Research and application of immersive panoramic VR image technology in the field of short video live broadcast

Yang Xia
Zhejiang University of Media and Communication, Hangzhou 310018, China

Abstract: With the continuous iteration and upgrading of panoramic VR image technology, panoramic VR has gradually entered the field of short video, and major short video platforms have gradually expanded the panoramic VR image content ecology. The immersive and interactive panoramic VR live broadcast gives the live broadcast a new user experience. However, compared with ordinary video, the data amount of VR video is far more than that of ordinary video, and the live bandwidth required by VR video is also far more than that of flat video. Therefore, the current network bandwidth is not enough to support high-quality VR live streaming. Therefore, this paper proposes a VR live broadcast scheme based on the user's field of view. By calculating and extracting the user's field of view area, the user's quality remains unchanged, the video quality of the non-field of view area is reduced, and the network bandwidth required for live broadcast is greatly reduced.

Key words: panoramic VR; Live streaming; Short video; Field Angle

1. The wide application of panoramic VR live broadcasting in various industries

With the continuous maturity of display and sensing technology, transmission and content production, panoramic VR live broadcast will gradually empower all walks of life. The panoramic VR live broadcast is used for the reporting of news events, and the application of VR technology allows the audience to have a richer visual sensory experience and a 360-degree viewing Angle, so as to better understand the current social hot events. When the panoramic VR live broadcast is used in games, users will enter the VR game world from the real world, project another "self" in the virtual world, and complete the behavior and actions of virtual characters through somatosensory interaction. The immersive participation experience of VR live broadcast broadens the development radius of the game e-sports live broadcast industry. Using panoramic VR live broadcast in traditional industries, VR has penetrated and applied to the manufacturing field in the era of Industry 4.0. VR has brought new development opportunities to the manufacturing information industry, and realized the simulation and optimization of product research and development.

2. The network bottleneck problem of panoramic VR live broadcast

The panoramic VR live broadcast picture is rich in information and strong interaction, so the data stream of its live broadcast is several times that of ordinary video live broadcast, and its bandwidth requirements for live broadcast also rise. In order to ensure the audience's visual experience, the overall resolution of VR video is generally 8K or even 12K, and the required bandwidth is even gigabit network. Most ordinary users in China can not meet the network requirements, which also causes great resistance to the promotion and popularization of VR live streaming.

3. VR live broadcast solutions based on user field of view Angle

According to the "National Network Speed and Quality Report 2022", the average download rate of domestic broadband networks in 2022 will reach 93.92Mbps, and the average upload rate will reach 38.20Mbps. Although the domestic network speed has been improving year by year, the current network speed is still the main factor limiting the perception of VR live broadcasting. Considering that the actual observation range of the human eye is limited, the horizontal field of view Angle of the human eyes can reach 200°, the area that our left and right eyes can coincide to see is about 120°, the best fixation area is 90° horizontal, 90° vertical, after exceeding the range, usually the way to rotate the head to adjust the Angle of view. The visual range of 90° is only a very small part of the VR spherical video. So in the data transmission, most of the data transmission efficiency is relatively low, the human eye can not see the part but occupy a lot of network resources. Therefore, a VR livestream solution based on field of view is proposed here. Different video quality adjustment is made between field of view area and non-field of view area. The video quality in the observation range of human eyes (field of view area) can be reduced to reduce the video quality in the non-field of view area, reduce the overall bandwidth, improve the real-time transmission efficiency of VR video, and optimize the user's experience.

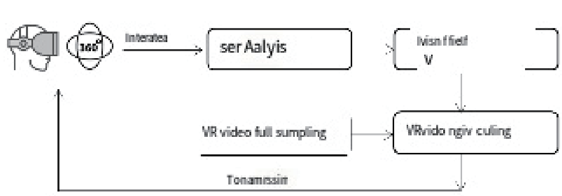

Figure 3-1 Process of VR live broadcast based on field of view Angle

As shown in Figure 3-1, the process of VR livestream based on field of view can be roughly divided into four steps: (1) user data analysis, (2) field of view area division, (3) VR video full sampling, and (4) VR video sub-area coding.

1. User data analysis

The client sends the user's head movement information to the server in real time, and extracts the user's field of view Angle from this data.

y:150r:-8 p:-60 (3-1)

If the data of Formula 3-1 is collected, y represents the yaw Angle of the user, that is, the left and right rotation value of the user's head; r stands for tangent bevel, that is, the Angle between the user's head and the horizontal line; p stands for pitch Angle, which is the Angle of rotation of the user's head in the pitch direction. According to Formula 3-2, the latitude and longitude of the user's viewpoint at this time can be calculated.

$$\begin{cases} \alpha = (y/180)*\pi \\ \beta = (p/180)*\pi \end{cases} \quad (3\text{-}2)$$

2. Field of view Angle area division

Through user data analysis, using Equi-Rectangular Projection(ERP), the three-dimensional sphere video can be developed into a two-dimensional plane video, and the position of the user's viewpoint in the two-dimensional plane video is represented by [x,z]. According to the general law of ERP projection, the farther away from the equator and the farther to the two polar regions, the larger the stretch of the unfolding graph will be. Therefore, the field of view Angle is set to a quarter of the VR video frame. The schematic diagram of the field of view Angle in three-dimensional space is shown in Figure 3-2.

In Figure 3-2, the longitude of the user's viewpoint is represented by (α) and latitude by (β); The horizontal user's viewpoint is represented by (ɛ). When the user's viewpoint moves from the middle equator to the North and South poles, the value of ɛ will change from 180° to 90°; The vertical user's Angle of view is expressed by (θ), which remains unchanged, or 90°. According to the ERP projection function method, there are generally three cases of the user's field of view Angle: the center area, the edge area, and the north and South Pole area. There will be different calculation methods for different situations.

(1) Center area

At this point the user is in the center of the three-dimensional sphere space, and the viewpoint is near the equator. According to the planar video diagram expanded by ERP projection, the position of the user's viewpoint is located in the central area of the image (equation 3-3), representing the field of view Angle, {} represents the rectangular area between vertices, the n and c subtables represent the number of rows and columns of the panoramic video image, and x' and z' are the two thresholds. f The area of the field of view Angle can be represented by a rectangle, as shown in Figure 3-3. $f = \{(x-x',z-z'),(x+x',z+z')\}$

(2)Edge area

On an ERP projection expansion two-dimensional plan, the left and right edges are actually connected in a three-dimensional sphere. Therefore, when the user's viewpoint is near the edge area on one side, the field of view Angle area will be divided into two areas, as shown in Figure 3-4. At this time, Formula 3-54 can be used to calculate the field of view Angle area. f f

$$f = \{(1, z-z'), (x-c+x', z+z')\} \cup \{(x-x', z-z'), (c, z+z')\}$$

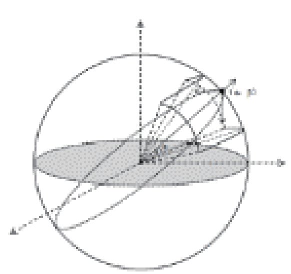

Figure 3-2 Schematic diagram of the field of view Angle in three-dimensional space

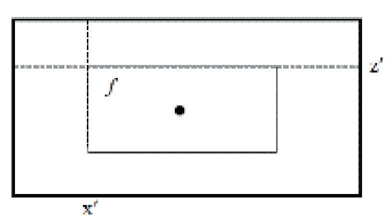

Figure 3-3 Schematic diagram of the field of view Angle in the center area f

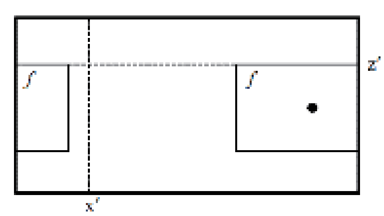

Figure 3-4 Schematic diagram of the field of view of the edge area f

(3) North and South Polar regions

As the user's viewpoint reaches the north and South polar regions, the stretch of the developed plane video map will increase. Because of the stretching of the image, Formula 3-5 is needed to calculate the field of view Angle region, as shown in Figure 3-5. f

$$f = \{(1,1), x-x', z'-z\} \cup \{(x+x',1),(c,z'-z)\} \cup \{(x-x',1),(x+x',z+z')\}$$

3. Full sampling of VR video

The main purpose is to get a basic panoramic frame. Considering the delay of the video bit rate change when the network delay and the field of view Angle change, the panoramic frame is combined with the field of view Angle calculated in real time as a security frame, and then the combined data is encoded. It can avoid the phenomenon of video missing or delay, and improve the user's experience.

4. VR video sub-region coding

There are three main ways to combine panoramic frames with real-time calculated field of view angles, as shown in Figure 3-6. (a) The figure is the combination of the user's view point in the central area. The view Angle is represented by a rectangle, and the subsampling factor is set. The view Angle and the edge of the panoramic frame are equal in length or width, and the two parts can be combined directly. f f (b) figure is the combination of the user's viewpoint in the edge area, the field of view Angle is composed of 2 rectangles, which can also be spangled into the pattern of (a) figure, so it can also be directly combined with the panoramic frame. *f* (c) Figure is the combination of the user's view point in the polar region, the field of view Angle is not composed of 1 or 2 rectangles, because the size of the entire field of view Angle is unified, so it can also be divided into multiple rectangles for combination. *f* *f* The 3 combination modes are essentially the same as the final image size.

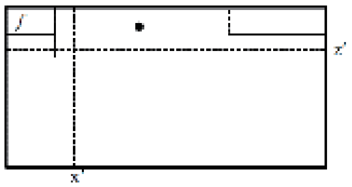

Figure 3-5 Schematic diagram of the field of view Angle of the polar regions *f*

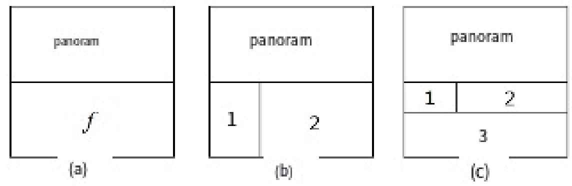

Figure 3-6 Schematic diagram of panorama frame and field of view Angle combination

4. Comparative experiment

In order to evaluate the actual effect, the following insta360pro2 is used as the shooting equipment to carry out the real shooting comparison experiment. insta360pro2 is composed of 6 fisheye lenses with 200° field of view, which can be combined in real time to push the live stream, and can also obtain the shooting data of 6 cameras and splice it by itself in the later stage. During the experiment, insta360pro2 was used to shoot the panoramic video of the laboratory. Two sets of videos were shot, one for motion video and the other for fixed video. The duration was 1 minute, the frame rate was 29.97fps, and the video resolution was 7680x3840. Use the H.264 coding standard. Design the normal motion trajectories of multiple user perspectives, and compare the VR calculation method based on field of view Angle with the normal VR calculation method based on full data. The comparison data are shown in Table 4-1.

Table 4-1 experimental video sizes

Case List	Normal video size (MB)	Field of View Angle Video size (MB)	Reduction (%)
Exercise video	1498.52	799.86	53.4
Fixed Video	1497.88	743.27	49.6

According to the experiment, the transmission method of the on-time field Angle can reduce the data of about 52% than that of normal VR video, which means that half of the network speed resources are saved. At the same time, the data size decreases, the efficiency of the corresponding video coding is improved, the pressure on the network is reduced, and the network delay is reduced. Structural Similarity (SSIM) is used to compare the quality of video, and the result is shown in Figure 4-1. The quality of the video in the field Angle area is the same as that of the original video, while the quality of the video in the non-field Angle area is significantly decreased. The overall quality of the reconstituted video is also lower than that of the field of view area, because the video in the field of view area of the reconstituted video only accounts for 1/4. The video quality of the reorganized video in the field of view area is not reduced, and the user experience is not reduced, but the transmission pressure is reduced.

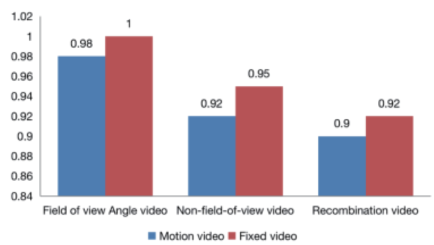

Figure 4-1 Structural Similarity (SSIM) comparison

5. Summary

Through comparative experiments, the effectiveness of the VR live broadcast scheme based on field of view Angle proposed in this paper is verified. By considering the actual observation range of human eyes, this scheme distinguishes the video quality of the field of view area and the non-field of view area, and greatly reduces the size of VR video on the premise of ensuring the user's experience, thus alleviating the current problems of high network transmission pressure and high network time delay during VR live broadcasting. It provides an important technical reference for the further popularization and promotion of VR live broadcasting on short video platforms.

References:

[1] YuehuaWang. Virtual Reality video production and application scenario analysis [J]. Audiovisual,2019(7):11-12

[2] YuhangMa. Characteristics and development trend of news client VR Live streaming [J]. Young Reporter, 2020(32):92-93.

[3] KainingDong,Jin Meng. Research and application of immersive VR in the field of animation education [J]. Animation Research,2018(1):6.

Fund Project: 2021 Open Project of Zhejiang Key Laboratory of Film and Television Media Technology Research (Serial number: CM2021006).