The following publication Z. Liang, J. Xu, D. Zhang, Z. Cao and L. Zhang, "A Hybrid l1-l0 Layer Decomposition Model for Tone Mapping," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 4758-4766 is available at https://doi.org/10.1109/CVPR.2018.00500.

A Hybrid ℓ_1-ℓ_0 Layer Decomposition Model for Tone Mapping

Zhetong Liang[1], Jun Xu[1], David Zhang[1], Zisheng Cao[2], Lei Zhang[1,*]
[1]The Hong Kong Polytechnic University, [2]DJI Co.,Ltd
zhetong.liang@connect.polyu.hk, csjunxu@comp.polyu.edu.hk, csdzhang@comp.polyu.edu.hk
zisheng.cao@dji.com, cslzhang@comp.polyu.edu.hk

Abstract

Tone mapping aims to reproduce a standard dynamic range image from a high dynamic range image with visual information preserved. State-of-the-art tone mapping algorithms mostly decompose an image into a base layer and a detail layer, and process them accordingly. These methods may have problems of halo artifacts and over-enhancement, due to the lack of proper priors imposed on the two layers. In this paper, we propose a hybrid ℓ_1-ℓ_0 decomposition model to address these problems. Specifically, an ℓ_1 sparsity term is imposed on the base layer to model its piecewise smoothness property. An ℓ_0 sparsity term is imposed on the detail layer as a structural prior, which leads to piecewise constant effect. We further propose a multiscale tone mapping scheme based on our layer decomposition model. Experiments show that our tone mapping algorithm achieves visually compelling results with little halo artifacts, outperforming the state-of-the-art tone mapping algorithms in both subjective and objective evaluations.

1. Introduction

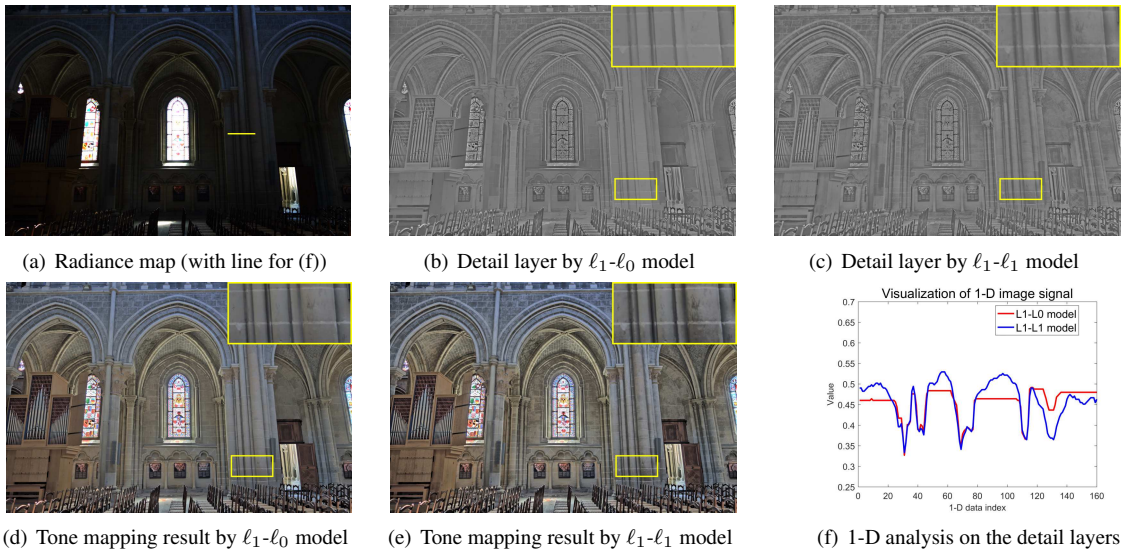
The real-world scenes could span a luminance dynamic range that significantly exceeds the response range of most imaging devices [4]. Thanks to the rapid development of high dynamic range (HDR) techniques in the past decade, the intact information of the scene can be recorded in a radiance map by bracketed exposure fusion technique [2, 7]. However, most of the display devices have a limited dynamic range and are not able to reproduce the information in the radiance map faithfully. Therefore, an effective tone mapping algorithm is needed to transform the HDR radiance map into a standard dynamic range (SDR) image without sacrificing the main visual information.

In the past two decades, a large number of tone mapping methods have been proposed in the literature. Despite the diversity in the design methodology, a large part of these tone mapping methods are based on layer decomposition [8, 14, 23, 29]. Specifically, an image is decomposed into a base layer and a detail layer and then processed separately. The detail layer with fine-grain details is preserved or boosted [8, 14], and the base layer with large spatial smoothness and high range variations is compressed. Although most layer-decomposition-based tone mapping algorithms could increase the visual interpretability of a radiance map to some extent, they still have limitations in obtaining natural and visually pleasing results. A typical problem is the over-enhancement of small scale textural details. This is because the existing works commonly ignore the spatial property of the detail layer, which has a significant impact on the tone mapped image. In addition, halo artifacts are also a problem in some tone mapping algorithms due to the lack of edge-preserving property for the base layer [14]. In order to obtain a natural and artifact-free reproduction of the radiance map, some proper priors must be incorporated into the layer decomposition framework.

Given the fact that a tremendous amount of information is recorded in an HDR radiance map, which part of the information should be assigned a high priority for visual perception is an important question for tone mapping. In psychology, it was found that human vision is more sensitive to edges [1, 13]. This visual mechanism facilitates the capturing of the main semantic information of the scene. In the research of intrinsic decomposition [3, 6], it is commonly assumed that the edges in the reflectance layer (a concept similar to the detail layer) is sparse, which also indicates the high importance of the structural information in an image. In view of the above observations, a tone mapping operator should address the structural reproduction in the first place. Since the spatial property of the detail layer in the layer decomposition framework largely affects the visual appearance of the tone mapped image, we consider to impose a structural sparsity prior on the detail layer.

While the use of spatial prior for detail layer has rarely been reported in tone mapping research, the ℓ_1 sparsity prior has long been adopted in Retinex decomposition [12, 25] to model the structural sparsity of the reflectance layer. Although the ℓ_1 term preserves edges in an image, its piecewise

*This work is supported by HK RGC GRF grant (PolyU 152124/15E), China NSFC grant (no. 61672446) and DJI donation.



(a) Radiance map (with line for (f)) (b) Detail layer by ℓ_1-ℓ_0 model (c) Detail layer by ℓ_1-ℓ_1 model

(d) Tone mapping result by ℓ_1-ℓ_0 model (e) Tone mapping result by ℓ_1-ℓ_1 model (f) 1-D analysis on the detail layers

Figure 1. Results by the proposed layer decomposition.

smoothness nature leads to a weak structural prior. On the other hand, the ℓ_0 sparsity term has shown great piecewise flattening property [34], and it seems to be a better choice for the structural prior.

In this paper, we propose a hybrid ℓ_1-ℓ_0 layer decomposition model for tone mapping. Specifically, an ℓ_0 gradient sparsity term is imposed on detail layer to model the structural prior. In this way, the detail layer will mostly contains structural information, which will be enhanced. Meanwhile, to reduce the halo artifacts, an ℓ_1 gradient sparsity term is imposed on the base layer to preserve edges. A multiscale tone mapping scheme is developed based on our decomposition model. Due to the use of proper priors in our layer decomposition, our tone mapper outperforms state-of-the-art algorithms in both subjective and objective evaluations.

This paper is organized as follows. Section 2 reviews some related work. Section 3 presents the proposed layer decomposition model. Our multiscale tone mapping algorithm is summarized in Section 4. Section 5 and Section 6 are experiments and conclusion, respectively.

2. Related Work

Our work is mainly related to tone mapping, Retinex-based layer decomposition and edge-aware filtering.

Tone mapping. Existing tone mapping algorithms can be categorized into global methods and local methods. Global tone mapping methods reproduce an SDR image with a single compressive curve [28, 32, 33]. In contrast, local tone mapping methods perform this task in a spatially variant manner and are better in detail enhancement. Local methods are commonly based on layer decomposition, where the base layer is first estimated by an edge-preserving filter and detail layer is the residual between base layer and the original image. Different local tone mapping algorithms mainly differ in the filter design techniques. At early stage, kernel-based filters were adopted. Reinhard *et al*. proposed to use a Gaussian-based filter with a spatially adaptive scale parameter [29]. Durand *et al*. adopted a bilateral filter to estimate the base layer [8]. Although this method can avoid halo artifacts to some extent, it over-enhances the image by boosting the small-scale details. Li *et al*. proposed a multiscale wavelet scheme for tone mapping [18]. Meylan *et al*. proposed a Retinex-based adaptive filter for tone mapping [23]. A weighted guided filter for tone mapping is proposed in [14], which also has the over-enhancement problem due to the excessive boosting of small scale details. Global optimization-based filters were also proposed for tone mapping. Farbman *et al*. proposed a weighted least square (WLS) filter [10]. This filter achieves excellent smoothing effect with strong edge-preserving property. Other tone mapping algorithms include globaly linear window method [30] and PCA-based method [17].

While the existing layer-decomposition-based tone mapping methods impose edge-preserving prior on base layer, they show little concern on detail layer. In contrast, our decomposition framework imposes a structural prior on the detail layer to improve the visual quality of the results.

Retinex-based decomposition. Though originally derived from visual constancy study [16], Retinex decomposition estimates the illumination and reflectance from a single image. Retinex decomposition is usually formulated as a variational model with different priors on reflectance and illumination. In the seminal work [15], Kimmel *et al*. proposed an ℓ_2-based Retinex decomposition model for contrast enhancement, where the illumination and reflectance are assumed to be globally smooth. Ng *et al*. assumed that the

reflectance layer is piecewise smoothness and replaced the ℓ_2 norm with a total variation term [25]. Liang *et al*. assumed that the illumination is piecewise-smooth and proposed a nonlinear diffusion based method for illumination estimation [19]. This method preserves edge in the illumination layer and suppresses halo artifact in the result. Recently, Fu *et al*. proposed an ℓ_1 term on the reflectance layer weighted by the luminance reciprocal to model the piecewise constant assumption of reflectance [12].

Edge-preserving smoothing. Edge-aware smoothing is a fundamental technique in image processing. The earliest edge-preserving filter is bilateral filter that considers local range variation of the image [9]. Min *et al*. proposed a fast global smoother based on weighted least square [24]. Other representative filters are ℓ_0-based filter in [34] and weighted-ℓ_1-based filter in [3].

3. Layer Decomposition Method

We first propose a hybrid ℓ_1-ℓ_0 layer decomposition model and give the solver. Then, we extend this decomposition method to a multiscale framework, where different components of an image can be manipulated for tone mapping.

3.1. Hybrid ℓ_1-ℓ_0 Layer Decomposition Model

To devise a suitable layer decomposition framework, we propose to impose the structural prior on the detail layer and the edge-preserving prior on the base layer. Denote by $\boldsymbol{S}$, $\boldsymbol{B}$ and $\boldsymbol{S}-\boldsymbol{B}$ the original image, the base layer, and the detail layer, respectively. The proposed layer decomposition optimization model is given as follows:

$$\begin{aligned}\min_{\boldsymbol{B}} \sum_{p=1}^{N} \Big\{ &(\boldsymbol{S}_p - \boldsymbol{B}_p)^2 + \lambda_1 \sum_{i=\{x,y\}} |\partial_i \boldsymbol{B}_p| \\ &+ \lambda_2 \sum_{i=\{x,y\}} F\big(\partial_i(\boldsymbol{S}_p - \boldsymbol{B}_p)\big) \Big\},\end{aligned} \tag{1}$$

where p is the pixel index, N is the number of pixels in the image. The first term $(\boldsymbol{S}_p - \boldsymbol{B}_p)^2$ forces the base layer to be close to the original image. The spatial property of the base layer is formulated as an ℓ_1 gradient sparsity term $|\partial_i \boldsymbol{B}_p|$, $i = x, y$, where ∂_i is the partial derivative operation along x or y direction. The spatial property of the detail layer is formulated as an ℓ_0 gradient sparsity term with an indicating function $F(x)$:

$$F(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}. \tag{2}$$

The merits of our layer decomposition model lie in the hybrid usage of the ℓ_1 and ℓ_0 regularizations. On one hand, due to the outlier-rejection nature of ℓ_1 sparsity term [20], the large gradients of the base layer are preserved. Thus, the base layer is piecewise smooth. On the other hand, it has been shown that the ℓ_0 sparsity term yields flattening effects [26, 34]. Our model applies ℓ_0 term to force small textural gradients of the detail layer to be zeros, while leaving the main structural gradients intact. This arrangement yields piecewise constant effect and successfully models the structural prior, as demonstrated in Fig. 1(b).

Another possible choice for the detail layer is ℓ_1 gradient sparsity prior, which has been reported in Retinex research [12, 25]. In [12], the ℓ_1 term is imposed on the reflectance/detail layer to gain piecewise constant effect. However, the ℓ_1 term has two drawbacks. First, its nature of piecewise smoothness [21] is not effective enough to produce piecewise constant result, as depicted in Fig. 1(c). Second, under the same parameter setting, the ℓ_1 term cannot strongly regularize the detail layer, which could lead to over-enhancement of the tone mapped image, as shown in Fig. 1(e). To show the difference between the ℓ_1 term and ℓ_0 term, the 1-D profile signals extracted from their resultant detail layers are shown in Fig. 1(f). The position of the signal is indicated by the yellow line in Fig. 1(a). We can see that the ℓ_0 term flattens the small trivial variations and preserves visually important edges, whereas the ℓ_1 term is not effective on this. As a result, the use of ℓ_0 term avoids the over-enhancement problem and increases the visual interpretability of an image, as demonstrated in Fig. 1(d).

3.2. Model Solver

The objective function (1) is nonconvex due to the ℓ_0 norm regularization. We adopt the Alternating Direction Method of Multipliers (ADMM) framework [5] to solve this optimization model. Due to the limited space, we only brief the solving of each subproblem. Please refer to the supplementary material for more detailed description.

For the sake of clarity, we firstly rewrite the objective function (1) in a matrix-vector form as:

$$\min_{\boldsymbol{b}} \tfrac{1}{2}\|\boldsymbol{s}-\boldsymbol{b}\|_2^2 + \lambda_1 \|\nabla \boldsymbol{b}\|_1 + \lambda_2 \mathbf{1}^\top F(\nabla(\boldsymbol{s}-\boldsymbol{b})), \tag{3}$$

where $\boldsymbol{s}, \boldsymbol{b} \in \mathbb{R}^N$ are the concatenated vector form of $\boldsymbol{S}, \boldsymbol{B}$ in (1), respectively, and $\mathbf{1} \in \mathbb{R}^{2N}$ is a vector of all ones. ∇ denotes the concatenation of two gradient operator matrices $\nabla = [\nabla_x^\top, \nabla_y^\top]^\top \in \mathbb{R}^{2N \times N}$. $F(\nabla(\boldsymbol{s}-\boldsymbol{b}))$ performs elementwise non-zero indication and outputs a binary vector. Now two auxiliary variables $\boldsymbol{c}_1, \boldsymbol{c}_2 \in \mathbb{R}^{2N}$ are introduced to replace $\nabla \boldsymbol{b}$, $\nabla(\boldsymbol{s}-\boldsymbol{b})$, respectively. The resultant augmented Lagrangian function of our model is

$$\begin{aligned}\mathcal{L}(\boldsymbol{b}, \boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{y}_1, \boldsymbol{y}_2) = &\tfrac{1}{2}\|\boldsymbol{s}-\boldsymbol{b})\|_2^2 + \lambda_1 \|\boldsymbol{c}_1\|_1 \\ &+ \lambda_2 \mathbf{1}^\top F(\boldsymbol{c}_2) + (\boldsymbol{c}_1 - \nabla \boldsymbol{b})^\top \boldsymbol{y}_1 \\ &+ (\boldsymbol{c}_2 - \nabla(\boldsymbol{s}-\boldsymbol{b}))^\top \boldsymbol{y}_2 \\ &+ \tfrac{\rho}{2}(\|\boldsymbol{c}_1 - \nabla \boldsymbol{b}\|_2^2 + \|\boldsymbol{c}_2 - \nabla(\boldsymbol{s}-\boldsymbol{b})\|_2^2),\end{aligned} \tag{4}$$

where $\boldsymbol{y}_i, i = 1, 2$ are the Lagrangian dual variables. At iteration k, the function (4) is optimized by minimizing several primal sub-problems and maximizing the dual problems alternatively.

(1) **Solving $\boldsymbol{b}^{k+1}$**:

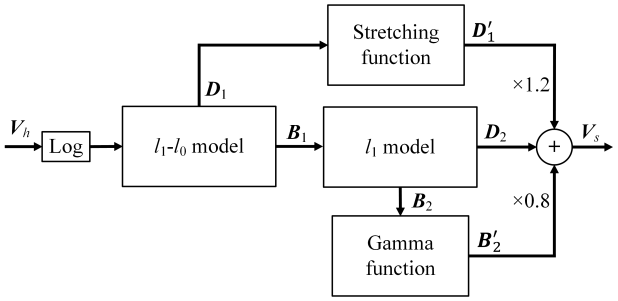


Figure 2. Flowchart of the proposed two-scale tone mapping scheme.

(a) One scale (b) Two scales

Figure 3. Tone mapping results by our model with one scale and two scales. Best viewed on screen with zoom-in.

If we split vector $\boldsymbol{c}_1^k$ into two equal-length pieces $\boldsymbol{c}_{1,1}^k$ and $\boldsymbol{c}_{1,2}^k$, split $\boldsymbol{c}_2^k$ into $\boldsymbol{c}_{2,1}^k$ and $\boldsymbol{c}_{2,2}^k$, split $\boldsymbol{y}_1^k$ into $\boldsymbol{y}_{1,1}^k$ and $\boldsymbol{y}_{1,2}^k$, and split $\boldsymbol{y}_2^k$ into $\boldsymbol{y}_{2,1}^k$ and $\boldsymbol{y}_{2,2}^k$, the objective function with respect to $\boldsymbol{b}^{k+1}$ is a quadratic programming problem, which can be solved efficiently via FFT transformation (Place refer to our supplementary material).

(2) **Solving** $\boldsymbol{c}_1^{k+1}$:

The objective function with respect to $\boldsymbol{c}_1^{k+1}$ can be solved via soft shrinkage

$$\boldsymbol{c}_1^{k+1} = \mathcal{T}_{\lambda_1/\rho^k}(\nabla \boldsymbol{b}^{k+1} - \boldsymbol{y}_1^k/\rho^k), \tag{5}$$

where $\mathcal{T}_\alpha(x) = \text{sign}(x) \cdot \max(|x| - \alpha, 0)$ is the soft-thresholding function.

(3) **Solving** $\boldsymbol{c}_2^{k+1}$:

According to the analysis of [34], the objective function with respect to $\boldsymbol{c}_2^{k+1}$ can be solved in a per-entry manner. This amounts to solving N independent scalar functions. Denote by subscript j the jth entry of a vector. The solution of $\boldsymbol{c}_2^{k+1}$ at entry j is

$$\boldsymbol{c}_{2,j}^{k+1} = \begin{cases} 0, & \text{if } (\boldsymbol{f}_j^k)^2 \le \frac{\lambda_2}{\rho^k} \\ \boldsymbol{f}_j^k, & \text{Otherwise} \end{cases}, \tag{6}$$

where

$$\boldsymbol{f}_j^k = \left(\nabla(\boldsymbol{s} - \boldsymbol{b}^{k+1}) - \boldsymbol{y}_2^k/\rho^k\right)_j, j = 1, ..., 2N. \tag{7}$$

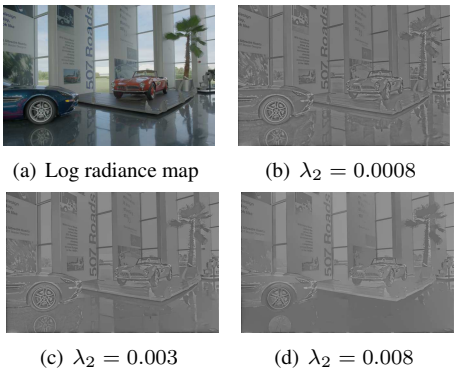
(4) **Dual ascent** for Lagrangian multipliers.

(5) **Update** ρ^{k+1} as $\rho^{k+1} = 2\rho^k$.

The ADMM is efficient to find the approximate solution for the base layer $\boldsymbol{B}$ variable within a few iterations (15 in our case). After $\boldsymbol{B}$ is obtained, the detail layer can be calculated by $\boldsymbol{S} - \boldsymbol{B}$.

3.3. Extension to Multiscale Decomposition

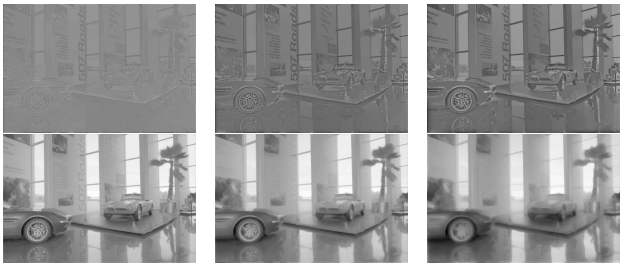
By applying the hybrid ℓ_1-ℓ_0 decomposition model (1) to the radiance map, we can produce a piecewise constant detail layer and a piecewise smooth base layer. While this single-scale scheme endows a standard framework for tone mapping, applying the decomposition to the base layer repeatedly leads to a multiscale decomposition, which can further improve the tone mapping results. In this way, different attributes of an image, represented by different scale layers, can be differently manipulated, which leads to a more flexible and effective tone reproduction. By leveraging the efficiency and effectiveness, we adopt a two-scale decomposition scheme for tone mapping, as depicted in Fig. 2. It will produce a scale-1 detail layer $\boldsymbol{D}_1$, a scale-2 detail layer $\boldsymbol{D}_2$ and a scale-2 base layer $\boldsymbol{B}_2$.



(a) Log radiance map (b) $\lambda_2 = 0.0008$

(c) $\lambda_2 = 0.003$ (d) $\lambda_2 = 0.008$

Figure 4. The effect of λ_2 on the detail layer when λ_1 is fixed to 0.3. The MLE values of (b)(c)(d) are 2.33, 1.55, and 0.97, respectively.



(a) $\lambda_1 = 0.06$ (b) $\lambda_1 = 0.3$ (c) $\lambda_1 = 1$

Figure 5. The effect of λ_1 on the two layers when λ_2 is fixed to $0.01\lambda_1$. The top three images are detail layer. The bottom three images are base layers. The MLE values of the base layers in (a)(b)(c) are 2.27, 2.18 and 2.09, respectively.

As discussed in Section 3.1, the spatial property of $\boldsymbol{D}_1$ largely affects the tone mapped image. We apply the proposed ℓ_1-ℓ_0 model (1) to the first scale decomposition:

$$\begin{aligned} \boldsymbol{B}_1 &= \text{model}_{\ell_1,\ell_0}(\boldsymbol{S}), \\ \boldsymbol{D}_1 &= \boldsymbol{S} - \boldsymbol{B}_1, \end{aligned} \tag{8}$$

where $\text{model}_{\ell_1,\ell_0}(\cdot)$ is the optimization model in (1). After the first level decomposition, the structural information remains in the detail layer $\boldsymbol{D}_1$ and the main textural information is transfered to the base layer $\boldsymbol{B}_1$.

For the second scale decomposition, a simplified model (1) is applied to $\boldsymbol{B}_1$, where the weight λ_2 of the ℓ_0 term is set to 0, leading to a total variation problem:

$$\boldsymbol{B_2} = arg\min_{\boldsymbol{B}} \sum_{p=1}^{N} \left\{ (\boldsymbol{B}_{1,p} - \boldsymbol{B}_p)^2 + \lambda_3 \sum_{i=\{x,y\}} |\partial_i \boldsymbol{B}_p| \right\},$$
$$\boldsymbol{D}_2 = \boldsymbol{B}_1 - \boldsymbol{B}_2. \tag{9}$$

This simplification is based on the fact that we aim to preserve the textural information of the image in the scale-2 detail layer $\boldsymbol{D}_2$. Thus, the ℓ_0-based structural prior is not applicable in this scale of decomposition. As a result, the layer $\boldsymbol{D}_2$ stores the majority of the textural information, and the layer $\boldsymbol{B}_2$ contains local mean brightness.

To summarize, our two-scale decomposition scheme produces three layers $\boldsymbol{D}_1$, $\boldsymbol{D}_2$ and $\boldsymbol{B}_2$, which satisfies:

$$\boldsymbol{S} = \boldsymbol{D}_1 + \boldsymbol{D}_2 + \boldsymbol{B}_2. \tag{10}$$

Fig. 3 shows the tone mapping results of our model with 1 scale and 2 scales (The details of our tone mapping algorithm will be discussed in Section 4). It can be seen that while the one-scale result is acceptable, the two-scale result preserves better the medium frequency component of an image and achieves more natural appearance.

Acceleration. The accuracy of the second scale decomposition (9) is not strictly required. Thus, we adopt an acceleration scheme. First, we linearly downsample the $\boldsymbol{B}_1$ layer by a factor of 4. Then the decomposition model in (9) is performed to get a low resolution image of $\boldsymbol{B}_2$, followed by a linear upsampling to the original resolution. Because the boundary regions in the image is slightly blurred due to the sampling scheme, we finally perform a fast joint bilateral filtering of $\boldsymbol{B}_2$ with the original $\boldsymbol{B}_1$ as the guidance image to recover the sharp boundary information [27].

4. Tone Mapping

Based on the outputs of the proposed layer decomposition, a tone mapping algorithm is developed, whose major steps include color transformation, multiscale decomposition, detail layer boosting, base layer compression, and recombination of the layers. While this framework is common in the tone mapping research, our approach differs in two aspects. First, our suit of layer decomposition models is discriminative in the spatial attributes of an image. As described in Section 3.3, our multiscale decomposition deploys the structural information, textural information and local mean brightness separately into different layers, whereas existing multiscale models merely perform progressive smoothing [10, 14]. Second, in our multiscale manipulation approach, we perform a layer-selective nonlinear processing, whereas other works only perform linear intensity scaling [10].

Since the dynamic range of an image is mostly embedded in the brightness domain, our core algorithm only processes the luminance channel and preserves the chromaticity components. Specifically, the input RGB radiance map is transformed to HSV space and only the V channel is tone mapped. At the reverse transformation stage, the saturation channel is multiplied by 0.6 to prevent from oversaturation.

Our tone mapping algorithm on the luminance channel of an radiance map is depicted in Fig. 2. The channel $\boldsymbol{V}_h$ of the radiance map is firstly converted to log domain and normalized to the range of (0, 1). This step mimics the response of human vision to the luminance and preliminarily reduces the dynamic range. Then our two-scale decomposition model using (8) and (9) is applied, yielding three layers $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{B}_2$. Since the base layer $\boldsymbol{B}_2$ can be considered as the local brightness level of the image, we compress it by a gamma function:

$$\boldsymbol{B}_2' = L \cdot (\tfrac{\boldsymbol{B}_2}{L})^{\frac{1}{\gamma}}, \tag{11}$$

where L is the largest brightness level ($L = 1$ in our case, due to the normalization). For the first-scale detail layer $\boldsymbol{D}_1$, we use a nonlinear stretching function to boost it:

$$\boldsymbol{D}_1' = \operatorname{sign} \boldsymbol{D}_1 \cdot \left(\frac{|\boldsymbol{D}_1|}{\max(|\boldsymbol{D}_1|)} \right)^{\alpha} \cdot \max(|\boldsymbol{D}_1|). \tag{12}$$

This function with the parameter α has a stretching effect for signals centering at 0. Smaller α yields larger stretching degree and vice versa. Since the structural prior is imposed on $\boldsymbol{D}_1$ by the decomposition model (1), the structural residual of the original image is boosted by the stretching function. This arrangement would result in a more visually appealing image. Then, a luminance SDR image is reconstructed by

$$\boldsymbol{V}_s = 1.2\boldsymbol{D}_1' + \boldsymbol{D}_2 + 0.8\boldsymbol{B}_2'. \tag{13}$$

Finally, the values of $\boldsymbol{V}_s$ at 0.5% and 99.5% intensity level are mapped to 0 and 1, respectively. Values out of this range are clipped.

5. Experiments and Analysis

This section presents several experiments to verify the performance of our hybrid ℓ_1-ℓ_0 layer decomposition model (1) and the proposed tone mapping algorithm. A HDR database with 40 radiance maps is collected from various sources[1] [2] for evaluation. These 40 images cover both indoor and outdoor scenes, with different types of objects, including plants, cars, sky and buildings.

5.1. Parameter Selection

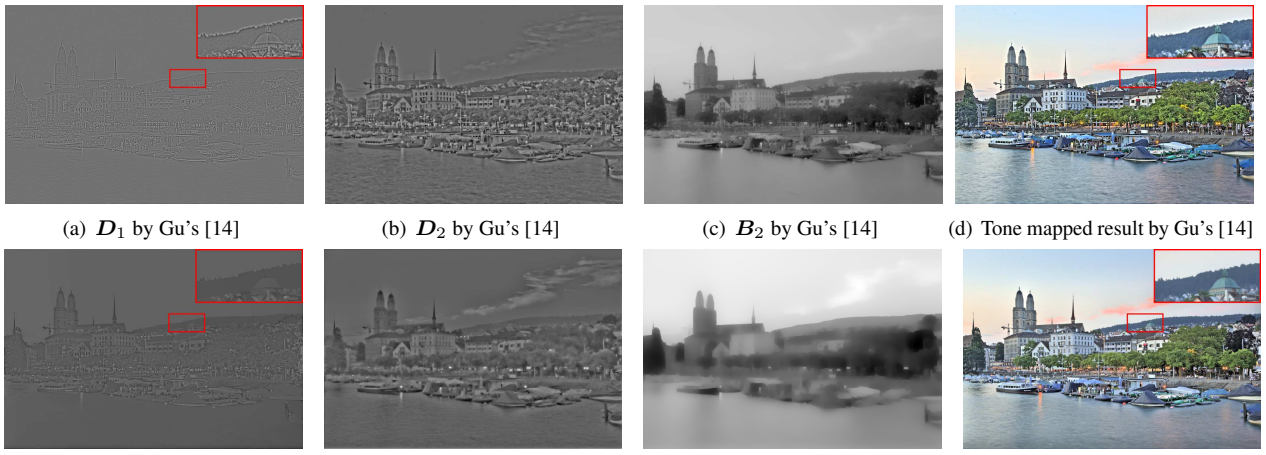
The major parameters that affect our ℓ_1-ℓ_0 decomposition model (1) are λ_1, and λ_2, which control the smoothness degree on the base layer and the detail later, respectively. In the following, along with visual inspection we use mean local entropy (MLE)[3] to objectively measure the smoothness of the two layers. Larger MLE indicates lower smoothness degree, i.e., more textures in the image, and vice versa.

Fig. 4 shows the effect of λ_2 on the detail layer when λ_1 is fixed. It can be seen from the graph that different values of λ_2 lead to different degrees of flattening/smoothness effect

[1] http://pfstools.sourceforge.net/hdr_gallery.html

[2] http://rit-mcsl.org/fairchild/HDR.html

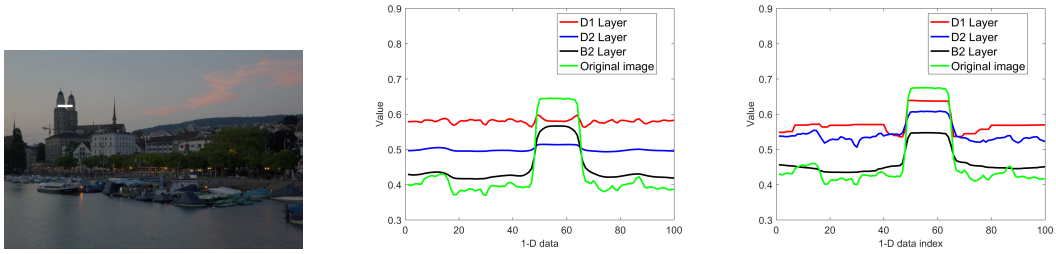
[3] We first calculate the local entropy in each 9×9 window and then average these entropy values.



(a) $\boldsymbol{D}_1$ by Gu's [14] (b) $\boldsymbol{D}_2$ by Gu's [14] (c) $\boldsymbol{B}_2$ by Gu's [14] (d) Tone mapped result by Gu's [14]

(e) $\boldsymbol{D}_1$ by Ours (f) $\boldsymbol{D}_2$ by Ours (g) $\boldsymbol{B}_2$ by Ours (h) Tone mapped result by ours

Figure 6. Comparison between the multiscale decomposition models in [14] and ours.



(a) Log radiance map (b) Gu's model [14] (c) Ours

Figure 7. 1-D profile signal analysis of multiscale decomposition. The location of the profile signals is indicated by the white line in (a).

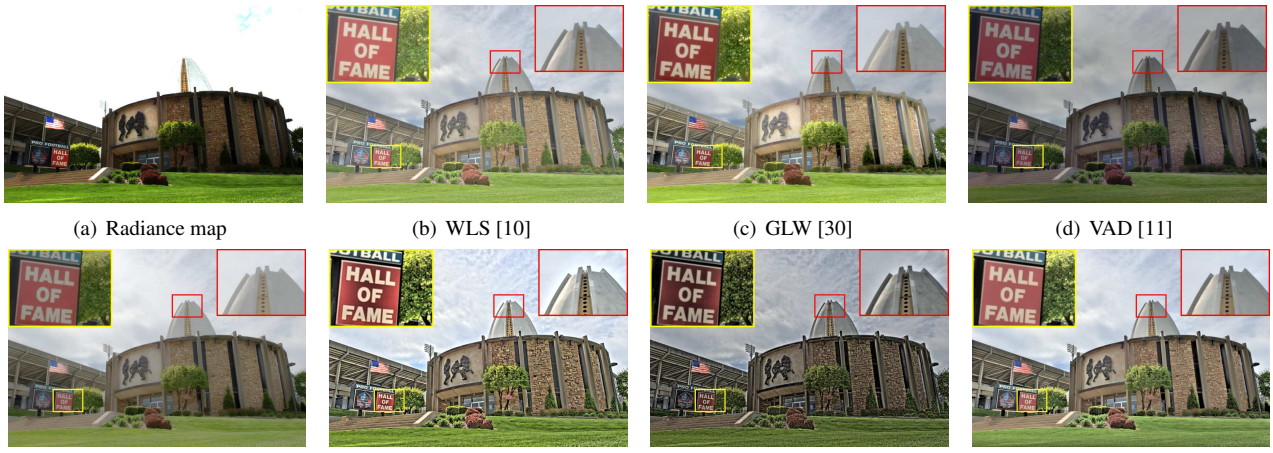
on $\boldsymbol{D}_1$. When λ_2 is excessively large (0.008), some structures are totally flattened, resulting in a low MLE (0.97). In contrast, when λ_2 is too small (0.0008), some small texture gradients appear in $\boldsymbol{D}_1$ with a large MLE (2.33), and the structural prior is less represented. We performed extensive experiments with our database and found that when λ_2 is set to $0.01\lambda_1$, the decomposition is consistently satisfactory. Fig. 5 presents the effect of parameter λ_1 when λ_2 is fixed to $0.01\lambda_1$. It can be seen that λ_2 controls mainly the signal magnitude of $\boldsymbol{D}_1$, but slightly the degree of piecewise smoothness of $\boldsymbol{B}_1$. We fix λ_1 to a moderate value of 0.3.

Other parameters to be determined are λ_3 in (9), γ in (11) and α in (12). λ_3 controls the degree of smoothness in the final base layer $\boldsymbol{B}_2$. We found that except some extreme settings, λ_3 does not affect much the tone mapped images. Hence λ_3 is fixed to 0.1. α mainly controls the stretching degree of the first detail layer $\boldsymbol{D}_1$. To avoid over-boosting effect, we set it to a moderate value of 0.8. Finally, the γ is set to 2.2 as a common practice in Retinex decomposition research [12, 15, 25].

5.2. The Decomposed Layers

To verify the multiscale decomposition performance of our tone mapping algorithm, we compare with Gu's multiscale tone mapper [14]. In Gu's model, a local guided filter weighted by gradient function is repeatedly applied to the original image to obtain a 2-scale decomposition (3 layers). Note that although Gu's model is claimed to have 3 scales (4 layers), the last scale base layer is a constant image. Thus the valid scale number is two. Gu's model enforces the edge-preserving property on the base layer without imposing any prior on the detail layer.

In Fig. 6, the multiscale decomposition results by Gu's model and our model are compared. A 1-D auxiliary analysis is shown in Fig. 7, where a piece of 1-D profile signal (the position is indicated by the white line in Fig. 7(a)) is extracted from the decomposed layers of each method. It can be seen from Fig. 7(b) that Gu's model performs progressive smoothing without considering the spatial property of the detail layer. Thus, the first detail layer (the red curve in Fig. 7(b)) is full of small fluctuations and the tone mapped image is over-enhanced, as depicted in Fig. 6(d). In addition, Gu's model does not strictly preserve edges due to the nature of local filtering. Thus the tone mapped result has halo artifact (see the zoom-in in Fig. 6(d)). In contrast, owe to the structural prior, our method distributes the small-scale variations in the second layer $\boldsymbol{D}_2$, and enforces the first layer $\boldsymbol{D}_1$ to be piecewise constant, as shown in Fig. 7(c). Meanwhile, our method is also edge-preserving. It not only avoids halo artifacts but also achieves visually compelling results, as



(a) Radiance map (b) WLS [10] (c) GLW [30] (d) VAD [11]

(e) BWC [22] (f) GF [14] (g) GR [31] (h) Ours

Figure 8. Comparison of tone mapping methods.

(a) Radiance map (b) WLS [10] (c) GLW [30] (d) VAD [11]

(e) BWC [22] (f) GF [14] (g) GR [31] (h) Ours

Figure 9. Comparison of tone mapping methods.

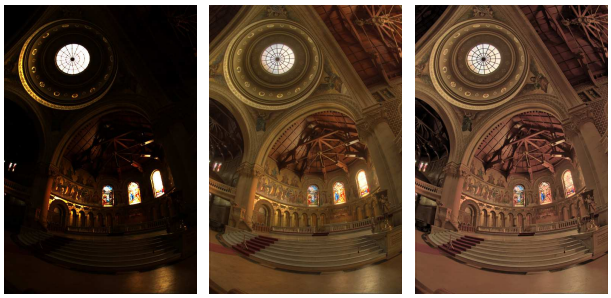
shown in Fig. 6(h).

5.3. Comparison of Tone Mapping

We compare our tone mapper with the state-of-the-art tone mappers [10, 11, 14, 22, 30, 31] on the collected database. These tone mappers include WLS-filter-based method (WLS) [10], globally linear-window method (GLW) [30], visual adaptation method (VAD) [11], backward-compatible method (BWC) [22], guided filter method (GF) [14], and gradient reconstruction method (GR) [31]. More comparison results can be found in the supplementary file. GF is implemented by us since the source code is not available. BWC is implemented with pfstool[4]. The others are implemented by the authors' source codes. All the tone mapping methods use the default parameters as provided in the original papers.

Subjective evaluation. Figs. 8, 9 show the comparison of tone mapping results on two images. We can see that our method achieves a good balance between detail enhancement and naturalness preservation. In contrast, other tone mappers suffer from different types of distortions. WLS loses local contrast and GLW suffers from brightness distortion. VAD has the color shift problem and BWC overly softens the

[4]http://pfstools.sourceforge.net/



(a) Radiance map (b) Photomatix (c) Ours

Figure 10. Comparison with Photomatix.

images. GR and GF have the over-enhancement problem and halo artifacts. In Fig. 10, our tone mapper is compared with the default tone mapper of Photomatix[5]. We can see that both the methods can obtain satisfactory results, while our method achieves higher visual interpretability due to the highlighting of structural information.

To further verify the performance of our tone mapper, we perform a subjective experiment on our HDR database. Specifically, 6 subjects, 3 males and 3 females, are requested to rate all the tone-mapped results of 40 HDR images by the 5 methods. The score ranges from 1 (the worst) to 8 (the best) spaced with 0.5. 2 of the 6 subjects are researchers in computer vision, while the others major in other fields. The tone mapped images are shown on a PA328 display with 32 inch (7680×4320), controlled by a Mac Pro PC with 2.9 GHz CPU. The mean opinion score statistics are illustrated in Fig. 11. Our tone mapper achieves the highest mean scores (6.43) and a tolerable standard deviation (1.20). The mean scores and standard deviations for other tone mappers are WLS (4.91, 1.02), GLW (4.24, 1.62), VAD (4.68, 1.48), BWC (5.11, 1.21), GF (5.31, 1.45), and GR (4.60, 1.60).

Objective evaluation. Aside from subjective evaluation, we use the Tone Mapped Image Quality Index (TMQI) [35] to perform an objective evaluation on the tone mappers. TMQI first evaluates the structural fidelity and naturalness of the tone mapped images. Then the two measures are adjusted by power function and averaged to give a final score ranging from 0 to 1. Larger values of TMQI indicate better quality of the tone mapped image, and vice versa. Table 1 illustrates the mean TMQI score of each tone mapper performed on our database with 40 HDR images. We can see that our method achieves not only the highest TMQI score (0.8851), but also the highest naturalness measure (0.5547). These excellent marks objectively indicate the high visual quality obtained by our algorithm. On the other hand, our tone mapper does not achieve a high fidelity score. This is because the fidelity measure computes the standard deviation in a local window on different scales. Our algorithm, however, regularizes the small scale details to avoid over-enhancement, which lowers the fidelity score.

[5]https://www.hdrsoft.com/

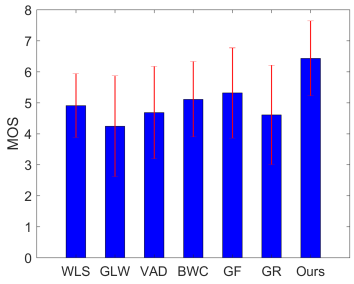


Figure 11. Comparison of mean opinion score statistics.

Table 1. Comparison of average TMQI scores

	TMQI	Fidelity	Naturalness
WLS[10]	0.8703	0.8513	0.4540
GLW[30]	0.8745	0.8172	0.5303
VAD[11]	0.8695	**0.8614**	0.4320
BWC[22]	0.8633	0.8498	0.4213
GF[14]	0.8692	0.8446	0.4508
GR[31]	0.8746	0.8303	0.5147
Ours	**0.8851**	0.8334	**0.5547**

Table 2. Comparison of running time. (M.: Matlab)

	WLS	GLW	VAD	BWC	GF	GR	Ours
Code	M.	M.	C++	C++	M.	M.	M.
Time	10.1	29.2	18.1s	0.7s	1.7s	77.6s	8.6s

Efficiency. The proposed tone mapper has a moderate computational complexity. The most complicated part is the FFT operation in the ADMM-based solver, which costs $\mathcal{O}(Nlog(N))$. Table 2 compares the running time of the 5 tone mappers on a 1333×2000 sized image (Fig. 8(a)). The testing environment is a PC with i7 6850k CPU, 16G RAM. It can be seen that our tone mapper has a moderate running time compared with other methods.

6. Conclusion

In this paper, a novel hybrid ℓ_1-ℓ_0 layer decomposition model was proposed to address the over-enhancement and halo artifact problems of tone mapping. This decomposition model effectively enforces a structural prior to the detail layer and the edge-preserving prior to the base layer. The ADMM algorithm was adopted to solve the decomposition model efficiently. Based on the ℓ_1-ℓ_0 layer decomposition outputs, a multiscale tone mapping algorithm was proposed. It performs dynamic range reduction in the base layer and structure boosting in the detail layer. Due to the proper use of the two priors, our multiscale tone mapping algorithm not only avoids halo artifact but also achieves more visually compelling tone mapping results than existing works.

References

[1] R. Arnheim. *Art and visual perception*. Stockholms Universitet, Institutionen för Konstvetenskap, 2001.

[2] N. Barakat, A. N. Hone, and T. E. Darcie. Minimal-bracketing sets for high-dynamic-range image capture. *IEEE Transactions on Image Processing*, 17(10):1864–1875, Oct. 2008.

[3] S. Bi, X. Han, and Y. Yu. An l1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition. *ACM Trans. Graph.*, 34(4):78:1–78:12, July 2015.

[4] R. Boitard, M. T. Pourazad, P. Nasiopoulos, and J. Slevinsky. Demystifying high-dynamic-range technology: A new evolution in digital media. *IEEE Consumer Electronics Magazine*, 4(4):72–86, Oct. 2015.

[5] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[6] J. Chang, R. Cabezas, and J. W. Fisher. *Bayesian Nonparametric Intrinsic Image Decomposition*, pages 704–719. Springer International Publishing, Cham, 2014.

[7] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '97, pages 369–378, New York, NY, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[8] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph.*, 21(3):257–266, July 2002.

[9] M. Elad. On the origin of the bilateral filter and ways to improve it. *IEEE Transactions on Image Processing*, 11(10):1141–1151, Oct. 2002.

[10] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.*, 27(3):67:1–67:10, Aug. 2008.

[11] S. Ferradans, M. Bertalmio, E. Provenzi, and V. Caselles. An analysis of visual adaptation and contrast perception for tone mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10):2002–2012, Oct. 2011.

[12] X. Fu, D. Zeng, Y. Huang, X. P. Zhang, and X. Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 2782–2790, June 2016.

[13] C. Grigorescu, N. Petkov, and M. A. Westenberg. Contour and boundary detection improved by surround suppression of texture edges. *Image and Vision Computing*, 22(8):609 – 622, 2004.

[14] B. Gu, W. Li, M. Zhu, and M. Wang. Local edge-preserving multiscale decomposition for high dynamic range image tone mapping. *IEEE Transactions on Image Processing*, 22(1):70–79, Jan. 2013.

[15] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel. A variational framework for retinex. *International Journal of Computer Vision*, 52(1):7–23, Apr 2003.

[16] E. H. Land and J. J. McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971.

[17] H. Li, X. Jia, and L. Zhang. Clustering based content and color adaptive tone mapping. *Computer Vision and Image Understanding*, pages –, 2017.

[18] Y. Li, L. Sharan, and E. H. Adelson. Compressing and companding high dynamic range images with subband architectures. *ACM Trans. Graph.*, 24(3):836–844, July 2005.

[19] Z. Liang, W. Liu, and R. Yao. Contrast enhancement by nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 25(2):673–686, Feb. 2016.

[20] C. Lu, J. Shi, and J. Jia. Online robust dictionary learning. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 415–422, June 2013.

[21] W. Ma and S. Osher. A tv bregman iterative model of retinex theory. *Ucla Cam Report*, pages 10–13, 2010.

[22] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, and W. Heidrich. Optimizing a tone curve for backward-compatible high dynamic range image and video compression. *IEEE Transactions on Image Processing*, 20(6):1558–1571, June 2011.

[23] L. Meylan and S. Susstrunk. High dynamic range image rendering with a retinex-based adaptive filter. *IEEE Transactions on Image Processing*, 15(9):2820–2830, Sept. 2006.

[24] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. N. Do. Fast global image smoothing based on weighted least squares. *IEEE Transactions on Image Processing*, 23(12):5638–5653, Dec. 2014.

[25] M. K. Ng and W. Wang. A total variation model for retinex. *SIAM Journal on Imaging Sciences*, 4(1):345–365, 2011.

[26] R. M. H. Nguyen and M. S. Brown. Fast and effective l0 Gradient minimization by region fusion. In *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, pages 208–216, Dec. 2015.

[27] S. Paris and F. Durand. A fast approximation of the bilateral filter using a signal processing approach. *International Journal of Computer Vision*, 81(1):24–52, Jan 2009.

[28] E. Reinhard and K. Devlin. Dynamic range reduction inspired by photoreceptor physiology. *IEEE Transactions on Visualization and Computer Graphics*, 11(1):13–24, Jan. 2005.

[29] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3):267–276, July 2002.

[30] Q. Shan, J. Jia, and M. S. Brown. Globally optimized linear windowed tone mapping. *IEEE Transactions on Visualization and Computer Graphics*, 16(4):663–675, July 2010.

[31] T. Shibata, M. Tanaka, and M. Okutomi. Gradient-domain image reconstruction framework with intensity-range and base-structure constraints. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, pages 2745–2753, June 2016.

[32] J. Tumblin and H. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics and Applications*, 13(6):42–48, Nov. 1993.

[33] G. Ward. A contrast-based scale factor for luminance display. *Graphics gems IV*, pages 415–421, 1994.

[34] L. Xu, C. Lu, Y. Xu, and J. Jia. Image smoothing via l0 gradient minimization. *ACM Trans. Graph.*, 30(6):174:1–174:12, Dec. 2011.

[35] H. Yeganeh and Z. Wang. Objective quality assessment of tone-mapped images. *IEEE Transactions on Image Processing*, 22(2):657–667, Feb. 2013.