Evaluation of Modifications to CPPPSNR in 360° Video Quality Assessment

Jayasingam Adhuran*
University of Surrey
United Kingdom
j.adhuran@surrey.ac.uk

Gosala Kulupana
University of Surrey
United Kingdom
g.kulupana@surrey.ac.uk

Chathura Galkandage*
c.galkandage@surrey.ac.uk
Anil Fernando*
w.fernando@surrey.ac.uk

Abstract*—360° videos are represented in spherical projection formats and the video quality of such videos is assessed using spherical objective quality metrics. Furthermore, the objective video quality between two different spherical projection formats can be evaluated using Cross projection metrics. Craster parabola, is a 2D cross projection format which is used by Craster Parabolic Peak Signal-to-Noise Ratio (CPPPSNR) metric. The existing CPPPSNR measurements do not consider the subsampling locations during the quality assessment to match the pixel density of a sphere. Nevertheless, it is vitally important to account for the oversampled projection formats and the sphere in order to be compatible with the existing video encoding architectures. To this end, the proposed improvements to the CPPPSNR locates the subsample points during craster parabolic projection and use nearest neighbor interpolation to assign pixels from the craster parabolic projection. Furthermore, in order to compensate the occurrence of oversampling, appropriate weights are applied to the corresponding pixels. The proposed method was tested with Shanghai Jiao Ton University (SJTU) Virtual Reality (VR) sequences for projection conversion. The comparison between Spherical PSNR (SPSNR) and existing CPPPSNR, validate the proposed CPPPSNR as an objective quality metric for cross projections.

Keywords—360° videos, projection, objective quality, craster parabola, CPPPSNR, ERP, Cubemap, CISP

I. INTRODUCTION

End-user Quality of Experience (QoE) towards multimedia consumption has become increasingly positive with the introduction of high resolution video formats such as Ultra High Definition (UHD)/4K, 8K, High Dynamic Range (HDR) and novel content types such as Augmented Reality (AR), Virtual Reality (VR), and omni directional videos. However, these novel video formats pose many challenges during their representations and subsequent processing which cannot be resolved using traditional techniques. As an example, the viewpoints in an omni directional video is subjected to change depending on its consumer. Assuming that the viewpoints are uniformly distributed, the observable 3D space of such videos becomes isotropic. Thus, unlike in traditional 2D videos, the observation space of omnidirectional videos could be defined as a spherical surface. These omni directional videos known as spherical or 360° videos (here after referred as) and can be represented by the parameters of the spherical coordinate system, latitude, longitude and radius (hereafter represented by θ ($\in$ $(-\pi/2, \pi/2)$), φ ($\in$ $(-\pi, \pi)$), and R (unit value) respectively). There have been numerous researches in the relevant literature which focus on defining representation formats for 360° videos.

The latest initiative of MPEG (Moving Picture Expert Group) called Versatile Video Coding (VVC) primarily focuses on improving the coding efficiency both in 2D video formats as well as 3D video formats such as 360° videos. One of the key research areas that is contributing to the coding efficiency of 360° videos in the VVC context is video/image projection techniques. 360° videos can be generated using multiple camera points in various directions to cover 360° space. However, VVC and many state-of-the-art video compression technologies require input videos in the 2D format prior to the encoding stage [1]. Many projection formats are being proposed in various researches to convert spherical 3D videos into 2D planer videos. Equirectangular Projection (ERP), Cubemap projection, pyramid projection, Rhombic Dodecahedron Map, Isohedran are some of them [2]-[4]. Few papers also discuss about variable quality mappings during the convertion process which assign different weights according to the user viewpoints in contrast to conventional uniform quality values [2]. Facebook which has 360° streaming facilities in the market uses tradition cubemap projection and pyramid projection in 360° video content delivery [5]. Moreover, offset-cubemap projection a variable quality mapping which has been implemented by Facebook F8 developers [6] and has also pioneered in view prediction techniques for 360° videos. Joint Video Experts Team (JVET) has used twelve projection methodologies, to support the new video coding standard, VVC and has established a 360Lib software package for 360° video coding and processing [4].

In addition to the above mentioned video projection formats, Craster Parabola is designed as a 2D cross projection format, to evaluate Craster Parabolic Peak Signal-to-Noise Ratio (CPPPSNR) metric. The role of CPPPSNR is crucially important when measuring objective video quality of two different video projection formats. The existing CPPPSNR downsamples the original projected video to match the pixel density of a sphere. However, in order to utilize the existing encoding architecture, it is important to consider the pixel density of a sphere as well as the oversampled projection formats. To this end, the proposed improvements to the CPPPSNR locates the subsample points during craster parabolic projection, assign pixels and apply weights to the corresponding to pixels to compensate the occurrence of oversampling.

The rest of this paper is organized as below. Section II describes the background work in the area, Section III describes the proposed improvements to the CPPPSNR, and Section IV illustrates the test results followed by the concluding remarks in Section V.

II. BACKGROUND

Objective quality evaluation is vital in video processing and compressing applications. A video quality metric supports to predict the subjectively perceived quality of the

978-9-0827-9703-9/19/$31.00

video. Spherical objective quality metrics are used for quality assessment of 360° videos. Commonly used metrics are PSNR based metrics due to their low computational complexity. Since PSNR does not address any spherical characteristics of the video, other PSNR based metrics such as Weighted Spherical PSNR (WSPSNR) [7], Average Weighted Spherical PSNR (AWSPSNR) [8], Craster Parabolic PSNR (CPPPSNR) [9], Spherical PSNR (SPSNR) [10] are also being implemented for objective quality evaluation of the 360° videos.

Cross projection metrics are used to evaluate between different projection formats of a 360° video. PSNR and WSPSNR which are non-cross projection metrics evaluate only between the same projection formats. Nevertheless, SPSNR and CPPPSNR can evaluate between different projections, i.e the objective evaluation of the constructed cubemap projected video with respect to an ERP video. SPSNR constructs a virtual sphere given samples of θ and φ for each frame of the video sequence. The 2D image frame would then be mapped on the sampled locations on the sphere and corresponding pixel values are obtained. SPSNR also use interpolations of pixels.

CPPPSNR uses craster parabolic representation of the sphere which has the least distortion among spherical projections [11]. The 2D image frame is down sampled to construct the craster parabola in order to represent the pixel density of a sphere. An advantage seen by CPPPSNR over SPSNR is that SPSNR requires pre-sample locations to construct the virtual sphere and CPPPSNR does not require new sampling locations to construct the craster parabola. Both SPSNR and CPPSNR along with other spherical objective quality metrics such as WSPSNR and PSNR have been adopted into 360 Library developed by the JVET group [4].

The metrics discussed above represent the spherical characteristics of the 360° video. However, the input sequences are generally in projection formats of the sphere. Therefore, when constructing the 360° videos from the projected videos, it is important to consider the sample points that have been used to construct the projected sequences. In addition, a compensation method must negate the additional sampling points introduced in the projection formats compared to that in a sphere. The proposed modifications to CPPPSNR will therefore focus on the introduction of additional sampling points and corresponding compensation method.

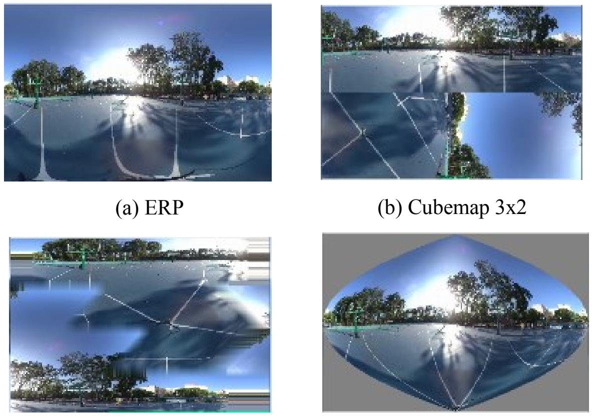
III. PROPOSED CPPPSNR

Craster parabola projection has lower integer sample locations than the 2D image frame. Thus, generated craster parabola would not have equally spaced sample positions as any rectangular image frame would have. In carrying out measurement with the available integer sampling positions and neglecting subsampling locations as implemented by the CPPPSNR, one to one mapping of the 2D image frame and the craster parabola projection is no longer available.

The proposed CPPPSNR locates pixel values in the subsample positions using nearest neighbor interpolation and compensate the oversampling by weighing down the pixels.

The main features of the proposed CPPPSNR:

- A spherical metric having same number of samples as any projected 2D image frame, while compensating for the oversampling in the projected 2D frame.



(a) ERP (b) Cubemap 3x2

(c) CISP (d) Craster Parabola

Fig 1. Basketball court sequence presented in various projection formats; (a) ERP,(b) Cubemap3x2, (c) CISP and (d) Craster Parabola

- It is a cross projection spherical metric which can spherically reference to state-of-the-art video codecs for cross projection applications owing to its ability to match sampling locations in 2D image.

The craster parabolic conversation between spherical (θ,φ) and cartesian coordinates (x,y) [4],[12] are given by equations 1 and 2.

$$x = \varphi(2cos\frac{2\theta}{3} - 1) \tag{1}$$

$$y = \pi sin\frac{\theta}{3} \tag{2}$$

Assume $(x_i`,y_j`)$ to be the sample location of an image A with width W and height H, which would have a total number of integer sample locations equal to W x H. However, after craster parabolic conversion, the total number of integer samples in the projected image P would be approximately two third of the original samples. Therefore, we further consider subsample locations in the proposed method. When allocating pixel locations for objective quality measurement, each $(x_i`,y_j`)$ sampling point in A would be mapped to spherical coordinates (φ_i, θ_j) using one to one mapping [4] and is then converted to (x_k, y_l) sampling point in P using the same relationship mentioned in equations 1 and 2.

$$(x_i`, y_j`) \rightarrow (\varphi_i, \theta_j) \rightarrow (x_k, y_l), \text{where } i,j \in Z \,\&\, k,l \in R$$

Although, craster parabola has an equal pixel density of the sphere, the construction of the parabola must consider nearest neighbor interpolation to stack the pixels placed on sub sample positions. The pixel values on sub sample position must be accounted, considering the fact that the encoder would count all sample positions. In order to compensate the error caused by the additional inclusion of the pixel values from the sub sample positions, a weight is applied to each pixel. In order to assign pixel weights, the Jacobian matrix is used.The Jacobian matrix in differential geometry is the transformation relationship between two Euclidean spaces. The craster parabola that is constructed from the spherical coordinates (φ_i, θ_j) would not have equal pixel density. Since Jacobian matrix (J) in differential geometry refers to the transformation relationship between two Euclidean spaces it has been used to weigh pixels error [7]. The Jacobian metric of the cartesian coordinates and the spherical coordinates is derived by equation 3.

$$J = \frac{\partial(x,y)}{\partial(\varphi,\theta)} = \pi/3\cos(\theta) \quad (3)$$

$$MSE = \frac{\sum_{j=0}^{H-1}\sum_{i=0}^{W-1}((x_l, y_l)_{Ori} - (x_k, y_l)_{Rec})^2 \cos(\theta_j)}{\sum_{j=0}^{H-1}\sum_{i=0}^{W-1}\cos(\theta_j)} \quad (4)$$

$$Proposed\ CPPPSNR = 10\log_{10}\frac{255^2}{MSE} \quad (5)$$

The proposed Mean Square Error (MSE) for each frame is calculated as shown by equation 4. $(x_k, y_l)_{Ori}$ and $(x_k, y_l)_{Rec}$ are the nearest neighbor values of the pixel values in P of the original and reconstructed frame respectively. Proposed CPPPSNR for each frame is calculated as shown by equation 5 and averaged over all the frames in the video sequence.

IV. Simulation Results and Discussion

The proposed method was tested using the 8k Virtual Reality (VR) sequences produced by Shanghai Jiao Ton University (SJTU) media lab [13] which are published to support next generation encoding performances and adaptive bitrate streaming and related researches. The tested video sequences were, Basketball Court (BC), Academic Building (AB), Sward (SW), South Gate (Night) (SG(N)), Siyuan Gate (SYG), East Gate (EG), Library (LB), Runners (RS), Bridge (Night) (BG(N)), South Gate (SG), Study Room (SR) and Administration Building (AD). The test sequences were in uncompressed ERP format with the resolution 8192x4096. The 360 library [4] provided by JVET was used to conduct the testing. PSNR values of luma samples for the following cross projection PSNR metrics were obtained during tests.

1. SPSNR_I – Spherical PSNR Bi-cubic Interpolated.
2. SPSNR_NN – Spherical PSNR Nearest Neighbour
3. CPPPSNR – Existing Craster parabolic PSNR.
4. Proposed CPPPSNR

The testing involves conversion between ERP and Cubemap3x2 and ERP and CISP. Cubemap 3x2 is the cubemap projection format with six square faces packed into a rectangular frame whereas CISP is the isohedran projection format with 20 triangular faces packed into rectangular frame. Cubemap and CISP are chosen for the testing sine cubemap is a frequently used and less complex projection where as CISP is a complex projection [14]. Therefore, there will be a definite difference in quality assessment. The craster parabola projection is obtained during the calculation of the CPPPSNR as an intermediate output. All projection formats used in this research are shown in Fig. 1 with respect to BC sequence. The test sequences were initially down sampled to 6k resolution to reduce the computation time. The 3 tests were conducted to validate the proposed CPPPSNR.

1. The 6k ERP sequences converted to 6k cubemap3x2 and 6k CISP, and back to 6k ERP (ERP6k–Cubemap6k-ERP6k & ERP6k-CISP6k-ERP6k).
2. The 6k ERP sequences projected to 6k cubemap3x2 and 6k CISP (ERP6k-Cubemap6k & ERP6k-CISP6k).
3. The original 8k ERP sequences projected to 6k cubemap3x2 and 6k CISP (ERP8k-Cubemap6k & ERP8k-CISP6k).

The Test 1 has a net conversion which is equal to zero since the original projection format is not changed. Therefore, other non-cross projection spherical metrics could be used for evaluation. WSPSNR is considered as the best objective metric to evaluate spherical videos represented in the same format [7] as it considers spherical distortions of the video and does not involve interpolations. Therefore, it has been used as reference to evaluate the performance of any cross-projection metric. Table 1 shows the results of Test 1. Although in Test 1 the original video was expected to be reconstructed from reverse conversions, all metrics including those accepted in 360 library by JVET [4] indicate an occurrence of distortion as the ideal case should yield an infinite value. However, there is no net format conversion. Therefore, the metrics are evaluated with respect to WSPSNR as mentioned earlier. The calculated Pearson Correlation Coefficient (PCC) for SPSNR_NN, SPSNR_I, CPPPSNR, proposed CPPPSNR with respect to WSPSNR are tabulated in Table2.

The coefficients in Table 2 show that all metrics are highly correlative. However, SPSNR_NN, SPSNR_I and the proposed CPPPSNR have much greater correlation than that of existing CPPPSNR. SPSNR_NN has the highest correlation compared all other metrics in all formats. Proposed CPPPSNR shows a greater correlation than that of SPSNR_I.

All metrics in Table 1 show similar variation between different sequences and different coding projections. (i.e., all metrics display lower PSNR values for BC sequence compared to that of AD sequence and all metrics display lower values for AD sequence for CISP projection than cube projection). Since CISP structure is complex than cubemap, the projection conversions to and from CISP is distorted more than cubemap.

The Test 2 is conducted to observe the effect of projection conversion and its results are shown in Table 3. WSPSNR metric is not applicable to this assessment, as this is a cross projection assessment. Test 3 is conducted to observe the combined effect of down sampling and projection conversion. The results are shown in the Table 4.

Proposed CPPPSNR and SPSNR_I have greater correlation indicated by PCC 0.9992, 0.9998, and 0.9995 in Test 2 and 0.9993, 0.9998, and 0.9995 in Test 3 for cube projection, CISP projection and overall respectively. CPPPSNR has considerably lower values compared to SPSNR_I and proposed CPPPSNR and weak correlation between them in both Test 2 and 3.

SPSNR_I, CPP_PSNR and proposed CPPPSNR values obtained for Test 2 and Test 3 are similar in comparison to the results obtained in Test 1. Also, similar variations between different sequences and between different projection formats for the same sequences is observed as observed in Test 1. However, the quality has dropped slightly from Test 1 to Test 2 and further dropped from Test 2 to Test 3. An average drop of 0.6412 dB, 0.1254 dB, and 0.6540 dB between Test 1 results and Test 2 results and an average drop of 0.8897 dB, 0.1909 dB and 0.8973 dB between Test 2 results and Test 3 results have been observed for SPSNR_I, CPP_PSNR and proposed CPPPSNR respectively. Test 2 induces a conversion distortion whereas Test 3 induces distortion from projection conversion and down sampling which has caused degradation of the quality going from Test 1 to Test 2 to Test3.

SPSNR_NN which had higher correlation with WSPSNR in Test 1 in assessment of overall conversion displayed very low PSNR numbers in Test 2 and Test 3 with average drop of 41.4704 dB and 41.4888 dB from Test 1 to Test 2 and from Test 2 to Test 3 respectively. SPSNR_NN show inconsistencies in sequences such as AD, SG(N), SYG, EG and LB where the numbers indicate an increase in quality

Table 1: Results for Test 1

	Coding Projection	WSPSNR	SPSNR_NN	SPSNR_I	CPPPSNR	Proposed CPPPSNR
Basketball Court	Cubemap	53.6197	53.6936	53.944	47.107	53.8058
Academic Building	Cubemap	55.523	55.5248	54.9845	47.3159	55.0523
Sward	Cubemap	49.7622	49.7417	50.0066	45.9567	50.2204
South Gate (Night)	Cubemap	58.0013	58.0123	57.0909	47.6132	57.1613
Siyuan Gate	Cubemap	56.6628	56.6942	56.1001	47.4857	56.1649
East Gate	Cubemap	55.5845	55.6014	55.1704	47.3473	55.258
Library	Cubemap	53.6139	53.6401	53.5742	47.0356	53.5786
Runners	Cubemap	56.4494	56.4749	55.8364	47.4497	55.9152
Bridge (Night)	Cubemap	57.8123	57.8215	56.9228	47.5949	57.0129
South Gate (Day)	Cubemap	53.6401	53.6697	53.5349	47.0249	53.6387
Study Room	Cubemap	56.145	56.1669	55.8641	47.4535	55.8745
Administration Building	Cubemap	52.885	52.9187	53.0123	46.901	53.0845
Basketball Court	CISP	50.0653	50.1591	50.4886	46.1056	50.303
Academic Building	CISP	52.9827	53.0752	52.702	46.8211	52.6526
Sward	CISP	45.8943	45.9538	46.3393	44.0809	46.3389
South Gate (Night)	CISP	55.5881	55.6361	54.8767	47.2948	54.8468
Siyuan Gate	CISP	54.0079	54.0216	53.6608	47.0565	53.6298
East Gate	CISP	52.6564	52.6687	52.4747	46.7701	52.4512
Library	CISP	50.5651	50.6949	50.7814	46.2275	50.6632
Runners	CISP	53.9847	54.0686	53.6214	47.0429	53.5843
Bridge (Night)	CISP	55.4531	55.5009	54.7469	47.2722	54.7266
South Gate (Day)	CISP	50.3802	50.4915	50.5796	46.1523	50.5159
Study Room	CISP	52.3945	52.4517	52.474	46.7636	52.3748
Administration Building	CISP	49.739	49.8269	50.0479	45.9372	50.0042

Table 2: Correlation between WSPSNR and other quality metrics in Test 1

Coding Projection	SPSNR_NN	SPSNR_I	CPPPSNR	Proposed CPPPSNR
Cubemap	0.999955257	0.99579327	0.963095	0.9983091
CISP	0.999926581	0.99790427	0.964266	0.9988346
Overall	0.999933874	0.99723497	0.952912	0.9973882

which is in contradiction with the other metrics. In addition, although other three metrics show a higher value for cubemap projection compared to CISP as similar to that in Test 1, SPSNR_NN show inconsistencies for some sequences such as AD, LB, SYG and RS where SPSNR_NN values for CISP are greater than that of cubemap. Since SPSNR_NN does not use interpolation to find the intermediate values and reliant on fixed sample positions to co-locate pixel positions, it considerably fails to find the co-location of pixel positions between the original frame and converted frame between different projections when frame resolution is very high such as 6k and 8k.

The comparison charts for Test 1, Test2 and Test3 for the proposed CPPPSNR with respect to cubemap and CISP projections are provided in Fig. 2 and Fig. 3 respectively. Charts in these figures confirm that the proposed CPPPSNR show decline in quality ranging from Test 1 to Test 3. This was the expected. This variation is observed as each test includes an addition quality degradation factor as mentioned above. Further, Fig. 2 and Fig. 3 showed that quality is

Table 3: Results for Test 2

	Coding Projection	SPSNR_NN	SPSNR_I	CPPPSNR	Proposed CPPPSNR
Basketball Court	Cubemap	12.262	52.5383	46.7827	52.4436
Academic Building	Cubemap	8.6568	53.6739	47.0591	53.6872
Sward	Cubemap	9.1865	49.3987	45.7066	49.5827
South Gate (Night)	Cubemap	15.0715	54.9898	47.3188	55.0145
Siyuan Gate	Cubemap	11.1218	54.4882	47.2278	54.5212
East Gate	Cubemap	12.7889	53.7289	47.0751	53.777
Library	Cubemap	11.0951	52.502	46.7761	52.4923
Runners	Cubemap	8.6389	54.3035	47.1915	54.3368
Bridge (Night)	Cubemap	21.8277	54.8626	47.2999	54.8998
South Gate (Day)	Cubemap	14.0348	52.6894	46.8233	52.7664
Study Room	Cubemap	11.3042	54.2132	47.1722	54.1911
Administration Building	Cubemap	10.0528	51.9336	46.6155	52.0073
Basketball Court	CISP	11.0189	50.124	45.9713	49.9467
Academic Building	CISP	9.8116	52.3637	46.7336	52.3254
Sward	CISP	8.1788	46.0357	43.9048	46.0477
South Gate (Night)	CISP	14.3011	54.324	47.1937	54.2996
Siyuan Gate	CISP	11.1155	53.2795	46.9726	53.2595
East Gate	CISP	12.4528	52.1339	46.6745	52.109
Library	CISP	12.3364	50.4464	46.1096	50.3393
Runners	CISP	8.6366	53.2323	46.9569	53.2063
Bridge (Night)	CISP	20.4088	54.2081	47.1718	54.1898
South Gate (Day)	CISP	13.288	50.2643	46.0437	50.2261
Study Room	CISP	11.6442	52.1008	46.6621	52.0154
Administration Building	CISP	9.5448	49.6463	45.7847	49.638

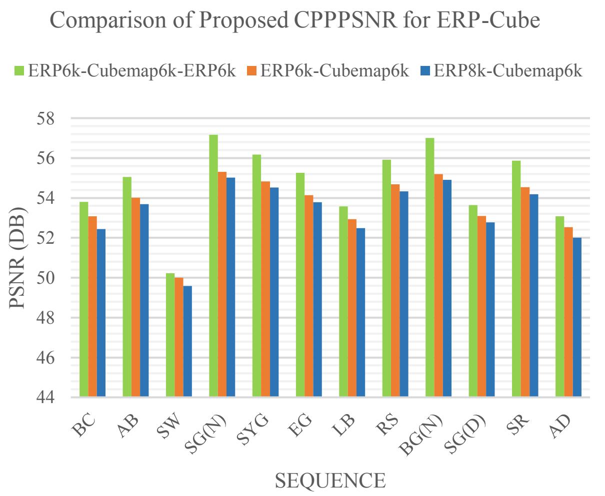


Fig. 2. Comparison of Proposed CPPPSNR for ERP-Cube conversion

Table 4: Results for Test 3

	Coding Projection	SPSNR_NN	SPSNR_I	CPPPSNR	Proposed CPPPSNR
Basketball Court	Cubemap	12.6051	53.1908	46.9437	53.0783
Academic Building	Cubemap	8.5477	53.9911	47.121	54.0128
Sward	Cubemap	9.5059	49.7988	45.8729	49.9989
South Gate (Night)	Cubemap	15.4038	55.2816	47.3687	55.3046
Siyuan Gate	Cubemap	10.512	54.8018	47.2842	54.8265
East Gate	Cubemap	12.6074	54.0821	47.1492	54.1305
Library	Cubemap	11.2339	52.9471	46.8913	52.9378
Runners	Cubemap	8.1735	54.6458	47.2569	54.6813
Bridge (Night)	Cubemap	21.9821	55.1649	47.3509	55.2014
South Gate (Day)	Cubemap	13.4295	53.0247	46.9072	53.0992
Study Room	Cubemap	11.8626	54.5608	47.2397	54.5431
Administration Building	Cubemap	10.6024	52.4852	46.769	52.5395
Basketball Court	CISP	11.8345	50.2522	46.0226	50.0744
Academic Building	CISP	9.5132	52.4604	46.7614	52.4214
Sward	CISP	8.49	46.271	44.0127	46.225
South Gate (Night)	CISP	14.5505	54.3673	47.2019	54.338
Siyuan Gate	CISP	10.5513	53.3379	46.9865	53.3123
East Gate	CISP	12.313	52.2334	46.7032	52.2047
Library	CISP	11.714	50.5686	46.1575	50.4634
Runners	CISP	8.1933	53.3143	46.9765	53.2837
Bridge (Night)	CISP	20.5086	54.2602	47.1823	54.2382
South Gate (Day)	CISP	13.2867	50.4024	46.0947	50.3559
Study Room	CISP	11.8398	52.1491	46.6759	52.0588
Administration Building	CISP	9.9576	49.8559	45.8702	49.8321

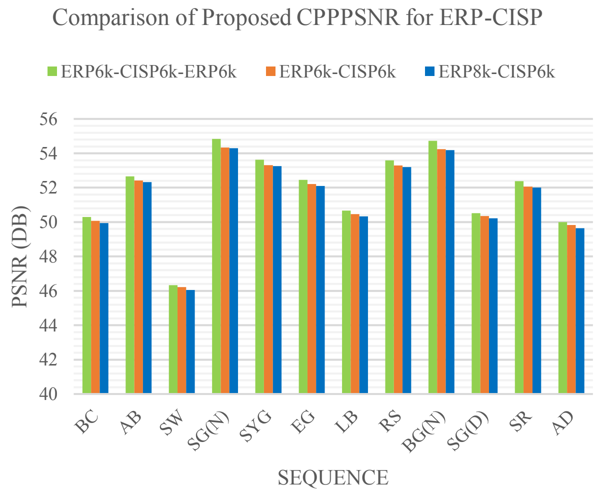


Fig. 3. Comparison of Proposed CPPPSNR for ERP-CISP conversion

decreased during ERP-CISP conversion more than ERP-cubemap conversion due to the complexity structure of CISP. Also, tests showed that the proposed CPPPSNR display similar results to that of SPSNR_I and there exists a high correlation between them. In addition, result variation between SPSNR_I, CPPPSNR and proposed CPPPSNR are similar. Proposed CPPPSNR also had a greater correlation with WSPSNR in Test 1. Since, SPSNR_I and CPPPSNR are accepted spherical metrics for cross projection [4], it is implied that the proposed CPPPSNR results are valid for objective quality assessment. Although SPSNR_NN is an accepted quality metric, many inconsistencies have found in its obtained results. Therefore, the metric cannot be used to validate the proposed CPPSNR.

V. Conclusion

The proposed CPPPSNR includes modifications to the existing CPPPSNR for objective quality assessment. The obtained results were compared against WSPSNR, SPSNR_I and existing CPPPSNR. From the analysis, it is conclusive that the proposed CPPPSNR metric is valid to be utilized as a spherical objective quality metric for cross projection evaluation. Therefore, the proposed modifications to the existing CPPPSNR can be incorporated. This proposed CPPPSNR can be used for spherical objective quality assessment inside the video encoders in the future.

VI. Acknowledgement

This research was aided by CONTENT4ALL which is a Horizon 2020 project funded by the European Commission (Grant Number 762021).

References

[1] El-Ganainy, Tarek, and Mohamed Hefeeda. "Streaming virtual reality content." arXiv preprint arXiv:1612.08350 (2016).

[2] Corbillon, Xavier, et al. "Viewport-adaptive navigable 360-degree video delivery." Communications (ICC), 2017 IEEE International Conference on. IEEE, 2017.

[3] Skupin, Robert, et al. "Standardization status of 360 degree video coding and delivery." 2017 IEEE Visual Communications and Image Processing (VCIP). IEEE, 2017

[4] Ye, Y., E. Alshina, and J. Boyce. "Algorithm descriptions of projection format conversion and video quality metrics in 360Lib." Joint Video Exploration Team of ITU-T SG 16 (2017).

[5] F. Qian, L. Ji, B. Han, and V. Gopalakrishnan, "Optimizing 360 video delivery over cellular networks," in Proceedings of the 5th Workshop on All Things Cellular: Operations, Applications and Challenges. ACM, 2016, pp. 1–6.

[6] El-Ganainy, Tarek, and Mohamed Hefeeda. "Streaming virtual reality content." arXiv preprint arXiv:1612.08350 (2016).

[7] Sun, Yule, Ang Lu, and Lu Yu. "Weighted-to-spherically-uniform quality evaluation for omnidirectional video." *IEEE signal processing letters* 24.9 (2017): 1408-1412.

[8] Xiu, Xiaoyu, et al. "An evaluation framework for 360-degree video compression." Visual Communications and Image Processing (VCIP), 2017 IEEE. IEEE, 2017.

[9] Zakharchenko, Vladyslav, Kwang Pyo Choi, and Jeong Hoon Park. "Quality metric for spherical panoramic video." *Optics and Photonics for Information Processing X*. Vol. 9970. International Society for Optics and Photonics, 2016.

[10] Yu, Matt, Haricharan Lakshman, and Bernd Girod. "A framework to evaluate omnidirectional video coding schemes." *2015 IEEE International Symposium on Mixed and Augmented Reality*. IEEE, 2015.

[11] White, Denis. "Display of pixel loss and replication in reprojecting raster ata from the sinusoidal projection." *Geocarto International* 21.2 (2006): 19-22.

[12] Calabretta, Mark R., and Eric W. Greisen. "Representations of celestial coordinates in FITS." *Astronomy & Astrophysics*395.3 (2002): 1077-1122.

[13] Liu, Xu, et al. "The SJTU UHD 360-Degree Immersive Video Sequence Dataset." *the 17th International Conference on Virtual Reality and Visualization (ICVRV2017)*. 2017.

[14] Pei, Qikai, et al. "COP: A New Continuous Packing Layout for 360 VR Videos." *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018.