Deep Controllable Backlight Dimming for HDR Displays

Lvyin Duan, Demetris Marnerides, Alan Chalmers, Zhichun Lei, and Kurt Debattista

***Abstract*—High dynamic range (HDR) displays with dual-panels are one type of displays that can provide HDR content. These are composed of a white backlight panel and a colour LCD panel. Local dimming algorithms are used to control the backlight panel in order to reproduce content with high dynamic range and contrast at a high fidelity. However, existing local dimming algorithms usually process low dynamic range (LDR) images, which are not suitable for processing HDR images. In addition, these methods use hand-crafted features to estimate the backlight values, which may not be suitable for many kind of images. In this work, a novel deep learning based local dimming method is proposed for rendering HDR images on dual-panel HDR displays. The method uses a Convolutional Neural Network (CNN) to directly predict backlight values, using as input the HDR image that is to be displayed. The model is designed and trained via a controllable power parameter that allows a user to trade off between power and quality. The proposed method is evaluated against seven other methods on a test set of 105 HDR images, using a variety of quantitative quality metrics. Results demonstrate improved display quality and better power consumption when using the proposed method compared to the best alternatives.**

***Index Terms*—High dynamic range, Local dimming, Displays**

I. Introduction

HDR technology is capable of capturing, storing and displaying a much wider dynamic range of luminance compared to the traditional standard or LDR technologies. HDR displays can significantly enhance viewing experiences and has been used in photography, gaming, films, medical and industrial imaging [1] [2]. However, due to the dynamic range limitation of widely available conventional displays, a lot of work has focussed on compressing the dynamic range OF HDR imagery to adapt to these displays [3]. With the development of HDR technologies, displays that can support HDR content are becoming a popular (and increasingly the only) choice for most consumers [4] [5] [6]. LED-based HDR displays, termed dual-panel display, are one of the predominant types of HDR displays available in the consumer electronics market, and the ones that offer the highest dynamic range currently. Such a display is composed of two panels, a backlight panel and a Liquid Crystal Display (LCD) panel, that are used for modulating the backlight luminance and maintaining colour and details respectively. Fig. 1 shows the structure of dual panel LC displays and their three main components: the backlight panel, the diffusion panel and the LC panel. The backlight panel is the lighting source for the LC panel, while the diffusion panel is used for smoothing and dispersing the backlight in order to avoid huge luminance gaps and mismatch between neighbouring pixels. The LC panel filters the backlight to create the three channel image output at a high resolution. Due to adopting backlight dimming (BLD) algorithms, HDR displays of this kind, are capable of presenting a significantly higher contrast ratio and lower power consumption compared to conventional displays, providing an enhanced viewing experience to the consumers.

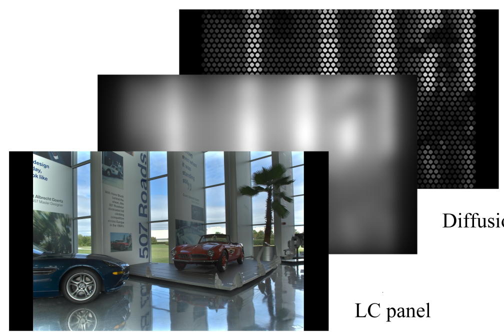


Fig. 1. Structure of LC displays.

BLD algorithms are designed for modulating the backlight of dual-panel displays according to the displayed image content. To date, many BLD algorithms have been proposed [7], which can enhance the contrast ratio and save power consumption to some extent. They can broadly be divided into three categories: statistical-based BLD methods, local characteristics-based BLD methods and optimisation-based BLD methods. Statistical-based local dimming algorithms obtain backlight values using straightforward mathematical operators. For instance, Funamoto et al. [8] proposed the use of maximum and average intensity of a given image segment. The maximum algorithm sets the intensity of each backlight value to the maximum pixel value of the corresponding image segment. The maximum approach is sensitive to noise, while the mean method tends to produce excessively dim backlighting and can lead to significant clipping artefacts. Local characteristics-based BLD methods assign a backlight value that depend on each local segment, rather than taking simple maximum or average values. Cho and Kwon [9] proposed a local characteristics-based BLD method to improve image quality using a correction term to adjust the average pixel intensity

This work was supported by the China Scholarship Council (CSC) under Grant 201806250187. (*Corresponding author: L. Duan*)

L. Duan and Z. Lei are with the School of Microelectronics, Tianjin University, Tianjin, 300072, China. (e-mail: duanlvyin@tju.edu.cn; zclei@tju.edu.cn)

D. Marnerides, A. Chalmers and K. Debattista are with the Warwick Manufacturing Group, University of Warwick, Coventry, CV4 7AL, United Kingdom. (e-mail: dmarnerides@gmail.com; Alan.Chalmers@warwick.ac.uk; k.debattista@warwick.ac.uk)

by considering the local difference between the maximum and average luminance. A similar method was developed by Zhang et al. [10] who also computed a correction term as the ratio of the difference of maximum and average luminance to obtain the backlight values. Lin et al. [11] inversed the cumulative distribution function (from a global histogram) to map a weighted mean of the maximum and average pixel values of each backlight segment for the resulting backlight values. Cho et al. [12] used an image metric to obtain the intensity of the backlight and refined these values by considering both local block lighting and the lighting from neighbouring blocks. Zhang et al. [13] extracted backlight values via calculating the backlight dimming gray (BDG) and obtained high contrast ratio and low power consumption of LCDs. Other BLDs were developed to preserve the image quality, including Kang and Kim [14] who considered the pixel distribution of an image using multiple histograms. Hsia et al. [15] proposed a method to improve the LCD image resolution by enhancing the weak edges of each image segment. In BLD methods, clipping artefacts are the most significant problem that effects the displayed image quality. To keep the balance between displayed image and backlight values, some optimisation-based BLD algorithms have been proposed. For instance, Zhang et al. [16] proposed an optimal method to maintain a balance between LCD image quality and power consumption. Cha et al. [17] presented an efficient optimised BLD method for edge-lit lighting-emitting diode backlight to reduce image quality fluctuation. There have also been other approaches, such as those introduced by Burini et al. [18] and Mantel et al. [19], which focus primarily on achieving a trade-off between clipping and leakage. Later, the authors [20] extended the method proposed by Mantel et al. [19] further to multiple viewers taking into account clipping and leakage as well as reflections of the ambient light. To keep the LCD image quality, Song et.al [21] proposed a pixel compensation algorithm based on deep learning for local dimming algorithms on the quantum-dot display.

The BLD methods introduced above, all target LDR images. To render HDR images on dual-panel displays, Seetzen et al. [22] created a method to solve this problem by splitting HDR images into two layers using square root of the image luminance channel. Later, Zerman et al. [23] proposed a method for HDR image rendering by minimising power consumption and maximising the fidelity to the target pixel values. Narwaria et al. [7] also proposed an HDR image rendering solution by minimising the difference between the theoretical backlight map and the computed light map. However, current BLD methods are mostly designed by display specialists and researchers using hand-crafted features or utilising real-time optimisation, which can be sub-optimal in the first case and may not be suitable for all kinds of images. Furthermore, most BLD algorithms are targeted at displaying LDR images on dual-panel displays to achieve an HDR-like display effect but very few studies have focused on displaying HDR images, and none of these used a CNN. Recently, CNNs have been used for addressing a large range of problems related to luminance processing because of their excellent performance and learning capabilities for analysing image characteristics. Hold-Geoffroy et al. [24] presented a CNN based technique to estimate high dynamic range outdoor illumination. A number of methods using CNNs have also been presented for Tone Mapping (HDR to LDR) and Inverse Tone Mapping (LDR to HDR) [25] [26] [27]. Inspired by the application of CNNs, Jo et al. [28] presented a local backlight dimming based on a CNN which addresses the lack of generalisation ability of hand-crafted features of other BLD algorithms. Later, Zhang et al. [29] proposed a deep CNN-based local dimming technology for dual-modulation display to improve contrast ratios and reduce power consumption. However, these two CNN-based models both address LDR images rather than HDR images. To the best of our knowledge, there are no local dimming methods using CNN architectures for HDR images. The differences between LDR and HDR, in general, are substantial as HDR is capable of handling the full range of lighting in a scene where LDR is considerably curtailed. Algorithms that are designed to work for LDR will not work for HDR. This is why there is a distinction in the literature between LDR and HDR techniques. In particular, there is a significant difference in the method of processing the target image. We remedy this situation by automatically adjusting the BLD values specifically for the HDR content to be displayed via the use of deep learning.

Recently, data driven methods, in particular deep learning, have been used for a wide range image processing applications due to their strong learning and representation capabilities and efficiency. In particular, CNNs form the basis for many current state-of-the-art models in classification, detection, image translation and synthesis [30]. Deep learning methods can bypass human expertise and heuristics by learning directly from data. In this paper, a novel local dimming algorithm based on a CNN architecture is proposed for displaying HDR images on dual-panel HDR monitors. The proposed CNN can efficiently predict the backlight values for each dimming area directly, providing a high-fidelity reproduction of the original content. This is the first paper to use CNNs for HDR backlight dimming, showing that it is possible to do so and that such a method outperforms other methods.

Currently, the high-power consumption of HDR displays remains a significant impediment in the adoption of HDR in many kinds of displays and mobile devices. In addition, low luminance displays with low-power consumption usually cannot offer high enough contrast for supporting HDR content. The proposed HDR local dimming algorithm considers the image quality and power consumption simultaneously via the use of an adaptive parameter that produces high image quality results, surpassing state-of-the-art, at relatively low power consumption; this will assist in the widespread adoption of HDR in consumer devices. From the consumers' point of view, the proposed algorithm provides one user parameter to balance the power consumption and quality, which delivers an improved visual experience and keep a relatively low power consumption by adjusting the user parameter. Results show that the proposed method outperforms other methods in terms of quality whilst maintaining relatively good power consumption at real-time rates.

The primary contributions of this work are: (a) the first learning-based local dimming method that uses a CNN model for rendering HDR images on a dual-panel HDR display; (b) one adaptive parameter, named power parameter, is introduced to adapt the power consumption of the predicted backlight values and suppress the clipping artefact; and (c) to prove the effectiveness of our method in actual consumer electronic device, we implement it on an HDR display. (d) A comprehensive objective evaluation of the proposed algorithm is conducted on an HDR image dataset. Experimental results demonstrate the effectiveness of our method.

II. Method

As discussed in the previous section, a variety of BLD algorithms have been proposed to date. More importantly, most methods are based on modeller expertise [8], with choices that can seem arbitrary and may not be optimal. Furthermore, non learning-based methods can ignore abstract and high level image features that are deemed important in many imaging applications.

The proposed Deep BLD method (DBLD) addresses these issues by using a parametric model to process an input HDR image and directly predict the backlight values. The model is optimised directly from data, avoiding modeller bias and heuristics. The parametric model of choice is a CNN, trained on a dataset of HDR images and optimised to maximise the fidelity of the displayed HDR image and can be controlled via a power parameter, p_a, that provides a balance between power consumption and quality.

As shown in Fig. 2, the proposed DBLD architecture includes four parts: UNet architecture, HDR reconstruction, loss function and optimisation. The UNet architecture is used for extracting luminance features of HDR images combining with the power parameter p_a, and it is optimized by minimizing a loss function. The HDR reconstruction aims to reconstruct the HDR luminance from the backlighting and LCD images.

A. Network Architecture

The CNN used in this work, is based on the UNet architecture [31], which is composed of two main parts, an encoder and a decoder, both composed of multiple convolutional layers. The encoder progressively downsamples the feature resolution until it reaches a low resolution bottleneck, which is then progressively upsampled by the decoder. At each resolution, features from the encoder are propagated directly to the decoder and concatenated, effectively combining multiple scales and speeding up convergence at optimisation. UNet architectures are the de-facto standard CNNs used for a variety of imaging problems. They can process information on multiple scales, have a large receptive field, whilst using lower computational power compared to other architectures due to the use of downsampling, which allows the bulk of the computation to happen on lower resolutions.

The encoder used is a residual network architecture [32] with 18 layers. Residual networks are formed from residual blocks, where the output of the main computation of each block is added to its input, thus allowing better gradient flow and improved training of deeper networks. The implementation is taken directly from the "resnet-18" architecture in the PyTorch model library [33]. The 18-layer resnet architecture is the most lightweight of the commonly implemented residual networks. It downsamples five times and uses 3×3 convolutions, except from the first layer which is of size 7×7 and the residual-connection convolutions that are of size 1×1 and are used to match the input-output feature sizes of each block when they differ.

The decoder consists of five upampling layers that use bilinear upsampling followed by blocks of $\{3 \times 3$ convolution - normalisation - activation - 3×3 convolution$\}$, matching the feature sizes of the encoder at each resolution. The ReLU activation [34] is used both in the encoder and the decoder, along with Instance Normalisationto help with convergence in the optimisation. Instance Normalisation is preferred to the more commonly used Batch Normalisationfor small batch sizes in gradient descent. In this work, the batch size consists of only one image at each iteration due to GPU memory constraints, since training is performed on Full-HD images. The model has a total of 13,782,031 parameters. Despite the large number of parameters, processing is quick, since most of the computation is performed on lower resolutions due to the use of the UNet architecture.

The network accepts a total of four channels of resolution 1,920 $\times$ 1,080, consisting of the RGB channels of the HDR image, I, in the [0, 1] range, along with a uniform single channel that holds the power parameter, $p_a \in [0, 1]$, which adapts the power consumption of the predicted backlight values. The output of the network, $\tilde{B} \in [0, 1]$, is a single channel image containing the backlight predictions at full resolution and is the result of a logistic (sigmoid) function following the final convolution. The final backlight prediction, B, is formed by selecting the $N (N = 2202)$ pixels corresponding to the number of LED lights in the backlight panel of the target display. These are selected as the central pixels of the corresponding areas of the image in $\tilde{B}$.

The final model for the backlight prediction, B, can be expressed as:

$$B(I)_{i,j} = \begin{cases} f_{\text{CNN}}(I, p_a)_{i,j}, & \text{if } (i,j) \in \mathcal{S}, \\ 0, & \text{otherwise}, \end{cases} \tag{1}$$

where $\mathcal{S}$ is the set of centres of the pixel neighbourhoods that correspond to the individual lights in the backlight panel.

B. HDR reconstruction

In theory, the resulting displayed image, $\tilde{I}$ is given by:

$$\tilde{I} = D \odot T, \tag{2}$$

where T is the transmittance of the LC panel, D is the simulation backlight intensity from the diffusion panel . $\odot$ denotes the (pixel-wise) Hadamard product operator, broadcasted channel-wise. The transmittance, T, is driven by the grey level of each pixel from every colour channel of the LCD image.

The diffusion panel output, D, can be estimated from the backlight values as the result of the convolution of the

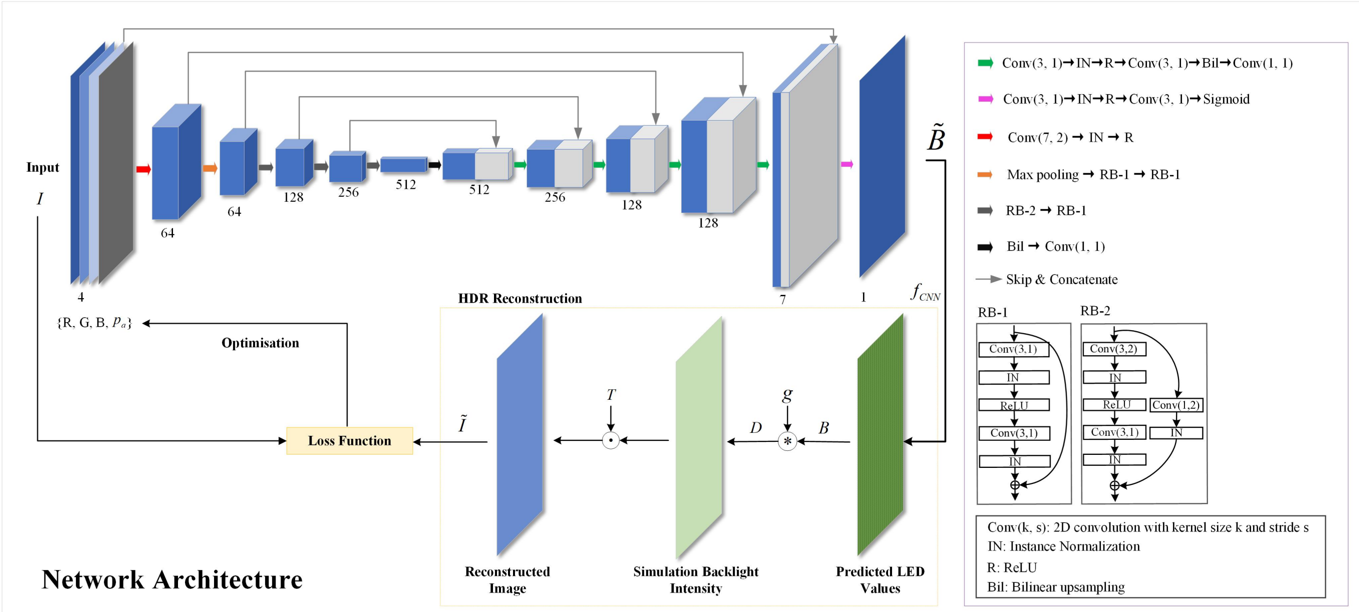


Fig. 2. Diagram of the DBLD CNN architecture.

displayed backlight image, B, with the Point Spread Function (PSF) [35], g, of the diffusion panel:

$$D = (g * B)_{i,j} = \sum_{x=-W_g/2}^{W_g/2} \sum_{y=-H_g/2}^{H_g/2} g_{x,y} B_{i-x,j-y}, \quad (3)$$

where W_g and H_g are the width and height of the PSF filter respectively. D is often referred to as the baseline luminance.

The loss function presented in Section II-C requires the reconstructed HDR image, $\tilde{I}$, which in turn requires evaluation of the baseline luminance, D. D is estimated by convolving the backlight prediction, B, with the PSF, g, following equation 3. However, the PSF for the modelled display is given as a single channel filter of size $1,000 \times 1,000$. Fast differentiable convolution with large filters is not directly implemented (at the time of writing) in modern deep learning libraries [36]. Most libraries optimise small convolutions, e.g. with 3×3 kernels, since almost all CNN architectures use relatively small kernels. Thus, the PSF convolution was implemented from scratch using base (differentiable) PyTorch operations [33].

In particular, the convolution is implemented using the convolution theorem, applied on B and g:

$$D = B * g = \mathcal{F}^{-1}\left(\mathcal{F}(B) \odot \mathcal{F}(g)\right), \quad (4)$$

where $\mathcal{F}$ is the Fourier Transform operator, in combination with the Discrete Fourier Transform (FFT):

$$S_{u,v} = \mathcal{F}(T) = \frac{1}{\sqrt{HW}} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} T(h,w) e^{-2\pi i\left(\frac{hu}{H} + \frac{wv}{W}\right)}, \quad (5)$$

where T is the input in coordinate space and S is the representation of the input in fourier space. H and W are the height and width of the image respectively. The Fourier transform is performed using the Fast Fourier Transform (FFT) algorithm. This implementation for convolutions with large kernels is much faster and uses less memory in contrast to the default optimised convolution based on the cudnn library that would get stuck and not complete the computation on the same machine [37].

C. Loss Function

The loss function, L, consists of two parts, a smooth L_1 regression loss, L_{reg}, and an additional magnitude regularisation term, L_{mag}, that also adapts power consumption by restricting the magnitude of the backlight predictions via the user-provided scalar power parameter, p_a. The smooth L1 loss is chosen as it is robust to outliers, which in this case are from the bright HDR pixels. The total loss is given by:

$$L(\tilde{I}, I) = L_{\text{reg}}(\tilde{I}, I) + p_a \beta L_{\text{mag}}(B), \quad (6)$$

where $\tilde{I}$ is the HDR image reconstructed from the backlight predictions of the model using the method described in Section II-B and I is the target HDR image. β is a hyper-parameter adjusting the magnitude of the regression loss that helps with levelling the gradient contribution of the two partial losses for improved convergence. The magnitude regularisation term, L_{mag}, is given by:

$$L_{\text{mag}}(B) = \frac{1}{M_{\text{max}}} \sum_{(i,j)} B_{i,j}, \quad (7)$$

where M_{max} is the maximum consumption, when all backlights take their maximum value. The magnitude regularisation term restricts power consumption by penalising large backlight values. The nonlearned user-provided power parameter, p_a, appears directly in the loss function, changing the form of the loss during training by adjusting the contribution of the magnitude term L_{mag}. Lower p_a values allow higher L_{mag} values in the loss, thus allowing higher power consumption.

D. Dataset

The training dataset consists of 958 HDR images with varying resolutions, up to 4K. None of the images contain absolute luminance values. The images are scaled keeping their aspect ratio (and zero padded if necessary) to Full-HD (1,920 $\times$ 1,080) resolution. The intensity range is randomly selected during training, with maximum intensity chosen uniformly in the interval [3,000, 5,000]. This random scaling works as a form of data augmentation and to help prevent overfitting. The images are then clipped at the maximum display intensity of 4,000 cd/m^2. The additional power-adaptation scalar is randomly chosen using a uniform $\mathcal{U}[0, 1]$ distribution for each mini-batch. The test dataset used for evaluation is formed from 105 HDR images from the Fairchild Photographic Survey [38]. These images contain calibrated absolute luminance values and are not used during training.

E. Optimisation

The network was optimised until convergence of the loss for approximately 500,000 iterations, with $\beta = 20$. The Adam optimiser [39] was used, with its default learning rate $\lambda = 1e-3$ and $\beta_1 = 0.9$, $\beta_2 = 0.99$. Training took 116 hours on an workstation with a high performance GPU using the PyTorch library.

III. Results

This section presents results comparing DBLD with seven other methods using quantitative analysis and qualitative visual inspection. In particular DBLD is compared against other methods: Avg and Max [8], LP [9], IMF [11], ZR [23], DM [7] and BDG [13].

A. Quantitative evaluation

DBLD is compared with the other methods using the evaluation scheme proposed by Duan et al. [40]. The authors demonstrated that there is a strong correlation between objective and subjective evaluation of different BLD algorithms. Fig. 3 shows the process of quantitative evaluation. A set of 105 HDR images from the Fairchild Photographic Survey database were used in the evaluation process. None of these 105 HDR images were used in the training of DBLD. The original HDR image and the recontructed HDR image are scaled to [0, 4000] to adapt the dynamic range of displays. The metrics used for comparison were the Perceptually Uniform (PU) [41] versions of PSNR, Multi-Scale SSIM [42], along with HDR-VDP-2.2 [43]. These metrics correlate significantly with subjective experiments [40] and therefore are expressive of an improved consumer experience. Higher values of these metrics mean better image quality. The Power Saving Patio (PSR) [44] corresponds to the percentage of power savings with respect to the maximum display power, with higher values representing further savings.

In the proposed method, the weights are chosen so that the gradients are on average of equal contribution/magnitude when training the network. This is a standard practice in the deep learning field. To investigate the effect of power parameters on image quality, we plot the results for the three quality

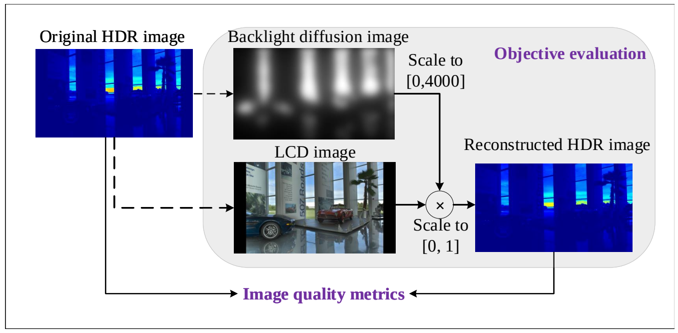


Fig. 3. The process of quantitative evaluation

metrics as a function of power saving ratio, as shown in Fig. 4. For DBLD multiple values are computed by adjusting p_a and can be seen in Fig. 4 as points on the curve. While DBLD was trained using p_a values $\in [0, 1]$, results are also shown for $p_a > 1$ via extrapolation, demonstrating how the method performs for very low power consumption. As can be seen, when the power parameter is in the range of 0.6 to 0.9, as the power saving ratio increases, the values of PU-PSNR and PU-MS-SSIM basically remain constant at 53.75 and 0.99, while the value of HDR-VDP declines slowly from 67.50 to 60.00. When the power parameter is in the range of 0.9 to 1.1, the value of PU-MS-SSIM still remains constant at 0.99, while the values of PU-PSNR and HDR-VDP have an obvious decrease. When the power parameter is in the range of 1.1 to 1.5, the values of PU-PSNR and HDR-VDP show a significant drop, while the value of PU-MS-SSIM decreases slightly in the range of 1.1 to 1.4, then it drops rapidly in the range of 1.4 to 1.5. Furthermore, under most circumstances, other methods are below the curve of the proposed method, demonstrating that DBLD provides better quality as a function of power usage. BDG uses up very little power relatively at the cost of overall image quality. From the analysis above, we can see that the image quality is relatively stable and the power consumption is relatively high when the power parameter is in the range of 0.6 to 0.9. Therefore, the range of 0.6 to 0.9 can be considered as the reasonable value range to balance the relationship between image quality and power consumption.

Fig. 5 illustrates the distribution of results across the 105 tested images for all the methods and the three quality metrics as well as the power saving ratio. As DBLD is adaptable to different outputs depending on p_a, we show distributions with values of p_a fixed to the values of 0.5 (DBLD.50), 0.65 (DBLD.65) and 0.9 (DBLD.90). These values of p_a were chosen to match the power consumption of popular methods. DBLD outperforms all others except for ZR for PU-PSNR and HDR-VDP-2.2, while for PU-MS-SSIM it achieves the first three positions.

B. Visual inspection

This section presents visual inspection results, including the backlight values distribution maps and LCD images, the HDR-VDP-2.2 visibility probability maps and display results via an actual HDR display for all the methods.

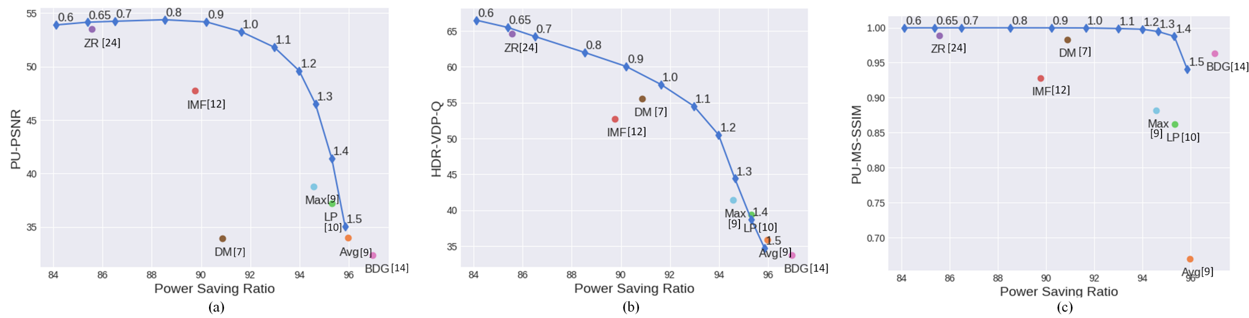


Fig. 4. Comparison of median values of PU-PSNR, HDR-VDP-2.2, and PU-MS-SSIM against PSR. Adjusting p_a allows for the proposed DBLD method (blue line) to adapt power consumption for improved quality.

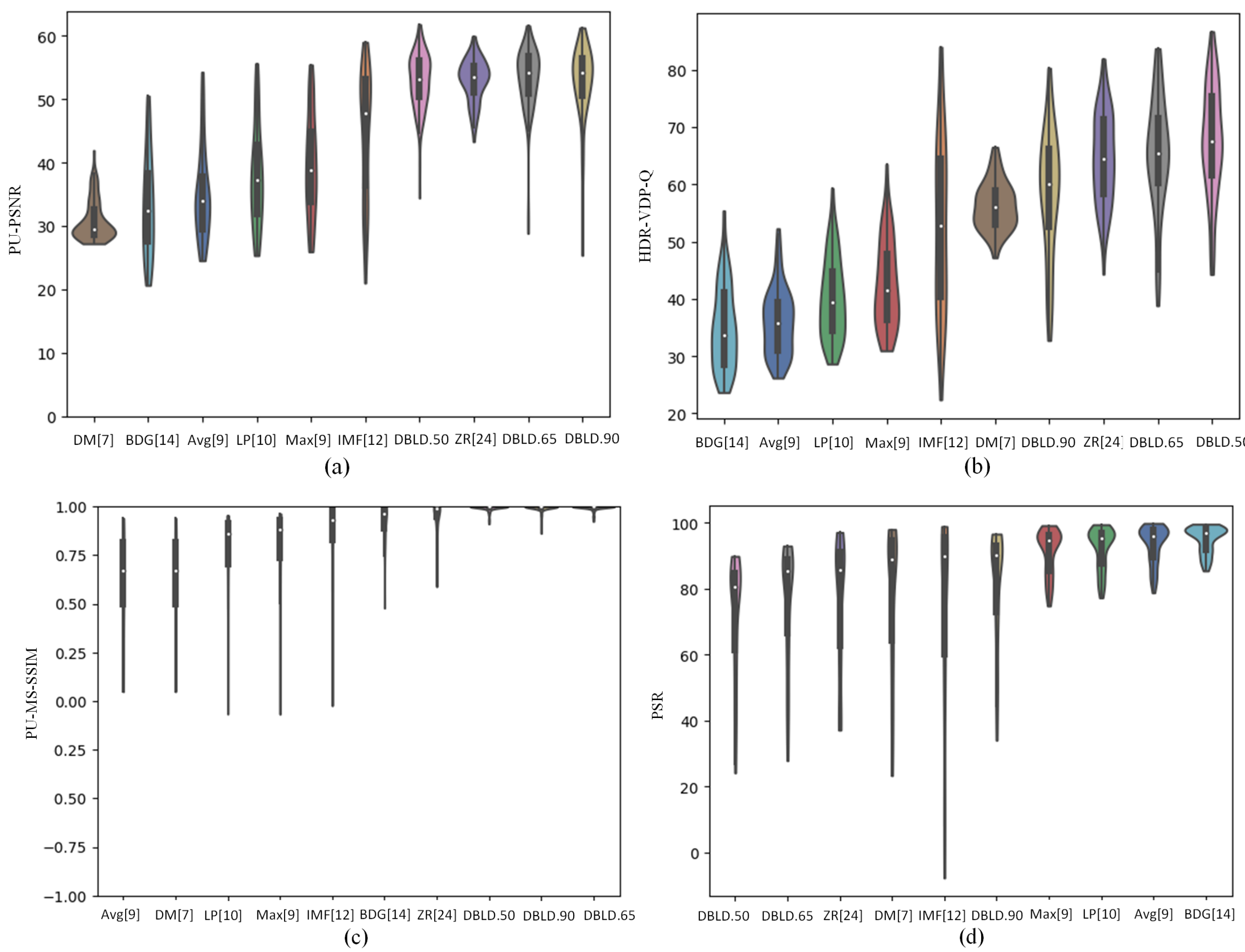


Fig. 5. Comparison of the distributions of PU-PSNR, HDR-VDP-2.2, PU-MS-SSIM and PSR for all methods. The proposed DBLD method is evaluated at different values of p_a (0.5, 0.65 and 0.9).

Fig. 6 shows the LCD images and their backlight values distribution maps for the different BLD methods. Compared with other methods, DBLD with different values of p_a can provide more details and suppress the clipping artefact in the LCD images, especially for the bright area where surrounded by the red boxes. Fig. 7 shows the HDR-VDP-2.2 visibility probability maps for all the methods for a selection of images from the testing dataset. The HDR-VDP-2.2 visibility probability maps describe how likely it is for a difference to be noticed by the average observer, at each pixel, between the reconstructed HDR and the target HDR that is being displayed. Red values indicate high probability, while blue values indicate low probability of noticeable difference. For DBLD, the same values of p_a used in Section III-A are considered. The results

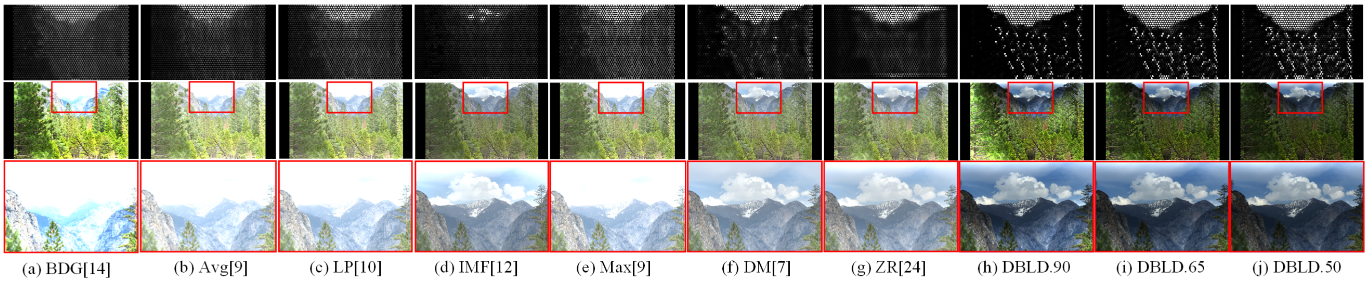


Fig. 6. The backlight values distribution maps and LCD images of KingsCanyou

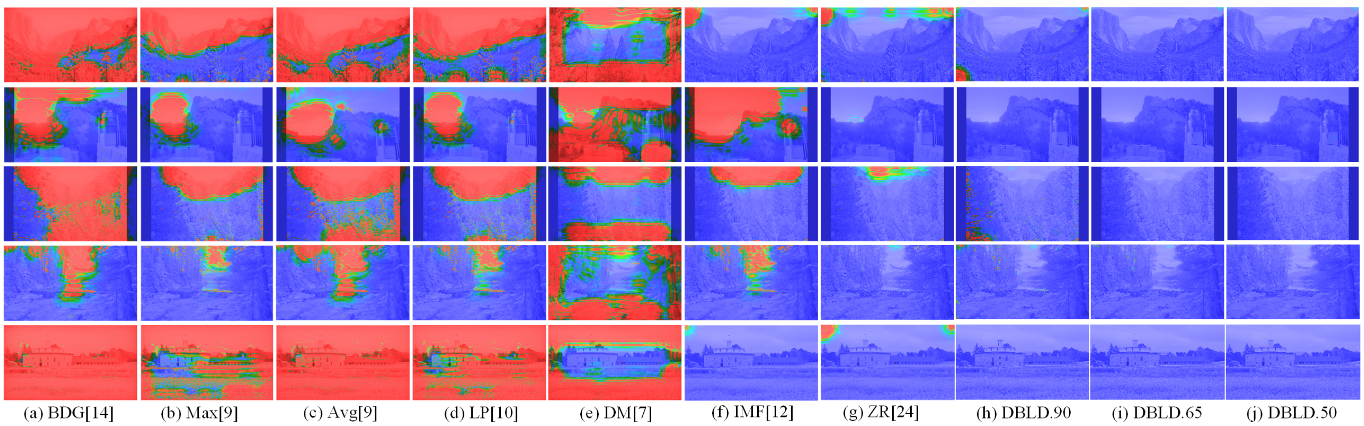


Fig. 7. HDR-VDP-2.2 visibility probability maps for reconstructions of TunnelView(2), MtRushmore(1), KingsCanyon, AmikeusBeaverDamPM2 and HancockSeedField using all methods. Blue indicates imperceptible differences, red indicates perceptible differences.

show that DBLD produces higher fidelity results than the other methods and the number of perceivable artefacts reduces as p_a decreases. In some methods, particularly the BDG, Avg, Max, LP and IMF methods, brighter areas appear overexposed due to the low backlight values. The ZR method can preserve more detail compared to these other methods.

C. Display effects on the HDR display

To examine how the results affect the display, these BLD algorithms were all tested on a state-of-the-art HDR display. This HDR display has 2202 LEDs as backlights and can achieve the maximum luminance with 4,000 cd/m^2. It allows users to control backlight values independently by writing backlight values into the displayed image. Fig. 8 shows the display results with capturing multiple exposures of the display. Due to the high dynamic range of the tested display showing images in the manuscript is difficult as an image at a single exposure cannot exhibit details in different luminance levels. Therefore, for each algorithm, we provide three photos of the display taken at three exposure levels (1/15s, 1/60s and 1/250s). As seen, the DBLD algorithm shows more details than other algorithms under different exposure levels, and its performance is consistent with the quantitative results shown in Section III-A.

D. Timings

DBLD takes an average of 0.061 seconds on the high performance Super GPU to render a Full-HD (1,920×1,080) image. It is worth noting that these are not optimised timings, using the model directly as implemented for training in Python.

IV. Conclusion and Future Work

Recently, HDR images/videos have become one of the most popular media. The efficient reproduction of HDR images on HDR displays is not straightforward, with different methods generating quite different results. High power consumption is also a key issue which can impede the widespread adoption of HDR in the consumer market. A CNN-based local dimming method is described herein to balance image quality and power consumption. The quality of the proposed CNN methods has been demonstrated to be superior than the other methods in our evaluations, and while objective, these evaluations correlate highly with participant-based experiments [40]. The relatively low power consumption and high image quality of our method makes it suitable for consumer devices, especially for mobile devices which have limited battery life and in-vehicle displays where the high contrast in the real-world requires HDR displays and where careful power management is essential. In addition, the method is applicable to current and future HDR displays. In particular, it will help promote future HDR displays with brighter luminance than currently available by providing a balance between power and quality. The results in this paper demonstrate the superiority of the method for potential viewers when compared to the other traditional methods. A limitation of the method is that the controllable parameter is not calibrated in a linear way, for example, being proportional

Fig. 8. The display effects. (1) Exposure time: 1/15s (2) Exposure time: 1/60s (3) Exposure time: 1/250s

to power consumption. This however can be adapted by the display settings, and calibrated against individual displays. Another limitation is that the trained model is trained on a relatively small HDR dataset, which might not cover a wide variety of scenes. This can be addressed by retraining with a larger dataset, as more HDR data becomes available. Future work will focus on further refinement of DBLD and extend it to process HDR videos directly and in real-time.

References

[1] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH Classes*, Los Angeles, CA, USA, Aug. 2008, pp. 1–10, doi:10.1145/1401132.1401174.

[2] C. Marchessoux, L. de Paepe, O. Vanovermeire, and L. Albani, "Clinical evaluation of a medical high dynamic range display," *Med. Phys.*, vol. 43, no. 7, pp. 4023–4031, Jun. 2016, doi:10.1118/1.4953187.

[3] S. Sen Gupta, S. Hossain, and K.-D. Kim, "HDR-like image from pseudo-exposure image fusion: A genetic algorithm approach," *IEEE Trans. Consum. Electron.*, vol. 67, no. 2, pp. 119–128, May 2021, doi:10.1109/TCE.2021.3066431.

[4] M. Azimi, R. Boitard, M. T. Pourazad, and P. Nasiopoulos, "Performance evaluation of single layer HDR video transmission pipelines," *IEEE Trans. Consum. Electron.*, vol. 63, no. 3, pp. 267–276, Aug 2017, doi:10.1109/TCE.2017.014982.

[5] Y. Liu, W. Hamidouche, O. Déforges, and F. Pescador, "A multi-modeling electro-optical transfer function for display and transmission of high dynamic range content," *IEEE Trans. Consum. Electron.*, vol. 63, no. 4, pp. 350–358, Nov 2017, doi:10.1109/TCE.2017.015068.

[6] D. Kobayashi, K. Nakamura, T. Onishi, H. Iwasaki, and A. Shimizu, "A 4k/60p HEVC real-time encoding system with high quality HDR color representations," *IEEE Trans. Consum. Electron.*, vol. 64, no. 4, pp. 433–441, Nov 2018, doi:10.1109/TCE.2018.2876677.

[7] M. Narwaria, M. P. Da Silva, and P. Le Callet, "Dual modulation for LED-backlit HDR displays," in *High Dynamic Range Video from Acquisition to Display and Applications*, F. Dufaux, P. Le Callet, R. K. Mantiuk, and M. Mrak, Eds. London, U.K.: Elsevier Inc., 2016, ch. 14, pp. 371–388, doi:10.1016/B978-0-08-100412-8.00014-0.

[8] T. Funamoto, T. Kobayashi, and T. Murao, "High-picture-quality technique for LCD television: LCD-AI," in *Proc. Int. Disp. Workshops(IDW)*, Kobe, Japan, 2000, pp. 1157–1158.

[9] H. Cho and O.-K. Kwon, "A backlight dimming algorithm for low power and high image quality LCD applications," *IEEE Trans. Consum. Electron.*, vol. 55, no. 2, pp. 839–844, May 2009, doi:10.1109/TCE.2009.5174463.

[10] X.-B. Zhang, R. Wang, D. Dong, J.-H. Han, and H.-X. Wu, "Dynamic backlight adaptation based on the details of image for liquid crystal displays," *J. Disp. Technol*, vol. 8, no. 2, pp. 108–111, Feb. 2012, doi:10.1109/JDT.2011.2165935.

[11] C. Y. Liao, F. C. Lin, H. P. D. Shieh, L. Y. Liao, S. C. Yeh, T. M. Wang, and Y. P. Huang, "Dynamic backlight gamma on high dynamic range LCD TVs," *J. Disp. Technol.*, vol. 4, no. 2, pp. 139–146, May 2008, doi:10.1109/JDT.2008.920175.

[12] S. I. Cho, H.-S. Kim, and Y. H. Kim, "Two-step local dimming for image quality preservation in LCD displays," in *Proc. Int. SoC Design Conf.*, Jeju, South Korea, Nov. 2011, pp. 274–277, doi:10.1109/ISOCC.2011.6138763.

[13] T. Zhang, Y. Wang, H. Wu, M. Li, and Z. Lei, "High-performance local-dimming algorithm based on image characteristic and logarithmic function," *J. Soc. Inf. Disp.*, vol. 27, no. 2, pp. 85–100, 2019, doi.org/10.1002/jsid.740.

[14] S.-J. Kang and Y. H. Kim, "Multi-histogram-based backlight dimming for low power liquid crystal displays," *J. Disp. Technol*, vol. 7, no. 10, pp. 544–549, Oct. 2011, doi:10.1109/JDT.2011.2158985.

[15] J. R. C. Chien, M. H. Sheu, S. K. Wang, and S. C. Hsia, "High-performance local dimming algorithm and its hardware implementation for lcd backlight," *J. Disp. Technol.*, vol. 9, no. 7, pp. 527–535, Mar. 2013, doi:10.1109/JDT.2013.2237755.

[16] T. Zhang, X. Zhao, X. Pan, X. Li, and Z. Lei, "Optimal local dimming based on an improved shuffled frog leaping algorithm," *IEEE Access*, vol. 6, pp. 40 472–40 484, Jul. 2018, doi:10.1109/ACCESS.2018.2858827.

[17] S. Cha, T. Choi, H. Lee, and S. Sull, "An optimized backlight local dimming algorithm for edge-lit LED backlight LCDs," *J. Disp. Technol.*, vol. 11, no. 4, pp. 378–385, Apr. 2015, doi:10.1109/JDT.2015.2401604.

[18] N. Burini, E. Nadernejad, J. Korhonen, S. Forchhammer, and X. Wu, "Image dependent energy-constrained local backlight dimming," in *Proc. 19th IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep. 2012, pp. 2797–2800, doi:10.1109/ICIP.2012.6467480.

[19] C. Mantel, N. Burini, E. Nadernejad, J. Korhonen, S. Forchhammer, and J. M. Pedersen, "Controlling power consumption for displays with backlight dimming," *J. Disp. Technol.*, vol. 9, no. 12, pp. 933–941, Dec. 2013, doi:10.1109/JDT.2013.2260131.

[20] C. Mantel *et al.*, "Viewpoint adaptive display of HDR images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*. IEEE, Beijing, China, Sep. 2017, pp. 1177–1181, doi:10.1109/ICIP.2017.8296467.

[21] S.-J. Song, Y. I. Kim, J. Bae, and H. Nam, "Deep-learning-based pixel compensation algorithm for local dimming liquid crystal displays of quantum-dot backlights," *Opt. Express*, vol. 27, no. 11, pp. 15 907–15 917, May 2019, doi:10.1364/OE.27.015907.

[22] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs, "High dynamic range display systems," *ACM Trans. Graph.*, vol. 23, no. 3, p. 760–768, Aug. 2004, doi:10.1145/1186562.1015797.

[23] E. Zerman, G. Valenzise, F. De Simone, F. Banterle, and F. Dufaux, "Effects of display rendering on HDR image quality assessment," in *Proc. SPIE*, vol. 9599, Sep. 2015, p. Art. no.95990R, doi:10.1117/12.2186674.

[24] Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, and J.-F. Lalonde, "Deep outdoor illumination estimation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR)*, Honolulu, HI, USA, Nov. 2017, pp. 7312–7321, doi:10.1109/CVPR.2017.255.

[25] G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, "HDR image reconstruction from a single exposure using deep CNNs," *ACM Trans. Graph. (TOG)*, vol. 36, no. 6, pp. 1–15, Nov. 2017, doi:10.1145/3130800.3130816.

[26] D. Marnerides, T. Bashford-Rogers, J. Hatchett, and K. Debattista, "Expandnet: A deep convolutional neural network for high dynamic range expansion from low dynamic range content," vol. 37, no. 2, pp. 37–49, May 2018, doi:10.1111/cgf.13340.

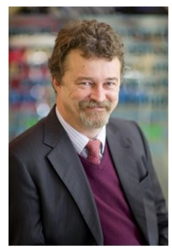
[27] K. Park, S. Yu, S. Park, S. Lee, and J. Paik, "An optimal low dynamic range image generation method using a neural network," *IEEE Trans. Consum. Electron.*, vol. 64, no. 1, pp. 69–76, Feb 2018, doi:10.1109/TCE.2018.2811257.

[28] J. Jo, J. W. Soh, J. S. Park, and N. I. Cho, "Local backlight dimming for liquid crystal displays via convolutional neural network," in *Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Virtual, Auckland, New zealand, Dec. 2020, pp. 1067–1074.

[29] T. Zhang, H. Wang, W. Du, and M. Li, "Deep CNN-based local dimming technology," *Appl. Intell.*, pp. 1–13, May 2021, doi:10.1007/s10489-020-02097-1.
[30] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi:10.1016/j.neunet.2014.09.003.
[31] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Lect. Notes Comput. Sci.*, Munich, Germany, Oct. 2015, pp. 234–241, doi:10.1007/978-3-319-24574-4_28.
[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Comput. Soc Conf. Comput. Vision Pattern Recognit.*, Las Vegas, NV, USA, Dec. 2016, pp. 770–778, doi:10.1109/CVPR.2016.90.
[33] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," in *Adv. neural inf. proces. syst.*, Vancouver, BC, Canada, Dec. 2019, pp. 8024–8035.
[34] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc.Int. Conf. Mach. Learn.(ICML)*, Haifa, Israel, Jun. 2010, doi:10.5555/3104322.3104425.
[35] S. Forchhammer, J. Korhonen, C. Mantel, X. Shu, and X. Wu, "HDR display characterization and modeling," in *High Dynamic Range Video: From Acquisition to Display and Applications*, F. Dufaux, P. Le Callet, R. K. Mantiuk, and M. Mrak, Eds. London, U.K.: Elsevier Inc., 2016, ch. 13, pp. 347–369, doi:10.1016/B978-0-08-100412-8.00013-9.
[36] Github, "Slow convolution with large kernels, should be using fft," https://github.com/pytorch/pytorch/issues/21462(accessed Jul.20, 2020).
[37] S. Chetlur, C. Woolley, P. Vandermersch, J. Cohen, J. Tran, B. Catanzaro, and E. Shelhamer, "cudnn: Efficient primitives for deep learning," *arXiv:1410.0759*, 2014.
[38] M. D. Fairchild, "The HDR photographic survey," *Color Imag. conf.*, vol. 2007, no. 1, pp. 233–238, Jan. 2007.
[39] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Int. Conf. Learn. Represent., ICLR - Conf. Track Proc.*, Y. Bengio and Y. LeCun, Eds., San Diego, CA, USA, May 2015, pp. 1–1.
[40] L. Duan, K. Debattista, Z. Lei, and A. Chalmers, "Subjective and objective evaluation of local dimming algorithms for HDR images," *IEEE Access*, vol. 8, pp. 51 692–51 702, Mar. 2020, doi:10.1109/ACCESS.2020.2980075.
[41] T. O. Aydın, R. Mantiuk, and H.-P. Seidel, "Extending quality metrics to full luminance range images," in *Proc. SPIE*, vol. 6806, Jan. 2008, p. 68060B, doi:10.1117/12.765095.
[42] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Conf. Rec. Asilomar Conf. Signals Syst. Comput.*, vol. 2, Pacific Grove, CA, USA, Nov. 2003, pp. 1398–1402, doi:10.1109/ACSSC.2003.1292216.
[43] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graphics (TOG)*, vol. 30, no. 4, p. 40, Jul. 2011, doi:10.1145/2010324.1964935.
[44] N. Burini, E. Nadernejad, J. Korhonen, S. Forchhammer, and X. Wu, "Modeling power-constrained optimal backlight dimming for color displays," *J. Display Technol.*, vol. 9, no. 8, pp. 656–665, Apr. 2013, doi:10.1109/JDT.2013.2253544.

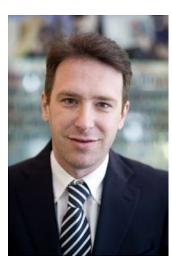
Lvyin Duan received her Ph.D. in Information and Communication Engineering at the Tianjin University. She is currently working in The National Police University for Criminal Justice, China. Her research interests include display technologies, deep learning, HDR image/video processing and medical image processing.

Demetris Marnerides received his BA in Physics at the University of Cambridge, UK in 2013, his MSc in Scientific Computing in 2015 at the University of Warwick, where he also completed his PhD in Engineering in 2019. His research topics include Machine Learning, Computer Vision and HDR Imaging.



Alan Chalmers received a Ph.D. in Computer Science from the University of Bristol, 1991 and an M.Sc. with distinction from Rhodes University, SA, 1984. He is currently a Professor of Visualisation at the University of Warwick. He has successfully supervised 51 PhD students and published over 260 papers in journals and international conferences on high-fidelity virtual environments, multi-sensory perception, and HDR imaging.

Zhichun Lei received a B.S. in Communication Engineering and a M.S in Electronics and System from the Tianjin University, and a Ph.D. in Communication Engineering from the Dortmund University of Technology, Germany, in1998. He is currently a Professor at the Tianjin University, China, and the Ruhr West University of Applied Sciences, North Rhine-Westphalia, Germany. His current research interests include image processing, multimedia technology, and electromagnetic pulse data acquisition and imaging technology.



Kurt Debattista received a B.Sc. in mathematics and computer science, an M.Sc. in psychology, an M.Sc. degree in computer science, and a Ph.D. from the University of Bristol. He is currently a Professor with WMG, at the University of Warwick. His research interests are high-fidelity rendering, HDR imaging, machine learning, and applied perception.