Research on the Methods of Panoramic Video Projection Mapping

Jiangeng Li*
Faculty of Information Technology
Beijing University of Technology
Beijing, China
* Corresponding author: lijg@bjut.edu.cn

Kai Xu
Faculty of Information Technology
Beijing University of Technology
Beijing, China
v587dxk@163.com

***Abstract*—In recent years, panoramic video has gradually emerged in the multimedia industry. As an important part of the panoramic video processing process, the choice of projection mapping method will affect the efficiency of panoramic video coding and decoding, transmission mode and user experience. There have been a lot of researches on the projection method from the spherical surface to the plane, and its important purpose is to reduce the distortion of the spherical video in the projection and improve the performance of panoramic video encoding. At present, projection methods are mainly divided into uniform quality projection and non-uniform quality projection. This paper compares several common projection mapping methods, discusses their respective advantages and disadvantages, and makes related summary at last.**

***Keywords*-Multimedia Industry; Panoramic Video; Projection Mapping**

I. Introduction

Panoramic video is a kind of video formed by 360° shooting with a 3D camera. Compared with traditional video, panoramic video allows viewers to watch videos from any angle within 360 degrees, which satisfies the viewer's vision requirements to the greatest extent [1]. With its wide field of view and strong sense of immersion, panoramic video has been widely used in surveillance, broadcasting, and video conferencing.

In the processing of panoramic video, projection mapping is a very important part. In order to facilitate the transmission through the network, the redundant information in the original spherical panoramic video sequence needs to be removed by encoding and compression [2]. The existing mainstream video coding standards cannot handle spherical video, and the spherical panoramic video projection needs to be mapped into a flat video before the next step can be processed [3]. The panoramic video player needs to decode the received video and do the inverse projection operation, and map the flat video to the spherical surface for rendering and display [4]. The choice of projection mapping format will affect the performance of panoramic video codec, transmission mode and even user experience [5].

At present, projection formats are mainly divided into uniform quality projection and non-uniform quality projection [6]. Uniform quality projection is to project each part of the spherical surface onto the plane with uniform quality, while non-uniform quality projection is to perform high-quality projection in the area within the human eye's field of view, and the area outside the field of view uses down-sampled low-quality projection [7]. The following content firstly introduces some common uniform quality projection such as ERP and CMP, then introduces one non-uniform quality projection format TSP, and carries out related summary at last.

II. Definition of Coordinate System

A. Three-dimensional Coordinate System

The coordinate system is shown in Figure 1.[4] A certain point on the sphere can be represented by (ϕ, θ).

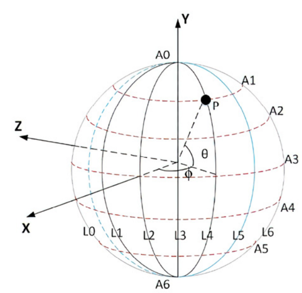

Figure 1. Coordinate system

The coordinate conversion relationship between (ϕ, θ) and (x, y, z) is as follows:

$$\begin{cases} x = \cos\phi\cos\theta \\ y = \sin\theta \\ z = -\sin\phi\cos\theta \end{cases} \tag{1}$$

B. Pixel Coordinates of the Mapped Surface

Define the number of the projection plane as f, for each mapping surface, set its resolution as $W \times H$, take the upper left corner of the mapping surface as the origin of the coordinates, and establish a rectangular coordinate system, then the position of each pixel in the mapping surface can be expressed in two-dimensional coordinates (m, n), where $m \in [0, W), n \in [0, H)$.

978-1-6654-3265-8/21/$31.00 ©2021 IEEE
DOI 10.1109/AIEA53260.2021.00041

C. (u,v) Coordinates in Projection Format

In order to facilitate the description of the relative position of the pixel on the projection plane, the (m,n) coordinates of the pixel can be converted to (u,v) coordinates, and the conversion relationship between them is as follows:

$$\begin{cases} u = \dfrac{m+0.5}{W}, m \in [0, W) \\ v = \dfrac{n+0.5}{H}, n \in [0, H) \end{cases} \tag{2}$$

III. UNIFORM QUALITY PROJECTION METHOD

A. ERP Projection

In the process of ERP projection mapping the panoramic video from the spherical surface to the plane, the warp lines on the spherical surface are mapped to vertical lines with equal spacing on the projection plane, and the latitude lines on the spherical surface are mapped to horizontal lines with equal spacing on the projection plane, using the same number of sampling points to save the data on each latitude [8].

The ERP projection mapping relationship is simple, and the two-dimensional plane video after projection is very intuitive and convenient for users to observe. However, the same number of sampling points is used for each latitude line of the spherical video, which will cause the more redundant sampling points to be on the latitude line closer to the poles. This method increases the space occupied by the video and brings great problems to the video transmission process. While the pixel density distribution is extremely uneven, it also introduces quite serious image stretching, among which the non-linear deformation is the most serious near the two poles [9]. Figure 2 [5] is a panoramic video image in ERP format. It can be clearly seen that the two poles are stretched seriously, causing great redundancy and increasing the coding burden.

Figure 2. ERP format panoramic video image

B. CMP Projection

CMP projection is another standard projection mapping format specified in the OMAF standard, which is actually a perspective projection format. As shown in Figure 3,[5] the video content of the sphere is projected onto the surface of the cube by circumscribing the cube, and then the cube is expanded into a plane. The generated two-dimensional plane image can be divided into 4x3 layout and 3x2 layout.

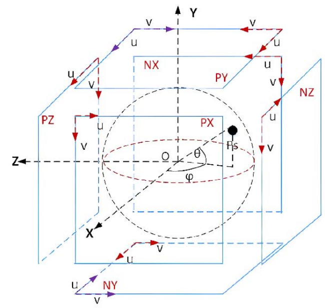

Figure 3. CMP projection diagram

The 4x3 layout is shown in Figure 4.[5] This layout maintains good image continuity between the various surfaces, which is conducive to video coding. However, there are six gray areas, which are invalid elements filled to make the layout a rectangle. These invalid elements account for half of the number of pixels in the entire two-dimensional plane image, which will increase the size of the space occupied by the video during storage, and will also lead to unsatisfactory coding efficiency [11].

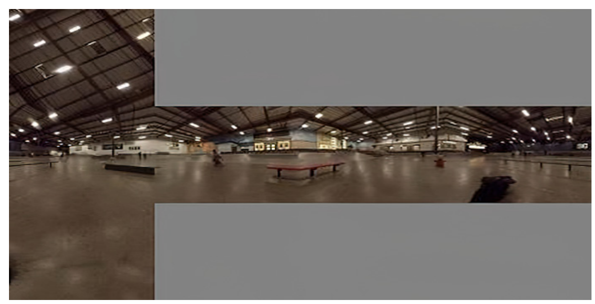

Figure 4. 4x3 Layout

The 3x2 layout is shown in Figure 5.[5] It can be clearly seen that the entire two-dimensional plane image is full of valid pixels. However, the image generated by this layout lacks continuity, which is not conducive to coding [12].

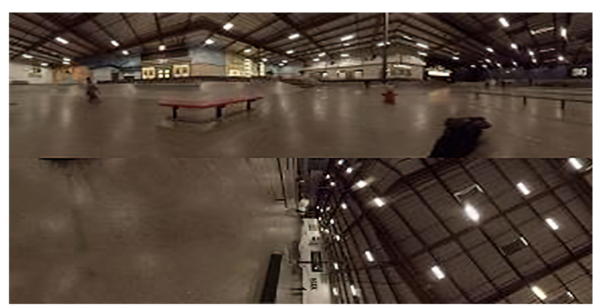

Figure 5. 3x2 Layout

Compared with the ERP projection format, the advantage of CMP projection is that there is no pixel redundancy between the two poles, and the image stretching distortion is small during the projection process [13]. But in the process of mapping from the sphere to the cube, the arcs of the same length on the circle are mapped to the square and the lengths are not equal. Therefore, for the same number of pixels on the sphere, the number of sampling pixels allocated to the edge area of the cube will be more than the number of sampling pixels allocated when projecting to the center area, that is, the edge area is sparse and the center area is dense [14].

C. *EAC Projection*

Google proposed the EAC projection, which is an optimization of CMP projection by adjusting the position of the sampling pixels on the cube corresponding to the spherical pixels to improve this non-uniform distribution [15].

The realization process of EAC projection is simply to take the angular uniformity of the pixel distribution into consideration. In the process of converting a point on a two-dimensional plane to a point in three-dimensional space, for a point (u,v) on a plane square, CMP projection can obtain the corresponding three-dimensional coordinates (x,y,z) according to its projection rules; for EAC projection, the point must first be aligned to obtain the coordinates (u',v'), and then the corresponding three-dimensional coordinates (x,y,z) are obtained according to the projection rules of CMP [16]. The conversion relationship of coordinates is as follows:

$$\begin{cases} u' = \tan\dfrac{u\cdot\pi}{4} \\ v' = \tan\dfrac{v\cdot\pi}{4} \end{cases} \tag{3}$$

The intuitive reflection is shown in Figure 6.[4]

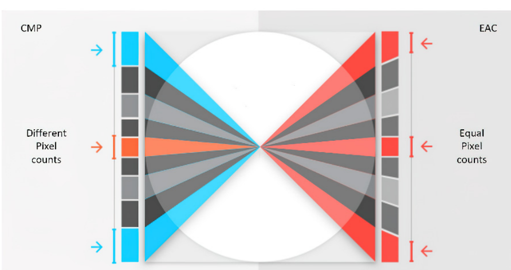

Figure 6. Pixel distribution strategy adopted by EAC

The equiangular cube projection adjusts the formula to make the sampling lengths approximately equal to form an approximately equiangular pixel distribution. Although there is still a small amount of distortion and pixel distribution unevenness, the effect has been significantly improved, as shown in Figure 7.[5]

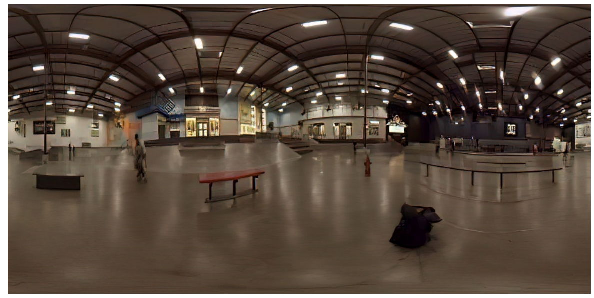

Figure 7. EAC format panoramic video image

D. *SSP Projection*

SSP uses 45° north latitude and 45° south latitude as segment boundaries, and divides the entire sphere into 3 sections: the north pole, the equatorial, and the south pole. The north pole area with spherical latitude greater than or equal to 45° and the south pole area with spherical latitude less than -45° are respectively vertically projected to the equatorial plane to obtain circular planes numbered 0 and 1, and other rectangular areas around the circular plane are filled with the default gray [17]. The projection method of the equator is the same as ERP. In order to make the size of each plane obtained by the projection the same, the plane obtained by the projection of the equatorial segment is divided into 4 planes of the same size, and the projection model is shown in Figure 8.[5]

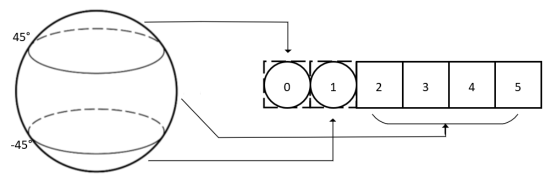

Figure 8. SSP projection model

SSP optimizes the problems existing in the projection of polar regions. The area near the two poles is directly projected vertically. The projection method is simple and the effect is relatively smooth [18]. The projection method of the equatorial section is the same as ERP, maintaining the advantages of ERP. However, the obtained plane has two discontinuous areas, and there are several gray default areas, which is a certain waste of storage space, as shown in Figure 9.[5]

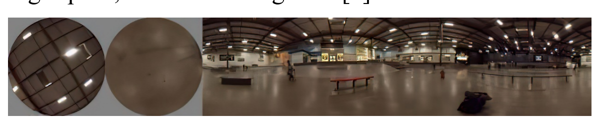

Figure 9. SSP format panoramic video image

IV. NON-UNIFORM QUALITY PROJECTION METHOD

This section introduces the non-uniform quality projection format of TSP. The projection model of TSP is a square prism. As shown in Figure 10,[5] the prism base (B) is the user's viewing area, and the pixels in this area are sampled and projected at the original resolution. The sides of the prism (R, L, T, D) are the area to be seen which are sampled at a reduced resolution and projected into a trapezoid. The top surface (F) is the area that is almost impossible to be seen next, which is projected at the lowest resolution and occupies a large square surface alone. TSP twists six cube faces into a compact frame.

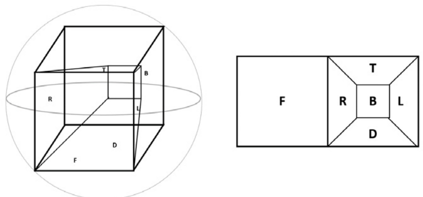

Figure 10. TSP projection model

The projection effect of TSP is shown in Figure 11.[5] The use of such a flexible sampling rate strategy is to reduce bandwidth pressure while ensuring that the video quality when the user is watching is not greatly reduced [19]. But on the other hand, although each surface of the TSP image is continuous, it can be seen that except for surface 0 and surface 1, the images of other surfaces have serious image distortion [20].

Figure 11. TSP format panoramic video image

V. CONCLUSION

Summarizing the current direction of spherical projection research, it can be divided into the following three types:

A. Use a Three-dimensional Model Closer to a Sphere

This type of method can directly improve the uniformity of pixel distribution of each surface after projection, and with it comes exponentially increasing computational complexity.

B. Use Heterogeneous Projection Model

The purpose of this type of method is more to reduce bandwidth pressure, and it often performs well when the bandwidth is limited.

C. Use Spherical Redistribution Strategy

This type of method can improve the image quality after projection with the introduction of minimal computational overhead.

REFERENCES

[1] Li Shiyan.Analysis of the application of 360° panoramic video in sports event reporting[J].News Research Guide,2021,12(03):253-254.

[2] M. Xu, C. Li, Y. Liu, X. Deng and J. Lu, "A subjective visual quality assessment method of panoramic videos," 2017 IEEE International Conference on Multimedia and Expo (ICME), 2017, pp. 517-522.

[3] R. G. Youvalari, A. Aminlou, M. M. Hannuksela and M. Gabbouj, "Efficient Coding of 360-Degree Pseudo-Cylindrical Panoramic Video for Virtual Reality Applications," 2016 IEEE International Symposium on Multimedia (ISM), 2016, pp. 525-528.

[4] Wang Zewei. Research on Rendering Technology of Panoramic Video[D]. Xidian University, 2019.

[5] Yuwen He et al., "360Lib Software Manual" Joint Video Exploration Team of ITU-T SG 16 WP3 and ISO/IEC JTC 1/SC 29/WG 11, JVET-D0071, 4th Meeting. 2016.

[6] Y J. Li, R. Sun, G. Wang and M. Fan, "Panoramic video live broadcasting system based on global distribution," 2019 Chinese Automation Congress, 2019, pp. 63-67.

[7] Y. Wang et al., "Polar square projection for panoramic video," 2017 IEEE Visual Communications and Image Processing, 2017, pp. 1-4.

[8] S. Zhang, F. Yang, S. Wan and P. Di, "Spherical Lanczos Interpolation in Planar Projection or Format Conversions of Panoramic Videos," in IEEE Access, vol. 8, pp. 9655-9667, 2020.

[9] Y. Wang, R. Wang, Z. Wang and W. Gao, "Asymmetric circular projection for dynamic virtual reality video stream switching," 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 2726-2730.

[10] Y. Duan, C. Han, X. Tao, B. Geng, Y. Du and J. Lu, "Panoramic Image Generation: From 2-D Sketch to Spherical Image," in IEEE Journal of Selected Topics in Signal Processing, vol. 14, no. 1, pp. 194-208, Jan. 2020.

[11] Yize Jin, Liming Wang, Qinwei Jiang, Xiaoyang Zeng and Yibo Fan, "An efficient spherical video sampling scheme based on Cube model," 2017 IEEE International Conference on Consumer Electronics (ICCE), 2017, pp. 211-214.

[12] B. Ray, J. Jung and M. Larabi, "A Low-Complexity Video Encoder for Equirectangular Projected 360 Video Content," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 1723-1727.

[13] A. Zare, A. Aminlou and M. M. Hannuksela, "Virtual reality content streaming: Viewport-dependent projection and tile-based techniques," 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 1432-1436.

[14] W. Chengjia, Z. Haiwu and S. Xiwu, "Octagonal Mapping Scheme for Panoramic Video Encoding," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 9, pp. 2402-2406, 2018.

[15] Y. Wang et al., "Polar square projection for panoramic video," 2017 IEEE Visual Communications and Image Processing, 2017, pp. 1-4.

[16] Y. Wang, D. Liu, S. Ma, F. Wu and W. Gao, "Spherical Coordinates Transform-Based Motion Model for Panoramic Video Coding," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 9, no. 1, pp. 98-109, March 2019.

[17] Y. Wang, L. Li, D. Liu, F. Wu and W. Gao, "A new motion model for panoramic video coding," 2017 IEEE International Conference on Image Processing (ICIP), 2017, pp. 1407-1411.

[18] Y. Liu, M. Xu, C. Li, S. Li and Z. Wang, "A novel rate control scheme for panoramic video coding," 2017 IEEE International Conference on Multimedia and Expo (ICME), 2017, pp. 691-696.

[19] H. Roodaki and S. Shirmohammadi, "Scalable multiview video coding for immersive video streaming systems," 2016 Visual Communications and Image Processing (VCIP), 2016, pp. 1-4.

[20] J. Zhang, Q. Gao and G. Zhang, "Edge Cache Replacement Strategy for SVC-Encoding Tile-Based 360-degree Panoramic Streaming," 2020 3rd International Conference on Hot Information-Centric Networking (HotICN), 2020, pp. 122-128.