

A Joint Intrinsic-Extrinsic Prior Model for Retinex

Bolun Cai¹ Xianming Xu^{1*} Kailing Guo¹ Kui Jia¹ Bin Hu² Dacheng Tao³

¹School of Electronic and Information Engineering, South China University of Technology, China

²Ubiquitous Awareness and Intelligent Solutions Lab, Lanzhou University, China

³UBTECH Sydney AI Centre, School of IT, FEIT, The University of Sydney, Australia

Abstract

We propose a joint intrinsic-extrinsic prior model to estimate both illumination and reflectance from an observed image. The 2D image formed from 3D object in the scene is affected by the intrinsic properties (*shape* and *texture*) and the extrinsic property (*illumination*). Based on a novel structure-preserving measure called **local variation deviation**, a joint intrinsic-extrinsic prior model is proposed for better representation. Better than conventional Retinex models, the proposed model can preserve the structure information by shape prior, estimate the reflectance with fine details by texture prior, and capture the luminous source by illumination prior. Experimental results demonstrate the effectiveness of the proposed method on simulated and real data. Compared with the other Retinex algorithms and state-of-the-art algorithms, the proposed model yields better results on both subjective and objective assessments.

1. Introduction

Retinex theory is a color perception model of human vision and is used to remove illumination effects in images. The primary goal of Retinex is to decompose the observed images into illumination and reflectance. In such a decomposition, the illumination represents the light intensity on the objects, and the reflectance represents the physical characteristics of objects. There are many applications derived from Retinex, such as backlit enhancement [13], low-light enhancement [14], and color correction [15].

It is an ill-posed problem to estimate illumination and reflectance from a single observed image. Based on the Retinex theory [28, 27], many algorithms have been proposed to handle this problem. Among them, path-based algorithms, PDE-based algorithms, and center/surround algorithms are three kinds of classical Retinex algorithms. Path-based algorithms [26, 4] consider that the reflectance depends on the multiplication of the ratios along random

walks, which bring along high computational complexity. In [12, 16], the path computation is replaced by a recursive matrix calculation to make the algorithm more efficient, but the number of iterations is not clearly defined and can strongly influence the final result. In the PDE-based algorithms [21, 29], a partial differential equation (PDE) is adopted to obtain the reflectance, which can be solved efficiently by fast fourier transformation (FFT). However, the non-sparse divergence free vector field results in the gradient without the expected piece-wise continuity. The center/surround algorithms, including single-scale Retinex (SSR) [23], multi-scale Retinex (MSR) [22] and multi-scale Retinex with color restoration (MSRCR) [33], constrain the illumination component to be smooth, and the constraint is contrary for the reflectance. Because Gaussian filter without structure-preserving is used to estimate the illumination, they are prone to have halo artifacts near structural edges.

Recently, several variational methods are introduced to estimate illumination and reflectance. The first variational framework (VF) for Retinex is introduced by Kimmel *et al.* [24], and its objective function is established based on the smooth illumination assumption. But the consideration of reflectance assumption is lacking in the variational framework. In [30], a total variation model (TVM) for reflectance decomposition is proposed by adopting Bregman iteration. However, the restored reflectance is over-smoothed and fine details are lost due to the side effect of the logarithmic transformation. Fu *et al.* [13] propose a probabilistic method for simultaneous reflectance and illumination estimation (SRIE) in the linear domain, which can better preserve the details than the logarithmic domain does. Furthermore, the most recent work in [15] proposes a weighted variational model (WVM) to reinforce the regularization of bright regions in the logarithmic domain. Both SRIE and WVM are based on the simplified assumption that illumination tends to change smoothly. However, different surfaces in different directions is faced with different illumination, which may result in incorrect estimation of illumination near edges. Moreover, the information of luminous source in the scene may be damaged due to the unconstrained

*X. Xu is the corresponding author. (Email: xmxu@scut.edu.cn)

isotropic smoothness assumption.

According to [11, 31], the photograph appearance of 3D object in the world is affected by intrinsic-extrinsic properties including *shape, texture, and illumination*. In this paper, a joint intrinsic-extrinsic prior model for Retinex is proposed to decompose both illumination and reflectance in the linear domain. To the best of our knowledge, this is the first time that shape prior information is incorporated in Retinex model. First, a novel structure-preserving measure called *local variation deviation* is proposed as a shape prior to preserve the structure. A texture prior is used to keep the reflectance with fine details and piece-wise continuity. Then a bright channel prior captures the luminous source based on illumination assumption. Finally, a block coordinate descent method based on iteratively re-weighted least square is adopted to optimize the illumination and reflectance simultaneously. Moreover, a fast solver called preconditioned conjugate gradient [2] is used to reduce the complexity from $O(N^3)$ to $O(N)$ in each iteration.

2. Local Variation Deviation

The illumination is piece-wisely smooth due to the shape of objects, which can be decomposed by a structure-preserving image smoothing. In this section, we first review some existing edge/structure-preserving smoothing method, and then propose a novel texture/structure selection measure for shape prior.

2.1. Existing Edge/Structure-Preserving Smoothing

Depending on the application, the edge/structure-preserving smoothing operators (in Fig. 1) may be manipulated separately in various ways, which can be divided into edge-aware filter, statistics-based method, and optimization-based method.

- Edge-aware filters, *e.g.* bilateral filter (BLF) [37] and rolling guidance filter (RGF) [44], are developed in different strategies. These edge-aware filters trade-off between details flattening and edge preservation in neighboring pixels. However, Gibbs phenomenon of local filters will result in ringing-effect near the edge.

- Statistics-based smoothing method, such as median filter (MED) [32] and local extrema filter (LEF)[36], can remove high-contrast details. Statistics-based methods calculate distribution mode rather than numerical methods in local patches, which can perfectly remove salt and pepper noise. But for high-frequency signals, local statistics still produce oscillating results.

- Weighted least squares (WLS) [9] and relative total variation (RTV) [42] are two representative optimization-based methods. These methods restore images by optimizing functions containing weighted L2 norm or total variation norm. However, they all focus on relatively small variance suppression and vulnerable to textures.

2.2. Our Approach

We address above problems in the mentioned existing methods by a novel texture/structure selection measure. A *local statistical magnitude* called local variation deviation (LVD) is proposed for structure-preserving smoothing by a *global optimization* function.

In statistics, the standard deviation is a measure that is used to quantify the consistency of a set of data. The local variation deviation provides a compact and natural way to identify different type of the variation with its statistical property. The local variation represents the gradient feature and its deviation represents the variation correlation in the local patch. Therefore, local variation deviation provides surprisingly strong discriminative power in distinguishing texture (weak correlation) and structure (strong correlation).

In mathematical terms, let $\mathcal{D}_{x/y}$ denote the local horizontal/vertical variation deviation extracted from I :

$$\mathcal{D}_{x/y} = \left| \nabla_{x/y} I - \frac{1}{|\Omega|} \sum_{\Omega} \nabla_{x/y} I \right|, \quad (1)$$

where $\nabla_{x/y}$ is the gradient operator, and Ω is a local patch with the size of $r \times r$ (r is set to 3 in this paper). To enhance the discrimination between the texture and structure, a relative deviation replaces the general deviation to amplify the distinction, which is rewritten as

$$\mathcal{R}_{x/y} = \left| \frac{\nabla_{x/y} I}{\frac{1}{|\Omega|} \sum_{\Omega} \nabla_{x/y} I + \epsilon} \right|, \quad (2)$$

where ϵ is a small number to avoid division by zero. As illustrated in Fig. 2, the local patches with texture and structure are described by significantly difference of local variation deviation. To remove the textures, the smoothing result I is expressed by an objective function with the LVD measure: $\arg \min_I \|I - S\|_2^2 + \alpha \|\mathcal{R}_x\|_1 + \alpha \|\mathcal{R}_y\|_1$. The first term $\|I - S\|_2^2$ is to minimize the distance between the smoothing result I and the input S . As demonstrated in Fig. 1(h), the proposed approach can effectively eliminate the texture without distorting the structure.

2.3. Analysis

The structure-preserving smoothing property of the local variation deviation can be explained intuitively as following (let the mean local variation value be $\overline{\nabla I} = \frac{1}{|\Omega|} \sum_{\Omega} \nabla I$ for terse expression):

- Case 1: **Flat**. If the patch I is almost constant, we have $\nabla I \approx 0$ and $\overline{\nabla I} \approx 0$, so $\overline{\mathcal{D}} \approx 0$ and $\overline{\mathcal{R}} \approx 0$.

- Case 2: **Texture**. If the patch I changes frequently, ∇I fluctuates more rapidly than $\overline{\nabla I}$, so $\overline{\mathcal{D}} > 0$ and $\overline{\mathcal{R}} \gg 1$.

- Case 3: **Structure**. If the patch I changes in accordance, the deviation of ∇I is vary small, so $\overline{\mathcal{D}} \approx 0$ and $\overline{\mathcal{R}} \approx 1$.

To quantitatively analyze the effectiveness of the LVD measure, we generate an 1D signal containing both texture

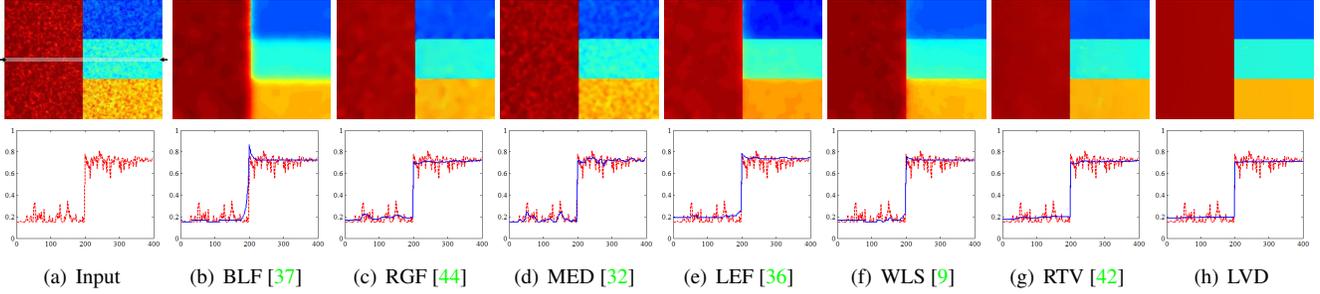


Figure 1: Comparison of different struct-preserving smoothing on a noisy image and a line of pixels extracted from it. (a) Input. (b) BLF ($\sigma_s = 12$, $\sigma_r = 0.45$). (c) RGF ($\sigma_s = 9$, $\sigma_r = 0.05$). (d) MED ($r = 10$). (e) LEF ($r = 3$). (f) WLS ($\lambda = 0.35$, $\alpha = 1.8$). (g) RTV ($\lambda = 0.015$, $\varepsilon_s = 0.02$). (h) LVD ($\alpha = 0.001$).

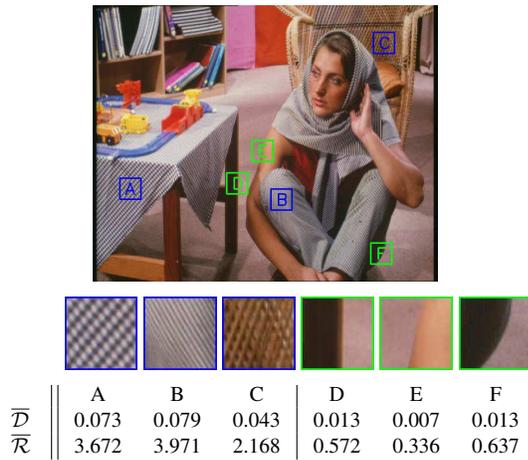


Figure 2: Local variation deviation for different patches. \overline{D} and \overline{R} are the averages of variation deviation in the local patches. The local variation deviation provides strong discriminative power in distinguishing texture (in blue bounding box) and structure (in green bounding box).

and structure in Fig. 3 – that is, the absolute variation $|\nabla I|$ contains full gradients including textures and structures; the mean variation $|\overline{\nabla I}|$ only contains obvious structures; the relative deviation \mathcal{R} captures the textures from the most prominent structures. In Fig. 4, the local variation deviation \mathcal{R} is analyzed on the noisy input. As shown in Fig. 4(a) and Fig. 4(b), the local mean filter inhibits the textural response, so the texture produces smaller values of $|\overline{\nabla I}|$ than $|\nabla I|$. Another intuitive explanation is that the structure in neighbor contributes more similar direction gradients than the texture with complex patterns.

3. A Joint Intrinsic-Extrinsic Prior Model

The physical model of Retinex can be described as $S = I \cdot R$, which means to decompose the observed image S into the illumination I and the reflectance R . According to

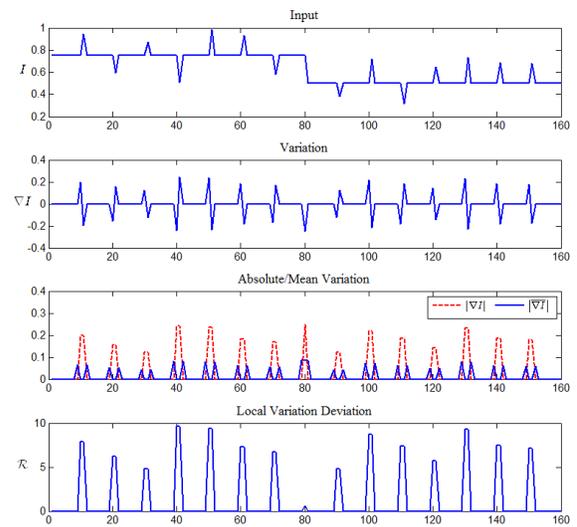


Figure 3: The analysis of LVD on an 1D simulated signal.

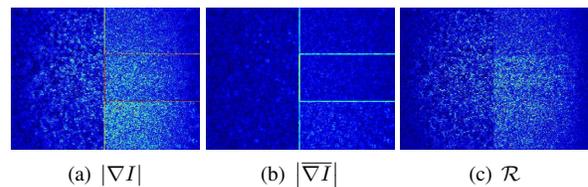


Figure 4: The LVD on the noisy input as in Fig. 1(a).

[11, 31], the appearance of objects is affected by intrinsic and extrinsic properties. The intrinsic properties of objects including shape and texture are independent of illumination, and they not only influence the reflectance but also the illumination conditions. The illumination in the scene is a extrinsic property, which causes the color appearance changes. In this section, an intrinsic-extrinsic prior model for Retinex is developed as an energy minimization problem with “*shape, texture, and illumination*” constraints.

3.1. Intrinsic Prior on Shape & Texture

Since illumination is unknown, a shape prior is proposed to seek the structure of illumination conditions. The prior on shape is motivated by the observation that the object shape tends to be oriented isotopically in the scene. That is, the different surfaces face different illumination with equal probability. This assumption is valid in many real-world environments, such as floors, walls, ground-plane, and sky (as Fig. 5). Thanks to the performance of local variation



Figure 5: The structure of illumination in the real-world environments. The different colored masks represent expresses the surfaces with different illumination conditions.

deviation, the prior on shape is simple and yet very effective to make structures stand out. We combine \mathcal{R}_x and \mathcal{R}_y to form an effective prior for structure decomposition. The energy function is finally expressed as

$$E_s(I) = \left\| \frac{\nabla_x I}{\left| \frac{1}{|\Omega|} \sum_{\Omega} \nabla_x I + \epsilon \right|} \right\|_1 + \left\| \frac{\nabla_y I}{\left| \frac{1}{|\Omega|} \sum_{\Omega} \nabla_y I + \epsilon \right|} \right\|_1 \quad (3)$$

Since possible noise previously hiding for the low-light regions is amplified simultaneously, an important problem of reflectance estimation is how to suppress noise in dark areas. Based on the well-known assumption that the reflectance contains fine texture and is piece-wise continuous [40], the distribution of gradients of reflectance is formulated with a Laplacian distribution, which corresponds to total variance sparsity. The texture prior enforces piece-wise continuous on the reflectance R is given as:

$$E_t(R) = \|\nabla_x R\|_1 + \|\nabla_y R\|_1 \quad (4)$$

3.2. Extrinsic Prior on Illumination

If illumination is only restrained to piece-wise smooth, the illumination of luminous sources and white objects cannot be accurately estimated. In this paper, we introduce a prior on illumination for Retinex. The illumination prior is based on the visual effect of inverted observed images $1 - S$, which is intuitively similar to haze images [8] as in Fig. 6:

$$(1 - S) = 1 - I \cdot R = (1 - R) \cdot I + (1 - I) \quad (5)$$

Let $H = 1 - S$, $J = 1 - R$, $T = I$ and $a = 1$, it is easy to explain the physical meaning of Eq. (5) as the



Figure 6: The inverted image of those shown in Fig. 5.

atmospheric scattering model $H = J \cdot T + a(1 - T)$, where H is the observed hazy image, J is the real scene to be recovered, T is the medium transmission, and a is the global atmospheric light. In most of haze-free patches, at least one color channel has some pixels whose intensity values are very low and even close to zero. A dark channel prior [20] is discovered to estimate the transmission map for dehazing:

$$T = 1 - \min_{\Omega} \left(\min_{c \in \{r, g, b\}} \frac{H^c}{a} \right) \quad (6)$$

During inference, the prior tries to capture the illumination by the local maximum value of three color channels:

$$I = 1 - \min_{\Omega} \left(\min_{c \in \{r, g, b\}} (1 - S)^c \right) = \max_{\Omega} \left(\max_{c \in \{r, g, b\}} S^c \right) \quad (7)$$

We define bright channel as $B = \max_{\Omega} (\max_c S^c)$, and minimize the L2 distance between estimated illumination and bright channel prior:

$$E_l(I) = \|I - B\|_2^2 \quad (8)$$

3.3. Joint Optimization

To efficiently estimate illumination and reflectance, the objective function is established by taking joint intrinsic-extrinsic prior into consideration:

$$E(I, R) = \|I \cdot R - S\|_2^2 + \alpha E_s(I) + \beta E_t(R) + \lambda E_l(I), \quad (9)$$

where α , β and λ are three positive parameters. The first term $\|I \cdot R - S\|_2^2$, which corresponds to L2 data fidelity, is to minimize the distance between estimated $I \cdot R$ and observed image S .

In this paper, a block coordinate descent [38] is adopted to find the optimal solution to the non-convex objective function (9). Since L1-norms in the shape prior E_s and texture prior E_t cause non-smooth optimization, an iteratively re-weighted least square [6] method is introduced and Eq. (3) and Eq. (4) are rewritten as:

$$\begin{cases} E_s(I) = u_x \|\nabla_x I\|_2^2 + u_y \|\nabla_y I\|_2^2 \\ E_t(R) = v_x \|\nabla_x R\|_2^2 + v_y \|\nabla_y R\|_2^2 \end{cases}, \quad (10)$$

$$\text{where } \begin{cases} u_{x/y} = \left(\left| \frac{1}{|\Omega} \sum_{\Omega} \nabla_{x/y} I \right| \left| \nabla_{x/y} I \right| + \epsilon \right)^{-1} \\ v_{x/y} = \left(\left| \nabla_{x/y} R \right| + \epsilon \right)^{-1} \end{cases}. \quad (11)$$

Therefore, two separated sub-problems are iteratively cycled through. In particular, for the k -th iteration:

$$(P1) I_k = \arg \min_I \|I \cdot R_{k-1} - S\|_2^2 + \alpha \left(u_x \|\nabla_x I\|_2^2 + u_y \|\nabla_y I\|_2^2 \right) + \lambda \|I - B\|_2^2$$

$$(P2) R_k = \arg \min_R \|I_k \cdot R - S\|_2^2 + \beta \left(v_x \|\nabla_x R\|_2^2 + v_y \|\nabla_y R\|_2^2 \right)$$

As can be seen, the problem now only involves quadratic terms. The two sub-problems have closed form global optimal solutions, and the algorithm is detailed as follows:

- 1) *Algorithm for P1:* Initializing $I_0 = S$, we use matrix notation to rewrite the loss function: $(\mathbf{I} \cdot \mathbf{R}_{k-1} - \mathbf{S})^T (\mathbf{I} \cdot \mathbf{R}_{k-1} - \mathbf{S}) + \alpha (\mathbf{I}^T \mathbf{D}_x^T \mathbf{U}_x \mathbf{D}_x \mathbf{I} + \mathbf{I}^T \mathbf{D}_y^T \mathbf{U}_y \mathbf{D}_y \mathbf{I}) + \lambda (\mathbf{I} - \mathbf{B})^T (\mathbf{I} - \mathbf{B})$. Here \mathbf{U}_x and \mathbf{U}_y are diagonal matrices containing the weights u_x and u_y , respectively, and \mathbf{D}_x and \mathbf{D}_y are the Toeplitz matrices from the discrete gradient operators with forward difference. The vector \mathbf{I} that minimizes (P1) is uniquely defined as the solution of the linear system:

$$\mathbf{I}_k = (\mathbf{R}_{k-1}^T \mathbf{R}_{k-1} + \alpha \mathbf{M}_k + \lambda \mathbf{1})^{-1} (\mathbf{R}_{k-1}^T \mathbf{S} + \lambda \mathbf{B}), \quad (12)$$

where $\mathbf{1}$ is an identity matrix and $\mathbf{M} = \mathbf{D}_x^T \mathbf{U}_x \mathbf{D}_x + \mathbf{D}_y^T \mathbf{U}_y \mathbf{D}_y$ is the weight matrix, which is a five-point positive definite Laplacian matrix [25].

- 2) *Algorithm for P2:* Since P2 is also a least squares problem, we initialize $\mathbf{R}_0 = \mathbf{S}/\mathbf{I}_1$ and update \mathbf{R}_k similarly to \mathbf{I}_k :

$$\mathbf{R}_k = (\mathbf{I}_k^T \mathbf{I}_k + \beta \mathbf{N}_k)^{-1} (\mathbf{I}_k^T \mathbf{S}), \quad (13)$$

where $\mathbf{N} = \mathbf{D}_x^T \mathbf{V}_x \mathbf{D}_x + \mathbf{D}_y^T \mathbf{V}_y \mathbf{D}_y$.

\mathbf{I} and \mathbf{R} is updated until $\|\mathbf{I}_k - \mathbf{I}_{k-1}\|/\|\mathbf{I}_{k-1}\| \leq \varepsilon$ or $\|\mathbf{R}_k - \mathbf{R}_{k-1}\|/\|\mathbf{R}_{k-1}\| \leq \varepsilon$. To reach $O(N)$ complexity, a fast solver called preconditioned conjugate gradient (PCG) [2] is used for speed up. The whole optimization process is summarized in Algorithm 1.

4. Experiments

In this section, we qualitatively and quantitatively compare the proposed model for Retinex. In our experiments, the empirical parameters α , β , λ and ε are set at 0.001, 0.0001, 0.25, and 10^{-2} , respectively.

4.1. Retinex Decomposition

To verify the effectiveness of joint intrinsic-extrinsic prior model for Retinex, the comparisons of Retinex decomposition are shown in Fig. 7. We compare the proposed

Algorithm 1 A Joint Intrinsic-Extrinsic Prior Model

Input: observed image \mathbf{S} , parameters α , β and λ , maximum iterations K and stopping parameters ε .

Output: illumination \mathbf{I} and reflectance \mathbf{R} .

- 1: initialize $\mathbf{I}_0 \leftarrow \mathbf{S}$
 - 2: **for** $k = 1$ to K **do**
 - 3: compute weights $u_{x/y}$ in Eq. (11)
 - 4: update \mathbf{I}_k using (12)
 - 5: **if** $k = 1$ **then**
 - 6: $\mathbf{R}_0 = \mathbf{S}/\mathbf{I}_1$
 - 7: **end if**
 - 8: compute weights $v_{x/y}$ in Eq. (11)
 - 9: update \mathbf{R}_k using (13)
 - 10: **if** $\|\mathbf{I}_k - \mathbf{I}_{k-1}\|/\|\mathbf{I}_{k-1}\| \leq \varepsilon$ or $\|\mathbf{R}_k - \mathbf{R}_{k-1}\|/\|\mathbf{R}_{k-1}\| \leq \varepsilon$ **then**
 - 11: **break**
 - 12: **end if**
 - 13: **end for**
-

model with two classical methods variational framework (VF) [24] and total variation model (TVM) [30], and two state-of-the-art methods simultaneous reflection & illumination estimation (SRIE) [13] and weighted variation model (WVM) [15]. The decomposition is applied for the V-channel in the HSV (Hue, Saturation and Value) space, and then transform it back to the RGB domain.

The illumination estimated by our model does not distort the structure and is smoother than the others. In other words, our result is more conform with the shape and illumination prior. The shading region on the statue's cheek is a challenge for illumination decomposition, because the changes of illumination are similar to the texture of reflectance. The illumination of VF and WVM are over smoothed and break the illumination information. Though TVM and SRIE both consider the reflectance sparse into the model to preserve illumination structure indirectly, the estimated reflectance lose fine details and look more blurry. Meanwhile, since the illumination is removed, details and noise in dark areas are revealed simultaneously in the reflectance, *i.e.* the background in Fig. 7(b) and Fig. 7(e). For our model, the prior on texture enforces piece-wise continuous and thus can suppress noise.

4.2. Illumination Adjustment

The illumination contains the lightness information, so removing or adjusting the illumination can generate visually pleasing results for dark/backlit images. Therefore, the Gamma correction operation is adopted to modify the estimated illumination. Following [13, 15], the Gamma correction of \mathbf{I} with an adjusting parameter is defined by: $\mathbf{I}' = \mathbf{I}^{\frac{1}{\gamma}}$, where the empirical parameter γ is set as 2.2. The enhanced image of the proposed method is computed by:

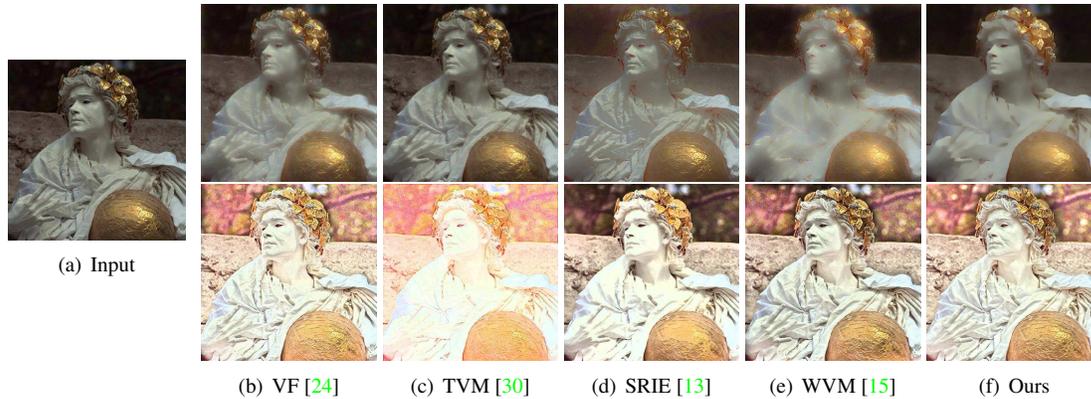


Figure 7: Comparison of Retinex decomposition. The illumination is shown in the first row, and the reflectance in the second.

$S' = I \cdot R$. To preserve color information, the Gamma correction is processed in the HSV domain.

Among the competitors, single-scale retinex (SSR) [23], multi-scale Retinex with color restore (MSRCR) [33], simultaneous reflection & illumination estimation (SRIE) [13] and weighted variation model (WVM) [15] are Retinex-based methods; naturalness preserved enhancement (NPE) [41], globally optimized linear windowed method (GOLW) [34], multi-deviation fusion method (MF) [14], low-light image enhancement (LIME) [19] are recent state-of-the-art image enhancement methods; histogram equalization (HE) [7] and brightness preserving dynamic fuzzy histogram equalization (BPDFHE) [35] are two classical histogram equalization methods used as baselines.

Fig. 8 shows some examples (collected in [41, 19, 13, 15]) of illumination adjustment by five state-of-art methods. The lightness is indeed increased, but the visual artifacts appear in the results obtained from HE and NPE, *i.e.* the sky of HE's and the roadbed of NPE's in the second image. The primary reason is that the nonlinear Gamma correction is carried out on each pixel individually without considering the relationship in the neighbor. Although our method also employs the Gamma correction, the illumination estimated by the joint prior is shape-aware to avoid artifacts. LIME can effectively enhance details, but it has over-enhancement and noise-amplification since the estimated illumination is directly removed (see the third image for an example).

Since all of the illumination adjustment algorithms can obtain effective brightness enhancement on general outdoor images, and the ground truth of the enhanced image is unknown. Following [13, 15], a blind image quality assessment called natural image quality evaluator (NIQE) is used to evaluate the enhanced results. The lower NIQE value represents a higher image quality. As shown in Fig. 8, our method has a lower value in agreement with our subjective experience. In addition, we focus on 35 identified challenging images with different illumination

conditions collected from [41, 34, 14, 19, 13, 15], which are identified can be enhanced effectively by those methods. Since NIQE is just for gray image assessment, we add a color image assessment called autoregressive-based image sharpness metric (ARISM) [18] for supplement. In Table 1, the proposed model has a lower average on NIQE/ARISM than the other state-of-art methods, which indicates that our model has a consistent good performance in most cases.

4.3. Color Correction

When the illumination is decomposed in the RGB-color space separately, the reflectance retains the original color information of the object, which means that the Retinex algorithms has the effect to correct color distortion. The color correction performance is shown in Fig. 9 to demonstrate the effect of the estimated illumination and reflectance. Compared with four other Retinex algorithms, the color correction effect of our model is more obvious in subjective view (*i.e.* the orange bottle in first image and the blue region of the book in the second).

The S-CIELAB color metric [45] based on spatial processing to measure color errors is adopted to verify the accuracy of color correction. The second and fourth rows in Fig. 9 are the spatial location of the errors between the ground truth and corrected results. As can be seen from the spatial locations of the errors, the green areas of our results are smaller than the others, which indicates that our result is more similar to the ground truth. In addition, we add a quantitative comparison with some state-of-the-art color constancy algorithms on Color-Checker dataset [17]. The global average of illumination is calculated simply in the RGB space separately as the estimated illumination. As is standard, we show the mean angle error in Table 2.

4.4. Prior Impact

Since the shape and illumination priors have not been adopted in previous Retinex algorithms, we analyze their

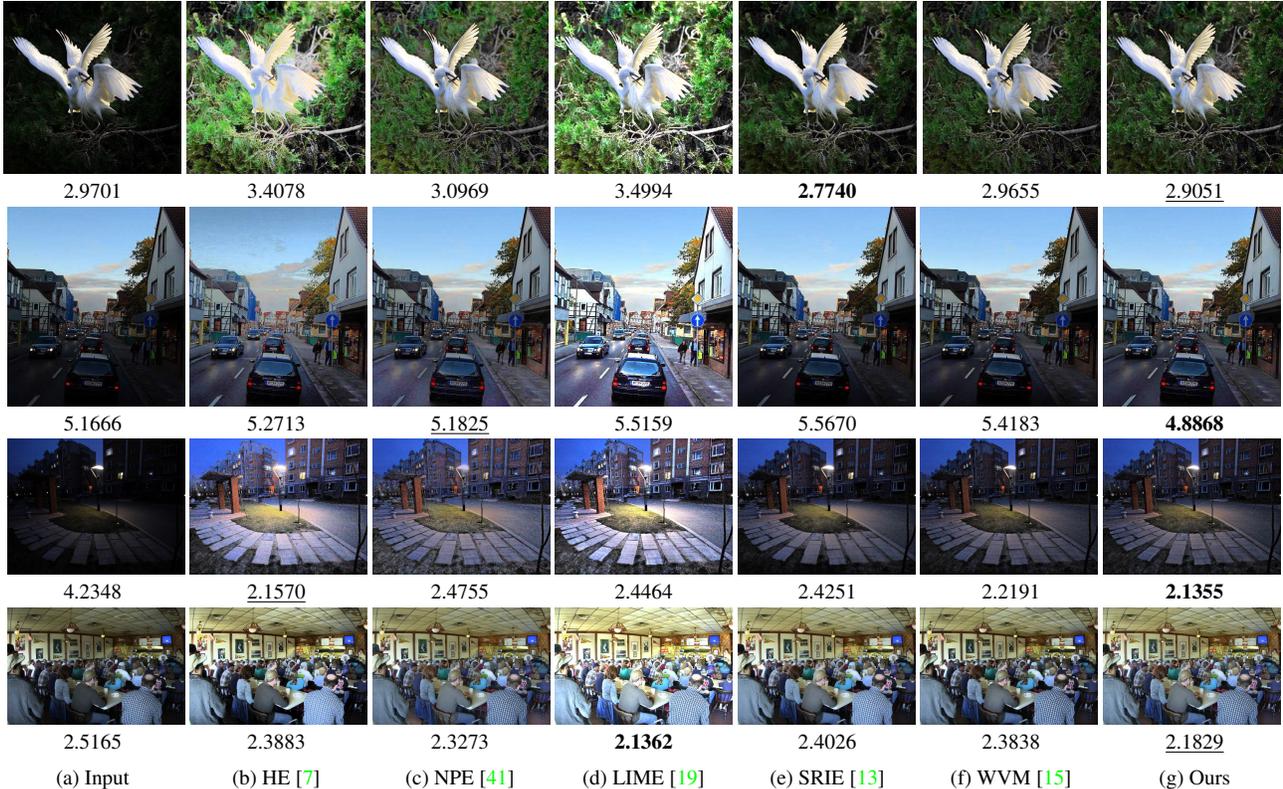


Figure 8: Comparison of illumination adjustment with NIQE.

Table 1: Quantitative performance comparison on 35 images with NIQE and ARISM

Method	HE [7]	BPDFHE [35]	SSR [23]	MSRCR [33]	NPE [41]	GOLW [34]	MF [14]	LIME [19]	SRIE [13]	WVM [15]	Ours
NIQE	3.4475	3.7267	3.3778	3.4295	3.4091	3.3647	3.5335	3.6155	3.4590	3.3594	3.3409
ARISM	3.2902	3.3275	3.0469	3.1014	3.0891	3.3243	3.0200	3.1753	2.9930	2.9958	2.9917

impacts in this experiment. To evaluate the impact of shape prior, the LVD measure of the shape prior is replaced by L2 smoothing regularization. The structure of estimated illumination is blurry without structure preserving (see Fig. 10(c)). To evaluate the impact of illumination prior, the regularization parameters λ is set to zero while keeping the others. As shown in Fig. 10(d), the estimated illumination is over-smoothed when the shape prior is replaced by L2 smoothing regularization. This is because bright channel prior is a strong assumption. The illumination estimation with joint prior captures the luminous source (*i.e.* the moonlight and the street lamps) more effectively than the result without illumination prior. By conducting experiment on the same dataset as in Sec. 4.2, the joint prior model generates more satisfactory results shown in Table 3.

Since effective priors will accelerate the optimization, the convergence rate with different prior is analyzed. Fig. 11 shows the relationship between the iterative error $\varepsilon_l = \|\mathbf{I}_k - \mathbf{I}_{k-1}\| / \|\mathbf{I}_{k-1}\|$, $\varepsilon_r = \|\mathbf{R}_k - \mathbf{R}_{k-1}\| / \|\mathbf{R}_{k-1}\|$ and

Table 3: The impact of priors on the 35 images.

Metric	NIQE	ARISM
Joint Prior	3.3409	2.9917
w/o Shape Prior	3.3964	2.9999
w/o Illumination Prior	3.3791	3.0024

the iteration number. As shown from the iterative error curves, the convergence rate of joint prior is faster than those without shape or illumination prior, because the shape and illumination restraints apply the prior constraints to reduce the solution space scale.

5. Conclusion

In this paper, we proposed a joint intrinsic-extrinsic prior model for Retinex, which is effective in estimating reflectance and illumination simultaneously. First, we proposed a novel measure called local variation deviation to remove textures effectively without distorting structure in

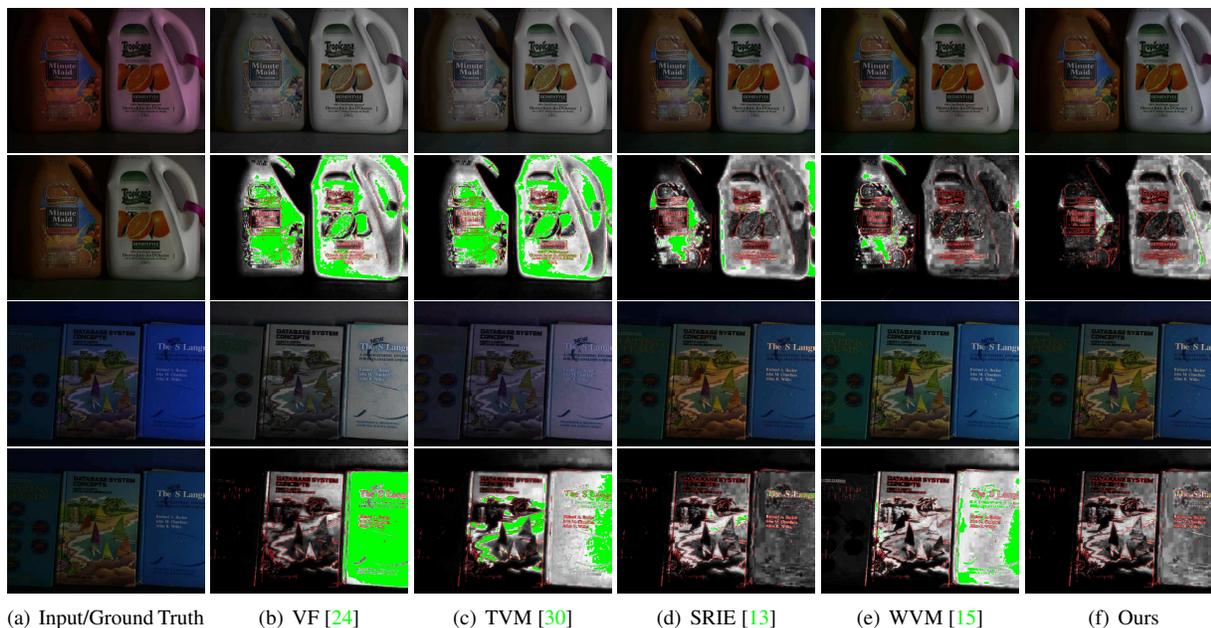


Figure 9: Comparison of color correction. Our results are more accurate when dealing with color distorted images.

Table 2: Comparison of color constancy with angle error on the Color-Checker Dataset [17].

Method	White-Patch [4]	Grey-World [5]	Gray-Edge [39]	Shades-Gray [10]	Bayesian [17]	CNNs [3]	Gray-World [1]	Grey-Pixel [43]	Ours
Mean ($^{\circ}$)	7.55	6.36	5.13	4.93	4.82	4.73	4.66	4.60	4.32

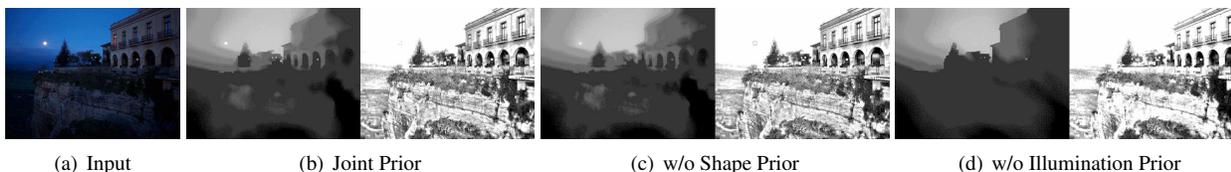


Figure 10: Comparison with different prior. In each group of sub-figure, the illumination and reflectance are shown in the first and second image respectively.

the illumination. In addition, a joint prior model with *shape*, *illumination*, and *texture* assumption is introduced to refine the regularization terms for better prior representation. To separate the illumination and the reflectance efficiently, the optimization problem is solved by a block coordinate descent method with iteratively re-weighted least square. Comprehensive experiments in terms of subjective and objective assessments are conducted. Compared with the state-of-the-art algorithms, the proposed model shows better results with satisfactory convergence rate.

Acknowledgements. This work is supported by National Natural Science Founding of China (61171142, U1636218), Fundamental Research Funds for Central Universities (2017MS045), Guangzhou Key Lab of Body Data Science (201605030011), and Australian Research Council Projects (FL-170100117, DP-140102164, LP-150100671).

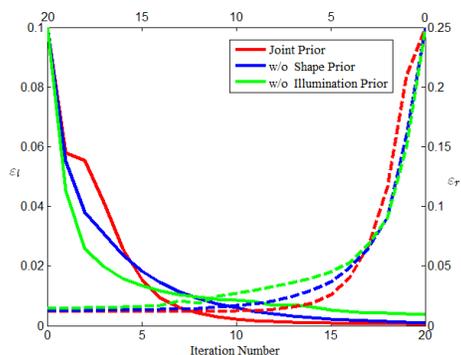


Figure 11: Convergence rate of different priors in 20 iterations. The solid and dash curves denote the relative error of illumination and reflectance, respectively.

References

- [1] K. Barnard, V. Cardei, and B. Funt. A comparison of computational color constancy algorithms. I: Methodology and experiments with synthesized data. *IEEE transactions on Image Processing*, 11(9):972–984, 2002. 8
- [2] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst. *Templates for the solution of linear systems: building blocks for iterative methods*. SIAM, 1994. 2, 5
- [3] S. Bianco, C. Cusano, and R. Schettini. Color constancy using cnns. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 81–89, 2015. 8
- [4] D. H. Brainard and B. A. Wandell. Analysis of the retinex theory of color vision. *JOSA A*, 3(10):1651–1661, 1986. 1, 8
- [5] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1):1–26, 1980. 8
- [6] E. J. Candes, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted l1 minimization. *Journal of Fourier analysis and applications*, 14(5-6):877–905, 2008. 4
- [7] H. Cheng and X. Shi. A simple and effective histogram equalization approach to image enhancement. *Digital Signal Processing*, 14(2):158–170, 2004. 6, 7
- [8] X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu. Fast efficient algorithm for enhancement of low lighting video. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1–6. IEEE, 2011. 4
- [9] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski. Edge-preserving decompositions for multi-scale tone and detail manipulation. In *ACM Transactions on Graphics (TOG)*, volume 27, page 67. ACM, 2008. 2, 3
- [10] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Color and Imaging Conference*, volume 2004, pages 37–41. Society for Imaging Science and Technology, 2004. 8
- [11] M. A. Fischler. Recovering intrinsic scene characteristics from images. Technical report, DTIC Document, 1982. 2, 3
- [12] J. A. Frankle and J. J. McCann. Method and apparatus for lightness imaging, 1983. US Patent 4,384,336. 1
- [13] X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding. A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation. *IEEE Transactions on Image Processing*, 24(12):4965–4977, 2015. 1, 5, 6, 7, 8
- [14] X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, and J. Paisley. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, 129:82–96, 2016. 1, 6, 7
- [15] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2782–2790, 2016. 1, 5, 6, 7, 8
- [16] B. Funt, F. Ciurea, and J. McCann. Retinex in matlab. *Journal of electronic imaging*, 13(1):48–57, 2004. 1
- [17] P. V. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp. Bayesian color constancy revisited. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. 6, 8
- [18] K. Gu, G. Zhai, W. Lin, X. Yang, and W. Zhang. No-reference image sharpness assessment in autoregressive parameter space. *IEEE Transactions on Image Processing*, 24(10):3218–3231, 2015. 6
- [19] X. Guo, Y. Li, and H. Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2016. 6, 7
- [20] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2011. 4
- [21] B. K. Horn. Determining lightness from an image. *Computer graphics and image processing*, 3(4):277–299, 1974. 1
- [22] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 6(7):965–976, 1997. 1
- [23] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell. Properties and performance of a center/surround retinex. *IEEE transactions on image processing*, 6(3):451–462, 1997. 1, 6, 7
- [24] 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, 2003. 1, 5, 6, 8
- [25] D. Krishnan, R. Fattal, and R. Szeliski. Efficient preconditioning of laplacian matrices for computer graphics. *ACM Transactions on Graphics (TOG)*, 32(4):142, 2013. 5
- [26] E. H. Land. Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences*, 80(16):5163–5169, 1983. 1
- [27] E. H. Land et al. *The retinex theory of color vision*. Citeseer, 1977. 1
- [28] E. H. Land and J. J. McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971. 1
- [29] J. M. Morel, A. B. Petro, and C. Sbert. A PDE formalization of retinex theory. *IEEE Transactions on Image Processing*, 19(11):2825–2837, 2010. 1
- [30] M. K. Ng and W. Wang. A total variation model for retinex. *SIAM Journal on Imaging Sciences*, 4(1):345–365, 2011. 1, 5, 6, 8
- [31] M. Olkkonen, T. Hansen, and K. R. Gegenfurtner. Color appearance of familiar objects: Effects of object shape, texture, and illumination changes. *Journal of Vision*, 8(5):13–13, 2008. 2, 3
- [32] S. Perreault and P. Hébert. Median filtering in constant time. *IEEE transactions on image processing*, 16(9):2389–2394, 2007. 2, 3
- [33] Z.-u. Rahman, D. J. Jobson, and G. A. Woodell. Retinex processing for automatic image enhancement. *Journal of Electronic Imaging*, 13(1):100–110, 2004. 1, 6, 7
- [34] 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, 2010. 6, 7

- [35] D. Sheet, H. Garud, A. Suveer, M. Mahadevappa, and J. Chatterjee. Brightness preserving dynamic fuzzy histogram equalization. *IEEE Transactions on Consumer Electronics*, 56(4), 2010. 6, 7
- [36] K. Subr, C. Soler, and F. Durand. Edge-preserving multi-scale image decomposition based on local extrema. *ACM Transactions on Graphics (TOG)*, 28(5):147, 2009. 2, 3
- [37] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Computer Vision, 1998. Sixth International Conference on*, pages 839–846. IEEE, 1998. 2, 3
- [38] P. Tseng. Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of optimization theory and applications*, 109(3):475–494, 2001. 4
- [39] J. Van De Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *IEEE Transactions on image processing*, 16(9):2207–2214, 2007. 8
- [40] L. Wang, L. Xiao, H. Liu, and Z. Wei. Variational Bayesian method for retinex. *IEEE Transactions on Image Processing*, 23(8):3381–3396, 2014. 4
- [41] S. Wang, J. Zheng, H.-M. Hu, and B. Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9):3538–3548, 2013. 6, 7
- [42] L. Xu, Q. Yan, Y. Xia, and J. Jia. Structure extraction from texture via relative total variation. *ACM Transactions on Graphics (TOG)*, 31(6):139, 2012. 2, 3
- [43] K.-F. Yang, S.-B. Gao, and Y.-J. Li. Efficient illuminant estimation for color constancy using grey pixels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2254–2263, 2015. 8
- [44] Q. Zhang, X. Shen, L. Xu, and J. Jia. Rolling guidance filter. In *European Conference on Computer Vision*, pages 815–830. Springer, 2014. 2, 3
- [45] X. Zhang and B. A. Wandell. A spatial extension of cielab for digital color-image reproduction. *Journal of the Society for Information Display*, 5(1):61–63, 1997. 6