Quality Assessment of Super-Resolved Omnidirectional Image Quality Using Tangential Views

Cagri Ozcinar[1] and Aakanksha Rana[2]
Samsung R&D[1], UK, Massachusetts Institute of Technology[2], USA

Abstract

Omnidirectional images (ODIs), also known as 360-degree images, enable viewers to explore all directions of a given 360-degree scene from a fixed point. Designing an immersive imaging system with ODI is challenging as such systems require very large resolution coverage of the entire 360 viewing space to provide an enhanced quality of experience (QoE). Despite remarkable progress on single image super-resolution (SISR) methods with deep-learning techniques, no study for quality assessments of super-resolved ODIs exists to analyze the quality of such SISR techniques. This paper proposes an objective, full-reference quality assessment framework which studies quality measurement for ODIs generated by GAN-based and CNN-based SISR methods. The quality assessment framework offers to utilize tangential views to cope with the spherical nature of a given ODIs. The generated tangential views are distortion-free and can be efficiently scaled to high-resolution spherical data for SISR quality measurement. We extensively evaluate two state-of-the-art SISR methods using widely used full-reference SISR quality metrics adapted to our designed framework. In addition, our study reveals that most objective metric show high performance over CNN based SISR, while subjective tests favors GAN-based architectures.

Introduction

High-resolution omnidirectional images (ODIs) are quintessential to provide a sufficiently high-quality immersive experience for head-mounted displays. Single image super-resolution (SISR), a fundamentally ill-posed problem, enhances the resolution of a low-resolution image. While no existing consumer cameras can capture very large resolution, such as $\geq$8K, SISR techniques have emerged as a viable option for ODI [1]. From traditional hand-crafted approaches to deep learning techniques, SISR methods for ODIs aim to construct the high-resolution ODIs using objective criteria adhering to SISR frameworks for standard images.

ODI processing systems can be optimized by leveraging the techniques of SISR. Similarly to the traditional image, quality assessment plays a vital role. However, no quality assessment framework exists in the literature that can measure and analyze the perceptual efficacy of the super-resolved ODIs while catering to its inherent spherical distortions. Without appropriate quality assessments, the viability of SISR for ODI techniques that promise immersive experience cannot be established.

Many recent learning-based SISR works have achieved remarkable improvements over traditional techniques. Majority of the existing SISR works aim to improve the super-resolved images' accuracy by minimizing the mean squared reconstruction error. With the invention of generative adversarial networks (GANs), state-of-the-art SISR methods claim to achieve higher perceptual quality and better realistic super-resolved images.

While recent SISR works have improved quantitative results, there is an inconsistency between quantitative and qualitative results. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) show great success for SISR. For instance, the super-resolution generative adversarial network (SR-GAN) [2] shows that the high PSNR values do not always accord with better perceptual quality. The reason of this situation is that GAN-based SISR methods often produce realistic yet fake details and textures on the generated super-resolved images [3], bringing a challenge for existing image quality assessment methods.

SISR for ODI techniques should produce images with high perceptual quality to have an enhanced immersive viewing experience. However, the lack of understanding of the distortion of super-resolved ODIs motivates us to develop a quality measurement framework, and analyze the distortions of SISR and the performance of quality metrics. In particular, in our previous work [1], we showed that widely used quality metrics (SSIM, PSNR) developed for standard 2D images do not perform well for SISR on ODIs. Hence, we believe that analyzing quality metrics is crucial for designing an optimal SISR method for ODIs. Nevertheless, ODI quality evaluation analysis for SISR was not considered in the literature. To this end, in this study, we target the following objectives: 1) designing a novel SISR quality assessment framework for ODI; 2) performing a thorough evaluation study using state-of-the-art SISR methods for ODIs over state-of-the-art image quality assessment frameworks; 3) investigating the qualitative and quantitative results of CNN- and GAN-based SISR results.

This paper systematically evaluates a broad set of full-reference image quality assessment models with tangential views of ODI. We first developed an objective full-reference ODI quality assessment framework and studied quality measurement for ODIs generated by SISR methods. The proposed quality assessment framework considers the spherical nature of a given ODI with tangential views. These distortion-free tangential views are used to perform a large-scale comparison with eleven perceptual image quality assessment models and state-of-the-art SISR methods for ODI. This study presents the first quality assessment framework to assess the perceptual quality of super-resolved ODIs generated by state-of-the-art SISR methods. Moreover, this is the first evaluation study showcasing the performance of perceptual image quality assessment models on state-of-the-art SISR methods for ODIs.

In this study, we present a straightforward comparison of objective and subjective analysis where we observe that GAN models shows superior performance in subjective tests over conventional CNN SISR architectures. However, contrary to it all the

objective metrics favor the CNN models except a few.

Related Work

We first describe the latest contributions mostly related to our work, which are focused on: 1) quality assessment; 2) projections; 3) super-resolution.

Quality Assessment

Traditional quality metrics for a standard 2D image do not have a high correlation with subjective scores because of inherent sphere-to-planar projection distortions of ODI [4, 5]. Several PSNR-based methods have been proposed to cope with such projection distortions, such as weighted-to-spherical (WS-PSNR) [6], spherical (S-PSNR) [7], and Craster parabolic projection (CPP-PSNR) [8]. In WS-PSNR, the PSNR is applied at each pixel sample of the planar representation of ODI. Each distortion value on the planar is weighted by the area covered by the area on the sphere. In S-PSNR, the PSNR is computed for the sampling points uniformly distributed on a spherical surface. Furthermore, in CPP-PSNR, the PSNR is calculated in the Craster parabolic projection, characterized by low projection distortions. However, the previous related works are based on PSNR, which are not correlated well with subjective scores [9, 10].

Although quality assessment techniques for ODI have been studied in recent years, the perceptual quality of ODI is still unexplored in image restoration techniques, such as super-resolution. In particular, most of the existing studies in this field consider only post-production [11], compression [12], and transmission artifacts [4].

ODI Projections

Sphere-to-planar projections inevitably introduce Sphere-to-planar projections inevitably introduce distortions which must be taken into account to estimate the ODI quality accurately. In particular, Sun et al. [13] proposed a multi-channel CNN for blind ODI quality assessment. They generated six viewport images, representing six cubemap projection faces, to minimize the geometric distortions on the equirectangular images. However, ODIs with the equirectangular or cubemap image representations demonstrates severe image distortions [14]. Also, [10] divides the ODV into a fixed number of planar patches based on the spherical Voronoi diagram.

Instead, in this work, we consider 20-sided regular icosahedron which is the lowest distortion spherical representation [15]. In particular, our quality assessment framework is based on the tangent image spherical representation [16]. With this approach, we can subdivide a given ODI to match the spherical image resolution to the number of low-distortion patches. This approach provides a scalable solution for high-resolution ODIs.

Super-resolution

Thanks to the advanced learning capability [17–19], CNNs have demonstrated superior performance for the super-resolution task compared to the traditional approaches. For instance, an enhanced deep super-resolution network (EDSR) [20] achieved the top PSNR rank in the NTIRE 2017 super-resolution challenge [21]. This method modified the conventional ResNet architecture to the SISR task. The wide activation for efficient and accurate image super-resolution (WDSR) [22] model demon-strated that expanding feature before activation leads to significant improvements for SISR. This network, WDSR, won the first places in NTIRE 2018 super-resolution challenge [23] in three realistic tracks. With the invention of GANs, state-of-the-art SISR methods are providing better perceptual reconstruction results. In SISR, GANs provide a framework to generate plausible-looking natural images with perceptual quality. In particular, SRGAN [2] inspired many follow-up SISR works using GANs like ESRGAN [24] and SFT-GAN [25].

We proposed an improved GAN-based SISR model for ODI and studied the problem of SISR in ODIs [1]. In this work, we observed that quantitative quality measurements do not represent the visual super-resolved ODI results.

Recent SISR methods claim to generate realistic textures and details incorporating feature loss and adversarial learning. However, they reveal their quantitative results with different full-reference quality metrics. Hence, Considering that there is a need to measure and understand super-resolved ODI quality, in this work, we performed subjective quality assessment experiments with diverse super-resolved ODIs, and extensive objective quality assessment using quality metrics on the tangential views.

Objective Quality Assessment Framework

We consider the full-reference objective quality assessment framework to measure super-resolved ODIs, as depicted in Fig. 1. Instead of learning a new model from scratch, our proposed framework allows us to use existing image quality assessment methods to ODI quality measurement. The proposed framework can take any ODI projection representations (e.g., equirectangular, cubemap, etc.) as super-resolved inputs and reference ODIs. We apply spherical tessellation with the icosahedron to subdivide the input ODIs. Tangential views are then extracted using the gnomonic projection as similar as [16]. We specifically focus on the full-reference quality metrics that take the human visual system's properties into account (e.g., SSIM, VIF, etc.). We apply the full-reference quality metrics developed for the traditional image to each tangential views. Finally, the metric scores for all tangential views are averaged to obtain the final objective quality score.

In the following sub-sections, we describes each component of the proposed objective quality assessment framework of super-resolved ODIs in detail.

Input Tangential Views

Inspired by Eder et al [16], we tessellate a sphere where every triangle of the icosahedron has exactly equal area to any desired resolution. For this, we render an image, called tangential view, onto an oriented pixel grids tangent to a subdivided icosahedron. With this approach, we can efficiently scale to represent high-resolution ODIs.

Our solution generates a number of low-distorted tangential views dependent on the resolution of the input ODI. The number of tangential views, N_t, is estimated by the faces of the base level of icosahedron as:

$$N_t = f(4^b), \tag{1}$$

where f is a constant number representing the number of faces of the icosahedron, which is 20. Also, b represents a level of the subdivided icosahedron. In this paper, we propose to use $b =$
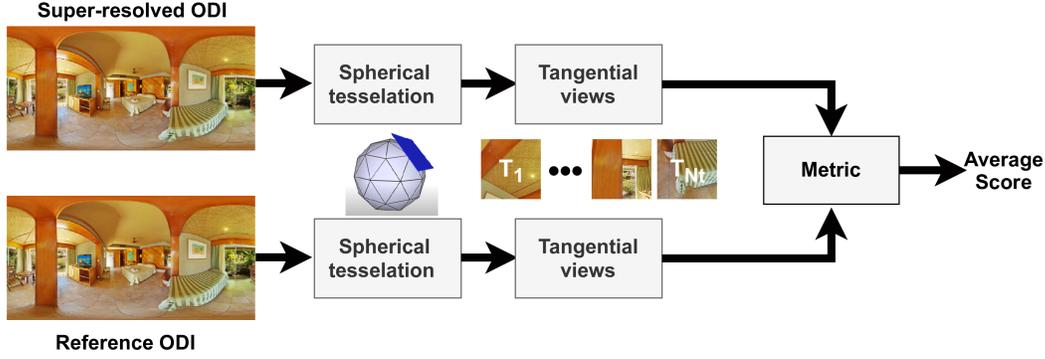
Figure 1. *An overview of the proposed objective quality assessment framework for SISR.*

1 which is the optimal number for many applications [16]. The spatial extent of this equation is a function of vertex resolution of the level $b-1$ icosahedron and the resolution of the image grid, as described in [16].

Quality measure using tangential views

The proposed framework extracts N_t number of tangential views from both the super-resolved ODI and the reference ODI. The traditional full-reference quality metrics are then applied to each tangent views. Full-reference quality metrics compare a given distorted image to the complete reference image. The final objective score, t-metric, is then obtained by computing the average of the tangent images as following:

$$t - metric = \frac{\sum_{i=1}^{N_t} T_i}{N_t}, \quad with \quad T_i = Q(\tilde{t}_i, t_i), \qquad (2)$$

where Q represent the quality score obtained by the full-reference quality metrics. Also, $\tilde{t}_i$ and t_i are i-th tangential views of the super-resolved ODI $\tilde{y}$ and the complete reference image y, respectively.

We extend eleven traditional full-reference image quality metrics to ODI. We call tangential metrics, t-metric, to these extended objective quality metrics. The full-reference objective quality assessment methods used in this study is listed in Table. 1. The selected quality metrics can be classified into four categories as:

- Error visibility methods: These methods measure the distance between the distorted and the reference image pixels. In particular, the classical error estimation solution of mean square error (l_2-norm) and mean absolute error (l_1-norm) are widely used for optimizing the super-resolution algorithms. In this category, we utilized the most apparent distortion (MAD) metric [26] and normalized Laplacian pyramid (NLPD) [27].
- Structural similarity measure: Structural similarity is used for measuring the similarity between the distorted and the reference image [28, 29]. These measurements extract three key features from the image, namely, luminance, contrast, and structure, and combined them for the final measure. In this category, we used structural similarity index (SSIM) [28], multiscale (MS-SSIM) [29], FSIM [30], Gradient magnitude similarity deviation (GMSD) [31], and visual saliency-induced index (VSI) [32].

Category	Metric
Error visibility	MAD [26]
	NLPD [27]
Structural similarity	SSIM [28]
	MS-SSIM [29]
	FSIM [30]
	GMSD [31]
	VSI [32]
Information theory	VIF [33]
	VIFs [33]
Learning-based	LPIPS [34]
	DISTS [35]

Table 1. The used qualitative quality assessment metrics

- Information theoretic estimation: These metrics measure the approximation of the mutual information between the perceived reference and distorted images. Here, we used visual information fidelity (VIF) [33] and spatial domain VIF (VIFs) metric that quantifies how much information from the reference image is preserved in the distorted image.
- Learning-based quality estimation: These approaches learn the quality score by training a set of distorted images and corresponding perceptual distance. In this category, we selected two quality measurements: learned perceptual image patch similarity (LPIPS) [34] model and deep image structure and texture similarity (DISTS) [35].

In this work, we do not consider the earliest energy-based full-reference quality assessment methods, such as mean-squared error and PSNR, which are already shown to be poorly correlated with human perception of super-resolved images.

Experimental Results

This section describes our super-resolution dataset for ODI, objective and subjective quality assessments, and experimental analysis.

Dataset

We selected a total of eight ODIs with a diverse content characteristic from the SUN 360 Panorama Database for objective and subjective testing [36]. This database is the only publicly available ODI dataset with many high-resolution 360-degree images with a wide variety of content, including indoor, outdoor, structures, and human faces. The selected eight ODIs represents four different categories: Structures, Daytime, Indoors, Night-time,

Figure 2. *Dataset sample. The selected ODIs from the SUN 360 Panorama Database [36].*

SISR Models				Metric
EDSR-CNN	WDSR-CNN	EDSR-GAN	WDSR-GAN	
93.05	93.67	113.16	107.62	MAD ↓
0.21	0.21	**0.33**	0.29	NLPD ↑
0.87	**0.87**	0.69	0.73	SSIM ↑
0.95	**0.95**	0.87	0.87	MS-SSIM ↑
0.93	**0.93**	0.88	0.89	FSIM ↑
0.09	0.09	**0.12**	0.11	GMSD ↑
0.98	**0.98**	0.96	0.97	VSI ↑
0.35	**0.35**	0.11	0.13	VIF ↑
0.32	**0.32**	0.19	0.22	VIFs ↑
0.29	**0.29**	0.37	0.36	LPIPS ↓
0.13	0.14	0.15	0.15	DISTS ↓

Table 2. Objective metric results of the proposed quality assessment framework. Here, ↑ means the higher the better, while ↓ means the lower the better. Best are marked in bold.

Textures and Humans. Each ODI is in ERP format with a resolution size of 1920×960. The selected ODIs are visualized in Fig. 2.

We built our SISR dataset for the selected ODIs using two different approaches: CNN-based and GAN-based. For CNN-based methods, we used two state-of-the-art SISR methods: EDSR [20] and WDSR [22]. For GAN-based methods, we fine-tuned EDSR and WDSR with the discriminator of the SR-GAN [2]. Hence, we analyzed the following SISR methods for ODI using our proposed quality assessment framework: EDSR-CNN, EDSR-GAN, WDSR-CNN, WDSR-GAN.

For training of EDSR-CNN and WDSR-CNN, we used the default settings of the two SISR [20, 22] methods with ×4 scaling factor. Our experiments were carried out on ×4 scaling factor between LR and HR. We obtained the LR images by downsampling HR images with bicubic interpolation. We used DIV2K dataset [37] during the training. Data augmentation techniques such as random cropping, flipping, and rotation were applied to scale up the training dataset. To train each model, we set the batch size to 16. The weights of all the layers were initialized randomly, and the network was trained from scratch. Further, to optimize the network, we used the ADAM solver with initial learning rates of $1e-4$ and $1e-3$ for EDSR and WDSR, respectively. We fine-tuned EDSR-CNN and WDSR-CNN using the discriminator of the SRGAN [2].

Objective quality assessment

We performed objective quality assessment experiments using the proposed objective quality assessment framework. In this experiment, we utilized eleven image quality metrics as described in Section : Input Tangential Views. For this, we used the PyTorch implementations released by Keyan et al. [38].

Table 2 shows the objective quality results using the proposed quality assessment framework. Objective quality results reveal that majority of the quality metrics shows higher results for CNN-based models (EDSR-CNN and WDSR-CNN) compared to GAN-based models (EDSR-GAN and WDSR-GAN). Interestingly, two quality metrics, NLPD and GMSD show the superiority of the EDSR-GAN model. The reason could be that these two metrics take the human visual system into account and are highly accurate for predicting perceptual quality on noisy images.

Subjective quality assessment

We performed a subjective pairwise comparison to validate the perceived quality of different super-resolved ODIs generated by the state-of-the-art methods. In particular, we are interested to investigate CNN-based versus GAN-based SISR results.

Test Environment and Setup

The study was performed in a crowd-sourced format by showcasing a video-like representation of the ODIs. A total of 20 people participated in this subjective study, with age range of 23-40 years, normal or corrected-to-normal vision. We generated a viewport video with 10 seconds for each ODI using the 360Lib software [39]. The viewport was generated by using the rectilinear projection along with the given field-of-view information. Dynamic viewport setting was selected to create a video-like viewport representation for each ODI.

Participants were able to access the experiments using their web-browser where two super-resolved ODIs corresponding to a scene is displayed for 10 seconds. A total of 48 pairwise combinations were generated from eight scenes and their four super-resolved images obtained using EDSR-CNN, EDSR-GAN, WDSR-CNN and WDSR-GAN models.

Procedure and Analysis

Procedure: A pair-wise subjective experimental study was performed where each observer was asked to choose an image by showing a pair of images side-by-side. We excluded the option of showing the same image type to force the users to choose one of the stimuli. Each participant was asked to select an image which is more natural and realistic-looking. Participants were provided with unlimited time to make their decision and record their selection. The experiment was divided into a training and test session, where training involved each participant being briefed to familiarize with the subjective quality evaluation task. Each observer compared a pair of image twice, having each super-resolved image displayed on both sides. Figure 4 shows the subjective experiment setup shown to participants.

Analysis: We compute the statistical analysis of pairwise subjective tests. To this end, we scaled the winning frequencies of each SISR model to the continuous quality-scores using the widely known Bradley-Terry (BT) model. The scaling is performed using the statistical analysis proposed in to determine whether the perceived visual quality difference of the compared models is statistically significant. The preference probabil-
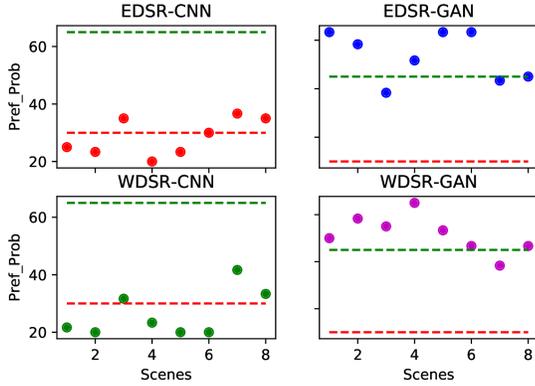
Figure 3. *Subjective Test Results. Preference probability of four techniques (EDSR-CNN, WDSR-CNN, EDSR-GAN, WDSR-GAN) for 8 ODI scenes.*

Figure 4. *ODI shown side by side during Subjective Test*

ity Pref-Prob($Pref - Prob_m$) is mathematically given as:

$$Pref - Prob_m = \frac{w_m}{n} + \frac{\tau}{2n}, \tag{3}$$

where w is the winning probability of the method m. n is the total number of participants and τ is the tie frequency. Given the statistical basis of the subjective study, each method in the pairwise comparison shares equal probability of occurrence (*i.e.*, 0.5) when compared with any other method and hence, the analysis follows a binomial distribution. Based on different state of the art subjective studies to establish initial hypothesis, we followed and performed a Binomial test on the set of collected data and the critical thresholds were determined by plotting the cumulative distribution function of Binomial distribution. We set 95% as the upper level of significance, if we receive $13(B(13,20,0.5) = 0.9423)$ or more preference votes for a given method over another competing method, we consider the given method to be significantly favored in terms of subjective quality with more natural appeal. Similarly, we set 5% as the lower significance level, if we receive $6(B(6,20,0.5) = 0.0577)$ or less votes for a given method, we consider the given method for an ODI to be least favored in terms of subjective quality.

The subjective tests results for four different methods are shown in Figure 3. The green and red line mark probabilities of high ($13/20 = .65$) and low ($6/20 = .30$) favor-abilities respectively. Our subjective tests results demonstrate that images generated with GANs are more perceptually appealing for different types of scenes compared to CNNs based models.

Comparison

We showcase the comparisons between objective and subjective analysis in Table 3 using the average preference probability of EDSR-CNN, EDSR-GAN, WDSR-CNN, WDSR-GAN models. These scores are computed using the equation 3 by performing a pairwise comparison between objective scores of 4 modalities of

SISR Models				Metric
EDSR-CNN	WDSR-CNN	EDSR-GAN	WDSR-GAN	
4.76	11.90	50	33.33	MAD ↓
7.14	7.14	50	33.33	NLPD ↑
35.71	38.09	0	16.67	SSIM ↑
42.85	35.71	0	16.67	MS-SSIM ↑
38.09	35.71	2.38	11.90	FSIM ↑
2.38	9.52	50	33.33	GMSD ↑
35.71	33.33	0	16.67	VSI ↑
42.85	40.48	0	16.67	VIF ↑
38.09	35.71	0	16.67	VIFs↑
7.14	7.14	42.86	40.48	LPIPS↓
14.28	23.81	33.33	26.19	DISTS ↓
14.27	13.23	36.67	35.83	Sub. Score

Table.3 Comparisons of Preference Probabilities of the objective and subjective studies computed over 8 scenes.

Figure 5. *Visual comparison between EDSR-GAN (left) and EDSR-CNN (right).*

a given scene and metric. All 8 scenes that are used in subjective experiments are utilized to obtain the average probability scores. For metrics where higher scores signify better performance, models with higher average probability scores are better and vice versa (shown with up and down arrows).

While GAN models show superior performance in subjective tests over conventional CNN architectures, all the objective metrics clearly favor the CNN models except the GMSD and NLPD metrics. GANs super-resolved output contains sharp details compared to blurred out CNN SISR network output, which appears realistic perceptually. Such sharp details are however contains visually-imperceptible noise, which hampers objective metric performance. Fig 5 shows the comparison between GAN-based (EDSR-GAN) and CNN-based (EDSR-CNN) SISR outputs.

GMSD and NLPD metric show high correlation with subjective analysis and hold potential for designing optimal SISR models for ODIs in future. Other metrics, however, appear to be sub-optimal to be used for quality assessment of super-resolved ODIs and need further introspection.

Conclusion

We presented an objective quality measurement framework which is able to measure super-resolution quality on tangential views. Distortion-free tangential views helps in comparing the subjective analysis and the results from objective quality metrics that are originally designed for 2D image compression methods. In addition, we investigated CNN-based and GAN-based super-resolution models in the objective and subjective experiments. Our analysis showcased how GAN-based super-resolution models show superior performance in subjective tests over conventional CNN architectures. Also, we found that all the objective metrics clearly favor the CNN models except the GMSD and NLPD metric. Our future work will focus on designing a robust SISR models for ODIs.

References

[1] C. Ozcinar, A. Rana, and A. Smolic, "Super-resolution of omnidirectional images using adversarial learning," in *2019 IEEE 21st*

International Workshop on Multimedia Signal Processing (MMSP), 2019, pp. 1–6.

[2] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[3] G. Jinjin, C. Haoming, C. Haoyu, Y. Xiaoxing, J. S. Ren, and D. Chao, "Pipal: a large-scale image quality assessment dataset for perceptual image restoration," in *European Conference on Computer Vision*. Springer, 2020, pp. 633–651.

[4] S. Croci, C. Ozcinar, E. Zerman, J. Cabrera, and A. Smolic, "Voronoi-based objective quality metrics for omnidirectional video," in *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2019, pp. 1–6.

[5] G. d. A. Roberto, N. Birkbeck, I. Janatra, B. Adsumilli, and P. Frossard, "A viewport-driven multi-metric fusion approach for 360-degree video quality assessment," in *2020 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2020.

[6] Y. Sun, A. Lu, and L. Yu, "Weighted-to-spherically-uniform quality evaluation for omnidirectional video," *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1408–1412, Sep. 2017.

[7] M. Yu, H. Lakshman, and B. Girod, "A framework to evaluate omnidirectional video coding schemes," in *2015 IEEE International Symposium on Mixed and Augmented Reality*, Sept 2015, pp. 31–36.

[8] V. Zakharchenko, K. P. Choi, and J. H. Park, "Quality metric for spherical panoramic video," *Proc.SPIE*, vol. 9970, pp. 9970 – 9970 – 9, 2016.

[9] E. Upenik, M. Rerabek, and T. Ebrahimi, "On the performance of objective metrics for omnidirectional visual content," in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, 2017, pp. 1–6.

[10] S. Croci, C. Ozcinar, E. Zerman, S. Knorr, J. Cabrera, and A. Smolic, "Visual attention-aware quality estimation framework for omnidirectional video using spherical voronoi diagram," *Quality and User Experience*, vol. 5, no. 1, pp. 1–17, 2020.

[11] S. Croci, S. Knorr, L. Goldmann, and A. Smolic, "A framework for quality control in cinematic vr based on voronoi patches and saliency," in *2017 International Conference on 3D Immersion (IC3D)*, 2017, pp. 1–8.

[12] H. T. T. Tran, N. P. Ngoc, C. M. Bui, M. H. Pham, and T. C. Thang, "An evaluation of quality metrics for 360 videos," in *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2017, pp. 7–11.

[13] W. Sun, X. Min, G. Zhai, K. Gu, H. Duan, and S. Ma, "Mc360iqa: A multi-channel cnn for blind 360-degree image quality assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 1, pp. 64–77, 2019.

[14] M. Eder and J.-M. Frahm, "Convolutions on spherical images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 1–5.

[15] J. A. Kimerling, K. Sahr, D. White, and L. Song, "Comparing geometrical properties of global grids," *Cartography and Geographic Information Science*, vol. 26, no. 4, pp. 271–288, 1999.

[16] M. Eder, M. Shvets, J. Lim, and J.-M. Frahm, "Tangent images for mitigating spherical distortion," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 12 426–12 434.

[17] A. Rana, G. Valenzise, and F. Dufaux, "Learning-Based Tone Mapping Operator for Image Matching," in *IEEE International Conference on Image Processing (ICIP)*. Beijing, China: IEEE, 2017.

[18] S. M. I. Zolanvari, S. Ruano, A. Rana, A. Cummins, R. Silva, M. Rahbar, and A. Smolic, "Dublincity: Annotated lidar point cloud and its applications," in *30th British Machine Vision Conference 2019, BMVC 2019, Cardiff, UK, September 9-12, 2019*.

[19] A. Rana, G. Valenzise, and F. Dufaux, "Learning-based Adaptive Tone Mapping for Keypoint Detection," in *IEEE International Conference on Multimedia & Expo (ICME)*, Hong Kong, China, Jul. 2017.

[20] B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings*

[21] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang *et al.*, "Ntire 2017 challenge on single image super-resolution: Methods and results," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, July.

[22] J. Yu, Y. Fan, J. Yang, N. Xu, Z. Wang, X. Wang, and T. Huang, "Wide activation for efficient and accurate image super-resolution," *arXiv preprint arXiv:1808.08718*, 2018.

[23] R. Timofte, S. Gu, J. Wu, and L. Van Gool, "Ntire 2018 challenge on single image super-resolution: Methods and results," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 852–863.

[24] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 0–0.

[25] X. Wang, K. Yu, C. Dong, and C. Change Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 606–615.

[26] E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of electronic imaging*, vol. 19, no. 1, p. 011006, 2010.

[27] V. Laparra, J. Ballé, A. Berardino, and E. P. Simoncelli, "Perceptual image quality assessment using a normalized laplacian pyramid," *Electronic Imaging*, vol. 2016, no. 16, pp. 1–6, 2016.

[28] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.

[29] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2. Ieee, 2003, pp. 1398–1402.

[30] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.

[31] W. Xue, L. Zhang, X. Mou, and A. C. Bovik, "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 684–695, 2013.

[32] L. Zhang, Y. Shen, and H. Li, "Vsi: A visual saliency-induced index for perceptual image quality assessment," *IEEE Transactions on Image processing*, vol. 23, no. 10, pp. 4270–4281, 2014.

[33] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on image processing*, vol. 15, no. 2, pp. 430–444, 2006.

[34] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.

[35] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Image quality assessment: Unifying structure and texture similarity," *arXiv preprint arXiv:2004.07728*, 2020.

[36] J. Xiao, K. A. Ehinger, A. Oliva, and A. Torralba, "Recognizing scene viewpoint using panoramic place representation," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 2695–2702.

[37] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 1122–1131.

[38] K. Ding, K. Ma, S. Wang, and E. P. Simoncelli, "Comparison of image quality models for optimization of image processing systems," *arXiv*, pp. arXiv–2005, 2020.

[39] Y. Ye, E. Alshina, and J. Boyce, "Algorithm descriptions of projection format conversion and video quality metrics in 360lib," *Joint Video E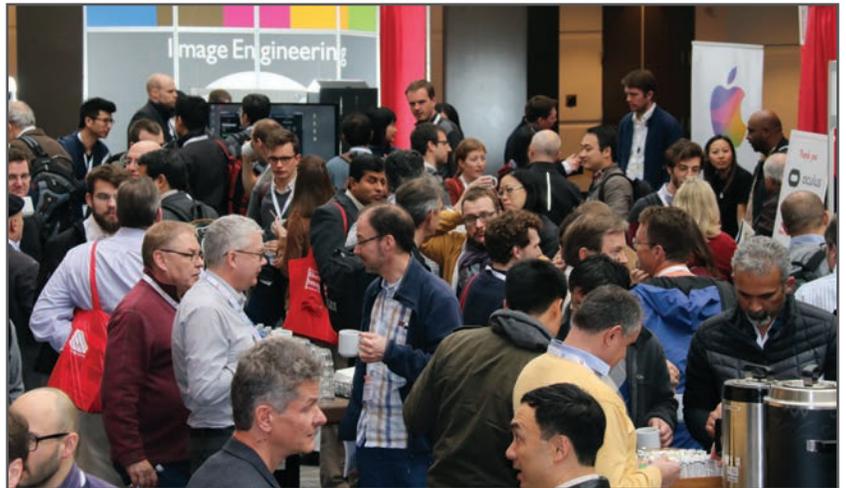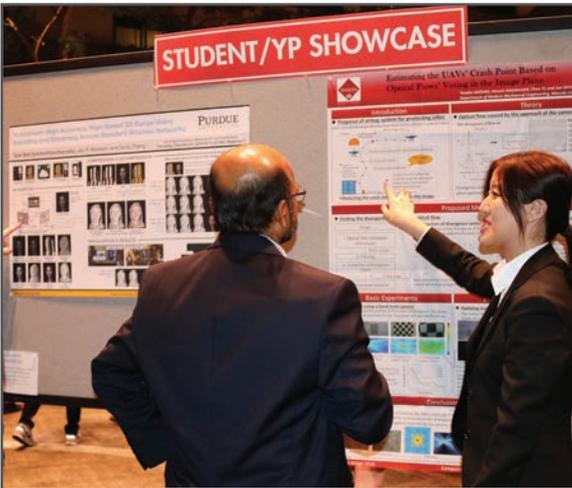xploration Team of ITU-T SG*, vol. 16, 2017.