A NOVEL RATE CONTROL SCHEME FOR PANORAMIC VIDEO CODING

Yufan Liu†, Mai Xu†, Chen Li†, Shengxi Li†, and Zulin Wang†‡*

†School of Electronic and Information Engineering, Beihang University, China
‡Collaborative Innovation Center of Geospatial Technology, Wuhan, China

ABSTRACT

The popularity of multi-view panoramic videos has been considerably increased for producing Virtual Reality (VR) content, due to its immersive visual experience. We argue in this paper that PSNR is less effective in assessing visual quality of compressed panoramic videos than Sphere-based PSNR (S-PNSR), in which sphere-to-plain mapping of panoramic videos is considered. Thus, the conventional rate control (R-C) schemes of 2-Dimensional (2D) video coding, which optimize on PSNR, are not suitable for panoramic video coding. To optimize S-PSNR , we propose in this paper a novel RC scheme for panoramic video coding. Specifically, we develop an S-PSNR optimization formulation with constraint on bit-rate. Then, a solution is provided to the developed formulation, such that bits can be allocated to each coding block for achieving optimal S-PSNR in panoramic video coding. Finally, the experiment results validate the effectiveness of the proposed RC scheme in improving S-PSNR of panoramic video coding.

Index Terms— Virtual reality, panoramic video, rate control

1. INTRODUCTION

Panoramic videos, as a form of Virtual Reality (VR) content, offer 360-degree viewing direction when playing videos. Recently, panoramic video has been increasingly popular due to its immersive visual experience. On the other hand, high resolution (4K or beyond) is required in panoramic video for pleasant visual experience, such that tremendous communication bandwidth is consumed by panoramic video. This may cause bandwidth-hungry issue in communications. Therefore, there is a desire demand for improving compression efficiency of panoramic video.

During the past decade, several approaches [1–12] have been proposed to enhance compression efficiency of panoramic video. For example, in the early time, user-driven interactive compression [1–3] was proposed for panoramic video coding. In particular, the panoramic video is compressed into multiple tile-based streams, and then a specific set of streams is transmitted, decoded and rendered, according to the user's present viewpoint at the decoder side. Although there exist a couple of advanced user-driven approaches in recent years [4, 5], user-driven approaches are not maturative for practical applications. It is because there exists large transmission delay on the feedback of the user's present viewpoint. Recently, encoder optimization approaches [6, 7, 11] have been proposed for panoramic video coding. For instance, Zheng *et al.* [6,7] proposed a novel motion compensated prediction approach to improve the quality of compressed panoramic video. Besides, Yu *et al.* [11] presented an H.264-complaint coding optimization approach, which adaptively adjusts different sizes and bit-rates of various tiles to enhance the overall quality of panoramic video. Instead of optimizing the encoder, some preprocessing approaches [8–10] have been proposed for panoramic video coding. In [8], Budagavi *et al.* proposed to utilize region adaptive smoothing before encoding panoramic videos, so that more bits can be saved in the regions close to poles which are rare to be viewed. Similarly, Youvalari [10] proposed a regional down-sampling preprocessing approach to decrease the extra bit-rate caused by over-stretched polar areas. However, the preprocessing approaches cannot control bit-rate for panoramic video coding, and thus compressed panoramic videos can hardly resume original quality at sufficient bit-rates. In fact, rate control (R-C) can be applied to optimize visual quality of compressed panoramic video at a given target bit-rate.

The existing 2-Dimensional (2D) video coding standards, such as the latest High Efficient Video Coding (HEVC) [13], have been used for panoramic video coding, and their R-C scheme may be applied as well. Most recently, the R-λ scheme [14] has been proposed for the latest HEVC, as the state-of-the-art RC scheme. The R-λ scheme performs well in both rate distortion (RD) optimization and control accuracy. Unfortunately, the RD optimization of the R-λ scheme mainly focuses on improving PSNR for 2D video coding, while Sphere-based PSNR (S-PSNR) [15] is more effective in measuring visual quality for panoramic video coding. Thus, it is necessary to control bit-rate of panoramic video coding, with optimization on S-PSNR. To the best of our knowledge, there is no RC work on optimizing S-PSNR for panoramic video coding.

In this paper, we propose an RC scheme for panoram-

*Corresponding author: Mai Xu (maixu@buaa.edu.cn). This work was supported by the NSFC projects under Grants 61573037, 61202139, and 61471022, and Fok Ying-Tong education foundation under grant 151061.

ic video coding, which optimizes S-PSNR at a given bit-rate. First, we evaluate subjective quality of some compressed panoramic videos by Differential Mean Opinion Score (D-MOS). Next, we find that DMOS is more correlated to S-PSNR than PSNR. Thus, S-PSNR is used as the optimization objective in our RC scheme. Accordingly, we propose a formulation to optimize S-PSNR with constraint on the target bit-rate, based on the R-λ RC scheme. In our formulation, the non-uniform sampling of sphere-to-plain map of S-PSNR is taken into account. Then, we develop a solution to our formulation, such that bits are optimally assigned to each Coding Tree Unit (CTU) according to S-PSNR. In summary, the main contributions of this paper are as follows: (1) We prove that S-PSNR is an effective metric for evaluating visual quality of panoramic video coding. (2) We propose the RC formulation for optimizing S-PSNR in panoramic video coding, with a solution achieving optimal CTU-level bit allocation.

2. PRELIMINARY

Since our RC scheme mainly concentrates on optimizing S-PSNR, we briefly review S-PSNR in Section 2.1, and verify its effectiveness in Section 2.2.

2.1. Overview of S-PSNR

S-PSNR aims at measuring distortion of impaired panoramic video given its reference video. Different from PSNR, S-PSNR is based on mean square error (MSE) at sphere domain, such that sphere-to-plain projection needs to be employed in calculating S-SPNR. Figure 1 shows the sphere-to-plain projection for both reference and impaired sequences, in which pixels at plain domain are non-uniformly sampled for computing pixel-wise S-PSNR.

To be more specific, we assume that pixels on the sphere are denoted as $\{\mathbf{g_n}\}_{n=1}^{N}$, where N is total number of pixels in the sphere of the panoramic video. Then, $\mathbf{g_n}$ on the sphere is projected to location (x_n, y_n) in the reference and impaired sequences. Given (x_n, y_n), S-PSNR can be computed based on the averaged sphere-based MSE (S-MSE) as follows,

$$\begin{aligned} \text{S-MSE} &= \frac{\sum_{n=1}^{N}(S(x_n,y_n)-S'(x_n,y_n))^2}{N}, \\ \text{S-PSNR} &= 10\log_{10}\frac{255^2}{\text{S-MSE}}, \end{aligned} \tag{1}$$

where $S(x_n, y_n)$ and $S'(x_n, y_n)$ are pixel values of the n-th pixel projected to reference and impaired sequences, respectively. However, the sphere-to-plain projection may make (x_n, y_n) as sub-pixels without knowing its pixel values. To solve such an issue, the *Lanczos* interpolation is used in [16] to estimate $S(x_n, y_n)$ and $S'(x_n, y_n)$, according to the values of its neighboring pixels. Finally, S-PSNR can be obtained by (1), with interpolated values of $\{(x_n, y_n)\}_{n=1}^{N}$.

2.2. Verification of S-PSNR

To verify the effectiveness of S-PSNR, we compare the performance of PSNR and S-PSNR, via evaluating their correlation with subjective quality over several panoramic video sequences. For comparison, 8 panoramic video sequences in raw format are selected from the standard test set of IEEE 1857 working group [17]. These sequences are compressed at 500 Kbps, 3Mbps and 100 Mbps. Then, PSNR and S-PSNR are measured for compressed sequences, with regard to their reference raw sequences.

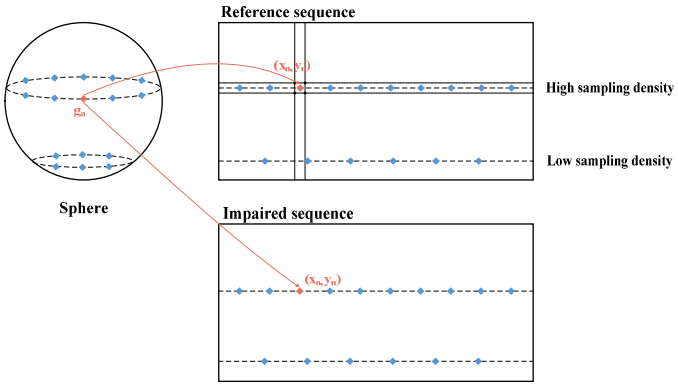


Fig. 1: Illumination of sphere-to-plain projection for reference and impaired sequences. Such projection leads to non-uniform sampling for computing S-PSNR in plain domain.

Next, we follow the way of [18] to conduct the subjective test, for quantifying subjective quality of those compressed sequences in terms of DMOS. In our subjective test, there are int total 17 subjects (10 males and 7 females, aging from 19 to 30) involved in rating raw quality scores of all sequences. Here, the VR headset *Oculus Rift DK2* is utilized to display the test panoramic videos at their original resolution. In addition, we leverage *LiveViewRift* as the Oculus Rift video player, with its default projection type of sphere. Finally, the D-MOS value of each compressed sequence can be obtained on the basis of their rated subjective scores, as reported in Table 1. Table 1 also tabulates PSNR and S-PSNR of each sequence.

Now, we measure the correlation between PSNR/S-PSNR and DMOS for each compressed sequence, using Pearson Correlation Coefficient (PCC), Spearman Rank Order Correlation Coefficient (SROCC), Root MSE (RMSE), and Mean Absolute Error (MAE)[1]. The correlation results, over all compressed sequences, are shown in Table 2. As observed from this table, S-PSNR is more correlated with DMOS, compared to PSNR. Besides, the error between DMOS and S-PSNR is smaller than those between DMOS and PSNR. Hence, S-PSNR is a more reasonable objective metric than PSNR, validating the effectiveness of S-PSNR.

3. PROPOSED RATE CONTROL SCHEME

The main goal of RC in panoramic video coding is to maximize S-PNSR at a given bit-rate, since the above section verifies the effectiveness of S-PSNR. Based on the definition of S-PSNR in (1), the sphere-based distortion is the sum of square

[1]Since DMOS and PSNR/S-PSNR have negative correlation, we use $(100 - \text{DMOS})$ instead of DMOS, to compute the RMSE and MAE.

Table 1: Subjective and objective metrics of sequences compressed at different bit-rates.

Name	Bit-rate	DMOS	PSNR	S-PSNR	Name	Bit-rate	DMOS	PSNR	S-PSNR	Name	Bit-rate	DMOS	PSNR	S-PSNR
Fengjing_1	500kbps	60.04	30.89	31.02	*Tiyu_1*	500kbps	57.54	36.35	36.01	*Yanchanghui_2*	500kbps	57.31	33.33	32.79
	3Mkbps	46.33	38.58	39.05		3Mkbps	37.48	42.93	42.94		3Mkbps	39.68	42.11	41.86
	10Mkbps	37.86	41.22	41.86		10Mkbps	35.37	45.18	45.32		10Mkbps	36.02	46.21	46.05
Dianying	500kbps	59.97	41.02	39.70	*Hangpai_1*	500kbps	68.89	27.19	26.82	*Hangpai_2*	500kbps	75.44	22.83	23.01
	3Mkbps	38.59	46.84	45.70		3Mkbps	42.02	35.57	35.39		3Mkbps	59.90	27.66	27.79
	10Mkbps	36.57	49.32	48.35		10Mkbps	35.36	39.92	39.85		10Mkbps	43.51	32.42	32.51
AerialCity	500kbps	64.53	30.79	30.17	*DrivingInCountry*	500kbps	71.83	26.47	25.69	Average	500kbps	66.64	39.28	38.71
	3Mkbps	39.66	37.92	37.74		3Mkbps	48.15	31.61	31.02		3Mkbps	49.54	45.22	44.77
	10Mkbps	36.26	39.13	39.18		10Mkbps	39.57	35.13	34.72		10Mkbps	44.08	47.94	47.62

Table 2: Correlation between DMOS and objective metrics.

Metrics	PCC	SROCC	RMSE	MAE
S-PSNR	-0.8526	-0.8526	0.1147	0.0935
PSNR	-0.8369	-0.8363	0.1210	0.1000

error between pixels sampled from sphere:

$$d_m = \sum_{n \in \mathbf{C}_m} (S(x_n, y_n) - S'(x_n, y_n))^2, \tag{2}$$

where $\mathbf{C}_m$ is the set of pixels belonging to the m-th CTU. Therefore, at target bit-rate R, the optimization on S-PSNR can be formulated by

$$\min(\sum_{m=1}^{M} d_m) \qquad \text{s.t.} \sum_{m=1}^{M} r_m = R. \tag{3}$$

In (3), r_m is the assigned bits at the m-th CTU, and M is the total number of CTUs in the current frame. To solve the above formulation, a Lagrange multiplier λ is introduced, and (3) can be converted to an unconstrained optimization problem:

$$\min_{\{r_m\}_{m=1}^{M}} \quad J = \sum_{m=1}^{M} (d_m + \lambda r_m). \tag{4}$$

Here, we define J as the value of RD cost. By setting derivative of (4) to zero, minimization on J can be achieved by

$$\begin{aligned} \frac{\partial J}{\partial r_m} &= \frac{\partial (\sum_{m=1}^{M} (d_m + \lambda r_m))}{\partial r_m} \\ &= \frac{\partial d_m}{\partial r_m} + \lambda \\ &= 0. \end{aligned} \tag{5}$$

Next, we need to model the relationship between distortion d_m and bit-rate r_m, for solving (5). Note that d_m and r_m are equivalent to S-MSE and bit per pixel (bpp) divided by the number of pixels in a CTU, respectively. Similar to [14], we use the Hyperbolic model to investigate the relationship between sphere-based distortion S-MSE and bit-rate bpp, on the basis of four encoded panoramic video sequences. Figure 2 plots the fitting RD curves using the Hyperbolic model, for these four sequences. In this figure, bpp is calculated by

$$\text{bpp} = \frac{R}{f \times W \times H}, \tag{6}$$

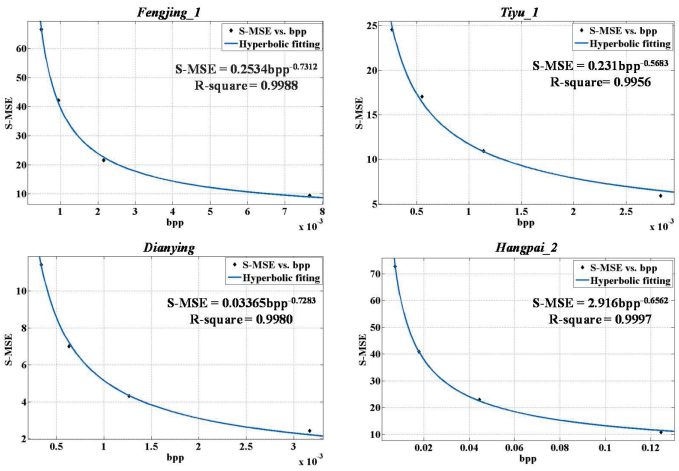


Fig. 2: R-D fitting curves using the hyperbolic model. Note that these four sequences are encoded by HM 15.0 with the default Low Delay P profile. The bit-rates are set as the actual bit-rates when compressing at 4 fixed QP (27, 32, 37, 42), to be described in Section 4.1.

where f means frame rate, and W and H stand for width and height of video, respectively. Figure 2 shows that the Hyperbolic model is capable of fitting on the relationship between S-MSE and bpp, and R-square for the fitting curves of four sequences are all more than 0.99. Therefore, the Hyperbolic model is used in our RC scheme as follows,

$$d_m = c_m \cdot (r_m)^{-k_m}, \tag{7}$$

where c_m and k_m are the parameters of the Hyperbolic model that can be updated for each CTU using the same way as [19].

The above equation can be rewritten by

$$-\frac{\partial d_m}{\partial r_m} = c_m \cdot k_m \cdot r_m^{(-k_m - 1)}. \tag{8}$$

Given (5) and (8), the following equation holds:

$$r_m = (\frac{c_m k_m}{\lambda})^{(\frac{1}{k_m + 1})}. \tag{9}$$

Moreover, according to (3), we have the following constraint:

$$\sum_{m=1}^{M} r_m = R. \tag{10}$$

Upon (9) and (10), the bit allocation for each CTU can be formulated as follows,

$$\sum_{m=1}^{M} r_m = \sum_{m=1}^{M} (\frac{c_m \cdot k_m}{\lambda})^{(\frac{1}{k_m + 1})} = R. \tag{11}$$

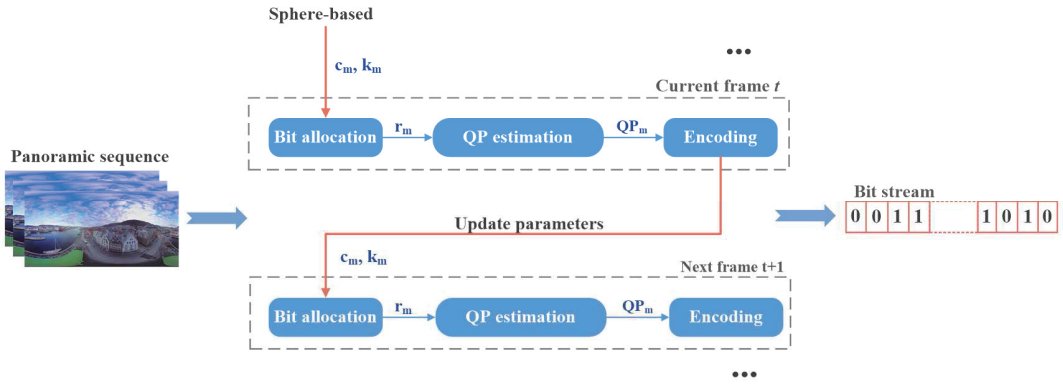


Fig. 3: The framework of the proposed RC scheme for panoramic video coding.

Therefore, once (11) is solved, target bit r_m can be obtained for each CTU, with maximization on S-PSNR. In this paper, we apply the latest Recursive Taylor Expansion (RTE) method [19] in solving (11) with the closed-form solution. Note that the computational speed of the RTE method [19] is rather fast (0.0015 ms for a CTU), such that our RC scheme takes little overhead on computational cost.

After obtaining the optimal bit-rate allocation, Quantization Parameter (QP) of each CTU can be estimated using the method of [14]. Figure 3 summarizes the overall procedure of our RC scheme for panoramic video coding. Note that our RC scheme is mainly applicable for the latest HEVC-based panoramic video coding, and it can be extended to other video coding standards by re-investigating the Hyperbolic model of bit-rate and distortion.

4. EXPERIMENT

In this section, experiments are conducted to validate the effectiveness of our RC scheme. Section 4.1 presents the settings for our experiments. Section 4.2 evaluates our approach from aspects of RD performance, Bjontegaard delta bit-rate (BD-rate) and Bjontegaard delta S-PSNR (BD-PSNR). Section 4.3 discusses the RC accuracy of our scheme.

4.1. Settings

Due to space limitation, 8 panoramic video sequences at 4K are chosen from the test set of IEEE 1857 working group [17], in our experiments. They are shown in Figure 4. These sequences are all at 30 fps with duration of 10 seconds. Figure 4 shows that the contents of these sequences, which vary from indoor to outdoor scenes and contain people and landscapes. Then, these panoramic video sequences are compressed by the HEVC reference software HM-15.0. Here, we implement our RC scheme in HM-15.0, and then compare our scheme with the latest R-λ RC scheme [14] that is default RC setting of HM-15.0. For HM-15,0 the Low Delay P setting is applied with the configuration file *encoder_lowdelay_P_main.cfg*. The same as [14], we first compress panoramic video sequences using the conventional HM-15.0 at four fixed QPs, which are 27, 32, 37, and 42. Then, the obtained bit-rates are used to set the target bit-rates of each sequence for both our and conventional [14] schemes. It is worth pointing out that we only compare with the state-of-the-art RC scheme [14] of HEVC for 2D video coding, since there exists no RC scheme for panoramic video coding.

4.2. Evaluation on RD performance

RD curves. We compare the RD performance of our and the conventional RC [14] schemes using S-PSNR in Y channel, since Section 2.2 has proved that S-PSNR is more effective than PSNR in evaluating the objective quality of panoramic videos. We plot in Figure 5 the RD curves of all test panoramic video sequences, for both our and the conventional RC schemes. We can see from these RD curves that our scheme achieves higher S-PSNR than [14] at the same bit-rates, for all test sequences. Thus, our RC scheme is superior to [14] in RD performance.

BD-PSNR and BD-rate. Next, we quantify RD performance in terms of BD-PSNR and BD-rate. Similar to the above RD curves, we use S-PSNR in Y channel for measuring BD-PSNR and BD-rate. Table 4 reports the BD-PSNR improvement of our scheme over [14]. As can be seen from this table, our scheme averagely improves 0.1613 dB in BD-PSNR over [14]. Such improvement is mainly because our scheme aims at optimizing S-PSNR, while [14] deals with optimization on PSNR. Table 4 also tabulates the BD-rate saving of our RC scheme with [14] being an anchor. We can see that our RC scheme is able to save 5.34% BD-rate in average, when compared with [14]. Therefore, our scheme has potential in relieving the bandwidth-hungry issue posed by panoramic videos.

Subjective quality. Furthermore, Figure 6 shows visual quality of one selected frame of sequence *Dianying*, encoded by HM-15.0 with our and conventional RC schemes at the same bit-rate. We can observe that our scheme yields better visual quality than [14], with smaller blurring effect and less artifacts. For example, both regions of fingers and light generated by our scheme is much more clear than those by [14]. Besides, the region of the leg encoded with our RC scheme has less blurring effect, compared to [14]. In summary, our scheme outperforms [14] in RD performance, evaluated by RD curves, BD-PSNR, BD-rate and subjective quality.

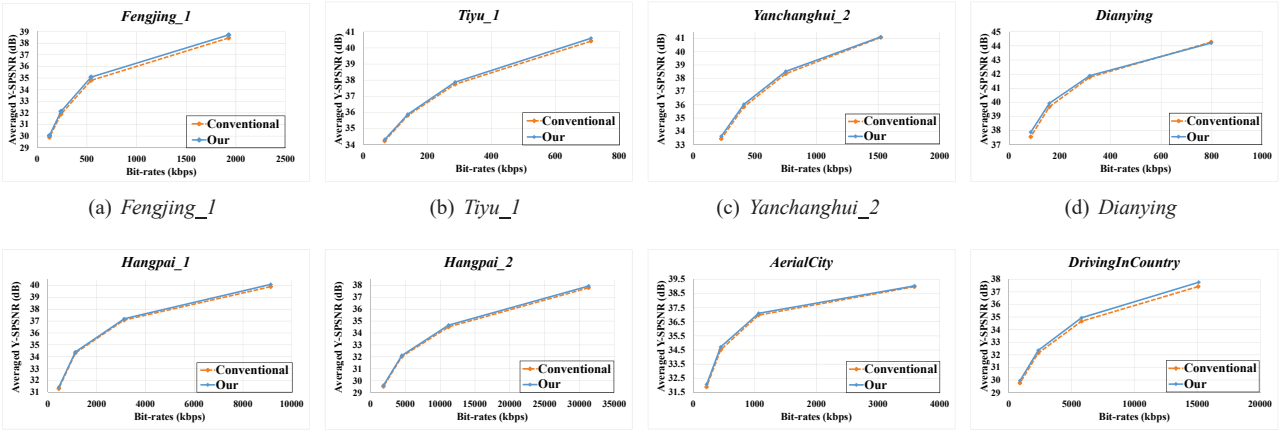
4.3. Evaluation on RC accuracy

Now, we evaluate the RC accuracy of our scheme. For such evaluation, Table 3 illustrates the error rate of actual bit-rate

(a) *Fengjing_1* (4096x2048) (b) *Tiyu_1* (4096x2048) (c) *Yanchanghui_2* (4096x2048) (d) *Dianying* (4096x2048)

(e) *Hangpai_1* (4096x2048) (f) *Hangpai_2* (4096x2048) (g) *AerialCity* (3840x1920) (h) *DrivingInCity* (3840x1920)

Fig. 4: Selected frames from all test panoramic video sequences.



(a) *Fengjing_1* (b) *Tiyu_1* (c) *Yanchanghui_2* (d) *Dianying*

(e) *Hangpai_1* (f) *Hangpai_2* (g) *AerialCity* (h) *DrivingInCountry*

Fig. 5: RD curves of all test sequences compressed by HM-15.0 with our and conventional RC [14] schemes.

Table 3: S-PSNR improvement and RC accuracy of our RC scheme, compared with the conventional scheme [14].

Name	Fixed QP	RC error (‰) (conventional)	RC error (‰) (our)	S-PSNR (dB) improvement	Name	Fixed QP	RC error (‰) (conventional)	RC error (‰) (our)	S-PSNR (dB) improvement	Name	Fixed QP	RC error (‰) (conventional)	RC error (‰) (our)	S-PSNR (dB) improvement
Fengjing_1	27	0.07	0.12	0.27	*Tiyu_1*	27	0.04	0.04	0.18	*Yanchanghui_2*	27	0.74	0.34	0.03
	32	0.06	0.05	0.28		32	0.46	0.08	0.13		32	0.34	0.52	0.17
	37	0.32	0.20	0.23		37	0.01	1.98	0.07		37	0.25	0.96	0.19
	42	0.31	0.23	0.15		42	1.37	3.02	0.10		42	0.54	1.89	0.20
Dianying	27	0.02	1.68	-0.07	*Hangpai_1*	27	0.05	0.19	0.19	*Hangpai_2*	27	0.02	0.18	0.15
	32	0.00	2.00	0.11		32	0.10	0.29	0.11		32	0.45	0.27	0.14
	37	0.26	0.49	0.25		37	0.06	1.46	0.09		37	0.42	0.50	0.09
	42	0.70	0.04	0.32		42	0.12	0.10	0.12		42	0.01	0.78	0.09
AerialCity	27	0.17	0.95	0.06	*DrivingInCountry*	27	0.02	2.75	0.34	Average	27	0.14	0.78	0.14
	32	0.28	0.74	0.12		32	0.04	0.24	0.27		32	0.21	0.52	0.17
	37	0.06	1.18	0.20		37	0.02	0.52	0.21		37	0.17	0.91	0.17
	42	0.09	4.43	0.18		42	0.05	1.24	0.19		42	0.40	1.47	0.17
Overall average		0.23	0.92	0.16										

Table 4: BD-rate saving and BD-PSNR enhancement for each test panoramic video sequence.

Name	*Fengjing_1*	*Tiyu_1*	*Yanchanghui_2*	*Dianying*	*Hangpai_1*	*Hangpai_2*	*AerialCity*	*DrivingInCountry*	Average
BD-rate saving	-7.63%	-4.39%	-3.96%	-4.81%	-3.87%	-4.04%	-5.41%	-8.63%	-5.34%
BD-PSNR (dB)	0.2527	0.1155	0.1619	0.1441	0.1143	0.1197	0.1356	0.2464	0.1613

with respect to target bit-rate, for both our and the conventional RC [14] schemes. We can see from this table that the average RC error rate is less than 1‰, comparable to the error rate of [14]. Besides, the maximum error rate for our RC scheme is 3.02‰for sequence *Tiyu_1*, and the error rate of [14] is up to 1.37‰. Although the RC accuracy of our scheme is smaller than [14], it is rather high and very close to 100% accuracy. Therefore, our scheme is effective and practical for controlling bit-rate of HEVC-based panoramic video coding. More importantly, our RC is capable of improving RC performance for panoramic video coding.

5. CONCLUSION

In this paper, we have proposed a novel RC scheme for panoramic video coding, which minimizes the sphere-based distortion at a target bit-rate. Specifically, we first verified the effectiveness of the latest sphere-based distortion measurement metric S-PSNR, by comparing with PSNR. Next, our

(a) Conventional RC scheme

(b) Our scheme

Fig. 6: Visual quality of *Dianying* compressed at 158 Kbps by HM-15.0 with our and conventional RC [14] schemes. Note that this figure shows the 68-th frame of compressed *Dianying*.

RC scheme was proposed to optimize S-PSNR, rather than PSNR, for panoramic video coding. As such, the coding efficiency of HEVC-based panoramic video coding, evaluated by S-PSNR can be improved, leading to better subjective quality. Finally, experimental results showed that our scheme is better than the state-of-the-art R-λ RC scheme, with averagely 0.16 dB S-PSNR improvement and 5.34% bit-rate saving.

6. REFERENCES

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

[2] S Heymann, A Smolic, K Mueller, Y Guo, J Rurainsky, P Eisert, and T Wiegand, "Representation, coding and interactive rendering of high-resolution panoramic images and video using mpeg-4," in *Proc. Panoramic Photogrammetry Workshop (PPW)*, 2005.

[3] Hideaki Kimata, Shinya Shimizu, Yutaka Kunita, Megumi Isogai, and Yoshimitsu Ohtani, "Panorama video coding for user-driven interactive video application," in *IEEE 13th International Symposium on Consumer Electronics*. IEEE, 2009, pp. 112–114.

[4] Vamsidhar Reddy Gaddam, Michael Riegler, Ragnhild Eg, Carsten Griwodz, and Pål Halvorsen, "Tiling in interactive panoramic video: Approaches and evaluation," *IEEE Transactions on Multimedia*, vol. 18, no. 9, pp. 1819–1831, 2016.

[5] Alireza Zare, Alireza Aminlou, Miska M Hannuksela, and Moncef Gabbouj, "Hevc-compliant tile-based streaming of panoramic video for virtual reality applications," in *Proceedings of the ACM on Multimedia Conference (ACM MM)*. ACM, 2016, pp. 601–605.

[6] Jiali Zheng, Yanfei Shen, Yongdong Zhang, and Guangnan Ni, "Adaptive selection of motion models for panoramic video coding," in *IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2007, pp. 1319–1322.

[7] Zheng Jiali, Zhang Yongdong, Shen Yanfei, and Ni Guangnan, "Panoramic video coding using affine motion compensated prediction," in *Multimedia Content Analysis and Mining*, pp. 112–121. Springer, 2007.

[8] Madhukar Budagavi, John Furton, Guoxin Jin, Ankur Saxena, Jeffrey Wilkinson, and Andrew Dickerson, "360 degrees video coding using region adaptive smoothing," in *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 750–754.

[9] Jisheng Li, Ziyu Wen, Sihan Li, Yikai Zhao, Bichuan Guo, and Jiangtao Wen, "Novel tile segmentation scheme for omnidirectional video," in *Image Processing (ICIP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 370–374.

[10] Ramin Ghaznavi Youvalari, "360-degree panoramic video coding," 2016.

[11] 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.

[12] Ivana Tosic and Pascal Frossard, "Low bit-rate compression of omnidirectional images," in *Picture Coding Symposium*. IEEE, 2009, pp. 1–4.

[13] Jens-Rainer Ohm and Gary J Sullivan, "High efficiency video coding: the next frontier in video compression [standards in a nutshell]," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 152–158, 2013.

[14] Bin Li, Houqiang Li, Li Li, and Jinlei Zhang, "Domain rate control algorithm for high efficiency video coding," *IEEE Transactions on Image Processing (TIP)*, vol. 23, no. 9, pp. 3841–3854, 2014.

[15] 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.

[16] Xiaoyu Xiu Yan Ye Yuwen He, Bharath Vishwanath, "Ahg8: Interdigital's projection format conversion tool," *Input Document to JVET*, 2016.

[17] IEEE 1857 working group, "http://www.ieee1857.org," 2016.

[18] Deng Xin, Maosheng Bai, and Wei Wei, "Draft of subjective evaluation methodology on vr video," *Output Document of Standard for Immersive Visual Content Coding in IEEE 1857*.

[19] Shengxi Li, Mai Xu, Zulin Wang, and Xiaoyan Sun, "Optimal bit allocation for ctu level rate control in hevc," *IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)*, 2016.