

Journal of Electronic Imaging

JElectronicImaging.org

Low-light color image enhancement via iterative noise reduction using RGB/NIR sensor

Hiroki Yamashita
Daisuke Sugimura
Takayuki Hamamoto



Hiroki Yamashita, Daisuke Sugimura, Takayuki Hamamoto, “Low-light color image enhancement via iterative noise reduction using RGB/NIR sensor,” *J. Electron. Imaging* **26**(4), 043017 (2017), doi: 10.1117/1.JEI.26.4.043017.

Low-light color image enhancement via iterative noise reduction using RGB/NIR sensor

Hiroki Yamashita,* Daisuke Sugimura, and Takayuki Hamamoto
Tokyo University of Science, Department of Electrical Engineering, Tokyo, Japan

Abstract. We propose a method to enhance the color image of a low-light scene using a single sensor that simultaneously captures red, green, blue (RGB), and near-infrared (NIR) information. Typical image enhancement methods require two sensors to simultaneously capture color and NIR images. In contrast, our proposed system utilizes a single sensor but achieves accurate color image restoration. We divide the captured multispectral data into RGB and NIR information based on the spectral sensitivity of our imaging system. Using the NIR information for guidance, we reconstruct the corresponding color image based on a joint demosaicking and denoising technique. Subsequently, we restore the estimated color image iteratively using the constructed guidance image. Our experiments demonstrate the effectiveness of our method using synthetic data, and real raw data captured by our imaging system. © 2017 SPIE and IS&T [DOI: [10.1117/1.JEI.26.4.043017](https://doi.org/10.1117/1.JEI.26.4.043017)]

Keywords: low-light scene; red, green, blue/near-infrared single sensor; denoising; demosaicking.

Paper 170156 received Mar. 2, 2017; accepted for publication Jul. 18, 2017; published online Aug. 19, 2017.

1 Introduction

Color image restoration is a major research subject in the field of image processing. In particular, color image restoration of low-light scenes has attracted significant interest. The removal of a large amount of noise from a captured color image is difficult in low-light scenes. Numerous denoising algorithms have been proposed to overcome this crucial problem.^{1–3}

The simplest approach is to use a flash unit to add light to the scene; thus, images can be captured with less noise. However, this approach changes the tone of the resulting restored image because the color temperatures between the ambient illuminations and the flash light are different. To mitigate the effects of using a flash unit, image enhancement methods^{4–6} have been proposed. They utilized a pair of images of the low-light scene: one with flash and one without flash. These methods reconstructed clear color images using the flash/no-flash image pairs. However, the applicability of these methods is restricted to static scenes because these approaches involve capturing the pair of images at different times.

To overcome this limitation, image enhancement methods that exploit images obtained with near-infrared (NIR) flash have been proposed.^{7–15} These methods enable the simultaneous capture of the flash/no-flash image pairs without changes in the color tone because the wavelength of the NIR flash is different from that of the visible flash. Thus, they can perform noise reductions of low-light color images in dynamic scenes. However, these methods require two sensors to simultaneously capture two types of images (color and NIR images). In this case, meticulous calibration is required to adjust the pixel positions of the two sensors. If even a small mismatch exists between the sensors, these methods fail to reconstruct clear color images. Therefore, a

calibration-free imaging system for the enhancement of low-light color images is desirable.

In this study, we propose a method to enhance low-light color images using a single sensor that simultaneously captures red, green, blue (RGB) and NIR information. Our single sensor that captures RGB and NIR information requires no severe calibrations to achieve low-light image enhancement.

In fact, RGB/NIR single sensors that can simultaneously capture RGB and NIR information using a specific color filter array (CFA) have also been proposed.^{16–23} However, the main objective of these studies is to separate the captured raw data accurately into color and NIR images. Thus, they offer no major contributions to noise reduction in low-light color images.

In contrast to these methods, our proposed method reconstructs a clear color image of a low-light scene using raw data containing RGB and NIR information captured by our single sensor. We separate the captured raw data into RGB and NIR information based on the spectral sensitivity of our imaging system. Using the extracted NIR information, we achieve image restoration using an iterative process to improve the input low-light image.

The main contribution of our study is as follows: unlike previously proposed methods^{7,9–11,15} that require two meticulously calibrated sensors, our system uses a calibration-free single sensor and achieves accurate color image restoration. Unlike methods that are aimed at separating the RGB/NIR raw data into color and NIR images,^{16–22} we achieve low-light color image enhancement using RGB/NIR raw data captured using our single sensor.

This study is the extended and more detailed version of our study presented previously.²⁴ The new contributions are as follows: first, we extend our color image restoration by introducing a scheme for the iterative restoration of the deteriorated color image. This processing allows us to

*Address all correspondence to: Hiroki Yamashita, E-mail: yamashita@isl.ee.kagu.tus.ac.jp

improve the output obtained by our joint demosaicking and denoising technique.²⁴ Second, we report the simulation evaluation results to observe the effectiveness of our method in a quantitative manner. Last, we show additional experimental results using numerous real raw data.

The rest of this paper is organized as follows: Section 2 presents the details of our imaging system. Sections 3–5 describe our restoration algorithm. In particular, Sec. 3 provides the construction details of our guidance image. Section 4 describes the method to separate the RGB and NIR information from the raw data. Section 5 presents our color image restoration algorithm. In Sec. 6, we present the experimental results using synthetic data, and real raw data captured using our imaging system. We conclude with a summary in Sec. 7.

2 Low-Light Color Image Restoration Using a Single Sensor

The goal of this study is to enhance a low-light color image using a single sensor that captures RGB and NIR information simultaneously. Figure 1 provides an overview of our system. We first interpolate the NIR image from raw data captured by our imaging system. Using the interpolated NIR image, we then construct a guidance image. Following this, we extract the RGB information from the raw data using such a NIR image. We synthesize a clear color image from the extracted noisy RGB information using the guidance image as an aid.

2.1 Imaging System

Our imaging system configures a normal CMOS image sensor with our specific CFA. Our CFA has four channels that record the R + NIR (R'), G + NIR (G'), B + NIR (B'), and NIR (I) components shown in Fig. 2. Further, our imaging system utilizes a dual bandpass filter (DBPF) that filters out incoming light in the wavelength range from 680 to 820 nm. The relative spectral sensitivity of our CFA and the transmittance characteristics of the DBPF are shown in Fig. 3(a). The relative spectral sensitivity of our imaging system (i.e., CFA with DBPF) is also shown in Fig. 3(b). This shows that the spectral sensitivity of each channel exhibits almost identical

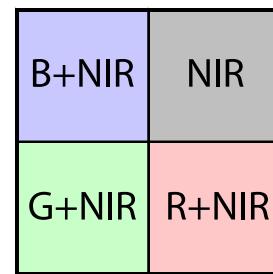


Fig. 2 CFA. Our CFA records the $R + \text{NIR}$, $G + \text{NIR}$, $B + \text{NIR}$, and NIR components.

characteristics for the wavelengths from 820 to 920 nm. This suggests that the RGB information can be extracted from the raw data by utilizing the characteristics of spectral sensitivity of the NIR channel; this is in accordance with the previous method.²¹

We utilize this imaging system with an NIR flash to acquire images of low-light scenes. As discussed in Sec. 1, the use of an NIR flash allows us to obtain NIR information with less noise, while low-intensity RGB information is obtained without large changes in the color tone.

3 Guidance Image Acquisition

According to previous methods,^{7,9–11,15} using NIR flash images as a guide is highly effective for enhancing color images of low-light scenes. Therefore, we construct a guidance image including NIR information from the captured raw data.

3.1 Demosaicking of Raw Data

Let the full resolution image be $\mathbf{y} = [\mathbf{y}_{R'}, \mathbf{y}_{G'}, \mathbf{y}_{B'}, \mathbf{y}_I]^T$, where R' , G' , B' , and I denote the $R + \text{NIR}$, $G + \text{NIR}$, $B + \text{NIR}$, and NIR components, respectively. To infer the optimal value of \mathbf{y} , we first interpolate each channel data based on our CFA by demosaicking processing; this is similar to the previous method.²⁵

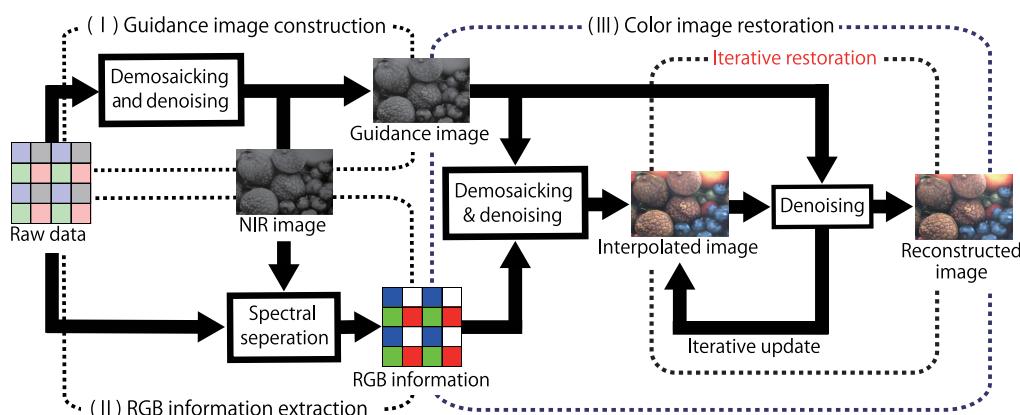


Fig. 1 Overview of our system: (I) Guidance image construction. We construct the NIR image using raw data captured by our imaging system. Using this NIR information, we construct a guidance image. (II) RGB information extraction. We extract the RGB information from the raw data using the constructed NIR image. (III) Color image restoration. We perform joint demosaicking and denoising of the noisy RGB information using the guidance image. Using the response from this joint demosaicking and denoising, we perform iterative restoration processing to acquire a final clear color image.

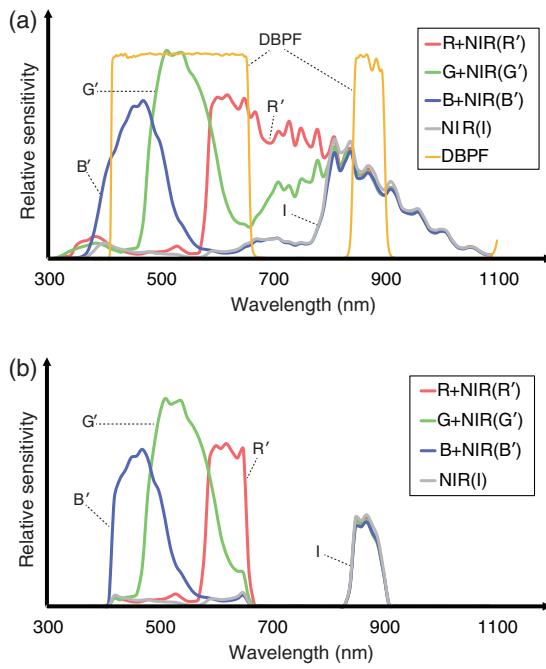


Fig. 3 Relative spectral sensitivity of our imaging system: (a) relative spectral sensitivity of our CFA and transmittance characteristics of the DBPF and (b) relative spectral sensitivity considering the DBPF.

3.1.1 Energy function

We minimize the following energy function with the observed raw data \mathbf{y}_S to obtain the latent full resolution image \mathbf{y} :

$$J(\mathbf{y}) = \|\mathbf{y}_S - \mathbf{H}_1 \mathbf{y}\|_p + \mu_1 J_1(\mathbf{y}) + \mu_2 J_2(\mathbf{y}) + \mu_3 J_3(\mathbf{y}), \quad (1)$$

where the control parameters in this energy function are defined as μ_1 , μ_2 , and μ_3 .

The first term in Eq. (1) represents the constraint for which \mathbf{y} tends to approximate \mathbf{y}_S . This term is configured by an ℓ_p norm ($0 < p < 1$) that facilitates noise reduction, as reported in the previous method.¹¹ The diagonal matrix $\mathbf{H}_1 = [\mathbf{H}_{R'}, \mathbf{H}_{G'}, \mathbf{H}_{B'}, \mathbf{H}_I]$ corresponds to our CFA sampling of the four channels. For example, the elements of $\mathbf{H}_{R'}$ equal 1 for pixels with captured R + NIR information, and 0 otherwise.

The second term characterizes the smoothness constraint for the individual channel; it is represented by a quadratic form: $J_1(\mathbf{y}) = \|\mathbf{F}_1 \mathbf{y}\|_2^2$. The matrix \mathbf{F}_1 represents the high-pass filter for the individual channel with respect to the vertical and horizontal directions. These respective filter coefficients are defined as $[-0.5, 1, -0.5]$ and $[-0.5, 1, -0.5]^T$.

For the remaining regularization terms (third and fourth terms), we describe the details in the subsequent sections because they contribute substantially to the performance enhancement of demosaicking of \mathbf{y} .

3.1.2 Spatially adaptive smoothness for individual channel

The third term $J_2(\mathbf{y})$ models a spatially adaptive regularization for the individual channel of \mathbf{y} . According to the previous method,²⁶ it is represented as

$$J_2(\mathbf{y}) = \|\mathbf{F}_1 \mathbf{y}\|_{\mathbf{W}_y}^2. \quad (2)$$

The diagonal matrix \mathbf{W}_y is introduced to adapt the penalty to the local features of \mathbf{y} . It contributes to the performance enhancement of image interpolation. The diagonal element of \mathbf{W}_y depends on the horizontal and vertical high-frequency components of \mathbf{y} . Specifically, \mathbf{W}_y is represented as

$$\mathbf{W}_y = \text{diag}(\mathbf{W}_{R',h}, \mathbf{W}_{R',v}, \mathbf{W}_{G',h}, \mathbf{W}_{G',v}, \mathbf{W}_{B',h}, \mathbf{W}_{B',v}, \mathbf{W}_{I,h}, \mathbf{W}_{I,v}), \quad (3)$$

where $\mathbf{W}_{K,h}$ and $\mathbf{W}_{K,v}$ ($K \in \{R', G', B', I\}$) represent the diagonal matrices characterizing the penalty for the horizontal and vertical components of the K 'th channel in \mathbf{W}_y , respectively. The respective j 'th elements of $\mathbf{W}_{K,h}$ and $\mathbf{W}_{K,v}$ are given as

$$\begin{aligned} \{\mathbf{W}_{K,h}\}_j &= \delta\left(\frac{\{\mathbf{e}_{K,h}\}_j}{\{\mathbf{e}_{K,h}\}_j + \{\mathbf{e}_{K,v}\}_j}\right), \\ \{\mathbf{W}_{K,v}\}_j &= \delta\left(\frac{\{\mathbf{e}_{K,v}\}_j}{\{\mathbf{e}_{K,h}\}_j + \{\mathbf{e}_{K,v}\}_j}\right), \end{aligned} \quad (4)$$

where $\{\mathbf{e}_{K,h}\}_j$ and $\{\mathbf{e}_{K,v}\}_j$ are the energies of the j 'th element of the horizontal and vertical high-frequency components of \mathbf{y}_K ; they are computed with filter coefficients $[-0.5, 1, -0.5]$. Further, $\delta(\cdot)$ is a function that is defined as

$$\delta(a) = \begin{cases} 0 & \text{if } a < \epsilon \\ \frac{a-\epsilon}{1-2\epsilon} & \text{if } \epsilon \leq a \leq 1-\epsilon \\ 1 & \text{if } a > 1-\epsilon, \end{cases} \quad (5)$$

with the parameter $\epsilon = 0.25$.

3.1.3 Regularization for signal residuals

In a low-light scene, our imaging system with NIR flash captures considerably more NIR than RGB information. Thus, a high NIR signal correlation is observed between any two channels. This correlation is subsequently utilized for image interpolation. Based on this, we model a regularization for the signal residuals in the fourth term $J_3(\mathbf{y})$; it is represented as

$$J_3(\mathbf{y}) = \sum_{(K,K') \in \Omega_y} \|\mathbf{F}_2 \mathbf{y}_K - \mathbf{F}_2 \mathbf{y}_{K'}\|_2^2, \quad (6)$$

where Ω_y is a combination set where (K, K') use two channels from R' , G' , B' , and I . Furthermore, \mathbf{F}_2 represents a high-pass filter for the signal residuals. According to the existing work,²⁶ the smoother filter coefficients are preferable in enforcing the smoothness of the signal residuals between different two channels. Thus, we set the filter coefficients of \mathbf{F}_2 for the vertical and horizontal directions as $[0.2, -0.5, 0.65, -0.5, 0.2]$ and $[0.2, -0.5, 0.65, -0.5, 0.2]^T$, respectively.

3.1.4 Energy minimization

The energy function $J(\mathbf{y})$ [Eq. (1)] becomes nonconvex owing to the inclusion of ℓ_p norm in the first term. Further,

\mathbf{W}_y in the third term makes $J(\mathbf{y})$ nonlinear. Thus, we explore the solution \mathbf{y} by minimizing $J(\mathbf{y})$ numerically.

We first approximate $J(\mathbf{y})$ so as to be convex using the iterative reweighted least squares (IRLS) method. The first term is approximated as the weighted square form

$$\|\mathbf{y}_S - \mathbf{H}_1\mathbf{y}\|_p \approx (\mathbf{y}_S - \mathbf{H}_1\mathbf{y})^T \mathbf{V} (\mathbf{y}_S - \mathbf{H}_1\mathbf{y}), \quad (7)$$

where \mathbf{V} is a diagonal matrix represented as

$$\mathbf{V} = \phi_p(\mathbf{y}_S - \mathbf{H}_1\mathbf{y}). \quad (8)$$

The j 'th element of the diagonal matrix $\phi_p(\mathbf{A})$ is defined by the IRLS method

$$\{\phi_p(\mathbf{A})\}_j = |\{\mathbf{A}\}_j|^{p-2}. \quad (9)$$

With the given \mathbf{W}_y , the third term is represented by the following weighted square form:

$$\|\mathbf{F}_1\mathbf{y}\|_{\mathbf{W}_y}^2 = (\mathbf{F}_1\mathbf{y})^T \mathbf{W}_y (\mathbf{F}_1\mathbf{y}). \quad (10)$$

By substituting Eqs. (7) and (10) into Eq. (1), $J(\mathbf{y})$ becomes linear. Thus, the solution \mathbf{y} can be obtained by applying a direct solver to Eq. (7). Based on the estimated solution \mathbf{y} , \mathbf{V} is updated using Eq. (9), while \mathbf{W}_y is computed using Eq. (4).

In the initial processing step, we perform denoising of \mathbf{y} using the BM3D method² to suppress the influences of the imposed noises to obtain reliable \mathbf{y} , similar to the method.²⁵

The above computation is iteratively processed until the value of $J(\mathbf{y})$ converges. In this manner, we acquire the interpolated full resolution images $\hat{\mathbf{y}} = [\hat{\mathbf{y}}_{R'}, \hat{\mathbf{y}}_{G'}, \hat{\mathbf{y}}_{B'}, \hat{\mathbf{y}}_I]^T$.

The summary of our demosaicking processing is shown in Algorithm 1.

Algorithm 1 Demosaicking of raw data.

Input: Raw data \mathbf{y}_S , parameters μ_1 , μ_2 , and μ_3

- 1: Initialize $\mathbf{W}_y \leftarrow \text{diag}(0, \dots, 0)$
- 2: Initialize $\mathbf{V} \leftarrow \text{diag}(1, \dots, 1)$
- 3: Estimate \mathbf{y} by minimizing Eq. (1) with a direct solver
- 4: Denoising of \mathbf{y} using BM3D²
- 5: **repeat**
- 6: Compute \mathbf{W}_y using Eq. (4)
- 7: Update \mathbf{V} using Eq. (9)
- 8: Estimate \mathbf{y} by minimizing Eq. (1) with a direct solver
- 9: **until** $J(\mathbf{y})$ converges

Output: Interpolated image $\hat{\mathbf{y}}$

3.2 Guidance Image Construction

Using $\hat{\mathbf{y}}$, we construct a guidance image \mathbf{z} . As mentioned, significantly more NIR signals than RGB ones tend to be included in all channels. Therefore, we can assume that $\hat{\mathbf{y}}$ is likely to contain less noise. Based on this concept, we construct a guidance image as $\mathbf{z} = (\hat{\mathbf{y}}_R + \hat{\mathbf{y}}_G + \hat{\mathbf{y}}_B + \hat{\mathbf{y}}_I)/4$. The integration of the images in all channels contributes to noise reduction in the guidance image, and thus a high-quality guidance image is obtained.

4 RGB Information Extraction from Raw Data

We extract the RGB components \mathbf{y}_C from the captured raw data \mathbf{y}_S using the constructed NIR image $\hat{\mathbf{y}}_I$. In fact, RGB and NIR information obtained using an RGB/NIR single sensor can be separated, as reported in previous works.^{16–22} We separate the RGB and NIR information using the method.²¹ Specifically, \mathbf{y}_C is acquired by

$$\mathbf{y}_C = \sum_{K \in R', G', B'} \mathbf{H}_K (\mathbf{y}_S - T_K \hat{\mathbf{y}}_I), \quad (11)$$

where $\mathbf{H}_K (K \in R', G', B')$ is a matrix whose elements are 1 for pixels where the component of K 'th channel is captured, and 0 otherwise. In addition, $T_K (K \in R', G', B')$ is weight that compensates for the differences in the NIR spectral sensitivities of each channel.

We determine T_K by measuring the ratio of the area of the spectral sensitivity curve of the K 'th channel to that of the NIR channel. Let the spectral sensitivity of the K 'th channel be $R_K(\lambda)$ and that of the NIR channel be $R_I(\lambda)$. Because $R_I(\lambda)$ has high sensitivity in the wavelength range from 820 to 920 nm, we compute T_K as

$$T_K = \frac{\int_{820}^{920} R_K(\lambda) d\lambda}{\int_{820}^{920} R_I(\lambda) d\lambda}, \quad K \in R', G', B'. \quad (12)$$

In this manner, we acquire the RGB information \mathbf{y}_C from the raw data.

Then, we perform histogram-stretching using the MATLAB® function “imadjust,” and a white-balance adjustment.²⁷

Figure 4 shows an example of the results of our RGB information extraction for a low-light scene (scene 1). We can see that the intensity of the obtained color image is very low [Fig. 4(c)]. Figure 4(d) shows the color image obtained by applying a histogram-stretching method in MATLAB® (imadjust), as well as the white-balance adjustment²⁷ of the low-light color image. We can see that the RGB information was extracted successfully. However, we can observe that the obtained color image was deteriorated owing to heavy noise.

5 Color Image Restoration Using Guidance Image

We reconstruct a clear color image \mathbf{x} from the noisy RGB information \mathbf{y}_C with the aid of the guidance image \mathbf{z} .

5.1 Likelihood of RGB Information in Raw Data

The latent clear color image $\mathbf{x} = [\mathbf{x}_R, \mathbf{x}_G, \mathbf{x}_B]^T$ tends to approximate the RGB components \mathbf{y}_C . We impose the constraint $E_1(\mathbf{x})$ using \mathbf{y}_C as

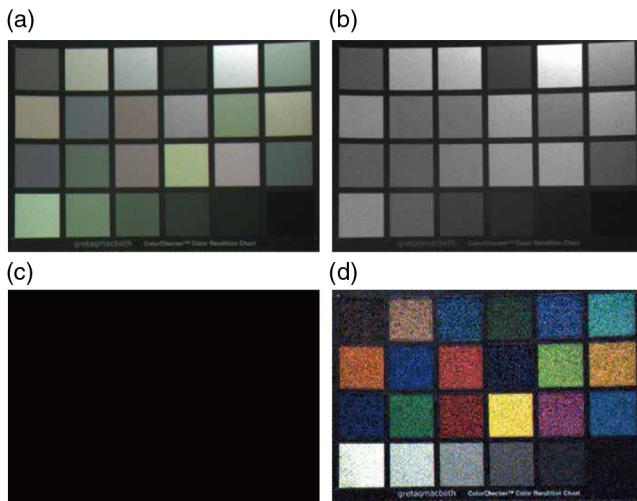


Fig. 4 RGB information extraction from the captured raw data (scene 1): (a) RGB + NIR raw image acquired using a bilinear interpolation, (b) constructed NIR image $\hat{\mathbf{y}}_l$, (c) low-light color image obtained by interpolating the extracted RGB information \mathbf{y}_C using bilinear interpolation, and (d) color image obtained by applying the histogram-stretching method in MATLAB (imadjust), and a white-balance adjustment.²⁷

$$E_1(\mathbf{x}) = \|\mathbf{y}_C - \mathbf{H}_2\mathbf{x}\|_p, \quad (13)$$

where $\mathbf{H}_2 = [\mathbf{H}_{R'}, \mathbf{H}_{G'}, \mathbf{H}_{B'}]$ are matrices corresponding to our CFA sampling of the three color channels (red, green, and blue).

5.2 Gradient Correlation

We impose a constraint on exploring the latent color image \mathbf{x} based on the structure of the guidance image \mathbf{z} . However, a structural discrepancy may exist between the color and NIR images. Such structural discrepancy influences the performance of color restoration. To address this problem, we employ a scale map¹¹ to compensate for the structural inconsistencies between the two images. Thus, the constraint $E_2(\mathbf{x}, \mathbf{s})$ is introduced using the gradient of \mathbf{x} and \mathbf{z} with a scale map $\mathbf{s} = [s_R, s_G, s_B]^T$ as

$$E_2(\mathbf{x}, \mathbf{s}) = \sum_{K \in R, G, B} \|\mathbf{F}_3\mathbf{x}_K - s_K \mathbf{F}_3\mathbf{z}\|_p, \quad (14)$$

where \mathbf{F}_3 denotes an operator that calculates the gradients with respect to the vertical and horizontal directions for a single channel with the filter coefficients $[-1, 1]$ and $[-1, 1]^T$.

To ensure the smoothness and sparsity of \mathbf{s} over the image, we impose the constraint $E_3(\mathbf{s})$ for \mathbf{s} . Recall

$$\begin{aligned} E_3(\mathbf{s}) = & \sum_{K \in R, G, B} \sum_j \frac{\eta^2}{\{\mathbf{F}_3\mathbf{z}\}_j^2 + 2\eta^2} \left(\frac{\{\mathbf{F}_3\mathbf{z}\}_j^T}{|\{\mathbf{F}_3\mathbf{z}\}_j|} \{\mathbf{F}_3\mathbf{s}_K\}_j \right)^2 \\ & + \frac{\{\mathbf{F}_3\mathbf{z}\}_j^2 + \eta^2}{\{\mathbf{F}_3\mathbf{z}\}_j^2 + 2\eta^2} \left(\frac{\{\mathbf{F}_3\mathbf{z}\}_j^{\perp T}}{|\{\mathbf{F}_3\mathbf{z}\}_j|} \{\mathbf{F}_3\mathbf{s}_K\}_j \right)^2, \end{aligned} \quad (15)$$

where $\{\mathbf{F}_3\mathbf{z}\}_j^\perp$ represents a vector perpendicular to $\{\mathbf{F}_3\mathbf{z}\}_j$ and η controls the isotropic smoothness (we set $\eta = 0.1$). For more details, see Yan et al.'s work.¹¹

5.3 Color Correlation in Wavelength Domain

A scale map \mathbf{s} cannot capture the structural inconsistencies in the nontextured regions of the guidance image (i.e., each element of \mathbf{s} is 0). In this region, an excessive smoothness is observed because the constraint in Eq. (14) produces a form equivalent to that of the gradient smoothing of latent color images.

To achieve adequate smoothness in such regions, we model the constraint $E_4(\mathbf{x})$ based on the color correlation in the wavelength domain. It is widely accepted that the gradients of a color image between nearby wavelengths are highly correlated. Therefore, we model the constraint $E_4(\mathbf{x})$ as

$$E_4(\mathbf{x}) = \sum_{(K, K') \in \Omega_x} \|\mathbf{F}_3\mathbf{x}_K - \mathbf{F}_3\mathbf{x}_{K'}\|_p, \quad (16)$$

where Ω_x is a combination set of (K, K') , whose K and K' correspond to two channels from R, G, and B. The constraint $E_4(\mathbf{x})$ compensates for the loss of the necessary texture of the color image \mathbf{x} owing to the gradient loss in the guidance image \mathbf{z} .

5.4 Iterative Color Image Restoration

We employ an iterative noise reduction scheme to enhance the performance of the noise reduction process.

We define the color image estimated at the $k-1$ 'th iteration as $\mathbf{x}^{(k-1)}$. We model the likelihood term $\tilde{E}_1^{(k)}(\mathbf{x})$ by considering $\mathbf{x}^{(k-1)}$ as the observed data; thus, it is represented as

$$\tilde{E}_1^{(k)}(\mathbf{x}) = \|\mathbf{x} - \mathbf{x}^{(k-1)}\|_p. \quad (17)$$

This likelihood term allows us to effectively reduce the remaining noise, primarily because $\mathbf{x}^{(k-1)}$ improves as the number of iterations increases.

In the initial step (i.e., $k = 0$), the color image estimated in the previous iteration is not given. Thus, we utilize the constraint Eq. (13) with the RGB components \mathbf{y}_C to obtain the initial color image. Consequently, we reformulate the likelihood term depending on the iteration number as

$$\tilde{E}_1^{(k)}(\mathbf{x}) = \begin{cases} \|\mathbf{y}_C - \mathbf{H}_2\mathbf{x}\|_p & \text{if } k = 0 \\ \|\mathbf{x} - \mathbf{x}^{(k-1)}\|_p & \text{otherwise.} \end{cases} \quad (18)$$

5.5 Energy Function

We model the energy function $E^{(k)}(\mathbf{x}, \mathbf{s})$ at the k 'th iteration. In fact, our CFA has pixels that constitute unobserved RGB information (i.e., the pixels where the NIR component is captured and RGB component is not captured). Thus, we minimize only $E^{(k)}(\mathbf{x}, \mathbf{s})$ for the pixels that captured RGB information.

The smoothness term described by Eq. (16) suppresses the loss of the necessary edges, where the elements of \mathbf{s} are 0. In the regions where the elements of \mathbf{s} are not 0, however, imposing the constraints of both Eqs. (14) and (16) tends to contribute excessive smoothness to \mathbf{x} . To avoid this, we apply Eqs. (14) and (16) adaptively to \mathbf{x} , depending on \mathbf{s} . Specifically, we control them using a binary mask $B(\mathbf{s})$. The j 'th element of $B(\mathbf{s})$ is set to $[B(\mathbf{s})]_j = 0$ where the j 'th element of \mathbf{s} is 0; otherwise, it is $[B(\mathbf{s})]_j = 1$.

Finally, we model $\tilde{E}^{(k)}(\mathbf{x}, \mathbf{s})$ as

$$\begin{aligned}\tilde{E}^{(k)}(\mathbf{x}, \mathbf{s}) = & \mathbf{U}\{\tilde{E}_1^{(k)}(\mathbf{x}) + \omega_1 B(\mathbf{s})E_2(\mathbf{x}, \mathbf{s}) + \omega_2 E_3(\mathbf{s}) \\ & + \omega_3 \bar{B}(\mathbf{s})E_4(\mathbf{x})\},\end{aligned}\quad (19)$$

where \mathbf{U} is a matrix whose elements are 1 for pixels where the RGB component is captured, and 0 otherwise. Furthermore, ω_1 , ω_2 , and ω_3 denote the control parameters.

In Eq. (19), we update the likelihood term $E_1^{(k)}(\mathbf{x})$ only, because the remaining terms are the regularization terms for the latent color image \mathbf{x} . In other words, these regularization terms are independent of the given color image, i.e., $\mathbf{x}^{(k-1)}$.

5.6 Numerical Solution

We consider the minimization of $\tilde{E}^{(k)}(\mathbf{x}, \mathbf{s})$ as a joint minimization problem with regard to \mathbf{x} and \mathbf{s} ; this is similar to previous works.^{11,15} The authors of these works^{11,15} solved joint minimization problems numerically by utilizing the IRLS method and the preconditioned conjugate gradient (PCG) method. We follow these approaches to solve the minimization problem of $\tilde{E}^{(k)}(\mathbf{x}, \mathbf{s})$.

We numerically solve the following two subproblems alternately and iteratively; the solutions at the k 'th iteration can be obtained as

$$\mathbf{x}^{(k)} = \underset{\mathbf{x}}{\operatorname{argmin}} \tilde{E}_x(\mathbf{x}), \quad (20)$$

$$\mathbf{s}^{(k)} = \underset{\mathbf{s}}{\operatorname{argmin}} \tilde{E}_s(\mathbf{s}), \quad (21)$$

where

$$\tilde{E}_x(\mathbf{x}) = \mathbf{U}\{\tilde{E}_1^{(k)}(\mathbf{x}) + \omega_1 B(\mathbf{s})E_2(\mathbf{x}| \mathbf{s}) + \omega_3 \bar{B}(\mathbf{s})E_4(\mathbf{x})\}, \quad (22)$$

$$\tilde{E}_s(\mathbf{s}) = \mathbf{U}\{\omega_1 B(\mathbf{s})E_2(\mathbf{s}| \mathbf{x}) + \omega_2 E_3(\mathbf{s})\}. \quad (23)$$

We solve Eq. (20) to obtain the latent color image $\mathbf{x}^{(k)}$ with the given scale map estimated in the previous iteration. Using the estimated $\mathbf{x}^{(k)}$, we solve Eq. (21) to obtain $\mathbf{s}^{(k)}$. This alternating iterative minimization procedure is executed until the value of $\tilde{E}^{(k)}(\mathbf{x}, \mathbf{s})$ converges.

The IRLS method approximates the ℓ_p norm constraints in Eqs. (13), (14), (16), and (17) as the weighted square forms in solving Eqs. (20) and (21)

$$\tilde{E}_1^{(k)}(\mathbf{x}) \approx \begin{cases} (\mathbf{y}_C - \mathbf{H}_2 \mathbf{x})^T \mathbf{V}_1 (\mathbf{y}_C - \mathbf{H}_2 \mathbf{x}) & \text{if } k = 0 \\ (\mathbf{x} - \mathbf{x}^{(k-1)})^T \mathbf{V}_1' (\mathbf{x} - \mathbf{x}^{(k-1)}) & \text{otherwise,} \end{cases} \quad (24)$$

$$E_2(\mathbf{x}, \mathbf{s}) \approx \sum_{K \in R, G, B} (\mathbf{F}_3 \mathbf{x}_K - \mathbf{s}_K \mathbf{F}_3 \mathbf{z})^T \mathbf{V}_2 (\mathbf{F}_3 \mathbf{x}_K - \mathbf{s}_K \mathbf{F}_3 \mathbf{z}), \quad (25)$$

$$E_4(\mathbf{x}) \approx \sum_{(K, K') \in \Omega_2} (\mathbf{F}_3 \mathbf{x}_K - \mathbf{F}_3 \mathbf{x}_{K'})^T \mathbf{V}_4 (\mathbf{F}_3 \mathbf{x}_K - \mathbf{F}_3 \mathbf{x}_{K'}), \quad (26)$$

where

$$\mathbf{V}_1 = \phi_p(\mathbf{y}_C - \mathbf{H}_2 \mathbf{x}), \quad (27)$$

$$\mathbf{V}_1' = \phi_p(\mathbf{x} - \mathbf{x}^{(k-1)}), \quad (28)$$

$$\mathbf{V}_2 = \phi_p(\mathbf{F}_3 \mathbf{x}_K - \mathbf{s}_K \mathbf{F}_3 \mathbf{z}), \quad (29)$$

$$\mathbf{V}_4 = \phi_p(\mathbf{F}_3 \mathbf{x}_K - \mathbf{F}_3 \mathbf{x}_{K'}). \quad (30)$$

Note that the diagonal matrix $\phi_p(\cdot)$ can be obtained using Eq. (9).

Based on the above approximations, we can solve Eqs. (20) and (23) using the PCG method because they become convex. Using the estimated solutions, the IRLS method updates their weights \mathbf{V}_1' , \mathbf{V}_2 , and \mathbf{V}_4 using Eq. (9).

Finally, following the convergence of $E^{(k)}(\mathbf{x}, \mathbf{s})$, we obtain the clear color image \mathbf{x}^* .

The summary of our color image restoration scheme is shown in Algorithm 2. The solvers for our subproblems are summarized in Algorithm 3 [solver for Eq. (20)] and Algorithm 4 [solver for Eq. (21)], respectively.

6 Experiments

In this section, we present experimental results that show the effectiveness of our method. However, it is difficult to compare our method directly with other methods. This is primarily because the CFA of our imaging system is specific; thus, it is difficult to apply existing demosaicking methods to the raw data captured by our imaging system. On the other hand, our color image restoration scheme [module (III) in Fig. 1] can be cast as a joint-upsampling problem; this is a research topic regarding full-resolution image

Algorithm 2 Color image restoration.

Input: RGB information \mathbf{y}_C , guidance image \mathbf{z} , parameters p , ω_1 , ω_2 , and ω_3

- 1: Initialize \mathbf{x} [i.e., $\mathbf{x}^{(-1)}$] by interpolating \mathbf{y}_C using bilinear interpolation
- 2: Initialize $\mathbf{s}^{(-1)} \leftarrow \mathbf{1}$ ($\mathbf{1}$ denotes a matrix whose all the elements are 1)
- 3: $k \leftarrow 0$
- 4: **repeat**
- 5: **if** $k = 0$ **then**
- 6: $\mathbf{U} \leftarrow \mathbf{H}_2$
- 7: **else**
- 8: $\mathbf{U} \leftarrow \mathbf{1}$
- 9: **end if**
- 10: Estimate $\mathbf{x}^{(k)}$ using Algorithm 3
- 11: Estimate $\mathbf{s}^{(k)}$ using Algorithm 4
- 12: $k \leftarrow k + 1$
- 13: **until** convergence

Output: Clear color image \mathbf{x}^*

Algorithm 3 Solver for Eq. (20) at the k 'th iteration.

Input: Color image $\mathbf{x}^{(k-1)}$, scale map $\mathbf{s}^{(k-1)}$, guidance image \mathbf{z} , parameters p , ω_1 , and ω_3

- 1: **if** $k = 0$ **then**
- 2: Initialize $\mathbf{V}_1 \leftarrow \text{diag}(1, \dots, 1)$, $\mathbf{V}'_1 \leftarrow \text{diag}(1, \dots, 1)$, $\mathbf{V}_2 \leftarrow \text{diag}(1, \dots, 1)$, and $\mathbf{V}_4 \leftarrow \text{diag}(1, \dots, 1)$
- 3: **else**
- 4: Update \mathbf{V}'_1 , \mathbf{V}_2 , and \mathbf{V}_4 using $\mathbf{x}^{(k-1)}$ and $\mathbf{s}^{(k-1)}$
- 5: **end if**
- 6: Estimate $\mathbf{x}^{(k)}$ by solving Eq. (20) using the PCG method

Output: Color image $\mathbf{x}^{(k)}$

Algorithm 4 Solver for Eq. (21) at the k 'th iteration.

Input: Color image $\mathbf{x}^{(k)}$, scale map $\mathbf{s}^{(k-1)}$, guidance image \mathbf{z} , parameters p , ω_1 , and ω_2

- 1: **if** $k > 0$ **then**
- 2: Update \mathbf{V}_2 using $\mathbf{x}^{(k)}$ and $\mathbf{s}^{(k-1)}$
- 3: **end if**
- 4: Estimate $\mathbf{s}^{(k)}$ by solving Eq. (21) using the PCG method

Output: Scale map $\mathbf{s}^{(k)}$

reconstruction by interpolating low-resolution images using a guidance image as an aid. Thus, we compared our color image restoration with state-of-the-art methods of joint-upsampling processing, using the same input signals, i.e., \mathbf{y}_C and \mathbf{z} . We utilized the following methods for these comparison experiments: (a) joint bilateral filtering (JBF),²⁸ (b) guided image filtering (GIF),²⁹ and (c) joint static and dynamic guidance filtering (SDF).³⁰

Our energy function consists of ℓ_p ($0 < p < 1$) norm constraints that involve an exponent value. This exponent value makes our energy function nonconvex; thus, the choice of p might affect the investigation of the optimal image. To observe the influences of p , we tested our method by varying p . In this preliminary experiment, we observed that the choice of p made little difference in the resulting color images. This was also reported in the previous work.¹⁵ In this experiment, we determined the exponent value as $p = 0.9$.

We conducted further preliminary experiments to determine optimal parameters using additional low-light images. Table 1 lists the parameters used for the performance evaluation. We use the same parameters for all experiments.

Our experiments were run on a Windows PC with an Intel Core i7-2600 3.4 GHz, and used MATLAB R2014a for our computation. In this computing environment, our algorithm cost ~ 230 s when processing a 1280×720 spatial resolution raw data of 12-bit pixels. In particular, the third step, color

Table 1 Parameters used for our algorithm. Note that the same parameters are used for all experiments.

Parameters in Eq. (1)			Parameters in Eq. (19)		
μ_1	μ_2	μ_3	ω_1	ω_2	ω_3
0.10	0.03	0.06	1.00	1.50	1.00

image restoration, is the most time-consuming, with a computational cost of ~ 220 s.

6.1 Results Using Synthetic Raw Data

First, we evaluated the performance of our color image restoration using synthetic data. We used a pair of RGB and NIR images provided in the dataset³¹ as the materials for our simulation. We used 20 pairs of RGB and NIR images from this dataset in this evaluation.

6.1.1 Producing synthetic raw data

As reported in previous works,^{32–35} the color image acquisition process in low-light scenes has been modeled with a mixed Poisson–Gaussian noise distribution. In contrast, our imaging system captures RGB and NIR information simultaneously. Thus, the existing image acquisition models^{32–35} are difficult to use directly for simulating raw data captured with our imaging system. In this simulation, we extend these existing image acquisition models^{32–35} to be applicable for simulating our imaging system.

We synthesize the raw data $\tilde{\mathbf{y}}_K$ ($K \in \mathbf{R}', \mathbf{G}', \mathbf{B}'$) based on a mixed Poisson–Gaussian noise model as

$$\tilde{\mathbf{y}}_K = \mathbf{K}' + \mathbf{I}' + \mathbf{N}_G, \quad (31)$$

with

$$\mathbf{K}' \sim \mathcal{P}(\lambda_K), \quad \mathbf{I}' \sim \mathcal{P}(\lambda_I), \quad \mathbf{N}_G \sim \mathcal{N}(0, \sigma^2), \quad (32)$$

where \mathbf{K}' and \mathbf{I}' denote the observed RGB and NIR information deteriorated by the photon-shot noise, respectively. Each piece of information was drawn from a Poisson distribution $\mathcal{P}(\lambda)$ with a parameter λ . Furthermore, a zero-mean Gaussian $\mathcal{N}(0, \sigma^2)$ models the readout noise of image sensor \mathbf{N}_G . Based on the characteristics of our imaging system, we approximate the parameters of Poisson distributions (λ_K and λ_I) as

$$\lambda_K = \zeta_V \mathbf{i}_K, \quad \lambda_I = \zeta_I T_K \mathbf{i}_I, \quad (33)$$

where \mathbf{i}_K and \mathbf{i}_I are the pixel values of the color and NIR images, respectively. In Eq. (33), the parameters ζ_V and ζ_I control the illumination levels of the color and NIR information, respectively. Because our imaging system utilizes an NIR flash unit when capturing low-light scenes, we set ζ_I to a larger value than ζ_V . Further, T_K represents the weights compensating the differences in spectral sensitivity, as described in Sec. 4.

Using the synthesized raw data $\tilde{\mathbf{y}}_K$, we construct the input guidance image \mathbf{z} and RGB information \mathbf{y}_C in accordance with modules (I) and (II) of the proposed method.

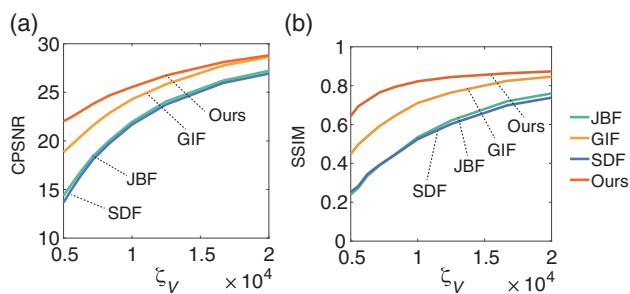


Fig. 5 Quantitative evaluation using synthetic raw data (a) CPSNR and (b) SSIM.

6.1.2 Quantitative comparisons

Using the synthesized raw data $\tilde{\mathbf{y}}_K$, we evaluate the proposed method in both quantitative and qualitative manners. In this paper, we show one example of the color restoration results by varying the simulation parameters.

The experimental setting is described as follows: the amount of readout noise of image sensor σ can be considered as a constant value, even when the illumination intensity of the target scenes varies. We maintain a high illumination intensity of NIR flash unit when capturing low-light scenes. Thus, we set ζ_1 to be a high constant value in this simulation evaluation. With this condition for our image acquisition

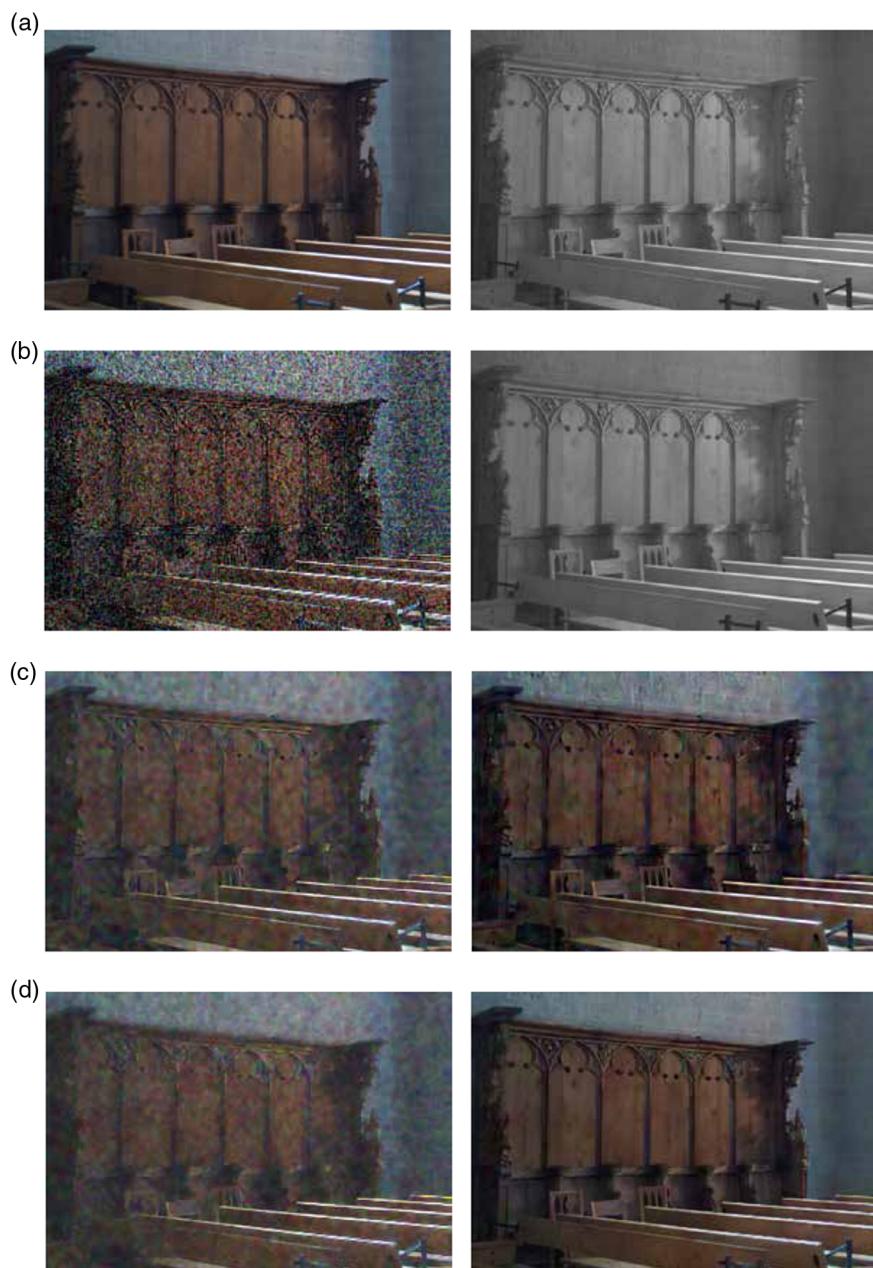


Fig. 6 Reconstruction results using synthetic raw data ($\zeta_V = 8.0 \times 10^3$): (a) original color and NIR images provided in the dataset,³¹ (b) simulated noisy color image and guidance image, (c) respective reconstruction results obtained using JBF²⁸ and GIF,²⁹ and (d) respective reconstruction results obtained using SDF³⁰ and our method.

process, we set σ and ζ_I such that the illumination intensity of the synthesized raw data is approximately the same as that of real raw data captured by our imaging system. Specifically, we set $\sigma = 5$ and $\zeta_I = 10^5$. On the other hand, ζ_V characterizes the illumination intensity of low-light scenes. Thus, we tested the performance of color image restoration with varying ζ_V .

Figure 5 shows changes in the color peak signal-to-noise ratio (CPSNR) and the structural similarity (SSIM) values with varying ζ_V . We observe that our color image restoration outperforms the other comparison methods. Figure 6 shows a visual comparison when $\zeta_V = 8.0 \times 10^3$. We can

qualitatively see that our method obtains better reconstruction results than the comparison methods.

6.2 Results Using Real Raw Data

We tested our algorithm using 20 real raw data captured by our imaging system. These raw data were taken in low-light scenes where the illumination intensity was ~ 0.2 to 0.3 lux.

We show four examples of our restoration results in this paper. Since it is difficult to obtain the ground truth data (clear color images) of the corresponding scenes, we evaluate the reconstruction performance in a qualitative manner.

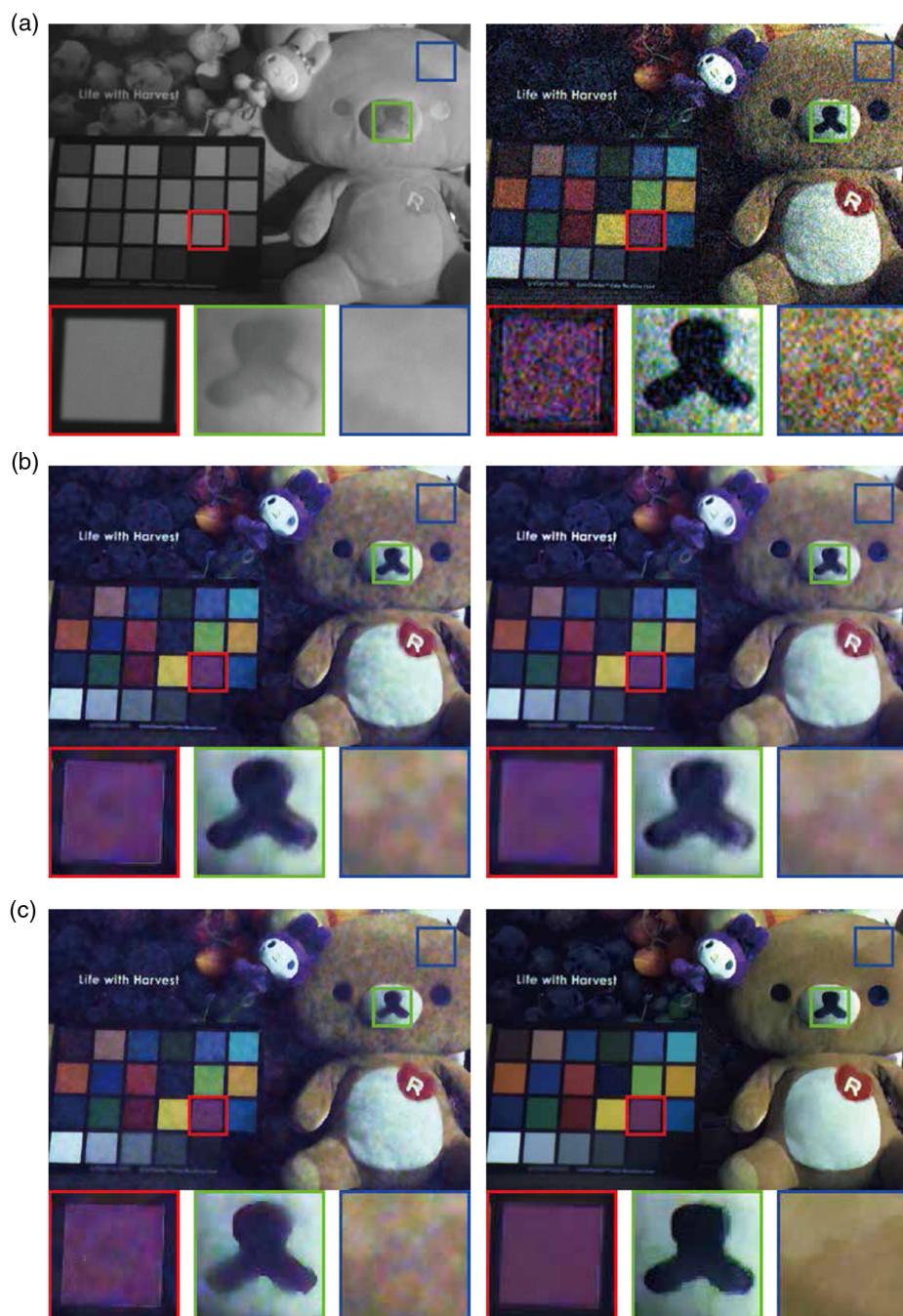


Fig. 7 Reconstruction results with comparison methods (scene 2): (a) guidance image and noisy color image, (b) respective reconstruction results obtained using JBF²⁸ and GIF²⁹ and (c) respective reconstruction results obtained using SDF³⁰ and our method.

6.2.1 Comparisons and discussions

Figures 7 and 8 show the reconstruction results for scenes 2 and 3, respectively. Note that the noisy color images were obtained by applying a histogram-stretching and a white-balance adjustment to the interpolated RGB information. The actual low-light color images had low-intensity pixel values (similar to Fig. 4). Figures 7 and 8 show qualitatively that our method obtains better results than those of the other methods (JBF, GIF, and SDF).

A discussion is required regarding the performance in color image restoration based on the results for scene 2. We can see that our method achieved effective noise

reduction in the nontexture regions (e.g., color chart). However, false contour artifacts were observed around the region of the stuffed bear's nose (green rectangular region). This is primarily because partial regions of the guidance image were not constructed with high quality. As described, our method reconstructs color images so as to be close to the structure of the guidance image. This indicates that the performance in color image restoration is highly dependent on the quality of the obtained guidance image. Although there are the strong gradients (edges of the stuffed bear's nose) in the green rectangular region in the input noisy color image, only weak spatial gradients were captured in the guidance

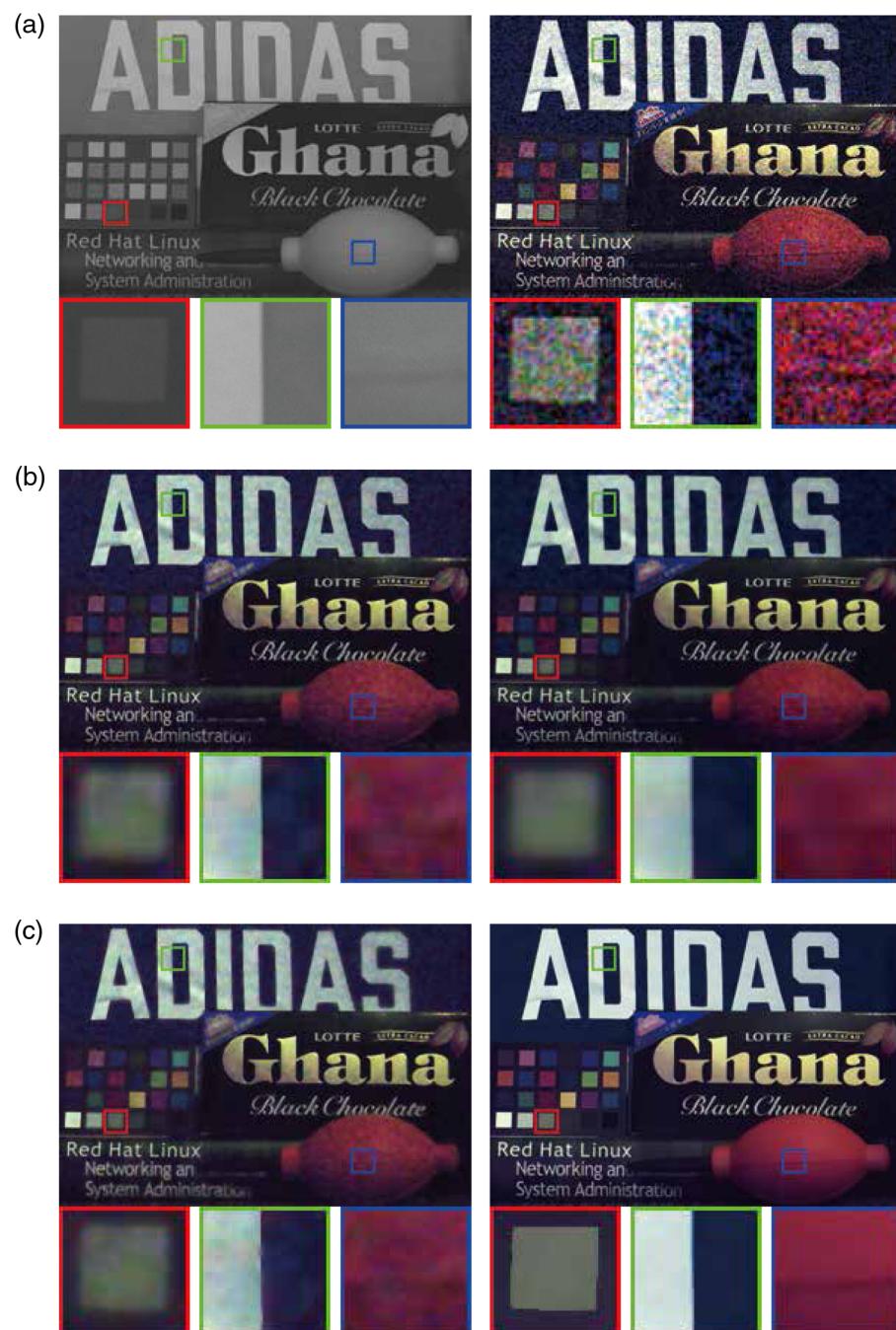


Fig. 8 Reconstruction results with comparison methods (scene 3): (a) guidance image and noisy color image, (b) respective reconstruction results obtained using JBF²⁸ and GIF,²⁹ and (c) respective reconstruction results obtained using SDF³⁰ and our method.

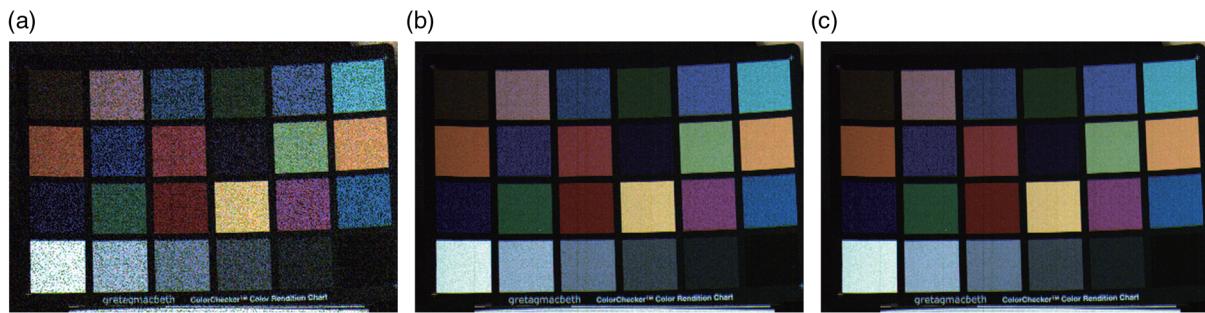


Fig. 9 Changes in the reference image with varying numbers of accumulation images [(a)–(c): $N = 1, 15, 30$].



Fig. 10 Comparison results with our previous method²⁴ (scene 4). (a) Constructed guidance image, (b) color image obtained by applying histogram-stretching and white balance adjustment,²⁷ (c) image enhancement result obtained using our previous method,²⁴ and (d) result obtained using our method.

image. Owing to such a result for the guidance image, boundaries between the edges and surrounding regions, which deteriorated with the green noise, were smoothed spatially. As a result, green false contour artifacts were generated in our result. In fact, such influences of the guidance image can also be observed in the results obtained using alternative methods. We can see that the reconstructed edge regions seem blurred because of the loss of the structural information of the guidance image. In addition, false textures observed in the guidance image were produced in the reconstructed color images. Therefore, to obtain better color image restoration results for all methods, it is important to investigate schemes to construct a better guidance image.

6.2.2 Comparison with multiple-shots accumulated image

In general, the accumulation of multiple shots for static scenes allows us to obtain an image with less noise. We discuss the performance in noise reduction by comparing the obtained results with the accumulated image. We name the accumulated image “reference” image. We captured the color checker board with the same exposure time as that for scene 2 (Fig. 7), under the similar illumination

environment with that in scene 2 (~ 0.2 lux). We then accumulated N shots to obtain the reference image.

Figure 9 shows changes in the reference image with varying N (the accumulation number of images). We can qualitatively see that our result is comparable with the reference image with $N = 30$. In contrast, noise reduction performance of the comparison methods appears insufficient, because visual artifacts owing to the noise can be observed in the color chart regions.

Consequently, our method is more effective than the comparison methods.

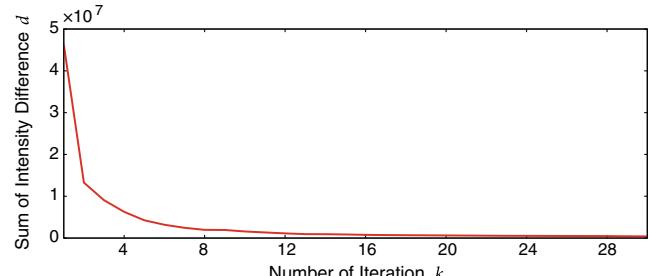


Fig. 11 Numerical convergence of our iterative restoration scheme with d (sum of intensity difference between $\mathbf{x}^{(k-1)}$ and $\mathbf{x}^{(k)}$) (scene 5).

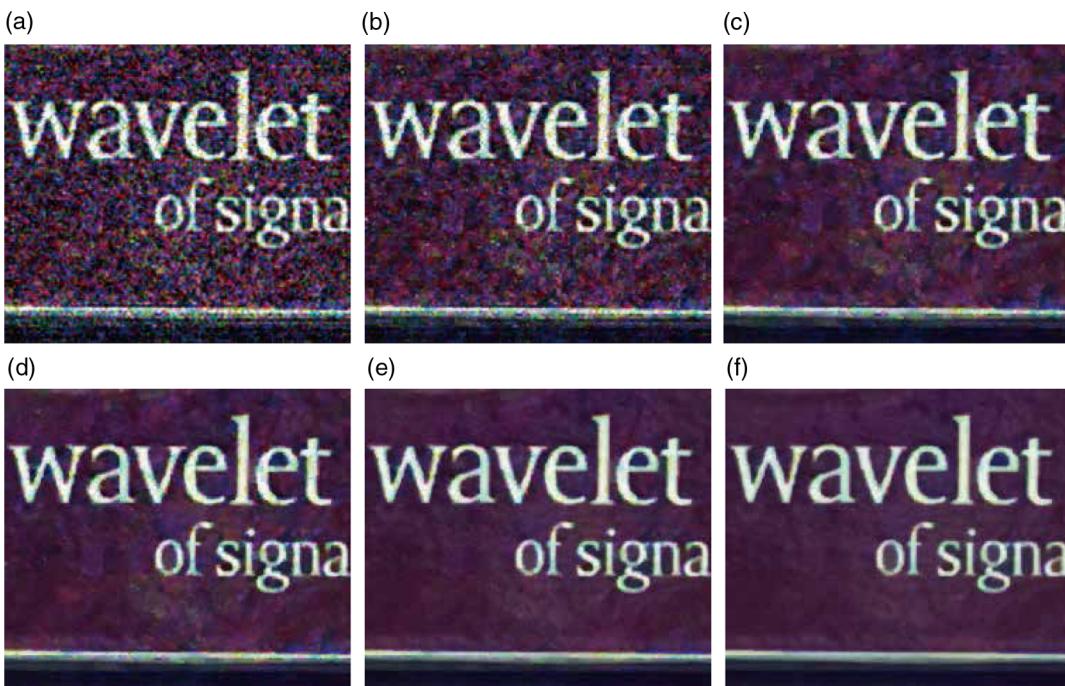


Fig. 12 Restoration results with the changes in the number of iterations (scene 5). (a)–(f) Respective reconstruction results obtained at iteration $k = -1$ (input image), 0, 3, 5, 10, 20.

6.2.3 Effects of iterative noise reduction

To observe the effects of our iterative image restoration, we compare our result with that obtained using our previous method (without the iterative processing for image restoration)²⁴ on another low-light scene (scene 4). Figure 10 shows this comparison result. We see that our current result is superior to that obtained using the previous method.²⁴

6.3 Numerical Convergence

As described in Sec. 5.4, our method estimates the optimal color image iteratively. Thus, it is important to demonstrate the numerical convergence of our method. To show the numerical convergence of our iterative restoration, we used a sum of intensity difference of the estimated image between two consecutive iterations, $d = |\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}|$.

Figure 11 shows the changes of d during iterations using another low-light raw data (scene 5). We can see that our iterative restoration numerically converges. In our procedure, 15 to 20 iterations are usually sufficient to generate visually compelling results. Figure 12 shows the changes in the estimated image $\mathbf{x}^{(k)}$ during the iterations. We can also see that the quality of the restored image improves with increasing the number of iterations.

7 Conclusion

We proposed and tested a system to recover a clear color image in a low-light scene, using a single sensor that captures RGB and NIR information simultaneously. In contrast to past image enhancement methods that require two sensors, our system allows us to obtain accurate color image restoration using a single sensor. To achieve image enhancement, the proposed method extracted RGB information from captured raw data based on the spectral sensitivity of our imaging system. Using the extracted RGB information, our

system reconstructed clear color images with the aid of a guidance image based on a joint demosaicking and denoising technique. To achieve further enhancement, we introduced an iterative restoration scheme.

Through the experiments, we demonstrated that our method was able to reconstruct a high-quality color image from raw data of low-light scenes. Moreover, we observed the superiority of our color image restoration over state-of-the-art methods, both qualitatively and quantitatively.

In this work, we determined the control parameters empirically, and then used them (Table 1). In fact, automatic estimation of optimal parameters has been considered a difficult task in related researches, as reported in previous works.^{11,15} We aim to conduct further research to explore solutions to the above remaining problems.

References

1. A. Buades and B. Coll, "A non-local algorithm for image denoising," in *IEEE Conf. on Computer Vision Pattern Recognition (CVPR)*, pp. 60–65 (2005).
2. K. Dabov et al., "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**(8), 2080–2095 (2007).
3. H. Liu et al., "Image denoising via adaptive soft-thresholding based on non-local samples," in *IEEE Conf. on Computer Vision Pattern Recognition (CVPR)*, pp. 484–492 (2015).
4. G. Petschnigg et al., "Digital photography with flash and no-flash image pairs," *ACM Trans. Graphics* **23**, 664–672 (2004).
5. E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," *ACM Trans. Graphics* **23**(3), 673–678 (2004).
6. A. Agrawal et al., "Removing photography artifacts using gradient projection and flash-exposure sampling," *ACM Trans. Graphics* **24**, 828–835 (2005).
7. E. P. Bennett, J. L. Mason, and L. McMillan, "Multispectral bilateral video fusion," *IEEE Trans. Image Process.* **16**, 1185–1194 (2007).
8. X. Zhang, T. Sim, and X. Miao, "Enhancing photographs with near infra-red images," in *IEEE Conf. on Computer Vision Pattern Recognition (CVPR)*, pp. 1–8 (2008).
9. S. Zhuo et al., "Enhancing low light images using near infrared flash images," in *IEEE Int. Conf. on Image Processing (ICIP)*, pp. 2537–2540 (2010).

10. S. Matsui et al., "Image enhancement of low-light scenes with near-infrared flash images," in *Asian Conf. on Computer Vision (ACCV)*, pp. 213–223, Springer (2009).
11. Q. Yan et al., "Cross-field joint image restoration via scale map," in *IEEE Int. Conf. on Computer Vision (ICCV)*, pp. 1537–1544 (2013).
12. C. Wu, R. Samadani, and P. Gunawardane, "Same frame rate IR to enhance visible video conference lighting," in *IEEE Int. Conf. on Image Processing (ICIP)*, pp. 1521–1524 (2011).
13. H. Honda, R. Timofte, and L. V. Gool, "Make my day—high-fidelity color denoising with near-infrared," in *IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVRW)*, pp. 82–90 (2015).
14. T. Shibata, M. Tanaka, and M. Okutomi, "Unified image fusion based on application-adaptive importance measure," in *IEEE Int. Conf. on Image Processing (ICIP)*, pp. 1–5 (2015).
15. D. Sugimura et al., "Enhancing color images of extremely low light scenes based on RGB/NIR images acquisition with different exposure times," *IEEE Trans. Image Process.* **24**(11), 3586–3597 (2015).
16. Z. Sadeghipoor, Y. M. Lu, and S. Süstrunk, "Correlation-based joint acquisition and demosaicing of visible and near-infrared images," in *IEEE Int. Conf. on Image Processing (ICIP)*, pp. 3165–3168 (2011).
17. Z. Sadeghipoor, Y. M. Lu, and S. Süstrunk, "A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1646–1650 (2013).
18. Z. Chen, X. Wang, and R. Liang, "RGB-NIR multispectral camera," *Opt. Express* **22**, 4985–4994 (2014).
19. H. Tang et al., "High resolution photography with an RGB-infrared camera," in *IEEE Int. Conf. on Computational Photography (ICCP)*, pp. 1–10 (2015).
20. Z. Sadeghipoor, Y. M. Lu, and S. Süstrunk, "Gradient-based correction of chromatic aberration in the joint acquisition of color and near-infrared images," *Proc. SPIE* **9404**, 94040F (2015).
21. M. Martinello et al., "Dual aperture photography: image and depth from a mobile camera," in *IEEE Int. Conf. on Computational Photography (ICCP)*, pp. 1–10 (2015).
22. H. Teranaka et al., "Single-sensor RGB and NIR image acquisition: toward optimal performance by taking account of CFA pattern, demosaicking, and color correction," in *SPIE Electronic Imaging* (2016).
23. C. Park and M. G. Kang, "Color restoration of RGBN multispectral filter array sensor images based on spectral decomposition," *Sensors* **16**(5), 719 (2016).
24. H. Yamashita, D. Sugimura, and T. Hamamoto, "Enhancing low-light color image using an RGB-NIR single sensor," in *IEEE Conf. on Visual Communications and Image Processing (VCIP)*, pp. 119–121 (2015).
25. D. Menon and G. Calvagno, "Joint demosaicking and denoising with space-varying filters," in *IEEE Int. Conf. on Image Processing (ICIP)*, Vol. 3, 477–480 (2009).
26. D. Menon and G. Calvagno, "Regularization approaches to demosaicking," *IEEE Trans. Image Process.* **18**(10), 2209–2220 (2009).
27. H.-K. Lam, O. C. Au, and C.-W. Wong, "Automatic white balancing using luminance component and standard deviation of RGB components," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 3, pp. 493–496 (2004).
28. J. Kopf et al., "Joint bilateral upsampling," *ACM Trans. Graphics* **26**, 96 (2007).
29. K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1397–1409 (2013).
30. B. Ham, M. Cho, and J. Ponce, "Robust image filtering using joint static and dynamic guidance," in *IEEE Conf. on Computer Vision Pattern Recognition (CVPR)*, pp. 4823–4831 (2015).
31. C. Fredembach and S. Süstrunk, "Colouring the near infrared," in *IS&T/SID Color Imaging Conf.*, pp. 176–182 (2008).
32. A. Foi et al., "Practical Poisson–Gaussian noise modeling and fitting for single-image raw-data," *IEEE Trans. Image Process.* **17**, 1737–1754 (2008).
33. F. Luisier, T. Blu, and U. Michael, "Image denoising in mixed Poisson–Gaussian noise," *IEEE Trans. Image Process.* **20**, 696–708 (2011).
34. A. Jeziorska et al., "An EM approach for Poisson–Gaussian noise modeling," in *EUSIP Signal Processing Conf.*, pp. 2244–2248 (2011).
35. Y. L. Montagner, E. D. Angelini, and J. C. Olivo-Marin, "An unbiased risk estimator for image denoising in the presence of mixed Poisson–Gaussian noise," *IEEE Trans. Image Process.* **23**, 1255–1268 (2014).

Hiroki Yamashita received his BS and MS degrees in electrical engineering from Tokyo University of Science, Tokyo, Japan, in 2015 and 2017, respectively. His research interests include image processing and computational imaging.

Daisuke Sugimura received his BS degree in engineering science from Osaka University, Osaka, Japan, in 2005, and his MS and PhD degrees in information science and technology from the University of Tokyo, Tokyo, Japan, in 2007 and 2010, respectively. Currently, he is a junior associate professor with the Department of Electrical Engineering, Tokyo University of Science. His research interests include computer vision, image processing, and pattern recognition.

Takayuki Hamamoto received his BS and MS degrees in electrical engineering from Tokyo University of Science, Japan, in 1992 and 1994, respectively, and his DEng degree in electrical engineering from the University of Tokyo, Japan, in 1997. Currently, he is a professor with the Department of Electrical Engineering and the dean of the Faculty of Engineering, Tokyo University of Science. His current research interests include image processing and computational image sensors.