BILATERAL HISTOGRAM EQUALIZATION FOR X-RAY IMAGE TONE MAPPING

Tahani Madmad and Christophe De Vleeschouwer

Institute of Information and Communication Technologies, Electronics and Applied Mathematics
UCLouvain, B-1348 Louvain-la-Neuve, Belgium

ABSTRACT

This paper introduces a novel tone mapping operator, designed to offer a good rendering of the local structures. The new operator fusions the multiple versions of a single HDR input obtained by clipping and normalizing its intensity based on a complete set of disjoint intervals. Defining the weight map associated to each version to be its clipping interval indicator function promotes contrast enhancement, but induces artifacts when neighboring pixels belong to distinct intervals. We thus propose to smooth out the indicators across neighboring pixels with similar intensity, using a standard cross-bilateral filter. With such weight maps, the fusion operator becomes equivalent to applying histogram equalization on the image regions on which the cross-bilateral filter diffuses the indicators, and is therefore referred to as Bilateral Histogram Equalization (BHE) operator. It compares favorably to previous tone mapping algorithms.

***Index Terms*—** Tone-mapping, high dynamic range, X-ray imaging, bilateral filter, histogram equalization

1. INTRODUCTION

High Dynamic Range imaging offers the ability to capture weakly contrasted details, due to a relatively fine quantization of the luminosity range. The visualization of those details however remains challenging given the perceptual capabilities of the human visual system, and the limited amount of intensity levels available on monitors. Therefore, tone mapping constitutes a major element of any high dynamic range imaging pipeline and has been the subject of numerous studies in the last two decades [1]. Tone mapping algorithms are responsible of compressing the extended luminosity range of High Dynamic Range (HDR) images to improve their display on a monitor with a limited dynamic range (LDR). The challenge is therefore to reduce the image bit depth, while preserving subtle and fine-grained local variation in the image. The traditional tone mapping operators (TMOs) have been widely studied and evaluated from a qualitative and quantitative point of view [1, 2, 3, 4], in most cases for natural image processing.

In this paper, we propose a TMO that has been designed for the different applications of X-ray images inspection. In contrast to conventional photography, which measures the light reflected by an object, X-ray imaging measures the amount of photons passing through the objects of interest. As a consequence, X-ray image signals generally mix-up a piecewise smooth component, reflecting the absorptive power of the objects in the scene, with fine-grained local variations, reflecting the texture/shape of the observed objects. The piecewise smooth component is easy to visualize, simply by using a linear mapping between the acquisition and rendering dynamic ranges. Our work primarily aims at visualizing the second component, which is often the most informative regarding the characterization/recognition of the observed objects. Therefore, it proposes a fusion-based strategy to approximate histogram equalization on each smooth segment of the first component. This is done in a seamless manner, without requiring explicit segmentation of the piecewise smooth signal. In short, multiple images, named slices in the following, are derived from the initial HDR image, by clipping and stretching distinct portions/segments of the HDR image histogram. The tone mapping is then defined locally by adapting the weights that are used to sum up those slices. In practice, the weights are spatially modulated to promote, in each pixel, the slices that amplify the local contrast, i.e. that improve the local structure visibility.

The rest of the paper is organized as follows. Section 2 briefly surveys the related works. Section 3 provides a formal description of our fusion-based tone mapping framework. Section 4 then introduces our proposed cross-bilateral filtering strategy to smooth out the weight maps associated to the fusion process, and explains why the resulting fusion-based tone mapping is equivalent to a local histogram equalization. Eventually, Section 5 considers the Tone-Mapped Image Quality Index (TMQI) [5] and the image entropy measure to compare our method to popular tone mapping approaches. Section 6 concludes.

2. RELATED WORKS

Unlike TMOs dedicated to X-ray medical images, our TMO does not exploit any prior information about the content of the HDR image. Previous generic (i.e. without prior) TMO works can be classified into global and local tone mapping operators. Global operators apply a predefined mapping function to all HDR pixel intensities. It means that equal intensity values in the input image are assigned to the exact same level of the dynamic range in the output image. Earliest methods aimed at preserving the brightness of the scene, or the just-noticeable contrast while displaying the scene on the screen [6, 7]. The method proposed by Drago [8] considers a logarithmic transformation of the luminance channel in order to account for the fact that the human visual system is characterized by a logarithmic response to the intensity variations of the real world. Pattanaik et al. [9] went one step further by taking visual adaptation into account. Overall, global operators are appreciated for their low computational cost and for their ability to preserve the global appearance of an image. Their major drawback lies in a poor rendering of weakly contrasted structures in the image [1]. In contrast, local operators process pixel intensities as a function of their local context. Local methods better preserve weakly contrasted details, compared to global approaches. They might suffer from higher computational cost, and poor global consistency of the mapping, inducing visual artifacts [1]. Local methods are usually based on a two-layer decomposition of the image[10]. The first layer is obtained by filtering out

Part of this work has been funded by the Belgian NSF, and the Walloon Region Pôle Mecatech.

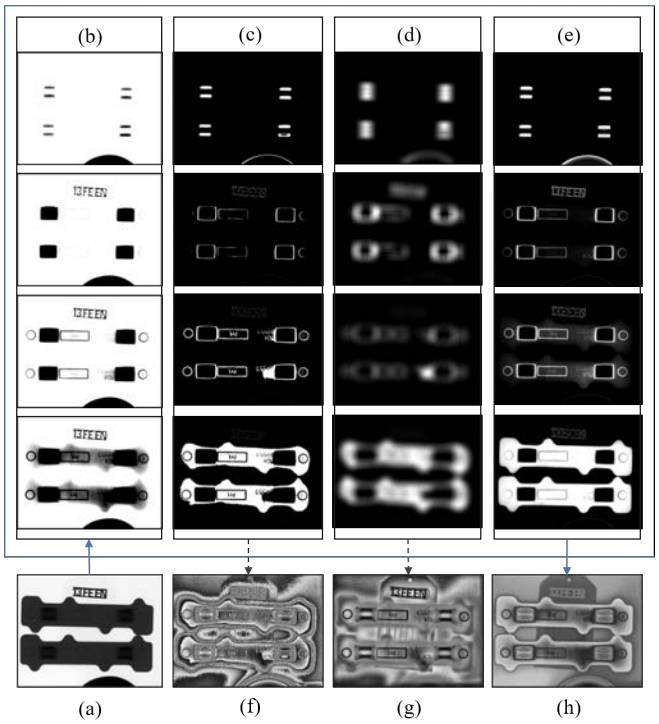

Fig. 1: Our fusion framework. The initial image (a) is split into slices (b). Results (f), (g), and (h) are computed based on Equation (2), with weight maps that respectively correspond to bin indicator functions (c), Gaussian filtered indicators (d), and cross-bilateral filtered indicators (e).

the key (generally low-frequency) structures in the HDR image. The second layer corresponds to the residues, obtained by subtracting the filtered image from the original. Tone mapping then scales down the first layer so as to fit the display dynamic range, while preserving the details by adding the residue to the scaled image. Several variants of this approach have been proposed, using different kinds of filters. Reinhard et al.[11] proposed to adapt spatially the scale of a Gaussian filter to preserve sharp edges in the first layer. Fattal's algorithm detects contrasts by computing the gradient of the image at several scales [12]. The operator of Ashikhmin [13] uses a pyramidal decomposition of the contrasts of the image to keep only contrasts that are visible by the human eye. Durand and Dorsey [14] introduced the non-linear bilateral filter [15] to drastically smooth out the image in regions of limited intensity variation, while preserving sharp edges. In this paper, we introduce a novel tone-mapping approach that builds on image fusion to implement a locally adaptive mapping.

3. FUSION-BASED TONE MAPPING FRAMEWORK

Our approach builds on the fusion of several versions of the original image, named slices, obtained by stretching histogram bins. Specifically, the dynamic range of the initial image is divided into uniform bins, i.e. into a sequence of adjacent non-overlapping intervals of equal size. Each bin is then stretched to amplify the contrast of the image in regions whose pixel values lie in that bin.

Formally, let x be the initial X-ray HDR image, with $0 \leq x(m,n) < 2^N$, $\forall 0 < m \leq H$ and $0 < n \leq W$. The dynamic range is divided in a sequence of K bins, denoted $\{B_k\}_{0<k\leq K}$, with

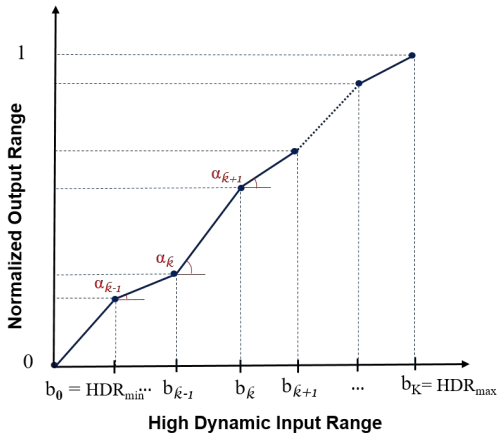

Fig. 2: Tone-mapping function defined by the normalized fusion-based framework in Equation (2). For a given K-dimensional weight vector $(w_1, w_2, ..., w_K)$, the mapping is piecewise linear, and $\alpha_k = w_k \cdot \frac{1}{b_k - b_{k-1}}$, $0 < k \leq K$.

$B_k = [b_{k-1}, b_k]$ and $b_k = k.2^N/K$. A normalized slice image s_k is then associated to each bin B_k as follows:

$$s_k(m,n) = \begin{cases} 0 & x(m,n) \leq b_{k-1}, \\ \frac{x(m,n)-b_{k-1}}{b_k - b_{k-1}} & b_{k-1} < x(m,n) < b_k, \\ 1 & x(m,n) \geq b_k. \end{cases} \quad (1)$$

Adopting a simple and conventional fusion framework, the normalized output image $o(m,n)$ is defined as:

$$o(m,n) = \sum_{k=1}^{K} w_k(m,n) \cdot s_k(m,n), \quad (2)$$

with $w_k(m,n)$ being the k^{th} weight map, associated to the k^{th} slice. To preserve the dynamic range of the output image, the weight maps are defined such that $\sum_k w_k(m,n) = 1$, $\forall(m,n)$. Moreover, in a given pixel, a large weight should be assigned to the slice associated to that pixel, because this slice is the one that stretches the bin associated to the pixel, and therefore amplifies the contrast around that pixel. Hence, a good candidate for the weight map is the k^{th} bin indicator function $i_k(m,n)$, defined as:

$$i_k(m,n) = \begin{cases} 1 & \text{if } b_{k-1} \leq x(m,n) \leq b_k. \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

Unfortunately, as illustrated by the result (f) in Figure 1, defining the weight maps to be equal to the bin indicator functions induces dramatic artifacts, due to the fact that distinct weight vectors, and thus distinct mapping functions (see Figure 2), are assigned to neighboring pixels when those pixels belong to distinct bins. In that case, there is no guarantee regarding the preservation, after the mapping, of the relative ordering of those pixels that are mapped with distinct functions. In order to avoid visual artifacts, typically by preserving the relative ordering of neighboring pixels, it is necessary to maintain similar weight vectors for neighboring pixels. This question is addressed in the next section.

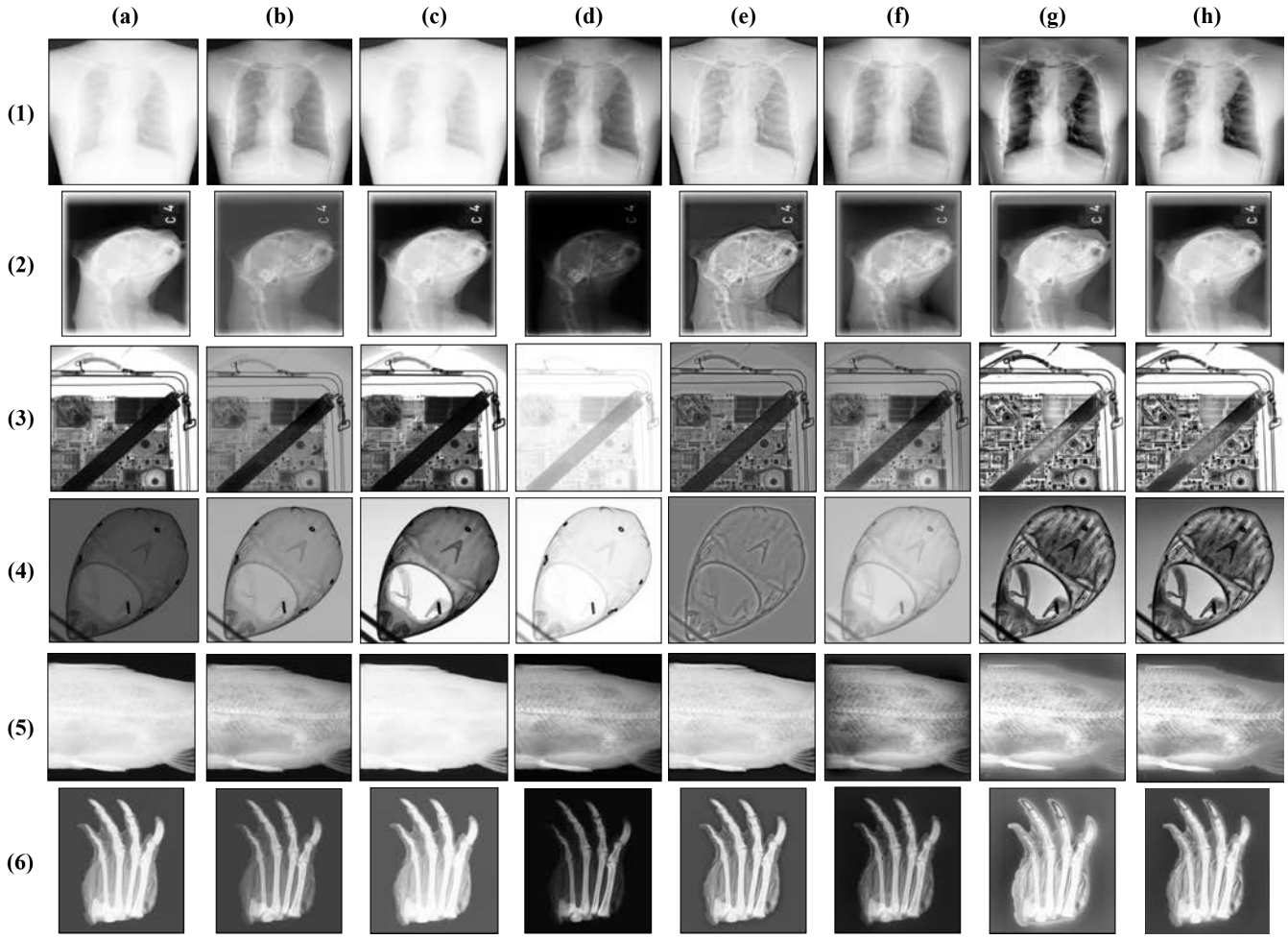

Fig. 3: (cl-)BHE results and comparison for a set of 6 X-ray HDR images from medical, industrial and archaeological applications. (a) Original HDR image, (b) Ashikhmin [13] , (c) Reinhard [11] , (d) Reinhard [16], (e) Durand [14], (f)Fattal [12] , (g)BHE , (h) cl-BHE.

4. BILATERAL HISTOGRAM EQUALIZATION

A natural approach to prevent variation of weight vectors across neighboring pixels consists in filtering the indicator functions with a low-pass Gaussian kernel

$$G_{\sigma_s}(p,q) = \frac{1}{2.\pi.\sigma_s^2}.\exp\frac{-(p^2+q^2)}{2\sigma_s^2}. \quad (4)$$

Figure 1 shows that Gaussian filtering indeed improves the output image quality. However, some ringing artifacts remain along the sharp edges that delimit uniform regions, due to the different mappings associated to the close-to-identical pixels of those regions.

Increasing the spatial support σ_s of the Gaussian kernel will obviously reduce this ringing artifact, but will also impact the spatial adaptation of the mapping, and thus the contrast in the output image. In the extreme case, i.e. when $\sigma_s \to \infty$, the kernel averages the indicator functions over the entire image, for all slices and in all pixels. Hence, in that case, the mapping associated to Equation (2) is the same for all pixel coordinates (m,n), and the weight $w_k(m,n)$ associated to the k^{th} slice is equal to the percentage β_k of image pixels that belong to the corresponding bin. Formally, for a pixel value $v = x(m,n)$ in bin k, the mapping $T(v) = o(m,n)$ is written:

$$T(v) = \sum_{j=0}^{k-1} \frac{\beta_j}{b_j - b_{j-1}} + \frac{v - b_{k-1}}{b_k - b_{k-1}}.\beta_k. \quad (5)$$

This mapping function computes the cumulative distribution of image pixel intensities, and therefore corresponds to a global (quantized) histogram equalization.

As such, averaging the indicator functions has a main drawback: it ends up in defining a global mapping function, which limits the local contrast amplification. This extreme case is however interesting, in that it reveals that diffusing the indicator functions to make them uniform on a spatial region turns the fusion operator defined by Equation (2) into an histogram equalization.

As a central contribution of our paper, we exploit this observation, and propose to control the diffusion of the indicator functions, so as to approximate histogram equalization on groups of pixels that are spatially connected and correspond to similar intensity ranges. Therefore, our work proposes to spread/diffuse the HDR segment indicator functions based on the cross-bilateral filter, also named joint-bilateral filter in the literature [17, 18]. Mathematically, the k^{th} indicator $i_k(m,n)$ is filtered to provide the bilateral weight map

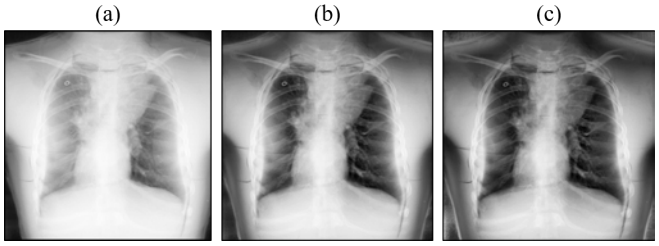

Fig. 4: The effect of the clip limit value on the cl-BHE results. (a) Clip limit = 0.1, (b) Clip limit = 0.2, (c) Clip limit = 1 (i.e no clipping)

$b_k(m,n)$ as follows:

$$b_k(m,n) = \kappa^{-1}(m,n) \cdot \sum_{(p,q)\in\Omega} G_{\sigma_s}(p,q) \cdot G_{\sigma_r}(|x(m,n) - x(p,q)|) \cdot i_k(m+p, n+q), \quad (6)$$

with Ω defining a spatial neighborhood, $G_{\sigma_r}(x) = \exp \frac{x^2}{2\sigma_r^2}$ being the Gaussian range kernel, and $\kappa(m,n) = \sum_{(p,q)\in\Omega} G_{\sigma_s}(p,q) \cdot G_{\sigma_r}(|x(m,n) - x(p,q)|)$ being the normalization constant. This cross-bilateral filtering strategy performs an an-isotropic diffusion of the indicator function, so as to smooth the weight maps while tolerating the assignment of different weight vectors to pixels that are distant from each other in terms of spatial distance or (and this is a specificity) in terms of intensity distance. Similarly to the uniform diffusion with large Gaussian kernel, we observe that the mapping defined by Equation (2) when $w_k(m,n)$ is set to $b_k(m,n)$ is equivalent to (local) histogram equalization when the indicator functions are spread uniformly on a region. This scenario is dominating the the diffusion process since the cross-bilateral filter tend to constraint the indicator functions smoothing within the sharp image edges. Therefore, we name our method Bilateral Histogram Equalization (BHE) method.

Due to its tight connection with histogram equalization, our fusion-based method naturally supports its contrast-limited variant, CLAHE [19, 20]. In BHE, the occurrence frequency of a bin in a pixel neighborhood (i.e. on its local diffusion support) is directly reflected by the bin weight. The BHE bin weight also defines the amplification level of its corresponding slice. Hence, the CLAHE principle, which is supposed to limit contrast amplification, can be transposed to BHE simply by clipping, $\forall(m,n)$, the $b_k(m,n)$ weights at some predefined threshold τ, and by redistributing the overhead equally among all non-zero weights in (m,n). In our experiments, this contrast-limited BHE (cl-BHE) is shown to effectively prevent excessive local contrast in the fusion result.

5. RESULTS AND EVALUATION

To assess our method, we have run our algorithm with small bins (of size 2^9), which gives BHE the opportunity to finely adjust the contrast amplification as a function of the local distribution of intensities. To control the indicator functions diffusion, σ_s has been set to 10 of the image size, while σ_r is set to an intermediate value of 0.3. In practice, the performance of the method remains reasonable constant as long as $\sigma_r \in [0.1, 0.6]$. To ensure a uniform diffusion of the indicator functions in large regions of smoothly changing intensities, iterations of the cross-bilateral filtering might be required. To prevent the computational overhead induced by iterations, we instead aggregate in connected components the neighboring pixels that are very similar (difference $\leq 10^{-5} \times$ HDR range), and distribute indicator functions uniformly on each of those components.

Figure 3 compares (cl-)BHE to a set of well-known tone mapping approaches on a representative set of images, corresponding to a variety of application fields (medical, industrial, security, and archaeology)[1]. This qualitative assessment is completed by the quantitative TMQI [5] and image entropy metrics, respectively provided in Tables 1 and 2. As in [21], we only consider the structural similarity component of TMQI, since the naturalness component is not relevant for X-ray images. We conclude that BHE generally performs better than conventional methods.

To complete our results, Figure 4 shows how (cl-)BHE evolves as a function of the clip limit. We observe that small clip values better preserves the global appearance of the image, at the cost of a reduced the contrast amplification.

Table 1: Structural similarity score of the TMQI metric

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Ashikhmin [13]	0.764	0.801	0.895	0.564	0.762	0.291
Reinhard [11]	0.843	0.894	0.905	0.620	0.661	0.289
Reinhard [16]	0.716	0.674	0.803	0.471	0.769	0.272
Durand [14]	0.717	0.810	0.861	0.519	0.750	0.295
Fattal [12]	0.815	0.856	0.905	0.576	0.775	0.354
BHE	**0.847**	**0.908**	**0.905**	**0.767**	**0.794**	**0.838**
cl-BHE	**0.844**	**0.939**	**0.903**	**0.760**	**0.803**	**0.672**

Table 2: Image entropy score

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Original	6.487	7.679	6.072	6.976	6.156	3.702
Ashikhmin [13]	7.564	6.959	6.092	5.542	6.635	3.278
Reinhard [11]	6.632	7.720	6.032	7.157	5.731	3.820
Reinhard [16]	7.573	6.287	4.971	4.866	6.652	2.587
Durand [14]	7.024	7.437	6.883	4.843	6.435	3.961
Fattal [12]	7.559	7.719	7.130	5.500	7.061	5.439
BHE	**7.782**	**7.767**	**7.206**	**7.708**	**7.097**	**6.953**
cl-BHE	**7.750**	**7.844**	**7.220**	**7.728**	**7.347**	**6.134**

6. CONCLUSION

This paper has introduced an original fusion-based strategy to implement a locally adaptive tone mapping that approximates histogram equalization on a support that is implicitly defined by the diffusion capabilities of a cross-bilateral filter. The method is shown to be competitive with common tone mapping approaches.

7. REFERENCES

[1] Gabriel Eilertsen, Rafal Konrad Mantiuk, and Jonas Unger, "A comparative review of tone-mapping algorithms for high dynamic range video," in *Computer Graphics Forum*. Wiley Online Library, 2017, vol. 36, pp. 565–592.

[2] Gabriel Eilertsen, Jonas Unger, and Rafal K Mantiuk, "Evaluation of tone mapping operators for hdr video," in *High Dynamic Range Video*, pp. 185–207. Elsevier, 2016.

[1] Matlab code, as well as the original HDR images will be made available with the camera-ready version

[3] Philippe Hanhart, Marco V Bernardo, Manuela Pereira, António MG Pinheiro, and Touradj Ebrahimi, "Benchmarking of objective quality metrics for hdr image quality assessment," *EURASIP Journal on Image and Video Processing*, vol. 2015, no. 1, pp. 39, 2015.

[4] Guanghui Yue, Chunping Hou, and Tianwei Zhou, "Blind quality assessment of tone-mapped images considering colorfulness, naturalness, and structure," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3784–3793, 2019.

[5] Hojatollah Yeganeh and Zhou Wang, "Objective quality assessment of tone-mapped images," *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 657–667, 2013.

[6] Greg Ward, "A contrast-based scalefactor for luminance display," *Graphics gems IV*, pp. 415–421, 1994.

[7] Jack Tumblin and Holly Rushmeier, "Tone reproduction for realistic images," *IEEE Computer graphics and Applications*, vol. 13, no. 6, pp. 42–48, 1993.

[8] Frédéric Drago, Karol Myszkowski, Thomas Annen, and Norishige Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," in *Computer Graphics Forum*. Wiley Online Library, 2003, vol. 22, pp. 419–426.

[9] Sumanta N Pattanaik, James A Ferwerda, Mark D Fairchild, and Donald P Greenberg, "A multiscale model of adaptation and spatial vision for realistic image display," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. ACM, 1998, pp. 287–298.

[10] Zhetong Liang, Jun Xu, Diwei Zhang, Zisheng Cao, and Lei Zhang, "A hybrid l1-l0 layer decomposition model for tone mapping," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4758–4766, 2018.

[11] Erik Reinhard, Michael Stark, Peter Shirley, and James Ferwerda, "Photographic tone reproduction for digital images," *ACM transactions on graphics (TOG)*, vol. 21, no. 3, pp. 267–276, 2002.

[12] Zeev Farbman, Raanan Fattal, Dani Lischinski, and Richard Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *ACM Transactions on Graphics (TOG)*. ACM, 2008, vol. 27, p. 67.

[13] Michael Ashikhmin and Jay Goyal, "A reality check for tone-mapping operators," *ACM Transactions on Applied Perception (TAP)*, vol. 3, no. 4, pp. 399–411, 2006.

[14] Frédo Durand and Julie Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *ACM transactions on graphics (TOG)*. ACM, 2002, vol. 21, pp. 257–266.

[15] Carlo Tomasi and Roberto Manduchi, "Bilateral filtering for gray and color images," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 839–846.

[16] Erik Reinhard and Kate Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Transactions on Visualization & Computer Graphics*, , no. 1, pp. 13–24, 2005.

[17] Georg Petschnigg, Richard Szeliski, Maneesh Agrawala, Michael Cohen, Hugues Hoppe, and Kentaro Toyama, "Digital photography with flash and no-flash image pairs," in *ACM transactions on graphics (TOG)*. ACM, 2004, vol. 23, pp. 664–672.

[18] Elmar Eisemann and Frédo Durand, "Flash photography enhancement via intrinsic relighting," in *ACM transactions on graphics (TOG)*. ACM, 2004, vol. 23, pp. 673–678.

[19] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

[20] Karel Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474–485, 1994.

[21] David Völgyes, Anne Martinsen, Arne Stray-Pedersen, Dag Waaler, and Marius Pedersen, "A weighted histogram-based tone mapping algorithm for ct images," *Algorithms*, vol. 11, no. 8, pp. 111, 2018.