Evaluation of H.265 and H.264 for Panoramas Video under Different Map Projections

Bijia Li[1,2], Li Song[1,2], Rong Xie[1,2], Nam Ling[3]
[1]Institute of Image Communication and Network Engineering, Shanghai Jiao Tong University
[2]Cooperative Medianet Innovation Center, Shanghai, China
[3]Department of Computer Engineering, Santa Clara University
Email: lovomi@sjtu.edu.cn, song_li@sjtu.edu.cn, xierong@sjtu.edu.cn, nling@scu.edu

***Abstract*—Various map projection methods are proposed to convert a VR omnidirectional video to a 2D planar video, including Mercator, Equirectangular, Equal-area, Cube projections. The BD-rate comparison of the different map projections has been conducted. However, the performance of the map projections has not been compared using different codec, like widely-used x265 and x264 encoders. In this paper, both x265 and x264 are used to test the performance among the map projections, while S-PSNR is set as metric to calculate the BD-rate between H.265 and H.264. Results show that the rate savings of H.265 over H.264 varies from 36.4% to 40.8% in different map projections and Equal-area has lowest bitrate savings of H.265 over H.264. What is more, this paper verifies the performance of the different map projections. Equal-area has the best performance while Mercator has intolerable performance. Last but not least, Equal-area has more bitrate savings relative to Equirectangular using x264 compared with x265.**

Keywords—VR; Map projection; BD-rate; Panoramas video; Omnidirectional video

I. INTRODUCTION

Virtual Reality (VR) applications have entered many industries including game, medical and live-streaming system with head-mounted displays (HMDs). Providing immersive experience for user is a significant reason of popularity of VR in recent years. Fig.1 illustrates the VR pipeline from video capturing to displaying. Several cameras are used to capture the wide range of ground truth scene from different angles. Thereafter, multiple different angle videos stitched together according to the overlap parts between the adjacent videos forms omnidirectional spherical VR videos. Coding the omnidirectional videos directly in the spherical domain results in loss in the video quality compared with that in the panoramas domain [1]. Therefore, various map projection methods from 3D spherical video to 2D panoramas video have been proposed to reduce the coding quality loss, e.g., cubic, Equirectangular, cylindrical [2], etc. After mapping and coding, most of the proposed methods evaluate the performance in the panoramas domain, which can not reflect the coding loss in the sphere domain. To compensate for the difference of coding loss between sphere and panoramas domain, Fu *et al*. [3] proposed a method to correct the value of PSNR by multiplying solid angle on the sphere. Thereafter, the panoramas video is inversely mapped to omnidirectional video, so that the HMDs can display the immersive 3D videos.

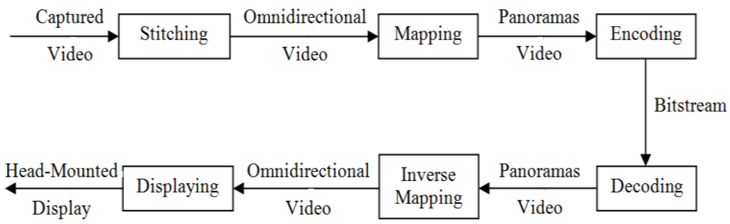


Fig.1. VR pipeline from capturing to displaying

Both H.265/HEVC and H.264/AVC coding standards are evaluated of different map projections in the experiments. The H.265/HEVC standard proposed to double the compression ratio of H.264/AVC without scarifying the visual quality. Plenty of previous works have been done to test the R-D performance between H.265/HEVC and H.264/AVC on the regular video sequences. Vanne *et al*. [4] analyzes the rate distortion complexity of High Efficiency Video Coding (HEVC) reference video codec (HM) and compares the results with AVC reference codec (JM). Nguyen *et al*. [5] evaluated the performance in intra coding of HEVC and H.264/AVC. Bitrate savings are compared of HEVC over H.264. However, since VR is a newly-developing technology, panoramas video sequences under different map projections have not been experimented to evaluate the performance of H.265 over H.264.

To address the above issue, this paper conducts experiments to evaluate performance of map projections using different codec, like widely-used x265 and x264 encoders. The remainder of the paper is structured around such comparison. Section II reviews several map projections including Mercator, Equirectangular, Equal-area and Cube projections. The evaluation methodology is introduced in Section III. In Section IV, experimental results and analysis are shown while conclusion is drawn in Section V.

II. MAP PROJECTION

Over these two years, VR map projection has become a hot research topic. Owing to the similarity between the omnidirectional video map projection and the cartography, some of the map projection methods can be directly migrated from cartography, such as Mercator, Equirectangular and Equal-area projections [6]. Fig.2 has presented the four panoramas images after mapping. As time goes on, many

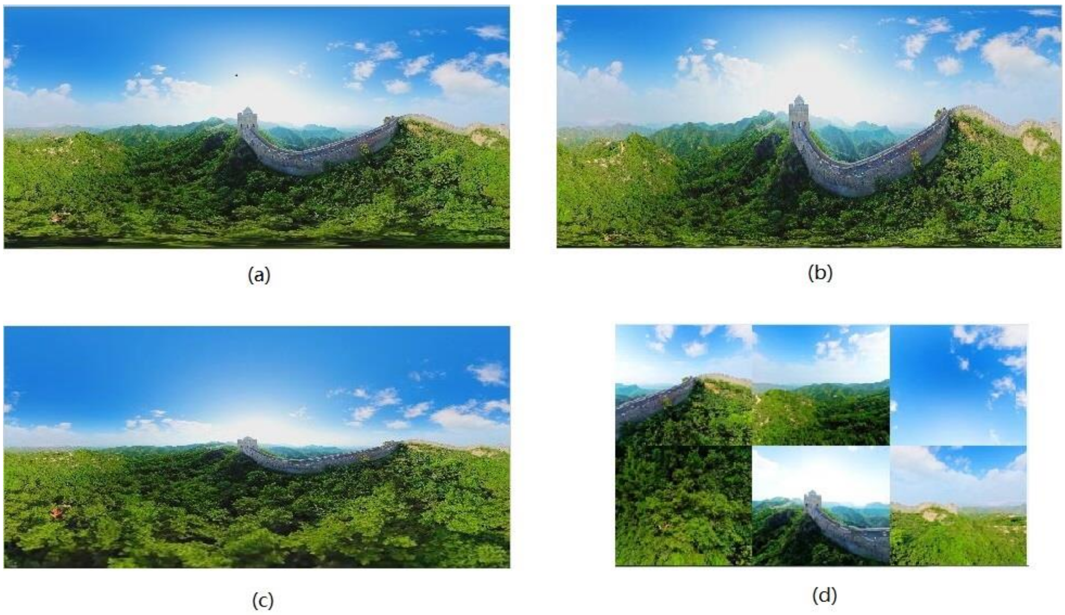


Fig.2. Demonstration of different map projections. (a).Equirectangular (b).Equal-area (c).Mercator (d).Cube

other map projection methods have been proposed, e.g., cube, dodecahedron, pyramid, etc. Four tested map projections are mainly discussed as follows, where others are simply introduced.

A. Mercator Projection

Mercator projection is a kind of cylindrical projection founded by Dutch cartography Mercator, whose basic idea is mapping the sphere to cylinder according to the equal angle rules. Mercator projection may generate great area distortion, leading to the unsatisfied visual experience. As for the image in Fig.2, the Great Wall is obviously flattened and trees are enlarged.

B. Equirectangular Projection

Equirectangular projection maps the sphere into a 2:1 rectangular based on the longitude and latitude. Its basic idea is to cut a slice of latitude and stretch it to the length of the equator. Because the slice of latitude has equal width, vertical sampling rate is uniform. Horizontal sampling rate is inconstant resulting from inconstant circumference corresponding to different latitude. Due to the limit slice length near the poles, one drawback of the Equirectangular projection is the severe distortion near the poles. Another drawback is that the area of the Equiretangular projection is 57% more than the sphere area, which means that there is almost one third of the mapping is redundant.

C. Equal-area Projection

Equal-area projection is a kind of equal-area cylindrical projection. Vertical sampling rate is set to compensate for the different horizontal sampling rate of Equirectangular so that the product of vertical and horizontal sampling rates is equal to one. Even though the distortion varies from equator to poles, the area occupied by pixels relatively stays the same, which indicates that the resolution of the video basically remains the same [6]. The Equal-area projection is thought to be a good cylindrical projection used in video coding.

D. Cube Projection

The sphere is placed internally tangent to a cubic with unit length sides in the cube projection. Six sides of the cubic contain entire information of omnidirectional videos. The intersection parts between the adjacent sides occupy smaller solid angle compared to the center parts of the sides, which leads to the greater distortion in proximity of the edge. By mapping into a panoramic view of the square, top and bottom views are compressed too much. There is more than 15% loss of pixels but 25% savings of the data.

Other than the above projections, plenty new map projection methods have been proposed in recent years. Fu *et al.* [3] proposed a rhombic dodecahedron map by subdividing the sphere into 12 rhombi which are constructed into a 4:3 panoramas video. Yu *et al.* [7] proposed tile segmentation methods by subdividing the sphere into several tiles based on latitude. RD performance is presented by contrast. Li *et al.* [8] proposed novel tile segmentation by mapping the poles into circles instead of tiles. Results show 20% BD-rate savings averagely compared to traditional Equirectangular projection.

III. Evaluation Methodology

A. Encoder Configuration

x265 [9] is an open source encoder using H.265/HEVC standard. Owing to the multi-threaded parallel acceleration, x265 encoder achieves great improvement in coding speed. Different kinds of presets are provided in x265 including placebo, slow, slower with different parameters. In order to balance between coding quality and speed, slower preset is chosen for the experiments.

x264 [10], one of the best H.264 encoders, is favored by enterprise in practical applications. The data structure of x264 has been optimized compared with JM, thus speeding up the data memory reading process. Slower preset is also chosen for the experiments.

B. VR Sequences Information

Four sequences with 4096x2048 resolutions are chosen for the test. Generally speaking, the test result has strong relationship with the video content. Temporal information (TI) and spatial information (SI) are measured in the experiments, which define inter and intra content complexities of the videos respectively [11]. TI and SI of four videos are demonstrated in TABLE I. Owing to the flashing background, sequence4 has more complex temporal and spatial information.

TABLE I. SI AND TI OF THE SEQUENCES

Sequence	Screenshot	SI	TI
Seq1		20.4	70.3
Seq2		16.0	74.0
Seq3		25.5	72.2
Seq4		26.8	85.0

C. Evaluation Metric

After mapping the omnidirectional video to 2D panoramas video, encoder is used to evaluate the coding loss of the panoramas video. Coding quality of the panoramas video is meaningless because the videos finally will be displayed on HMDs in spherical form. In this way of thinking, some metrics are proposed to evaluate the coding loss in the spherical domain instead of panoramas domain. Yu *et al*. [12] has proposed a metric called S-PSNR to compute the coding quality in the sphere domain. S-PSNR is improved by multiplying the corresponding solid angle on the sphere based on the PSNR. Higher S-PSNR usually indicates higher visual quality. Different QPs are chosen to obtain different S-PSNR values in the experiments. The average bitrate difference between the rate-distortion curves is summarized as BD-rate metric [13].

IV. EXPERIMENTS AND ANALYSIS

In this paper, experiments were conducted with both open source encoder x265 and x264 on their slower settings. Ops used for calculating the BD-rate are set to 22, 27, 32, 37, while S-PSNR is chosen as the video quality metric. Rate control for both encoders is off. Four 4096x2048 VR videos provided by LatinVR are tested in the experiments. As for convenience, Equal-area, Equirectangular, Cube and Mercator projections can be abbreviated as Eqar, Rect, Cube, Merc, respectively.

The comparison of results is divided into three parts. In the first part, the experiments show rate savings of H.265 over H.264 under different map projections, while map projection comparison is illustrated in the second part. In the third part, different rate savings among the sequences are deeply analyzed with spatial information and temporal information.

A. Codec Comparison

Subjective evaluation has concluded that H.265/HEVC obtains approximately 50% rate savings than H.264/AVC. However, the result may not be suitable when it comes to the panoramas videos using S-PSNR as metric.

TABLE II shows the BD-rate comparison of x265 relative to x264. Rate savings mean at the same video quality, proposed scheme needs much less bitrate than the anchor scheme. Overall, the average rate savings of Equal-area projection is 36.4%, while other projections vary from 39.6% to 40.8%. In the best situation, H.265 can achieve a significant coding gain up to 59.9% for Equirectangular projection. It can be found that Equal-area projection may achieve less rate savings, which can be considered to be less sensitive to the different codecs, compared to other projections.

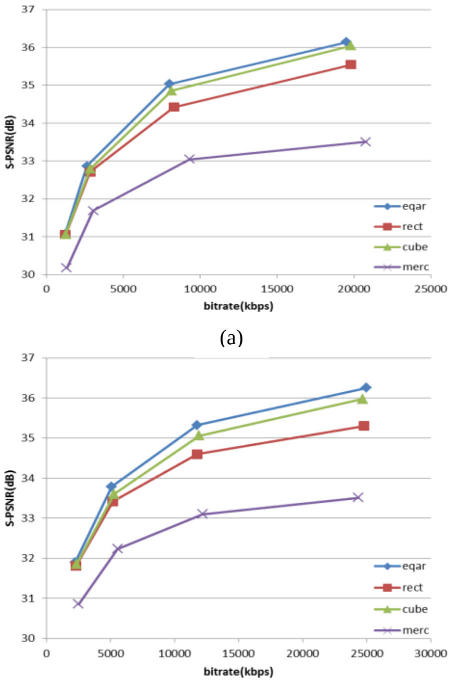
TABLE II. BD-RATE OF X265 RELATIVE TO X264

Projection	Seq1	Seq2	Seq3	Seq4	Average
Eqar	-55.8%	-22.3%	-33.2%	-34.3%	-36.4%
Rect	-59.9%	-24.0%	-33.6%	-45.6%	-40.8%
Merc	-58.3%	-25.2%	-38.6%	-38%	-40.0%
Cube	-57.9%	-27.3%	-32%	-39.6%	-39.2%

Fig.3 and Fig.4 illustrate RD-curves of four map projections in two sequences using x265 and x264 respectively. What should be paid attention to is that two sequences show the same trend of the gap between Equal-area projection and other projections. Considering the case of Equal-area and Equirectangular, when using the H.265 standard, the line connected by last two points is almost parallel between Equirectangular and Equal-area, which indicates that the gap between them stays the same as the bit rate increases. However, the gap becomes larger in the proximity of higher bit rate when using H.264. In other words, Equal-area projection is more adaptable to H.264 and has significantly better performance for H.264.

B. Map Projection Comparison

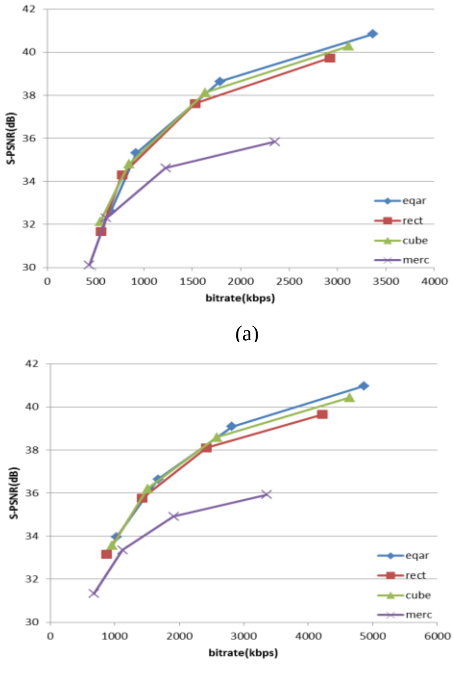
In this section, four map projections including Equal-area, Equirectangular, Cube and Mercator projections are experimented. RD-curve of sequence2 coded by x265 is presented in Fig.3 (a), which plots the S-PSNR values against bit rate of each projection. Generally speaking, S-PSNR values increase with the bit rate for the four projections. Obviously seen from the figure, a large gap is formed between the Mercator and other projections, which indicates that Mercator projection generates intolerant distortion. As for Equal-area



(a)

(b)

Fig.3. RD curves of sequence2 for different projections using x265 and x264. (a). x265 (b). x264



(a)

(b)

Fig.4. RD curves of sequence2 for different projections using x265 and x264. (a). x265 (b). x264

projection, S-PSNR value climbs rapidly to 33dB. Thereafter, S-PSNR increases slowly to 35dB, before reaching a peak at 36dB. For Equirectangular and Cubic, the figures are lower but show a similar trend throughout the bit rate axis. It is worth mentioning that the S-PSNR gap between Equal-area and other projections becomes larger as the bit rate increases. It can be deduced that Equal-area projection performs better at higher bit rate.

TABLE III and TABLE IV list the BD-rate of Equal-area and Cube relative to the Equirectangular with x265 and x264 encoders respectively. Because the performance of the Mercator projection is far below other projections, the result is not listed in the table. On the whole in TABLE III, Equal-area achieves about 10.9% rate savings, while Cube gains 7.4% rate savings compared to the Equirectangular projection. TABLE IV shows the similar result as x265 with slight fluctuations. It can be observed that the Equal-area projection has higher bitrate savings than other projections using x264, which verifies the above results.

TABLE III. BD-RATE OF MAP PROJECTIONS USING X265

	Eqar	Cube
Seq1	-4.5	-0.9
Seq2	-21.9	-13.2
Seq3	-13.1	-11.7
Seq4	-3.9	-3.7
Avg	-10.9	-7.4

TABLE IV. BD-RATE OF MAP PROJECTIONS USING X264

	Eqar	Cube
Seq1	-4.5	-1.1
Seq2	-23.9	-13.7
Seq3	-15.8	-11.1
Seq4	-4.9	-4.0
Avg	-12.3	-7.5

C. *Content Analysis*

The spatial information of the sequence plays a crucial role in rate savings among the projections. S-PSNR tends to be more sensitive to sequences' spatial complexity than temporal complexity. As for the low level of spatial information sequence2 in TABLE III, Equal-area projection obtains peak rate savings at 21.9%. On the contrary, rate savings value plunges to a low of 3.9% when meeting the high SI value sequence4. High level of the spatial information exerts pressure on the encoder, reducing the effects of the different methods of mapping, thus the rate savings value decreases. On the other hand, scenes with simple detail mainly depend on the differences among the map projections, as it will not push hard on the encoder. To sum up, there is a threshold of the value of spatial information, exceeding which encoder will dominate the performance instead of the discrepancy among the map projections.

V. Conclusion

In this paper, both x265 and x264 encoders are used to test the performance of several map projections. Firstly, Equal-area has lowest bitrate savings of H.265 over H.264. What is more, Equal-area has better performance for H.264. Secondly, map comparison indicates that Equal-area has best performance among the tested projections and gains approximately 12.3% bitrate savings over Equirectangular, while Mercator has significantly worse performance. Thirdly, Equal-area and Cube projections achieve considerable rate savings relative to Equirectangular in less complex spatial information sequences.

Acknowledgment

This work was supported by NSFC (61521062, 61527804, 61420106008), the 111 Project (B07022 and Sheitc No.150633) and the Shanghai Key Laboratory of Digital Media Processing and Transmissions.

References

[1] I. Bauermann, M. Mielke, and E. Steinbach, "H.264 based coding of omnidirectional video," in Proceedings of IEEE International Conference on Computer Vision and Graphics, pp. 209-215, 2004.

[2] C. Gruenheit, A. Smolic, and T.Wiegand, "Efficient representation and interactive streaming of high-resolution panoramic views." in Proceedings of IEEE International Conference on Image Processing, pp. 209-212, 2002.

[3] C. Fu, L. Wan, T. Wong, and C. Leung, "The Rhombic Dodecahedron Map: An Efficient Scheme for Encoding Panoramic Video," IEEE Trans. Mutimedia, vol. 11, no. 4, 2009.

[4] J. Vanne, M. Viitanen, and T.D. Hamalainen, "Comparative rate-distortion-complexity analysis of HEVC and AVC video codecs," IEEE Trans. Circuits and Systems for Video Technology, vol. 22, no. 12, pp.1885-1898, 2012.

[5] T. Nguyen, and D. Marpe, "Performance analysis of HEVC-based intra coding for still image compression," in Proceedings of IEEE Picture Coding Symposium, pp. 233-236, 2012.

[6] A. Smolic and D. McCutchen, "3DAV exploration of video-based rendering technology in MPEG," IEEE Trans. Circuits and Systems for Video Technology, vol. 14, no. 3, pp. 348–356, 2004.

[7] M. Yu, H. Lakshman, and B. Girod, "Content adaptive representations of omnidirectional videos for cinematic virtual reality," in Proceedings of ACM the 3rd International Workshop on Immersive Media Experiences, pp. 1-6, 2015.

[8] J.Li, Z.Wen, S.Li, Y.Zhao, B.Guo, and J.Wen, "Novel tile segmentation scheme for omnidirectional video," in Proceedings of IEEE International Conference on Image Processing, 2016.

[9] x265 Latest Reference Software [Online]. Website: http://x265.org/

[10] x264 Latest Reference Software [Online]. Website: http://www.videolan.org/developers/x264.html

[11] ITU-T, "Subjective video quality assessment methods for multimedia applications," Recommendation ITU-R P 910, Sep.1999.

[12] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omnidirectional video coding schemes," in Proceedings of IEEE International Symposiumon Mixed and Augmented Reality, pp. 31–36, 2015.

[13] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T VCEG-M33, Apr. 2001.