware imaging technology. Its cinematic 360° camera product has been successfully used for immersive video shooting that requires high image quality.

As one of the world's leading LED manufacturing country, China has successfully created classic immersive projects such as "Riyadh Season" in Saudi Arabia, and assisted in the virtual videography of famous films and TV series such as "The Mandalorian" and "Westworld". The sound-transmitting LED spherical screen, develops by China, has been successfully applied in planetariums in Nagoya and Yokohama, Japan. The system has solved the problems of sound focusing in spherical presentation space and limited speaker's installation.

China's enterprises continue to carry out technology research, development and innovation, and have developed 3D audio recorders and supporting softwares, content creation tools and material libraries, 3D audio decoders, processors, etc. These products have formed a closed loop from 3D audio collection, production to playback, and have been applied in the fields of film, TV, music, VR, Metaverse, automotive and consumer electronics.

8. Future

As an important part of the high-tech video format, immersive video has enabled a series of new technologies, new products, new applications and new consumption scenarios, and will continue to contribute to the digital development of emerging fields such as tourism, museums and art exhibitions, cultural entertainment. And it will be gradually applied to people's daily consumption such as catering, shopping, and entertainment.

In the future, immersive video technology will be used to build a virtual world that complements the real world, creating a Metaverse entrance that integrates virtual and reality, and real-time interaction, boosting the further development of the Metaverse industry.