EUROGRAPHICS 2026 / B. Masia and J. Thies
(Guest Editors)

COMPUTER GRAPHICS *forum*
Volume 45 (2026), Number 2

ML-PEA: Machine Learning-Based Perceptual Algorithms for Display Power Optimization

Kenneth Chen[1], Nathan Matsuda[2], Thomas Wan[2], Ajit Ninan[2], Alexandre Chapiro[2], Qi Sun[1]

[1]New York University, USA
[2]Meta Reality Labs, USA

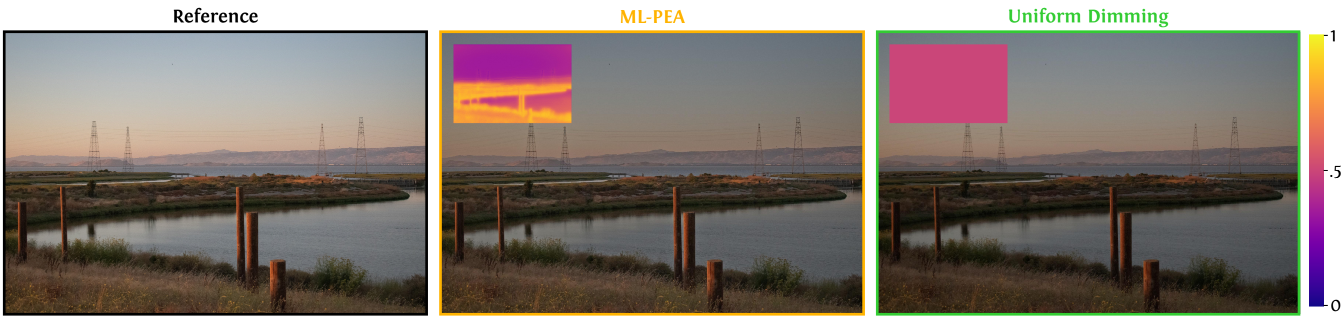

Figure 1: *Our pipeline generates images which consume less power than the original when shown on a display, while minimizing perceptual impact. Here, we show an example of an image generated with our technique (●) compared to the reference (●) and its uniformly dimmed version (●). The corresponding dimming maps are shown in the insets, with the multiplicative scaling factor presented in the color bar on the right. Note that both the uniformly dimmed image and the image generated with our technique in this figure consume the same amount of display power: 52.1% of the reference.*

Abstract

Image processing techniques can be used to modulate the pixel intensities of an image to reduce the power consumption of the display device. A simple example of this consists of uniformly dimming the entire image. Such algorithms should strive to minimize the impact on image quality while maximizing power savings. Techniques based on heuristics or human perception have been proposed, both for traditional flat panel displays and modern display modalities such as virtual and augmented reality (VR/AR). In this paper, we focus on developing and evaluating display power-saving techniques that use machine learning (ML) in VR displays. We developed a U-Net-based technique paired with perceptual and power optimization loss functions that generates spatially varying dimming maps. These dimming maps are used to modulate input images, per-pixel, to generate a power-efficient image. Our pipeline was validated via quantitative analysis using image quality metrics and through a subjective study. Our subjective validation provides results scaled in perceptual just-objectionable-difference (JOD) units. This data, when rescaled, allows for comparisons of our technique with recent studies on VR display power optimization. Our results show that participants prefer our technique over a uniform dimming baseline for high target power saving conditions. This model and study serve as a template and baseline for future applications of deep learning to display power optimization. Model training code and data can be found at kenchen10.github.io/projects/mlpea/index.html.

CCS Concepts

• ***Computing methodologies*** → ***Virtual reality; Mixed / augmented reality; Perception; Machine learning;***

1. Introduction

The power requirements of display devices are ever-increasing due to higher resolutions, greater peak luminances, and more. Displays are very power-hungry, and can account for up to 40% of a device's entire power budget [TOMB13; ATS*11]. This issue is magnified for novel display modalities like VR/AR, which typically consist of displays within a compact head-mounted form factor, and are untethered from a constant power supply. And, while a wide body of work in the machine learning community has studied low-power and efficient neural networks [HTH22; CGZH20; WCB*18], little work on how machine learning can be applied to optimize the power efficiency of media devices has been explored.

Power saving image processing techniques are often lossy operations and introduce a tradeoff in image quality, for example by reducing a display's peak luminance. Imaging algorithms that optimize power while minimizing perceptual impact are critical to maintaining acceptable power usage with high visual quality in such devices, while also reducing the carbon footprint of display devices. Many proposals have been made for such methods. We describe a number of these in Section 2, and they range from simple dimming-based methods to complex gaze-contingent color modulation algorithms [DCT*22; CDU*23].

In this work, we focus on employing machine learning (ML) techniques to achieve this goal, specifically for a VR scenario. In Section 2, we discuss existing works that use ML for power savings, but find that they are insufficiently described to replicate, and limited in terms of validation. To address this lack, we employ the framework introduced in the recent PEA-PODs work [CWM*24], where a study was conducted to measure perceptual impact of a number of VR display power optimization techniques on the same unified perceptual scale. We conduct a study comparing our **m**achine **l**earning-based **pe**rceptual power saving **a**lgorithm (**ML-PEA**) against a uniform dimming baseline, with results scaled in a perceptual just-objectionable-difference (JOD) scale. Finally, we compare ML-PEA's performance in the context of the previously measured power saving techniques from PEA-PODs by re-scaling JOD scores, demonstrating competitive performance that beats all other methods for key applications. In summary, our main contributions include

- an ML technique that produces images that reduce VR display power consumption, while maintaining visual quality,
- a user study to assess it, and
- analysis of results in context of state-of-the-art power saving algorithms, showing performance improvements for key use cases.

2. Background & Related Work

In order to develop a display power optimization algorithm, it is important to understand how power consumption varies depending on the target display. The power consumed by a display device depends on both the pixel intensity distribution of the content being displayed and the device's underlying display architecture. This means that the power profile of the same image shown on one display architecture can be very different from another. Display power models are used to estimate the power consumption of an image. We briefly summarize common display types and their respective power models here.

Two classes of flat panel display are common in today's commercial landscape; LC and organic light-emitting diode (OLED) displays. LC displays consist of an array of LEDs, or backlight unit (BLU), below a color filter array and diffusers. In an OLED display, each pixel typically consists of three LEDs, one for each color primary. The power consumed by any display is typically dominated by the devices that emit light. As such, the power consumed by an LED display can be modeled as a function of the sum of pixel intensities, weighted differently for each primary,

$$\mathcal{P}(\mathcal{I}) = \sum_{\mathbf{c}\in\mathcal{I}} w_R\mathbf{c}_R + w_G\mathbf{c}_G + w_B\mathbf{c}_B, \tag{1}$$

where $\mathbf{c}$ is a single pixel with red, green, and blue component LEDs. The weights, $w_{(\cdot)}$, can be modeled or determined through experimental measurements of display power [DZ12; DCT*22]. In an LC display, power is determined by LEDs in the BLU, which is typically of much lower resolution than the image itself,

$$\mathcal{P}(\mathcal{I}) = \frac{1}{N}\sum_{d\in B(\mathcal{I})} d, \tag{2}$$

where $B(\cdot)$ is the BLU array which is a function of the displayed image and d is the intensity of individual LEDs in the BLU. N is the number of LEDs in the BLU (notably, all BLU LEDs in a global dimming display have the same intensity). A detail that is omitted here is the way in which the BLU intensities are determined, given the input image. This is typically simple to determine for an LC display with *global dimming*, whereby all BLU LED intensities are set to the maximum pixel value in the image,

$$d = \max(\mathbf{c}_R, \mathbf{c}_G, \mathbf{c}_B | \mathbf{c}\in\mathcal{I}). \tag{3}$$

In a *local dimming* scheme, some image processing is usually employed, with consideration of the display's point spread function and BLU LED arrangement [THW*07]. Power consumption predictions for a local dimming display require knowledge of the display device's unique processing algorithm. Of note, however, is that power consumption in LC displays depends entirely on the intensity of BLU LEDs, whereas the power consumption in an LED-based display depends on the sum of RGB pixel intensities. As such, display power optimization methods which modulate color typically only work for LED-based displays. In this work, we focus specifically on optimizing power for OLED displays, using the model in Equation (1). We note that in prior works, a constant offset value, δ, is added to the expected display power to account for the circuit [CWM*24]. Because this value does not vary with image content, we discard it here. Furthermore, our work only considers power consumed by the display; we do not consider power consumed by components such as the CPU.

Heuristic Techniques The simplest of techniques for display power reduction is perhaps uniform dimming, where all pixel values are scaled by the same factor. Uniform dimming is implemented on many commercial laptops and smartphones and is a surprisingly effective technique [CWM*24]. Shye et al. [SSM09] found that gradual dimming can improve user satisfaction, Gatti et al. [GABR02] dim the display more in dark compared to bright ambient conditions, Park et al. [PS16] reduce the luminance channel of videos, and Choi et al. [CSC02] compensate for backlight dimming by letting more light through the display's color filter array. [KD06] introduced a clipping curve, which decreases highlights and preserves mid-tones. A number of works also modulate color in OLED displays [DZ11; DZ12]. These methods, however, do not adapt to the specific image content being displayed.

Perception-Guided Approaches A number of previous works have proposed techniques that take advantage of human limitations in luminance or color perception. For example, Yan et al. [YSLX18] dim the display over time based on a temporal brightness adaptation model. For wide field of view displays like VR and AR, Kim et al. [KL20] developed an eye-tracked peripheral dimming technique and Duinkharjav et al. [DCT*22] proposed a peripheral color modulation technique for OLED displays. Surace et al. [SCD25] developed a temporal uniform dimming for VR that attempts to minimize detail loss. Recent work has explored power optimizations in high dynamic range (HDR) VR via studies of human contrast and luminance preferences [CMM*25]. Despite taking into account human perception, these techniques do not consider the spatial distribution of the displayed content.

Optimization-Based Methods Another line of work optimizes an objective to produce power-efficient imagery [SDL24]. For example, Hadizadeh et al. [Had17] used a saliency model as an objective to optimize display power and Lee et al. [LLLK12] used a histogram-equalizing term as an objective. In this work, we focus on machine learning techniques for display power optimization. Most of these works train a neural network that takes as input an image and outputs a *dimming map*, which when applied to the input image results in an output that requires less display power to show [AR22; LDB23; ITU24]. Techniques have been proposed that can invert the power-optimized image to recover the original [MD23]. Ameur et al. [ADMM25] include a display power-aware block in an encoding pipeline. However, we found that the quality gains compared to a simple uniform dimming baseline were minimal, and that validation criteria could be improved. For example, Le Meur et al. [LDB23] found that, at most, their technique improved upon uniform dimming by 0.05 SSIM score and at worst was the same as uniform dimming in terms of peak signal-to-noise ration (PSNR).

Evaluation Strategies The way in which the previously-described techniques are evaluated is not standardized. This lack of standardization leads to difficulty in comparing different techniques, especially when considering different display architectures. For example, Kerofsky et al. [KD06] used metrics to validate their highlight clipping curve, Dong et al. [DZ11] conducted user surveys to evaluate their UI color modulation, Kim et al. [KL20] ran a psychophysical study to measure thresholds for their peripheral dimming, and Wee et al. [WCB18] conducted a task-based study for several techniques. Previous ML-based techniques for display power optimization were evaluated by comparing the average of metric results (e.g. PSNR and SSIM, which may not correlate well with subjective judgments) over a number of test images. None of the prior ML works, from what we can find, have evaluated their techniques in a subjective study. This makes drawing comparisons between methods challenging. The PEA-PODs project [CWM*24] provides a unified framework for subjective measurement of the perceptual impact of power saving techniques on a single perceptual JOD scale. In this work, we conducted the first subjective study to evaluate a deep learning power saving technique using this framework, and then rescaled our results to those of PEA-PODs, allowing us to make comparisons with other methods on the same JOD scale.

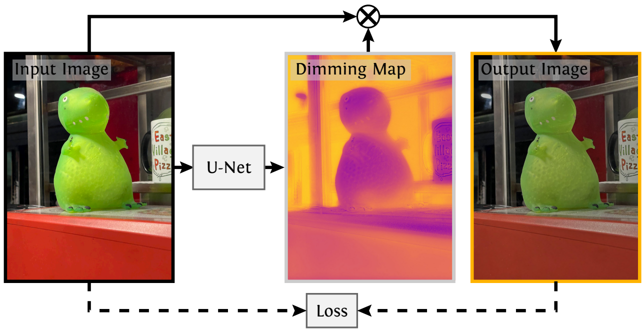

Figure 2: *Here, we show our pipeline for generating power-optimized images. First, an input image is passed to a model to generate a per-pixel dimming map. The input image is then multiplied by the dimming map to create the output. A loss between the input and output is computed to optimize the network weights.*

3. Display Power Optimization Algorithm

Here, we define a simple paradigm that can serve as a baseline for future ML-based power optimization algorithms for VR displays. An input image to be displayed, $\mathcal{I}$, is modified by a neural network, $\mathcal{M}$, to produce a more power-efficient image, $\mathcal{I}^*$. The neural network's output is a per-pixel dimming map $\mathcal{D} = \mathcal{M}(I)$, which is applied to the input image $\mathcal{I}$ using some element-wise operator, f, to produce a power-optimized output image,

$$\mathcal{I}^* = f(\mathcal{I}, \mathcal{D}). \tag{4}$$

A visualization of our pipeline is shown in Figure 2; notably, the dimming map $\mathcal{D}$ consists of a single-channel. Alternative schemes, such as a 3-channel dimming map or using different operations f led to worse performance, including significant chromatic distortions and halo artifacts (see Appendix Section 9.2).

3.1. Loss Functions

The model is optimized with a combination of perceptual and display power loss functions,

$$\mathcal{I}^* = \arg\min_{\mathcal{I}^*} \mathcal{L}(\mathcal{I}, \mathcal{I}^*) + \mathcal{L}_{\mathcal{P}}(\mathcal{I}, \mathcal{I}^*, \alpha). \tag{5}$$

This is inherently a self-supervised scheme, in which there is no ground truth power-efficient image. Instead, the quality of generated images is a function of the losses used to train the neural network. As such, loss functions that align with human perception should be used to produce images that human users prefer. In this work, we found that a combination of VGG [JAF16] and SSIM [WBSS04] losses works well for this task,

$$\mathcal{L}(\mathcal{I}, \mathcal{I}^*) = \lambda_{\text{VGG}} \cdot \mathcal{L}_{\text{VGG}}(\mathcal{I}, \mathcal{I}^*) + \lambda_{\text{SSIM}} \cdot \mathcal{L}_{\text{SSIM}}(\mathcal{I}, \mathcal{I}^*), \tag{6}$$

where $\lambda_{(\cdot)}$ are weights on each loss term. In this work, we use $\lambda_{\text{VGG}} = 0.5$ and $\lambda_{\text{SSIM}} = 5.0$. See Section 4.2 for an ablation on these weights. The $\mathcal{L}_{\text{VGG}}$ is a perceptual loss function based on the features extracted from a pre-trained VGG-19 network, and $\mathcal{L}_{\text{SSIM}}$ is a perceptual metric that takes into account structure, luminance

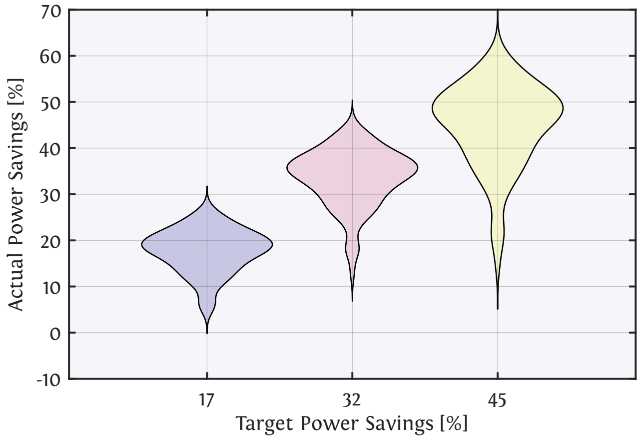

Figure 3: *We plot target vs. actual power savings as violin plots. x-axis is target and y-axis is actual power savings from ML-PEA.*

and contrast. Similar to prior art, we included a power loss function,

$$\mathcal{L}_{\mathcal{P}}(\mathcal{I},\mathcal{I}^*,\alpha) = \lambda_{\mathcal{P}} \cdot \left(\mathcal{P}(\mathcal{I}^*) - \alpha \cdot \mathcal{P}(\mathcal{I})\right)^2, \tag{7}$$

that computes the difference between the output image's power consumption and that of the input image, scaled by α. We set $\lambda_{\mathcal{P}} = 50.0$ in our experiments. Here, α is used to control the target power saving rate and is set to values of $\alpha = \{0.55, 0.68, 0.83\}$ in our quantitative evaluation and user study. Our work optimized the power consumption for an OLED display, and the power model in this loss, $\mathcal{P}(\cdot)$, is the OLED model from Duinkharjav et al. [DCT*22] which has the same form as Equation (1). As such, our model works best for an OLED display with per-primary weights similar to this display. We show in Figure 8 that this is the case, where the power savings for a different display architecture are lower. A new model would have to be trained for a new display modality, using the appropriate power model and loss function.

3.2. Model Architecture and Training

We train a U-Net architecture [RFB15], which has not been used in prior works for the display power optimization task. The U-Net has the same structure as described in the original paper from Ronneberger et al. [RFB15]. During training, the input image, $\mathcal{I}$, to the U-Net model is of resolution $256 \times 256 \times 3$, which are randomly cropped from the ground truth images. A sigmoid function is applied to the output dimming map to constrain the output to the range $[0,1]$. This means that the output pixel intensities are necessarily less than or equal to those of the original, because they are multiplied by a factor between 0–1.

The U-Net is trained on the DIV-2K dataset [AT17], which consists of 800 training images and 100 test images. The dataset contains diverse, high-quality photographs, and are 2K resolution (at least one axis, height or width, is greater than 2K pixels). We trained models for 60 epochs, and used the Adam optimizer [KB14] with a learning rate of 2e-4. We selected a batch size of 1, the same as in Ronneberger et al. [RFB15], to allow for larger input images during training and to maximize GPU usage. All networks were trained on an HPC cluster using a single NVIDIA V100 GPU.

Remark We note that our primary goal was to *evaluate* ML techniques for display power optimization, and in doing so have defined a simple pipeline that acts as a baseline for this task and in addition performs well compared to prior art (see Results Section 4).

Table 1: *Average scores are tabulated for both methods, at the three target power saving rates studied. We omit actual power saved for uniform dimming because it is matched to ML-PEA.*

Target	Method	PSNR (dB) ↑	SSIM ↑	CVVDP (JOD) ↑	Saved
17%	Uniform Dim	28.22	0.99	9.99	-
	ML-PEA	27.94	0.99	9.93	17.72%
32%	Uniform Dim	21.96	0.96	9.97	-
	ML-PEA	21.82	0.98	9.81	33.61%
45%	Uniform Dim	19.08	0.93	9.92	-
	ML-PEA	18.81	0.96	9.55	44.91%

4. Results

First, we show evaluations of ML-PEA via metrics, as well as several qualitative visualizations of ML-PEA. Note that, unless otherwise stated, power savings are computed using the OLED display power model from Duinkharjav et al. [DCT*22], which is the same model used to train the network. Furthermore, while images during training are randomly cropped (via PyTorch's RandomCrop), all inference is done on full-resolution images. Structure is preserved due to U-Net's skip connections that help recover spatial detail that may have been lost during pooling, as can be seen in example outputs. This also helps to avoid the vanishing gradient problem.

4.1. Quantitative Evaluation

We evaluated the performance, plotted in Appendix Figure 8, of our ML-PEA pipeline on the generated images. Results on the DIV2k test set are displayed for three metrics: PSNR, SSIM [WBSS04], and ColorVideoVDP [MHA*24]. These metrics are run on images generated by ML-PEA at three target power saving rates, 17%, 32%, and 45% (the power saving rates used in our subjective study, described in Sec. 5), as well as uniformly dimmed images at a matched power saving rate. Refer to Table 1 for a tabulated version of these results and Appendix Section 8 for additional discussion.

Note that even though we trained ML-PEA on three unique target power saving rates, the model's output does not match these targets exactly. We show the extent of this in Figure 3, where individual violin plots represent the distribution of power savings for ML-PEA run on the 100 test images in the DIV2k dataset for a specific target power saving rate, α. From this plot, it seems that, as α increases, the variation in actual power savings increases. We hypothesize that this could be due to the fact that metrics perform worse for high power saving rates, and so the tradeoff is poor performance on the power optimization task. A potential solution to this could be to increase the weight on the power loss for higher target power rates. The output also depends in some way on image statistics (see Appendix Section 11.1).

The distribution of scores is very different across the three metrics. PSNR between uniform dimming and ML-PEA are nearly

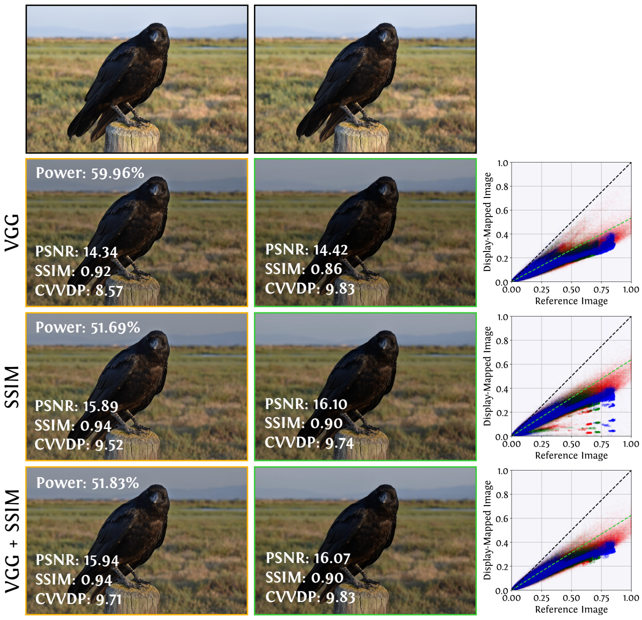

Figure 4: *We show ML-PEA (orange) and uniform dimming (green) images for different loss functions, all of which have a 45% target power saving rate. Plots show linear RGB values for the reference versus the ML-PEA image for each color channel. The identity is plotted as a black dash and the uniformly-dimmed image as a green dash. Reference images are shown at top.*

Figure 5: *Example input images outside of the training dataset are shown here. We show the reference (●), uniform dimming (●), and ML-PEA (●) images as well as the dimming maps (see Figure 1 for scale). Plots of the mapping (similar to Figure 4) between reference and ML-PEA outputs are shown in the top right of each dimming map. Matched power savings are shown in the inset plots.*

identical, whereas SSIM scores are largely better for ML-PEA. JOD scores predicted by ColorVideoVDP are almost always significantly lower (worse quality) for ML-PEA than they are for uniform dimming. This is to say that the proper metric for evaluating ML display power optimization algorithms is yet to be determined; SSIM and PSNR metrics used in prior works can yield wildly different conclusions. We analyze a wider set of metrics in Appendix Section 10, however, and suggest an improved evaluation protocol. In prior art [LDB23], metric results are compared with their corresponding *target* power savings. If actual power savings greatly differ from the target, these metric comparisons may be unfair. For example, a high metric score could correspond to actual power savings much lower than the target, which is undesirable. We recommend that comparisons of metric scores be computed on images with matched power savings. As such, in this and the subjective evaluation, we compare with *uniform dimming* because it is a simple method that performs surprisingly well [CWM*24] and has predictable power savings.

4.2. Ablation Study

A sweep of the weights on each loss function, SSIM and VGG, was conducted. As shown in Figure 4, with only a VGG loss, generated images have halo artifacts around edges (see around bird), whereas images generated with SSIM do not show these halos, but exhibit high-frequency distortions which are made clear in the right scatter plot. The combination of the two loss terms greatly mitigates both. Metric scores are included in Figure 4, where scores for the VGG+SSIM condition are higher, with a better match to the target power rate. See Appendix Sections 9.1 to 9.3 for ablations on the number of dimming map channels, the application function f, and comparisons to prior art.

4.3. Generalization

We tested the ability of our model to generalize to images not part of the DIV2K training dataset. We collected a number of photos taken with two different smartphones and a DSLR camera. Images are variable in resolution, ranging from 1400x933 to 6000x4000. We found that despite the different characteristics of these images, ML-PEA produced good outputs. Qualitative comparisons with uniform dimming are shown in Figure 5, all generated with ML-PEA trained at target 45% power savings ($\alpha = 0.55$). Examples with zoomed insets are shown in Fig. 6. Additional results are shown in Appendix Section 12.

5. User Study

Prior art in ML-based display power savings used metrics exclusively to evaluate the effectiveness of their algorithms. However, as shown in Sec 4.1, metrics often have inconsistencies, and consequently this analysis is insufficient. Here, we describe the first subjective study validating a deep learning-based method for display power savings. In this study, we use the same study protocol and stimuli as PEA-PODs [CWM*24] to allow comparisons between our results and the display power optimization techniques

Figure 6: *We show more examples, with zoomed insets. Power savings are 54% and 49% for left and right images, respectively.*

studied in their work. Specifically, we compare ML-PEA with uniform dimming because it is easy to implement and yields good subjective results in a VR setting [CWM*24]. Furthermore, uniform dimming is a good baseline because it yields controllable power savings (equal to the percentage of dimming).

Participants We recruited 10 users (8M/2F, 22-34 years of age), all of which had normal or corrected-to-normal vision. We note that the gender of participants was unbalanced; we confirmed, via outlier analysis (see Section 4), that results for female participants were not significantly different from those of the overall participant pool. An Institutional Review Board (IRB) approved the study, and participants gave informed consent before starting the experiment.

Stimuli User study stimuli were displayed on a Meta Quest Pro commercial VR head-mounted display (HMD). The display has a resolution of 1800×1920, and a vertical/horizontal field of view of 95.57° and 106°, respectively.

We used 5 stereoscopic videos, all of which are frustums of the 360° videos from the PEA-PODs study [CWM*24]. These include three real-world videos captured by a stereo 360° camera from the LIVE-FBT-FCVR database [JCG*19; JCB*20; JCG*21], and two scenes that represent scrolling and video watching applications. Frustums were sampled at the most salient gaze positions for each 360° scene, as defined by gaze data collected in PEA-PODs. Videos have a resolution of 1800×1920, the same as that of the VR hardware, and were 10-12 seconds long. Head-tracked, 360° videos were not used because our method was not implemented to run in real time.

Uniform dimming and ML-PEA were applied to each video. The techniques were applied at target power saving rates of 17%, 32%, and 45%. ML-PEA is unable to match the exact target power saving rate, as described in Section 4.1. In order to make comparisons fair, we match the power saving rate of uniform dimming by scaling frames by the average savings across all frames of the stimuli with ML-PEA applied. We show the power savings of representative frames for each scene for both techniques in Figure 7.

In total, the study consisted of 5 videos $\times$ 2 display optimization techniques $\times$ 3 power saving rates = 30 conditions. A full design would result in $\binom{30}{2} = 435$ trials. Assuming a participant takes about 15 seconds to complete a trial, this naive design would take around 2 hours to complete. To speed up data collection, we used an active sampling protocol, *ASAP* [MWP*21], which schedules comparisons so that each trial provides optimal information gain. This allows us to sample one trial per study condition, reducing the number of trials to 30. Our final study consisted of two repeats of this protocol, for a total of 60 trials per user.

Experiment Procedure Participants are first read the instructions shown in Appendix Figure 14. We follow the two-interval forced choice (2IFC) protocol, which has been shown to produce less noisy results and an easier task for users [MTM12]. Participants were seated for the duration of the study, and are first presented with a reference, unmodified, video to start each trial. They can then freely swap between this reference and two test videos by clicking on keyboard buttons. A test video can either be the reference video, the uniformly dimmed video, or the video with ML-PEA applied. Users are asked to select the video which contains fewer distortions and is closer to the reference. A 500 ms gray blank is shown when users switch stimuli so that participants cannot make direct comparisons. The results of the user study were anonymized.

5.1. Study Results

We scale our user study results to a perceptual just-objectionable-difference (JOD) scale using the *pwcmp* algorithm described by Perez-Ortiz et al. [PM17]. Notably, JOD units can be converted to intuitively interpretable percentage preference values. For instance, a difference in 1 JOD between method A and B indicates that participants selected method A 75% of the time over method B (see Appendix Section 13.2 for more). The scaled results are shown (as stars) in Figure 8 for both uniform dimming and ML-PEA. Outliers were detected, via *pwcmp*, if observers had an inter-quartile-normalised score above 1.5. We found 1 outlier participant using this procedure, who was subsequently removed from our analysis. Error bars are computed via bootstrapping. In Figure 8, the reference condition (black shaded area) itself has an error bar because it was a condition in our study.

An N-way analysis of variance (ANOVA) was run on the results. We found that the main effect of display power optimization technique (ML-PEA vs. uniform dimming) on JODs is significant ($p = 0.0007$), and that the main effect of method strength (α) on JOD scores is significant ($p \ll 0.05$). The effect of scene on JODs was not found to be significant ($p = 0.08$). Mean JODs of ML-PEA and uniform dimming applied at the same strength level are only significantly different at level 3 (45% target power savings).

5.2. Re-scaling to PEA-PODs Data

In order to make comparisons between the results of our study and those of PEA-PODs, we have to *rescale* the JOD scores of our study to the scores of their dataset. We define rescaling as an operation

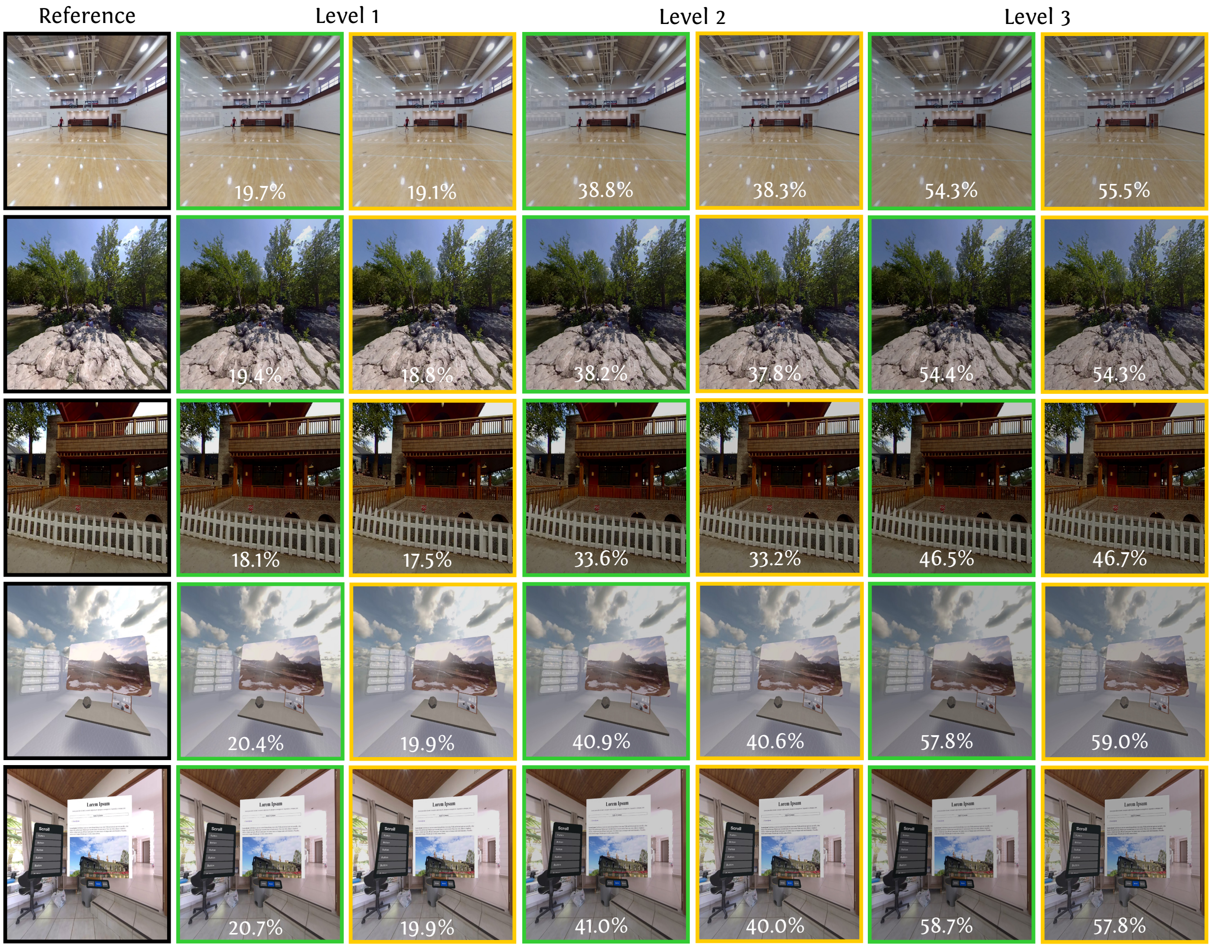

Figure 7: *All scenes used in the user study with both techniques, ML-PEA (●) and uniform dimming (●), applied are shown here, as well as the corresponding power savings for the displayed frames. The levels (1, 2, and 3) correspond to the strengths of application of the power-saving techniques. They correspond to roughly 17%, 32%, and 45% target power savings (uniform dimming power savings were averaged across frames to match those of ML-PEA).*

that maps quality scores from one study to the range of another. This operation is required because, for the same distortion applied to a piece of content, quality scores can vary in magnitude between different studies (even though quality scores may have the same relative order within studies). To start, our results for uniform dimming are rescaled to PEA-PODs using a linear model,

$$\mathrm{JOD}_{\text{PEA-PODs}} = a \cdot \mathrm{JOD}_{\text{ML-PEA}} + b, \tag{8}$$

where a and b are parameters to be optimized. Uniform dimming was selected for this purpose because it is a distortion contained in both our study and that of Chen et al. [CWM*24]. This linear function is applied to all JOD scores collected in our study ($\mathrm{JOD}_{\text{ML-PEA}}$), mapping them to the JOD scale of PEA-PODs.

Note that the power saving rates for uniform dimming differ from those used in the PEA-PODs study, because we match them with ML-PEA across scenes in our study. To account for this mismatch in power savings, we first computed the cross-scene mean power savings for both uniform dimming and ML-PEA, and applied a transfer function, Ψ: a non-linear model fit to uniform dimming power saving rates (α) versus JOD data, $\mathrm{JOD}_{\text{PEA-PODs}_i} = \Psi(\alpha)$, using the Weibull distribution function.

The results of this procedure are shown in Figure 8, where our results are compared to the techniques studied by Chen et al. [CWM*24]. To allow for easier comparison of techniques across different power saving rates, we evaluated $\Psi(\alpha)$ for each rate independently (visualized as dotted lines). We use the function parameters from PEA-PODs for all methods except the ones studied in this work (uniform dimming and ML-PEA), for which we computed a linear fit.

Discussion The technique offering the best power savings for the OLED display (left plot in Figure 8) is the *Brightness Rolloff*. However, this technique requires active eye tracking, which itself incurs significant power costs [CWM*24]. Next, *Dichoptic Dimming*, *ML-PEA*, and *Uniform Dimming* have similar performance. Notably, *ML-PEA* performs the best of any non-eye tracked method for

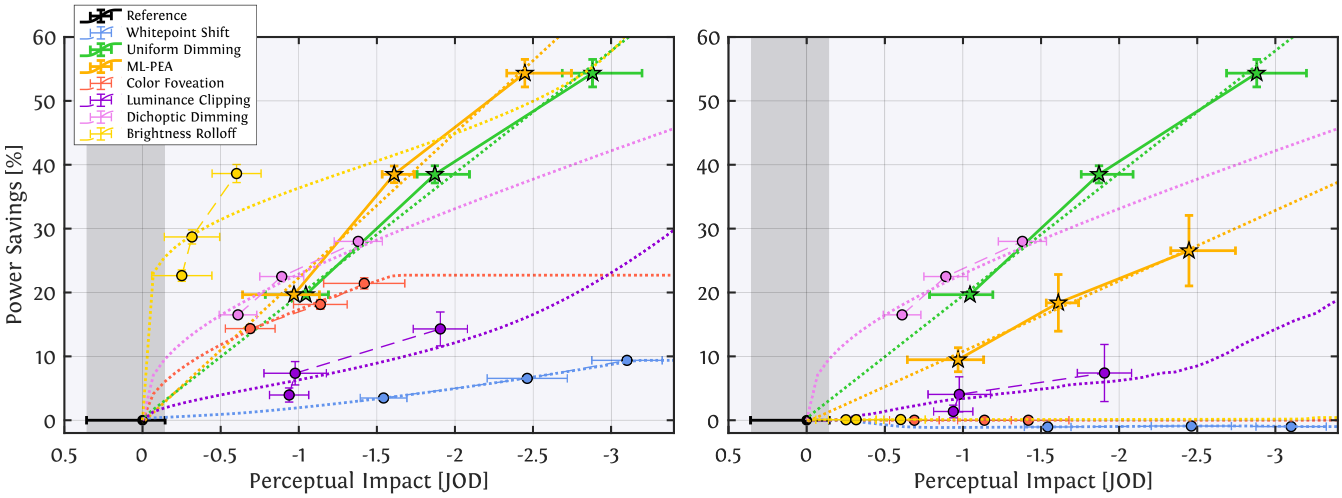

Figure 8: *We plot the user study results, scaled to JOD units, on the x-axis. The y-axis represents power saved. Note that this plot shows increasing negative perceptual impact on the x-axis, so points at the top left of the plot are better. Solid lines with star markers are conditions studied in this work. The reference condition is plotted as a shaded area at 0 JODs. Error bars are 95% confidence intervals. Dotted lines are transfer functions fit to the data.*

power savings of $\sim$ 26% and more, and is preferred over uniform dimming by 61.5% of users (0.43 JODs) for high power saving rates. In LC displays with a global dimming backlight (right plot in Figure 8), brightness rolloff yields no power savings because the central region (at the eye gaze position) remains unchanged. Furthermore, the color modulation techniques yield no power, as discussed in Section 2. The top two techniques are *Dichoptic Dimming* and *Uniform Dimming*. Our *ML-PEA* technique has the next best performance. Note that the difference in power savings between OLED and LC displays is due to the specific power model used in the power loss function that our model is optimized with. The power consumed by a global dimming LC display is related to the maximum pixel intensity in the image, which is not optimized for in our pipeline.

6. Limitations and Future Work

While ML-PEA provides optimal results for key power saving conditions, it also has some limitations. First, our analysis does not cover the computational cost or speed of the explored techniques, focusing only on display power. This is sufficient for applications that involve pre-rendered content, such as video streaming. However, it is not appropriate in a power-constrained, real-time environment like those present in VR gaming, where the costs due to inference computation would impact the expected savings rate. Future work evaluating and optimizing the inference cost and speed of ML-PEA would help extend it to these applications. VR systems that already contain ML modules, e.g. for foveated rendering [KSL*19], may be able to implement our pipeline in addition. Current commercial systems already show promise in producing real-time ML-based vision systems on-device [Cad24].

Although our model does not explicitly incorporate time-varying effects, no temporal artifacts were observed during development and evaluation of ML-PEA. This may be due, in part, to lack of extreme frame-by-frame changes in the reference videos, as well as the non-stochastic nature of U-NET. If present, temporal inconsistencies would lead to lower JOD scores in a subjective study like the one presented in this work. Future work could explore loss functions or models that incorporate spatial and temporal features. In addition, the implementation of ML-PEA described in this work is trained on specific target power saving rates. The method could be further extended by incorporating the target savings rate as an input to the model.

7. Conclusion

We proposed a deep-learning-based solution for optimizing display power savings while maintaining image quality. We demonstrated that the evaluation methodology used by previous ML-based methods is likely insufficient, and proposed robust alternatives based on recent subjective study protocols. We validated our method's performance using these techniques, and found that it has competitive performance with state-of-the-art techniques, obtaining best-in-class results in some cases. Our work serves as a baseline and provides a subjective evaluation framework that can be used for future development of ML-based display power saving models.

References

[ADMM25] Ameur, Zoubida, Demarty, Claire-Hélène, Ménard, Daniel, and Meur, Olivier Le. "3R-INN: How to Be Climate Friendly While Consuming/Delivering Videos?": *Computer Vision – ECCV 2024*. Ed. by Leonardis, Aleš, Ricci, Elisa, Roth, Stefan, et al. Cham: Springer Nature Switzerland, 2025, 146–163. ISBN: 978-3-031-73226-3 3.

[AR22] Adi Nugroho, Kuntoro and Ruan, Shanq-Jang. "R-ACE Network for OLED Image Power Saving". *2022 IEEE 4th Global Conference on Life Sciences and Technologies (LifeTech)*. 2022, 284–285. DOI: 10.1109/LifeTech53646.2022.9754748 3.

[AT17] Agustsson, Eirikur and Timofte, Radu. "Ntire 2017 challenge on single image super-resolution: Dataset and study". *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2017, 126–135 4.

[ATS*11] Anand, Bhojan, Thirugnanam, Karthik, Sebastian, Jeena, et al. "Adaptive display power management for mobile games". *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*. MobiSys '11. Bethesda, Maryland, USA: Association for Computing Machinery, 2011, 57–70. ISBN: 9781450306430. DOI: 10.1145/1999995.2000002. URL: https://doi.org/10.1145/1999995.2000002 1.

[Cad24] Cadenas, Cesar. "The Quest 3 is getting a personal assistant courtesy of Meta AI". *TechRadar* (2024). URL: https://www.techradar.com/computing/virtual-reality-augmented-reality/the-quest-3-is-getting-a-personal-assistant-courtesy-of-meta-ai?utm_source=chatgpt.com 8.

[CDU*23] Chen, Kenneth, Duinkharjav, Budmonde, Ujjainkar, Nisarg, et al. "Imperceptible Color Modulation for Power Saving in VR/AR". *ACM SIGGRAPH 2023 Emerging Technologies*. SIGGRAPH '23. 2023. URL: https://doi.org/10.1145/3588037.3595388 2.

[CGZH20] Cai, Han, Gan, Chuang, Zhu, Ligeng, and Han, Song. "Tinytl: Reduce memory, not parameters for efficient on-device learning". *Advances in Neural Information Processing Systems* 33 (2020), 11285–11297 2.

[CMM*25] Chen, Kenneth, Matsuda, Nathan, McElvain, Jon, et al. "What is HDR? Perceptual Impact of Luminance and Contrast in Immersive Displays". *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*. SIGGRAPH Conference Papers '25. New York, NY, USA: Association for Computing Machinery, 2025. ISBN: 9798400715402. DOI: 10.1145/3721238.3730629. URL: https://doi.org/10.1145/3721238.3730629 3.

[CSC02] Choi, Inseok, Shim, Hojun, and Chang, Naehyuck. "Low-power color TFT LCD display for hand-held embedded systems". *Proceedings of the International Symposium on Low Power Electronics and Design*. 2002, 112–117. DOI: 10.1109/LPE.2002.146722 2.

[CWM*24] Chen, Kenneth, Wan, Thomas, Matsuda, Nathan, et al. "PEA-PODs: Perceptual Evaluation of Algorithms for Power Optimization in XR Displays". *ACM Trans. Graph.* 43.4 (July 2024). ISSN: 0730-0301. DOI: 10.1145/3658126. URL: https://doi.org/10.1145/3658126 2, 3, 5–7.

[DCT*22] Duinkharjav, Budmonde, Chen, Kenneth, Tyagi, Abhishek, et al. "Color-Perception-Guided Display Power Reduction for Virtual Reality". *ACM Trans. Graph. (Proc. SIGGRAPH Asia)* 41.6 (2022), 144:1–144:16 2–4.

[DZ11] Dong, Mian and Zhong, Lin. "Chameleon: a color-adaptive web browser for mobile OLED displays". *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services*. MobiSys '11. Bethesda, Maryland, USA: Association for Computing Machinery, 2011, 85–98. ISBN: 9781450306430. DOI: 10.1145/1999995.2000004. URL: https://doi.org/10.1145/1999995.2000004 2, 3.

[DZ12] Dong, Mian and Zhong, Lin. "Power Modeling and Optimization for OLED Displays". *IEEE Transactions on Mobile Computing* 11.9 (2012), 1587–1599. DOI: 10.1109/TMC.2011.167 2.

[GABR02] Gatti, Franco, Acquaviva, Andrea, Benini, Luca, and Ricco', Bruno. "Low Power Control Techniques For TFT LCD Displays". *Proceedings of the 2002 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems*. CASES '02. Grenoble, France: Association for Computing Machinery, 2002, 218–224. ISBN: 1581135750. DOI: 10.1145/581630.581664. URL: https://doi.org/10.1145/581630.581664 2.

[Had17] Hadizadeh, Hadi. "Energy-Efficient Images". *IEEE Transactions on Image Processing* 26.6 (2017), 2882–2891. DOI: 10.1109/TIP.2017.2690523 3.

[HTH22] Hsiung, Hao-Lun Sun1 Lei, Tsung, Nandhini Chandramoorthy2 Pin-Yu Chen, and Ho, Yi. "NeuralFuse: Improving the Accuracy of Access-Limited Neural Network Inference in Low-Voltage Regimes". *Annual Conference on Neural Information Processing Systems*. 2022 2.

[ITU24] ITU, International Telecommunication Union Radiocommunication Sector. "Display energy reduction through image signal processing". *Report ITU-R BT. 2540-0* (2024) 3.

[JAF16] Johnson, Justin, Alahi, Alexandre, and Fei-Fei, Li. "Perceptual Losses for Real-Time Style Transfer and Super-Resolution". *Computer Vision – ECCV 2016*. Ed. by Leibe, Bastian, Matas, Jiri, Sebe, Nicu, and Welling, Max. Cham: Springer International Publishing, 2016, 694–711. ISBN: 978-3-319-46475-6 3.

[JCB*20] Jin, Yize, Chen, Meixu, Bell, Todd Goodall, et al. "Study of 2D foveated video quality in virtual reality". *Applications of Digital Image Processing XLIII*. Ed. by Tescher, Andrew G. and Ebrahimi, Touradj. Vol. 11510. International Society for Optics and Photonics. SPIE, 2020, 1151007. DOI: 10.1117/12.2568883. URL: https://doi.org/10.1117/12.2568883 6.

[JCG*19] Jin, Yize, Chen, Meixu, Goodall, Todd, et al. "LIVE-Facebook Technologies-Compressed Virtual Reality (LIVE-FBT-FCVR) Databases". *Online: http://live. ece. utexas. edu/research/LIVEFBTFCVR/index. html* (2019) 6.

[JCG*21] Jin, Yize, Chen, Meixu, Goodall, Todd, et al. "Subjective and objective quality assessment of 2d and 3d foveated video compression in virtual reality". *IEEE Transactions on Image Processing* 30 (2021), 5905–5919 6.

[KB14] Kingma, Diederik P and Ba, Jimmy. "Adam: A method for stochastic optimization". *arXiv preprint arXiv:1412.6980* (2014) 4.

[KD06] Kerofsky, Louis and Daly, Scott. "Brightness preservation for LCD backlight dimming". *Journal of the Society for Information Display* 14.12 (2006), 1111–1118 2, 3.

[KL20] Kim, Jeong-Sik and Lee, Seung-Woo. "Peripheral Dimming: A New Low-Power Technology for OLED Display Based on Gaze Tracking". *IEEE Access* 8 (2020), 209064–209073. DOI: 10.1109/ACCESS.2020.3038428 3.

[KSL*19] Kaplanyan, Anton S., Sochenov, Anton, Leimkühler, Thomas, et al. "DeepFovea: neural reconstruction for foveated rendering and video compression using learned statistics of natural videos". *ACM Trans. Graph.* 38.6 (Nov. 2019). ISSN: 0730-0301. DOI: 10.1145/3355089.3356557. URL: https://doi.org/10.1145/3355089.3356557 8.

[LDB23] Le Meur, Olivier, Demarty, Claire-Hélène, and Blonde, Laurent. "Deep-Learning-Based Energy Aware Images". *2023 IEEE International Conference on Image Processing (ICIP)*. 2023, 590–594. DOI: 10.1109/ICIP49359.2023.10222188 3, 5.

[LLLK12] Lee, Chulwoo, Lee, Chul, Lee, Young-Yoon, and Kim, Chang-Su. "Power-Constrained Contrast Enhancement for Emissive Displays Based on Histogram Equalization". *IEEE Transactions on Image Processing* 21.1 (2012), 80–93. DOI: 10.1109/TIP.2011.2159387 3.

[MD23] Meur, Olivier Le and Demarty, Claire-Hélène. "Invertible Energy-Aware Images". *IEEE Signal Processing Letters* 30 (2023), 1392–1396. DOI: 10.1109/LSP.2023.3320558 3.

[MHA*24] Mantiuk, Rafal K., Hanji, Param, Ashraf, Maliha, et al. "ColorVideoVDP: A visual difference predictor for image, video and display distortions". *ACM Trans. Graph.* 43.4 (July 2024). ISSN: 0730-0301. DOI: 10.1145/3658144. URL: https://doi.org/10.1145/3658144 4.

[MTM12] MANTIUK, RAFAŁ K., TOMASZEWSKA, ANNA, and MANTIUK, RADOSŁAW. "Comparison of Four Subjective Methods for Image Quality Assessment". *Computer Graphics Forum* 31.8 (2012), 2478–2491. DOI: https://doi.org/10.1111/j.1467-8659.2012.03188.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2012.03188.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2012.03188.x 6.

[MWP*21] MIKHAILIUK, ALIAKSEI, WILMOT, CLIFFORD, PEREZ-ORTIZ, MARIA, et al. "Active Sampling for Pairwise Comparisons via Approximate Message Passing and Information Gain Maximization". *2020 IEEE International Conference on Pattern Recognition (ICPR)*. Jan. 2021 6.

[PM17] PEREZ-ORTIZ, MARIA and MANTIUK, RAFAL K. "A practical guide and software for analysing pairwise comparison experiments". *arXiv preprint arXiv:1712.03686* (2017) 6.

[PS16] PARK, MINYOUNG and SONG, MINSEOK. "Saving Power in Video Playback on OLED Displays by Acceptable Changes to Perceived Brightness". *Journal of Display Technology* 12.5 (2016), 483–490. DOI: 10.1109/JDT.2015.2501318 2.

[RFB15] RONNEBERGER, OLAF, FISCHER, PHILIPP, and BROX, THOMAS. "U-Net: Convolutional Networks for Biomedical Image Segmentation". *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by NAVAB, NASSIR, HORNEGGER, JOACHIM, WELLS, WILLIAM M., and FRANGI, ALEJANDRO F. Cham: Springer International Publishing, 2015, 234–241. ISBN: 978-3-319-24574-4 4.

[SCD25] SURACE, LUCA, CONDOR, JORGE, and DIDYK, PIOTR. "Temporal Brightness Management for Immersive Content". *Eurographics Symposium on Rendering*. Ed. by WANG, BEIBEI and WILKIE, ALEXANDER. The Eurographics Association, 2025. ISBN: 978-3-03868-292-9. DOI: 10.2312/sr.20251183 3.

[SDL24] SAMPAIO, EMMANUEL, DEMARTY, CLAIRE-HÉLÈNE, and LE MEUR, OLIVIER. "How to make images less power-hungry: An objective benchmark study". *Proceedings of the Second International ACM Green Multimedia Systems Workshop*. GMSys '24. Bari, Italy: Association for Computing Machinery, 2024, 1–7. ISBN: 9798400706172. DOI: 10.1145/3652104.3652525. URL: https://doi.org/10.1145/3652104.3652525 3.

[SSM09] SHYE, ALEX, SCHOLBROCK, BENJAMIN, and MEMIK, GOKHAN. "Into the wild: studying real user activity patterns to guide power optimizations for mobile architectures". *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture*. MICRO 42. New York, New York: Association for Computing Machinery, 2009, 168–178. ISBN: 9781605587981. DOI: 10.1145/1669112.1669135. URL: https://doi.org/10.1145/1669112.1669135 2.

[THW*07] TRENTACOSTE, MATTHEW, HEIDRICH, WOLFGANG, WHITEHEAD, LORNE, et al. "Photometric image processing for high dynamic range displays". *Journal of Visual Communication and Image Representation* 18.5 (2007), 439–451 2.

[TOMB13] TAN, KIAT WEE, OKOSHI, TADASHI, MISRA, ARCHAN, and BALAN, RAJESH KRISHNA. "FOCUS: a usable & effective approach to OLED display power management". *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. UbiComp '13. Zurich, Switzerland: Association for Computing Machinery, 2013, 573–582. ISBN: 9781450317702. DOI: 10.1145/2493432.2493445. URL: https://doi.org/10.1145/2493432.2493445 1.

[WBSS04] WANG, ZHOU, BOVIK, A.C., SHEIKH, H.R., and SIMONCELLI, E.P. "Image quality assessment: from error visibility to structural similarity". *IEEE Transactions on Image Processing* 13.4 (2004), 600–612. DOI: 10.1109/TIP.2003.819861 3, 4.

[WCB*18] WANG, NAIGANG, CHOI, JUNGWOOK, BRAND, DANIEL, et al. "Training deep neural networks with 8-bit floating point numbers". *Advances in neural information processing systems* 31 (2018) 2.

[WCB18] WEE, TAN KIAT, CUERVO, EDUARDO, and BALAN, RAJESH. "FocusVR: Effective 8 Usable VR Display Power Management". *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2.3 (Sept. 2018). DOI: 10.1145/3264952. URL: https://doi.org/10.1145/3264952 3.

[YSLX18] YAN, ZHISHENG, SONG, CHEN, LIN, FENG, and XU, WENYAO. "Exploring Eye Adaptation in Head-Mounted Display for Energy Efficient Smartphone Virtual Reality". *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*. HotMobile '18. Tempe, Arizona, USA: Association for Computing Machinery, 2018, 13–18. ISBN: 9781450356305. DOI: 10.1145/3177102.3177121. URL: https://doi.org/10.1145/3177102.3177121 3.