# Saliency Driven Perceptual Quality Metric for Omnidirectional Visual Content

Evgeniy Upenik and Touradj Ebrahimi

evgeniy.upenik@epfl.ch

Multimedia Signal Processing Group
Ecole Polytechnique Fédérale de Lausanne

26th IEEE International Conference on
Image Processing, ICIP
Taipei, Taiwan, 25 September 2019

EPFL

- Omnidirectional visual content

- Interactivity and non-determinism of immersive multimedia

- Visual attention as statistical approach

- Objective perceptual quality metric for immersive content
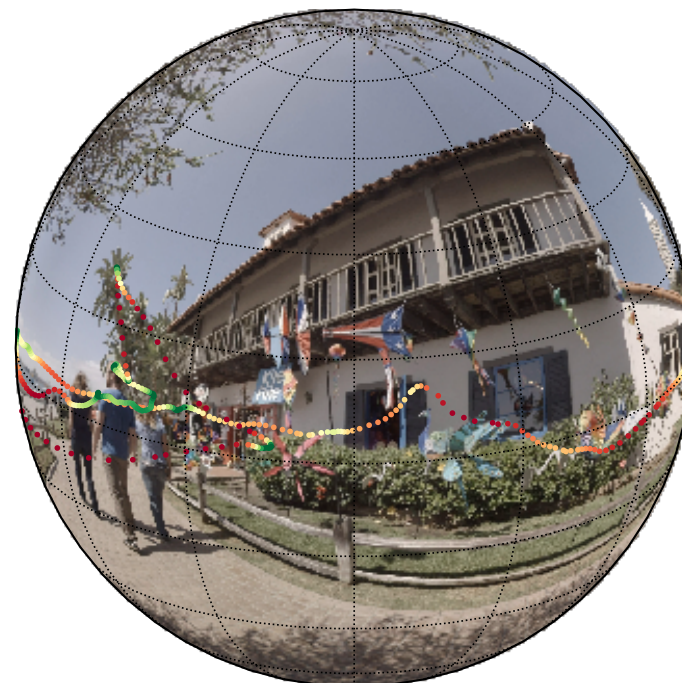
At any given moment a viewer sees only a subset of the entire omnidirectional image. Thus, there may exist a case when an observer does not explore every part of an image. Hence, **particular regions acquire more significance and provide higher impact on perceived visual quality**, whilst other regions contribute less to it.
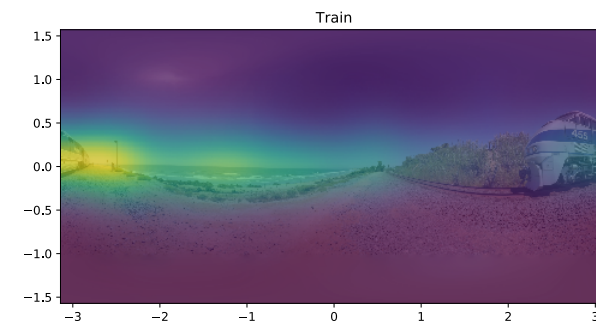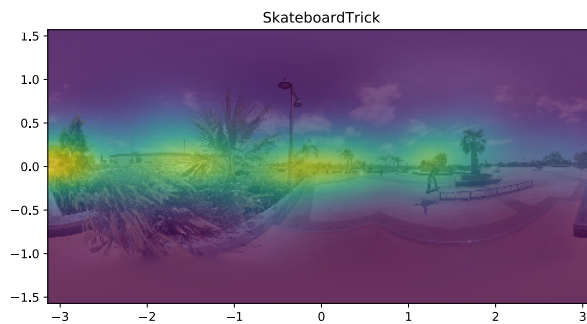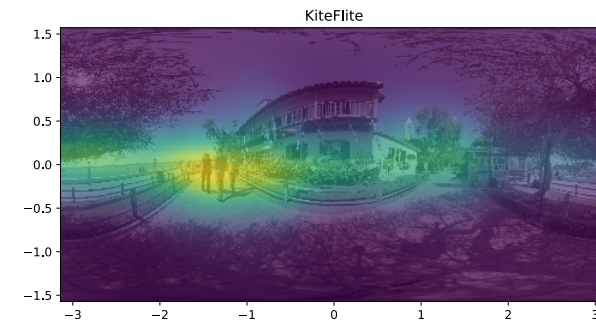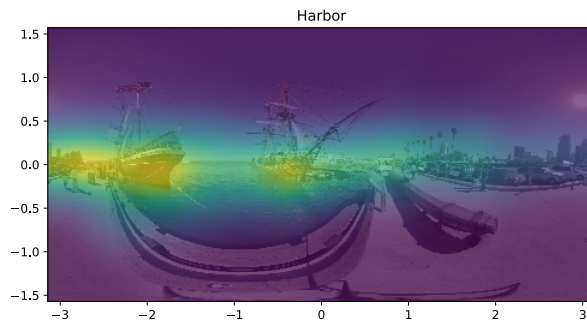
One way to take this factor into consideration is to collect statistical data of user interactions in order to estimate **visual attention** or saliency.
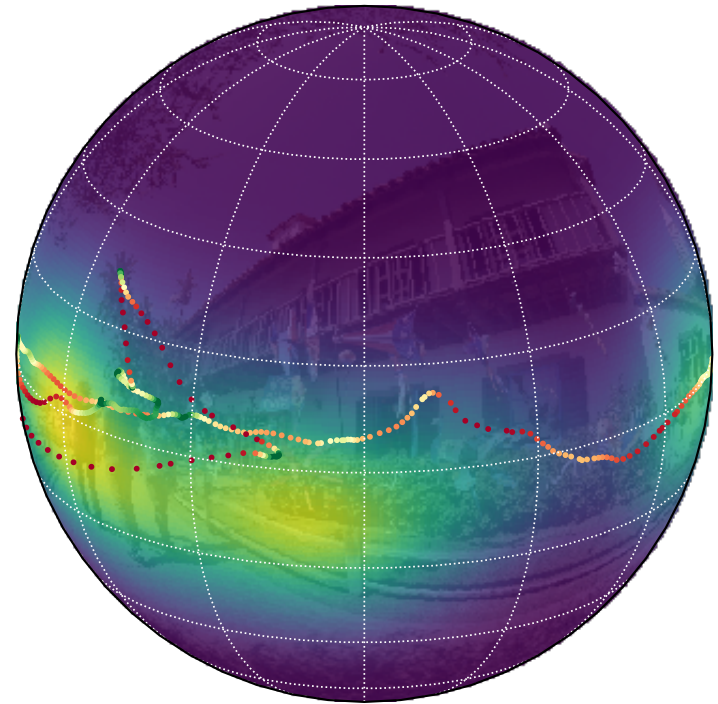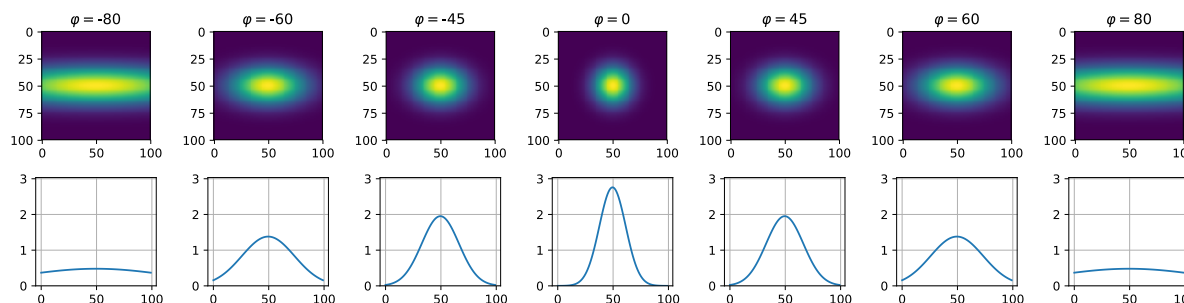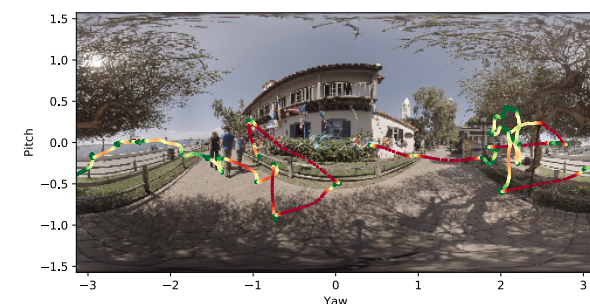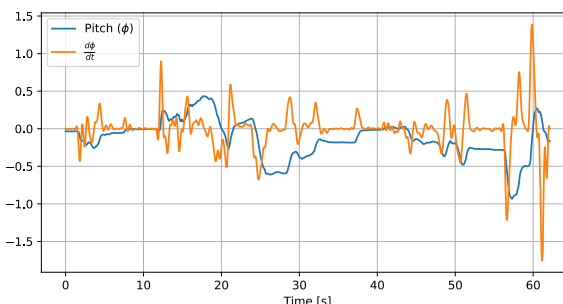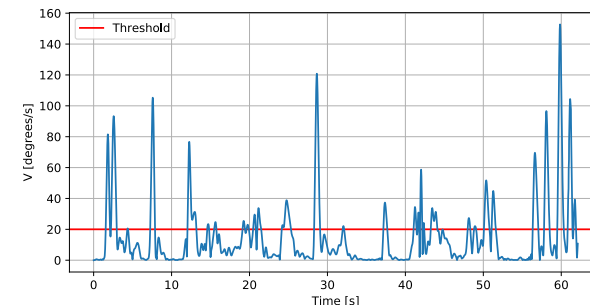
Saliency map can be defined as

$$h_{i,j} \in [0,1], i = 0, 1, ..., H, j = 0, 1, ..., W$$

where each pixel of $\boldsymbol{h_{i,j}}$ provides a visual attention value for each corresponding pixel of an image

Equirectangular projections of omnidirectional images overlaid with visual attention heatmaps
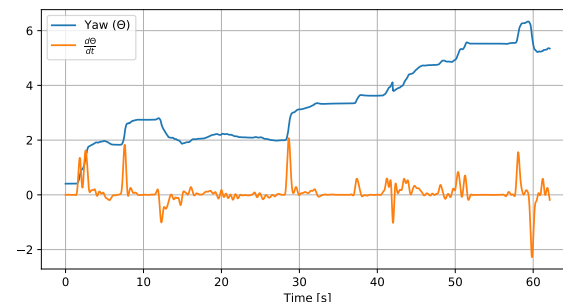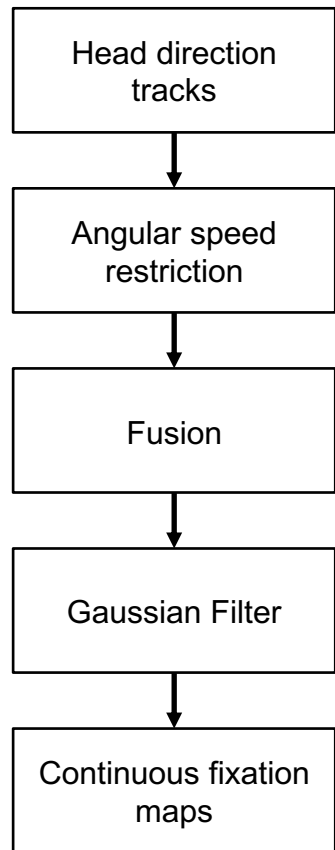
E. Upenik and T. Ebrahimi, "A simple method to obtain visual attention data in head mounted virtual reality," ICME 2017

Peak Signal-to-Noise Ratio

$$PSNR = \frac{MAX_I^2}{MSE}$$

where

$$MSE = \frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \left(I(i,j) - \hat{I}(i,j)\right)^2}{H * W}$$

PSNR is a classic objective metric
to estimate perceptual visual quality

Applying weighting to PSNR

$$h_{i,j} \in [0,1], i = 0, 1, ..., H, j = 0, 1, ..., W$$

$$MSE_{VA} = \frac{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} \left(I(i,j) - \hat{I}(i,j)\right)^2 h_{i,j}}{\sum_{i=0}^{H-1} \sum_{j=0}^{W-1} h_{i,j}}$$

$$PSNR_{VA} = \frac{MAX_I^2}{MSE_{VA}}$$

Multimedia Signal Processing Group
Ecole Polytechnique Fédérale de Lausanne

26th IEEE International Conference on
Image Processing, ICIP
Taipei, Taiwan, 25 September 2019

EPFL

Evaluation

- Absolut Category Rating with Hidden Reference (ACR-HR)
- Five-level scale: 5 - Excellent, 4 - Good, 3 - Fair, 2 - Poor, and 1 - Bad
- 19 subjects, 9 females, median age of 24.5

Exploration

- Same set up as for evaluation
- Subjects were asked to evaluate aesthetic value of the pictures and only uncompressed stimuli were used
- Subjective scores were discarded, only head direction tracks were collected
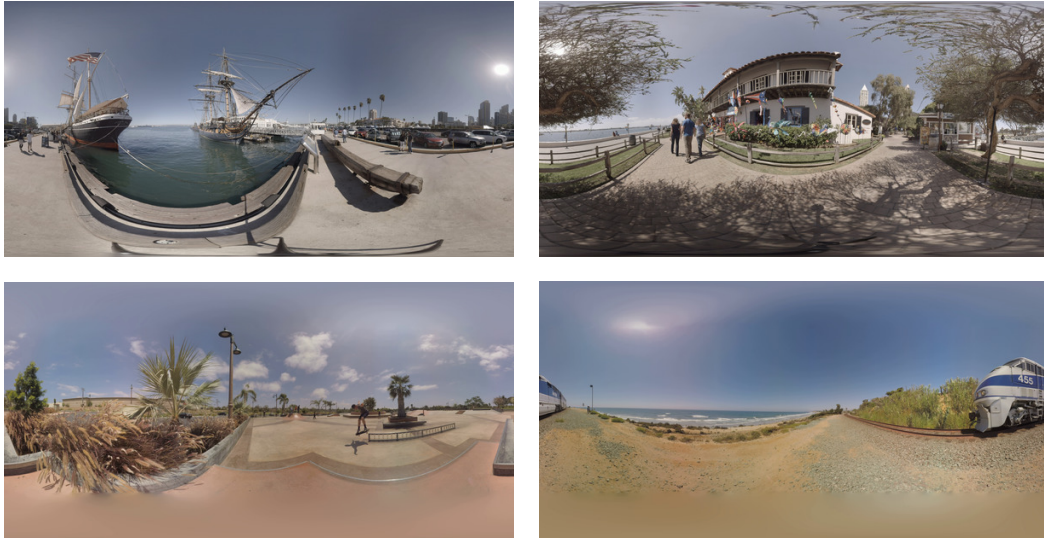- 17 participants, 10 were females, median age of 24.3

Evaluation. Time spent on a stimulus.

Mean: 18.88s
Median: 14.12s
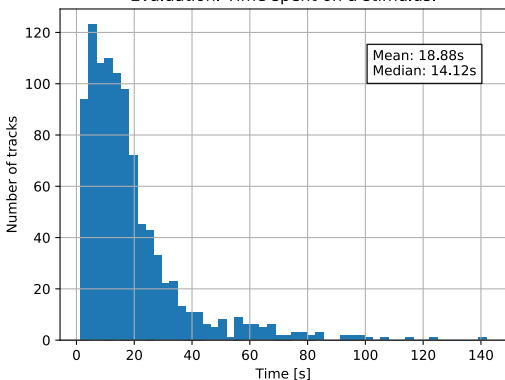
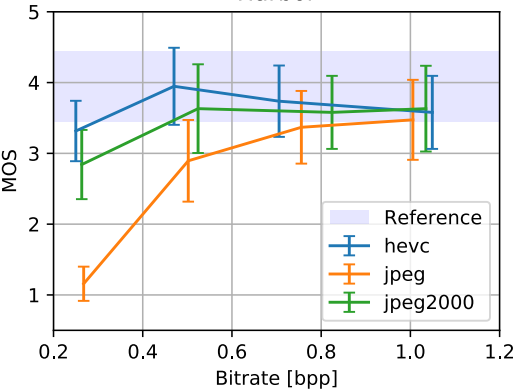Exploration. Time spent on a stimulus.
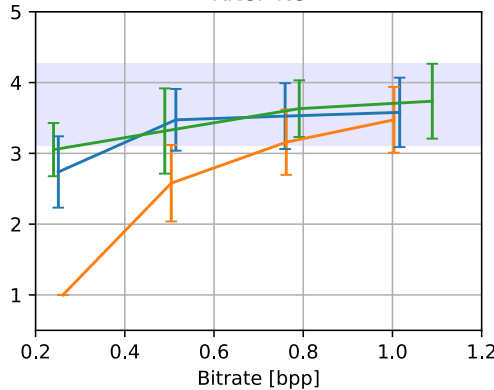
Mean: 25.28s
Median: 22.87s
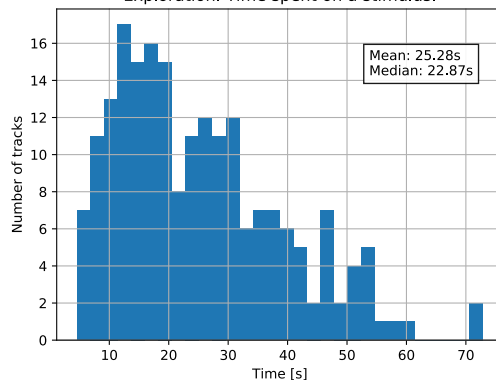
Subjective mean opinion scores (MOS) with 95% confidence intervals. The area filled with transparent purple color depicts the 95% confidence interval of the hidden reference.
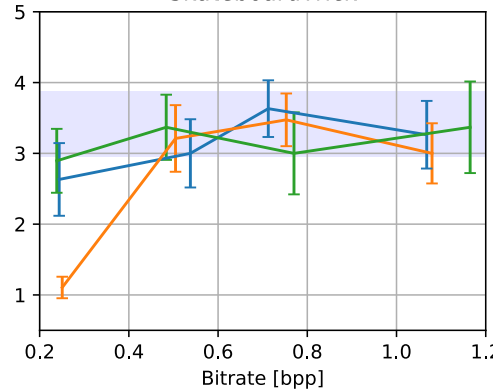
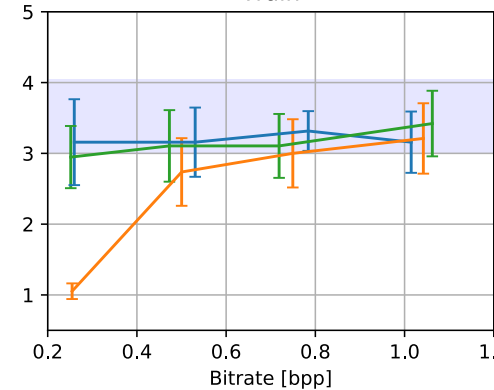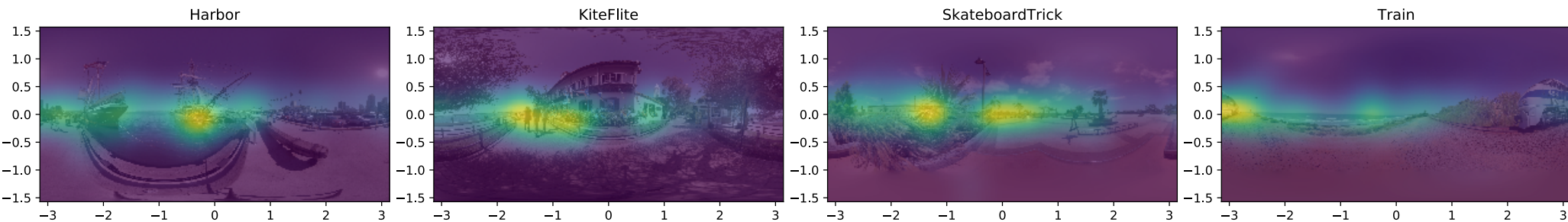- Evaluation experiment: All quality levels (Eval)

- Evaluation experiment: Stimuli rated as high quality (Eval-Refs)

- Evaluation experiment: Stimuli rated as low quality (Eval-LowQ)

- Exploration experiment (Expl)

Exploration experiment heatmaps (Expl)
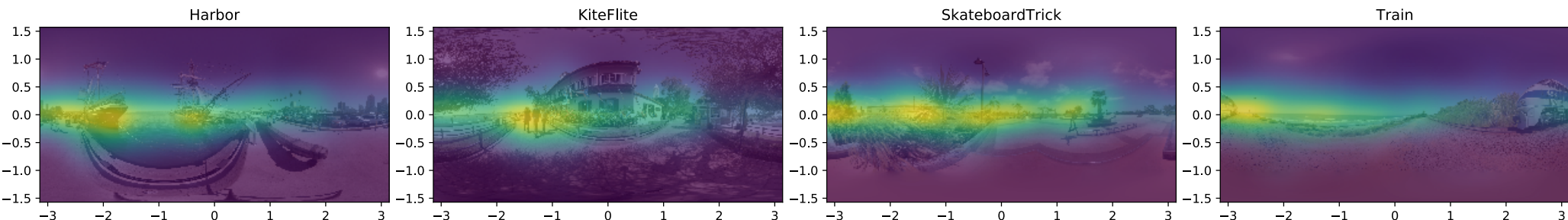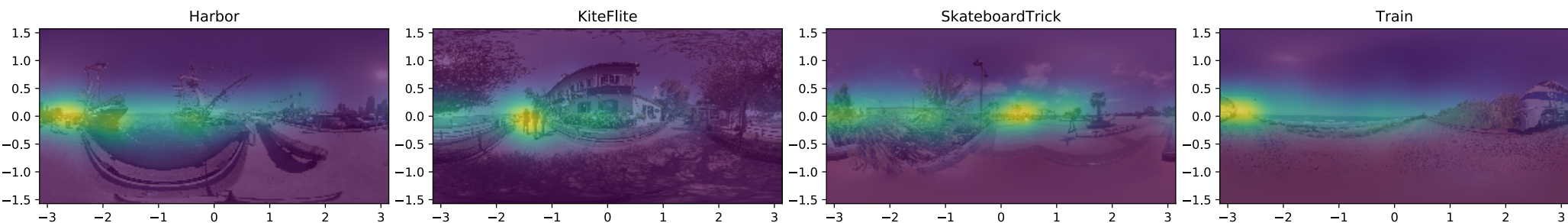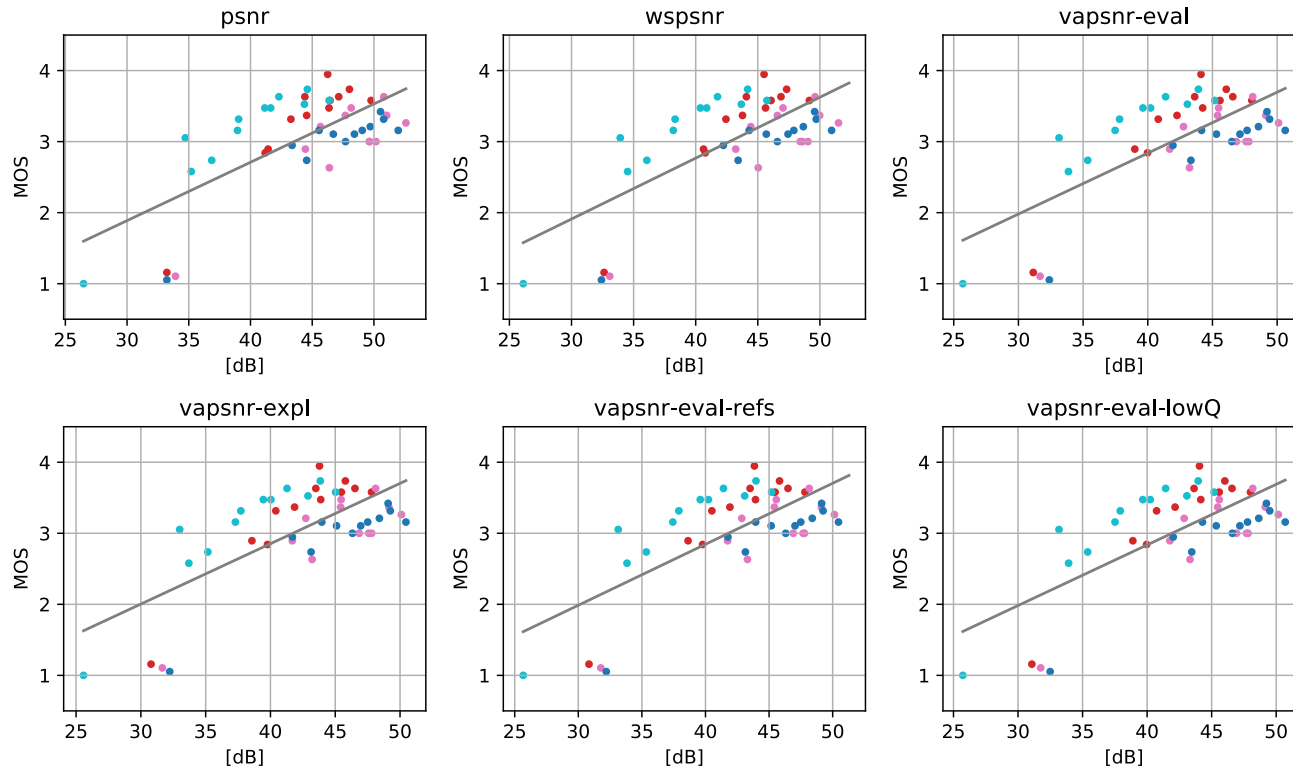
Evaluation experiment heatmaps: all quality levels (Eval)
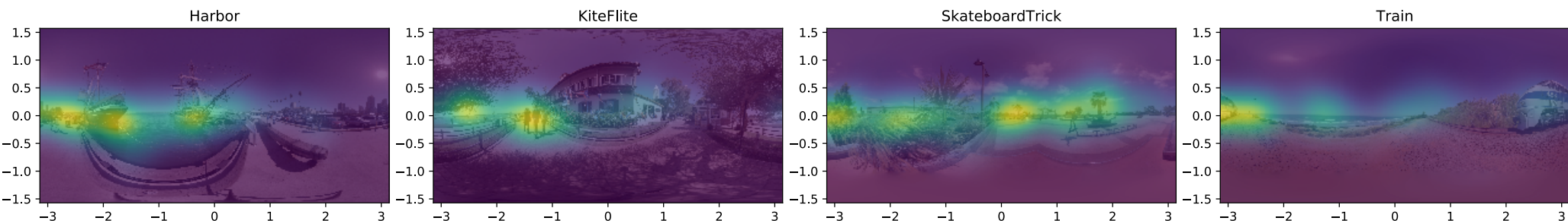
Evaluation experiment heatmaps: stimuli rated by subjects as high quality (Eval-Refs)

Evaluation experiment heatmaps: stimuli rated by subjects as low quality (Eval-LowQ)

Mapping of objective scores to subjective ratings. Grey line depicts linear fitting. Different colors represent different contents: blue - Train, red - Harbor, cyan - SkateboardTrick, magenta - KiteFlite.

Standard performance indexes: Pearson linear correlation coefficient (PLCC), Spearman rank correlation coefficient (SRCC), and Kendall rank correlation coefficient (KRCC).

	PSNR	WS-PSNR [19]	VA-PSNR Eval	VA-PSNR Expl	VA-PSNR Eval-Refs	VA-PSNR Eval-LowQ
PLCC	0.6959	0.7106	0.7107	0.7074	**0.7114**	0.7083
SRCC	0.3706	0.4131	0.4131	0.4075	**0.4163**	0.4080
KRCC	0.2706	0.2976	**0.3012**	0.2904	0.2976	0.2958

- A new method called VA-PSNR which estimates perceptual quality of omnidirectional content considering visual attention has been proposed

- The method is validated against subjective MOS and benchmarked against state-of-the-art objective metrics

- VA-PSNR shows better performance when compared to alternative approaches based on PSNR

Multimedia Signal Processing Group
Ecole Polytechnique Fédérale de Lausanne

26th IEEE International Conference on
Image Processing, ICIP
Taipei, Taiwan, 25 September 2019

EPFL

Thank you!

Evgeniy Upenik

evgeniy.upenik@epfl.ch

Multimedia Signal
Processing Group
EPFL

http://mmspg.epfl.ch/