Trained perceptual transform for quality assessment of high dynamic range images and video

Nanyang Ye, María Pérez-Ortiz, Rafał K. Mantiuk
Department of Computer Science and Technology, University of Cambridge

UNIVERSITY OF CAMBRIDGE

erc European Research Council
Established by the European Commission

Abstract

We propose a trained perceptual transform for quality assessment of high dynamic range (HDR) images and video. The transform is used to convert absolute luminance values found in HDR images into perceptually uniform units, which can be used with any standard-dynamic-range metric. The new transform is derived by fitting the parameters of a previously proposed perceptual encoding function to 4 different HDR subjective quality assessment datasets using Bayesian optimization. The new transform combined with a simple peak signal-to-noise (PSNR) ratio measure achieves better prediction performance in cross-dataset validation than existing transforms.

Motivation

There is a need for low complexity, accurate and interpretable quality metric for HDR images and video.

Previous work

1. HDR quality metrics

HDR-VDP-2 [Mantiuk *et al.* 11] – visibility and quality metric based on models of low-level vision (contrast sensitivity, masking, etc.).

HDR-VDP-2.2 [Narwaria *et al.* 15] - optimized frequency pooling weights in HDR-VDP-2 for better quality prediction.

HDR-VQM [Narwaria *et al.* 15] – video metric for HDR video, accounting for the spatial-temporal effects.

2. Perceptually Uniform (PU) transform

PU encoding [Aydın *et al.* 08] transforms the absolute luminance into perceptually uniform (PU) values related to the just-noticeable-differences in luminance.

PQ-EOTF [Miller *et al.* 12] transformed the PQ-EOTF to achieve similar goals as PU encoding but for coding.

Trained perceptual transform encoding

The application of trained perceptual transform (T-PT) for quality prediction is shown in Figure 1.

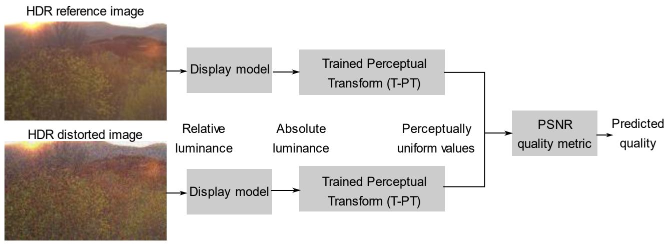


Figure 1 The application of T-PT for HDR quality assessment.

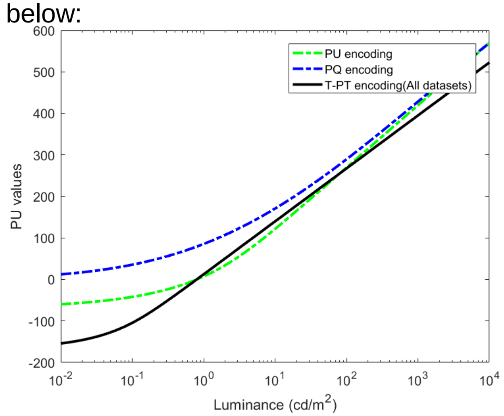
T-PT transforms absolute luminance L into perceptually uniform units. It is defined as:

$$P(L) = \int_{L_{\min}}^{L} \frac{1}{T(l)} dl \qquad T(L) = S \cdot \left(\left(\frac{C_1}{L} \right)^{C_2} + 1 \right)^{C_3},$$

where L_{min} is the minimum luminance to be encoded. The detection threshold $T(L)$ models the maximum difference in luminance that is just invisible. C_1, C_2 and C_3 are parameters to train.

Bayesian Optimization of parameters

We use Bayesian optimization to find C_1, C_2 and C_3. For training, we use T-PT-PSNR and SROCC as the loss. The optimized values are: $C_1 = 0.14249$, $C_2 = 2.192$, $C_3 = 0.30499$. The T-PT encoding function is shown below:



Experiment results

Four training and evaluation datasets: *Nawaria2013*, *Nawaria2014*, *Korsunov2015*, *Emin2017*.

1. Cross-dataset evaluation:

To avoid over-fitting, we train T-PT-PNSR on 3 datasets and leave one dataset out for cross-dataset evaluation. The evaluation criteria is SROCC.

Test dataset	T-PT-PSNR	PQ-PSNR	PU-PSNR	HDR-VDP2.2	HDR-VQM
#1 Narwaria2013[8]	**0.6024**	0.58478	0.5898	*0.8911*	*0.8874*
#2 Narwaria2014[9]	**0.4887**	0.38043	0.3605	*0.5727*	*0.8126*
#3 Korsunov2015[11]	**0.8908**	0.8751	0.8833	*0.9503*	*0.9572*
#4 Emin2017 [15]	**0.8673**	0.81347	0.8249	*0.9298*	*0.9193*

2. Results of T-PT trained on all datasets

To test the performance of the T-PT transform, we evaluate the T-PT transform trained on all datasets.

T-PT-PSNR – trained perceptual transform and PSNR as a quality metric

Test Dataset	T-PT-PSNR
#1 Narwaria2013[8]	0.6186
#2 Narwaria2014[9]	0.5230
#3 Korsunov2015[11]	0.8906
#4 Emin2017 [15]	0.8669

T-PT-SSIM – trained perceptual transform and SSIM as a quality metric

To test the generalization ability of T-PT, we test the performance of T-PT-SSIM with T-PT trained with PSNR and compare it with PU-SSIM and PQ-SSIM. T-PT-SSIM also shows generally better performance than untrained PU-SSIM.

Test Dataset	T-PT-SSIM	PU-SSIM	PQ-SSIM
#1 Narwaria2013[8]	0.6838	0.6969	**0.7348**
#2 Narwaria2014[9]	0.6145	0.5149	**0.8292**
#3 Korsunov2015[11]	**0.9268**	0.9239	0.8728
#4 Emin2017 [15]	**0.8864**	0.8430	0.8022

References

Aydın *et al. Extending quality metrics to full luminance range images*. *Proceedings of HVEI, 68060B-10,* 2008.

Mantiuk *et al.* HDR-VDP-2: *A calibrated visual metric for visibility and quality predictions in all luminance conditions*. *ACM Transactions on Graphics, 30 (4),* 2011.

Miller *et al. Perceptual signal coding for more efficient usage of bit codes*. SMPTE Motion Imaging Journal, 122 (4), 2013.