VR IQA NET:

DEEP VIRTUAL REALITY IMAGE QUALITY ASSESSMENT USING ADVERSARIAL LEARNING

ICASSP 2018, Calgary, Canada

Heoun-taek Lim, Hak Gu Kim, and Yong Man Ro
Image and Video sYstems (IVY) Lab.
Korea Advanced Institute of Science and Technology (KAIST)

Introduction of Virtual Reality

YouTube

STEAMVR™

PlayStation.VR

vimeo

oculus

Introduction of Virtual Reality

Spherical image

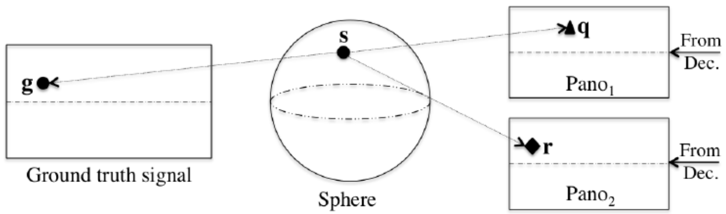
projection

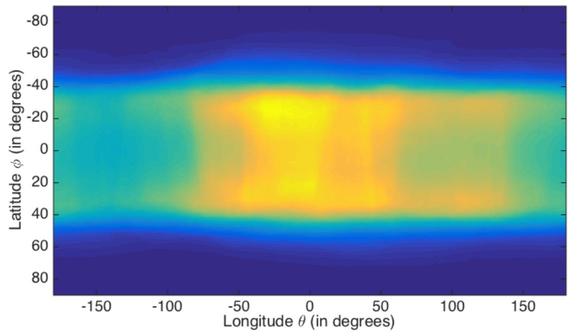
Omnidirectional (360 degree) image

Compression based on conventional 2D IQA model

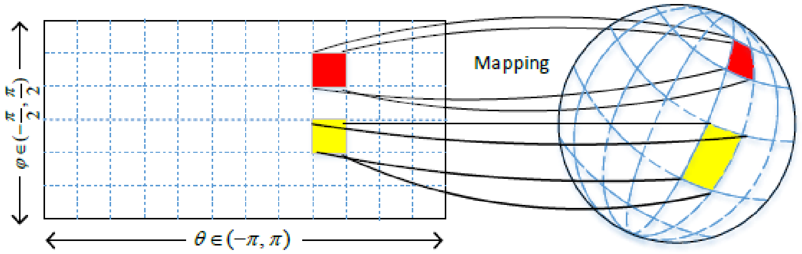
Image Quality Assessment in VR

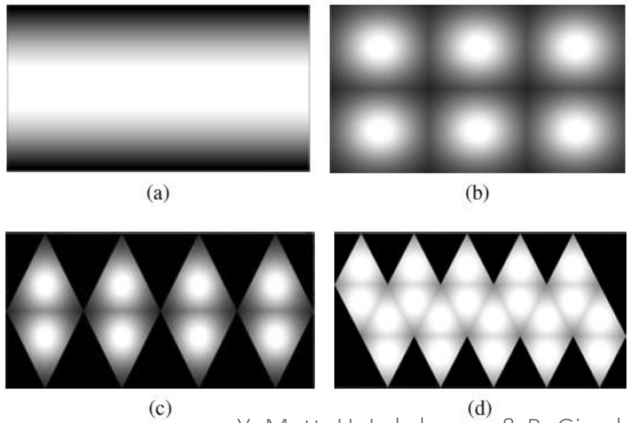
S-PSNR



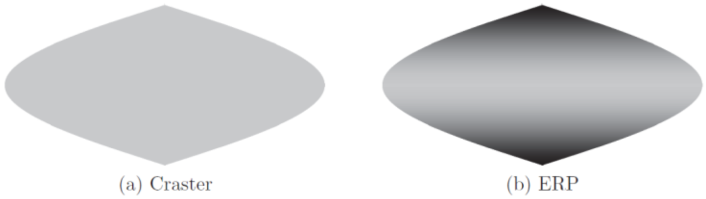


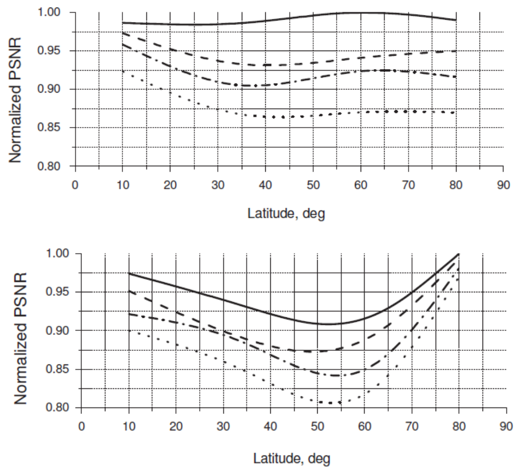
WS-PSNR





CPP-PSNR





Y. Matt, H. Lakshman, & B. Girod, "A framework to evaluate omnidirectional video coding schemes". ISMAR 2015.

S. Yule, A. Lu, & Y. Lu, "Weighted-to-Spherically-Uniform Quality Evaluation for Omnidirectional Video". Signal processing letters, 2017.

V. Zakharchenko, K. P. Choi, and J. H. Park, "Quality metric for spherical panoramic video". SPIE 9970, Optics and Photonics for Information Processing X, 2016.

Limitation of Image Quality Assessment in VR

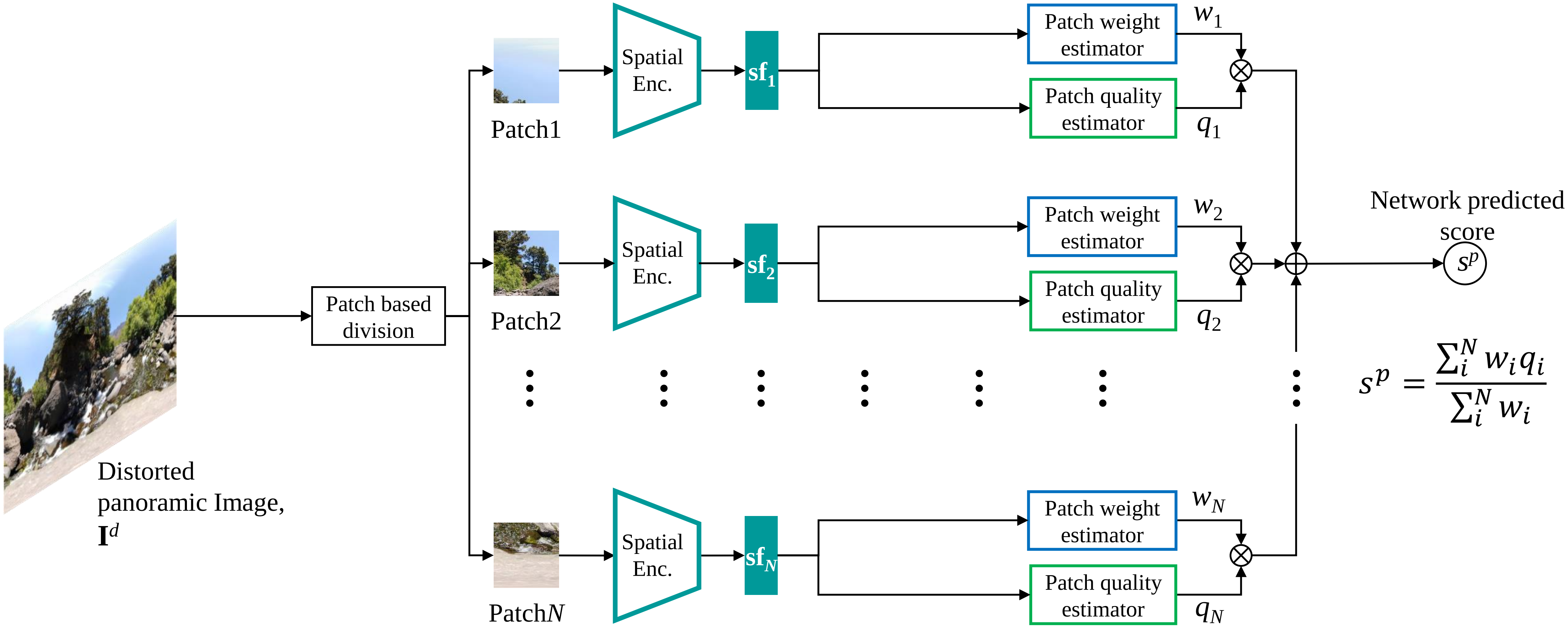
Metric	PLCC				SROCC				RMSE				OR			
	A	*B*	*C*	*D*	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*
PSNR	0.8714	0.8437	0.8553	0.9487	0.7176	0.7731	0.7567	**0.8909**	0.4804	0.5103	0.5008	0.2929	0.4375	0.4375	0.4167	**0.2396**
SSIM	0.8898	0.8632	0.8740	0.9459	0.7365	0.7927	0.7709	0.8821	0.4464	0.4790	0.4689	0.3050	**0.3958**	0.4583	0.4167	0.2812
MSSSIM	0.9059	0.8661	0.8860	0.9123	0.7539	0.7796	0.7814	0.8394	0.4143	0.4755	0.4483	0.3887	0.4583	0.4167	0.4271	0.3229
VIFP	**0.9116**	**0.8875**	**0.8994**	0.9319	**0.7608**	**0.8029**	**0.7953**	0.8538	**0.4025**	**0.4374**	**0.4221**	0.3395	**0.3958**	**0.3958**	0.4167	0.3125
S-PSNR	0.8766	0.8482	0.8392	0.9168	0.7376	0.7836	0.7307	0.8214	0.4715	0.5035	0.5257	0.3705	0.4583	0.4375	0.4271	0.3021
WS-PSNR	0.8748	-	-	**0.9583**	0.7297	-	-	0.8648	0.4746	-	-	**0.2544**	0.4375	-	-	0.2500
CPP-PSNR	0.8800	0.8521	0.8658	0.9467	0.7403	0.7745	0.7697	0.8843	0.4654	0.4975	0.4838	0.2966	0.4375	0.4167	**0.4062**	0.2500

E. Upenik, M. Rerabek, and T. Ebrahimi, "On the performance of objective metrics for omnidirectional visual content“. QoMEX 2017.

Goals of VR IQA NET

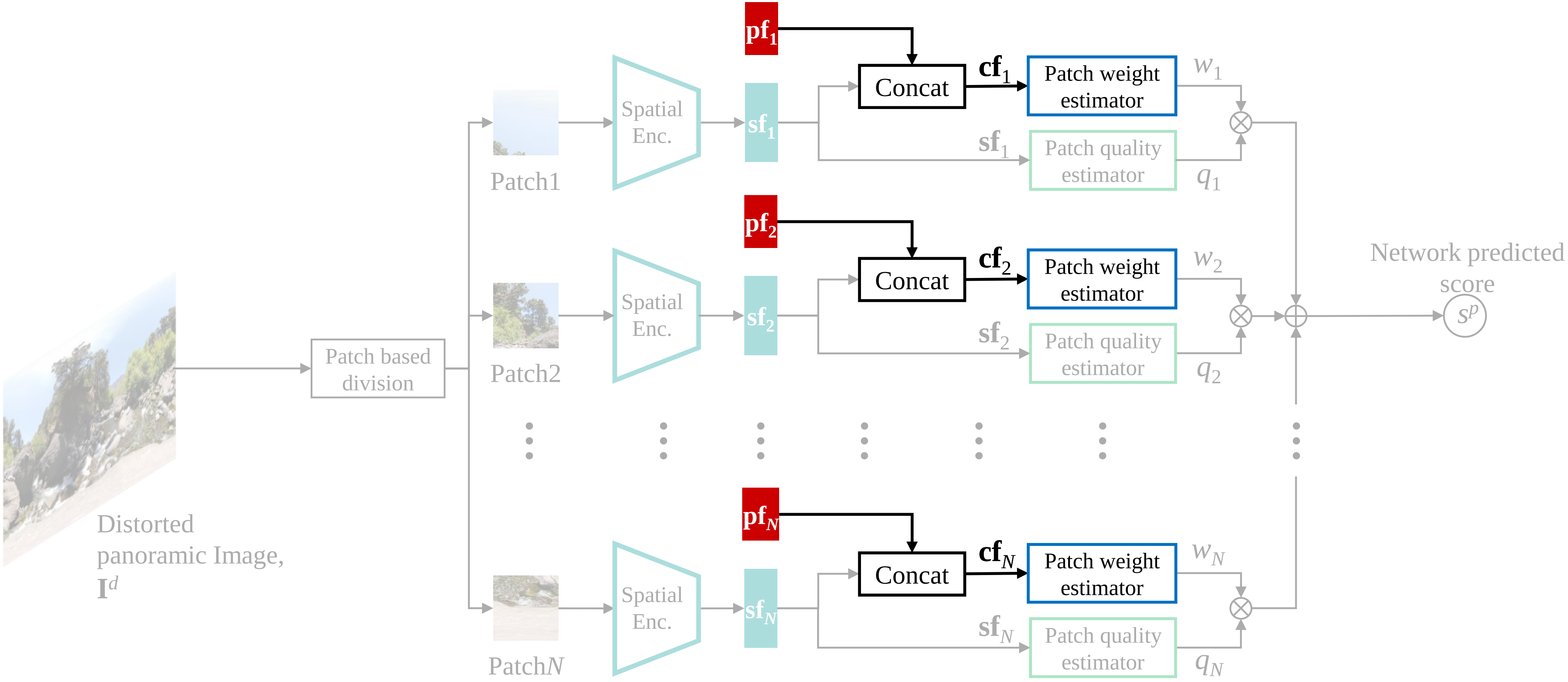
- Yield the strongly correlated results with human perception
  - ✓ Propose a deep learning-based framework
- Consider the characteristics of the VR projection
  - ✓ Adopt the patch position feature
- Enhance the no-reference performance using the reference images in training phase
  - ✓ Adopt the adversarial learning with reference images

VR IQA NET: Deep Learning-based Framework

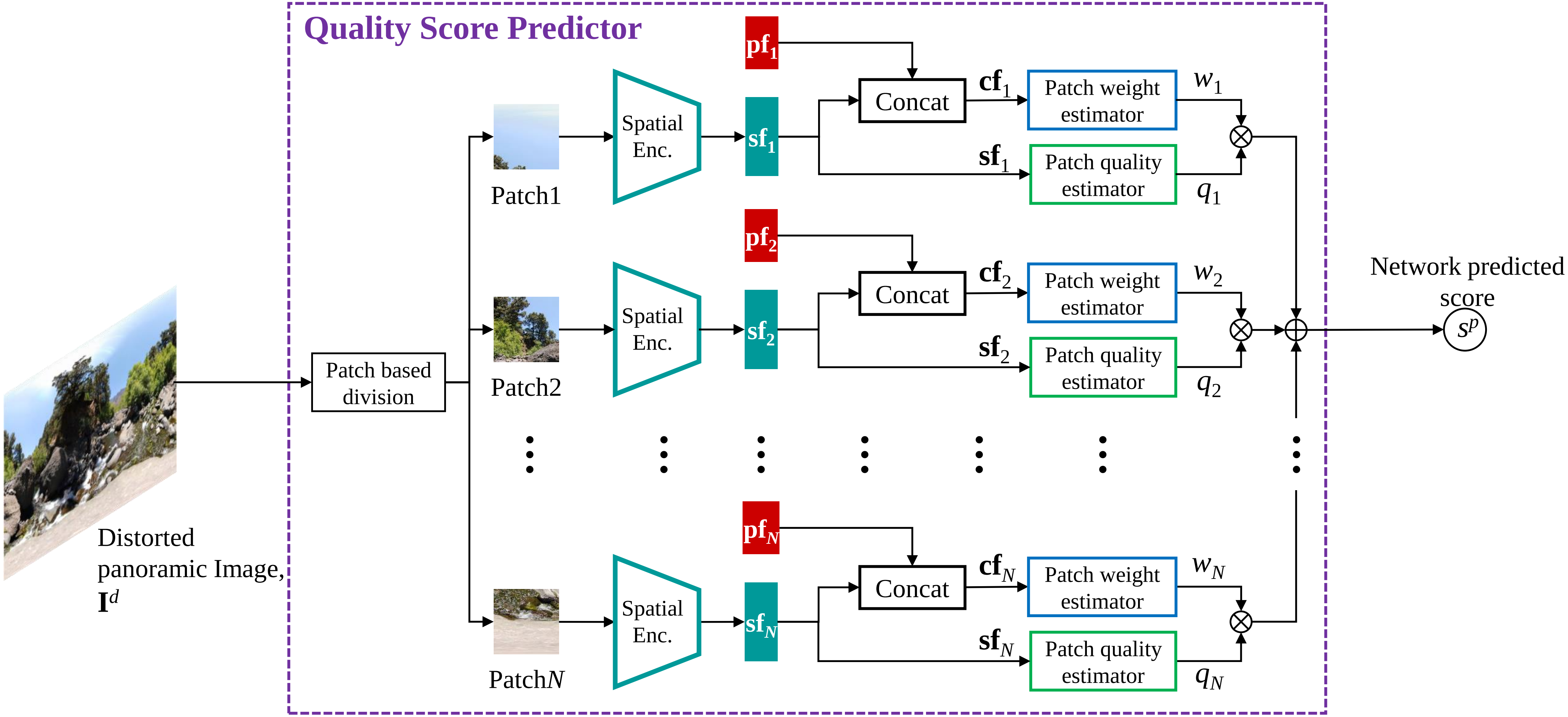


S. Bosse, D. Maniry, K. R. Müller, T. Wiegand, and W. Samek, "Neural network-based full-reference image quality assessment". PCS 2016.

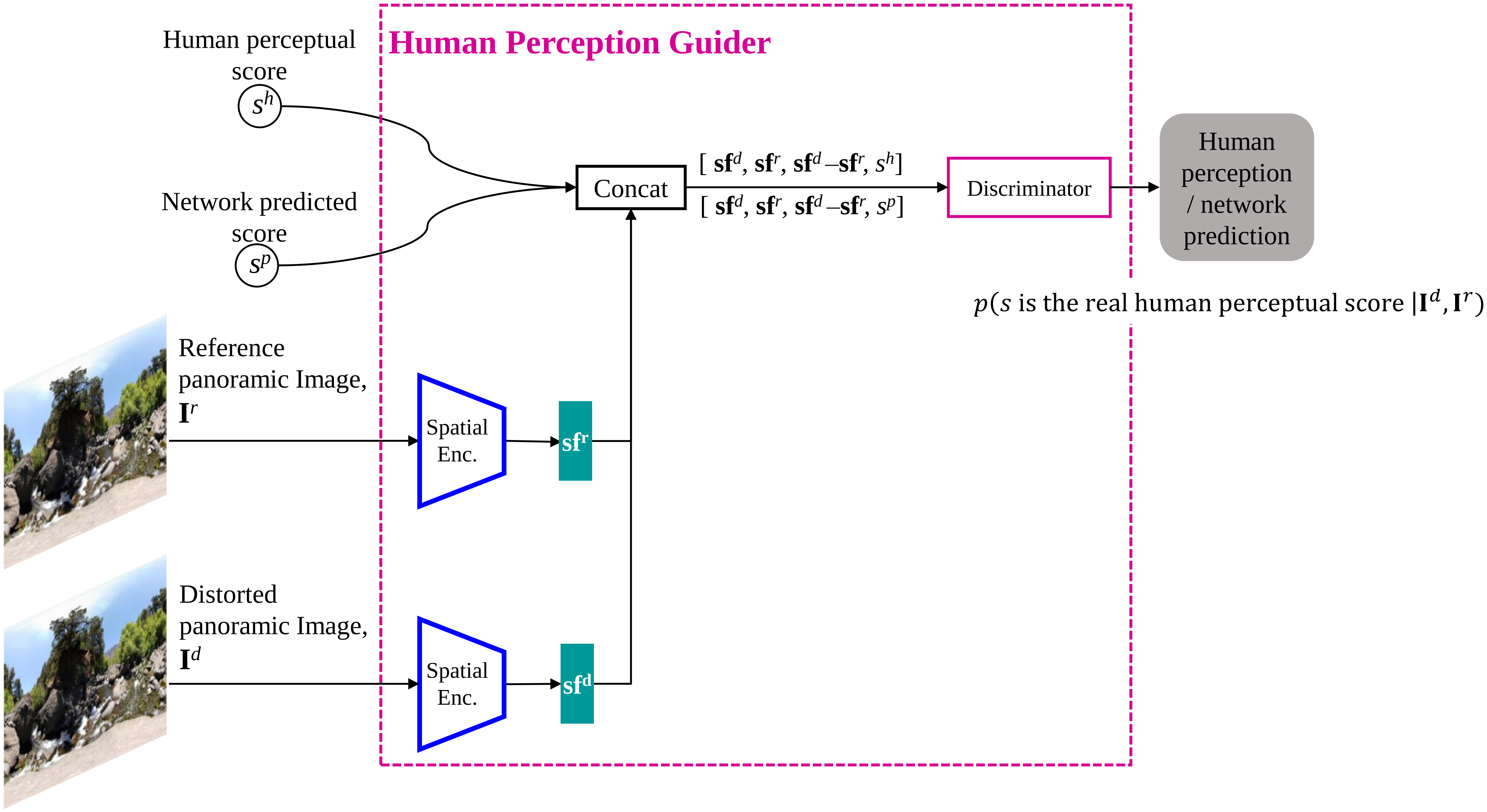
VR IQA NET: Patch Position Feature



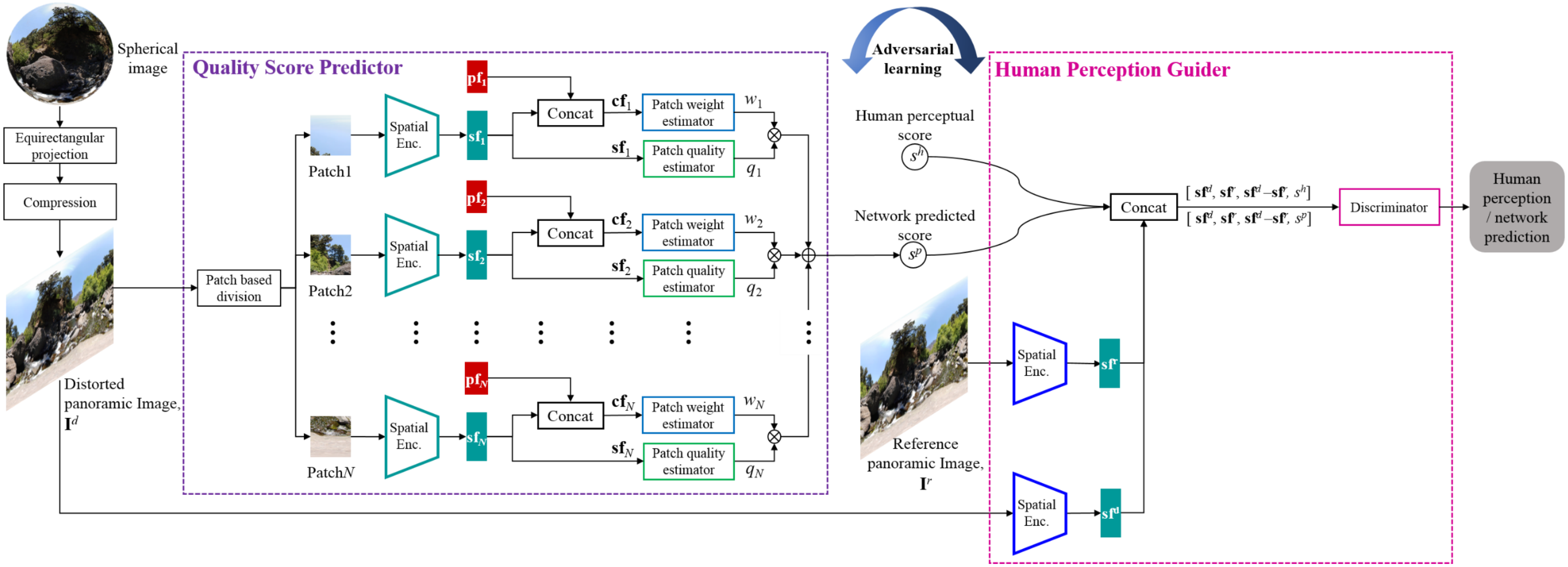
VR IQA NET: Quality Score Predictor



VR IQA NET: Human Perception Guider



VR IQA NET: Overall Framework & Adversarial Learning



$$\min_P \max_D V(P,D) = \min_P \max_D [(s^p - s^h)^2 - \lambda\{J(D(s^h|I^d,I^r),1) + J(D(s^p|I^d,I^r),0)\}$$

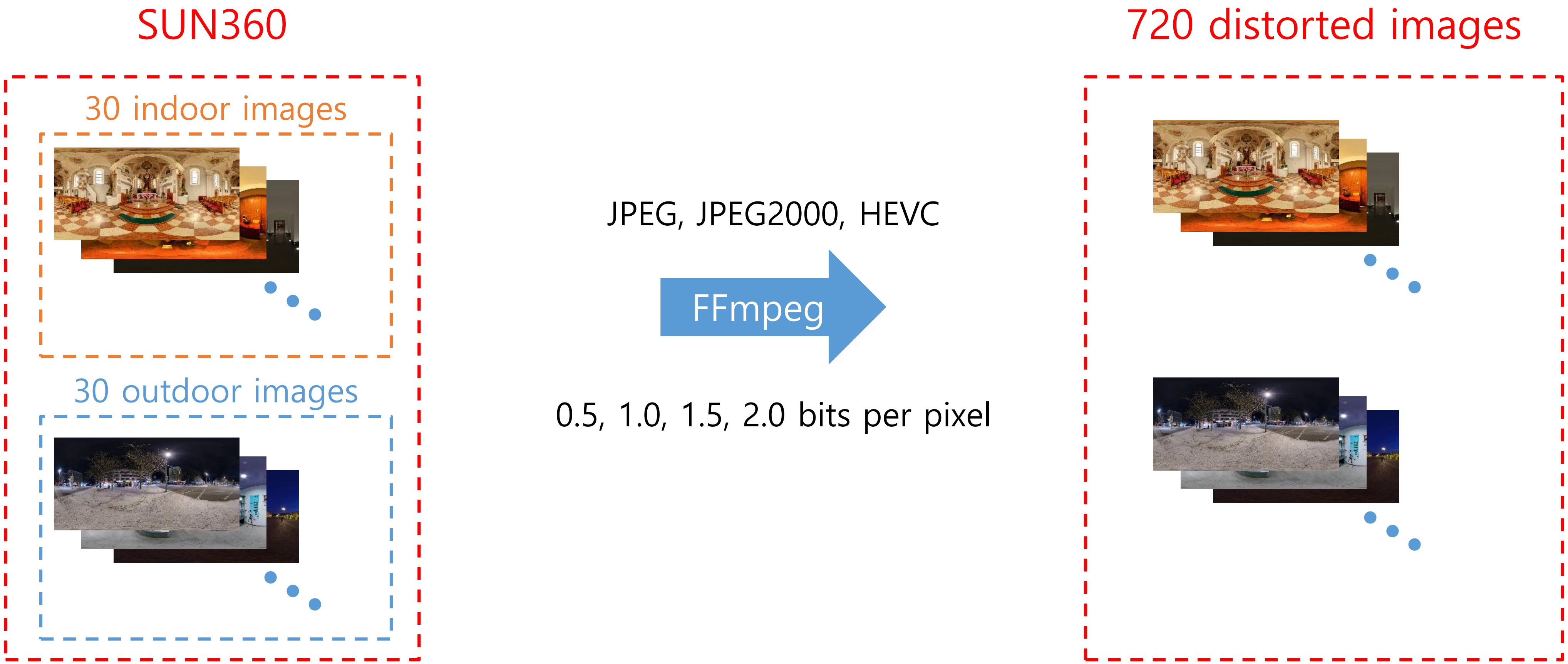
Regression loss

Adversarial learning loss

$$\min_P [(s^p - s^h)^2 + \lambda J(D(P(\mathbf{I}^d)|\mathbf{I}^d,\mathbf{I}^r),1)$$

$$\min_D J(D(s^h|\mathbf{I}^d,\mathbf{I}^r),1) + J(D(P(\mathbf{I}^d)|\mathbf{I}^d,\mathbf{I}^r),0)$$

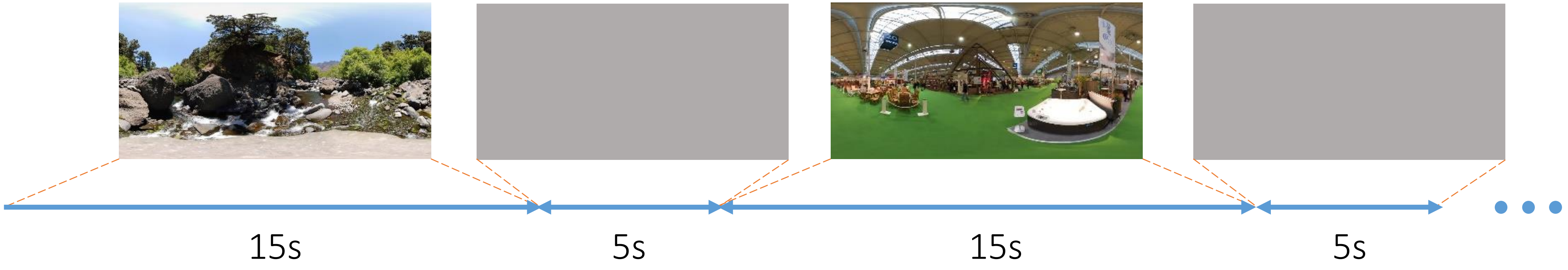
Experiments: Dataset Generation



J. Xiao, K. A. Ehinger, A. Oliva, and A. Torralba, “Recognizing Scene Viewpoint using Panoramic Place Representation”. CVPR 2012.

Experiments: Subjective Assessment Experiment

- Oculus Rift CV1 and *Oculus 360 photos*
- 15 subjects participated
- Single stimulus continuous quality evaluation (SSCQE)
- Quality score in the continuous scale range of 0-100
  - Excellent(100-80), good(80-60), fair(60-40), poor(40-20), and bad(20-0)



x 6 sessions

Results: Prediction Performances

Table 1. Prediction performances comparison

Objective metrics		PLCC	SROCC	RMSE
2-D IQA	PSNR	0.6983	0.6794	12.8791
	SSIM	0.7301	0.7259	12.2954
	MS-SSIM	0.7383	0.7516	12.1364
	VIFp	0.8124	0.7907	10.4909
VR IQA	S-PSNR	0.5316	0.5470	15.2403
	WS-PSNR	0.5270	0.5172	15.2917
	CPP-PSNR	0.5185	0.5347	15.3850
DL based VR IQA	**Proposed method without Critic**	**0.8516**	**0.8227**	**9.4313**
	Proposed method with Critic	**0.8721**	**0.8522**	**8.8048**

Visualization: The area that most affect the predicted score

HEVC with 2.0 bpp (MOS: 70.85)

Visualization: The area that most affect the predicted score

JPEG with 0.5 bpp (MOS: 7.14)

Conclusions

- We proposed a novel no-reference VR IQA NET using adversarial learning for automatically assessing the image quality of VR content.
- The proposed quality score predictor could reliably predict the quality score of the omnidirectional images by considering the spatial characteristics of the projection from sphere to rectangle domain.
- The proposed human perception guider could make the predictor to more correct by comparing the predicted score of the predictor and the human perceptual score with adversarial learning.
- Experimental results showed that the proposed VR IQA method was strongly correlated with human perception.

Thank you