

Temporal Brightness Management for Immersive Content

Luca Surace¹ , Jorge Condor¹ , Piotr Didyk¹ 

¹Università della Svizzera italiana, Switzerland

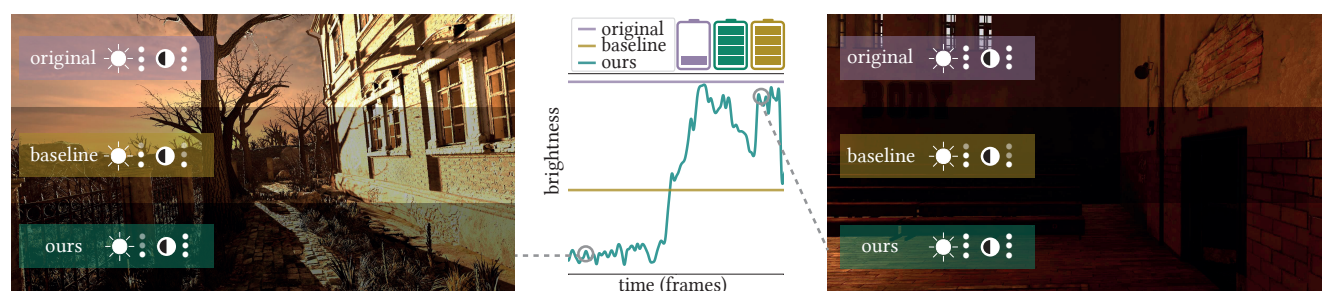


Figure 1: Our method adjusts the screen brightness based on the content of each frame. In the outdoor scene on the left, characterized by high brightness, our technique darkens the frame to a greater extent compared to a constant reduction (baseline) without significant losses in the visibility of image details. Conversely, in the darker indoor space shown on the right, where dimming the screen can easily lead to loss of visible image details, our technique increases the brightness to better preserve contrast. Throughout the entire sequence, our technique utilizes the same power budget as the baseline.

Abstract

Modern virtual reality headsets demand significant computational resources to render high-resolution content in real-time. Therefore, prioritizing power efficiency becomes crucial, particularly for portable versions reliant on batteries. A significant portion of the energy consumed by these systems is attributed to their displays. Dimming the screen can save a considerable amount of energy; however, it may also result in a loss of visible details and contrast in the displayed content. While contrast may be partially restored by applying post-processing contrast enhancement steps, our work is orthogonal to these approaches, and focuses on optimal temporal modulation of screen brightness. We propose a technique that modulates brightness over time while minimizing the potential loss of visible details and avoiding noticeable temporal instability. Given a predetermined power budget and a video sequence, we achieve this by measuring contrast loss through band decomposition of the luminance image and optimizing the brightness level of each frame offline to ensure uniform temporal contrast loss. We evaluate our method through a series of subjective experiments and an ablation study, on a variety of content. We showcase its power-saving capabilities in practice using a built-in hardware proxy. Finally, we present an online version of our approach which further emphasizes the potential for low level vision models to be leveraged in power saving settings to preserve content quality.

CCS Concepts

• **Computing methodologies** → **Perception**; **Virtual reality**;

1. Introduction

Power consumption is a fundamental axis of hardware design, particularly in edge and portable devices. Battery life and computational power form a careful balance, and efficient use of power resources can directly contribute to increased computational power budget for more realistic and interactive applications.

In virtual reality (VR) devices, dimming the screen is a common approach to energy preservation, but it encompasses a trade-off

between power consumption and image quality. Recent work has highlighted this problem [CWM*24], comparing various power-saving techniques and their respective impacts on perceived quality. Their analysis shows that, on OLED architectures with a 30% power-saving target, uniform dimming ranks as the second-best, in terms of perceptual impact measured by Just-Objectable Differences (JOD). Furthermore, when the power-saving target exceeds 60%, uniform dimming becomes the most effective approach for

maintaining visual quality. For more common Liquid Crystal Displays (LCDs), uniform dimming is the only method capable of reducing power consumption by 50% without incurring a significant perceptual quality loss. Therefore, refining global dimming techniques is a promising avenue for advancing power-efficient display designs across various architectures. However, substantial corrections to brightness levels risk reducing content visibility and luminance contrast.

Previous works address this problem with frame-dependent brightness scaling techniques, evaluated using standard image quality metrics [PDP*19]. Additionally, user studies have assessed the acceptability of brightness degradation [PS15]. Integrating post-processing steps with power-saving measures, including contrast enhancement [PMP18] and luminance retargeting [WM14] is a relatively simple approach to partly mitigate the degradation in quality. Unlike previous techniques, however, our method explicitly models luminance adaptation, resulting in imperceptible brightness modulations within VR environments.

In this paper, we present a novel technique for dynamically adjusting screen brightness to reduce display power consumption. Our method is formulated as an optimization problem based on a low-level human vision model that accounts for perceived luminance contrast. The optimization process aims to achieve a balance among three objectives. First, it aims to maintain power consumption at a specified input level. Second, it seeks to ensure consistent luminance contrast loss throughout the entire sequence. Finally, it guarantees that any variations in brightness over time are not perceived as artifacts, such as visible flickering. Here, we specifically design our methods for virtual reality displays, which allows us to leverage the fact that users immersed in a virtual environment without access to real-world stimuli exhibit low sensitivity to global luminance changes. We thoroughly validate our technique through subjective human experiments and measure its power savings on actual hardware. We demonstrate that our method not only improves perceived quality compared to a baseline solution which reduces the brightness uniformly over the sequence duration, but can also lead to improved task performance. Apart from the offline optimization, we also propose a simple online solution which realizes the same goals.

2. Related work

Power reduction strategies vary depending on display technology. LCDs depend on a backlight filtered by a color matrix, making backlight dimming the main method for reducing power consumption. Conversely, OLED panels are self-emissive, allowing more precise control since the intensity and color of each pixel impact overall power use. Irrespective of the technology, reducing power can result in a perceived loss of image quality. Advanced techniques aim to balance power efficiency with image quality. Recently, Chen et al. [CWM*24] established a standard framework for quantifying the perceptual impacts and power savings of different solutions, which highlights global dimming as one of the most promising avenues for imperceptible power use reduction.

Display dimming The most dominant method for reducing power consumption is display dimming [CP04; CSC02; GABR02;

CWM*24; PS15]. In addition to various models of power consumption and dimming strategies, some approaches focused on explicitly addressing the trade-off between power consumption and quality loss by using simple quality metrics like PSNR or SSIM as constraints [Kan15a; Kan15b; CXS16]. More advanced methods integrate quality metrics as objectives within optimization processes. For instance, Pagliari et al. [PDP*19] employed a simplified SSIM metric in an optimization framework that determines the appropriate brightness for each displayed frame. In the context of immersive wide-field-of-view display, it is possible to apply brightness modulation in the periphery, where it is less noticeable [KL20]. These simple metrics, however, are not capable of capturing the full complexity of the human visual system (HVS) and can lead to temporal artifacts or smaller potential power savings [PDP*19].

Color modulation In non-quantum dot self-emissive displays, like OLED panels, it is possible to enhance power efficiency by remapping pixel colors based on the efficiencies of primary colors. In certain cases, changing the color scheme or profile can be a viable method to save energy [DZ11b; DZ11a]. However, when it comes to natural content, remapping colors may produce undesirable effects. It has been demonstrated that this strategy can still be effectively applied in wide-field-of-view displays by taking advantage of the limitations of the HVS in peripheral vision and modulating colors only in those regions [DCT*22].

Compensation methods Dimming the display can significantly impact the quality of image details, color, and depth perception. A number of techniques have been dedicated to compensate for such undesired changes. An example of a simple compensation strategy is an increase in image brightness prior to display, which preserves the visibility of image details in dark image regions but introduces clipping to bright regions. Kerofsky and Dally [KD06] proposed to counteract this effect by smoothly attenuating bright image values to preserve bright image details. Such an approach is similar to other methods that apply global contrast adjustment to correct for the lost details [CSC02; CP04; CKK13]. The loss of spatial details can be more effectively addressed by local contrast enhancement methods. Such methods have been extensively studied in the context of tone mapping, with examples such as, [MDK08; KMS07; EMU16]. They usually rely on accurate modeling of luminance perception and frameworks for local contrast processing. In the context of power-efficient displays, existing works proposed retargeting methods for matching image appearance between dark and bright conditions [WM14; AMMW22]. It is important to note that while the visibility of image details is the primary visual cue affected by reduced brightness, others that rely on luminance patterns can also be inhibited. An example of such a cue is stereopsis [DRE*12]. The problem was investigated in [WZMM22], where the authors proposed a method for compensating side effects of reduced luminance conditions on depth perception. Another issue related to brightness modulation can be temporal coherence. Rapid changes in luminance can lead to flickering artifacts, while slower modulation rates may result in brightness incoherence over time. Closest to our method are the techniques that focus on preserving temporal brightness coherency in the context of tone mapping [BBC*12; BCTB14]. It is important to note that any post-processing effort to recover lost contrast is orthogonal to our work,

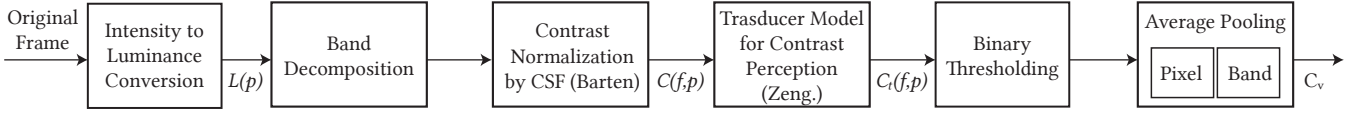


Figure 2: Pipeline of our method. The conversion to luminance is achieved using the luminance curve of the display and applying the display model from Mantiuk et al. (Equation 1). The self-parameters for masking remain consistent with the original model proposed by Zeng et al. [ZDL00]: $\alpha = 0.7$ and $\beta = 0.2$. The output value C_v represents the visible contrast in the frame.

and could be integrated into our optimization procedure to further reduce power consumption.

Discussion The method presented in this paper aims to enhance power efficiency through display dimming, aligning with existing methods in this category. However, there are significant differences between our approach and previous techniques. First, our method explicitly models luminance adaptation, which is crucial for understanding the reduction in perceived contrast due to decreased image brightness. Second, we rely on a precise perceptual model of perceived luminance contrast, in contrast to some general-purpose quality metrics that do not provide accurate information about visibility of luminance contrast. Finally, our method focuses on brightness modulation that remains imperceptible to users. While many previous techniques were designed for general screens, such as cell phones, our approach specifically targets VR devices. This allows us to leverage the fact that, while immersed in a virtual environment without access to real-world stimuli, users exhibit very low sensitivity to gradual brightness changes. Through careful calibration of our method, we ensure that the brightness modulation remains imperceptible.

3. Method

When reducing the screen brightness, the visibility of luminance contrast in the image changes. The effect can be directly attributed to the reduction of the sensitivity of the HVS to luminance contrast when adaptation luminance decreases [WM14]. Consequently, screen dimming may result in a loss of visible luminance contrast, leading to poorer visibility of image details. Our goal is to achieve a specific average power budget for the display by intelligently dimming the screen based on the content being shown, preserving visible contrast. Since such temporal brightness modulation may lead to visible temporal artifacts, we constrain the modulation rate to remain smooth. Our method quantifies the loss of visible contrast by analyzing perceived luminance contrast modeled using multi-band frequency decomposition (Section 3.1). In addition to quantifying the magnitude of contrast loss, our method also identifies the portions of the image where details become invisible due to the brightness reduction (Section 3.2). An overview of this pipeline is shown in Figure 2. We use this per-frame prediction to optimize the brightness modulation for an entire video sequence offline (Section 3.3).

3.1. Perceived Contrast

In order to locally measure where contrast is perceivable by a human observer, and quantify its magnitude per frame, we adapt Tursun et al. [TAW*19] approach. We analyze the contrast in a frame,

as defined by Michelson et al. [Mic95], using a frequency band decomposition of the luminance image. We first convert image intensity to luminance using Mantiuk et al. [MDK08] display model:

$$L(p) = (L')^\gamma \cdot (L_{max} - L_{black}) + L_{black} + L_{refl}, \quad (1)$$

where L is the displayed luminance, L' is the pixel value at location p , and γ is the display gamma. L_{max} denotes the peak display luminance, while L_{black} is the display's black level. The term L_{refl} refers to ambient light reflected from the display surface. Next, we compute the Laplacian pyramid decomposition [BE83] of the luminance image, obtaining the band-limited luminance difference, $\Delta L(f, p)$, at given frequencies f and image locations p . To compute perceived contrast, we normalize the pyramid by the average luminance in the area $L_a(f, p)$, which is provided by the corresponding point in the Gaussian pyramid two levels down the resolution. Finally, contrast is normalized by the Contrast Sensitivity Function (CSF) [Bar03]. The final perceived contrast value can be then computed as:

$$C(f, p) = \frac{\Delta L(f, p)}{L_a(f, p) + \epsilon} \cdot S_{CSF}(f, L_a(f, p)), \quad (2)$$

where the additional ϵ denotes a small value that prevents singularities in dark image regions. To incorporate the effect of the visual masking, we apply the transducer model of Zeng et al. [ZDL00] with their proposed default parameters and obtain the final perceived luminance contrast as

$$C_t(f, p) = \frac{\text{sign}(C(f, p)) \cdot |C(f, p)|^{0.7}}{1 + \frac{1}{N} \sum_{q \in N(p)} |C(f, q)|^{0.2}}, \quad (3)$$

where $N(p)$ denotes a 5×5 neighborhood in the same band. Differently from the work of Tursun et al. [TAW*19], we do not model retinal eccentricity, relying on a simpler model for the CSF, since we target common global-backlit panels. This may be of interest, however, in gaze-contingent scenarios using self-emitting displays like OLED or QD-LED, and panels with local area dimming.

3.2. Contrast Loss

We use the perceived luminance contrast measure to compute contrast loss due to screen dimming. To do so, we compare perceived contrast pyramids of the reference luminance image I and its dimmed version $D = b \cdot I$, where b is a dimming factor. We obtain the portion of the contrast pyramids containing visible contrast using

$$C_v = \frac{1}{N_t} \sum_f \sum_p \frac{\mathbb{I}(C_t(f, p) > 1)}{N_p^f}, \quad (4)$$

where summation over f iterates over pyramid levels, discarding the coarsest level containing the average value of the image; while the summation over p iterates over all pixels in the level. Additionally, N_l denotes the number of levels, N_p^f is the number of pixels in the layer corresponding to frequency f , and $\mathbb{I}$ is the indicator function which returns 1 if the condition given as the argument is true, and 0 otherwise. Since we normalize by the CSF, we set the threshold value within the argument to 1, as it corresponds to the smallest contrast level that the HVS can reliably perceive, serving as a lower limit for visibility under ideal conditions. Figure 3 illustrates the effect of $\mathbb{I}$ on one contrast pyramid level. Finally, we compute the value of Equation 4 both for the reference images as C_v^I and for the dimmed version C_v^D and define the contrast lost due to the dimming I by a factor b as

$$\mathcal{L}_c(I, b) = 1 - \frac{C_v^D}{C_v^I}, \quad \text{where } D = b \cdot I. \quad (5)$$

3.3. Optimization

Our optimization, based on the idea that spatial contrast affects the perceived quality of image details, as supported by previous work in tone mapping [MMS06], is designed with three goals:

- quantify the spatial contrast loss and reduce its impact, to ensure a consistent visual experience during playback;
- optimizing power consumption, by ensuring that a predetermined power budget is maintained;
- mitigate temporal artifacts, by limiting the maximum brightness variation rate tolerated by the human visual system, as determined through the perceptual experiment (Section 3.4).

However, there is a trade-off between these goals, as it is not possible to satisfy all of them simultaneously, e.g. achieving both a perfectly flat contrast loss and the desired power consumption. In our case, we formulate the model to prioritize meeting a target average power budget on a specific display, treating it as a hard constraint. This constraint effectively determines the total contrast loss. By fixing the power budget, the task becomes how to best distribute contrast loss across the sequence. Therefore, our objective function focuses on minimizing variations in contrast loss throughout the sequence.

The input to the optimization is a video sequence represented by a set of N luminance images $\{I_i\}$ and a target power budget $\mathcal{P}_{base}$. The goal of the optimization is to find a set of per-frame brightness reduction factors $\{b_i\}$, fulfilling our design goals outlined above. More formally, we define the following optimization as:

$$\begin{aligned} & \underset{\{b_i\}}{\text{minimize}} && \sum_{i=1}^N \sum_{j=i+1}^N |\mathcal{L}_c(I_i, b_i) - \mathcal{L}_c(I_j, b_j)| \\ & \text{subject to} && \frac{dL_a(I_i, b_i)}{dt} < \Delta L(b_i) \\ & && \sum_{i=1}^N \frac{\mathcal{P}_i}{N} = \mathcal{P}_{base}, \\ & && \forall_{0 < i \leq N} \quad 0 < b_i < 1, \end{aligned} \quad (6)$$

where $\frac{dL_a(I_i, b_i)}{dt}$ denotes the derivative of the average luminance of each frame and ΔL is the maximum undetectable rate of luminance

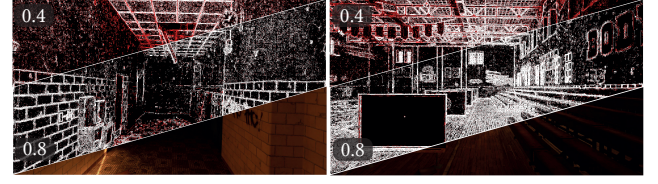


Figure 3: Visible contrast (white) at the highest frequency level of the band decomposition after binary thresholding, and regions where contrast loss occurs between the dimmed frame and the original frame (red), for target average brightness levels of 0.4 (top) and 0.8 (middle). The original frame is displayed at the bottom of the image. The lower the backlight brightness, the higher the amount of detail lost.

change, which we derive in Section 3.4. Finally, $\mathcal{P}_i$ is the power consumption required for the display to show the frame I_i with the dimming factor b_i , and it is dependent on the specific display.

3.4. Brightness modulation rate calibration

We find the maximum rate ΔL at which brightness is modulated according to human perception. We evaluate human sensitivity to global luminance changes by conducting a subjective two-alternative forced choice (2AFC) experiment. We then use the resulting data to ensure that our modulation does not lead to excessively steep transitions, minimizing the risk of producing visible artifacts. Each specific display requires its own calibration.

In the experiment, 10 participants were presented with a white screen of constant brightness, which was linearly decreased over a five-second interval, which we assumed to be a representative time for reliable adaptation in most display usage conditions [WMP19]. This was followed by an additional second of constant brightness. Initial brightness varied from a predefined set of brightness factors $B = \{0.2, 0.4, 0.6, 0.8, 1.0\}$. Additionally, different slopes were tested for each value. Participants were instructed to indicate whether they perceived a change in brightness by responding with either "yes" or "no". The threshold for detecting brightness changes was determined as the value corresponding to a 75% detection rate for each factor, calculated by fitting

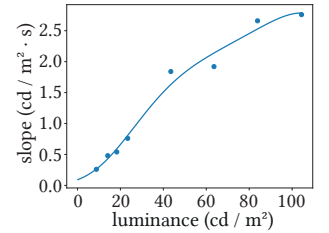


Figure 4: Correlation between luminance and the slope ΔL at the visibility threshold (probability of 75% of detecting the change). The curve is a fourth-degree polynomial fitted to the data points.

a logistic function to the answers. After determining the thresholds, we measured the displayed luminance corresponding to the initial and final brightness levels with a luminance meter. We repeated the procedure using a gray level to cover the relevant luminance range.

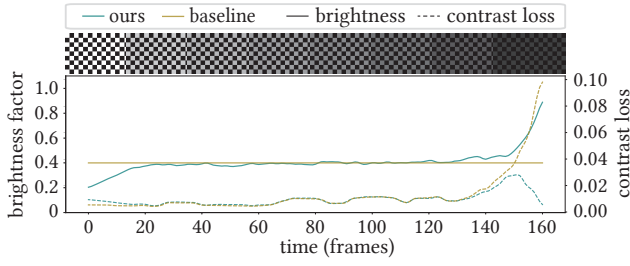


Figure 5: We apply our method to a sequence of checkerboard patterns (top, input frames), where the white squares gradually transition to black through various shades of grey. Our method compensates for the increasing loss of visible contrast by enhancing the brightness accordingly.

In Figure 4 we show the curve obtained from the experimental data.

3.5. Implementation

We minimize the objective function in Eq. 6 using the Sequential Least-Squares method [CMK83]. Additionally, we parameterize our display model (Equation 1) setting ambient light reflected from the display surface to zero, which is a fair assumption on VR setups. We initialize the optimization process with a sequence of constant brightness factors corresponding to the desired target power reduction, perturbed by a small random deviation, δ . Since $\mathcal{L}_c(I_i, b_i)$ is repeatedly evaluated, for each image I_i , we precomputed the term for a set of brightness factors $\{0.2, 0.4, 0.6, 0.8, 1.0\}$ and linearly interpolate between the obtained values during the optimization for efficiency reasons. We ran all our optimizations until convergence with a tolerance value of 10^{-8} . An optimization of 3000 frames takes around three minutes single-threaded on CPU (Intel Core i7-9700K @ 3.60GHz).

4. Results

We test our method with different video sequences exhibiting a variety of scenery, both synthetic (Figures 1, 7) and real (Figure 6).

We visually showcase our algorithm in Figure 5, applied to checkerboard patterns. Initially, the tiles are black and white, maximizing contrast, which allows the brightness factor to remain low. As the white tiles slowly approach black, the algorithm compensates for the loss of contrast by increasing the brightness factor. In a second, more realistic scenario, we run our optimization on a time-lapse video sequence of a sunrise (Figure 6). The video is reversed at half of its duration. During the darker intervals, when the sun is not visible, brightness is high. Conversely, brightness decreases when the sun rises above the horizon.

In order to validate our approach, we ran an extensive set of experiments on human subjects to measure the impact of our modulation on perceived quality and contrast. For all our perceptual experiments, we used a Varjo XR-3 headset. We include details on its hardware capabilities on Supplementary materials.

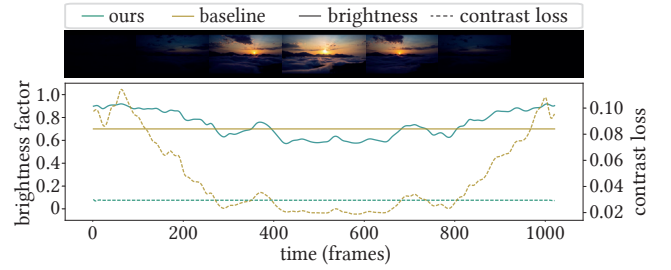


Figure 6: Time-lapse sequence capturing the Sun's movement (top, input frames). Our technique optimizes power allocation by spending more of the "brightness budget" during the initial and final parts of the video, where the frames are darker. At the same time, during the middle section, the dimming can be more aggressive.

Stimuli We created four different camera paths (LIBRARY, SCHOOL, LIVING ROOM, BASEMENT) from two high-quality 3D environments containing a variety of indoor, outdoor and differently illuminated spaces (Figure 7). We generated four video sequences from these paths, each targeting a different average brightness factor: 30%, 40%, 50%, and 70% of the maximum brightness. We selected different brightness factors on different scenes to showcase flexibility of the method while keeping the length of our perceptual experiments in check. We include additional modulation curves targeting different power consumption in the supplementary material. We compute with our method (Figure 2) the optimal brightness modulation (referred to as *brightness factor*) for the four sequences. For comparison, we render a BASELINE alternative, which simply applies constant dimming to achieve the same target power consumption.

4.1. Detail Preservation

The first experiment aimed to assess the quality of perceived details in different environments. In theory, with our modulation, detail visibility should be preserved when brightness is reduced, as opposed to substantial contrast loss when constant-dimming.

Study A total of 17 people naive to the study with normal vision participated in the experiment (ages 23-31). For each of the four video described above, the two alternatives OURS and BASELINE are shown in a randomized order from one subject to the next. The task of the participants was to give a score 1-5 to rate the quality and saliency of the details found in the scene for each alternative, where 1 indicates low quality and 5 indicates high quality.

Results Table 1 presents the results obtained for all four sequences. Our method consistently outperformed the baseline, receiving higher ratings for perceived detail quality in both scenarios.

4.2. Contrast Preservation through a Performance Experiment

In order to quantitatively measure the preservation of contrast provided by our approach in the previous sequences, we designed a task performance experiment based on finding various Landolt

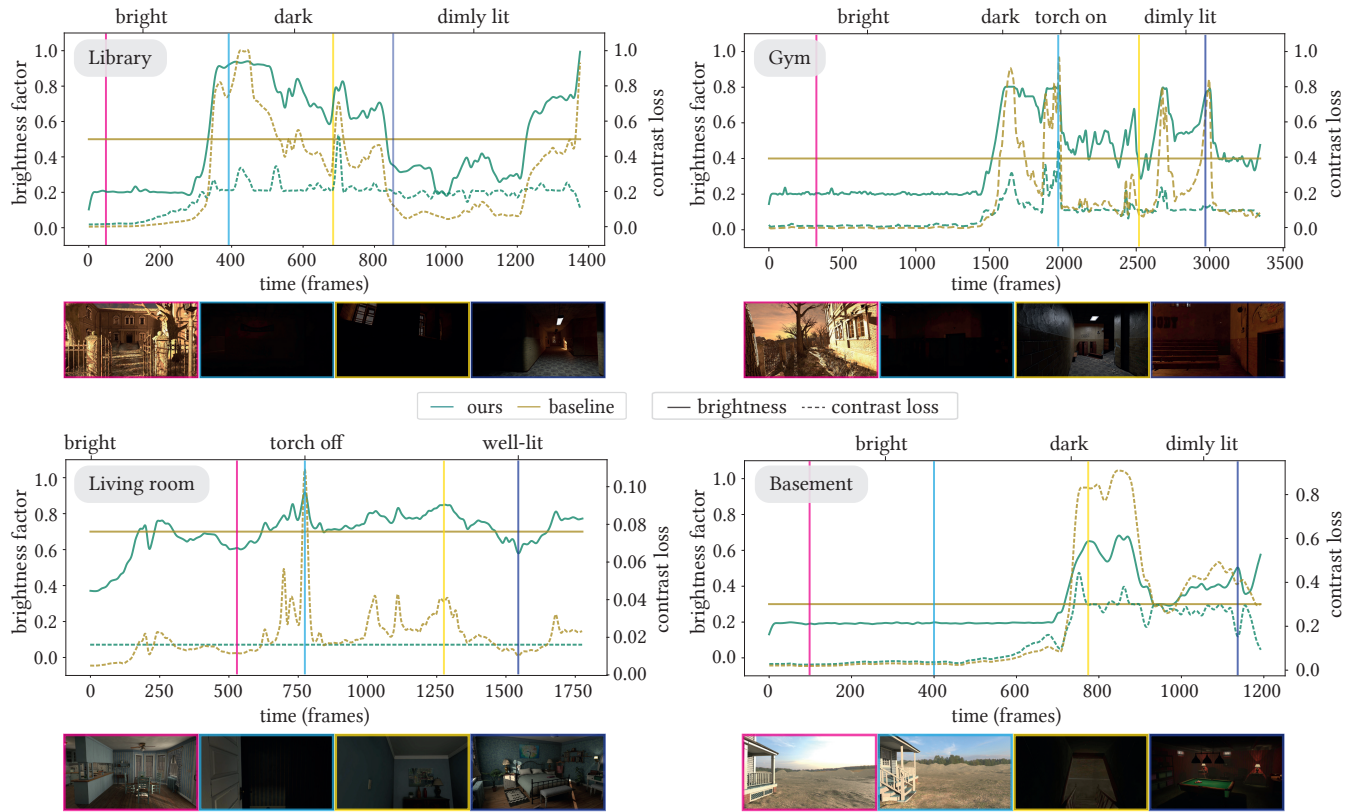


Figure 7: Brightness and loss of contrast plots for the four sequences LIBRARY, GYM, LIVING ROOM and BASEMENT. The sequences target an average power consumption as a fraction of the maximum screen power (0.5 in LIBRARY, 0.4 in GYM, 0.7 in LIVING ROOM and 0.3 in BASEMENT, with 0 being absolute black and 1 the original image without dimming). For each sequence, we annotate labels describing the lighting conditions to provide insight into the algorithm's behavior. As an example, in the LIBRARY sequence, the brightness initially starts at a low level and gradually increases when entering indoor spaces with dimmer lighting. Subsequently, the camera navigates through a corridor with moderate illumination before reaching the dark library shelves. Although the modulation rate appears quite steep, it remains within the limit imposed by the calibration (Section 3.4). Vertical colored lines indicate the timestamps of some example frames shown below.

rings [CC*88] (in the SCHOOL and HOUSE scenes). We select specific keypoints in these sequences, locating the user in these positions, and maintain the predicted brightness modulation from the previous experiment. The objective is to quantitatively measure (via detection time) if our brightness modulation technique enhances the visibility of environmental details compared to the baseline method. We include extensive details on the stimuli and study in the Supplemental.

Results Figure 8 presents the average detection times for each location, along with the standard error of the mean. We can observe that detection times were significantly reduced in low-brightness environments. Furthermore, in bright areas, our modulation did not significantly impact detection time, while saving energy with respect to the baseline via reduced brightness. In rare cases (e.g. *Corridor*), however, our method reduces brightness too aggressively to maintain effective contrast. Although our algorithm achieves consistent contrast losses, it operates globally, which may lead to specific features being compromised, such as the Landolt in this case. This is further discussed in Section 7.

4.3. Imperceptible Brightness Modulation

We also verify that our adjustments in brightness are subtle or imperceptible. While our method primarily focuses on preserving spatial contrast, ensuring temporal coherence is important to prevent artifacts resulting from excessively steep brightness modulation rates. The stimuli utilized in this subjective experiment were identical to those described in Section 4.1.

Study The same 17 participants who took part in the detail preservation experiment also participated in this study. Their task was to identify, using a 2AFC protocol, the stimulus that they believed showed a change in brightness. The two alternatives were OURS (modulated) and BASELINE (constant dimming) for each sequence examined. Before the experiment, participants were shown an example of a brightness change on the VR headset and they were informed that modulation was only present in one of the two alternatives.

Results The probability of selecting our method was reported to be nearly 50% for all sequences (Table 1, last column), indicating no

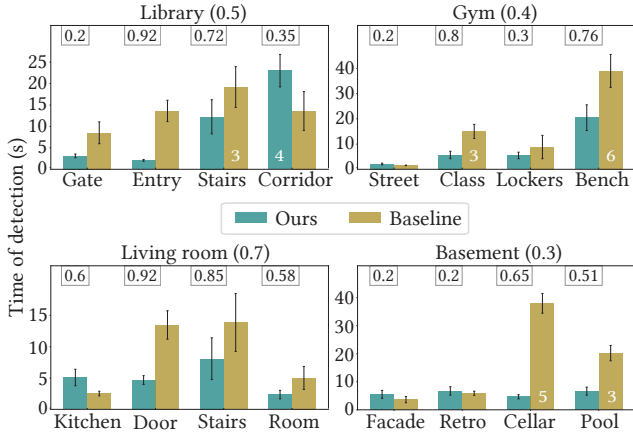


Figure 8: Results of the performance experiment (Section 4.2). For each location in the four sequences, we report the average detection time. The white numbers indicate trials where the ring was undetected, and the number above each column shows the modulated brightness used. The baseline brightness, in brackets, is constant across all locations in each sequence.

Table 1: Results of both the detail quality experiment and the brightness modulation rate experiment (Sections 4.1 and 4.3, respectively). The first two columns show the comparison of quality ratings for perceived details between OURS and the BASELINE, along with standard error of the mean. Higher values indicate better perceived details. The last column is the probability of detecting brightness changes in the sequence where our algorithm, modulating the brightness, is applied. A result of 0.5 indicates imperceptible modulation.

Sequence	Baseline	Ours	% artifacts det.
Library	3.59 ± 0.22	3.76 ± 0.16	53%
Gym	3.23 ± 0.26	3.71 ± 0.19	47%
Living room	3.82 ± 0.19	4.18 ± 0.17	47%
Basement	3.47 ± 0.15	4.29 ± 0.18	53%

substantial perceptual difference between OURS and the BASELINE. This suggests that our method achieves a brightness modulation rate that is imperceptible.

4.4. Ablation on Visual Adaptation

By modeling the temporal behavior of photoreceptors, simulating luminance adaptation, we verify that our modulation aligns with the continuous adaptation process of the human visual system. Using Pattanaik’s method [PTYG00], we model both the fast, symmetric neural adaptation effects and the slower, asymmetric effects resulting from pigment bleaching, regeneration, and saturation. We compute the static response [Hun05], which assumes viewers have achieved a steady-state of adaptation, and the dynamic time-varying response, that eventually reaches the same state given enough time, indicating full luminance adaptation. For our analysis, we focus exclusively on cones, which are primarily re-

Table 2: For each sequence, we report the average difference between the static and dynamic cone responses with (constrained) and without (unconstrained) the temporal change limit ΔL determined in our calibration experiment (Section 3.4). Results confirm that the constrained modulations closely align with the adaptation of the HVS. In the LIVING ROOM sequence, the two differences are very similar to each other, indicating that when the reduction factor is not severe (70% of brightness), our method effectively handles adaptation even in an unconstrained fashion.

Sequence	Brightness factor	Constr.	Unconstr.
Living room	0.7	0.6055	0.6151
Library	0.5	0.4189	0.7311
Gym	0.4	0.4537	0.6286
Basement	0.3	0.3368	1.4984

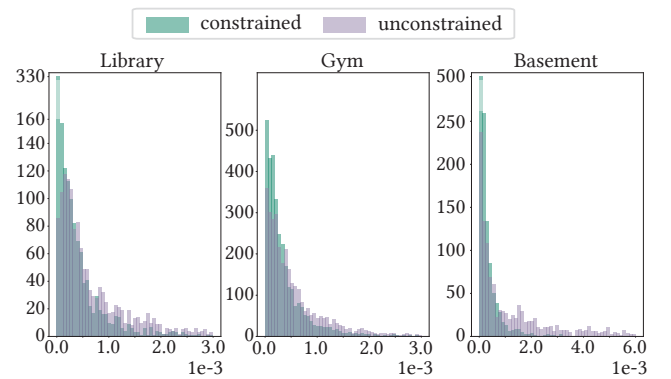


Figure 9: Histogram of the differences between static and dynamic cone responses. Under constrained conditions, the majority of differences are minimal, whereas the unconstrained approach results in larger discrepancies. This demonstrates that our calibration is crucial for maintaining effective visual adaptation throughout the sequences.

sponsible for visual perception when viewing displays under typical daylight conditions. We integrate the values of the cone responses over time and report the results in Table 2. Furthermore, we show the histogram of differences between static and dynamic cone responses in Figure 9.

As we can see, in the tested sequences, the two responses closely align with each other when applying our modulation rate constraint. In contrast, optimizing without this constraint results in a larger divergence between the two responses. This demonstrates that our calibration helps maintaining a continuous visual adaptation of the human eye while watching the displayed content.

5. Measuring Real-life Power Savings

Isolating the power consumption of individual elements in a real VR setup can prove difficult without direct access to the hardware. In order to realistically assess the potential power savings of our approach, and in a similar fashion to closely related works [CP04;

DCT*22], we crafted a representative setup consisting of an off-the-shelf LCD panel (see Supplemental for the technical sheet), a custom screen driver and a power measuring circuit. A detailed schematic can be seen in the Supplemental. To characterize the energy consumption of the display, we measured power consumption at different brightness levels and fit the measurements to establish the relationship between luminance and power consumption, resulting in the following model:

$$\mathcal{P}(L) = 0.001858L + 0.2945 \quad (7)$$

where L is the luminance value; a linear correlation between backlight brightness and power consumption is typical in global-backlight LCD displays. Thanks to the linear correlation, brightness plots throughout our paper can also provide sufficient information on instantaneous power consumption (e.g. Figure 7). This function is then used to compute power consumption per frame, $\mathcal{P}_i$, in our optimization (Equation 6).

In Table 3, we report measured average screen power consumption throughout Section 4.1's experiment. We also report average power consumption when running the sequences without dimming the display (full brightness). Our approach achieves the desired power consumption target while substantially improving perceived quality, as demonstrated before.

Table 3: Average screen power consumption throughout the duration of the sequences running on our setup. As imposed by the optimization, OURS and BASELINE consume the same amount of power.

Sequence	Baseline	Ours	Full brightness
Library	0.959 W	0.975 W	1.800 W
Gym	0.800 W	0.793 W	1.797 W
Living room	1.258 W	1.259 W	1.789 W
Basement	0.646 W	0.644 W	1.789 W

6. Extension to Real-Time rendering

While optimal perceived-quality-aware brightness modulation requires knowledge over future content to be displayed, we can still leverage our contrast-luminance sensitivity model to diminish the perceived impact on quality loss when attempting to reduce screen power consumption. Inspired by classic control theory, we developed an alternative formulation of our algorithm that dynamically adjusts screen brightness based on a user-defined maximum tolerable contrast loss. We implemented a proof of concept on GPU with HLSL and Compute Shaders in Unity, which runs real-time.

Method motivation A key insight from our approach in Section 3 is that it can be seen as closed-loop negative feedback system, where measured difference of contrast losses between subsequent frames influences the brightness of the following frame, attempting to rectify deviations. This conception of the system allows us to leverage classic control theory methods to design optimal real-time brightness modulation. In particular, we deploy a Proportional-Integral-Derivative (PID) controller (Figure 10) in conjunction with

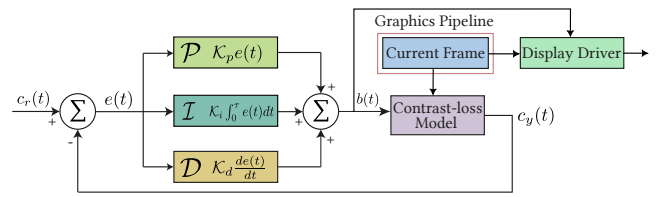


Figure 10: Block diagram of our PID-based control scheme. $c_r(t)$ is the user-defined reference contrast loss, which could be fixed or changed over time. The controller's duty is to achieve this contrast loss by minimizing the error $e(t)$ between the current contrast loss $c_y(t)$ and the desired one. The resulting action from our controller (brightness, $b(t)$) is directly applied to the display driver and is used as well to predict the contrast loss induced by the current displayed frame using the model introduced in Section 3.2. We empirically found $K_p = 0.2$, $K_i = 0.01$, $K_d = 0.05$ to work well for a variety of content.

our previously introduced perceived luminance-contrast sensitivity model, managing average contrast loss near desired values. While directly controlling power consumption would be more desirable, this is unfeasible within our framework in an online/real-time environment, where how the player interacts with the content is outside of our control. In practice, our model saves energy (on average) while maintaining consistent quality; a straightforward alternative (i.e. constant dimming of the screen) attempting to obtain similar energy savings results in degraded quality depending on displayed content (which we specifically test in the *Results* section below).

Implementation details Each rendered frame is first processed with the same pipeline described in Section 3. The algorithm computes the band decomposition and from it the loss of visible contrast with respect to the previous frame. The controller operates with a predefined target contrast loss, adjusting brightness to achieve this target. A higher target contrast loss generally enables more aggressive dimming. With regards to the controller, we can see an overview in Figure 10. The PID's constants K_p , K_i , K_d were empirically determined across varied content to ensure stability and quick reaction time while avoiding large, sudden brightness changes. Broadly, the proportional component K_p regulates the reaction time of the system to changes in the state: larger values will typically allow it to quickly rectify deviations in the target contrast loss; but too large and it can produce significant overshoots, temporal artifacts and even slower response times. The derivative component K_d on the other hand, regulates the instantaneous modulation rate and allows us to constrain it in a similar fashion to our temporal calibration in Section 3.4. The integral component K_i smooths the response ensuring that on average we reach our quality target. We include the values we employed throughout our experiments in Figure 10, and speculate about optimal control schemes in Section 7.

Results To verify that visible contrast is preserved, we run a game session of about three minutes in our environment HOUSE while using our online approach. We computed average brightness and compared it to a baseline method that applies a constant dimming

factor. Targeting a contrast loss of 0.25, our modulation results in an average brightness factor $\bar{b} = 0.52$ and a measured average contrast loss of $\bar{\mathcal{L}}_c = 0.255 \pm 0.119$. The baseline, constantly dimmed with $b = 0.52$, exhibits $\bar{\mathcal{L}}_c = 0.259 \pm 0.182$. Therefore, our control scheme yields slightly lower contrast loss and maintains it more stable throughout the session, improving perceived quality.

We also study the impact in practical power savings of different contrast loss targets. We run the BASEMENT sequence with our real-time approach setting increasingly higher contrast loss tolerances (Figure 11). As expected, higher target contrast losses result in substantially lower power usage. The question of where this threshold should be for losses to remain imperceptible or acceptable remains an open avenue for future work. We also include a comparison between the offline and the real-time approach in the supplementary material.

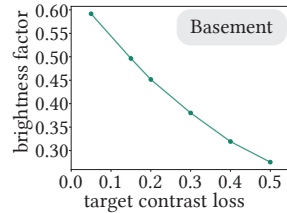


Figure 11: The relationship between target contrast loss and brightness factor.

In terms of performance, our implementation in Unity has an overhead of around ~ 25 ms on a computer with an Intel Core i7-9700K and an NVIDIA RTX 2080. It should be possible however to drastically reduce this overhead to under 5ms by further optimizing the GPU implementation, as shown in similar luminance-contrast models [TAW*19; TMP*23].

7. Limitations and future work

The effectiveness of our method depends on the content. Sequences with little brightness and detail variation will benefit less from our approach. An interesting extension to our method would involve guiding the optimization with additional information about important regions within the sequence. This would enable prioritizing contrast preservation in specific frames while enabling more aggressive dimming in frames without crucial content. Our method can also be combined with contrast enhancement techniques. We did not explore this in our work and view these methods as independent, leaving the combination for future exploration.

The online alternative of our method has not been perceptually validated with subjective experiments involving participants. While promising as a proof of concept, a more careful controller design could substantially improve its behavior and runtime performance. For example, there are many different approaches, e.g., [Eva50], to define optimal control schemes based on a set of constraints (i.e., overshoot, peak transient response time). The challenge in applying them is their requirement for a robust mathematical model of the system. Alternatively, natural image statistics or reinforcement learning could be used to learn the optimal parametrization for our PID-controlled system.

Our experiments used a power consumption model derived for an LCD panel with a uniform backlight. However, this model is not applicable to all display technologies. Our method can still be

used across various technologies; however, the power consumption model should be adjusted, and the exact amount of savings will vary depending on the specific technology. An interesting direction for future work is to extend our method to exploit the local dimming capabilities of displays.

Another exciting venue for future research is exploring saccadic suppression. Similar to other perceptual techniques [SPW*18; GMG*24] that utilize saccadic eye movements or blinks to hide content manipulation, it may be possible to perform quicker updates to display brightness during such events. This could enable more aggressive dimming and, as a result, greater power savings.

8. Conclusions

Power efficiency presents a significant challenge for emerging VR technologies, which demand increasingly higher image quality and performance. In this paper, we have introduced a novel technique for managing image brightness over time to achieve a desired power budget while maintaining uniform temporal contrast loss. Our work enhances detail visibility and perceived quality without introducing perceivable artifacts. Furthermore, we have extensively validated our approach, which has been shown to align with visual adaptation models, through various perceptual experiments and corroborated our power saving claims through a custom hardware proxy. Additionally, by tracing connections to control theory, we have introduced a proof-of-concept real-time formulation of our approach, achieving very promising results and increasing the applicability of our method. We believe our framework could spark further interest in adaptive brightness modulation featuring more robust low-level vision models, and it could be of interest to varying fields ranging from video streaming and playback to VR games and cinema.

References

- [AMW22] ASHRAF, MALIHA, MANTIUK, RAFAEL K, MARTINOVIC, JASNA, and WUERGER, SOPHIE. “Suprathreshold contrast matching between different luminance levels”. (2022) 2.
- [Bar03] BARTEN, PETER GJ. “Formula for the contrast sensitivity of the human eye”. *Image Quality and System Performance*. Vol. 5294. SPIE. 2003, 231–238 3.
- [BB*12] BOITARD, RONAN, BOUATOUCH, KADI, COZOT, REMI, et al. “Temporal coherency for video tone mapping”. *Applications of Digital Image Processing XXXV*. Vol. 8499. SPIE. 2012, 113–122 2.
- [BCTB14] BOITARD, RONAN, COZOT, RÉMI, THOREAU, DOMINIQUE, and BOUATOUCH, KADI. “Zonal brightness coherency for video tone mapping”. *Signal Processing: Image Communication* 29.2 (2014), 229–246 2.
- [BE83] BURT, PETER and EDWARD, EH. “ADELSON, The Laplacian pyramid as a compact image code”. *IEEE transactions on Communications* 31.4 (1983) 3.
- [CC*88] COLENBRANDER, AUGUST, COMMITTEE, CONSILIUM OPHTHALMOLOGICUM UNIVERSALE VISUAL FUNCTIONS, et al. “Visual acuity measurement standard”. *Ital J Ophthalmol* 2.1 (1988), 1–15 6.
- [CKK13] CHO, SUNG IN, KANG, SUK-JU, and KIM, YOUNG HWAN. “Image quality-aware backlight dimming with color and detail enhancement techniques”. *Journal of display technology* 9.2 (2013), 112–121 2.
- [CMK83] CARAYANNIS, GEORGE, MANOLAKIS, D, and KALOUPITSIDIS, NICHOLAS. “A fast sequential algorithm for least-squares filtering and prediction”. *IEEE transactions on acoustics, speech, and signal processing* 31.6 (1983), 1394–1402 5.

- [CP04] CHENG, WEI-CHUNG and PEDRAM, MASSOUD. "Power minimization in a backlit TFT-LCD display by concurrent brightness and contrast scaling". *IEEE Transactions on Consumer Electronics* 50.1 (2004), 25–32 [2](#), [7](#).
- [CSC02] CHOI, INSEOK, SHIM, HOJUN, and CHANG, NAEHYUCK. "Low-power color TFT LCD display for hand-held embedded systems". *Proceedings of the 2002 international symposium on Low power electronics and design*. 2002, 112–117 [2](#).
- [CWM*24] CHEN, KENNETH, WAN, THOMAS, MATSUDA, NATHAN, et al. "Pea-pods: Perceptual evaluation of algorithms for power optimization in xr displays". *ACM Transactions on Graphics (TOG)* 43.4 (2024), 1–17 [1](#), [2](#).
- [CXS16] CHANG, T.-C, XU, S.S.-D, and SU, S.-F. "SSIM-Based Quality-on-Demand Energy-Saving Schemes for OLED Displays". 46 (May 2016), 623–635. DOI: [10.1109/TSMC.2015.2455020](#) [2](#).
- [DCT*22] DUINKHARJAV, BUDMONDE, CHEN, KENNETH, TYAGI, ABHISHEK, et al. "Color-perception-guided display power reduction for virtual reality". *ACM Transactions on Graphics (TOG)* 41.6 (2022), 1–16 [2](#), [8](#).
- [DRE*12] DIDYK, PIOTR, RITSCHER, TOBIAS, EISEMANN, ELMAR, et al. "A luminance-contrast-aware disparity model and applications". *ACM Transactions on Graphics (TOG)* 31.6 (2012), 1–10 [2](#).
- [DZ11a] 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*. 2011, 85–98 [2](#).
- [DZ11b] DONG, MIAN and ZHONG, LIN. "Power modeling and optimization for OLED displays". *IEEE Transactions on Mobile Computing* 11.9 (2011), 1587–1599 [2](#).
- [EMU16] EILERTSEN, GABRIEL, MANTIUK, RAFAL K, and UNGER, JONAS. "Real-time noise-aware tone-mapping and its use in luminance retargeting". *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE. 2016, 894–898 [2](#).
- [Eva50] EVANS, WALTER R. "Control system synthesis by root locus method". *Transactions of the American Institute of Electrical Engineers* 69.1 (1950), 66–69 [9](#).
- [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*. 2002, 218–224 [2](#).
- [GMG*24] GROTH, COLIN, MAGNOR, MARCUS, GROGORICK, STEVE, et al. "Cybersickness Reduction via Gaze-Contingent Image Deformation". *ACM Transactions on Graphics (TOG)* 43.4 (2024) [9](#).
- [Hun05] HUNT, ROBERT WILLIAM GAINER. *The reproduction of colour*. John Wiley & Sons, 2005 [7](#).
- [Kan15a] KANG, SUK-JU. "Image-quality-based power control technique for organic light emitting diode displays". *Journal of Display Technology* 11.1 (2015), 104–109 [2](#).
- [Kan15b] KANG, SUK-JU. "Perceptual quality-aware power reduction technique for organic light emitting diodes". *Journal of Display Technology* 12.6 (2015), 519–525 [2](#).
- [KD06] KEROFISKY, LOUIS and DALY, SCOTT. "Brightness preservation for LCD backlight dimming". *Journal of the Society for Information Display* 14.12 (2006), 1111–1118 [2](#).
- [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 [2](#).
- [KMS07] KRAWCZYK, GRZEGORZ, MYSZKOWSKI, KAROL, and SEIDEL, HANS-PETER. "Contrast restoration by adaptive countershading". *Computer Graphics Forum*. Vol. 26. 3. Wiley Online Library. 2007, 581–590 [2](#).
- [MDK08] MANTIUK, RAFAL, DALY, SCOTT, and KEROFISKY, LOUIS. "Display adaptive tone mapping". *ACM SIGGRAPH 2008 papers*. 2008, 1–10 [2](#), [3](#).
- [Mic95] MICHELSON, ALBERT ABRAHAM. *Studies in optics*. Courier Corporation, 1995 [3](#).
- [MMS06] MANTIUK, RAFAL, MYSZKOWSKI, KAROL, and SEIDEL, HANS-PETER. "A perceptual framework for contrast processing of high dynamic range images". *ACM Transactions on Applied Perception (TAP)* 3.3 (2006), 286–308 [4](#).
- [PDP*19] PAGLIARI, DANIELE JAHIER, DI CATALDO, SANTA, PATTI, EDOARDO, et al. "Low-overhead adaptive brightness scaling for energy reduction in OLED displays". *IEEE Transactions on Emerging Topics in Computing* 9.3 (2019), 1625–1636 [2](#).
- [PMP18] PAGLIARI, DANIELE JAHIER, MACII, ENRICO, and PONCINO, MASSIMO. "LAPSE: Low-overhead adaptive power saving and contrast enhancement for OLEDs". *IEEE Transactions on Image Processing* 27.9 (2018), 4623–4637 [2](#).
- [PS15] 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 (2015), 483–490 [2](#).
- [PTYG00] PATTANAIK, SUMANTA N, TUMBLIN, JACK, YEE, HECTOR, and GREENBERG, DONALD P. "Time-dependent visual adaptation for fast realistic image display". *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. 2000, 47–54 [7](#).
- [SPW*18] SUN, QI, PATNEY, ANJUL, WEI, LI-YI, et al. "Towards virtual reality infinite walking: dynamic saccadic redirection". *ACM Transactions on Graphics (TOG)* 37.4 (2018) [9](#).
- [TAW*19] TURSUN, OKAN TARHAN, ARABADZHIYSKA-KOLEVA, ELENA, WERNIKOWSKI, MAREK, et al. "Luminance-contrast-aware foveated rendering". *ACM Transactions on Graphics (TOG)* 38.4 (2019), 1–14 [3](#), [9](#).
- [TMP*23] TARIQ, TAIMOOR, MATSUDA, NATHAN, PENNER, ERIC, et al. "Perceptually adaptive real-time tone mapping". *SIGGRAPH Asia 2023 Conference Papers*. 2023, 1–10 [9](#).
- [WM14] WANAT, ROBERT and MANTIUK, RAFAL K. "Simulating and compensating changes in appearance between day and night vision". *ACM Transactions on Graphics (TOG)* 33.4 (2014), 1–12 [2](#), [3](#).
- [WMP19] WERNIKOWSKI, MAREK, MANTIUK, RADOŚŁAW, and PIÓRKOWSKI, RAFAL. "Preferred Model of Adaptation to Dark for Virtual Reality Headsets". *MultiMedia Modeling: 25th International Conference, MMM 2019, Thessaloniki, Greece, January 8–11, 2019, Proceedings, Part I* 25. Springer. 2019, 118–129 [4](#).
- [WZMM22] WOLSKI, KRZYSZTOF, ZHONG, FANGCHENG, MYSZKOWSKI, KAROL, and MANTIUK, RAFAL K. "Dark stereo: improving depth perception under low luminance". *ACM Transactions on Graphics (TOG)* 41.4 (2022), 1–12 [2](#).
- [ZDL00] ZENG, WENJUN, DALY, SCOTT, and LEI, SHAWMIN. "Point-wise extended visual masking for JPEG-2000 image compression". *Proceedings 2000 International Conference on Image Processing (Cat. No. 00CH37101)*. Vol. 1. IEEE. 2000, 657–660 [3](#).