Transformed-Domain Robust Multiple-Exposure Blending With Huber Loss

RYO MATSUOKA[1], (Member, IEEE), **SHUNSUKE ONO**[2], (Member, IEEE),
AND MASAHIRO OKUDA[3], (Senior Member, IEEE)
[1]Department of Artificial Intelligence, Kyushu Institute of Technology, Fukuoka 820-8502, Japan
[2]Department of Computer Science, School of Computing, Tokyo Institute of Technology, Tokyo 152-8552, Japan
[3]Faculty of Environmental Engineering, The University of Kitakyushu, Fukuoka 808-0135, Japan

Corresponding author: Ryo Matsuoka (matsuoka@ai.kyutech.ac.jp)

ABSTRACT Pixel-domain weighting methods for multiple-exposure blending can efficiently remove noise and under-/over-exposed pixels simultaneously in *high dynamic range* (HDR) image generation. Various types of noise such as non-Gaussian noise, e.g., Poisson, impulse noise, and pixel saturation, are often superimposed to multiple-exposure images taken with a high ISO setting in a low-light condition. Because almost all existing methods assume Gaussian noise, these methods cannot sufficiently reduce these types of noise. To achieve high-quality HDR image generation in such difficult conditions, we propose a novel multiple-exposure blending method in which image blending is performed in a wavelet domain so as to enhance the denoising performance. In addition, the Huber loss function is utilized as a fidelity measure in blending to make the method robust against outliers. We also introduce an efficient algorithm based on a primal-dual splitting method for solving our optimization problem. The experimental results demonstrate the advantages of the proposed method over several conventional methods.

INDEX TERMS Exposure blending, wavelet transform, convex optimization, total variation, primal-dual splitting.

I. INTRODUCTION

A. BACKGROUND

The objective of *high dynamic range* (HDR) imaging is to represent the amount of light in a scene with a broad dynamic range, and it is applied to various fields such as computer graphics, medical imaging, in-vehicle cameras, and surveillance systems. Multiple-exposure blending is the most standard approach for HDR imaging because it can easily generate HDR images by blending a set of multiple-exposure images captured with a consumer camera [1]–[14].

High ISO shooting is a promising approach in taking multiple-exposure images for HDR image generation. This approach enables us to take ghost-free images, but the images are contaminated by heavy noise. Hence, image-blending methods with noise removal have been studied [15]–[22].

However, these blending methods have a limitation: they assume Gaussian noise. In a low-light condition, observed images taken with a high ISO setting would be degraded by various types of noise like Gaussian noise, Poisson noise,

impulse noise, and under-/over-saturated pixels, in which the statistical properties are different according to exposure time. The noise included in an HDR image obtained by blending these images occurs according to an unknown distribution, and it is difficult to suitably model by a specific distribution and design the fidelity of an HDR image. In addition, the noise removal enhances the dynamic range of an image, which is usually defined by the ratio between the maximum achievable signal intensity and the maximum level of camera noise, and thus sensor noise reduces the dynamic range. Therefore, noise removal techniques play an important role in HDR imaging.

B. RELATED WORK

Simply taking the mean of multiple images can help to reduce the amount of random noise. Several authors have investigated more effective exposure-blending methods based on pixel-domain weighting [4]–[10], [12], [14]. To avoid blur and ghost artifacts, the use of burst images taken with a short-exposure and a high ISO setting is an effective approach [7]–[10]. However, dozens of images are required to generate a high-quality image. A pixel-domain weight

FIGURE 1. Examples of multiscale image-blending methods applied to noisy multiple-exposure images taken with ISO 12800 and −2, 0, and 2 EV difference: (from left to right) exposure fusion [25], wavelet-based fusion [29], and ours.

optimization method for exposure blending was proposed in [13]. In this method, an HDR image is generated by the weighted sum of a few input images with optimal weights defined for each pixel, where the weights are determined by alternately solving an optimization problem with total variation regularization [23].

However, when input images are taken with a high ISO setting in a low-light condition, almost none of the existing methods yield desired results due to the inappropriate modeling discussed in the previous section. If the luminance of each exposure image is uniformly increased (or decreased) due to noise, a noise-free desired pixel cannot be obtained. This is because a blended pixel is obtained by the convex combination of noisy inputs, and thus it is impossible to sufficiently reduce noise, especially in the case of a small number of images.

Multiscale image-blending methods have been introduced to generate natural and high-contrast images from multiple-exposure images [11], [24]–[27] or multi-modal medical images [28], [29]. The weight maps for blending are heuristically calculated without optimization in most methods. The purpose of them is to enhance the amplitude of image gradients and contrast, and thus noise tends to be enhanced (see Fig. 1).

There have been several attempts to handle the noise removal problem by deep convolutional neural networks (CNN) [30]–[33]. Learning-based methods are fragile to degradation not found in a training data set. When removing noise from HDR images using a learning (CNN)-based method, it is difficult and impractical to construct learning data sets and networks. For example, since exposure time, noise variance and characteristics vary greatly depending on scenes, it is necessary to construct a plurality of networks suitable for each scene and appropriately replace them according to an input scene.

C. CONTRIBUTION

In this paper, we propose a novel multiple-exposure blending method for HDR image generation. Our method compounds multiple-exposure images in a wavelet domain. Exposure blending in the wavelet domain can increase the degree of freedom in designing weights and enhance the denoising capability. The subbands yielded by the wavelet

transform are combined as a weighted sum, where the weights used in blending are determined by solving a proposed convex optimization problem. One of the features in our method is that we evaluate data fidelity by using the Huber loss function [34], which reduces the influence of various outliers [35]–[37]. The optimal solution of the proposed optimization problem can be found via the *primal-dual splitting* (PDS) algorithm [38].

The contributions of this paper are as follows:

1) Weight optimization in the transform domain: Input images are decomposed by the wavelet transform, and then each subband is independently combined by a weighted sum. The weights are defined for all the coefficients of each subband, and then its optimal values are estimated by solving a proposed convex optimization problem so as to reduce noise. Our method significantly increases the degree of freedom in the weighted sum thanks to weight optimization, leading to much better blending.

2) Robust data fidelity: As a data fidelity term, we use the Huber loss function, which is robust to outliers. Various types of noise, such as Gaussian, Poisson, impulse noise, and pixel-saturation, are contaminated in an HDR image, especially in a low-light condition. The ℓ_2 data fidelity used in previous methods, e.g., [13], is sensitive to such non-Gaussian noise, and so it cannot reduce noise sufficiently. In contrast, the use of the Huber loss function enables us to generate a high-quality HDR image in such cases.

The remainder of the paper is organized as follows. In Section II, we present mathematical preliminaries and the PDS algorithm. In Section III, we propose a novel multiple-exposure blending method.[1] In Section IV, we show several examples to confirm the effectiveness of our method, where we compare it with conventional exposure-blending methods and denoising methods. Finally, Section V concludes the paper.

II. PRELIMINARIES
A. PROXIMAL TOOLS
The *proximity operator* [40] is a key tool of proximal splitting techniques. Let $\mathbf{x} \in \mathbb{R}^N$ be an input vector. For any $\gamma > 0$, the proximity operator of a proper lower semi-continuous convex function f over $\mathbb{R}^N$ is defined by

$$\text{prox}_{\gamma f}(\mathbf{x}) := \arg \min_{\mathbf{y} \in \mathbb{R}^N} f(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{y}\|_2^2. \quad (1)$$

For a given non-empty closed convex set $\mathcal{C}$, the *indicator function* of $\mathcal{C}$ is defined in [41], [42] by

$$\iota_{\mathcal{C}}(\mathbf{x}) := \begin{cases} 0, & \text{if } \mathbf{x} \in \mathcal{C}, \\ +\infty, & \text{otherwise.} \end{cases} \quad (2)$$

---

[1]The preliminary versions of this work, without the extension of transformed domain exposure blending, several generalizations, mathematical details, new applications, or comprehensive experiments, have appeared in conference proceedings [39].
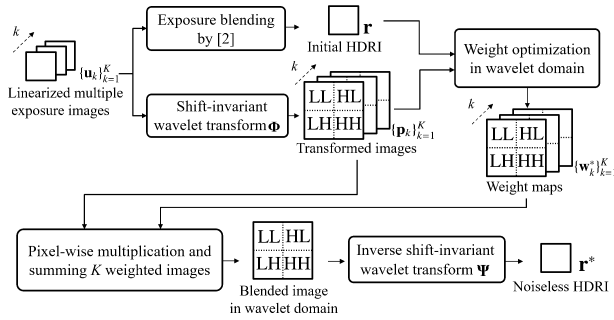
FIGURE 2. The flow of the proposed method.

The proximity operator of this function is expressed as

$$\text{prox}_{\iota_{\mathcal{C}}}(\mathbf{x}) = \arg \min_{\mathbf{y} \in \mathbb{R}^N} \iota_{\mathcal{C}}(\mathbf{y}) + \frac{1}{2\gamma} \|\mathbf{x} - \mathbf{y}\|_2^2. \qquad (3)$$

The solution of (3) should be in the set $\mathcal{C}$ and minimize the term $\|\mathbf{x} - \mathbf{y}\|_2^2$. Thus, for any index $\gamma > 0$, the proximity operator of $\iota_{\mathcal{C}}$ is a projection onto $\mathcal{C}$, i.e., $\mathcal{P}_{\mathcal{C}}(\mathbf{x}) = \text{prox}_{\gamma \iota_{\mathcal{C}}}(\mathbf{x})$.

B. PRIMAL-DUAL SPLITTING ALGORITHM
The PDS method [38] is well known as one of the most flexible solvers for convex optimization. It finds an optimal solution of a convex optimization problem of the form

$$\min_{\mathbf{x}} F(\mathbf{x}) + G(\mathbf{x}) + H(\mathbf{L}\mathbf{x}), \qquad (4)$$

where F, G, and H are proper lower semi-continuous convex functions and $\mathbf{L}$ is a linear operator. The PDS algorithm iteratively computes the two proximity operators

$$\begin{cases} \mathbf{x}^{(\tau+1)} = \text{prox}_{\gamma_1 G} \left(\mathbf{x}^{(\tau)} - \gamma_1 \nabla F \left(\mathbf{x}^{(\tau)} \right) - \gamma_1 \mathbf{L}^* \mathbf{y}^{(\tau)} \right), \\ \mathbf{y}^{(\tau+1)} = \text{prox}_{\gamma_2 H^*} \left(\mathbf{y}^{(\tau)} + 2\gamma_2 \mathbf{L}\mathbf{x}^{(\tau+1)} - \gamma_2 \mathbf{L}\mathbf{x}^{(\tau)} \right), \end{cases} \quad (5)$$

where ∇F is the gradient of F and $\mathbf{L}^*$ is the adjoint of $\mathbf{L}$. The operator $\text{prox}_{\gamma_2 H^*}$ can be computed by prox_{H/γ_2} as

$$\text{prox}_{\gamma_2 H^*}(\mathbf{z}) := \mathbf{z} - \gamma_2 \text{prox}_{H/\gamma_2} \left(\frac{\mathbf{z}}{\gamma_2} \right). \qquad (6)$$

Under appropriate conditions for γ_1 and γ_2, the sequence $(\mathbf{x}^{(\bullet)})_{\bullet \in \mathbb{N}}$ weakly converges to an optimal solution of (4) (see [38] for more details).

III. PROPOSED METHOD
A. OUTLINE
Figure 2 shows the flow of the proposed method. Our objective is to simultaneously perform exposure blending and noise removal. The proposed exposure blending is performed in a wavelet domain.[2] To obtain an initial HDR image, we use the conventional exposure-blending method [2]. We assume linearized multiple-exposure images as input, which are appropriately scaled by exposure ratio. Optimal weight maps that efficiently reduce noise are obtained by solving a proposed weight optimization problem with an initial HDR image

---

[2] Note that other transformation techniques, e.g., *discrete cosine transform* (DCT) and *curvelet transform* [43], can be used in the same manner.

and wavelet-transformed images. After image blending with these weight maps and applying an inverse wavelet transform to it, we obtain a noiseless HDR image.

B. TRANSFORMED-DOMAIN EXPOSURE-BLENDING MODEL
Let $\mathbf{u}_k \in \mathbb{R}^{3N}$ $(k = 1, \ldots, K)$ and $\mathbf{r} \in \mathbb{R}^{3N}$ be linearized color multiple-exposure images (N and K are the number of pixels and images, respectively) and a color HDR image. In addition, let $\mathbf{\Phi} \in \mathbb{R}^{12N \times 3N}$ be a forward transform, which performs a first-level *shift-invariant discrete wavelet transform* [44] for a color image, and $\mathbf{\Psi} \in \mathbb{R}^{3N \times 12N}$ be an inverse transform, which performs an *inverse shift-invariant discrete wavelet transform*, that satisfies $\mathbf{\Psi}\mathbf{\Phi} = \mathbf{I}$, where $\mathbf{I} \in \mathbb{R}^{3N \times 3N}$ is the identity matrix. Our exposure-blending model defined in the irradiance domain is expressed as

$$\mathbf{r} = \mathbf{\Psi}\mathbf{W} \left[(\mathbf{\Phi}\mathbf{u}_1)^{\top} \ldots (\mathbf{\Phi}\mathbf{u}_K)^{\top} \right]^{\top}, \qquad (7)$$

$$\mathbf{W} := \left[\text{diag}(\mathbf{w}_1) \ldots \text{diag}(\mathbf{w}_K) \right] \left(\in \mathbb{R}^{12N \times 12KN} \right), \quad (8)$$

where $\mathbf{w}_k \in \mathbb{R}^{12N}$ $(k = 1, \ldots, K)$ are the weight maps, which are denoted by $\mathbf{w}_k := \left[\mathbf{w}_{k,LL}^{\top} \ \mathbf{w}_{k,LH}^{\top} \ \mathbf{w}_{k,HL}^{\top} \ \mathbf{w}_{k,HH}^{\top} \right]^{\top}$ ($\in \mathbb{R}^{12N}$, $\mathbf{w}_{k,\star} \in \mathbb{R}^{3N}$, $\star \in \{LL, LH, HL, HH\}$), and $\text{diag}(\mathbf{w}_k) \in \mathbb{R}^{12N \times 12N}$ $(k = 1, \ldots, K)$ are its diagonal matrices. Note that these are normalized as $\sum_k w_{k,i} = 1 \forall_i$. When all the elements of $\mathbf{w}$ are $\frac{1}{K}$, the weighted sum of them equals the simple average of the K images. The transpose of a vector or a matrix is defined by $(\cdot)^{\top}$.

Letting $\mathbf{v}_k \in \mathbb{R}^{3N}$ $(k = 1, \ldots, K)$ be observed multiple-exposure images, we consider the following observation model

$$\mathbf{v}_k = D_k \left(\mathbf{u}_k + \mathbf{n}_k \right), \qquad (9)$$

where $\mathbf{n}_k \in \mathbb{R}^{3N}$ and D_k $(k = 1, \ldots, K)$ are added white Gaussian noise (AWGN) and non-Gaussian noise contamination processes.

In the existing pixel-domain weighting methods [1]–[3], [5], weight maps are calculated by using the weight function, which is specified to be small for pixel values near 0 and 1 and high for the middle intensities, where the dynamic range of images are normalized in [0, 1]. A typical weight function $w_c(\cdot)$ introduced in [2] is given as

$$w_c(x) := \begin{cases} 2x, & \text{if } x \leq 0.5, \\ 2(1-x), & \text{if } 0.5 < x. \end{cases} \qquad (10)$$

The role of the weight is to discard saturated pixels. The region where the pixel values are close to 0 or 1 is backed up by other exposure images with larger weights. In the conventional methods, the weighting functions are built based on the assumption that middle intensities around 0.5 have high reliability for irradiance estimation. In contrast, our method defines the weights for subbands in the wavelet domain, which are determined so as to sufficiently reduce noise.

C. PROBLEM FORMULATION

By introducing $\mathbf{p}_k := \mathbf{\Phi}\mathbf{u}_k \ \left(\in \mathbb{R}^{12N}\right) \ (k = 1, \ldots, K)$, our exposure-blending model (7) is rewritten as

$$\mathbf{r} = \mathbf{\Psi}\mathbf{P}\mathbf{w} := \mathbf{\Psi}\left\{\sum_{k=1}^{K} \mathbf{p}_k \otimes \mathbf{w}_k\right\}, \qquad (11)$$

where $\mathbf{w} = \left[\mathbf{w}_1^\top \ \ldots \ \mathbf{w}_N^\top\right]^\top$ is the vectorized form of $4K$ weight maps for the four subbands of the K images, and $\mathbf{P} \in \mathbb{R}^{12N \times 12KN}$ is denoted by $\mathbf{P} := \left[\text{diag}(\mathbf{p}_1) \ \ldots \ \text{diag}(\mathbf{p}_K)\right]$. The notation $\otimes$ denotes the element-wise multiplication.

Our aim is to find a set of weight maps $\mathbf{w}^*$ such that the noise of the reconstructed image $\mathbf{r}$ is relieved. The proposed optimization problem for a set of weight maps $\mathbf{w}$ in the wavelet domain is defined by

$$\min_{\mathbf{w}} \ \rho_{\delta,\mathbf{r}}(\mathbf{\Psi}\mathbf{P}\mathbf{w}) + \alpha\|\mathbf{\Psi}\mathbf{P}\mathbf{w}\|_{\text{TV}},$$

$$\text{s.t. } \mathbf{w} \in \mathcal{C}, \quad \sum_{k=1}^{K} \mathbf{w}_{k,LL} \in \mathcal{E},$$

$$\sum_{k=1}^{K} \mathbf{w}_{k,LH} \in \mathcal{D},$$

$$\sum_{k=1}^{K} \mathbf{w}_{k,HL} \in \mathcal{D}, \quad \sum_{k=1}^{K} \mathbf{w}_{k,HH} \in \mathcal{D}, \qquad (12)$$

where α is the balancing weight for the two terms. We name this problem as the *transformed-domain Huber loss TV* (TD-HLTV) weight optimization.

The first term in (12) is a data-fidelity term, which is characterized by the Huber loss function [34]:

$$\rho_{\delta,\mathbf{r}}(\mathbf{x}) := \sum_{n=1}^{3N} H_\delta(x_n - r_n), \qquad (13)$$

where

$$H_\delta(x) := \begin{cases} \dfrac{x^2}{2}, & \text{if } |x| \leq \delta, \\ \delta|x| - \dfrac{\delta^2}{2}, & \text{otherwise.} \end{cases} \qquad (14)$$

This function acts as a quadratic function for small values and a linear function for large values, where δ is a thresholding parameter.

The second term in (12) is the *total variation* (TV) [23], [45]–[48] regularization term, which represents the characteristic of local smoothness. By letting $\mathbf{D}_v$ and $\mathbf{D}_h \in \mathbb{R}^{N \times N}$ be the vertical and horizontal first-order differential operators, respectively, with Neumann boundaries, the differential operator is expressed by $\mathbf{D}_1 := [\mathbf{D}_v^\top \ \mathbf{D}_h^\top]^\top \ \left(\in \mathbb{R}^{2N \times N}\right)$ for a vectorized gray image with N pixels and $\mathbf{D} := \text{diag}(\mathbf{D}_1, \mathbf{D}_1, \mathbf{D}_1) \ \left(\in \mathbb{R}^{6N \times 3N}\right)^3$ for a vectorized color image $\mathbf{x} \in \mathbb{R}^{3N}$, and thus the TV is defined as [49]–[52]

$$\|\mathbf{x}\|_{\text{TV}} := \|\mathbf{D}\mathbf{x}\|_{1,2}. \qquad (15)$$

---

³The function $\text{diag}(\mathbf{D}_1, \mathbf{D}_1, \mathbf{D}_1)$ denotes the block-diagonal matrix.
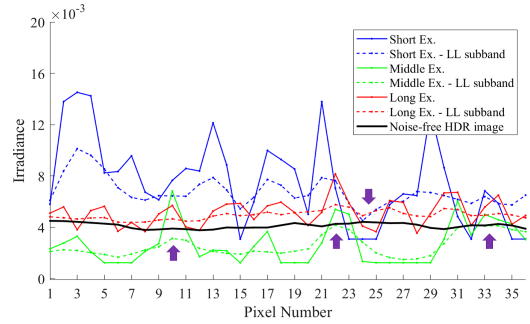
FIGURE 3. Example of multiple-exposure blending: The blue, green, and red lines indicate noisy multiple-exposure image signals, and its *LL*-subband signals are shown by dotted lines with the same colors. The black line indicates a noise-free HDR image signal. The values in the HDR image (black plot) denoted by the arrows are not restored by the weighted sum of the input images in the image domain because the sign of noise in the three signals are same, and all the weight values are positive numbers. In contrast, those values can be restored by the weighted sum of the *LL*-subbands.

The convex sets $\mathcal{C}$, $\mathcal{E}$, and $\mathcal{D}$ are defined as

$$\mathcal{C} := \{\mathbf{x} \in \mathbb{R}^{12KN} \mid x_n \in [0, 1] \ (n = 1, \cdots, 12KN)\},$$
$$\mathcal{E} := \{\mathbf{x} \in \mathbb{R}^{3N} \mid |x_n - 1| \leq \xi \ (n = 1, \cdots, 3N)\},$$
$$\mathcal{D} := \{\mathbf{x} \in \mathbb{R}^{3N} \mid x_n \in [0, 1] \ (n = 1, \cdots, 3N)\},$$

where the first and third convex sets are defined with different-dimension vectors that are essentially the same. By the constraints in (12), all the weights are restricted to fall within the range of $[0, 1]$, and furthermore, the sum of the K weights are restricted to be in the range of $[0, 1]$ for the subbands LH, HL, and HH. Those constraints implicitly induce the energy reduction of high-pass components for the blended image. That is, it allows the derivation of the weights that hardly reflect the high-frequency energy of all K input pixels corresponding to the ideally smooth regions, yielding a smooth image. For the LL-subband, the sum of the K weights should ideally be close to 1 to preserve the energy. In our approach, a tolerable error ξ is set in consideration of calculation error.

Remark (on the objective function (12))

- *Weight optimization in the transform domain:* We consider multiple-exposure blending in the transform (wavelet) domain to overcome the drawback of the blending in an image domain [1]–[3], [13]. The major drawback of the image domain approach is that it strictly limits the possible range of pixel values in $\mathbf{r}$, e.g., if all the K pixels to be combined have negative (or positive) noise values, an ideal pixel value cannot be recovered. Figure 3 shows an example of exposure blending with three exposure images. In Fig. 3, noisy multiple-exposure image signals and a noise-free HDR image signal are shown, in which the noisy images were taken with ISO 12800, and the HDR image was reconstructed by applying the simple weighted sum with the weight function (10) to a set of input noise-free exposure images taken with ISO 100. The values in the HDR image (black line) denoted by the arrows cannot
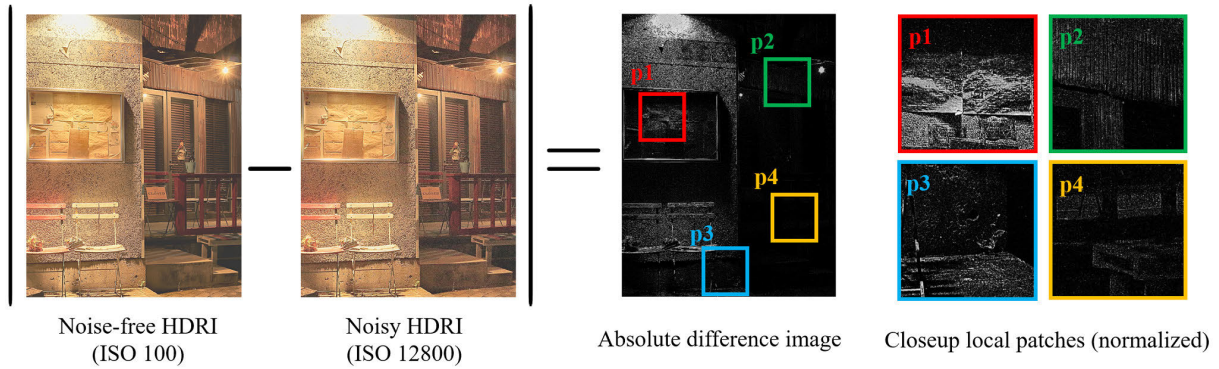
FIGURE 4. Multiple-exposure blending in a low-light condition: Absolute difference image was obtained by calculating the absolute difference between a noise-free and a noisy HDR image, in which each HDR image was reconstructed by the weighted sum of a set of noise-free (or noisy) multiple-exposure images taken with ISO 100 (or 12800). Some closeup local patches are shown, which were normalized by the maximum absolute value of each patch.

be restored by the weighted sum of the noisy exposure images because the sign of noise in the three signals are same, and all the weight values are positive numbers. Thus, the simple convex combination cannot represent signals outside the input signal range. If image blending is performed in a wavelet domain, those values can often be represented by the weighted sum of *LL*-subbands (see Fig. 3). In addition, by independently combining the scaling and wavelet coefficients, the representation ability of complex details is enhanced in image blending.

- *Huber loss-based reconstruction error:* We assume that a set of multiple-exposure images $\mathbf{u}_k$ $(k = 1, \dots, K)$ is given and has already been linearized by photometric calibration. As an input, we prepare a noisy HDR image $\mathbf{r}$ by applying the simple weighted sum with the weight function (10) to the image set $\mathbf{u}_k$ $(k = 1, \dots, K)$. We also assume that the HDR image $\mathbf{r}$ contains additive Gaussian noise in the regions where all the K pixels to combine are not saturated. For regions with saturated pixels in at least a single image, reconstruction errors may become large. Pixel values in bright regions are likely to saturate due to noise contamination and clipping. Furthermore, unknown non-Gaussian noise is superimposed to an image in photographing under a low-light condition. Figure 4 shows an example of the absolute difference between noise-free and noisy HDR images. One observes from the figure that it is obviously inappropriate to simply approximate the reconstruction errors as additive white Gaussian noise. Thus, penalizing the error by the ℓ_2-norm, which is based on a likelihood function of the Gaussian distribution in the Bayesian context, is not appropriate because this norm is significantly influenced by large errors (i.e., outliers). Instead of the ℓ_2-norm, the proposed method uses the Huber loss function [34] to effectively penalize the errors (the Huber loss function is depicted in Fig. 5). Because the function involves a quadratic function near the origin and a linear one away from it, it is robust to outliers.
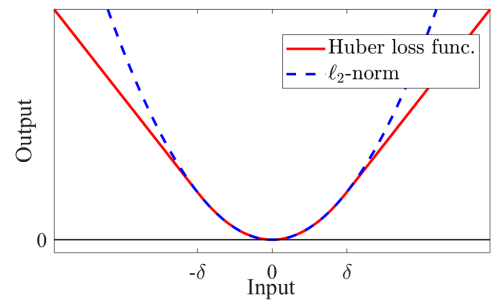

FIGURE 5. Fidelity functions: The blue dashed line and the red line indicate the ℓ_2-norm and the Huber loss function, respectively, where δ is a parameter of the Huber loss function, which determines the boundary between quadratic and linear parts.

D. REFORMULATION AND OPTIMIZATION

By using the indicator functions of $\mathcal{C}$, $\mathcal{E}$, and $\mathcal{D}$, the TD-HLTV weight optimization problem (12) is reformulated into an unconstrained problem

$$\min_{\mathbf{w}} \ \rho_{\delta,\mathbf{r}}(\mathbf{\Psi P w}) + \alpha\|\mathbf{D\Psi P w}\|_{1,2} + \iota_{\mathcal{C}}(\mathbf{w})$$
$$+ \iota_{\mathcal{E}}(\mathbf{E}_{LL}\mathbf{w}) + \sum_{z \in \{LH, HL, HH\}} \iota_{\mathcal{D}}(\mathbf{E}_z\mathbf{w}), \quad (16)$$

where a matrix $\mathbf{E}_{LL} \in \mathbb{R}^{3N \times 12KN}$ computes the sum of the K subbands for *LL*, and the same notation is used for subbands, *LH*, *HL*, and *HH*, which are defined as

$$\mathbf{E}_{LL} := [\mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \dots \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O}],$$
$$\mathbf{E}_{LH} := [\mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \dots \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O}],$$
$$\mathbf{E}_{HL} := [\mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \dots \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O}],$$
$$\mathbf{E}_{HH} := [\mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d \ \dots \ \mathbf{O} \ \mathbf{O} \ \mathbf{O} \ \mathbf{I}_d],$$

where $\mathbf{I}_d \in \mathbb{R}^{3N \times 3N}$ and $\mathbf{O} \in \mathbb{R}^{3N \times 3N}$ are the identity and the zero matrices. For solving the convex optimization problem (16), we use the PDS method described in Section II-B. The correspondence with each term of the objective function defined in (4) is given as follows:

$$F(\mathbf{x}) := \rho_{\delta,\mathbf{r}}(\mathbf{\Psi P x}),$$

FIGURE 6. Input scene images for artificial noise removal experiments: (from left to right) scenes 1, 2, 3, 4, and 5.

| | 18.31 | 22.94 | 20.74 | 20.37 | **23.49** |

Scene 1
(a)

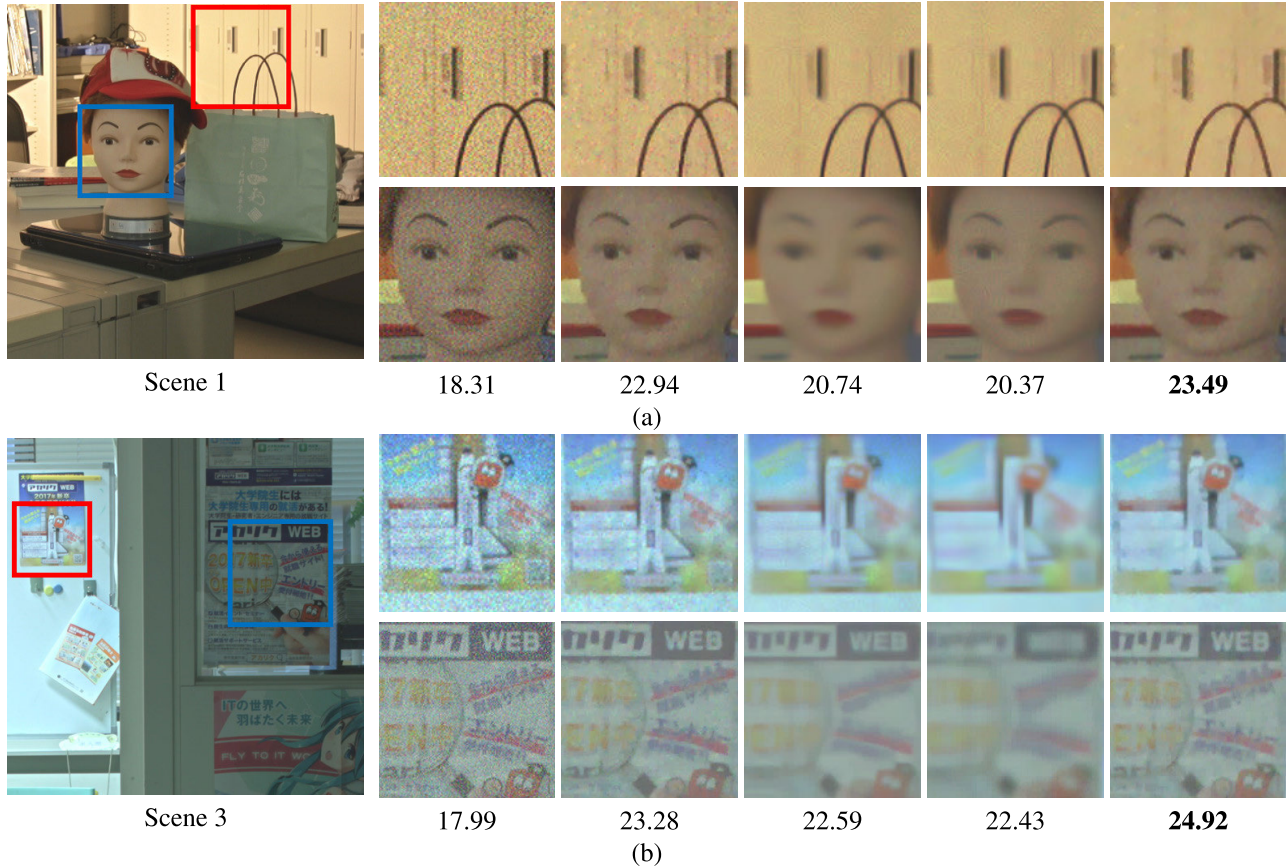| | 17.99 | 23.28 | 22.59 | 22.43 | **24.92** |

Scene 3
(b)

FIGURE 7. Blending results and their NSNR for Gaussian noise with $\sigma^2 \mathcal{D} 8.0 \cdot 10^{-3}$: (from left to right) Debevec and Malik [2], conventional weight optimization method [13], "BM3D + HDR", "BM3D + MEI", and ours.

$$G(\mathbf{x}) := \iota_{\mathcal{C}}(\mathbf{x}),$$

$$H(\mathbf{Lx}) := \alpha \|\mathbf{D\Psi Px}\|_{1,2} + \iota_{\mathcal{E}}(\mathbf{E}_{LL}\mathbf{x}) + \sum_{z \in \{LH, HL, HH\}} \iota_{\mathcal{D}}(\mathbf{E}_z \mathbf{x}),$$

$$\mathbf{L} := \begin{bmatrix} \mathbf{D\Psi P} \\ \mathbf{E}_{LL} \\ \mathbf{E}_{LH} \\ \mathbf{E}_{HL} \\ \mathbf{E}_{HH} \end{bmatrix} \quad (\in \mathbb{R}^{18N \times 12KN}).$$

From Section II-B, an optimal solution can be obtained by alternately calculating the following equations:

$$\begin{vmatrix} \mathbf{w}^{(t+1)} = \text{prox}_{\gamma_1 \iota_{\mathcal{C}}} \left(\mathbf{w}^{(t)} - \gamma_1 \mathbf{P}^\top \mathbf{\Phi} \nabla \rho_{\delta, \mathbf{r}} \left(\mathbf{\Psi P w}^{(t)} \right) - \gamma_1 \mathbf{L}^\top \mathbf{y}^{(t)} \right), \\ \mathbf{y}^{(t+1)} = \text{prox}_{\gamma_2 H^*} \left(\mathbf{y}^{(t)} + 2\gamma_2 \mathbf{L w}^{(t+1)} - \gamma_2 \mathbf{L w}^{(t)} \right), \end{vmatrix}$$
$$(17)$$

where $\mathbf{y} := [\mathbf{y}_1^\top \quad \mathbf{y}_2^\top \quad \mathbf{y}_3^\top \quad \mathbf{y}_4^\top \quad \mathbf{y}_5^\top]^\top$ ($\mathbf{y}_1 \in \mathbb{R}^{6N}$ and $\mathbf{y}_j \in \mathbb{R}^{3N}$ for $j = 2, \dots, 5$). The Huber loss function $\rho_{\delta, \mathbf{r}}$ is differentiable, and the computation of its gradient in (17) is given for $n = 1, \dots, 3N$ by [37] as

$$[\nabla \rho_{\delta, \mathbf{r}}(\mathbf{x})]_n = \begin{cases} x_n - r_n, & \text{if } |x_n - r_n| \le \delta, \\ \delta, & \text{if } x_n - r_n > \delta, \\ -\delta, & \text{if } x_n - r_n < -\delta. \end{cases} \quad (18)$$

The operator $\text{prox}_{\gamma_2 H^*}$ in (17) is independently computed w.r.t. $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4$, and $\mathbf{y}_5$ as follows:

$$\mathbf{y}_1^{(t+1)} = \text{prox}_{\gamma_2 \|\cdot\|_{1,2}^*} \left(\mathbf{y}_1^{(t)} + \gamma_2 \mathbf{D\Psi P} \left(2\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right) \right), \quad (19)$$

$$\mathbf{y}_2^{(t+1)} = \text{prox}_{\gamma_2 t_{\mathcal{E}}^*} \left(\mathbf{y}_2^{(t)} + \gamma_2 \mathbf{E}_{LL} \left(2\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right) \right), \quad (20)$$
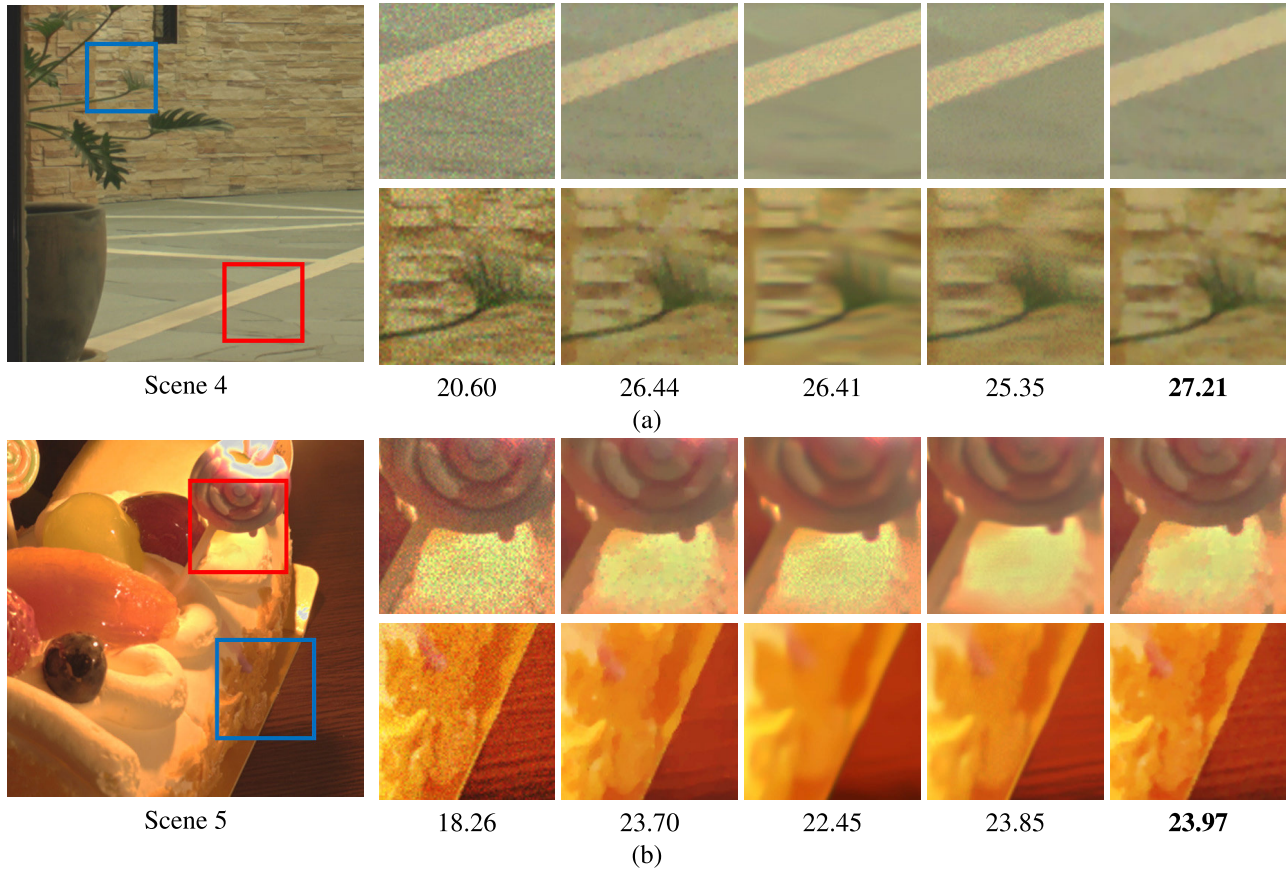
FIGURE 8. Blending results and their NSNR for poisson noise: (from left to right) Debevec and Malik [2], Conventional weight optimization method [13], "BM3D + HDR", "BM3D + MEI", and ours.

$$\mathbf{y}_3^{(t+1)} = \text{prox}_{\gamma_2 \iota_{\mathcal{D}}^*} \left(\mathbf{y}_3^{(t)} + \gamma_2 \mathbf{E}_{LH} \left(2\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right) \right), \quad (21)$$

$$\mathbf{y}_4^{(t+1)} = \text{prox}_{\gamma_2 \iota_{\mathcal{D}}^*} \left(\mathbf{y}_4^{(t)} + \gamma_2 \mathbf{E}_{HL} \left(2\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right) \right), \quad (22)$$

$$\mathbf{y}_5^{(t+1)} = \text{prox}_{\gamma_2 \iota_{\mathcal{D}}^*} \left(\mathbf{y}_5^{(t)} + \gamma_2 \mathbf{E}_{HH} \left(2\mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right) \right). \quad (23)$$

The proximity operator for the $\ell_{1,2}$-norm is given by a group soft-thresholding, for $i = 1, \ldots, mN$, as [50], [52]:

$$[\text{prox}_{\gamma_2 \| \cdot \|_{1,2}}(\mathbf{x})]_i = \max \left\{ 1 - \gamma \left(\sum_{j=0}^{m-1} x_{i+jN}^2 \right)^{-\frac{1}{2}}, 0 \right\} x_i, \quad (24)$$

where m is the number of grouping pixels, and we set $m = 6$ (3 channels × 2 directions of the gradient) in our case. The proximity operators for $\iota_{\mathcal{C}}(\mathbf{x})$ is calculated for $i = 1, \ldots, 12KN$ by

$$[\text{prox}_{\gamma_2 \iota_{\mathcal{C}}}(\mathbf{x})]_i = \begin{cases} 0, & \text{if } x_i < 0, \\ x_i, & \text{if } 0 \le x_i \le 1, \\ 1, & \text{if } x_i > 1, \end{cases} \quad (25)$$

and the proximity operator for $\iota_{\mathcal{E}}(\mathbf{x})$ is calculated for $i = 1, \ldots, 3N$ by

$$[\text{prox}_{\gamma_2 \iota_{\mathcal{E}}}(\mathbf{x})]_i = \begin{cases} 1 - \xi, & \text{if } x_i < 1 - \xi, \\ x_i, & \text{if } 1 - \xi \le x_i \le 1 + \xi, \\ 1 + \xi, & \text{if } x_i > 1 + \xi, \end{cases} \quad (26)$$

where ξ is a specific tolerable error. The proximity operator for $\iota_{\mathcal{D}}(\mathbf{x})$ is also computed by (25) for $i = 1, \ldots, 3N$.

Once the optimal solution $\mathbf{w}^*$ is obtained, we generate a noiseless HDR image by $\mathbf{r}^* = \mathbf{\Psi P w}^*$.

Let us discuss the computational cost of our algorithm. At the update of $\mathbf{w}^{(t+1)}$, the products by $\mathbf{\Psi P}$ and $\mathbf{P}^\top \mathbf{\Phi}$ can be computed by using the shift-invariant wavelet transform with an $\mathcal{O}(N \log N)$ cost (see [44], [53], [54]). The operator $\text{prox}_{\gamma_1 \iota_{\mathcal{C}}}$ and the gradient of the Huber loss function $\nabla \rho_{\mathbf{r}}$ are computed with an $\mathcal{O}(N)$ cost. Hence, the cost of the update of $\mathbf{w}^{(t+1)}$ is $\mathcal{O}(N \log N)$. At the update of $\mathbf{y}^{(t+1)}$, we need to calculate the proximity operators of the $\ell_{1,2}$-norm and the indicator functions $\iota_{\mathcal{E}}$ and $\iota_{\mathcal{D}}$. These costs are $\mathcal{O}(N)$. Hence, the cost of each iteration of our algorithm is $\mathcal{O}(N \log N)$.

IV. EXPERIMENTAL RESULTS
To show the robustness and validity of the proposed multiple-exposure blending method, we applied it to

| Scene 1 | 19.33 | 23.19 | 22.54 | 25.09 | 25.79 | **26.05** |

(a)

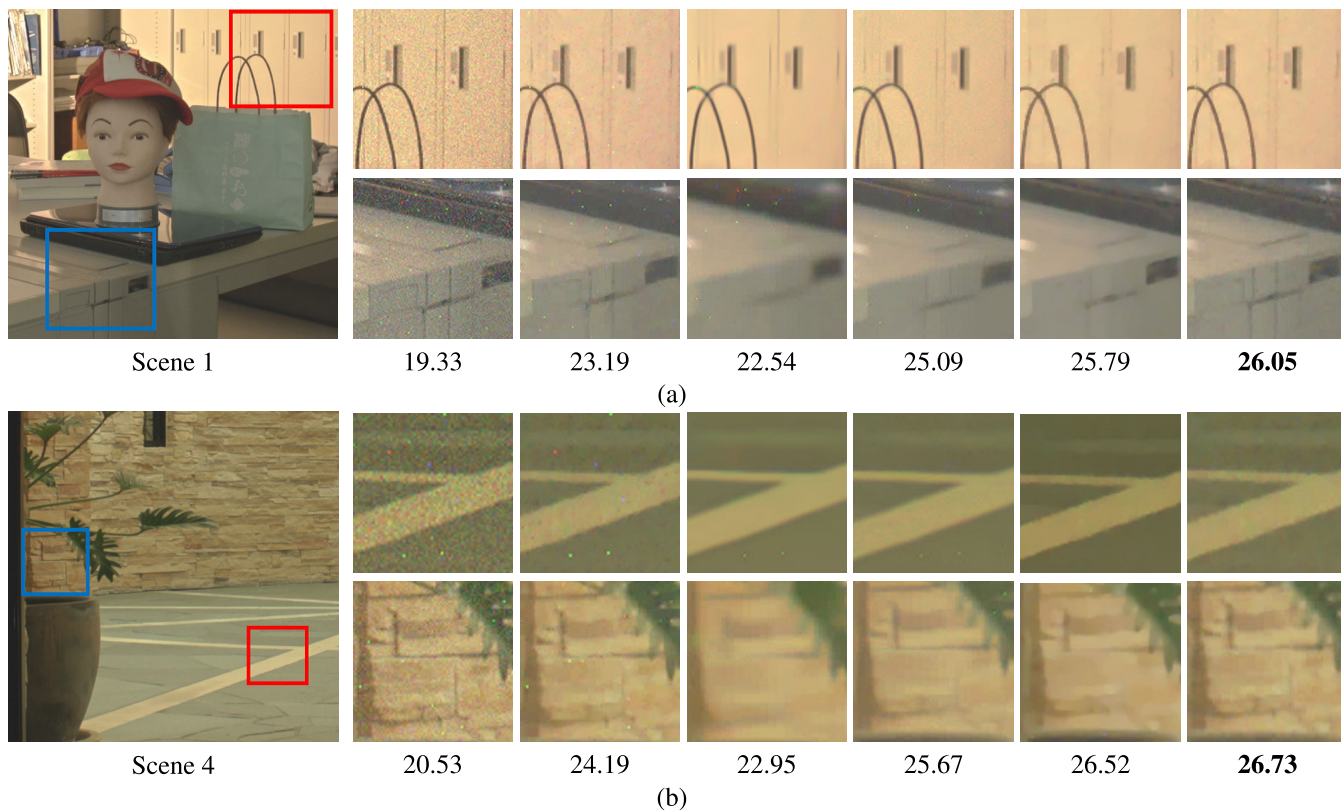| Scene 4 | 20.53 | 24.19 | 22.95 | 25.67 | 26.52 | **26.73** |

(b)

FIGURE 9. Blending results and their NSNR for mixed Gaussian-impulse noise: (from left to right) Debevec and Malik [2], Conventional weight optimization method [13], "BM3D + HDR", "BM3D + MEI", PD-HLTV weight optimization, and TD-HLTV weight optimization (ours).

multiple-exposure images artificially degraded by various types of noise that may occur when taking photographs under low-light conditions and compared our method with several conventional methods. Furthermore, we show the denoising performance of the proposed method to apply it multiple-exposure images taken with a high ISO setting.[4] Note that we assumed scenes were static and carefully photographed input multiple-exposure images so as not to capture moving objects and scene luminance variations.

A. ARTIFICIAL NOISE REMOVAL

For quantitative evaluation, we generated "Ground truth" HDR images as follows:

(i) Acquisition of input multiple-exposure images: To suppress sensor noise, we took multiple-exposure images with ISO 100 and a long shutter speed. Three images, i.e., short-, middle-, and long-exposure images, were obtained by varying the shutter speed with $-2, 0,$ and 2 EV difference while other camera parameters were fixed. To avoid camera shake, we used a tripod.

(ii) Generation of noise-free HDR images: The images obtained in step (i) were simply combined by Debevec

and Malik's method [2] with the weight function (10), yielding noise-free HDR images.

Figure 6 shows some of the Ground truth images.[5] The purpose of the first and second experiments was to clarify the denoising performance of the proposed method by quantitative and qualitative evaluation using the Ground truth images.

For the quality metric, we used the nonlinear SNR (NSNR) of the generated HDR images. Because of the wider dynamic range of an HDR image, the SNR evaluation is susceptible to errors in bright regions, whereas human perception is insensitive to differences in those regions. Therefore we adopted the nonlinear SNR to evaluate the HDR images. This is the SNR of the HDR image tone-mapped by Reinhard *et al.*'s local operator [55]. We compared our method with the two conventional exposure-blending methods based on pixel-domain weighting; one was Debevec and Malik's method [2], i.e., an HDR image was obtained by the weighted sum with (10), and the other was the conventional weight optimization method [13]. Moreover, our results were compared with the state-of-the-art denoising method BM3D [56].[6] For a fair comparison, we performed BM3D before and after exposure blending. In the case of applying BM3D before

---

[4]In these experiments, since all input images were taken with RAW data format, we did not perform *photometric calibration* to them. This is because the RAW images have linear pixel values w.r.t. scene irradiance.

[5]In this paper, instead of directly showing an HDR image, a low dynamic range image tone-mapped by Reinhard *et al.*'s local operator [55] is shown.

[6]We used the source code provided by the authors at http://www.cs.tut.fi/ foi/GCF-BM3D
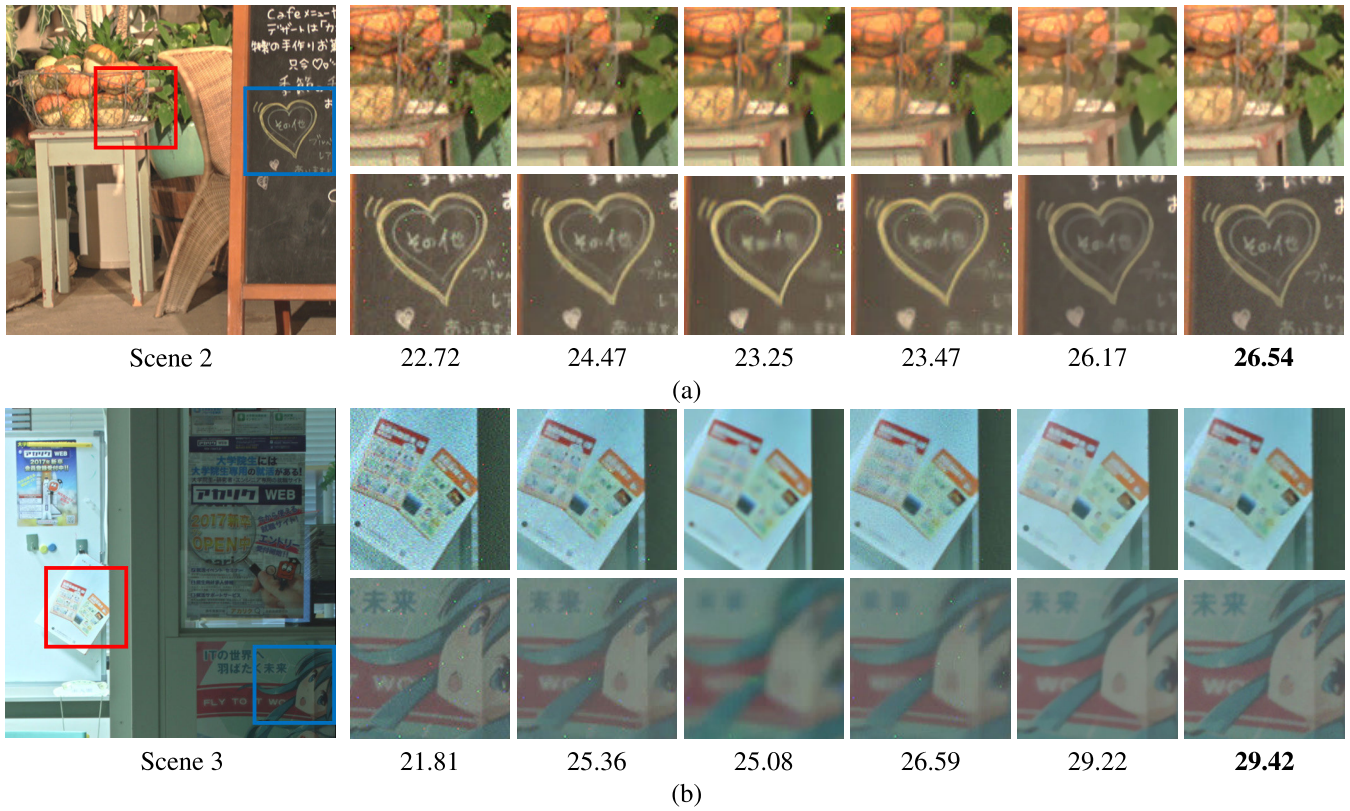
FIGURE 10. Blending results and their NSNR for mixed Poisson-impulse noise: (from left to right) Debevec and Malik [2], Conventional weight optimization method [13], "BM3D + HDR", "BM3D + MEI", PD-HLTV weight optimization, and TD-HLTV weight optimization (ours).

exposure blending, the denoised images obtained by applying BM3D to each input image were blended by [2], yielding a noise-free HDR image. In the case of applying BM3D after exposure blending, BM3D was directly applied to a noisy HDR image obtained by [2]. We call the two approaches as "BM3D + MEI" and "BM3D + HDR", respectively.

All experiments were performed using MATLAB on a Windows 8.1 (64 bit) desktop computer with an Intel Xeon E3-1220 v3 3.1GHz CPU and 16 GB RAM. To accelerate the computation time of the proposed method, we used an NVIDIA GeForce GTX 1080 GPU. For the parameters of our method, we set $\xi = 1.0 \cdot 10^{-5}$, $\gamma_1 = 7.0 \cdot 10^{-4}$, and $\gamma_2 = (1/2000\gamma_1)$ in all experiments. The thresholding parameter of the Huber loss function and the balancing weight were set to $\delta \in [0.8, 4.3]$[7] and $\alpha \in [0.3, 1.2]$, respectively. In each method, we adjusted the degree of noise removal so as to obtain the visually best restoration results, i.e., maximizing smoothness while keeping the edges of the images as much as possible. For the parameter setting of BM3D, we found the visually best results from many resulting images obtained by sequentially changing the standard deviation from a small to a large value. The other parameters were close to the default values shown in Table 1 of [56]. Note that we increased the number of group blocks to enhance the smoothing degree of

---

[7]We recommend that δ be set to a large value in the case of low-intensity noise and to a small value in the case of high-intensity noise.

TABLE 1. NSNR comparison for Gaussian noise, Deb.: Debevec and Malik [2], Conv.: Conventional weight optimization [13], "BM3D + HDR": Applying BM3D [56] to the HDR image obtained by [2], "BM3D + MEI": the multiple-exposure images denoised by BM3D and blended by [2], and ours.

Scene	σ^2	Deb.	Conv.	BM3D + HDR	BM3D + MEI	Ours
1	$4.0 \cdot 10^{-3}$	21.33	25.90	23.01	21.82	**25.92**
	$6.0 \cdot 10^{-3}$	19.56	24.33	21.74	21.01	**24.59**
	$8.0 \cdot 10^{-3}$	18.31	22.94	20.74	20.37	**23.49**
2	$4.0 \cdot 10^{-3}$	21.67	**24.80**	22.13	21.73	24.53
	$6.0 \cdot 10^{-3}$	19.64	**23.26**	20.19	20.01	22.94
	$8.0 \cdot 10^{-3}$	18.29	21.97	19.50	19.44	**22.18**
3	$4.0 \cdot 10^{-3}$	20.84	26.30	24.24	23.61	**27.01**
	$6.0 \cdot 10^{-3}$	19.15	24.55	23.16	22.55	**25.71**
	$8.0 \cdot 10^{-3}$	17.99	23.28	22.59	22.43	**24.92**
4	$4.0 \cdot 10^{-3}$	21.78	25.94	24.90	26.18	**26.95**
	$6.0 \cdot 10^{-3}$	20.10	25.12	23.66	24.99	**25.93**
	$8.0 \cdot 10^{-3}$	18.90	24.50	22.77	24.19	**25.16**
5	$4.0 \cdot 10^{-3}$	21.21	**25.86**	25.16	24.47	25.80
	$6.0 \cdot 10^{-3}$	19.69	23.30	23.90	23.04	**24.89**
	$8.0 \cdot 10^{-3}$	18.62	22.06	23.07	22.50	**23.96**

BM3D, i.e., an upper bound on the number of grouped blocks and the threshold value for evaluating similarity were set to larger values than the default values.

1) SINGLE NOISE CASE

In the first experiment, we independently added AWGN and Poisson noise to the input multiple-exposure images.
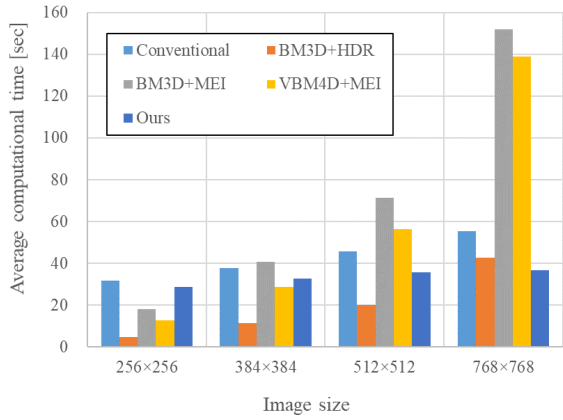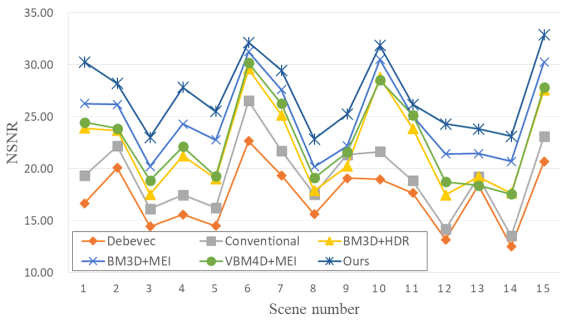
FIGURE 11. Computational time comparison.

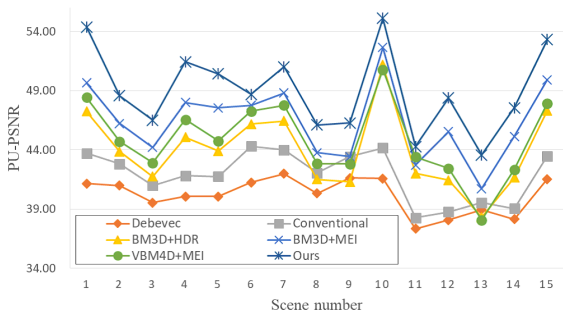**FIGURE 12.** Comparison of NSNR in the case of datasets [59].

FIGURE 13. Comparison of PU-PSNR [60] in the case of datasets [59].

FIGURE 14. Comparison of PU-SSIM [60] in the case of datasets [59].

FIGURE 15. Comparison of HDR-VDP 2.2 [61] in the case of datasets [59].

TABLE 2. NSNR comparison for poisson noise.

Scene	Deb.	Conv.	BM3D + HDR	BM3D + MEI	Ours
1	19.07	24.30	23.36	23.21	**25.27**
2	19.55	23.19	20.60	21.41	**23.32**
3	18.78	23.23	23.88	24.19	**26.26**
4	20.60	26.44	26.41	25.35	**27.21**
5	18.25	23.70	22.45	23.85	**23.97**

For the case of AWGN, we set three types of intensities, *i.e.*, $\sigma^2 = 4.0 \cdot 10^{-3}, 6.0 \cdot 10^{-3}$, and $8.0 \cdot 10^{-3}$. For the case of Poisson noise, we generated each k-th noisy exposure image by $\hat{\mathbf{u}}'_k = \frac{1}{\lambda} D_P(\lambda \hat{\mathbf{u}}_k)$, where $\hat{\mathbf{u}}_k, \hat{\mathbf{u}}'_k$ are the k-th noise-free and noisy images, respectively, and D_P is the Poisson distribution (see [57], [58]). The parameter λ is a scaling parameter that determines the noise intensity, and we set $\lambda = 0.2$ in this experiment.

Tables 1 and 2 show the NSNR comparison for AWGN and Poisson noise, respectively. From Table 1, the NSNR of our method is higher than that of the conventional methods in most cases with the exception of scene 2 with $\sigma^2 = 4.0 \cdot 10^{-3}$ and $6.0 \cdot 10^{-3}$, and scene 5 with $\sigma^2 = 4.0 \cdot 10^{-3}$, in which the conventional weight optimization method has
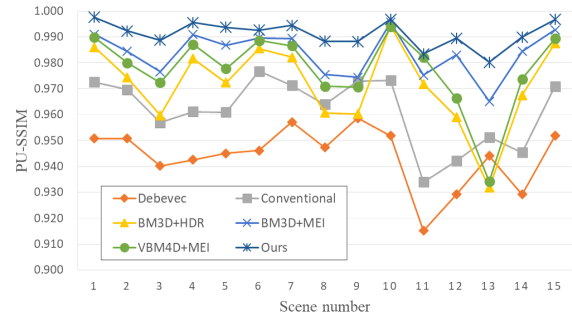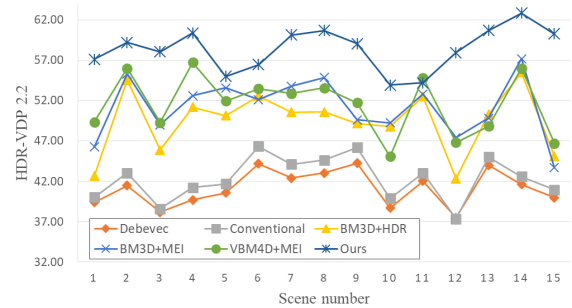
the highest NSNR. In the case of $\sigma^2 = 8.0 \cdot 10^{-3}$, our method outperforms other methods. One can observe from Table 2 that our method has the highest NSNR compared with the conventional methods in all the scenes.

Next, some closeup HDR images in the cases of AWGN with $\sigma^2 = 8.0 \cdot 10^{-3}$ and Poisson noise are shown in Figs. 7 and 8, respectively. One can observe from both figures that our method effectively removes noise compared with the conventional methods, and it preserves details in both dark and bright regions. Although the methods based on BM3D produce smooth images, details in dark regions cannot be preserved, especially in Fig. 7(b). The conventional weight optimization method cannot remove noise in bright regions, especially in Fig. 8(b).

2) MIXED NOISE CASE
In photographing under a low-light condition, in addition to Gaussian/Poisson noise, non-Gaussian noise often occurs in sensor outputs. We show the robustness of the proposed method through mixed noise removal in the second experiment. Mixed noise considered in this experiment included mixed Gaussian-impulse noise and mixed Poisson-impulse
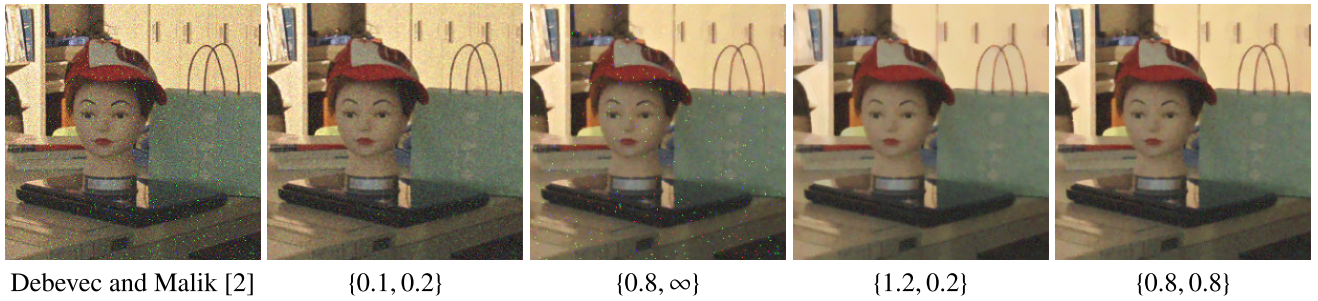
| Debevec and Malik [2] | {0.1, 0.2} | {0.8, ∞} | {1.2, 0.2} | {0.8, 0.8} |

FIGURE 16. Examples of our method with some parameter sets $\{\alpha, \delta\}$.

TABLE 3. Comparison of NSNR in the case of mixed Gaussian-impulse noise.

Scene	Deb.	Conv.	BM3D + HDR	BM3D + MEI	PD-HLTV	TD-HLTV (Ours)
1	19.33	23.19	22.54	25.09	25.79	**26.05**
2	20.59	23.27	20.28	22.42	24.26	**24.90**
3	19.80	24.39	23.85	25.38	26.71	**26.95**
4	20.53	24.19	22.95	25.67	26.52	**26.73**
5	20.79	24.72	24.97	26.02	**26.08**	26.02

TABLE 4. Comparison of NSNR in the case of mixed Poisson-impulse noise.

Scene	Deb.	Conv.	BM3D + HDR	BM3D + MEI	PD-HLTV	TD-HLTV (Ours)
1	21.23	24.22	24.05	25.07	27.80	**28.04**
2	22.72	24.47	23.25	23.47	26.17	**26.54**
3	21.81	25.36	25.08	26.59	29.22	**29.42**
4	22.99	25.37	24.78	26.81	28.87	**29.24**
5	22.01	25.46	25.61	25.26	27.32	**27.45**

noise. The variance of AWGN was set to $\sigma^2 = 4.0 \cdot 10^{-3}$. Poisson noise was generated in the same manner as the first experiment with scaling parameter $\lambda = 0.3$. We considered a salt-and-pepper noise as impulse noise and added it to the images degraded by AWGN and Poisson noise. Note that the probability of impulse noise was set to $8.0 \cdot 10^{-4}$. In this experiment, we compared our method with the preliminary version of the TD-HLTV weight optimization method, named the *pixel-domain Huber loss TV* (PD-HLTV) weight optimization [39]. This method solves a Huber loss TV weight optimization problem in the pixel domain.

We show the NSNR comparison for mixed Gaussian-/Poisson-impulse noise in Tables 3 and 4, respectively. One observes that TD-HLTV has the best NSNR in most scenes. In Table 3-Scene 5, although the NSNR of TD-HLTV is lower than that of PD-HLTV, the difference is slight, but the value is the same as that of "BM3D + MEI".

In Figs. 9 and 10, some closeups of the resulting HDR images are shown for mixed Gaussian-impulse and Poisson-impulse noise, respectively. One observes that TD-HLTV effectively removed Gaussian/Poisson noise and impulse noise simultaneously, and yielded noise-free HDR images while preserving the edges. Because we employed the ℓ_2-norm for data fidelity, the conventional weight optimization method hardly removed impulse noise. Although those types of noise are almost removed by the BM3D-based methods, details are over-smoothed; complex textures on the wall in particular are lost in Fig. 9(b). Both PD-HLTV and TD-HLTV removed mixed noise well by using the Huber loss function. Because PD-HLTV blends pixels in each local region, the degree of smoothing is enhanced. However, it tends to lose the rich textures and sharp edges of an image by over-smoothing, as shown in Figs. 9(a) upper and (b) lower and in Fig. 10(a) lower. This is because different color

TABLE 5. Average value comparison of NSNR, PU-PSNR, PU-SSIM [60], and HDR-VDP 2.2 [61] in the case of mixed Gaussian-impulse noise applied to the fifteen scenes of datasets [59], Deb.: Debevec and Malik [2], Conv.: Conventional weight optimization [13], "BM3D + HDR": Applying BM3D [56] to the HDR image obtained by [2], "BM3D + MEI"/"VBM4D + MEI": the multiple-exposure images denoised by BM3D/VBM4D [62] and blended by [2], and ours.

Metric	Var. of Gauss. noise/ Prob. of Imp. noise	Deb.	Conv.	BM3D +HDR	BM3D +MEI	VBM4D +MEI	Ours
NSNR	0.003 / 0.0013	17.58	20.50	22.26	24.19	22.94	**25.13**
	0.006 / 0.0014	15.87	19.32	21.12	23.16	21.58	**23.28**
	0.002 / 0.0018	17.49	19.79	22.32	24.61	23.62	**25.91**
	0.001 / 0.0023	17.30	19.27	22.19	24.69	22.80	**27.11**
PU-PSNR	0.003 / 0.0013	40.51	43.03	43.72	45.57	44.74	**46.95**
	0.006 / 0.0014	38.52	41.75	42.26	44.34	42.74	**44.89**
	0.002 / 0.0018	40.38	42.75	43.93	46.22	45.63	**47.84**
	0.001 / 0.0023	40.18	41.88	43.95	46.42	44.87	**49.05**
PU-SSIM	0.003 / 0.0013	0.943	0.971	0.972	0.982	0.981	**0.989**
	0.006 / 0.0014	0.912	0.964	0.967	0.981	0.976	**0.983**
	0.002 / 0.0018	0.945	0.967	0.973	0.984	0.983	**0.990**
	0.001 / 0.0023	0.944	0.962	0.972	0.984	0.978	**0.991**
HDR-VDP 2.2	0.003 / 0.0013	42.13	43.31	49.35	51.03	53.42	**56.07**
	0.006 / 0.0014	41.97	43.02	49.27	51.73	52.65	**53.51**
	0.002 / 0.0018	41.38	42.51	49.70	51.50	53.03	**56.81**
	0.001 / 0.0023	41.14	42.32	49.46	51.16	51.55	**58.39**

pixels are often blended around textures and edges. TD-HLTV suppressed these artifacts by blending images in the wavelet domain, which is effective for texture preservation.

3) USE OF A MULTIPLE-EXPOSURE IMAGE DATASET [59]
To show the validity of the proposed method, we added four types of intensities of mixed Gaussian-impulse noise to fifteen scenes[8] of a multiple-exposure image dataset provided in [59] and evaluated these results by using four quality metrics, NSNR, PU-PSNR, PU-SSIM [60], and HDR-VDP 2.2 [61]. The average values of the quality metrics are shown in Table 5. Here, we added the comparison with

[8]Scenes consist of three exposure images with -2, 0, and 2 EV difference.
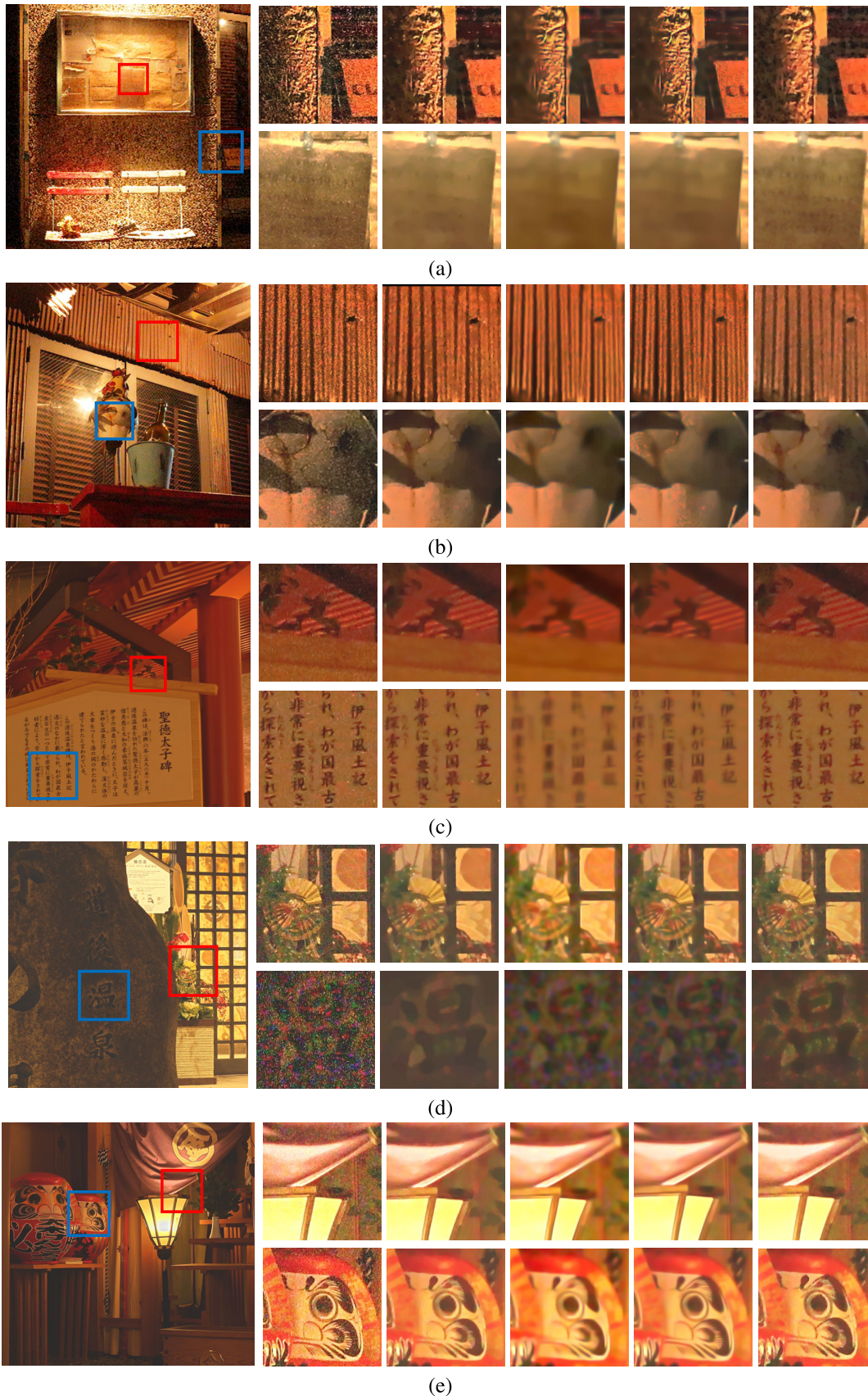
FIGURE 17. Blending results for sensor noise: (from left to right) Debevec and Malik [2], Conventional weight optimization method [13], "BM3D + HDR", "BM3D + MEI", and ours.

VBM4D [62]. Specifically, VBM4D was applied to input multiple-exposure images, and the resulting images were blended by [2] (which is called "VBM4D + MEI"). One sees that the proposed method has the best values of all metrics in most cases. In the fourth case (i.e., 0.001 / 0.0023), our method significantly outperforms the conventional methods. We show these quantitative evaluation results for each scene in Figs. 12, 13, 14, and 15, respectively.

Figure 11 shows the average computational time of the various methods for four different image sizes. One observes that our method has almost constant execution time, regardless of the image size thanks to parallel computing whereas the execution times of the other methods tend to increase as the image size increases. In the case of 768×768, our method is the fastest and about one-quarter that of BM3D + MEI and VBM4D + MEI.

B. REAL-WORLD EXAMPLE

In the final experiment, we applied the proposed method to multiple-exposure images taken with a high ISO setting. Photographs of five scenes were taken with ISO 12800 and -2, 0, and 2 EV difference and used as inputs. Figure 17 shows the resulting images. One observes that our method outperforms the conventional methods. In Figs. 17(a), (d), and (e), although the conventional weight optimization method and "BM3D + HDR" remove strong noise, the details of the bright regions are over-smoothed. The details of the resulting images obtained by "BM3D + MEI" are lost, whereas noise remains on the lower right of the upper closeup image in Figure 17(a). In addition, over-smoothing artifacts occur in both a dark and a bright region. Our method can remove sensor noise efficiently and preserve the image details in both dark and bright regions. Figures 17(b) and (c) also show that our method can remove strong noise while preserving image details in both dark and bright regions. From the results of both the artificial and sensor noise removal experiments, it was confirmed that the proposed blending method is more robust than the existing methods.

C. PARAMETER SETTING COMPLEXITY

The conventional weight optimization method [13] iteratively solves two optimization problems; one is total variation regularization and the other is weight optimization in the image domain. The method requires the careful adjustment of regularization parameters. Because BM3D has various parameters, they should be carefully adjusted for hard-thresholding, block matching, and Wiener filtering. Moreover, BM3D needs to know the standard deviation of noise for achieving high-quality denoising, which is often unknown in practical situations. In the case of a high dynamic range scene, images often have saturated pixels, and the standard deviation of noise adaptively varies in saturated/unsaturated regions. These problems are more serious for "BM3D + MEI" because it is applied to each input image and the parameters should be individually optimized. Furthermore, it is difficult to intuitively improve the image quality of

HDR images because the denoising process is individually carried out on each input image. In contrast, the proposed blending method only has two parameters related to the degree of noise removal, the regularization parameter α and the parameter δ of the Huber loss function. Note that the parameters of the PDS algorithm are independent of input scenes and do not affect the blending results.

Next, we conducted an additional experiment on mixed Gaussian-impulse noise removal. The variance of AWGN was set to $\sigma^2 = 8.0 \cdot 10^{-3}$ and the probability of impulse noise was set to $5.0 \cdot 10^{-3}$. We blended a noisy multiple-exposure image set by our method with some parameter sets $\{\alpha, \delta\}$. The results are shown in Fig. 16. One sees that the noise removal characteristics of our method depend on the sets of $\{\alpha, \delta\}$. By setting α to large values, the smoothing degree of our method is enhanced, yielding smoother images; however, too large a value of α often causes over-smoothing, and thus details and textures are lost. In the cases of small values of δ, non-Gaussian noise, i.e., outliers, is well removed, but little Gaussian noise is removed with small values of α, especially in the case of $\{\alpha, \delta\} = \{0.1, 0.2\}$. In the case of $\{\alpha, \delta\} = \{0.8, \infty\}$, Gaussian noise is well removed, but outliers usually remain. This is because the Huber loss function set to $\delta = \infty$ is equivalent to the ℓ_2 data fidelity from (14), and thus it is sensitive to such non-Gaussian noise. In contrast, in the case of $\{\alpha, \delta\} = \{0.8, 0.8\}$, because δ was set so that the balance between Gaussian and non-Gaussian noise removal was appropriate, each noise could be effectively removed simultaneously. By adjusting the parameter δ, it was possible to flexibly remove mixed noise.

V. CONCLUSION

In this paper, we propose a novel multiple-exposure blending technique for HDR image generation based on wavelet decomposition and the Huber loss function. The proposed exposure blending is performed in the wavelet domain, in which the optimal weights for scaling and wavelet coefficients can be estimated by solving a proposed weight optimization problem that can robustly eliminate various types of noise, including mixed noise contamination. The optimization problem is defined as a convex optimization problem that is solved by an efficient algorithm based on the *primal-dual splitting* method. By considering the estimation of an optimal weight map in the wavelet domain, our method can find more suitable weight maps for exposure blending, which sufficiently reduce noise, than can pixel-domain weighting methods and a highest level denoising method. In addition, to show the validity of using the Huber loss function as data fidelity for exposure blending, we applied the proposed method to noisy multiple-exposure images degraded by artificial noise or sensor noise. Through the all experiments, it was confirmed that the proposed method can robustly remove various types of noise.

In future works, we will attempt to improve the computational efficiency of the iterative estimation algorithm by employing stochastic gradient descent algorithms, and

we will apply the weight optimization scheme for other image restoration problems, i.e., deblurring, demosaicing, and bimodal image blending.

REFERENCES

[1] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, Jun. 1999, pp. 374–380.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM Trans. Graph.*, Aug. 1997, pp. 369–378.

[3] T. Jinno and M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 358–365, Jan. 2012.

[4] R. Matsuoka, T. Baba, M. Okuda, and K. Shirai, "High dynamic range image acquisition using flash image," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2013, pp. 1612–1616.

[5] S. Mann and R. W. Picard, "On being undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *Proc. IST*, 1995, pp. 428–442.

[6] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting* (The Morgan Kaufmann Series in Computer Graphics). San Mateo, CA, USA: Morgan Kaufmann, 2005.

[7] T. Buades, Y. Lou, J. M. Morel, and Z. Tang, "A note on multi-image denoising," in *Proc. Local Non-Local Approx. Image Process. Int. Workshop*, Aug. 2009, pp. 1–15.

[8] L. Zhang, A. Deshpande, and X. Chen, "Denoising vs. Deblurring: HDR imaging techniques using moving cameras," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2010, pp. 522–529.

[9] Z. Liu, L. Yuan, X. Tang, M. Uyttendaele, and J. Sun, "Fast burst images denoising," *ACM Trans. Graph.*, vol. 33, no. 6, Nov. 2014, Art. no. 232.

[10] S. W. Hasinoff, D. Sharlet, R. Geiss, A. Adams, J. T. Barron, F. Kainz, J. Chen, and M. Levoy, "Burst photography for high dynamic range and low-light imaging on mobile cameras," *ACM Trans. Graph.*, vol. 35, no. 6, Nov. 2016, Art. no. 192.

[11] R. Matsuoka, T. Jinno, and M. Okuda, "Multiple exposure integration with image denoising," in *Proc. APSIPA Annu. Summit Conf. (APSIPA ASC)*, Dec. 2012, pp. 1–4.

[12] W. Wang and M. K. Ng, "A variational method for multiple-image blending," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1809–1822, Apr. 2012.

[13] R. Matsuoka, T. Yamauchi, T. Baba, and M. Okuda, "Weight optimization for multiple image integration and its applications," *IEICE Trans. Inf. Syst.*, vol. E99–D, no. 1, pp. 228–235, Jan. 2016.

[14] R. Matsuoka, K. Shirai, and M. Okuda, "Reference-based local color distribution transformation method and its application to image integration," *Signal Process., Image Commun.*, vol. 76, pp. 231–242, Aug. 2019.

[15] K. Shirai, M. Okamoto, and M. Ikehara, "Noiseless no-flash photo creation by color transform of flash image," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2011, pp. 3437–3440.

[16] T. Baba, R. Matsuoka, S. Ono, K. Shirai, and M. Okuda, "Flash/no-flash image integration using convex optimization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2014, pp. 1185–1189.

[17] T. Baba, R. Matsuoka, K. Shirai, and M. Okuda, "Misaligned image integration with local linear model," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2035–2044, May 2016.

[18] C.-H. Son and X.-P. Zhang, "Layer-based approach for image pair fusion," *IEEE Trans. Image Process.*, vol. 25, no. 6, pp. 2866–2881, Jun. 2016.

[19] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," *ACM Trans. Graph.*, vol. 26, no. 3, Jul. 2007, Art. no. 1.

[20] M. Tico and K. Pulli, "Image enhancement method via blur and noisy image fusion," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Nov. 2009, pp. 1521–1524.

[21] S. Zhuo, X. Zhang, X. Miao, and T. Sim, "Enhancing low light images using near infrared flash images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2010, pp. 2537–2540.

[22] S. Takeyama, S. Ono, and I. Kumazawa, "Image restoration with multiple hard constraints on data-fidelity to blurred/noisy image pair," *IEICE Trans. Inf. Syst.*, vols. E100–D, no. 9, pp. 1953–1961, Sep. 2017.

[23] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenomena*, vol. 60, nos. 1–4, pp. 259–268, 1992.

[24] C. Wen and J. Chen, "Multi-resolution image fusion technique and its application to forensic science," *Forensic Sci. Int.*, vol. 140, nos. 2–3, pp. 217–232, Mar. 2004.

[25] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion," in *Proc. 15th Pacific Conf. Comput. Graph. Appl.*, Oct. 2007, pp. 382–390.

[26] W. Zhang, X. Liu, W. Wang, and Y. Zeng, "Multi-exposure image fusion based on wavelet transform," *Int. J. Adv. Rob. Syst.*, vol. 15, no. 2, Apr. 2018, Art. no. 1729881418768939.

[27] W. Liu, Q. Wang, Y. Liu, and N. Li, "High dynamic tone mapping algorithm based on wavelet domain image fusion," in *Proc. 13th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, May/Jun. 2018, pp. 1945–1950.

[28] H. Chen, "A multiresolution image fusion based on principle component analysis," in *Proc. Int. Conf. Image Graph. (ICIG)*, Aug. 2007, pp. 737–741.

[29] R. Singh and A. Khare, "Multiscale medical image fusion in wavelet domain," *Sci. World J.*, vol. 2013, Sep. 2013, Art. no. 521034.

[30] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[31] Y. Chen and T. Pock, "Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1256–1272, Jun. 2017.

[32] B. Mildenhall, J. T. Barron, J. Chen, D. Sharlet, R. Ng, and R. Carroll, "Burst denoising with kernel prediction networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 2502–2510.

[33] S. Lefkimmiatis, "Universal denoising networks: A novel CNN architecture for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3204–3213.

[34] P. J. Huber, "Robust estimation of a location parameter," *Ann. Math. Statist.*, vol. 35, no. 1, pp. 73–101, 1964.

[35] P. Petrus, "Robust Huber adaptive filter," *IEEE Trans. Signal Process.*, vol. 47, no. 4, pp. 1129–1133, Apr. 1999.

[36] M. Nikolova, "Minimizers of cost-functions involving nonsmooth data-fidelity terms. application to the processing of outliers," *SIAM J. Numer. Anal.*, vol. 40, no. 3, pp. 965–994, 2002.

[37] S. Ono and I. Yamada, "Signal recovery with certain involved convex data-fidelity constraints," *IEEE Trans. Signal Process.*, vol. 63, no. 22, pp. 6149–6163, Nov. 2015.

[38] L. Condat, "A primal–dual splitting method for convex optimization involving Lipschitzian, proximable and linear composite terms," *J. Optim. Theory Appl.*, vol. 158, no. 2, pp. 460–479, Aug. 2013.

[39] R. Matsuoka, S. Ono, and M. Okuda, "High dynamic range image generation based on convolutional weight optimization robust to mixed noise removal," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2018, pp. 1066–1070.

[40] J. J. Moreau, "Fonctions convexes duales et points proximaux dans un espace Hilbertien," *C. R. Acad. Sci. Paris A, Math.*, vol. 255, pp. 2897–2899, Dec. 1962.

[41] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, H. H. Bauschke, R. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Eds. New York, NY, USA: Springer-Verlag, 2011, pp. 185–212.

[42] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.*, vol. 20, no. 3, pp. 681–695, Mar. 2011.

[43] E. Candès, L. Demanet, D. Donoho, and X. Ying, "Fast discrete curvelet transforms," *Multiscale Model. Simul.*, vol. 5, no. 3, pp. 861–899, Sep. 2006.

[44] S. Mallat, *A Wavelet Tour of Signal Processing: The Sparse Way*, 3rd ed. New York, NY, USA: Academic, 2008.

[45] P. Blomgren and T. F. Chan, "Color TV: Total variation methods for restoration of vector-valued images," *IEEE Trans. Image Process*, vol. 7, no. 3, pp. 304–309, Mar. 1998.

[46] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. Imag. Vis.*, vol. 20, no. 1, pp. 89–97, Jan. 2004.

[47] M. V. Afonso, J.-M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010.

[48] L. Condat, "A generic proximal algorithm for convex optimization—Application to total variation minimization," *IEEE Signal Process. Lett.*, vol. 21, no. 8, pp. 985–989, Aug. 2014.

[49] X. Bresson and T. F. Chan, "Fast dual minimization of the vectorial total variation norm and applications to color image processing," *Inverse Problems Imag.*, vol. 2, no. 4, pp. 455–484, Nov. 2008.

[50] P. L. Combettes and J.-C. Pesquet, "A proximal decomposition method for solving convex variational inverse problems," *Inverse Problems*, vol. 24, no. 6, Nov. 2008, Art. no. 065014.

[51] V. Duval, J.-F. Aujol, and L. A. Vese, "Mathematical modeling of textures: Application to color image decomposition with a projected gradient algorithm," *J. Math. Imag. Vis.*, vol. 37, no. 3, pp. 232–248, Jul. 2010.

[52] S. Ono and I. Yamada, "Color-line regularization for color artifact removal," *IEEE Trans. Comput. Imag.*, vol. 2, no. 3, pp. 204–217, Sep. 2016.

[53] G. Beylkin, "On the representation of operators in bases of compactly supported wavelets," *SIAM J. Numer. Anal.*, vol. 29, no. 6, pp. 1716–1740, 1992.

[54] M. Shensa, "The discrete wavelet transform: Wedding the a trous and Mallat algorithms," *IEEE Trans. Signal Process.*, vol. 40, no. 10, pp. 2464–2482, Oct. 1992.

[55] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, Jul. 2002.

[56] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[57] R. Puetter, T. Gosnell, and A. Yahil, "Digital image reconstruction: Deblurring and denoising," *Annu. Rev. Astron. Astrophys.*, vol. 43, no. 1, pp. 139–194, Aug. 2005.

[58] S. Ono and I. Yamada, "Poisson image restoration with likelihood constraint via hybrid steepest descent method," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2013, pp. 5929–5933.

[59] C. Bloch, *The HDRI Handbook: High Dynamic Range Imaging for Photographers CG Artists +DVD*. San Rafael, CA, USA: Rocky Nook, 2007.

[60] E. Zerman, G. Valenzise, and F. Dufaux, "An extensive performance evaluation of full-reference hdr image quality metrics," *Qual. User Exper.*, vol. 2, no. 1, pp. 1–16, Apr. 2017.

[61] M. Narwaria, R. K. Mantiuk, M. P. Da Silva, and P. Le Callet, "HDR-VDP-2.2: A calibrated method for objective quality prediction of high-dynamic range and standard images," *Proc. SPIE*, vol. 24, no. 1, Jan. 2015, Art. no. 010501.

[62] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video denoising, deblocking, and enhancement through separable 4-D nonlocal spatiotemporal transforms," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 3952–3966, Sep. 2012.

SHUNSUKE ONO (S'11–M'15) received the B.E. degree in computer science, in 2010, and the M.E. and Ph.D. degrees in communications and computer engineering, in 2012 and 2014, respectively, from the Tokyo Institute of Technology. From April 2012 to September 2014, he was a Research Fellow (DC1) of the Japan Society for the Promotion of Science (JSPS). He is currently an Associate Professor with the Department of Computer Science, School of Computing, Tokyo Institute of Technology. Since 2016, he has also been a Researcher with the Precursory Research for Embryonic Science and Technology (PRESTO), Japan Science and Technology Corporation (JST), Tokyo, Japan. His research interests include image processing, signal processing, mathematical optimization, and data science.

Dr. Ono received the Young Researchers' Award from the IEICE, in 2013, the Excellent Paper Award from the IEICE, in 2014, the Outstanding Student Journal Paper Award from the IEEE SPS Japan Chapter, in 2014, the Yasujiro Niwa Outstanding Paper Award from Tokyo Denki University, in 2015, the Seiichi Tejima Doctoral Dissertation Award from the Tokyo Institute of Technology, in 2016, the TELECOM System Technology Award from the Telecommunications Advancement Foundation, in 2016, and the Funai Research Award from the Funai Foundation, in 2017. He has been an Associate Editor of the IEEE Transactions on Signal and Information Processing over Networks, since 2019.

RYO MATSUOKA (S'12–M'16) received the B.E., M.E., and Ph.D. degrees from The University of Kitakyushu, Kitakyushu, Japan, in 2011, 2013, and 2016, respectively. He has been an Assistant Professor with the Faculty of Engineering, Electronics and Information Engineering, Kagawa University, from 2016 to 2019. He is currently an Assistant Professor with the Department of Artificial Intelligence, Kyushu Institute Technology. His research interests include multiple image-based image processing and digital filter design. He received the Excellent Award of the Hibikino Award for master's thesis and the Excellent Student Award of the IEEE Fukuoka Section, in 2013.

MASAHIRO OKUDA received the B.E., M.E., and Dr. Eng. degrees from Keio University, Yokohama, Japan, in 1993, 1995, and 1998, respectively. He was with the University of California at Santa Barbara, Santa Barbara, CA, USA, and Carnegie Mellon University, Pittsburgh, PA, USA, as a Visiting Scholar, in 1998 and 1999, respectively. From 1996 to 2000, he was a Research Fellow of the Japan Society for the Promotion of Science. From 2000 to 2010, he was an Associate Professor with the Faculty of Environmental Engineering, The University of Kitakyushu, Kitakyushu, Japan. He is currently a Professor with faculty of The University of Kitakyushu. His research interests include image restoration, high dynamic range imaging, multiple image fusion, and digital filter design. He received the SIP Distinguished Contribution Award, in 2013, the IE Award, and the Contribution Award from IEICE, in 2017.

• • •