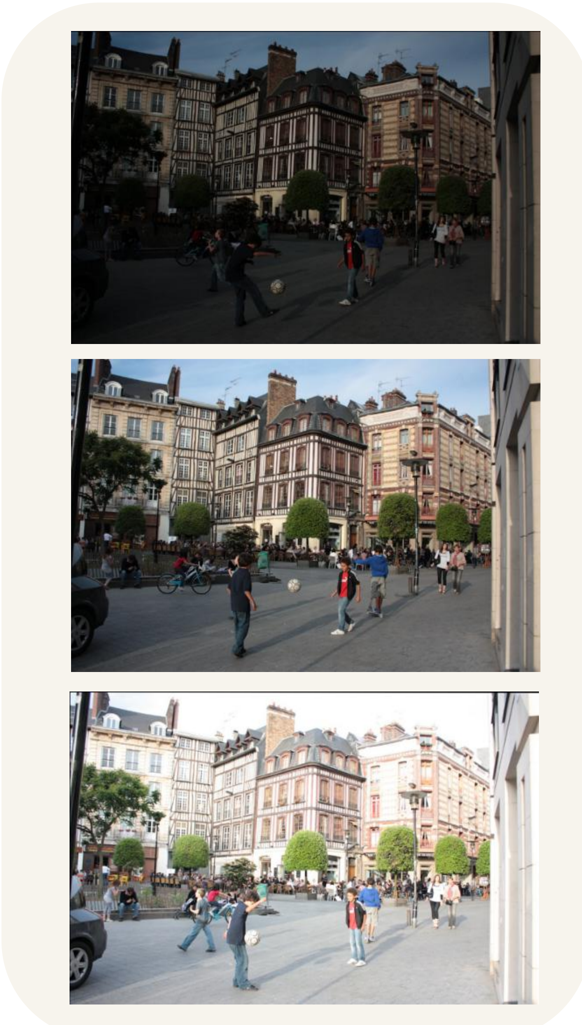
Part IV: Real-World Applications

- High-dynamic-range imaging (image fusion)
- Color-to-gray conversion
- Image retargeting
- Stereoscopic images
- Omnidirectional images
- Screen content images
- Natural videos

High Dynamic Range (HDR) imaging



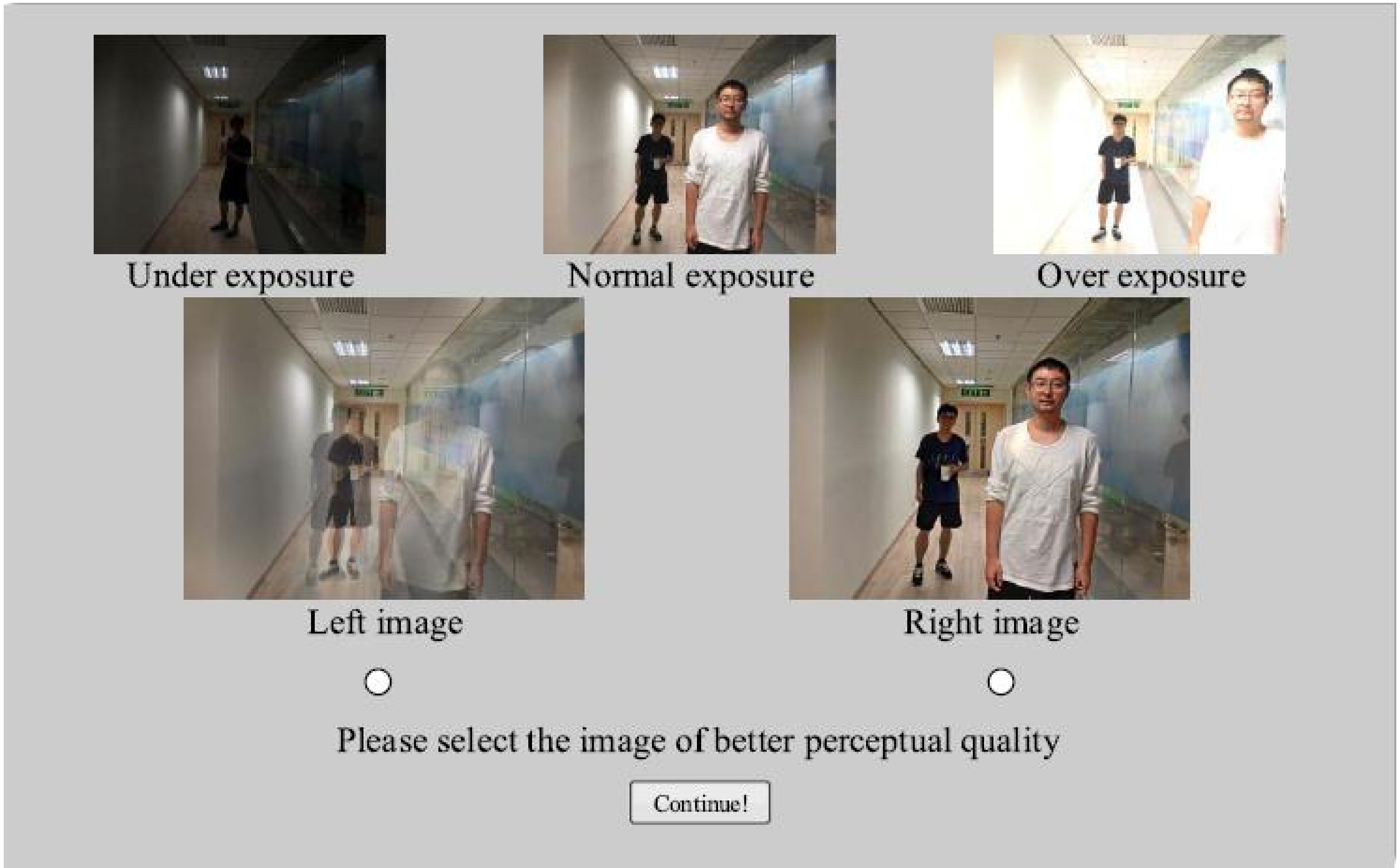
Multi-exposure image fusion

Source multi-exposure image sequence

Multi-Exposure Image Fusion (MEF) Quality Assessment Database

- The database contains 20 source sequences with multiple exposure levels (>= 3)
- Nine multi-exposure image fusion (MEF) methods are adopted to generate 180 images

Subjective Quality Assessment

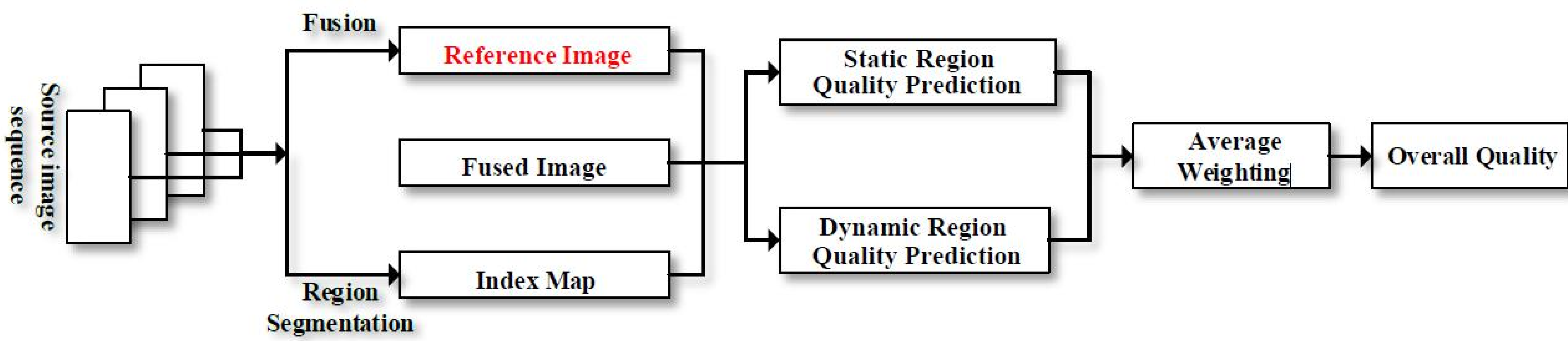


60 subjects join in the experiment and each image pair is compared exactly 20 times.

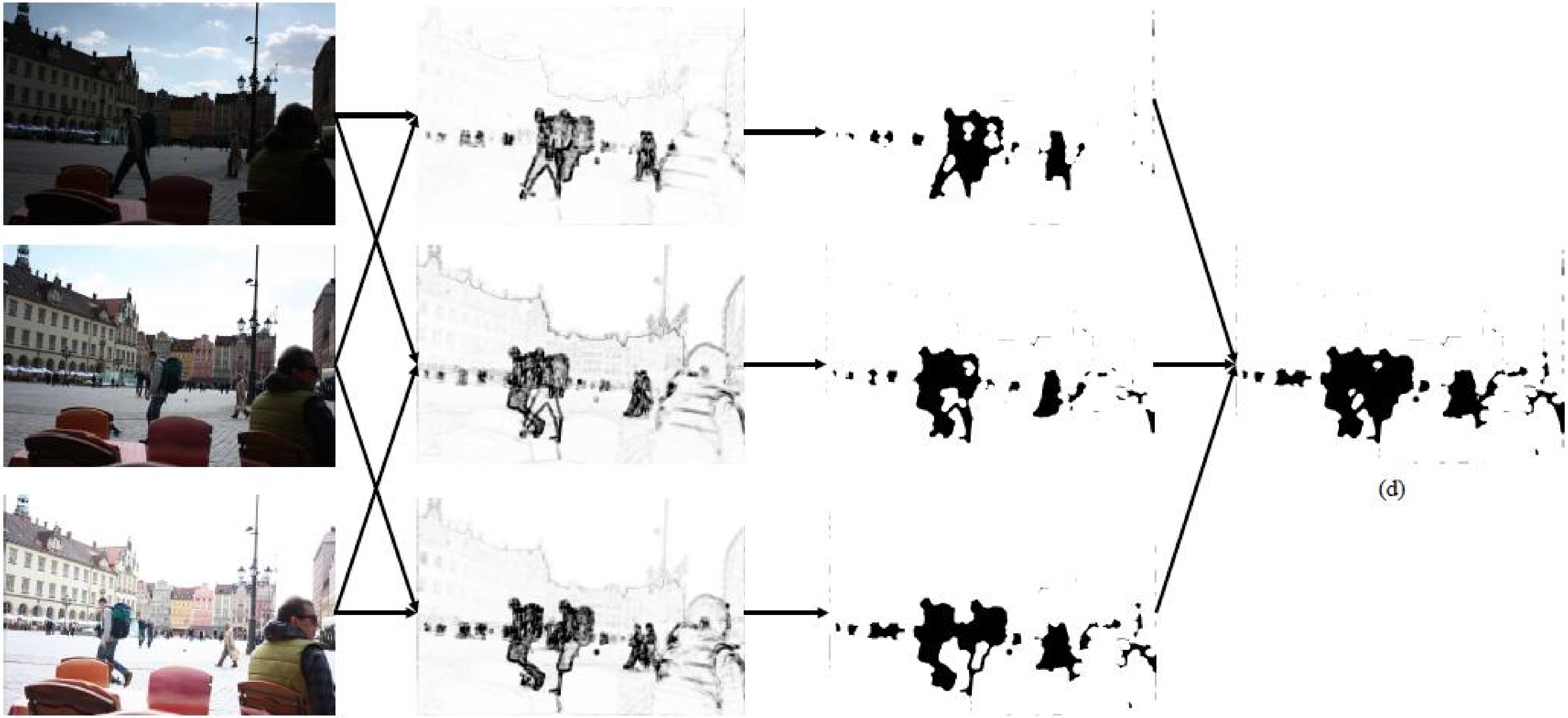
Data Analysis & Objective Quality Metric

Algorithm	Li12	Lee14	Photomatix	Qin15	Pece10	Sen12	Hu13	Li14	SPD-MEF	SUM
Li12 [19]	0	134	86	68	52	60	52	54	15	521
Lee14 [14]	266	0	137	119	141	71	99	73	53	959
Photomatix [38]	314	263	0	189	184	115	108	102	59	1334
Qin15 [21]	332	281	211	0	211	143	123	109	52	1431
Pece10 [10]	348	259	216	220	0	127	160	117	94	1541
Sen12 [2]	340	329	285	257	273	0	157	177	91	1909
Hu13 [20]	348	301	292	277	240	243	0	159	111	1971
Li14 [23]	346	327	298	291	283	223	241	0	111	2120
SPD-MEF [11]	385	347	341	348	306	309	289	289	0	2614

Workflow of the Proposed model:

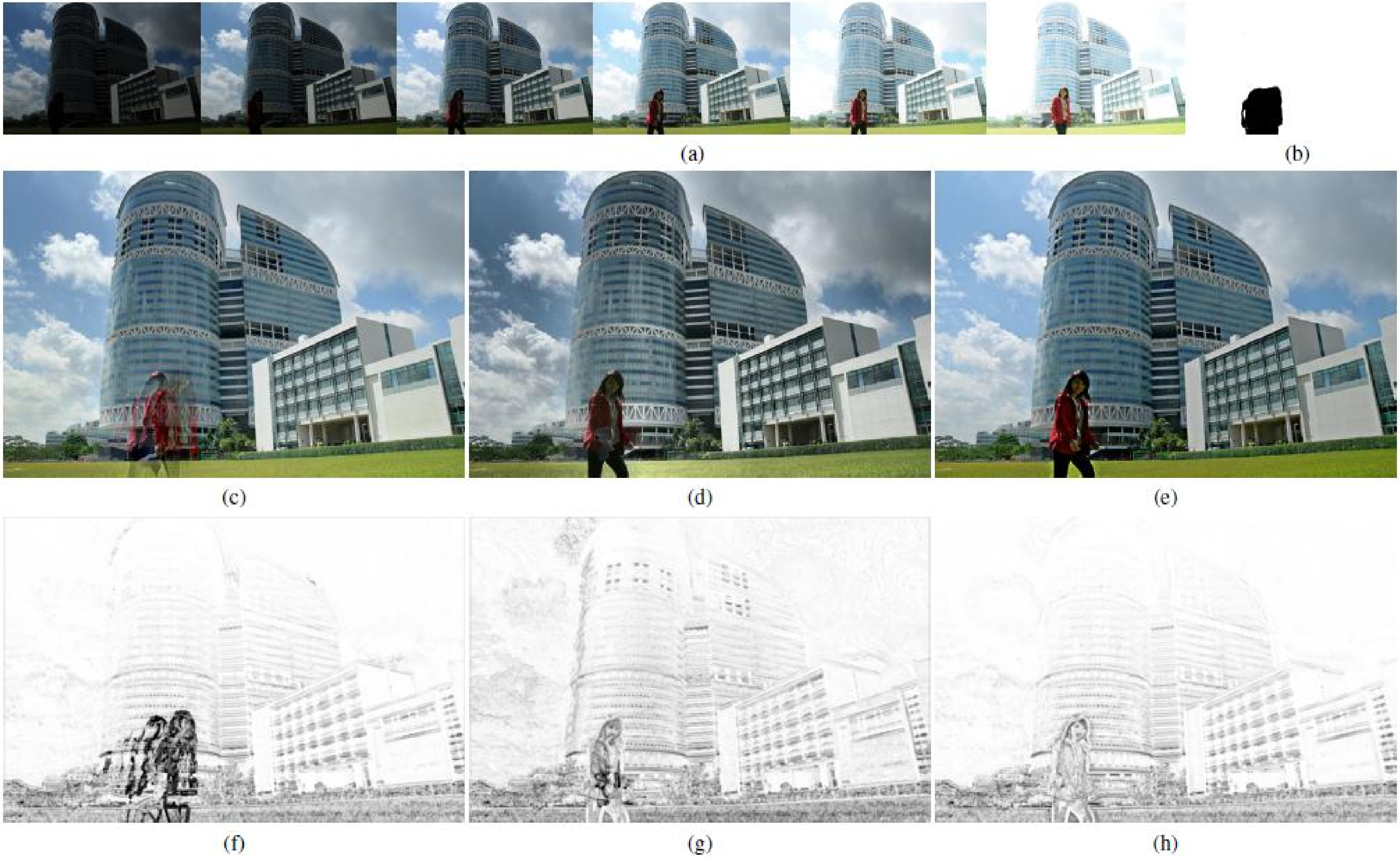


Region Segmentation



(a) (b) (c) (d)

Computing the structure consistency across exposures using patch decomposition strategy



(a) Input multi-exposure sequence. (b) Binary map for region segmentation. (c) Fused image by Pece10 [10]. (d) Fused image by Lee14 [14]. (e) Fused image by SPD-MEF [11]. (f) Quality map of (c) with $q^s = 0.937$, $q^d = 0.558$, and q = 0.748. (g) Quality map of (d) with $q^s = 0.908$, $q^d = 0.768$, and q = 0.838. (h) Quality map of (e) with $q^s = 0.939$, $q^d = 0.829$, and q = 0.884. Higher brightness in the quality map indicates better quality.

Experimental Results

Sequence	Zheng07 [31]	Cvejic06 [27]	Chen07 [34]	Piella03 [33]	MEF-SSIM [6]	Xydeas00 [29]	Wang08 [30]	Hossny08 [28]	Wang04 [44]	MEF-SSIM$_d$
Men	-0.250	-0.066	-0.133	0.166	0.500	0.483	0.600	0.666	0.583	0.866
Arch	-0.666	-0.283	-0.350	0.533	0.766	0.816	0.766	0.400	0.400	0.533
Llandudno	-0.350	-0.166	-0.366	0.533	0.566	0.183	0.400	0.333	0.400	0.800
Square	-0.466	-0.016	0.316	0.033	-0.050	-0.066	-0.033	0.466	0.466	0.933
Tate3	-0.433	0.133	0.133	-0.016	-0.033	0.016	0.216	0.850	0.766	0.667
Forest	-0.283	-0.633	-0.416	0.566	0.233	0.733	0.600	0.666	0.666	0.783
Horse	-0.366	0.166	0.166	-0.133	-0.300	-0.100	0.000	0.650	0.650	0.667
Corridor	-0.250	0.300	-0.066	0.533	0.450	0.333	0.400	0.750	0.750	0.700
Office	-0.550	-0.133	-0.433	0.283	0.433	-0.066	-0.183	0.583	0.600	0.350
Russ1	-0.400	-0.583	-0.266	0.216	0.366	0.466	0.616	0.866	0.866	0.833
Puppets	-0.833	-0.150	0.433	0.250	0.066	0.066	0.250	0.616	0.616	0.783
Cliff	-0.400	0.333	-0.716	0.283	0.616	0.266	0.233	0.583	0.600	0.466
Sculpture	-0.300	-0.583	-0.233	0.300	0.150	0.200	0.016	0.716	0.550	0.683
Wroclav	-0.150	-0.816	0.000	-0.150	-0.250	-0.116	-0.133	0.533	0.650	0.383
ProfJeon	-0.150	-0.416	-0.216	0.333	0.016	0.350	0.433	0.750	0.800	0.867
NoiseCam	-0.283	-0.033	-0.400	0.450	0.416	0.766	0.683	0.516	0.750	0.767
Campus	-0.133	0.216	-0.333	-0.266	-0.150	0.060	0.300	0.216	0.216	0.933
Brunswick	-0.616	-0.300	-0.033	-0.100	0.100	0.133	0.366	0.500	0.533	0.883
YWFusion	-0.066	0.216	0.083	-0.233	-0.116	-0.100	0.083	0.617	0.616	0.917
Lady	0.200	-0.516	-0.316	0.133	0.250	0.066	0.033	0.883	0.883	0.817
Average	**-0.338**	**-0.167**	**-0.158**	**0.186**	**0.202**	**0.225**	**0.283**	**0.608**	**0.618**	**0.730**

Summary

- We created an MEF database and conducted a subjective experiment to collect human opinions of fused image quality.
- we design a objective quality model, which successfully captures the ghosting artifacts, resulting in the best quality prediction performance.

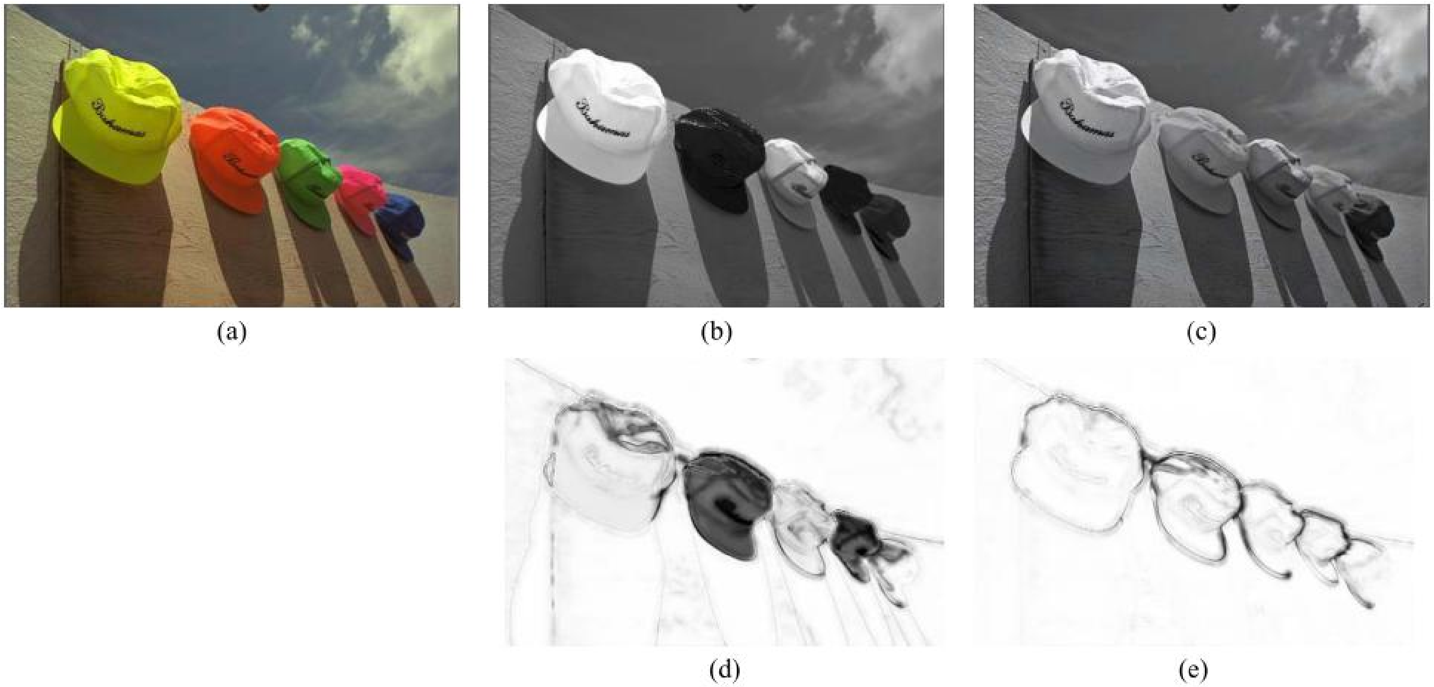
Database & Models:

https://github.com/h4nwei/MEF-SSIMd

Yuming Fang, H. Zhu, **Kede Ma**, *et al.*, Perceptual evaluation for multi-exposure image fusion of dynamic scene, ***IEEE T-IP***, 2020.

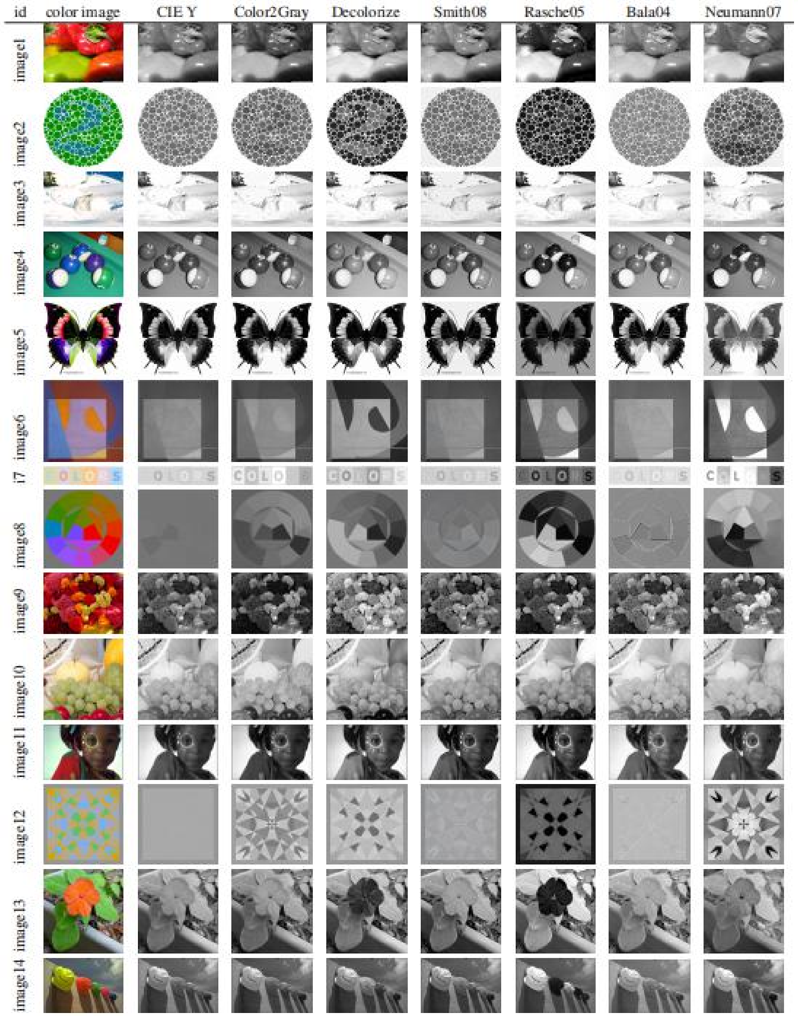
Yuming Fang, *et al.*, Superpixel-based quality assessment of multi-exposure image fusion for both static and dynamic scenes, ***IEEE T-IP***, 2021.

Color-to-gray (C2G) Conversion



(a) Reference color image. (b), (c) , (d) , (e) C2G images

The Test Database

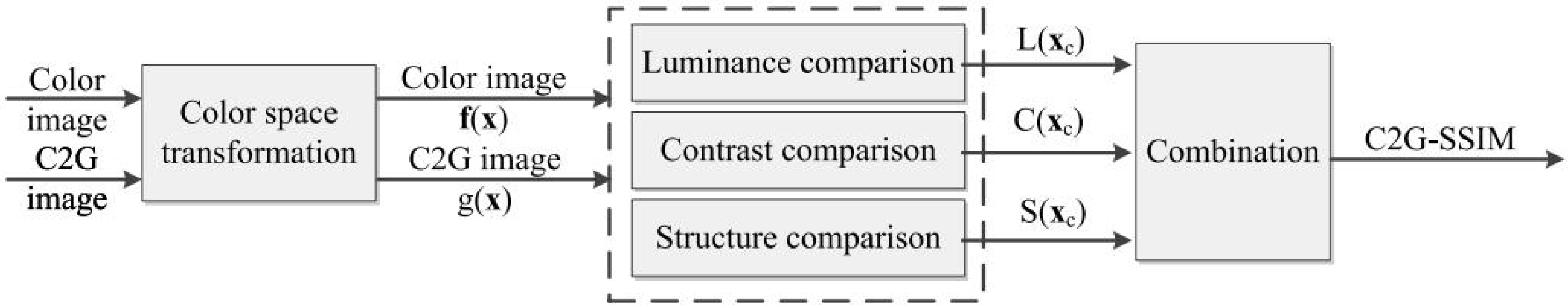


- Reference images: 24.
- Algorithms:7.
- Total images: 24×7=169.

- For *Accuracy* labels: the grayscale images are presented along with the original color image.
- For *Preference* labels: two grayscale images without any reference are rated by subjects.

P. Cadik, Perceptual evaluation of color-to-grayscale image conversions, *CGF*, 2008.

The proposed C2G-SSIM



(a) Framework of the proposed C2G-SSIM

$$q(\mathbf{X}_c) = L(\mathbf{X}_c)^{\alpha} \cdot C(\mathbf{X}_c)^{\beta} \cdot S(\mathbf{X}_c)^{\gamma}$$

where $\alpha > 0,\ \beta > 0, \gamma > 0.$

Kede Ma, *et al*., Objective quality assessment for color-to-gray image conversion, *IEEE T-IP*, 2015.

Results

Performance comparison of C2G-SSIM with existing metrics for *Accuracy* test

Image set		SRCC				KRCC			
		RWMS	E-score	Subject	C2G-SSIM	RWMS	E-score	Subject	C2G-SSIM
PI	Image1	-0.1071	-0.5357	0.5697	0.7143	0.0476	-0.2381	0.4512	0.6190
	Image3	0.4286	0.3214	0.5204	0.6071	0.3333	0.1429	0.4150	0.4286
	Image4	-0.2500	0.0000	0.6667	0.8214	-0.1429	-0.0476	0.5601	0.7143
	Image9	0.3929	0.3929	0.4323	0.7500	0.2381	0.3333	0.3484	0.6190
	Image10	0.5000	0.5357	0.4906	0.6429	0.3333	0.4286	0.3784	0.4286
	Image11	0.5000	0.5714	0.4787	0.9643	0.3333	0.4286	0.3868	0.9048
	Image13	-0.5357	-0.2857	0.6015	0.4286	-0.4286	-0.2381	0.4987	0.4286
	Image14	-0.1071	-0.3571	0.6147	0.9643	-0.1429	-0.3333	0.5288	0.9048
	Image15	0.6071	0.6429	0.5376	0.9643	0.5238	0.4286	0.4286	0.9048
	Image16	0.4286	0.7500	0.5969	0.8571	0.3333	0.6190	0.4739	0.7143
	Image19	0.1071	0.5357	0.6429	0.9286	0.1429	0.2381	0.5038	0.8095
	Image22	0.6071	0.3929	0.7538	0.5714	0.5238	0.2381	0.6642	0.3333
	Image23	0.2500	0.0714	0.7194	0.9286	0.1429	0.0476	0.6100	0.8095
	Image24	0.6071	0.5357	0.6523	0.8214	0.5238	0.4286	0.5188	0.6190
PI Average		**0.2449**	**0.2551**	**0.5912**	**0.7832**	**0.1973**	**0.1769**	**0.4833**	**0.6599**
SI	Image2	0.0357	0.8571	0.8853	0.5714	0.0476	0.7143	0.8045	0.4286
	Image5	0.2143	0.8214	0.8010	0.8929	0.1429	0.6190	0.6689	0.8095
	Image6	0.5714	0.9643	0.7801	0.9286	0.4286	0.9048	0.6541	0.8095
	Image7	0.6429	0.0714	0.5752	0.7500	0.4286	0.0476	0.4586	0.5238
	Image8	0.3571	0.8214	0.8402	0.8571	0.3333	0.6190	0.7043	0.7143
	Image12	0.2143	0.7143	0.8327	0.8571	0.1429	0.6190	0.7193	0.7143
	Image17	0.2857	0.2143	0.6616	0.3929	0.1429	0.0476	0.5465	0.3333
	Image18	0.1071	0.1786	0.5697	0.9286	0.1429	0.0476	0.4286	0.8095
	Image20	0.5357	0.6071	0.8233	0.7500	0.4286	0.5238	0.7043	0.6190
	Image21	-0.2143	0.6786	0.7379	0.8214	-0.1429	0.5238	0.6217	0.7143
SI Average		**0.2750**	**0.5929**	**0.7507**	**0.7750**	**0.2095**	**0.4667**	**0.6311**	**0.6476**
Overall		**0.2574**	**0.3958**	**0.6577**	**0.7798**	**0.2024**	**0.2976**	**0.5449**	**0.6548**

Kede Ma, *et al*., Objective quality assessment for color-to-gray image conversion, *IEEE T-IP*, 2015.

Results

Performance comparison of C2G-SSIM with existing metrics for *Preference* test

Image set		SRCC				KRCC			
		RWMS	E-score	Subject	C2G-SSIM	RWMS	E-score	Subject	C2G-SSIM
PI	Image1	0.4286	-0.1071	0.6143	0.6786	0.3333	0.0476	0.5143	0.5238
	Image3	0.0714	0.1429	0.4982	0.7143	0.0476	0.0476	0.3810	0.5238
	Image4	-0.2500	0.0000	0.8750	0.8214	-0.1429	-0.0476	0.7714	0.7143
	Image9	0.4643	0.5000	0.5771	0.7500	0.4286	0.3333	0.4837	0.6190
	Image10	0.4286	0.4643	0.7870	0.7500	0.2381	0.3333	0.6824	0.5238
	Image11	0.2857	0.2143	0.5977	0.7500	0.2381	0.1429	0.4586	0.6190
	Image13	0.1786	0.3214	0.4388	0.6071	0.2381	0.2381	0.3333	0.3333
	Image14	-0.0357	0.1429	0.5017	0.6071	0.0476	0.2381	0.4014	0.5238
	Image15	0.5000	0.5357	0.5561	0.8571	0.3333	0.4286	0.4603	0.7143
	Image16	0.1429	0.6786	0.6933	0.5357	0.0476	0.5238	0.5605	0.4283
	Image19	0.2143	0.6429	0.7619	1.0000	0.1429	0.4286	0.6281	1.0000
	Image22	0.5357	0.3929	0.7519	0.7143	0.3333	0.2381	0.6441	0.5238
	Image23	0.4643	0.1429	0.7179	1.0000	0.3333	0.0476	0.6286	1.0000
	Image24	0.8571	0.7500	0.5969	0.4643	0.7143	0.6190	0.4649	0.4286
PI Average		**0.3061**	**0.3444**	**0.6406**	**0.7321**	**0.2381**	**0.2585**	**0.5295**	**0.6054**
SI	Image2	0.2143	0.9286	0.9492	0.6429	0.1429	0.8095	0.8947	0.5238
	Image5	0.4286	0.6071	0.8321	0.8214	0.3333	0.4286	0.7048	0.6190
	Image6	0.3571	0.8929	0.8553	0.8571	0.3333	0.8095	0.7243	0.7143
	Image7	0.6786	0.2143	0.7279	0.8214	0.5238	0.1429	0.5964	0.6190
	Image8	0.2857	0.7500	0.8797	0.8214	0.2381	0.5238	0.7744	0.6190
	Image12	0.5000	0.5714	0.8384	0.8214	0.3333	0.4286	0.6916	0.7143
	Image17	-0.0357	0.0714	0.7161	0.2143	-0.0476	0.0476	0.6000	0.1429
	Image18	0.4643	0.2500	0.5018	0.8571	0.2381	0.1429	0.3857	0.7143
	Image20	0.4643	0.5714	0.8095	0.7143	0.3333	0.4286	0.6780	0.5238
	Image21	-0.0714	0.7857	0.8008	0.9286	-0.0476	0.6190	0.6942	0.8095
SI Average		**0.3286**	**0.5643**	**0.7911**	**0.7500**	**0.2381**	**0.4381**	**0.6744**	**0.6000**
Overall		**0.3155**	**0.4360**	**0.7033**	**0.7396**	**0.2381**	**0.3333**	**0.5898**	**0.6032**

Kede Ma, *et al*., Objective quality assessment for color-to-gray image conversion, *IEEE T-IP*, 2015.

Image Retargeting

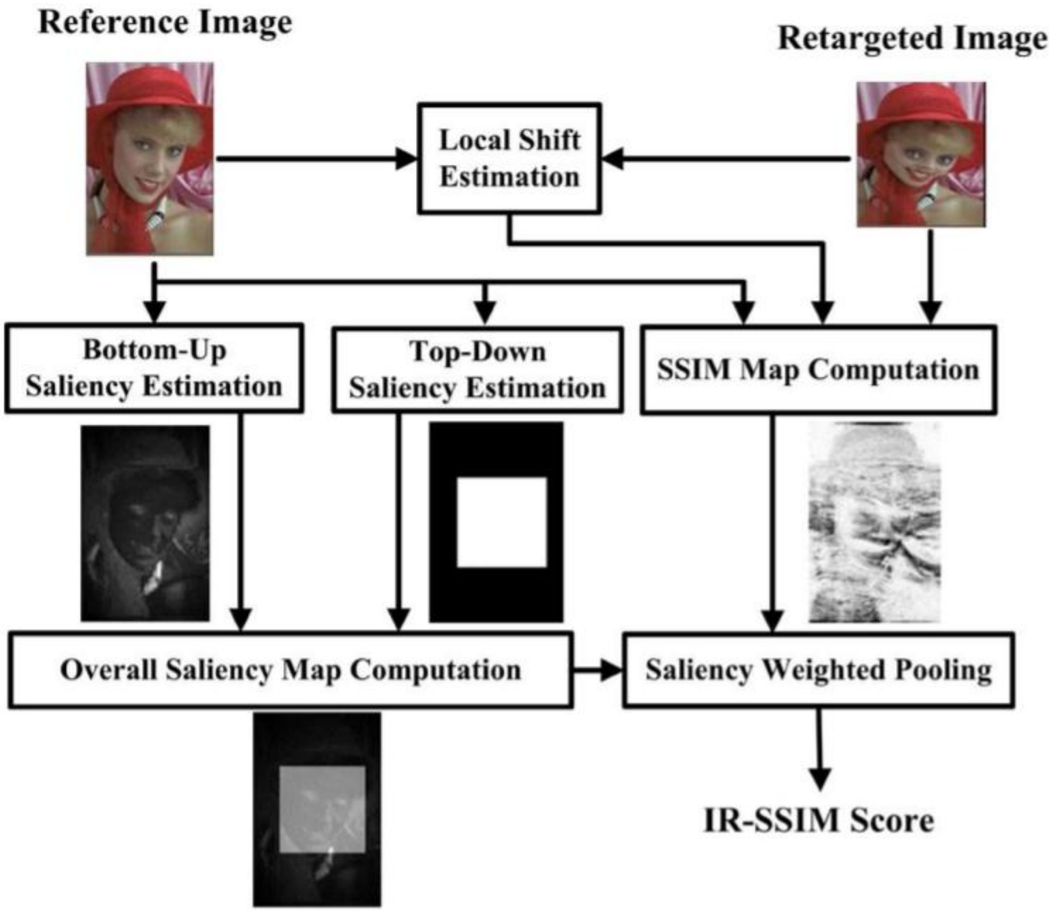
- The pixel correspondence is lost

retargeting

(a) source image (b) resized image

Yuming Fang, *et al*., Saliency detection in the compressed domain for adaptive image retargeting, *IEEE T-IP*, 2012.
Yuming Fang, *et al*., Optimized multioperator image retargeting based on perceptual similarity measure, IEEE *T-SMCS, 2016*

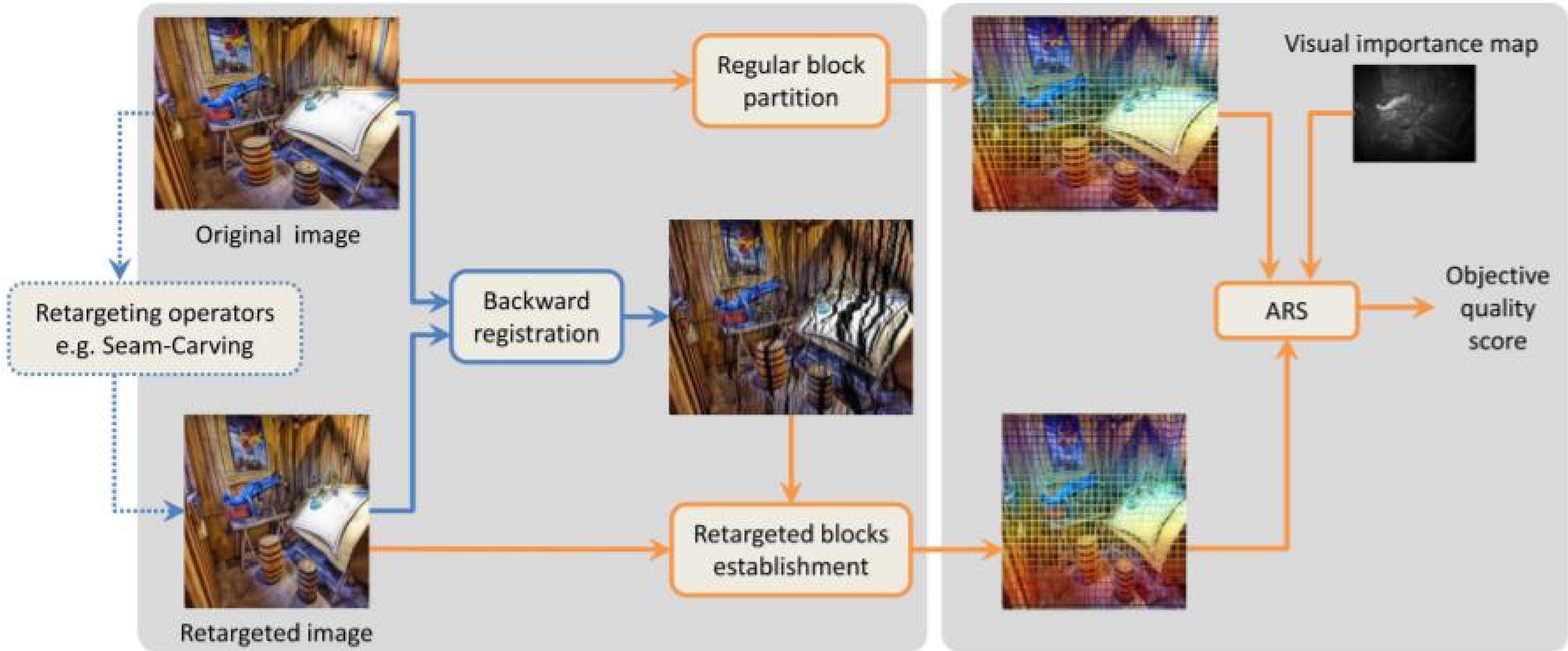
Image Retargeting



(a) Framework

Yuming Fang, *et al*., Objective quality assessment for image retargeting based on structural similarity, *IEEE JESTCS*, 2014.

Image Retargeting

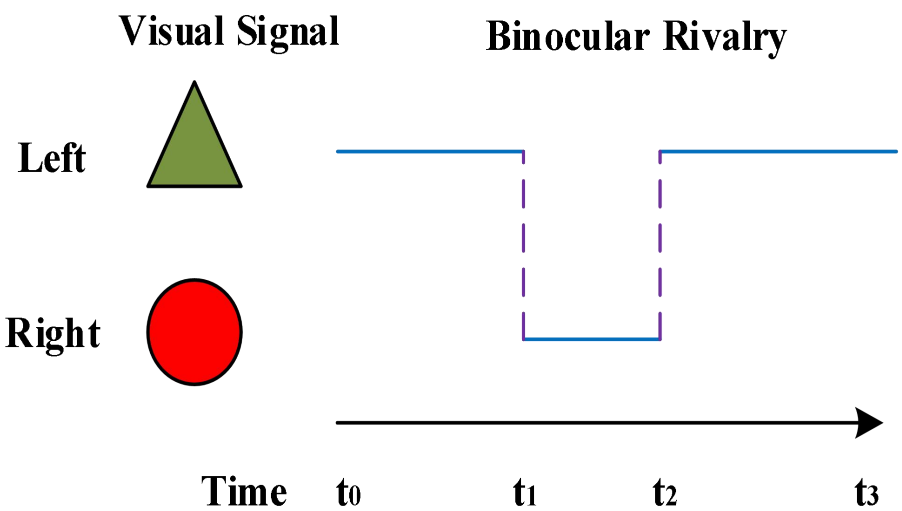


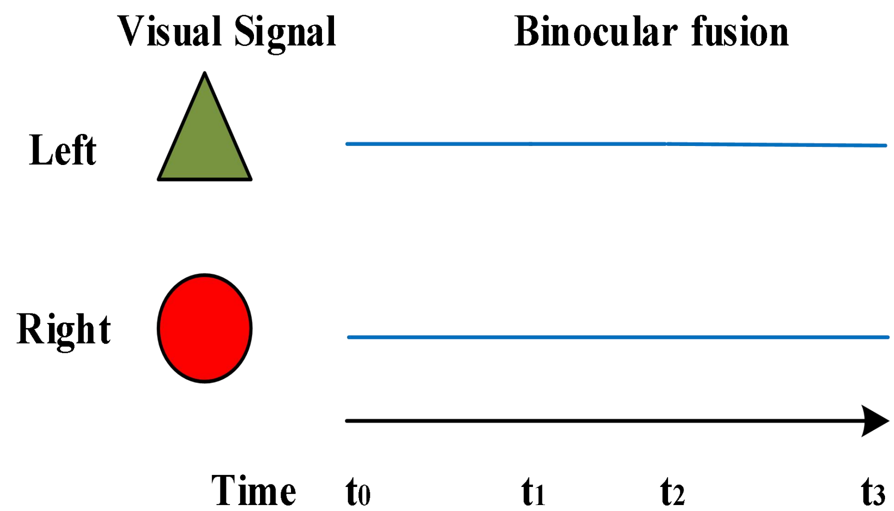
(a) Framework

Y. Zhang, **Yuming Fang**, *et al*., Backward registration-based aspect ratio similarity for image retargeting quality assessment, *IEEE T-IP*, 2016.

Stereoscopic Images

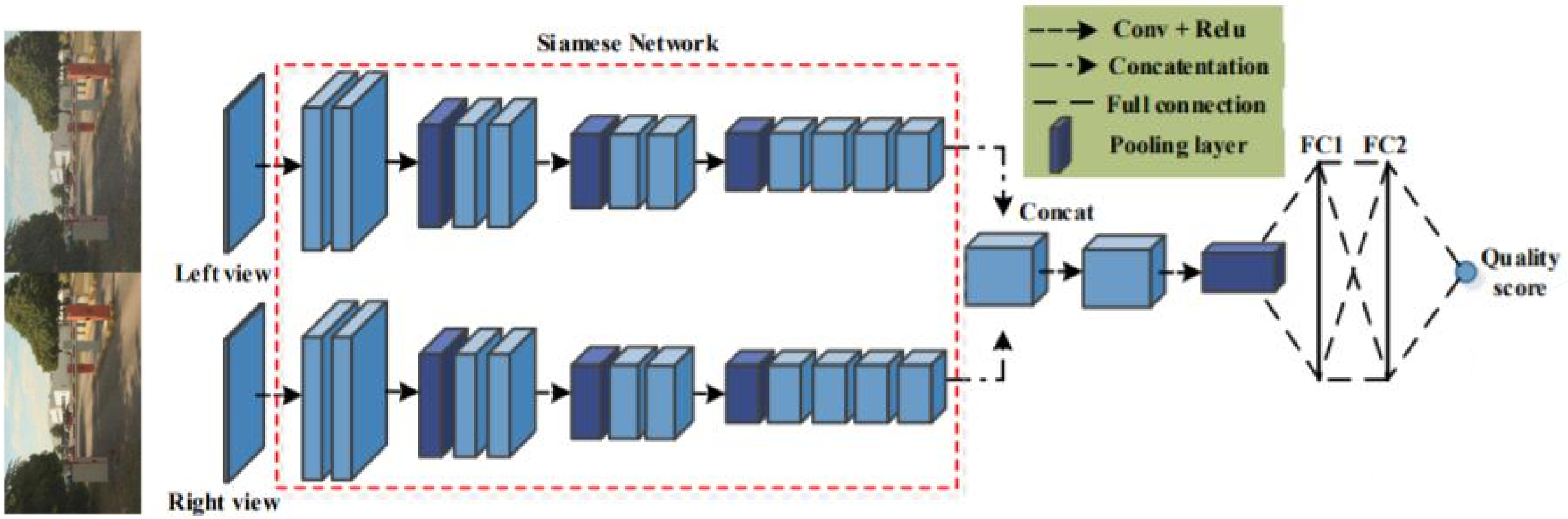
- Consideration of two characteristics of the human visual system
  - Binocular mechanism





M. J. Chen, *et. al.*, Full-reference quality assessment of stereopairs accounting for rivalry, *SPIC*, 2013.
Yuming Fang, *et al.*, Stereoscopic image quality assessment by deep convolutional neural network, *JVCIR*, 2019.

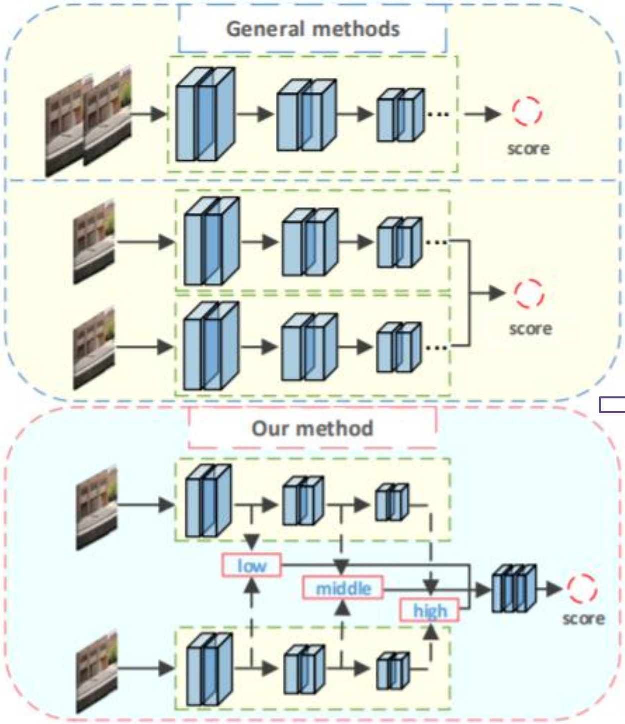
Stereoscopic Images



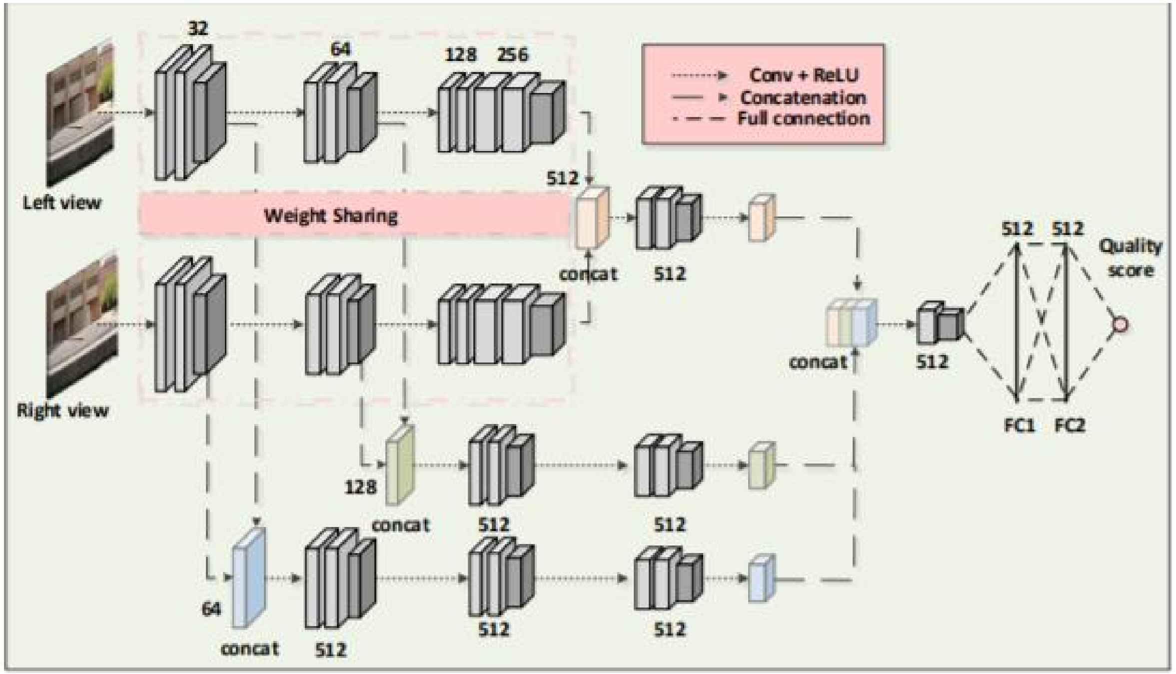
(a) Framework

Yuming Fang, *et al.*, Stereoscopic image quality assessment by deep convolutional neural network, *JVCIR*, 2019.

Stereoscopic Images



(a) Motivation

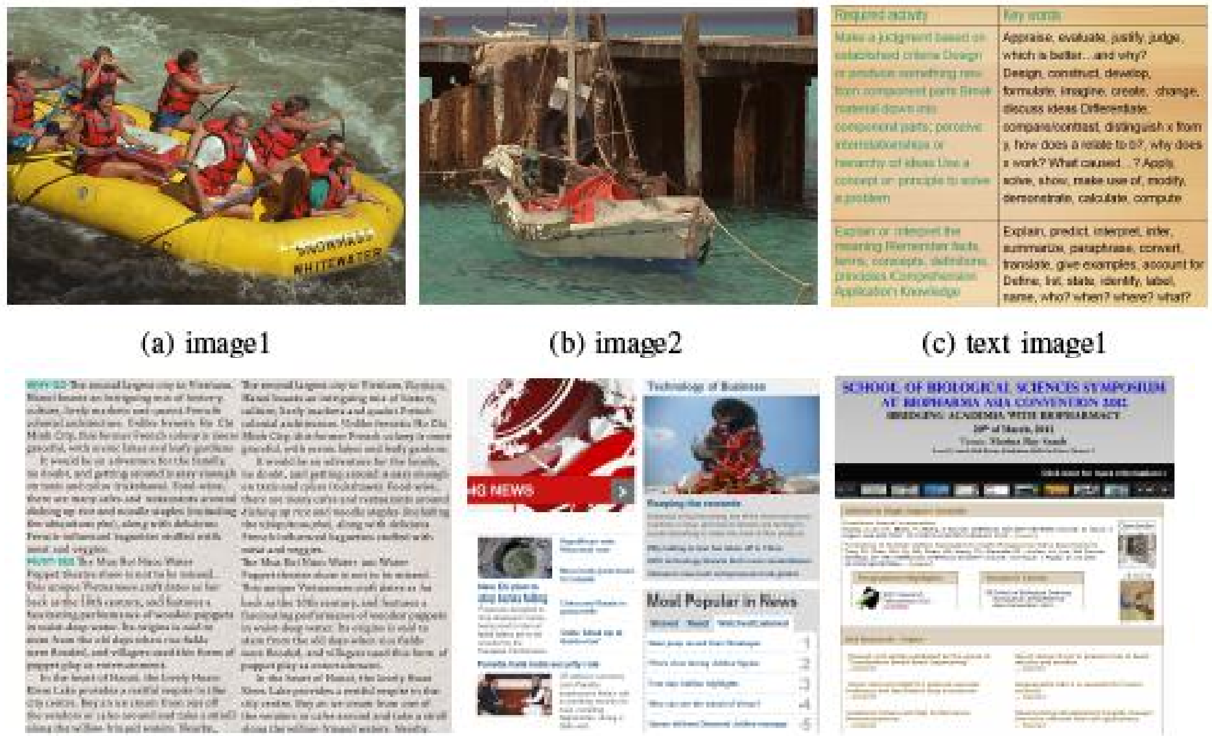


(b) Framework

J. Yan, **Yuming Fang**, *et al.*, Blind stereoscopic image quality assessment by deep convolutional neural network of multi-level feature fusion, in *ICME*, 2020.

VQA for Screen Content Images

- Screen content images
  - Non-natural image statistics should be extracted

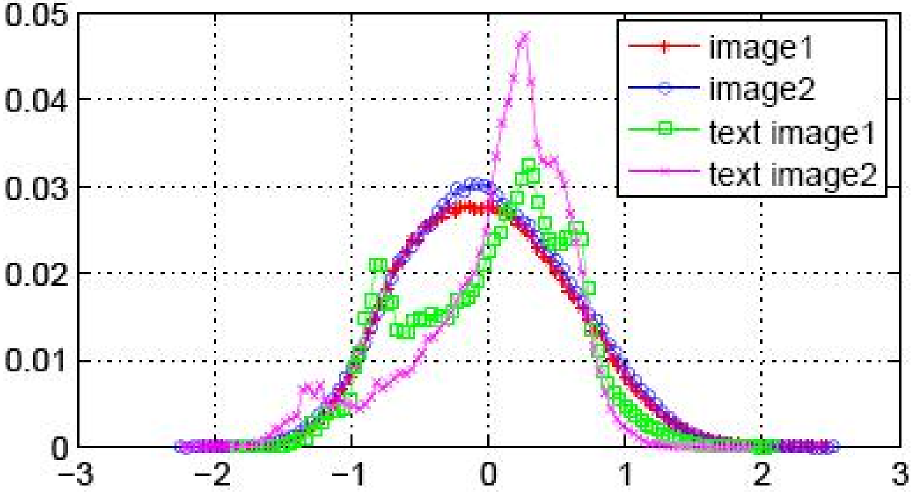


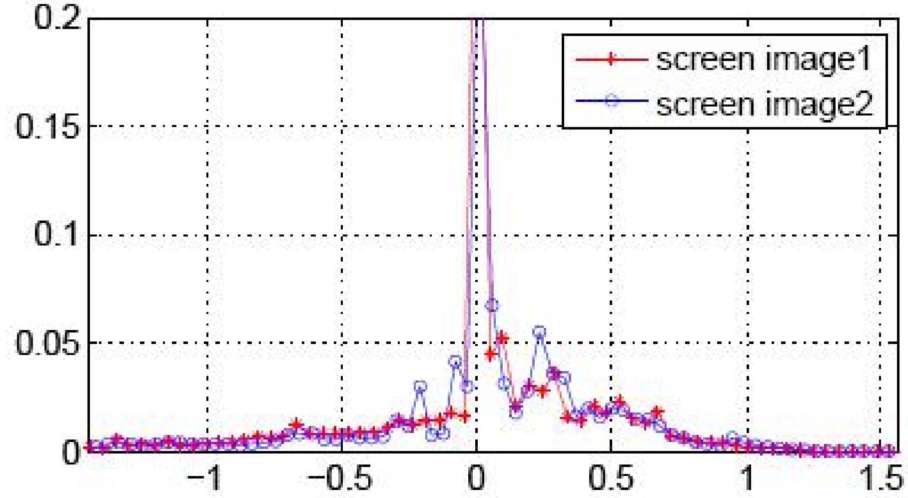
(a) image1 (b) image2 (c) text image1

(d) text image2 (e) screen image1 (f) screen image2

Naturalness

$$I'(i,j) = \frac{I(i,j) - \mu(i,j)}{\sigma(i,j) + 1}$$





Subjective Evaluation of SCIs

- Screen Image Quality Assessment Database (SIQAD)
  - Reference SCIs
    - 20 SCIs
    - Various layout style, including different text sizes, positions, and ways of textual/pictorial combination
    - Diverse content
  - Distorted SCIs (980)
    - Seven distortion types: Gaussian noise, Gaussian blur, motion blur, contrast change, JPEG, JPEG2000, layer segmentation based coding
    - Seven degradation levels (from slight to high annoying)
  - Display setting
    - ITU-R BT.500-13. viewing conditions in laboratory environment
  - Human subjects
    - 96 subjects involved in this user study

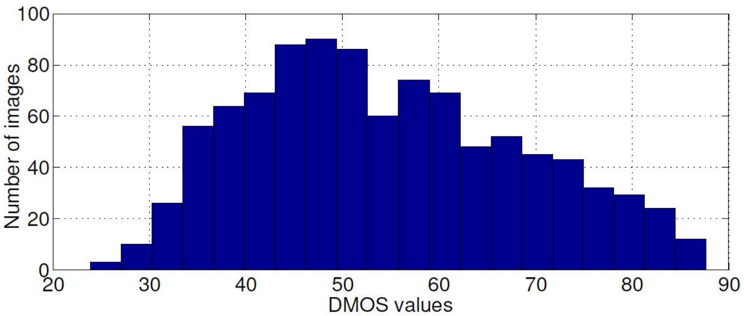
Subjective Evaluation of SCIs

- Subjective testing methodology
  - Absolute category rating (ACR)
  - Three subjective scores (content recognizability, content clarity and viewing comfort)
    - Quality of entire region (QE)
    - Quality of textual region (QT)
    - Quality of pictorial region (QP)
- Data organization
  - Each session <= 30 mins
  - Each image has at least 30 scores

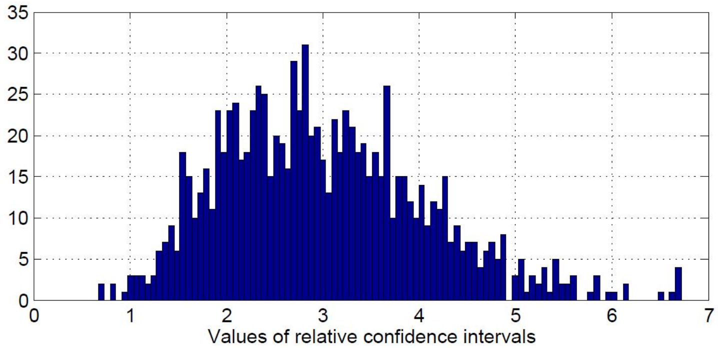
Reliability of DMOS Values

DMOS: Difference Mean Opinion Score

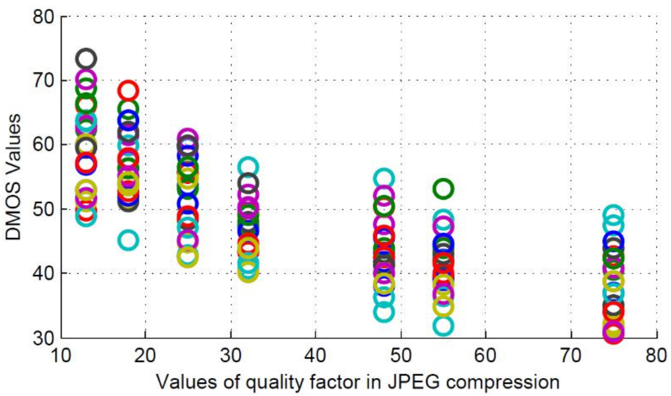
- Outlier detection and rejection
- Scale realignment



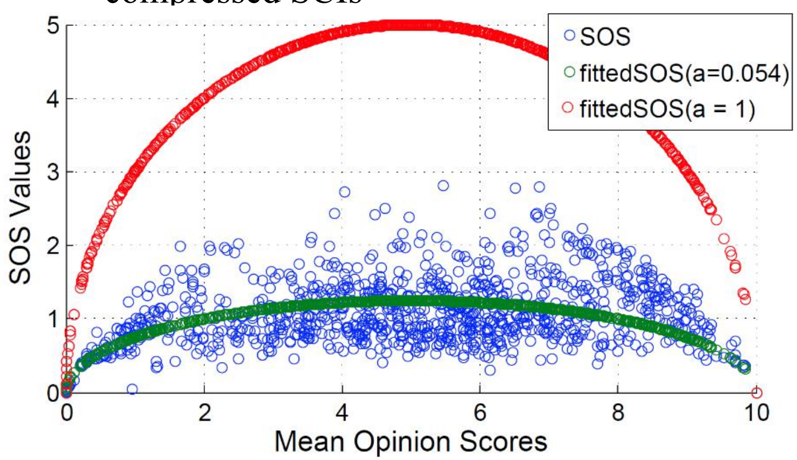
Histogram of overall DMOS values



Distribution of relative confidence intervals of overall DMOS values. The quality scale for all images is (0,100). Smaller Values indicate higher reliability.



Distribution of DMOS values of JPEG compressed SCIs



Standard deviation of Opinion Scores (SOS) hypothesis for the subjective scores. Higher value of α indicates larger diversity of subject's judgment.

Observations from the Subjective Testing

- Correlations of different kinds of DMOS values

Correlation between QE and QT, QE and QP

	QE and QT			QE and QP		
Distortions	PLCC	SROCC	RMSE	PLCC	SROCC	RMSE
GN	0.9424	0.9367	4.9915	0.8958	0.8819	6.6295
GB	0.9268	0.9234	5.7006	0.8889	0.8916	6.9530
MB	0.9042	0.9057	5.5528	0.8513	0.8526	6.8218
CC	0.8332	0.7580	6.9558	0.8405	0.8030	6.8150
JPEG	0.8548	0.8488	4.8765	0.7493	0.7162	6.2226
JPEG2000	0.8474	0.8521	5.5185	0.8058	0.7821	6.1554
LSC	0.7701	0.7755	5.4432	0.6914	0.6923	6.1647
Overall	0.9040	0.8958	6.1204	0.8389	0.8336	7.7899

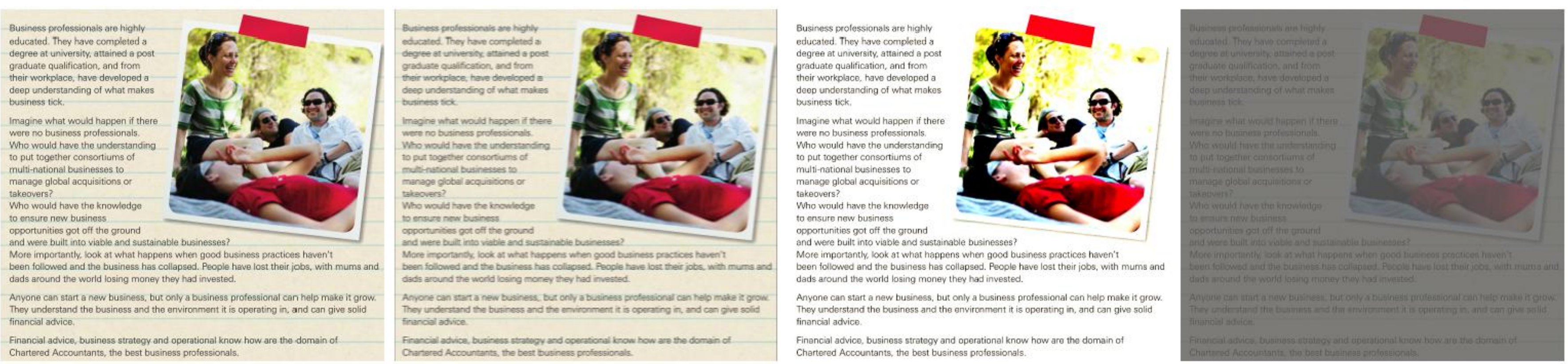
QE: Quality of Entire region

QT: Quality of Textual region

QP: Quality of Pictorial region

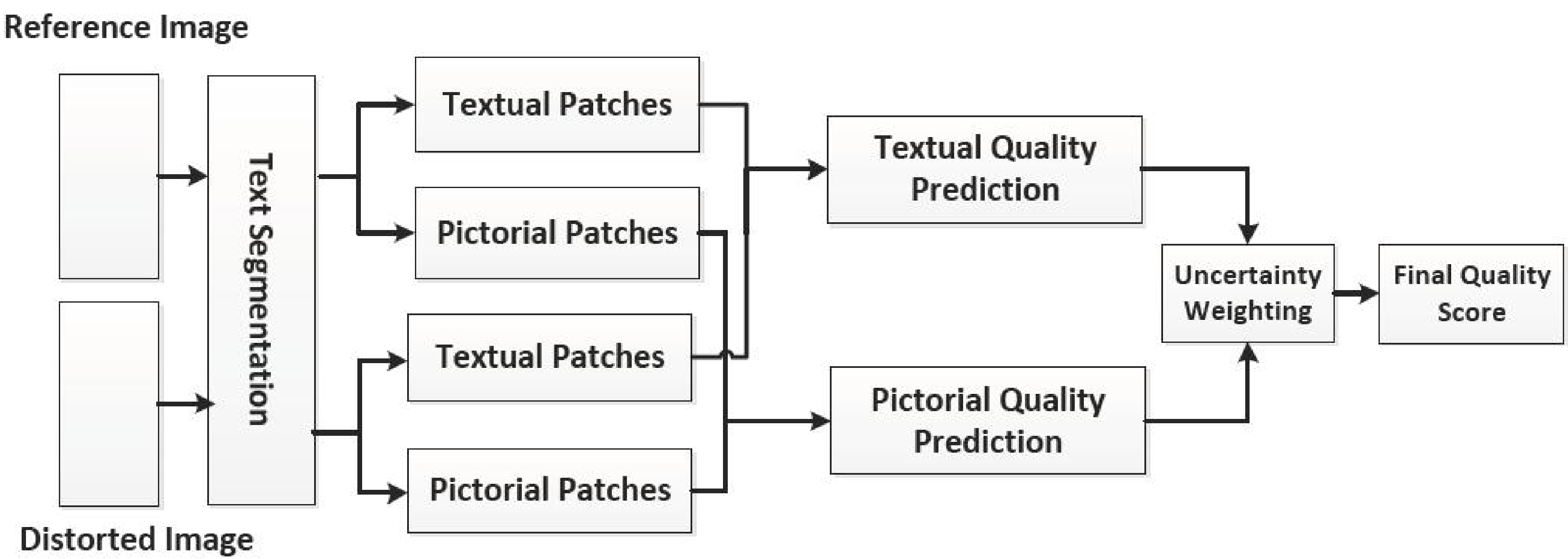
PLCC: Pearson Linear Correlation Coefficient
RMSE: Root Mean Squared Error
SROCC: Spearman rank-order correlation coefficient

- Different visual perception to textual and pictorial regions



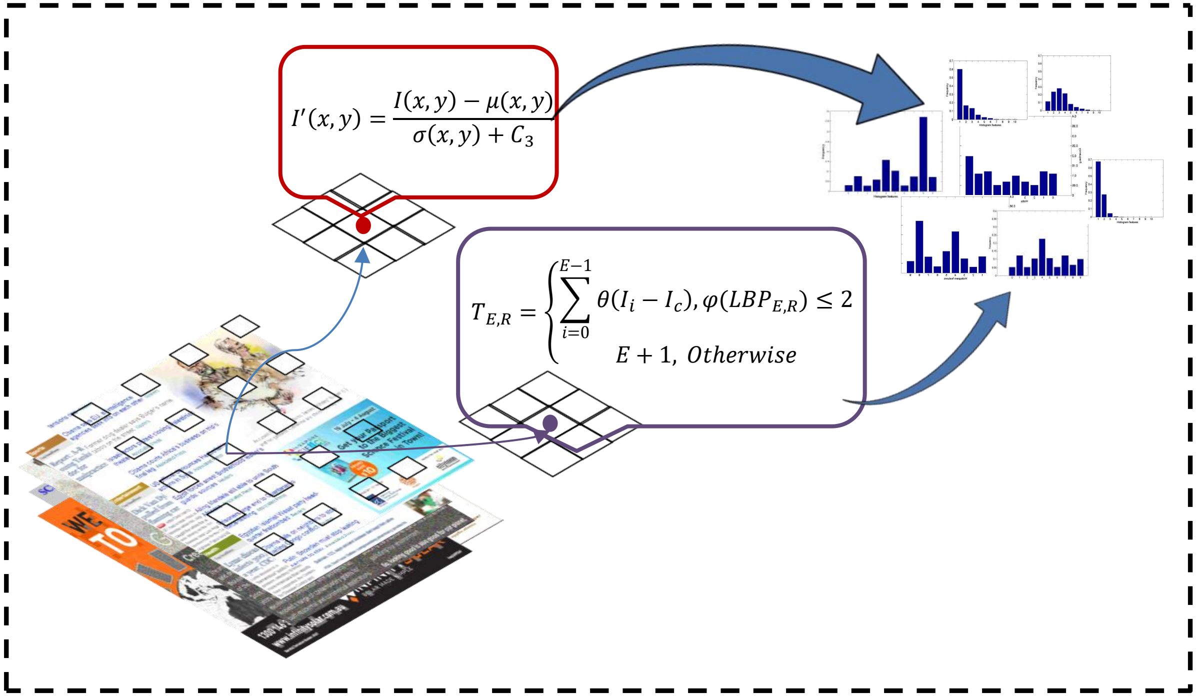
(a) Reference image: cim11 (b) cim11_3_5, DMOS:63.98 (c) cim11_4_7, DMOS:37.50 (d) cim11_4_1, DMOS:76.54

FR VQA for SCIs



Yuming Fang, *et al.*, Objective quality assessment of screen content images by uncertainty weighting, ***IEEE T-IP***, 2017.
S. Wang, L. Ma, **Yuming Fang**, *et al.*, Just noticeable difference estimation for screen content images, ***IEEE T-IP***, 2016.
H. Yang, **Yuming Fang**, *et al.*, Perceptual quality assessment for screen content images, ***IEEE T-IP**, 2015*.

NR VQA for SCIs



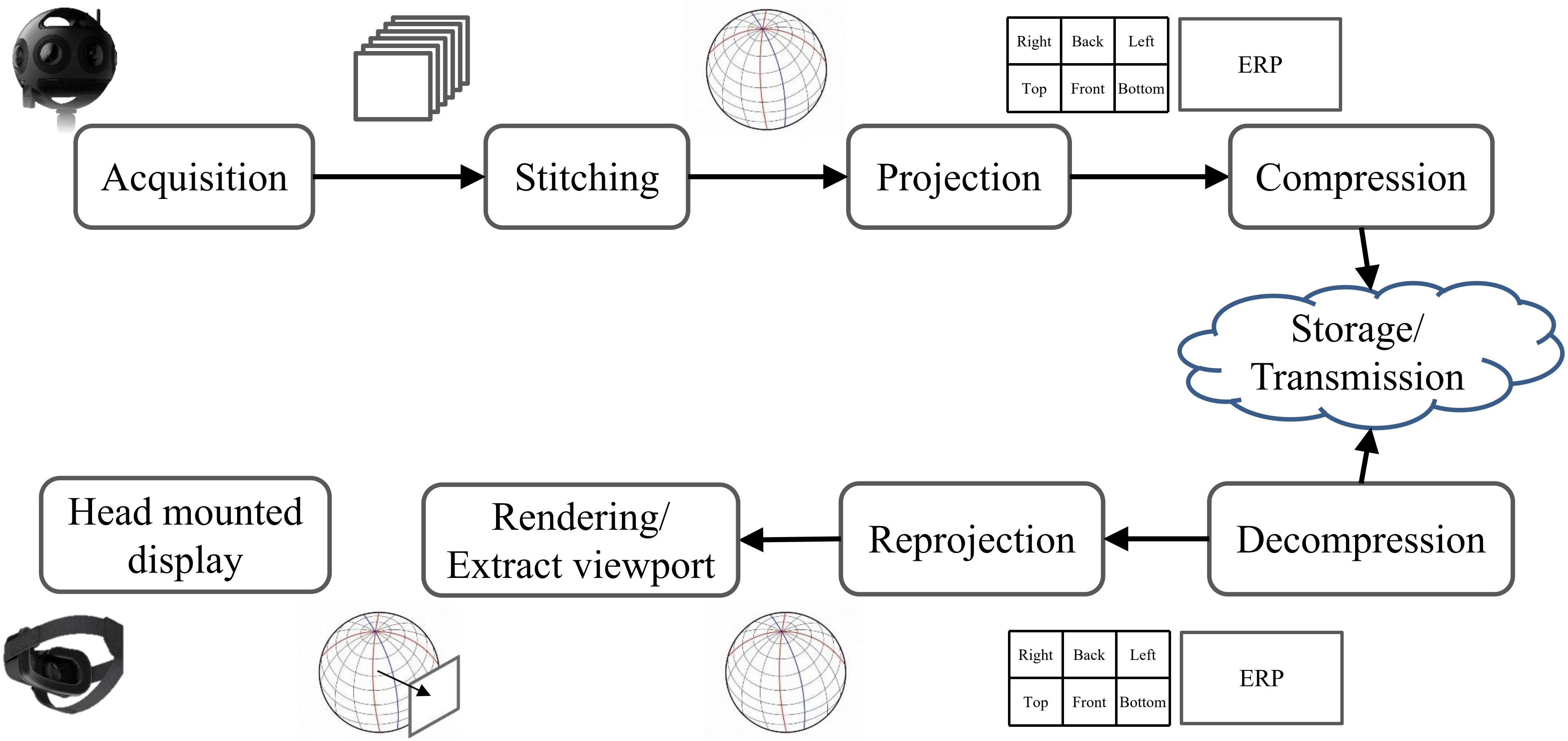
Yuming Fang, *et al.*, No reference quality assessment for screen content images with both local and global feature representation, ***IEEE T-IP***, 2017.

Summary

- We propose the first subjective database for SCIs, where a comprehensive study regarding to the sensitivity of the human visual system on the texture part and pictorial part is conducted.
- We propose an effective FR-IQA method for SCIs by uncertainty weighting, where two specific metrics are designed to capture quality degradation of textual and pictorial parts, and a uncertainty weighting is devised to fuse the quality scores of textual and pictorial parts.
- We propose a NR-IQA method for SCIs by incorporating statistical luminance and texture features with both local and global feature representation.

Yuming Fang, *et al.*, No reference quality assessment for screen content images with both local and global feature representation, ***IEEE T-IP***, 2017.
Yuming Fang, *et al.*, Objective quality assessment of screen content images by uncertainty weighting, ***IEEE T-IP***, 2017.
S. Wang, L. Ma, **Yuming Fang**, *et al.*, Just noticeable difference estimation for screen content images, ***IEEE T-IP***, 2016.
H. Yang, **Yuming Fang**, *et al.*, Perceptual quality assessment for screen content images, ***IEEE T-IP**, 2015*.

Panoramic Video Processing



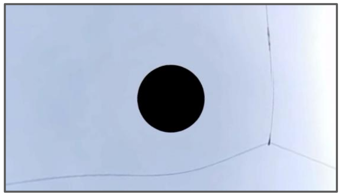
Visual Distortion in Panoramic Photography

Projection distortions

Equirectangular

Cubemap

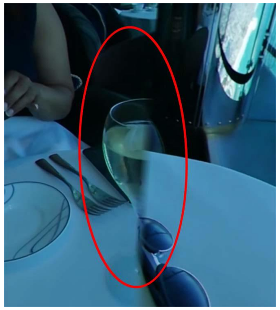
Post-processing



Post-processing on the poles

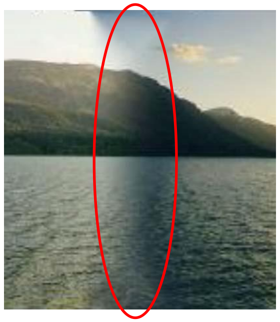
Stitching distortions

Broken edges



Missing information

Ghosting

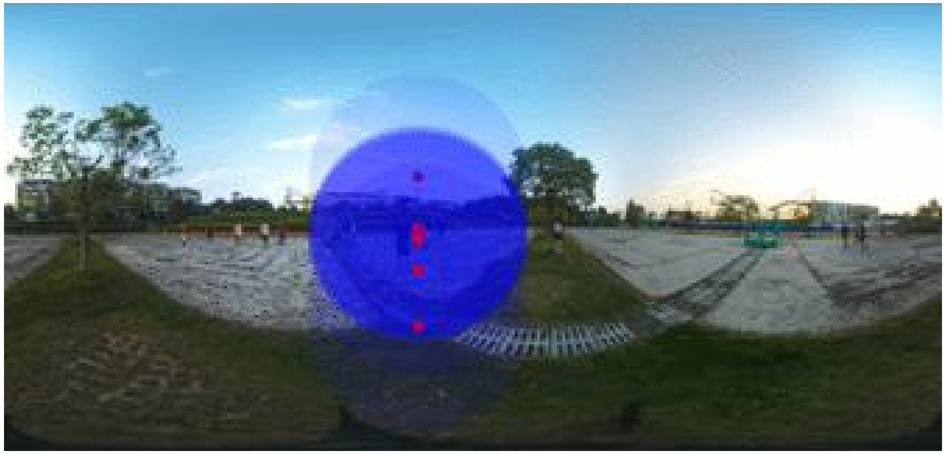


Exposure

Geometrical distortions

VQA for Panoramic Images

Diversity of observer behavior：
Different viewing conditions (starting point and viewing time)



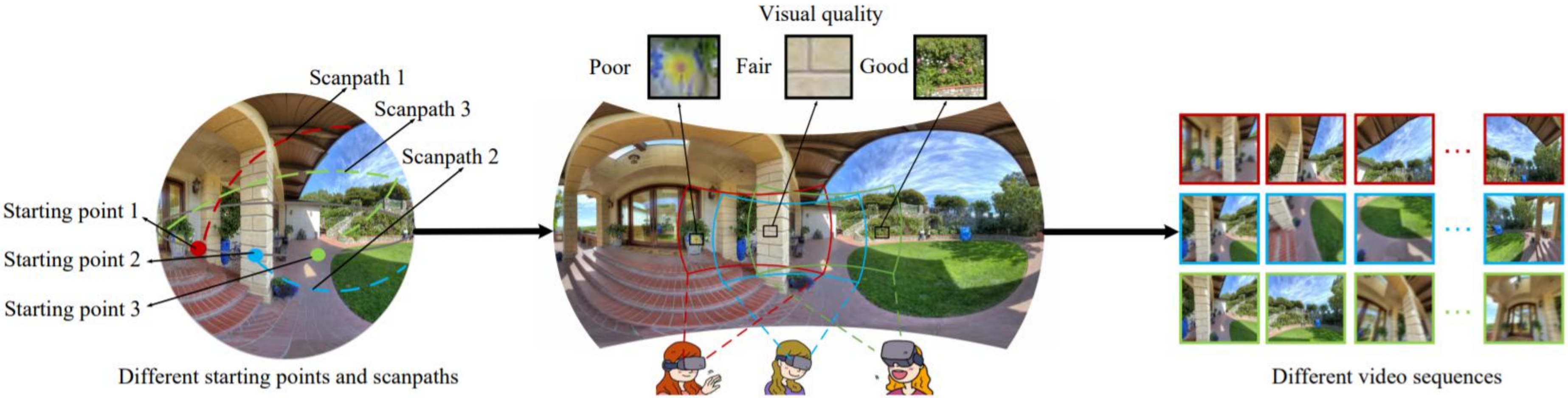
Starting Point 1

Starting Point 2

VQA for Panoramic Images

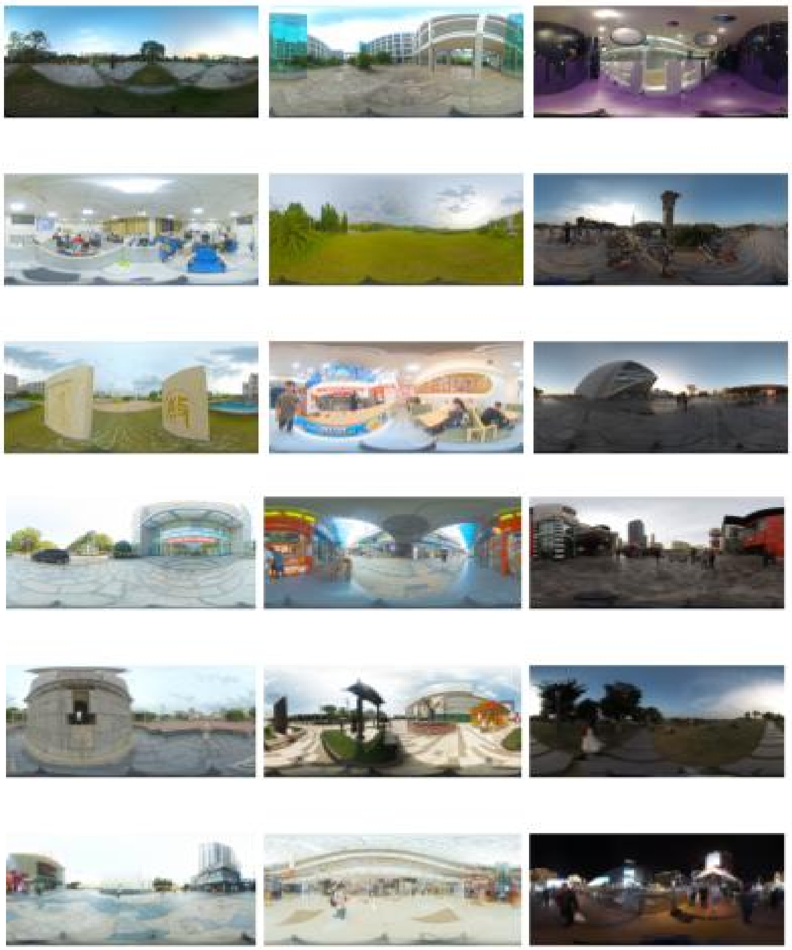
Diversity of observer behavior：
Different viewing conditions (starting point and viewing time)

VQA for Panoramic Images



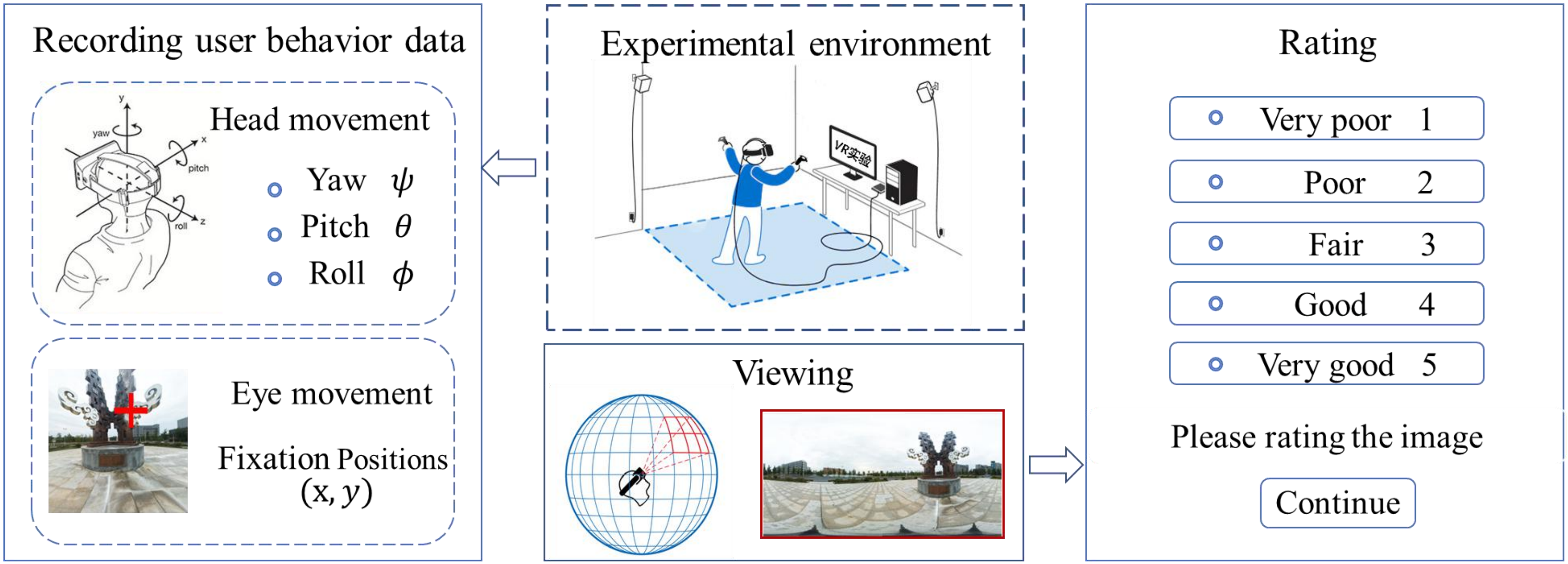
- Each user may have different viewing behaviors (i.e., scanpaths) under different viewing conditions, giving rise to different video representations of the same 360 image/video with varying perceived quality.
- We consider two types of viewing conditions - the starting point and the exploration time are important in influencing the perceived quality of 360 image/video.

Panoramic Images Quality Assessment Database



- The database contains 36 source panoramic images and the corresponding distorted panoramic images. The types of distortions include H.265 compression and stitching. Besides, 2 VR viewing conditions (i.e., the starting point and the exploration time) are adopted.

Subjective Quality Assessment



22 subjects join in the experiment and rating 36 distorted panoramic images.

Data Analysis

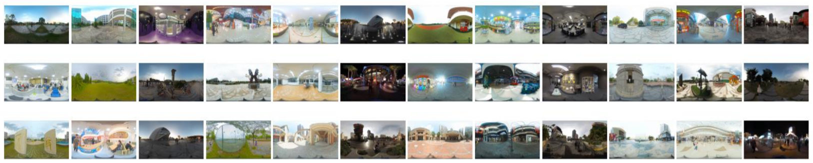


Fig. 3: The sample of reference images in the database.

Fig. 6: Consistency of user viewing behaviors under different starting points. The exploration time is fixed to 5 seconds. (a) Initial viewport that contains a passerby, which attracts human visual attention and leads to a higher PLCC of 0.935. (b) Initial viewport that exhibits symmetrical image structures with no eye-catching event, leading to a much lower PLCC of 0.187.

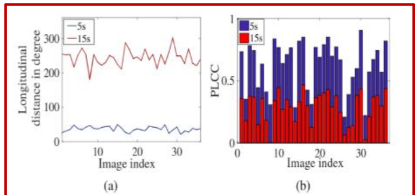


Fig. 7: (a) Farthest longitudinal distance to the starting point averaged across viewers for 5 and 15 seconds of exploration. (b) Consistency between scanpaths from different users in terms of PLCC for 5 and 15 seconds of exploration.

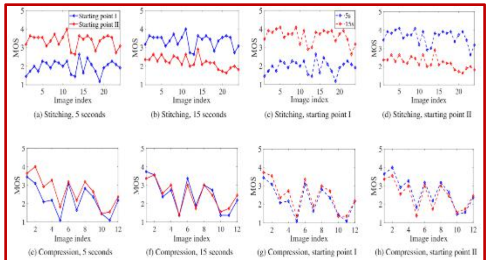


Fig. 8: MOSs of 360° images in the proposed database under different viewing conditions.

Source of variation	*SS*	*d.f.*	*MS*	*F*	*p*
Starting point	2.76	1	2.76	9.60	≈0
Exploration time	0	1	0	0.01	0.91
Distortion type	3.49	1	3.49	12.14	≈0
Starting point× Exploration time	19.06	1	19.10	66.28	≈0
Starting point× Distortion type	0.08	1	0.08	0.29	0.59
Exploration time× Distortion type	0	1	0	0.01	0.9048
Starting point× Exploration time× Distortion type	9.45	1	9.45	32.86	≈0
Residual	39.11	136	0.2876		
Total	88.31	143			

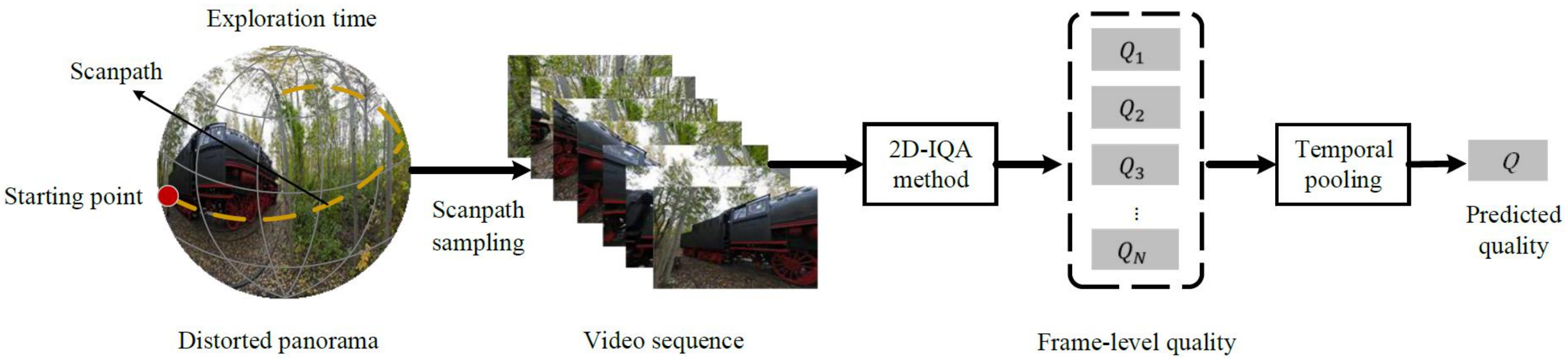
Table 2: The results of multi-factorial ANOVA test. '*SS*' denotes Sum of Squares, '*d.f*' indicates Degrees of Freedom, '*MS*' means Mean Square, '*F*' denotes F value, and '*p*' is p-value for the null hypothesis.

Conclusion 1: Viewing conditions have a important effect in influencing the user's viewing behavior, which may further affect the perceptual quality.

Conclusion 2: When the panoramic images are locally distorted, viewing conditions have a significant impact on the perceptual quality.

Conclusion 3: The recency effect is clearly observed when the users explore locally distorted panoramic images.

Objective Quality Models



We propose a general computational framework for panoramic IQA, where user viewing conditions and behaviors are incorporated naturally by treating panoramic images as moving camera videos

Experimental Results

	Proposed database						OIQA database									
	ST		H.265		Overall		JPEG		JP2K		GB		GN		Overall	
	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
S-PSNR	0.151	-0.113	0.931	0.890	0.225	-0.103	0.890	0.847	0.886	0.887	0.784	0.780	0.915	0.881	0.763	0.751
S-SSIM	0.149	0.055	0.922	**0.932**	0.018	-0.031	0.922	0.903	0.930	0.931	0.869	0.870	0.955	0.941	0.828	0.823
WS-PNSR	0.153	-0.116	0.931	0.893	0.215	-0.104	0.890	0.847	0.886	0.886	0.785	0.781	0.915	0.881	0.764	0.751
CPP-PNSR	0.129	-0.054	0.930	0.906	0.215	-0.079	0.891	0.849	0.885	0.885	0.767	0.764	0.914	0.878	0.757	0.747
PSNR	0.165	-0.114	0.924	0.893	0.231	-0.102	0.891	0.848	0.891	0.893	0.759	0.754	0.925	0.895	0.744	0.733
V-PSNR	0.148	-0.049	0.928	0.893	0.241	-0.077	0.905	0.898	0.897	0.896	0.835	0.831	0.913	0.884	0.795	0.779
O-PSNR	0.583	0.516	**0.933**	0.911	0.597	0.467	0.905	0.891	0.901	0.901	0.884	0.886	0.914	0.881	0.797	0.780
SSIM	0.148	0.057	0.910	**0.932**	0.036	-0.030	0.910	0.893	0.924	0.926	0.849	0.845	0.951	0.937	0.809	0.802
V-SSIM	0.149	0.044	0.930	0.916	0.038	-0.038	0.924	0.905	0.932	0.931	0.891	0.891	0.942	0.929	0.850	0.844
O-SSIM	0.468	0.495	0.923	0.881	0.579	0.435	0.938	0.922	0.941	0.939	0.918	0.921	0.942	0.930	0.866	0.862
VIF	0.111	0.057	0.920	0.872	0.356	0.331	0.916	0.900	0.955	0.956	0.960	0.958	0.950	0.921	0.871	0.862
V-VIF	0.151	0.046	0.923	0.861	0.493	0.342	0.929	0.915	0.960	0.962	0.957	0.954	0.947	0.916	0.883	0.873
O-VIF	**0.605**	**0.555**	0.893	0.843	0.617	0.496	0.937	0.923	0.969	0.968	0.965	0.965	0.947	0.917	0.889	0.880
NLPD	0.012	-0.009	0.907	0.870	0.244	-0.063	0.925	0.945	0.919	0.947	0.849	0.893	0.952	0.947	0.854	0.844
V-NLPD	0.069	-0.017	0.895	0.892	0.244	-0.065	0.964	0.954	0.954	0.954	0.933	0.933	0.970	0.957	0.911	**0.907**
O-NLPD	0.479	0.534	0.898	0.857	0.311	0.472	**0.972**	**0.958**	0.964	0.962	0.942	0.945	**0.974**	**0.963**	**0.912**	**0.907**
DISTS	0.079	0.025	0.867	0.861	0.450	0.299	0.863	0.915	0.939	0.952	0.959	0.956	0.951	0.944	0.837	0.830
V-DISTS	0.055	0.069	0.900	0.910	0.512	0.402	0.942	0.937	0.961	0.959	0.965	0.957	0.963	0.949	0.883	0.875
O-DISTS	0.489	0.518	0.916	0.903	**0.660**	**0.613**	0.955	0.942	**0.971**	**0.969**	**0.973**	**0.969**	0.966	0.952	0.882	0.875

Summary

- We conduct a psychophysical experiment to study the interplay among the VR viewing conditions, the user viewing behaviors, and the perceived quality of panoramic images. Thorough analysis of the collected human data validates that viewing conditions have an important impact on the perceived quality of panoramic images.
- We propose a computational framework for objective quality assessment of distorted panoramas, incorporating viewing conditions and behaviors.

Database & Models:

https://github.com/xiangjieSui/img2video

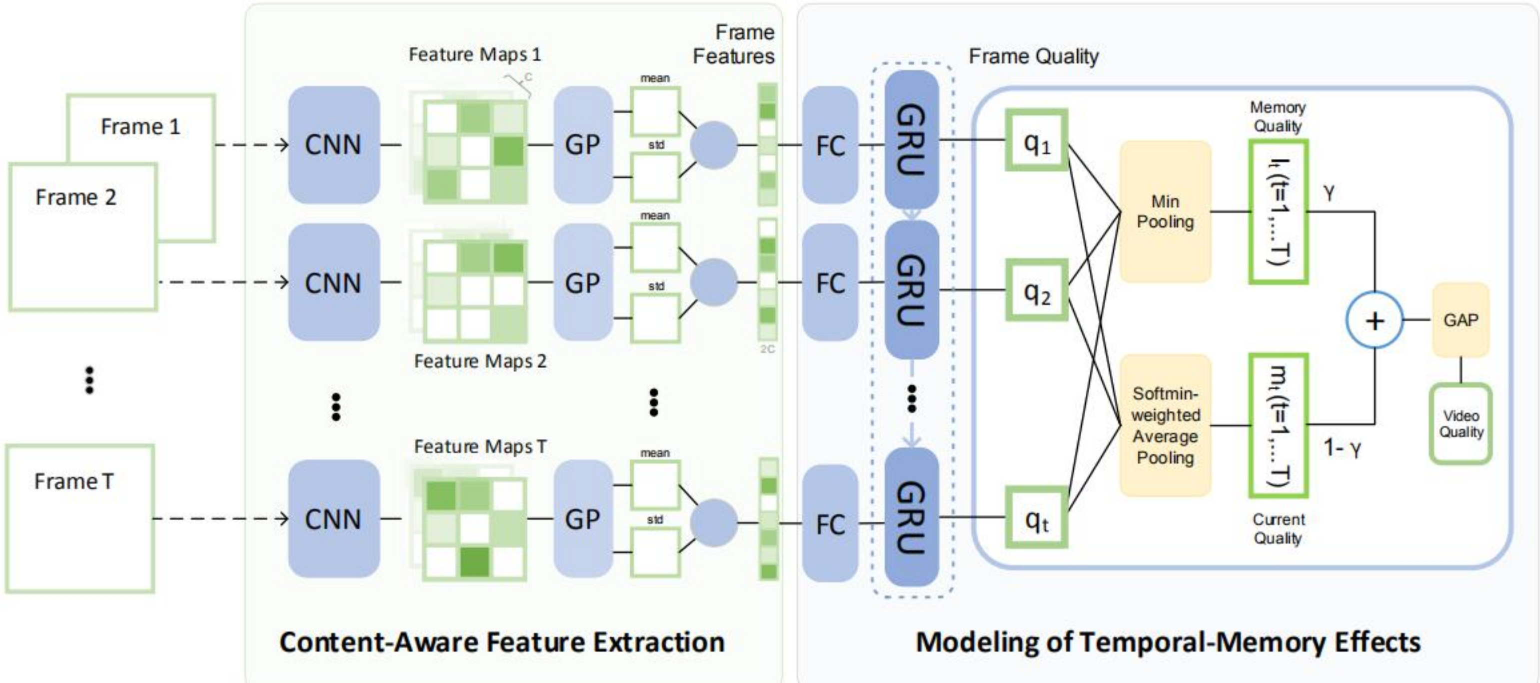
X. Sui, **Kede Ma**, Y. Yao, and **Yuming Fang**, Perceptual quality assessment of omnidirectional images as moving camera videos, ***IEEE T-VCG***, 2021.

Video Quality Assessment

- How to effectively capture spatial distortion.
- How to measure spatiotemporal degradation.

J. Park. K. Seshadrinathan, S. Lee, A. C. Bovik, Video quality pooling adaptive to perceptual distortion severity, ***IEEE T-IP***, 2013.
Yuming Fang, *et al*., Asymmetrically distorted 3D video quality assessment: From the motion variation to perceived quality, ***SP***, 2021.

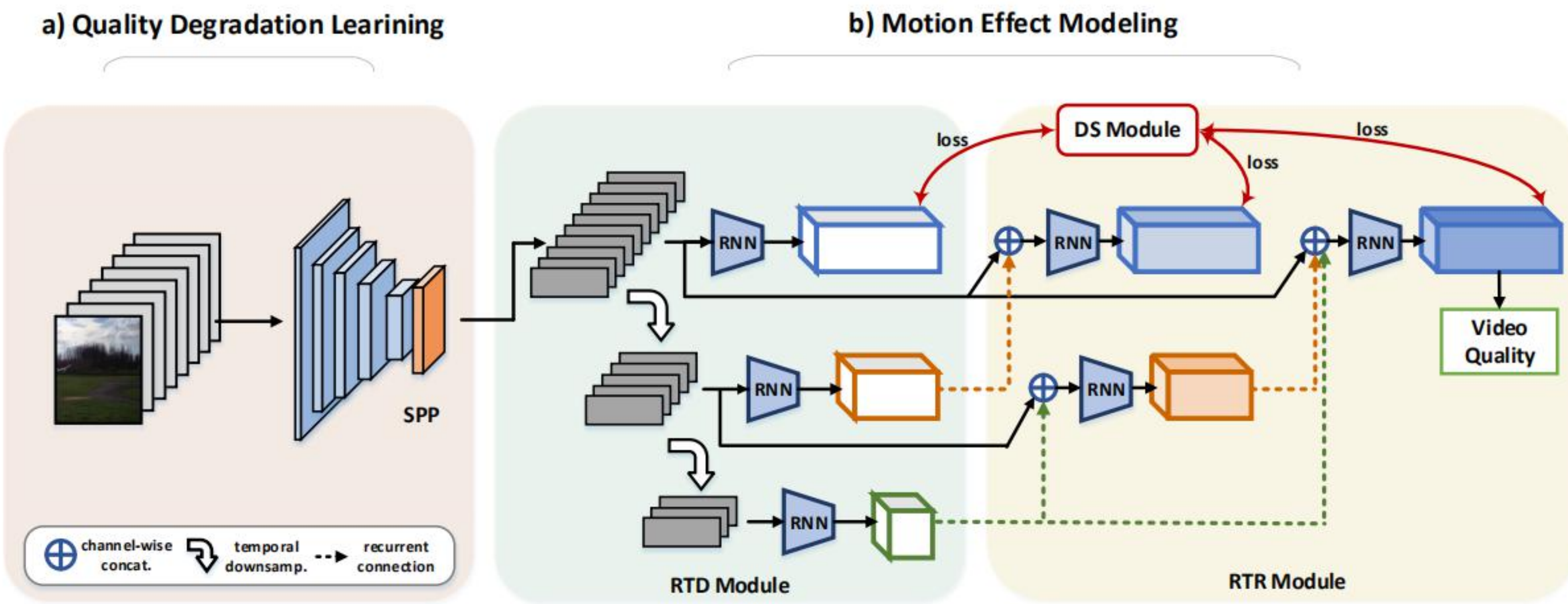
Quality Assessment of In-the-Wild Videos



The framework which consists of two modules. (a) *Content-aware feature extraction* is a pre-trained CNN with effective global pooling serving as a feature extractor. (b) *Modeling of temporal-temporal effects*: a GRU network and a subjectively-inspired temporal pooling layer.

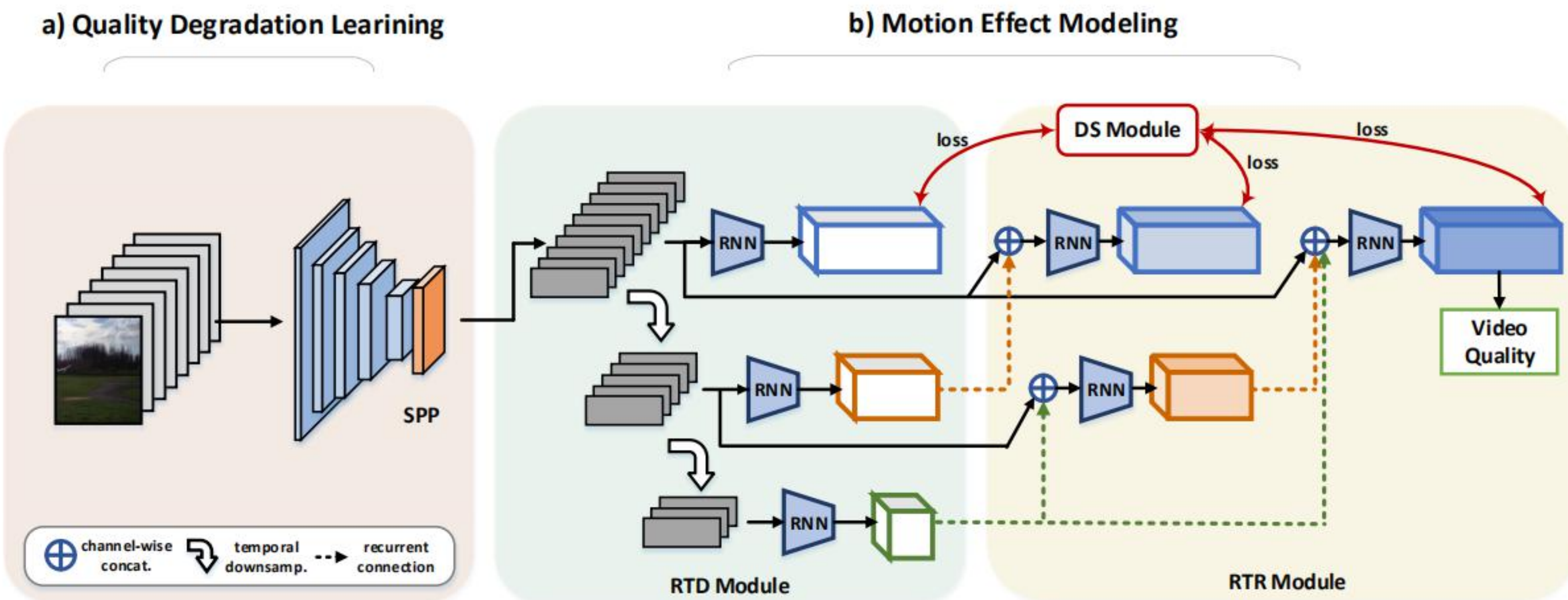
D. Li, T. Jiang, M. Jiang, Quality assessment of in-the-wild videos, in ***ACM MM***, 2019.

RIRNet: Recurrent-In-Recurrent Network for Video Quality Assessment



F. Chen, L. Li, L. Ma, J. Wu, G. Shi, RIRNet: Recurrent-in-recurrent network for video quality assessment, in *ACM MM*, 2020.

RIRNet: Recurrent-In-Recurrent Network for Video Quality Assessment



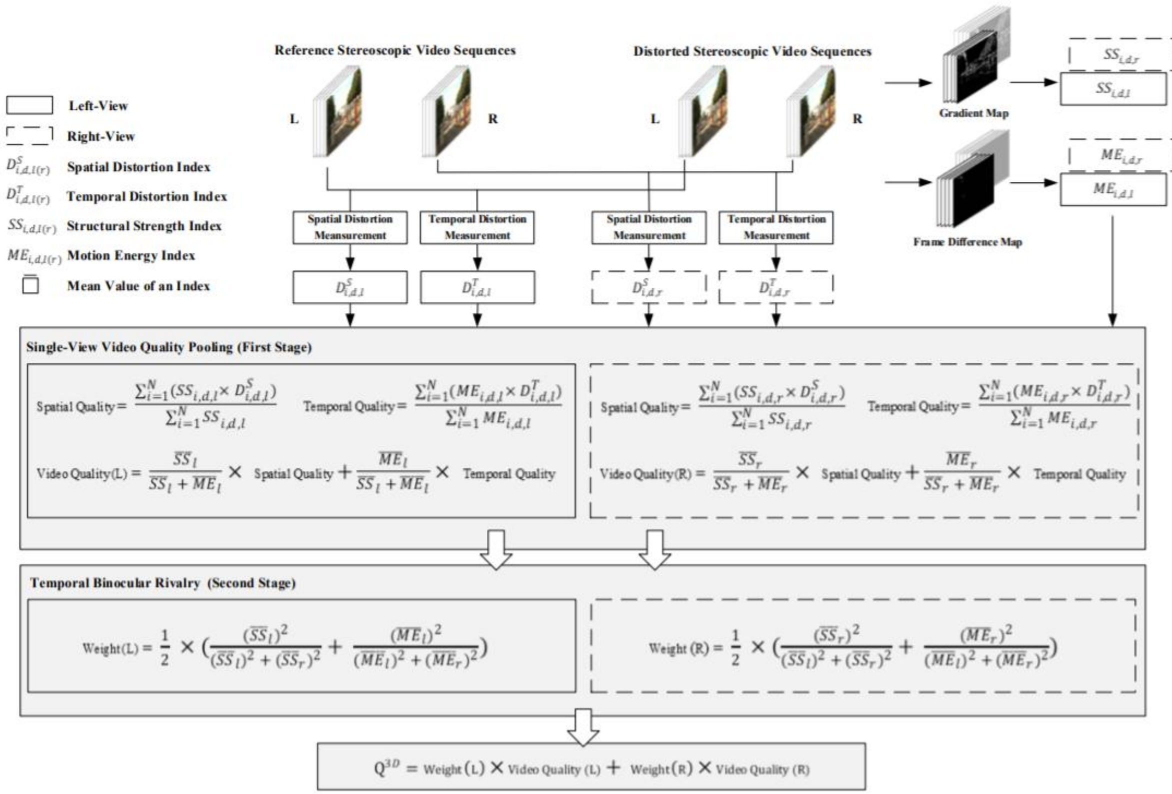
Two main modifications

(a) Temporal down-sampling.

(b) Fusing the multiple motion information with different temporal frequencies.

F. Chen, L. Li, L. Ma, J. Wu, G. Shi, RIRNet: Recurrent-in-recurrent network for video quality assessment, in *ACM MM*, 2020.

Perceptual Quality Assessment for Asymmetrically Distorted Stereoscopic Video by Temporal Binocular Rivalry



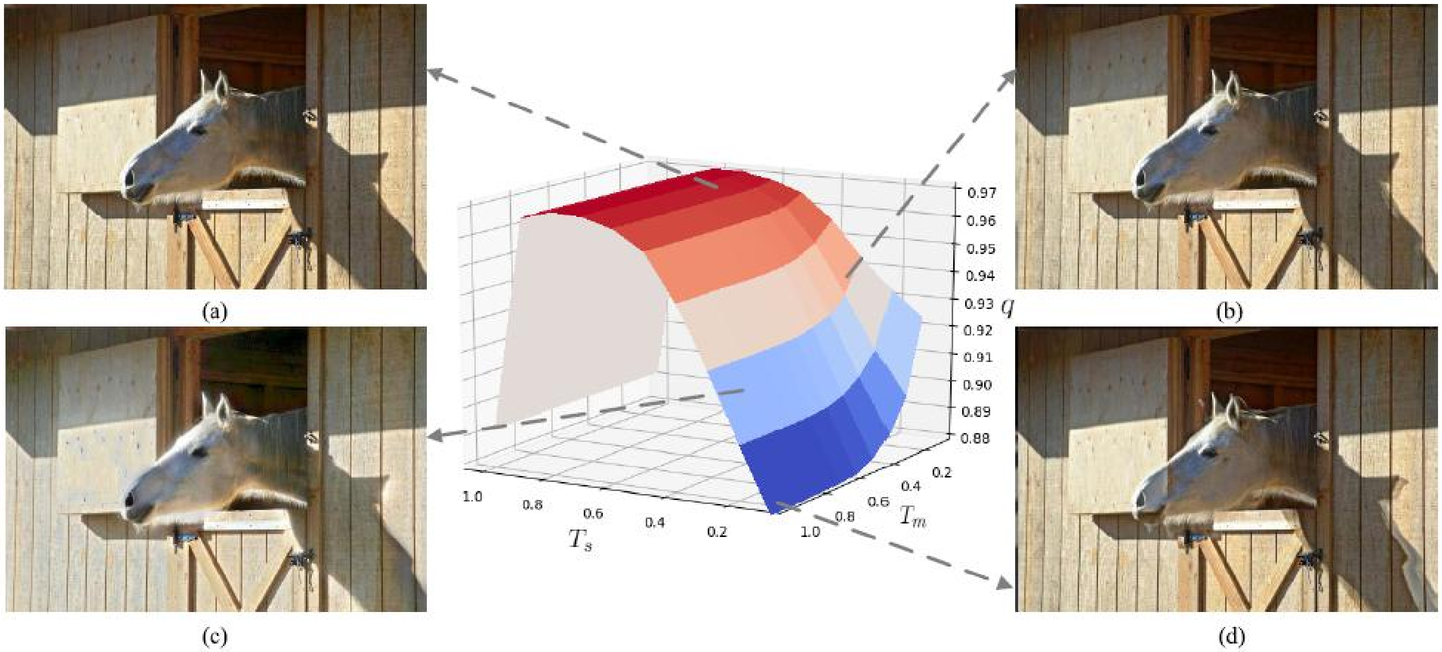
A two-stage framework

(a) Stage 1: single-view video quality prediction.

(b) Stage 2: Stereoscopic video quality prediction by considering temporal binocular rivalry.

Yuming Fang, *et al.*, Perceptual quality assessment for asymmetrically distorted stereoscopic video by temporal binocular rivalry, ***IEEE T-CSVT***, 2021.

Application I: Parameter Tuning



Warmer color in the surface plot indicates better predicted quality of SPD-MEF. (a) q = 0.971. (b) q = 0.934. (c) q = 0.901. (d) q = 0.885.

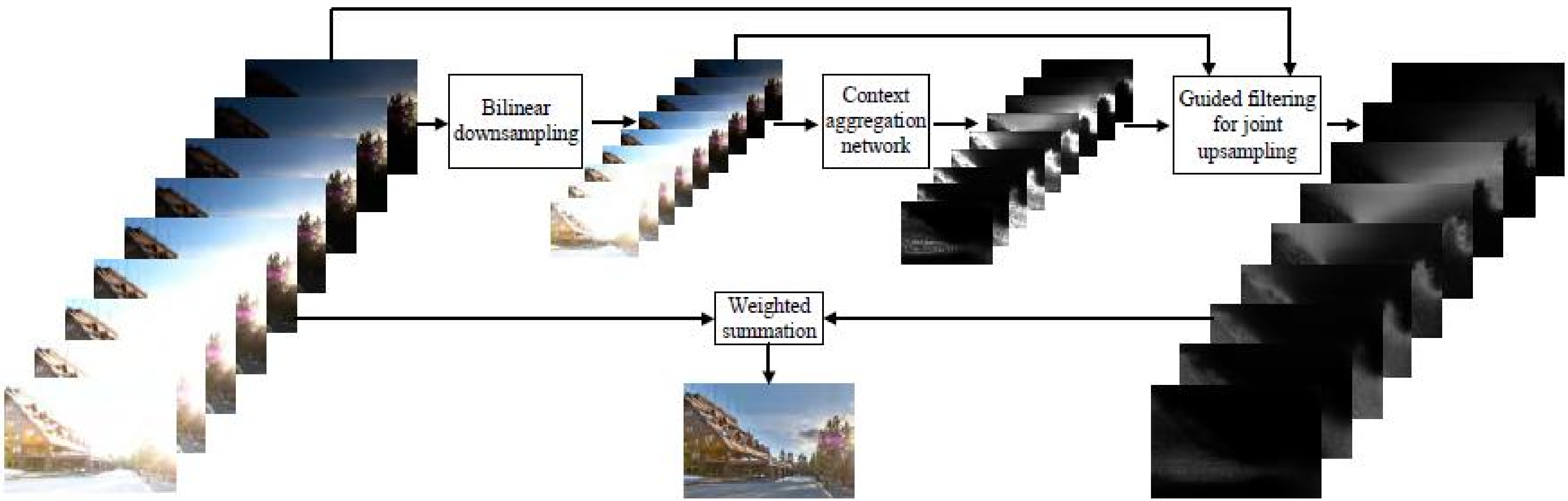
Yuming Fang, H. Zhu, **Kede Ma**, *et al.*, Perceptual evaluation for multi-exposure image fusion of dynamic scene, ***IEEE T-IP***, 2020.

Yuming Fang, *et al.*, Superpixel-based quality assessment of multi-exposure image fusion for both static and dynamic scenes, ***IEEE T-IP***, 2021.

Application II: Perceptual Optimization

Perceptual optimization for the proposed MEF quality metric Q:

$$Y_{opt} = \arg\max_{Y} Q(\{X_k\}, Y)$$



Multi-Exposure Image Fusion

Kede Ma, H. Zhu, **Yuming Fang**, *et al.*, Deep guided learning for fast multi-exposure image fusion, ***IEEE T-IP***, 2020.

Context Aggregation Network

Specification of the CAN in MEF-Net for low-resolution weighting map predication

Layer	1	2	3	4	5	6	7
Convolution	3×3	3×3	3×3	3×3	3×3	3×3	1 × 1
Dilation	1	2	4	8	16	1	1
Width	24	24	24	24	24	24	1
Bias	✗	✗	✗	✗	✗	✗	✓
Adaptive normalization	✓	✓	✓	✓	✓	✓	✗
Nonlinearity	✓	✓	✓	✓	✓	✓	✗
Receptive field	3×3	7×7	15×15	31×31	63×63	65×65	65×65

Adaptive normalization:

$$\mathrm{AN}(\mathbf{Z}) = \lambda_n \mathbf{Z} + \lambda'_n \mathrm{IN}(\mathbf{Z})$$

Where $\lambda_n, \lambda'_n \in \mathbb{R}$ are learnable scalar weights, **Z** indicates the intermediate representations, and IN() stand for the instance normalization operators.

Kede Ma, H. Zhu, **Yuming Fang**, *et al.*, Deep guided learning for fast multi-exposure image fusion, ***IEEE T-IP***, 2020.

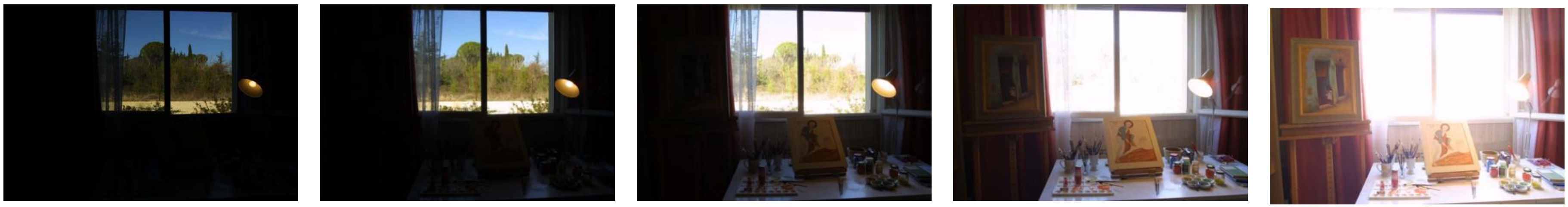
Qualitative Comparison

Source image sequence

Learned weight maps

Fused image of the proposed method

Qualitative Comparison



Source Image Sequence

Mertens09

SPD-MEF

Ours

Kede Ma, H. Zhu, **Yuming Fang**, *et al.*, Deep guided learning for fast multi-exposure image fusion, ***IEEE T-IP***, 2020.

Qualitative Comparison

Source Image Sequence

GGIF

Li13

Ours

Kede Ma, H. Zhu, **Yuming Fang**, *et al.*, Deep guided learning for fast multi-exposure image fusion, ***IEEE T-IP***, 2020.

Quantitative Comparison

Average MEF-SSIM and MEF-VIF scores of different MEF methods

MEF method	Mertens09 [2]	Li13 [4]	SPD-MEF [5]	GGIF [7]	DeepFuse [6]	MEF-Opt [8]	MEF-Net
MEF-SSIM [9]	0.923	0.945	0.953	0.958	0.862	**0.978**	**0.964**
MEF-VIF [42]	**0.969**	0.967	0.956	**0.972**	0.926	0.952	0.967

Average MEF-SSIM score as a function of input resolution, depth, and width of CAN in MEF-Net. The default setting is heighted in bold

Input res		32	64	**128**	256	
MEF-SSIM		0.950	0.960	0.964	**0.967**	
Depth	4	5	6	**7**	8	9
MEF-SSIM	0.961	0.963	0.963	0.964	**0.965**	**0.965**
Width	8	16	**24**	32	48	64
MEF-SSIM	0.953	0.963	0.964	0.966	**0.967**	**0.967**

Average MEF-SSIM score as a function of the regularization parameter λ_a and the radius r in the guided filter. The default setting is heighted in bold

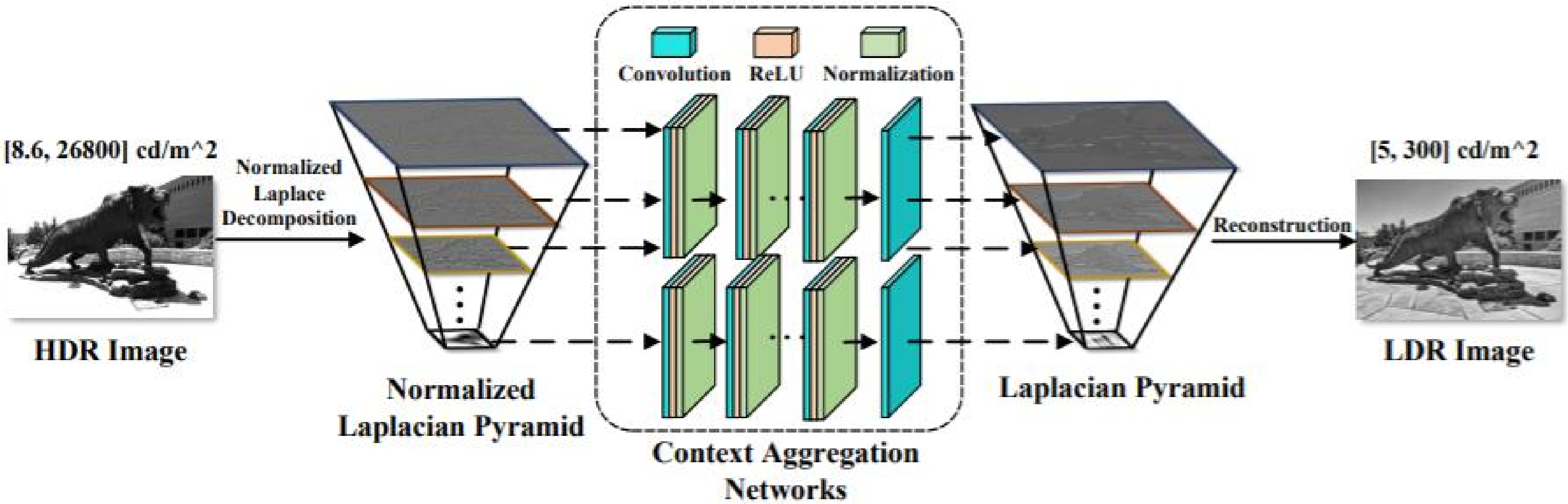
λ_a	10^{-1}	10^{-2}	$\mathbf{10^{-4}}$	10^{-6}	10^{-8}
MEF-SSIM	0.961	0.961	**0.964**	0.962	0.961
r	**1**	2	4	8	16
MEF-SSIM	**0.964**	0.963	0.959	0.956	0.950

Kede Ma, H. Zhu, **Yuming Fang**, *et al.*, Deep guided learning for fast multi-exposure image fusion, ***IEEE T-IP***, 2020.

Application II: Perceptual Optimization

Perceptual optimization for the quality metric NLPD:

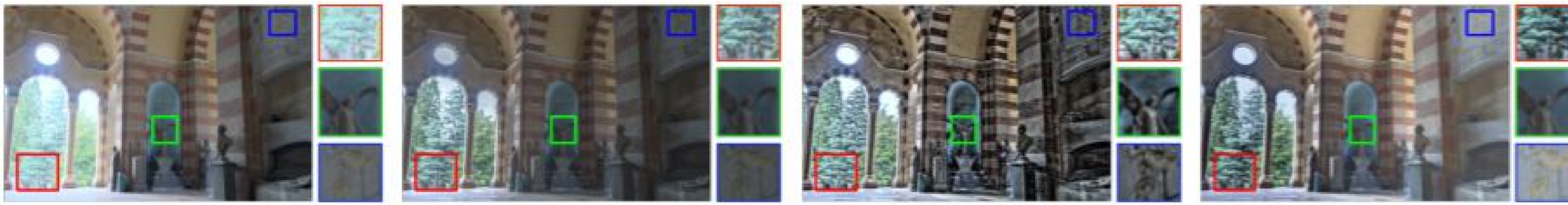
$$L_{opt} = \arg\max_{L} Q(H, L)$$



Tone mapping

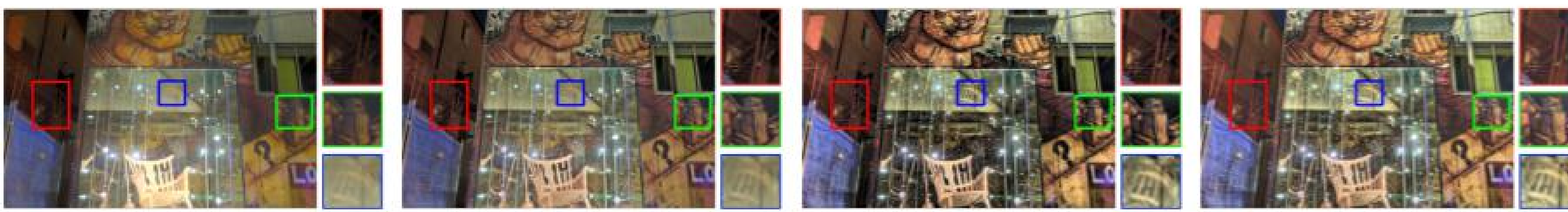
C. Le, J. Yan, **Yuming Fang**, **Kede Ma**,, Deep guided learning for fast multi-exposure image fusion, in ***ICVRV***, 2021.

Application II: Perceptual Optimization



(a) Drago03 [10]　(b) LLF [13]　(c) GR [15]　(d) Ours

Fig. 3: Tone mapping results of the “Architecture” image courtesy of Nemoto Hiromi.



(a) Kim08 [12]　(b) Liang18 [16]　(c) NLPD-Opt [20]　(d) Ours

Fig. 4: Tone mapping results of the “Night Building” image courtesy of Nemoto Hiromi.

C. Le, J. Yan, **Yuming Fang**, **Kede Ma**,, Deep guided learning for fast multi-exposure image fusion*,* in ***ICVRV**,* 2021.

Application II: Perceptual Optimization

(a) One-level (b) Two-level (c) Three-level (d) Four-level (e) Five-level

Fig. 5: Tone mapping results of the "Workshop" image with different input pyramid levels. Image courtesy of Nemoto Hiromi.

(a) MAE-optimized (b) SSIM-optimized [26] (c) TMQI-optimized [18] (d) NLPD-optimized [20]

Fig. 6: Tone mapping results of the "Man" image with different objective functions. Image courtesy of Nima Khademi Kalantari.

C. Le, J. Yan, **Yuming Fang**, **Kede Ma**,, Deep guided learning for fast multi-exposure image fusion, in ***ICVRV***, 2021.

Application II: Perceptual Optimization

Full-reference IQA models for perceptual optimization :

$$y^* = \arg\min_y D(x,\ y)$$

were D denotes a full-reference IQA measure with a lower score indicating higher predicted quality, and y^* is the recovered image.

- **Tasks**
  - Image denoising.
  - Blind image deblurring.
  - Single image super-resolution.
  - Lossy image compression.

K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV***, 2021.

Application II: Perceptual Optimization

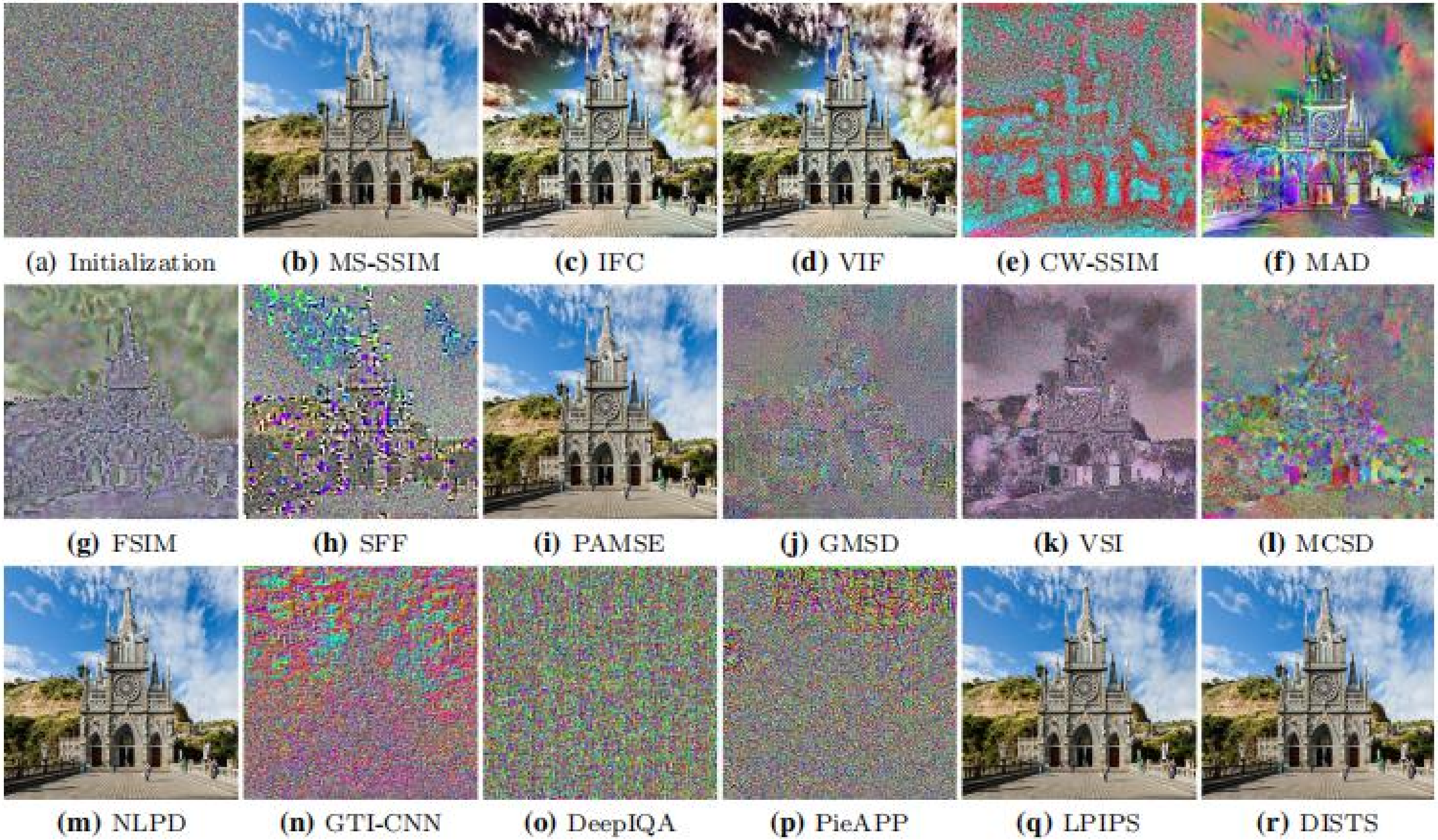


Fig. 1 Reference image recovery test. Starting from (**a**) a white Gaussian noise image, we recover images by optimizing the predicted quality relative to a reference image, using different IQA models (**b**)–(**r**)

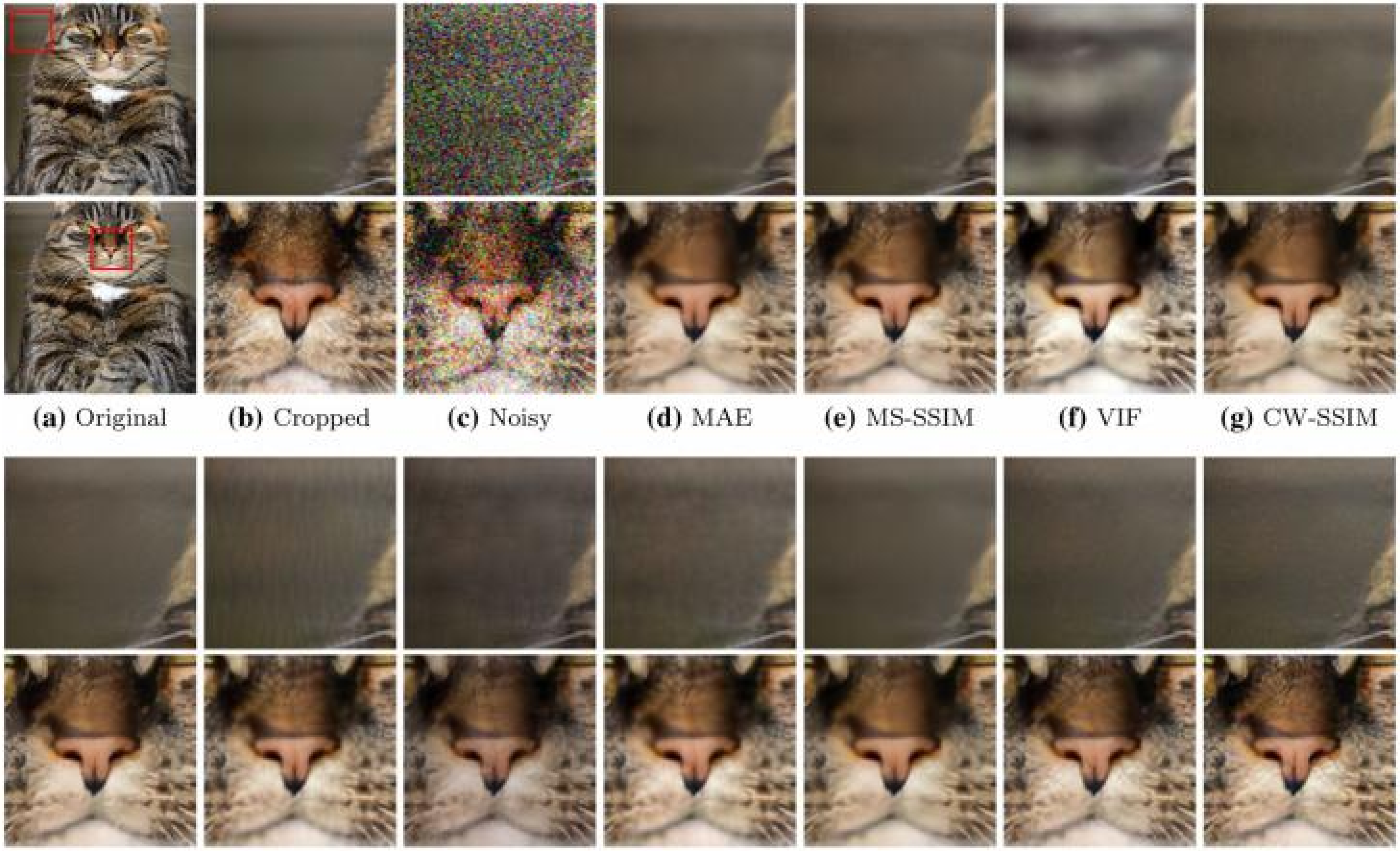
K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV***, 2021.

Application II: Perceptual Optimization

Fig. 2 Reference image recovery test. Staring from (**a**) a JPEG compressed version of a reference image, we recover images by optimizing the predicted quality relative to the reference image, using different IQA models (**b**)–(**r**)

K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV***, 2021.

Application II: Perceptual Optimization



(a) Original (b) Cropped (c) Noisy (d) MAE (e) MS-SSIM (f) VIF (g) CW-SSIM

(h) MAD (i) FSIM (j) GMSD (k) VSI (l) NLPD (m) LPIPS (n) DISTS

Fig. 10 Denoising results on two regions cropped from an example image, using a DNN optimized for different IQA models

K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV**,* 2021.

Application II: Perceptual Optimization

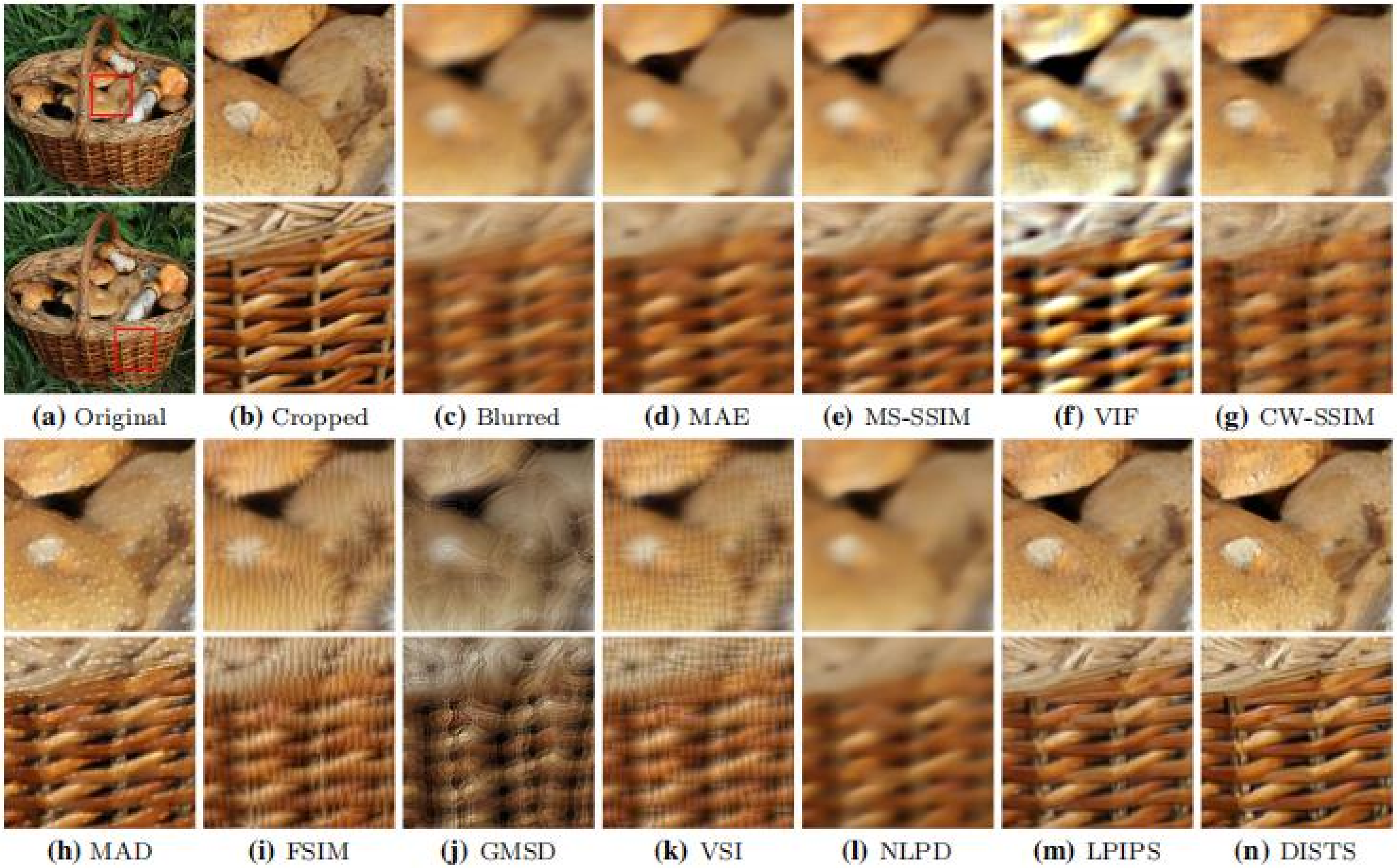


Fig. 11 Deblurring results for two regions cropped from an example image, using a DNN optimized for different IQA models

K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV***, 2021.

Application II: Perceptual Optimization

Fig. 13 Compression results for two cropped regions from an example image, using a DNN optimized for different IQA models

K. Ding, **Kede Ma**, *et al.*, Comparison of full-reference image quality models for optimization of image processing systems, ***IJCV***, 2021.

Summary

- **Current status**: many IQA methods designed specifically for different contents have been proposed, and many efforts have been put on perceptual optimization. On the whole, researchers have achieved giant and excited success in the filed of IQA.
- **Outlook**
  - Robust, feasible, generalizable IQA models.
  - Deeper and wider application in the field of image processing and computer vision, etc.
  - Looking forward to more interesting works.

Thanks For Your Attention！