Assessor360: Multi-sequence Network for Blind Omnidirectional Image Quality Assessment

Tianhe Wu[1,*], Shuwei Shi[1,2*], Haoming Cai[3], Mingdeng Cao[2], Jing Xiao[4], Yinqiang Zheng[2], Yujiu Yang[1†]

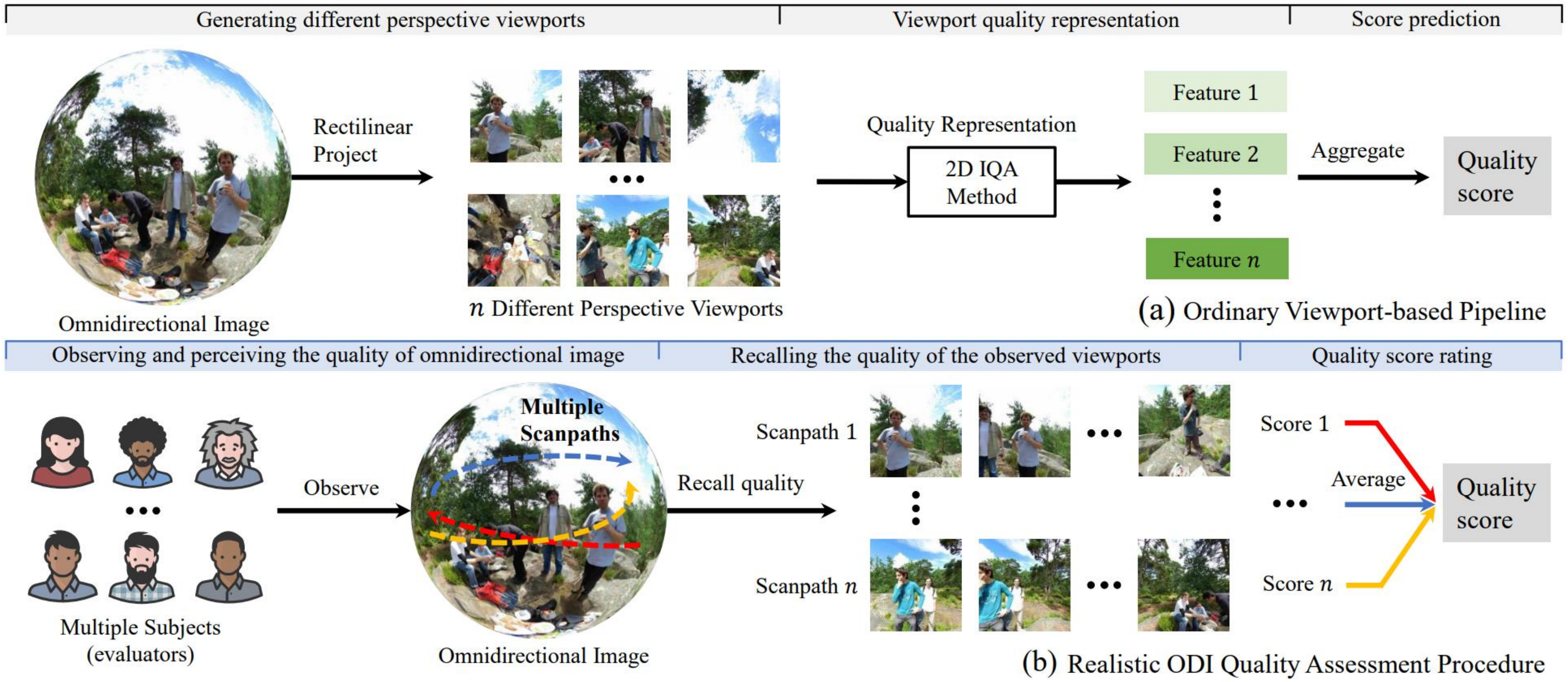
[1] Shenzhen International Graduate School, Tsinghua University, [2] The University of Tokyo,
[3] University of Maryland, College Park, [4] Pingan Group

https://github.com/TianheWu/Assessor360

Speaker: Tianhe Wu
Email: wth22@mails.tsinghua.edu.cn
2023.11.10

Introduction

- Evaluate the quality of panoramas (VR images)
- Lacks the modeling of the observer's browsing process



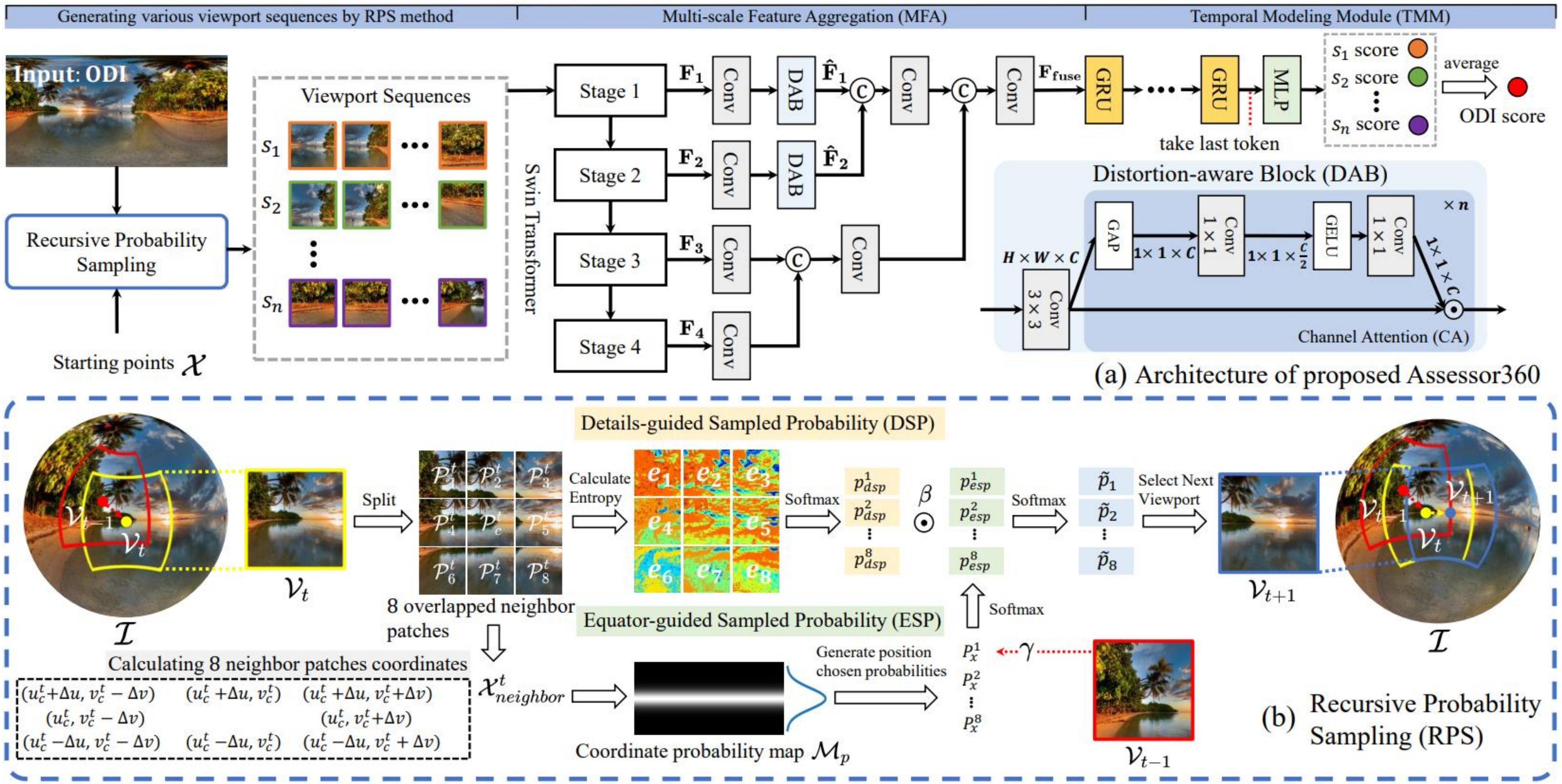
Overall Pipeline

- Generating viewport sequences **(step 1)**
- Representing distorted features and regressing quality score **(step 2 and step 3)**

$$\mathcal{Q}_{\mathcal{I}} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{H}(\mathcal{F}(\mathcal{G}(\mathcal{I}, \mathcal{X}; \Theta_g); \Theta_f); \Theta_h)$$

- Recursive Probability Sampling
  - Equator-guided Sampled Probability (ESP)
  - Details-guided Sampled Probability (DSP)
- Multi-scale Feature Aggregation
- Temporal Modeling Module

Methods



(a) Architecture of proposed Assessor360

(b) Recursive Probability Sampling (RPS)

Performance and Generalization

Table 1: Quantitative comparison of the state-of-the-art methods and proposed Assessor360. The best are shown in **bold**, and the second best (except ours) are underlined. Two baselines w/ ERP and w/ CMP mean that we replace input viewport sequences generated by RPS with ERP and CMP.

Type	Method	MVAQD		OIQA		IQA-ODI		CVIQD	
		SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
FR-IQA methods	PSNR	0.8150	0.7591	0.3929	0.3893	0.4018	0.4890	0.8015	0.8425
	SSIM [45]	0.8272	0.7202	0.3402	0.2307	0.5014	0.5686	0.6737	0.7273
	MS-SSIM [46]	0.8032	0.7136	0.5750	0.5084	0.7434	0.8389	0.9218	0.9272
	WS-PSNR [43]	0.8152	0.7638	0.3829	0.3678	0.3780	0.4708	0.8039	0.8410
	WS-SSIM [67]	0.8236	0.5328	0.6020	0.3537	0.5325	0.7098	0.8632	0.7672
	VIF [60]	0.8687	0.8436	0.4284	0.4158	0.7109	0.7696	0.9502	0.9370
	DISTS [9]	0.7911	0.7440	0.5740	0.5809	0.8513	0.8723	0.8771	0.8613
	LPIPS [61]	0.8048	0.7336	0.5844	0.4292	0.7355	0.7411	0.8236	0.8242
NR-IQA methods	NIQE [30]	0.6785	0.6880	0.8539	0.7850	0.6645	0.5637	0.9337	0.8392
	BRISQUE [29]	0.8408	0.8345	0.8213	0.8206	0.8171	0.8651	0.8269	0.8199
	PaQ-2-PiQ [56]	0.3251	0.3643	0.1667	0.2102	0.0201	0.0419	0.7376	0.6500
	MANIQA [54]	0.5531	0.5718	0.4555	0.4171	0.2642	0.2776	0.6013	0.6142
	MUSIQ [22]	0.5436	0.6117	0.3216	0.3087	0.0565	0.0983	0.3483	0.3678
	CLIP-IQA [44]	0.5862	0.4941	0.2330	0.2531	0.0927	0.1929	0.4884	0.4347
	LIQE [63]	0.6837	0.7539	0.7634	0.7419	0.8551	0.9020	0.8594	0.8086
	SSP-BOIQA [64]	0.7838	0.8406	0.8650	0.8600	-	-	0.8614	0.9077
	MP-BOIQA [20]	0.8420	0.8543	0.9066	0.9206	-	-	0.9235	0.9390
	MC360IQA [42]	0.6605	0.6977	0.9071	0.8925	0.8248	0.8629	0.8271	0.8240
	SAP-net [52]	-	-	-	-	0.9036	0.9258	-	-
	VGCN [49]	0.8422	0.9112	0.9515	0.9584	0.8117	0.8823	0.9639	0.9651
	AHGCN [16]	-	-	0.9647	0.9682	-	-	0.9617	0.9658
	baseline w/ ERP	0.9076	0.9240	0.8961	0.8857	0.9098	0.9196	0.9330	0.9485
	baseline w/ CMP	0.8966	0.9324	0.9216	0.9170	0.9105	0.9122	0.9390	0.9412
	Assessor360	**0.9607**	**0.9720**	**0.9802**	**0.9747**	**0.9573**	**0.9626**	**0.9644**	**0.9769**

Comparing with SoTA methods

Table 2: Cross-dataset validation SRCC and PLCC results of SOTA methods. These models (except WS-PSNR [43] and WS-SSIM [67]) are trained on CVIQD [41], OIQA [12] and MVAQD [21] datasets (80% set) and tested on three other datasets (full set).

Method	CVIQD			OIQA			MVAQD		
	OIQA	IQA-ODI	MVAQD	CVIQD	IQA-ODI	MVAQD	CVIQD	OIQA	IQA-ODI
SRCC									
WS-PSNR	0.5027	0.4360	0.7225	0.7638	0.4360	0.7225	0.7638	0.5027	0.4360
WS-SSIM	**0.5442**	0.5032	**0.7930**	0.6625	0.5032	**0.7930**	0.6625	0.5442	0.5032
MC360IQA	0.4189	0.7114	0.0296	0.7044	0.5687	0.4081	0.0373	0.0025	0.0486
VGCN	0.2361	0.2875	0.2452	0.6932	0.3873	0.4682	0.4650	0.6227	0.3921
Assessor360	0.4597	**0.8610**	0.5640	**0.8430**	**0.8751**	0.6417	**0.8756**	**0.7765**	**0.8646**
PLCC									
WS-PSNR	0.4701	0.5468	**0.6962**	0.7895	0.5468	**0.6962**	**0.7895**	0.4701	0.5468
WS-SSIM	0.4363	0.5941	0.6246	0.6536	0.5941	0.6246	0.6536	0.4363	0.5941
MC360IQA	0.4295	0.7872	0.0404	0.7368	0.5930	0.4238	0.0430	0.0202	0.0646
VGCN	0.2582	0.3127	0.2467	0.5929	0.3551	0.2419	0.3420	0.4642	0.3870
Assessor360	**0.5332**	**0.9032**	0.5824	**0.8636**	**0.9137**	0.6565	0.7232	**0.7287**	**0.8541**

Generalization

Table 3: Cross-dataset validation SRCC results of SOTA methods. These models are trained on MVAQD [21] (full set) and tested on CVIQD [41] and OIQA [12] all distortion data except MP-BOIQA (removing AVC on CVIQD).

Testing set	MC360IQA	MP-BOIQA	**Assessor360**
OIQA	0.2542	0.5043	**0.6658**
CVIQD	0.4749	0.7992	**0.8994**

Generalization

RPS Ablation Studies

Table 5: Quantitative comparison of different generation methods (**RPS vs ScanGAN360 [28] and ScanDMM [38]**) with metrics of scanpath prediction task on JUFE [14] and JXUFE [39] datasets with authentic scanpaths.

Published Time	Generation Method	JUFE [14]			JXUFE [39]		
		LEV↓	DTW↓	REC↑	LEV↓	DTW↓	REC↑
-	Random Baseline (lower bound)	35.21	1707.45	0.38	35.08	1695.93	0.38
TVCG22	ScanGAN360 [28]	32.53	1448.65	1.07	31.89	1427.55	1.14
CVPR23	ScanDMM [38]	31.23	**1434.36**	1.21	31.48	1438.29	1.12
-	RPS w/o DSP (Ours)	29.54	1471.82	2.14	29.99	1463.38	1.94
-	**RPS (Ours)**	**29.48**	1454.03	**2.21**	**29.66**	**1422.85**	**2.07**
-	Human Baseline (upper bound)	23.85	1309.29	3.78	26.73	1302.15	2.88

Performance on scanpath prediction task

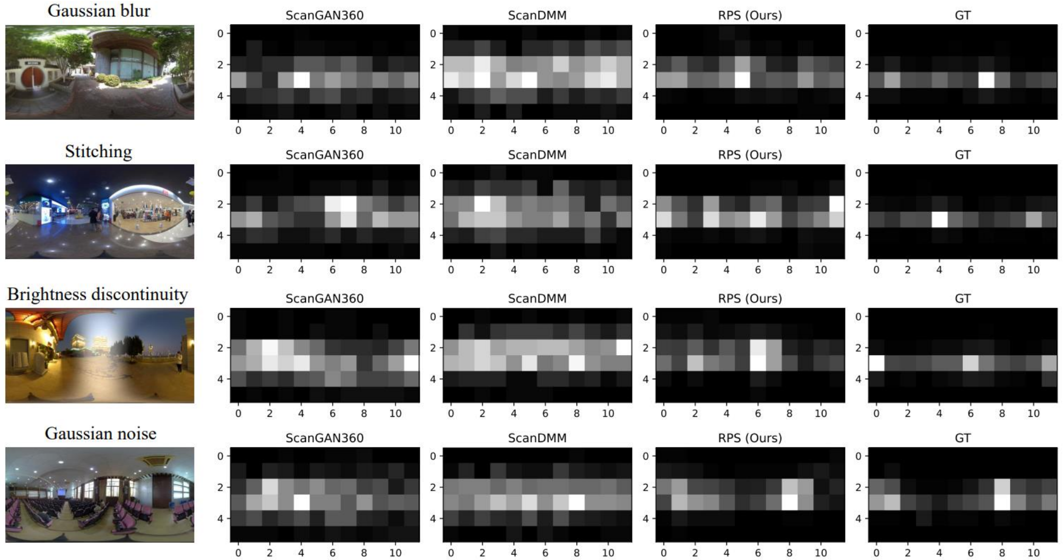


Table 7: Quantitative comparison of using original VGCN sampling method and proposed RPS on IQA-ODI [52] and MVAQD [21] datasets.

Method	IQA-ODI		MVAQD	
	SRCC	PLCC	SRCC	PLCC
VGCN	0.8117	0.8823	0.8422	0.9112
VGCN-RPS	**0.8382**	**0.8883**	**0.9122**	**0.9273**

RPS in VGCN method

Table 4: Quantitative comparison of using different viewport sequence generation methods on OIQA [12] and MVAQD [21].

Generation Method	OIQA		MVAQD	
	SRCC	PLCC	SRCC	PLCC
Random Generation	0.9461	0.9444	0.9359	0.9543
ScanGAN360	0.9705	0.9670	0.9493	0.9694
ScanDMM	0.9652	0.9634	0.9558	0.9612
RPS (Ours)	**0.9802**	**0.9747**	**0.9607**	**0.9720**

Different generation methods

清華大学 Tsinghua University

平安科技 PINGAN TECHNOLOGY

東京大学 THE UNIVERSITY OF TOKYO

UNIVERSITY OF MARYLAND

Other Ablation Studies

Table 6: Quantitative comparison of different starting point positions on MVAQD [21] dataset.

Position (latitude, longitude)	SRCC	PLCC
(0°, 0°), (0°, 0°), (0°, 0°)	**0.9607**	**0.9720**
(60°, 0°), (60°, 0°), (60°, 0°)	0.9106	0.9312
(0°, 120°), (0°, 0°), (0°, −60°)	0.9599	0.9660
(60°, 120°), (60°, 0°), (60°, −60°)	0.9174	0.9455

Impact of the initialization of the starting point

Table 8: Ablation studies of each component in proposed Assessor360 on MVAQD [21] dataset.

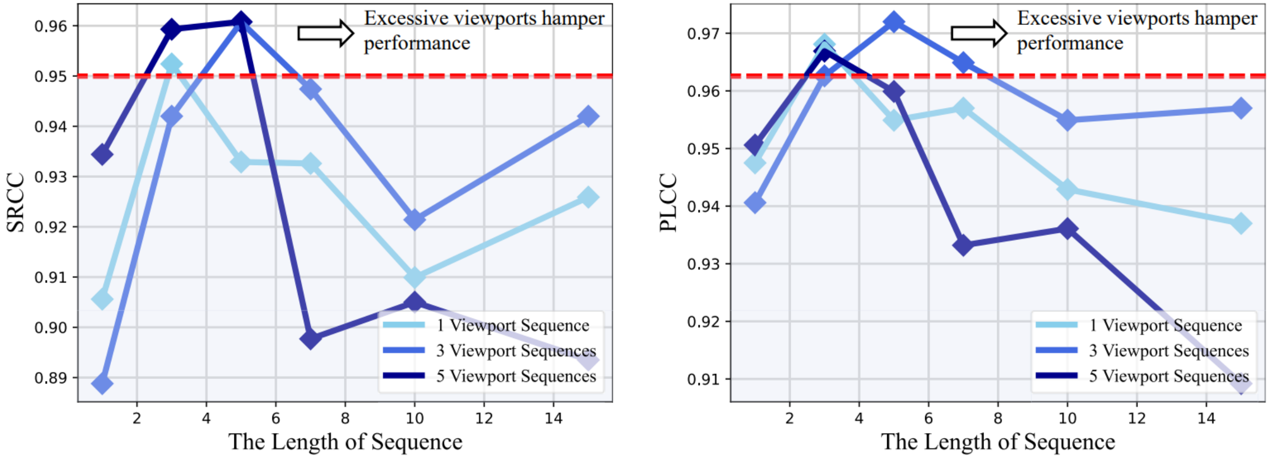
Method	Para (M)	SRCC	PLCC
Assessor360 w/o MFA	88.53	0.8514	0.9171
Assessor360 w/o DAB	88.86	0.8437	0.8779
Assessor360 w/ GAP	88.69	0.9393	0.9587
Assessor360 w/ GRU	89.28	**0.9607**	**0.9720**

Each method component

Table 9: Quantitative comparison of using GT sequences and sequences generated by RPS on JUFE [14] dataset. Starting Point (SP).

Viewport Sequence	Good SP SRCC	Good SP PLCC	Bad SP SRCC	Bad SP PLCC
RPS (Ours)	0.6623	0.6365	0.5044	0.4946
GT Sequence	**0.7158**	**0.7013**	**0.5400**	**0.5377**

Comparing with GT sequences



Impact of the number and length of the viewport sequence

Conclusion

- We propose a novel BOIQA network aligning to the realistic observation procedure.
- We propose RPS method which is competitive to the SoTA scanpath prediction methods.
- Our proposed pipeline can provide long-term valuable insights for future OIQA task.

Thanks for your listening

Tianhe Wu[1,*], Shuwei Shi[1,2*], Haoming Cai[3], Mingdeng Cao[2], Jing Xiao[4], Yinqiang Zheng[2], Yujiu Yang[1†]

[1] Shenzhen International Graduate School, Tsinghua University, [2] The University of Tokyo,
[3] University of Maryland, College Park, [4] Pingan Group

https://github.com/TianheWu/Assessor360

Speaker: Tianhe Wu
Email: wth22@mails.tsinghua.edu.cn
2023.11.10