# Polar Square Projection for Panoramic Video

Yueming Wang, Ronggang Wang, Zhenyu Wang, Kui Fan
Shenzhen Graduate School, Peking University
Shenzhen, Guangdong, China
wymmail@foxmail.com, {rgwang, wangzhenyu}@pkusz.edu.cn, kuifan@pku.edu.cn

Yufan Deng, Shensian SYU, Ming-Jong Jou
Shenzhen China Star Optoelectronics Semiconductor Display Technology Co. Ltd.
Guangdong, China
{dengyufan, Valexsyu, mingjongjou}@tcl.com

***Abstract*—Panoramic video provides an immersive experience by presenting a 360° spherical video content. Due to the limitations of coding and storage technology, the spherical panoramic video needs to be projected onto the two-dimensional plane for storage and encoding. In this paper, we propose a polar square projection scheme. We project the area near the poles of the sphere into two square planes and a latitude circle on sphere is projected to a square circle on squares plane, in addition, the rest of area on sphere is projected into a rectangle by means of equal area projection. Experimental results show our proposed projection can obtain a gain of 11.63% BD-rate compared to the equirectangular projection.**

***Keywords—virtual reality, panoramic video, spherical video projection, aqui-angle projection, equal area projection***

## I. Introduction

Panoramic video is a typical content for virtual reality applications, and panoramic video provides an immersive visual experience by presenting a 360° video. Users usually watch panorama video through head mounted displays (HMDs). HMD simulates a sphere of real space, where the viewer is at the center of the sphere, and HMD can track users head direction and render corresponding video content to provide a sense of immersion.

Unfortunately, existing encoding techniques do not support the processing of spherical video. Therefore, the spherical panoramic video needs to be projected onto the two-dimensional plane for storage and encoding. And since panoramic video provides 360 degrees of video content, panoramic video usually requires a higher resolution than traditional video. With the development of virtual reality video capture devices, the resolution of panoramic video is increasing to 8k and even higher-resolution. In order to address the above issues, Both ISO/IEC MPEG and IEEE 1857 have setup adhoc workgroups to standardize efficient panoramic video projections and coding methods [6].

There are many projection methods to map a spherical panoramic video onto a two-dimensional plane [1][2], and the coding efficiency varies with different projection methods. Equirectangular projection [3] (ERP) is one of the most basic and common projection methods, and is widely used in panoramic video. However, ERP has severe oversampling in high latitude regions and the coding efficiency is low. In order to improve the coding efficiency of panoramic image, many projection methods were proposed. Cube projection is a typical method, which projects the sphere onto the cube by perspective projection [4]. Afterwards, octahedron projection (project the sphere onto the octahedron by perspective projection) [5] and icosahedron projection (ISP, project the sphere onto the icosahedron by perspective projection) [6][7] were proposed to obtain a better approximation of the sphere.

In [8], Yu et al proposed content adaptive representations of omnidirectional videos, which divide video of equirectangular format into tiles, then the tiles are down-sampled based on the content of the video. Moreover, Yu et al obtains an average division and down sampling configuration according to the test data set.

In [9], a novel tile segmentation projection was proposed, according to latitude, the author divides the sphere into 3 or 5 pieces then the polar pieces are projected into circles, and other pieces are projected into rectangles.

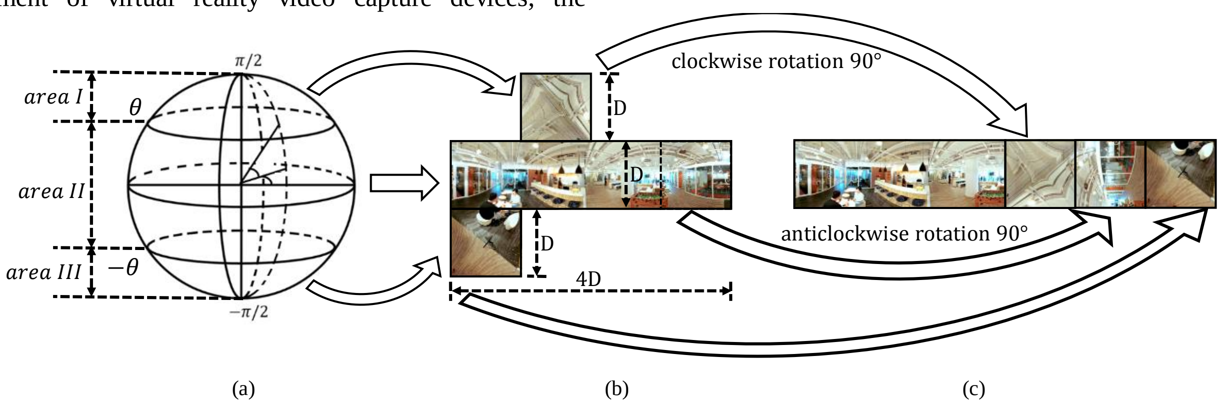


Fig.1 The schematic diagram of our proposed projection method in this paper

978-1-5386-0462-5/17/$31.00 ©2017 IEEE. VCIP 2017, Dec. 10 – 13, 2017, St Petersburg, U.S.A.

In [10], the author proposed doublering projection, which use two latitude lines ($\pm 30°$) to divide the sphere into three areas. Then the middle area is sampled at equal intervals based on longitude and latitude. The polar areas are further divided into eight sub areas, and the eight sub areas are projected onto two-dimensional planes by means of equi-angle projection, radiological transformation, and so on.

There are many studies on how to evaluate the quality of panoramic video. Since panoramic video is theoretically a spherical video, the method of evaluating 2D video quality is not suit for panoramic video. In [11], the author designed weighted SPSNR to evaluate the quality of the panoramic video, but this method does not apply to different formats or different sizes of panoramic video. In [12], Yu et al proposed two methods (S-PSNR and L-PSNR) to evaluate the quality of panoramic video. The S-PSNR weights all position on the sphere with equal weights while the L-PSNR weights all position on the sphere with different weights according to their latitudes and access frequency. Both ISO/IEC MPEG and IEEE 1857 workgroup have adopted S-PSNR as an evaluation of the quality of panoramic video.

In this paper, we proposed a projection method called polar square projection. Our method firstly divides the sphere into three areas, then the polar areas are projected to squares and the middle area is projected to a rectangle to address the over-sampling issue, finally, these three areas are combined into one rectangular. Moreover, our projection method uses the equi-angular projection and equal area projection of longitude and latitude to obtain a more uniform sampling density. Experimental results show that our proposed method obtained a better coding performance compared with some state-of-the-art methods.

The rest of this paper is organized as follows. Section 2 presents the proposed asymmetric circular projection method. Section 3 provides the experimental results to testify the effectiveness of the proposed method. At last, this paper is concluded in Section 4.

## II. PROPOSED METHOD

We divide a sphere video into three areas according to latitude. As shown in Fig.1 (a), area I is the Arctic area and its corresponding latitude is $\theta \sim 90°$, area II is the low latitude area and its corresponding latitude is $-\theta \sim \theta$, area III is the Antarctic region and its corresponding latitude is $-90° \sim -\theta$. As shown in Fig.1 (b) and Fig.1 (c), we project the polar areas to squares and the middle region is projected into rectangle, then the projected two-dimensional planes are arranged so that the contents of the two-dimensional planes can be kept as continuous as possible.

### A. Polar areas projection

As shown in Fig.2 (a), we projected the bipolar areas (area I and area III) into the rectangular areas in the two-dimensional planes, in order to make use of the area of the two-dimensional plane effectively, we have projected the rings at the same latitude on the spherical polar into a square ring on the two-dimensional plane.

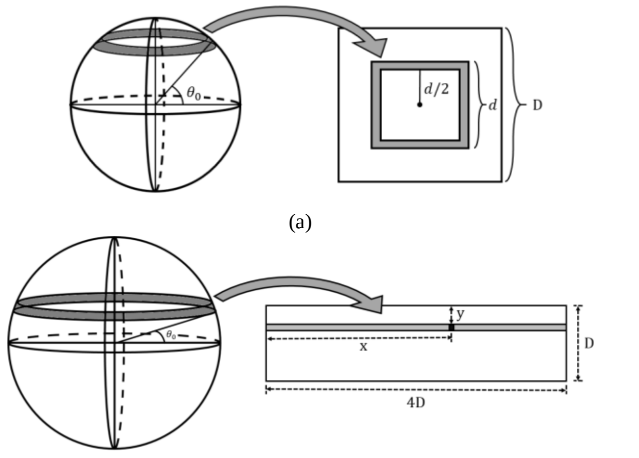


Fig.2 (a) A schematic diagram of the process of projecting spherical bipolar areas into two-dimensional planes; (b) A schematic diagram of the process of projecting the low latitude area of the sphereinto a two-dimensional plane

We obtain a more uniform sampling density by equi-angular projection of longitude and latitude. So, the relation between the latitude of the ring on the sphere and the diameter of the square ring on the two-dimensional plane is as follows:

$$\frac{d}{2} = C_1\theta_0 + C_2 \qquad (1)$$

In the Arctic region, $C_1$ and $C_2$ are solved by the following boundary condition:

$$0 = C_1\frac{\pi}{2} + C_2 \text{ and } D = C_1\theta + C_2 \qquad (2)$$

And in the Antarctic region, $C_1$ and $C_2$ are solved by the following boundary condition:

$$0 = C_1\left(-\frac{\pi}{2}\right) + C_2 \text{ and } D = C_1(-\theta) + C_2 \qquad (3)$$

### B. Low latitude region projection

As shown in Fig.2 (b), we projected the area II into a rectangular area in the two-dimensional plane. The most common form of projection is the equidistant sampling of the sphere based on the longitude and latitude (for example ERP). But the sampling density of this projection is uneven, and the sampling density in the longitude direction increases with latitude absolute value.

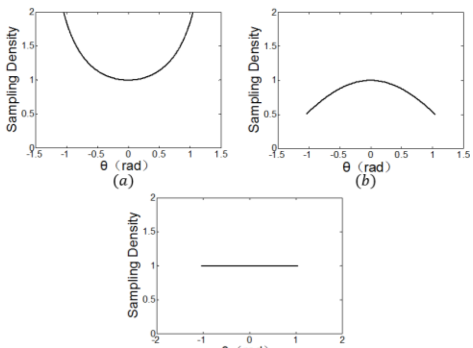


Fig.3 The variation curve of sampling density with latitude in equal area projection; (a)The variation curve of sampling density in the longitude direction; (b) The variation curve of sampling density in the latitude direction; (c) The variation curve of total sampling density

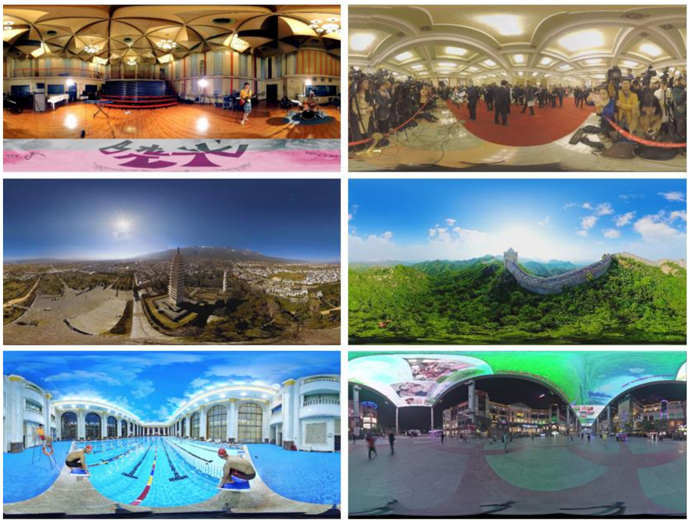

Fig.4 The sample frames of the test sequences

In order to obtain a uniform sampling density, we use the projection method of equal-area projection. The equal-area projection reduces the oversampling in the high latitude region by reducing the sampling density in latitude. As shown in Fig.3, as the absolute value of latitude increases, the sampling density of longitude increases gradually, while the latitudinal sampling density decreases gradually, so that the total sampling density remains unchanged.

The relation between the latitude ($\theta_0$) and longitude ($\varphi_0$) of the point on the sphere and the coordinate $(x, y)$ of the point the two-dimensional plane is as follows:

$$\varphi_0 = (2\frac{x}{4D} - 1)\pi \quad (4)$$

$$\theta_0 = \arcsin\left(\sin\theta(1 - 2\frac{y}{D})\right) \quad (5)$$

### C. Optimization of parameter $\theta$

In our proposed method, the coding efficiency of two-dimensional video is related to the choice of parameter $\theta$. So we optimize the parameters of $\theta$ to achieve higher coding efficiency.

In the two-dimensional plane, the area corresponding to the Arctic area accounts for one-sixth of the total area. So, we choose the appropriate parameter $\theta$ so that the area of the spherical Arctic area (area I) is one-sixth of the sphere.

The surface areas of the spherical Arctic area and sphere are calculated as follows:

$$S_{\text{spherical crown}} = 2\pi R^2(1 - sin\theta) \quad (6)$$

$$S_{sphere} = 4\pi R^2 \quad (7)$$

We make their area ratio of one to six and solve the parameters $\theta = \arcsin(2/3)$.

## III. EXPERIMENTS

We evaluate the coding efficiency of our method using the test sequence set provided by LETIN VR [13], which have been adopted as the test sequences of IEEE 1857.9-Immersive Video Content Coding[14]. The test sequence set contains 8 panorama sequences in ERP format with 4096x2048 resolution and 4 panorama sequences in ERP format with 8192x4096 resolution.

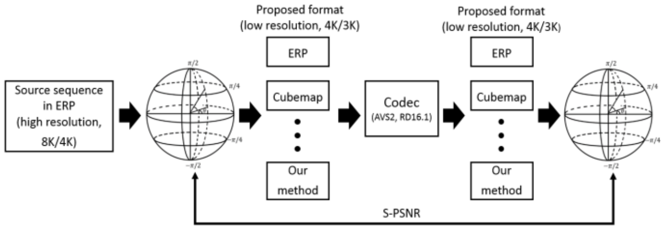


Fig.5 Experimental flow diagram

The test sequence frame rate and frame number are 30 and 300, respectively. These sequences contain a wide range of video contents, including indoor, outdoor, landscape, sports, concerts and so on. Fig.4 shows some sample frames of the test sequence set.

### A. Experimental procedures

The experimental procedure is shown in Fig.5. We use high resolution panoramic sequences (8192x4096 and 4096x2048) in ERP format as ground truth. Then we then projected them into lower resolution sequences in different projection formats, the data amount of lower resolution sequences in different projection formats are 25% of original ERP sequence in 8K case and 56.25% in 4K case. We use RD-VR16.1[14] to encode and decode the lower resolution sequences in different projection formats. Finally, we compute the S-PSNR between the decoded sequences and the original sequences.

### B. S-PSNR and RD-VR16.1

At present, S-PSNR is widely used in the evaluation of panoramic video distortion. Both IEEE 1857.9 and ISO/IEC MPEG workgroup have adopted S-PSNR as an evaluation method of the quality of panoramic video. S-PSNR calculate the PSNR of two panoramic videos on the sphere, and it can calculate the distortion of panoramic videos in different projection formats or in different resolution.

RD-VR16.1 is the reference encoding software for IEEE 1857.9-Immersive Visual Content Coding Workgroup. The RD-VR16.1 reference software adds syntax extensions of VR on the basis of the AVS2 reference software RD16.1. In the experiment, we use the default configuration of random access to encode and decode the test sequences.

### C. Experimental results

In our experiments, we use ERP as an experimental anchor. We tested the coding efficiency of cube projection, octahedron projection (OHP), icosahedron projection (ISP), doublering projection and our proposed method. The results of experiment are showed in Tab.1, and Fig. 6 gives the RD curve of the sequence Fengjing_1.

Our proposed method achieves the optimal coding efficiency in most sequences, and the average coding efficiency of our proposed method is 11.63% higher than that of ERP.

Tab.1 BD-rate comparison of various projection method

| Source resolution | Sequence | Cube projection | Octahedron projection | Icosahedron projection | Doublering projection | Our proposed method |
|---|---|---|---|---|---|---|
| 4096x2048 | Fengjing_1 | -19.81% | -22.16% | -25.40% | -25.17% | -30.20% |
| | Fengjing_3 | -3.35% | -5.55% | -14.34% | -17.35% | -28.04% |
| | Hangpai_1 | -8.09% | -3.52% | -7.16% | -9.28% | -11.37% |
| | Hangpai_2 | -14.73% | -10.33% | -12.45% | -14.88% | -13.36% |
| | Hangpai_3 | -7.08% | -3.46% | -7.87% | -9.30% | -10.74% |
| | Xinwen_1 | -0.83% | 6.38% | 3.07% | 0.25% | -1.59% |
| | Xinwen_2 | -0.37% | 10.74% | 3.88% | 3.62% | -1.13% |
| | Yanchanghui_2 | 3.73% | 8.85% | 0.64% | -1.59% | -5.98% |
| 8192x4096 | THEPLACE | 16.14% | 11.87% | 0.29% | 5.74% | 3.20% |
| | Natatorium | -1.94% | -0.23% | -9.86% | -10.26% | -12.27% |
| | Highway | -13.75% | -19.43% | -26.87% | -25.00% | -25.77% |
| | Canolafield | 6.02% | -1.83% | -4.58% | -0.15% | -4.61% |
| 4K average | | -6.31% | -2.38% | -7.45% | -9.21% | -12.52% |
| 8K average | | 1.62% | -2.40% | -10.25% | -7.42% | -9.86% |
| overall | | -3.67% | -2.39% | -8.39% | -8.61% | -11.63% |

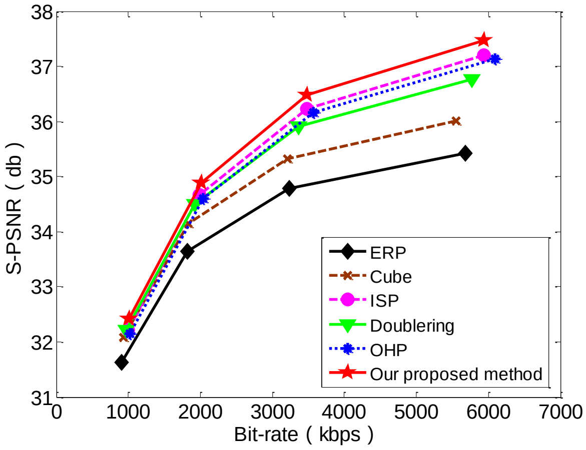


Fig.6 The RD curves of the various projection methods on sequence Fengjing_1

## IV. Conclusion

In this paper, we propose a new projection scheme. We divide the sphere of VR video into three areas according to latitude. Then, the middle and polar areas of the sphere are projected into two-dimensional planes by using equal-area and equi-angular methods respectively. Finally, we rotate and stitch the two-dimensional plane so that its video content is as continuous as possible. Experimental results testified the effectiveness of our proposed projection.

## Acknowledgment


This work was supported by Shenzhen Peacock Plan, National Natural Science Foundation of China 61672063，61370115，China 863 project of 2015AA015905, Shenzhen Research Projects of JCYJ20150331100658943 and JCYJ20160506172227337，and Guangdong Province Projects of 2014B010117007 and 2014B090910001 for funding.


## References


[1] Denis Zorin and Alan H Barr, “Correction of geometric perceptual distortions in pictures,” in Proceedings of the 22nd annual conference on Computer graphics and interactive techniques. ACM, 1995, pp. 257–264.

[2] Snyder, John P. Flattening the earth: two thousand years of map projections. University of Chicago Press, 1997.

[3] Aljoscha Smolic and David McCutchen, “3DAV exploration of video-based rendering technology in MPEG,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 14, no. 3, pp. 348–356, 2004.

[4] King-To Ng, Shing-Chow Chan, and Heung-Yeung Shum, “Data compression and transmission aspects of panoramic videos,” Circuits and Systems for Video Technology, IEEE Transactions on, vol. 15, no. 1, pp. 82–95, 2005.

[5] Hung-Chih Lin, Chia-Ying Li, Chao-Chih Huang, Jian-Liang Lin, Shen-Kai Chang, and Chi-Cheng Ju, “An efficient compact layout for octahedron projection”,IEEE1857.9-M1038/AVSVR-M1038, Guiyang, China, June, 2016

[6] Jie Chen, Jun Zhang, Vladyslav Zakharchenko, Elena Alshina, Kwang Pyo Choi, “Mapping CE: Icosahedral projection”, IEEE1857.9-M1038/AVSVR-M1038, Guiyang, China, June, 2016

[7] Jie Chen, Jun Zhang, Vladyslav Zakharchenko, Elena Alshina, Kwang Pyo Choi, “An improved layout for Icosahedral projection”, IEEE1857.9-M1089/AVSVR-M1089, Beijing, China, March, 2017

[8] Matt Yu, Haricharan Lakshman, and Bernd Girod, “Content adaptive representations of omnidirectional videos for cinematic virtual reality,” in Proceedings of the 3rd International Workshop on Immersive Media Experiences. ACM, 2015, pp. 1–6.

[9] Jisheng Li, Ziyu Wen, Sihan Li, Yikai Zhao, Bichuan Guo, Jiangtao Wen, “Novel tile segmentation scheme for omnidirectional video,” IEEE International Conference on Image Processing. IEEE, 2016, pp. 370-374.

[10] Haiwu ZHAO, Guozhong WANG, Guoping LI, Guowei TENG, Chang LIN, Xiao GU, Chengjia WU, Dapeng CHEN, Fang CHEN, Zhixian CHEN, Cheng ZHENG, Zhenglong YANG, “A new projection method of omnidirectional video”, IEEE1857.9-M0025/AVSVR-M0025, Guiyang, China, June, 2016

[11] Yule Sun, Ang Lu, Lu Yu， “WS-PSNR for VR objective quality evaluation”, IEEE1857.9-M0025/AVSVR-M0025, Dalian, China, August, 2016

[12] Matt Yu, Haricharan Lakshman, and Bernd Girod, “A framework to evaluate omnidirectional video coding schemes,” in Mixed and Augmented Reality (ISMAR), 2015 IEEE International Symposium on. IEEE, 2015, pp. 31–36.

[13] http://www.air360china.com

[14] http://www.ieee1857.org/1857.9.asp