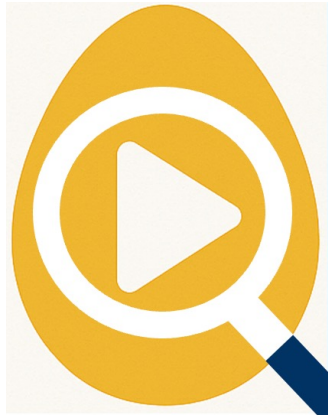


AIGCOIQA2024: Perceptual Quality Assessment of AI Generated Omnidirectional Images



Liu Yang, Huiyu Duan, Long Teng, Yucheng Zhu, Xiaohong Liu, Menghan Hu, Xiongkuo Min, Guangtao Zhai, and **Patrick Le Callet**

(Shanghai Jiao Tong University, East China Normal University, **Nantes Université**)



VQEG meeting Spring 2025

AIGCOIQA2024: Perceptual Quality Assessment of AI Generated Omnidirectional Images



Liu Yang, Huiyu Duan, Long Teng, Yucheng Zhu, Xiaohong Liu, Menghan Hu, Xiongkuo Min, Guangtao Zhai, and **Patrick Le Callet**

(Shanghai Jiao Tong University, East China Normal University, **Nantes Université**)



VQEG meeting Spring 2025

AI Generated Omnidirectional Images (AIGODIs):

AI generated omnidirectional images hold significant potential for Virtual Reality (VR) and Augmented Reality (AR) applications.

However, many omnidirectional images generated by AI are of low quality and cannot meet the visual expectations of the users

=> IQA studies for AI-generated omnidirectional images which have significant difference compared with classical image in terms of formats, characteristics, and applications



300 AIGODIs generated by:

- 25 scene prompts (12 indoor and 13 outdoor)
- 5 omnidirectional generation models

Database: <https://github.com/IntMeGroup/AIGCOIQA>

25 prompts?

Source: 25 omnidirectional images from the MVDiffusion¹ training dataset and SUN360² as **natural instances**

Preliminary prompts: using BLIP2³ to annotate the images
Refined using GPT4 to add details and polish

=> prompts containing abundant scene details, making the generated images more distinguishable in terms of text-image correspondence

1. Tang and al. MVDiffusion: Enabling holistic multi-view image generation with correspondence-aware diffusion,” NeurIPS 2023
2. Xiao and al. “Recognizing scene viewpoint using panoramic place representation,” in CVPR 2012
3. Li and al., “BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models,” arXiv 2023

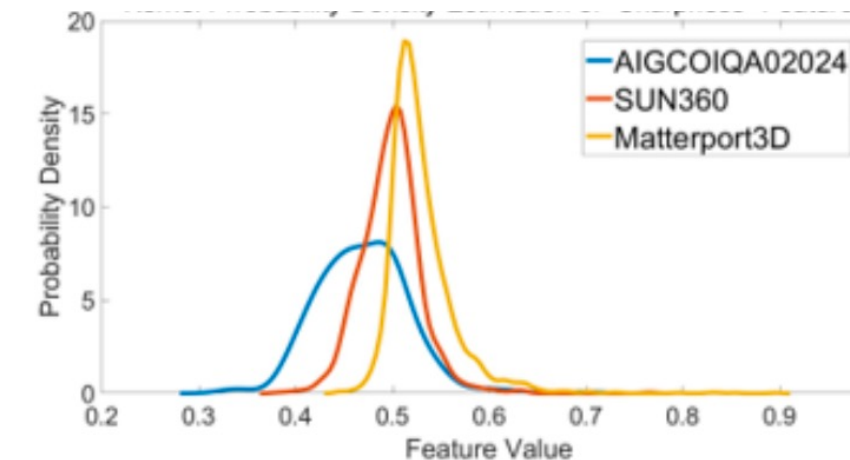
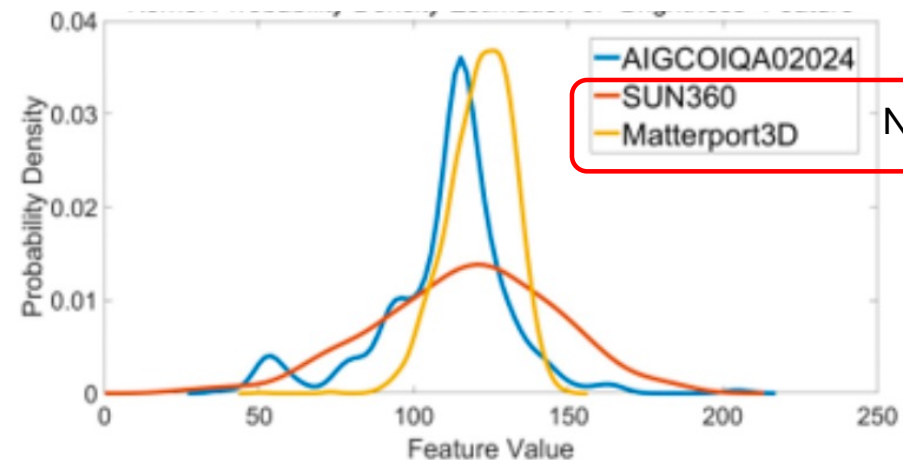
5 omnidirectional generation models?

- MVDiffusion
- Text2Light
- DALLE
- Omni-inpainting
- and a fine-tuned Stable Diffusion model to generate omnidirectional images.
 - => fine-tuning Stable Diffusion model: 2000 and 4000 BLIP2- labeled indoor/outdoor ERP omnidirectional images to fine-tune the Unet module

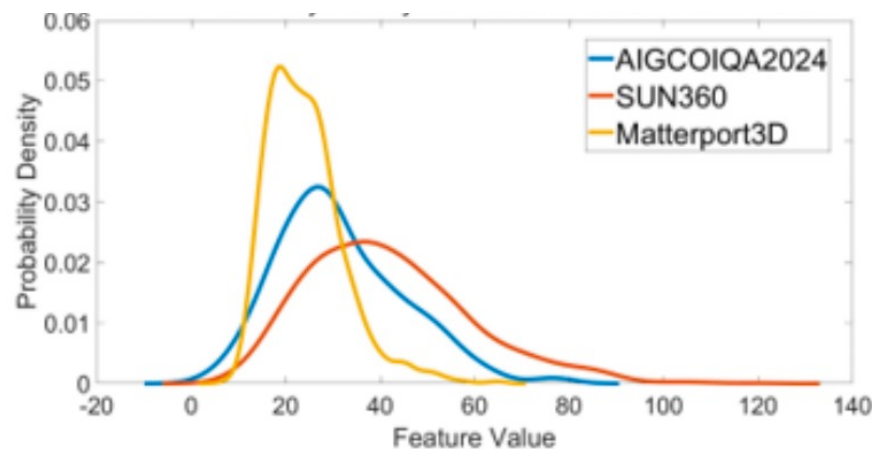
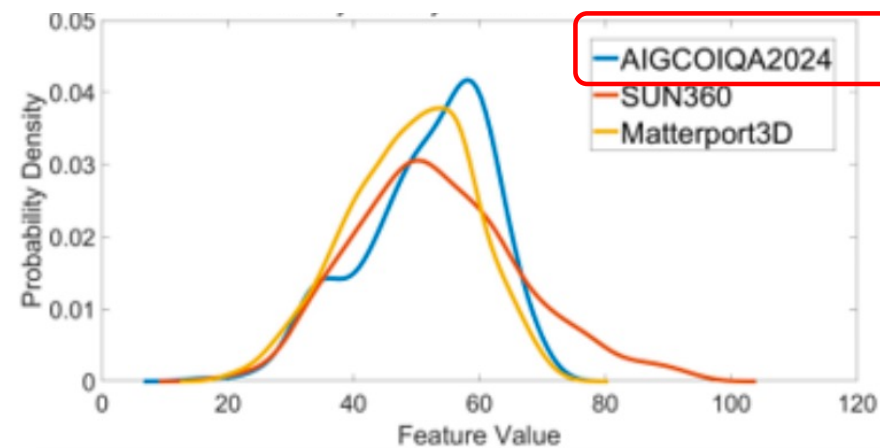
1. Tang and al. MVDiffusion: Enabling holistic multi-view image generation with correspondence-aware diffusion,” NeurIPS 2023
2. Xiao and al. “Recognizing scene viewpoint using panoramic place representation,” in CVPR 2012
3. Li and al., “BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models,” arXiv 2023

Statistical analysis of generated images (PVS)

Distribution of 4 features: sharpness, contrast, colorfoulness, brightness



Natural



Generated

Subjective scores?

Subjective cores: 3 dimensions

“Quality”: comprehensive score of low-level visual qualities such as color, lighting and clarity, etc.

“Comfortability” (VR related), subjects are asked to give an overall score for the authenticity level, structure deformation level, as well as the comfort level.

“Text-imagecorrespondence” (Prompt related) correspondence between text and

Preference from 3 Perspectives: Quality, Comfortability, Correspondence



(a) Prompt: “A panorama view of an urban street with a brick building, empty paved roads, street lamps, and clear skies.”

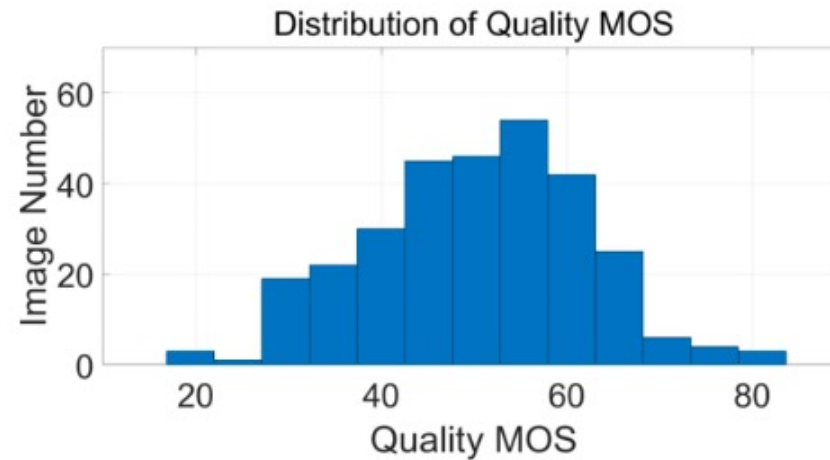
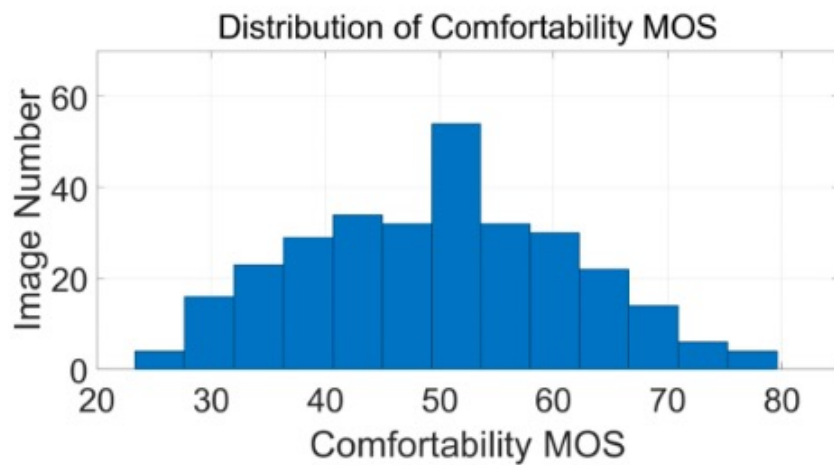
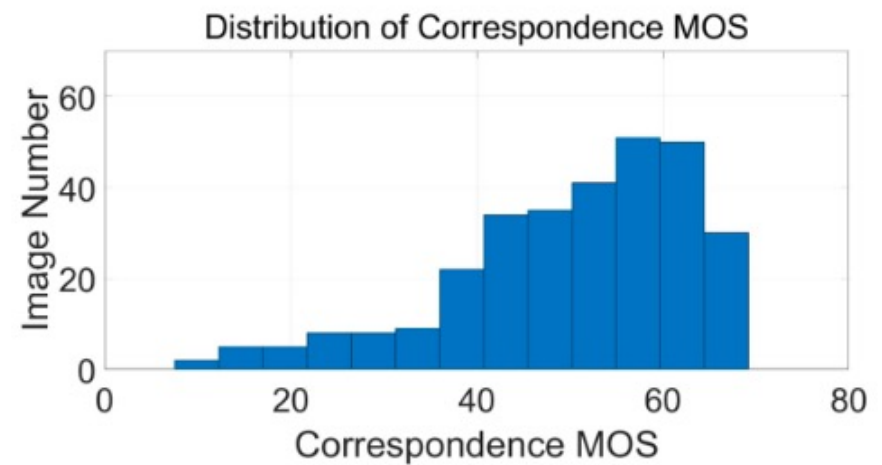


(b) Prompt: “A panorama view of an expansive parking lot with a red van hitched to a trailer under a bright sun with fluffy clouds overhead.”



(c) Prompt: “A panorama view of an opulent room with a crystal chandelier, a large painting, and a wooden display cabinet filled with collectibles.”

MOS distribution on each dimension



Objective assessment?
(NR metrics)

Dimension	Quality		
Model	SRCC	KRCC	PLCC
QAC [27]	0.1802	0.1197	0.0933
BMPRI [28]	0.4285	0.3010	0.5612
NIQE [29]	0.6858	0.4814	0.6236
ILNIQE [30]	0.0434	0.0293	0.0617
HOSA [31]	0.7114	0.5154	0.7175
BPRI-PSS [32]	0.3673	0.2465	0.4299
BPRI-LSSs [32]	0.3018	0.2094	0.4398
BPRI-LSSn [32]	0.3604	0.2300	0.5095
BPRI [32]	0.3553	0.2503	0.4849
FISBLIM [33]	0.6472	0.4588	0.6124
CLIPScore [11]	0.3308	0.2241	0.3320
CNNIQA [34]	0.7066	0.5127	0.6345
Resnet18 [35]	0.7722	0.5852	0.7286
Resnet34 [35]	0.6343	0.4632	0.6478
VGG16 [36]	0.7956	0.6047	0.7351
VGG19 [36]	0.7628	0.5731	0.7105
HyperIQA [37]	0.8354	0.6405	0.7769
MANIQA [38]	0.8038	0.6160	0.7682
TReS [39]	0.8355	0.6378	0.7803

Handcrafted-based models

Vision-language pretrained: score= cosine similarity
between text and
image embeddings

Deep learning-based: retrained 3:2 ratio

Dimension	Quality			Comfortability		
Model	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC
QAC [27]	0.1802	0.1197	0.0933			
BMPRI [28]	0.4285	0.3010	0.5612			
NIQE [29]	0.6858	0.4814	0.6236			
ILNIQE [30]	0.0434	0.0293	0.0617			
HOSA [31]	0.7114	0.5154	0.7175			
BPRI-PSS [32]	0.3673	0.2465	0.4299			
BPRI-LSSs [32]	0.3018	0.2094	0.4398			
BPRI-LSSn [32]	0.3604	0.2300	0.5095			
BPRI [32]	0.3553	0.2503	0.4849			
FISBLIM [33]	0.6472	0.4588	0.6124			
CLIPScore [11]	0.3308	0.2241	0.3320	0.1752	0.1192	0.1666
CNNIQA [34]	0.7066	0.5127	0.6345	0.5715	0.4068	0.5709
Resnet18 [35]	0.7722	0.5852	0.7286	0.6537	0.4728	0.6047
Resnet34 [35]	0.6343	0.4632	0.6478	0.6551	0.4747	0.6177
VGG16 [36]	0.7956	0.6047	0.7351	0.7074	0.5309	0.6616
VGG19 [36]	0.7628	0.5731	0.7105	0.7310	0.5439	0.6893
HyperIQA [37]	0.8354	0.6405	0.7769	0.7477	0.5564	0.7516
MANIQA [38]	0.8038	0.6160	0.7682	0.7837	0.5906	0.7796
TReS [39]	0.8355	0.6378	0.7803	0.7768	0.5857	0.7763

Dimension	Quality			Comfortability			Correspondence		
Model	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC
QAC [27]	0.1802	0.1197	0.0933						
BMPRI [28]	0.4285	0.3010	0.5612						
NIQE [29]	0.6858	0.4814	0.6236						
ILNIQE [30]	0.0434	0.0293	0.0617						
HOSA [31]	0.7114	0.5154	0.7175						
BPRI-PSS [32]	0.3673	0.2465	0.4299						
BPRI-LSSs [32]	0.3018	0.2094	0.4398						
BPRI-LSSn [32]	0.3604	0.2300	0.5095						
BPRI [32]	0.3553	0.2503	0.4849						
FISBLIM [33]	0.6472	0.4588	0.6124						
CLIPScore [11]	0.3308	0.2241	0.3320	0.1752	0.1192	0.1666	0.3809	0.2641	0.4915
CNNIQA [34]	0.7066	0.5127	0.6345	0.5715	0.4068	0.5709	0.5935	0.4222	0.5291
Resnet18 [35]	0.7722	0.5852	0.7286	0.6537	0.4728	0.6047	0.6709	0.4801	0.6829
Resnet34 [35]	0.6343	0.4632	0.6478	0.6551	0.4747	0.6177	0.5414	0.3848	0.6407
VGG16 [36]	0.7956	0.6047	0.7351	0.7074	0.5309	0.6616	0.6598	0.4747	0.6936
VGG19 [36]	0.7628	0.5731	0.7105	0.7310	0.5439	0.6893	0.6773	0.4927	0.7054
HyperIQA [37]	0.8354	0.6405	0.7769	0.7477	0.5564	0.7516	0.6506	0.4701	0.7052
MANIQA [38]	0.8038	0.6160	0.7682	0.7837	0.5906	0.7796	0.7198	0.5303	0.7598
TReS [39]	0.8355	0.6378	0.7803	0.7768	0.5857	0.7763	0.7036	0.5174	0.6996

AIGCOIQA2024: Perceptual Quality Assessment of AI Generated Omnidirectional Images

Database: <https://github.com/IntMeGroup/AIGCOIQA>

VQ



Dimension	Quality			Comfortability			Correspondence		
Model	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC	SRCC	KRCC	PLCC
QAC [27]	0.1802	0.1197	0.0933	0.2665	0.1823	0.2483	0.1756	0.1175	0.0023
BMPRI [28]	0.4285	0.3010	0.5612	0.2652	0.1776	0.3728	0.2705	0.1780	0.3092
NIQE [29]	0.6858	0.4814	0.6236	0.5741	0.4128	0.6072	0.6672	0.4786	0.5776
ILNIQE [30]	0.0434	0.0293	0.0617	0.0298	0.0186	0.0902	0.2959	0.1964	0.3055
HOSA [31]	0.7114	0.5154	0.7175	0.4793	0.3282	0.5020	0.7200	0.5262	0.6738
BPRI-PSS [32]	0.3673	0.2465	0.4299	0.2153	0.1427	0.2216	0.5782	0.4101	0.6491
BPRI-LSSs [32]	0.3018	0.2094	0.4398	0.2435	0.1668	0.3516	0.1004	0.0672	0.1536
BPRI-LSSn [32]	0.3604	0.2300	0.5095	0.1506	0.0825	0.3268	0.5532	0.3878	0.5433
BPRI [32]	0.3553	0.2503	0.4849	0.2866	0.1986	0.4092	0.1863	0.1269	0.2319
FISBLIM [33]	0.6472	0.4588	0.6124	0.4425	0.3015	0.3492	0.6836	0.4929	0.5493
CLIPScore [11]	0.3308	0.2241	0.3320	0.1752	0.1192	0.1666	0.3809	0.2641	0.4915
CNNIQA [34]	0.7066	0.5127	0.6345	0.5715	0.4068	0.5709	0.5935	0.4222	0.5291
Resnet18 [35]	0.7722	0.5852	0.7286	0.6537	0.4728	0.6047	0.6709	0.4801	0.6829
Resnet34 [35]	0.6343	0.4632	0.6478	0.6551	0.4747	0.6177	0.5414	0.3848	0.6407
VGG16 [36]	0.7956	0.6047	0.7351	0.7074	0.5309	0.6616	0.6598	0.4747	0.6936
VGG19 [36]	0.7628	0.5731	0.7105	0.7310	0.5439	0.6893	0.6773	0.4927	0.7054
HyperIQA [37]	0.8354	0.6405	0.7769	0.7477	0.5564	0.7516	0.6506	0.4701	0.7052
MANIQA [38]	0.8038	0.6160	0.7682	0.7837	0.5906	0.7796	0.7198	0.5303	0.7598
TReS [39]	0.8355	0.6378	0.7803	0.7768	0.5857	0.7763	0.7036	0.5174	0.6996