An Introduction to the Spherical Structural Similarity Index for Omnidirectional Video Quality Assessment

VQEG Immersive Media Group (IMG)

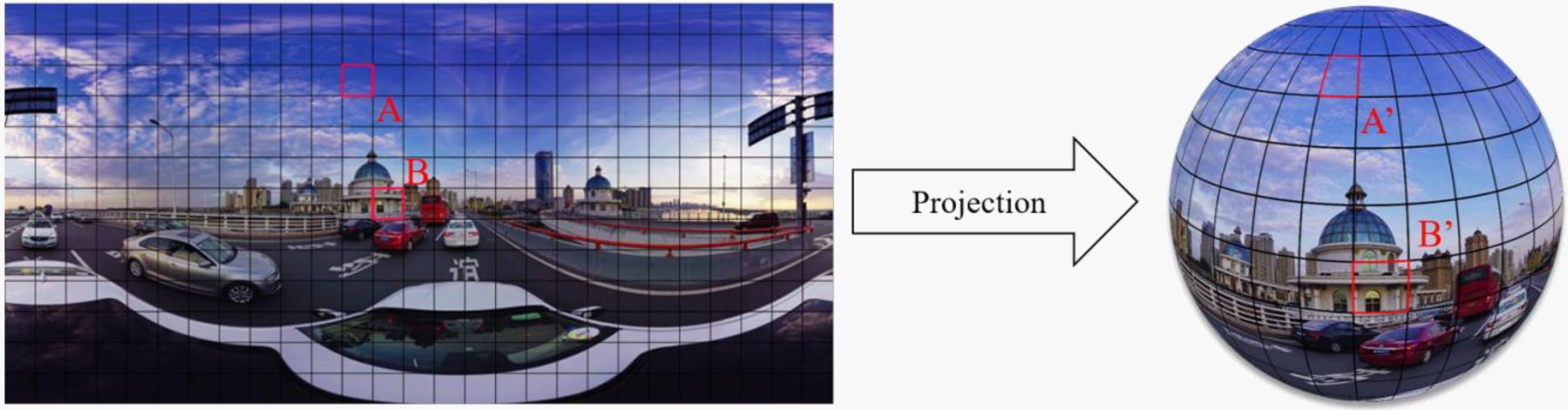
Wuhan University

Zhenzhong Chen, **Yingxue Zhang**

zzchen@whu.edu.cn
grace@whu.edu.cn

Background

Projection between plane and sphere



- Spherical information projected onto the plane results in **visual mismatches between the two domains.**
- The **relationship between the plane and spherical domain** needs to be taken into account for the objective video quality assessment.

[See more details in *Sijia Chen, Yingxue Zhang, Yiming Li, Zhenzhong Chen, and Zhou Wang*, “Spherical Structural Similarity Index for Objective Omnidirectional Video Quality Assessment,” ICME 2018]

Background

SSIM

SSIM computes the luminance, contrast and structural similarities between the distorted and original images based on the local patterns of pixel intensities that have been normalized, and combines these three comparisons to describe the overall structural similarity between the distorted and the original images as an estimate of the quality of the distorted image.

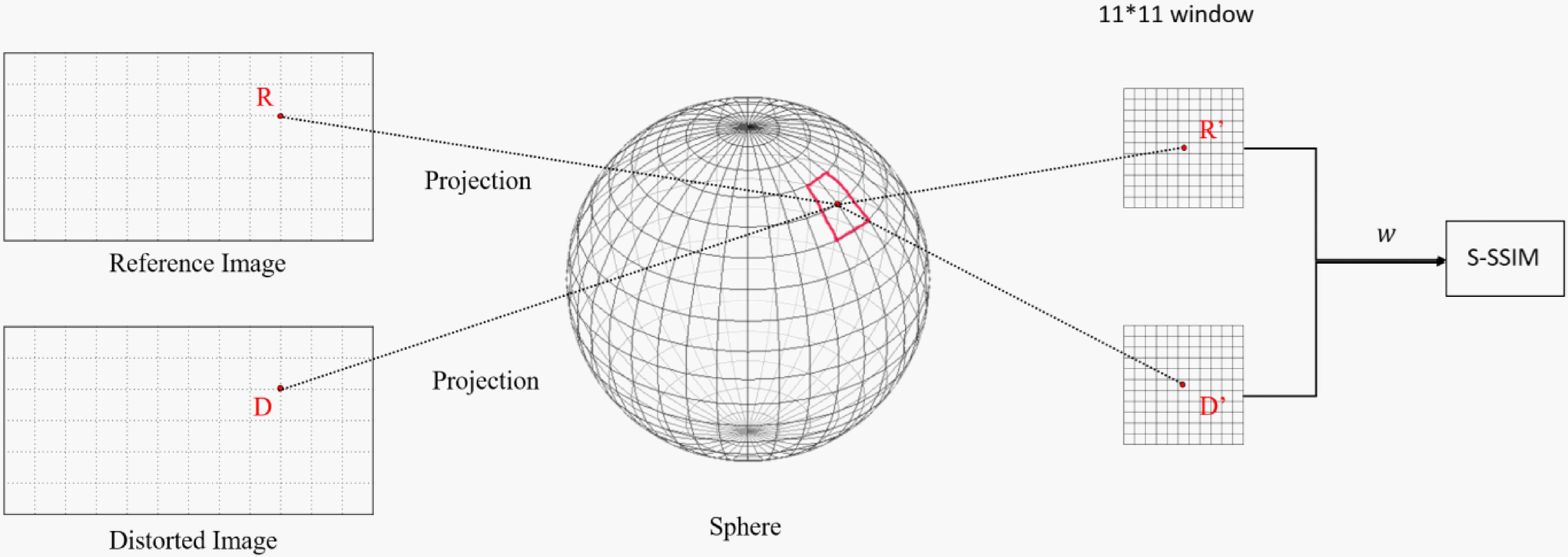
SSIM is more consistent with subjective quality evaluation than PSNR.

[Wang, Z., et al. "Image quality assessment: from error visibility to structural similarity. " *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society* 13.4(2004):600-612.]

2 The Proposed Method

SSIM Components in the Spherical Domain
Similarity Relationship on the Sphere and Projected Plane
S-SSIM in Spherical Domain

The Overall Framework



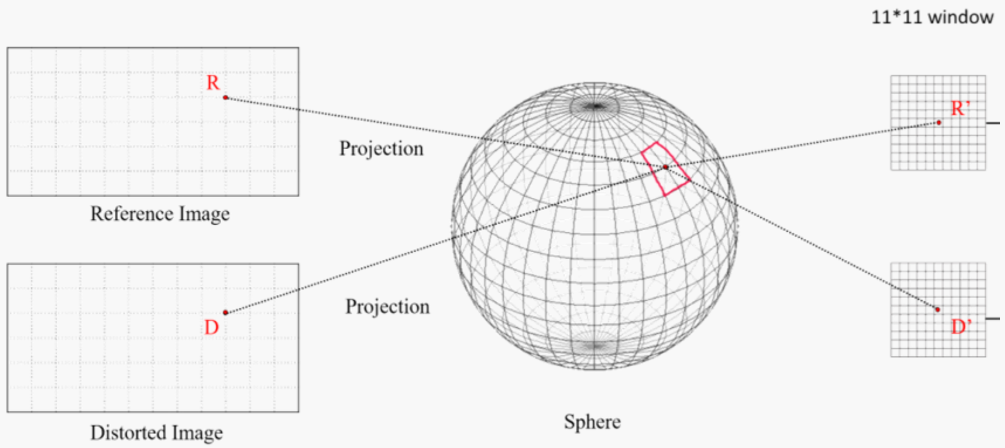
SSIM in the Spherical Domain

Pixels on the plane are projected to the sphere to compute the SSIM in the spherical domain.

- For each projected pixel located at *(i, j)* on the sphere, an *11×11* symmetric window are used to calculate the similarity measure in the spherical domain:

$$S-SSIM(i,j)=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}$$

  - *x* and *y*: pixels from the reference and distorted images.
  - μ_x, σ_x *and* σ_{xy}: the local mean, variance and the local covariance between the two regions.



Similarity Relationship between the Sphere and Projected Plane

For each location in 2-D plane, its similarity level on the sphere and the projected plane is not equal due to stretching during the projection.

- Let $S-SSIM_a$ denote the structural similarity measure near pixel a on the sphere. For the corresponding image patch ε_x with N pixels on the sphere, the local structural similarity measure is given by:

- $$S-SSIM_{\varepsilon_x} = \frac{\sum_{a=1}^{N} S\text{−}SSIM_a \cdot Area(\varepsilon_a)}{Area(\varepsilon_x)} = \frac{\sum_{a=1}^{N} S\text{−}SSIM_a \cdot Area(D_a) \cdot \frac{Area(\varepsilon_a)}{Area(D_a)}}{\sum_{a=1}^{N} Area(D_a) \cdot \frac{Area(\varepsilon_a)}{Area(D_a)}} = \frac{\sum_{a=1}^{N} S\text{−}SSIM_a \cdot w_a}{\sum_{a=1}^{N} w_a}$$

  - $Area(\varepsilon_a)$ and $Area(D_a)$: the area of a pixel in the spherical domain and the projected plane, respectively.
  - $Area(\varepsilon_x)$: the area of the image patch ε_x on the sphere.

$w_a = \frac{Area(\varepsilon_a)}{Area(D_a)}$, the scaling factor

$Area(D_a)$ = unit area

S-SSIM in Spherical Domain

- For an $M \times N$ image, the final similarity measure is defined as:

$$S-SSIM = \frac{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}(S-SSIM_{(m,n)} \cdot w_{(m,n)})}{\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} w_{(m,n)}}$$

  - $w_{(m,n)}$: the scaling factor from the 2-D plane to the sphere
  - The scaling factor is dependent on the projection methods.

$$\Downarrow$$

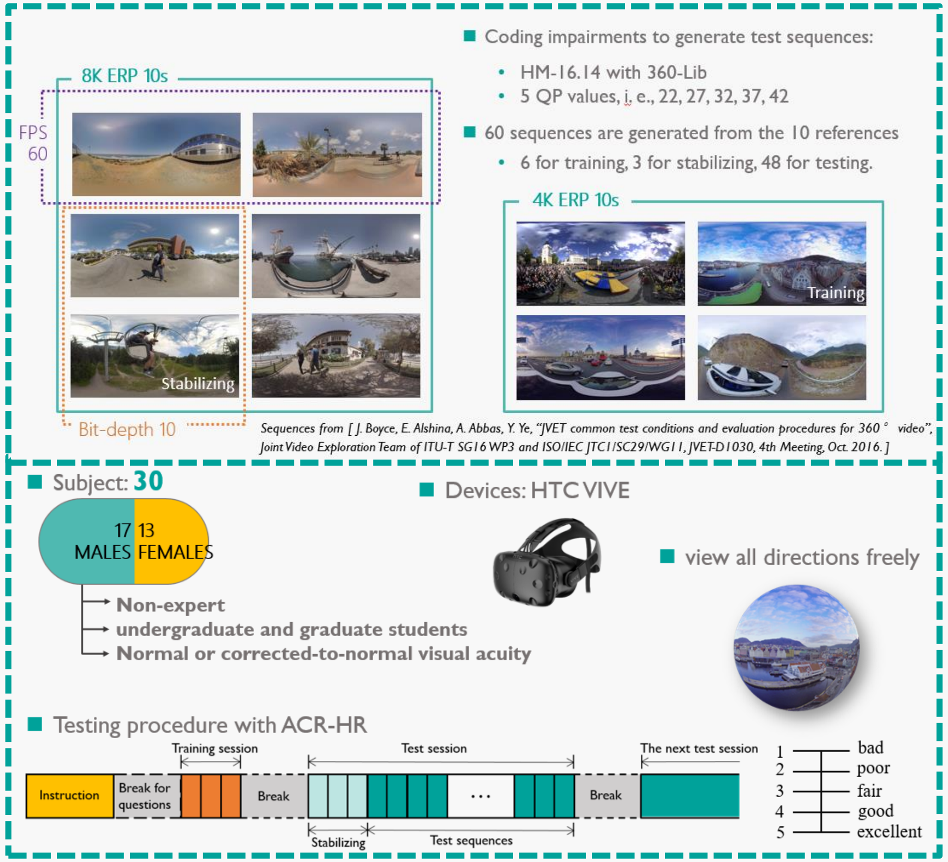
  - The proposed metric can be easily adapted to various types of projections.

3 Experimental Results

Correlation with Subjective data
Objective Score versus MOS

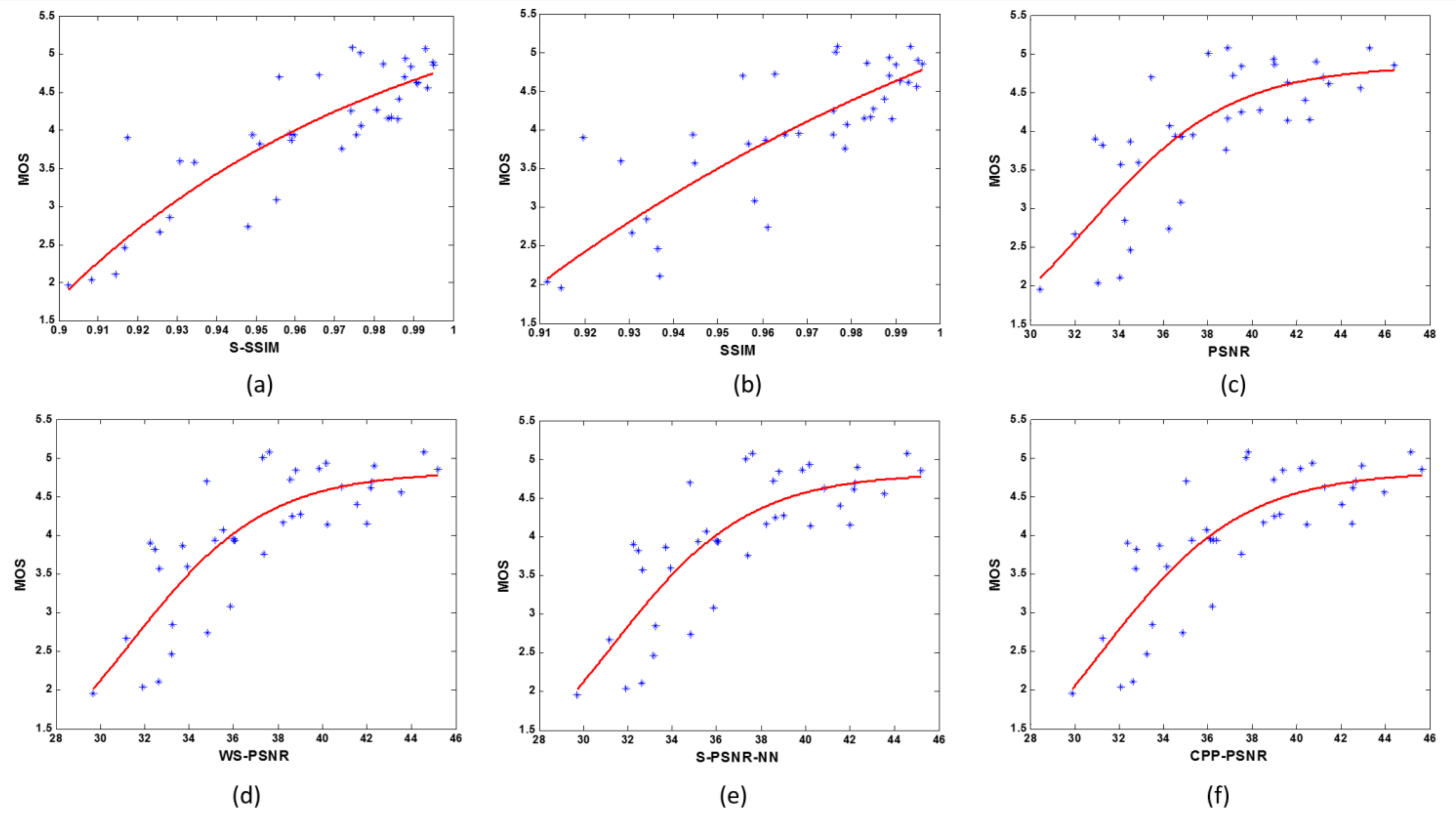
Correlation with Subjective data

Metric	SROCC↑	KROCC↑	PLCC↑	RMSE↓
S-SSIM	**0.8211**	**0.6509**	**0.8635**	**0.4428**
SSIM	0.7749	0.5915	0.8038	0.5223
PSNR	0.7825	0.5834	0.7741	0.5558
WS-PSNR	0.7937	0.6050	0.7971	0.5301
S-PSNR-NN	0.7937	0.6050	0.7963	0.5310
CPP-PSNR	0.8088	0.6185	0.8002	0.5265



See [Z. Chen, Y. Zhang, "Towards Subjective Quality Assessment for Panoramic Video", VQEG Krakow Meeting] for more detail on the subjective database.

Objective scores versus MOS



Conclusion

- We analyze the relationship of structural similarity between the plane and the spherical domain, and propose an SSIM based VQA algorithm for omnidirectional video, i.e., S-SSIM.
- The proposed metric is verified on a subjective video quality assessment database and compared with state-of-the-art objective quality evaluation metrics, indicating the superior performance.
- The framework can be easily generalized for various projection methods.


Thanks!
iip.whu.edu.cn