

GHOST-FREE IMAGE USING BLUR AND NOISE ESTIMATION

Ji-Hye Kim¹, Rae-Hong Park¹, and SoonKeun Chang²

¹Department of Electronic Engineering, School of Engineering, Sogang University
35 Baekbeom-ro (Sinsu-dong), Mapo-gu, Seoul 121-742, Korea

²Samsung Electronics Co., Ltd., Suwon, Gyeonggi-do 443-742, Korea

ABSTRACT

This paper presents an efficient image enhancement method by fusion of two different exposure images in low-light condition. We use two degraded images with different exposures: one is a long-exposure image that preserves the brightness but contains blur and the other is a short-exposure image that contains a lot of noise but preserves object boundaries. The weight map used for image fusion without artifacts of blur and noise is computed using blur and noise estimation. To get a blur map, edges in a long-exposure image are detected at multiple scales and the amount of blur is estimated at detected edges. Also, we can get a noise map by noise estimation using a denoised short-exposure image. Ghost effect between two successive images is avoided according to the moving object map that is generated by a sigmoid comparison function based on the ratio of two input images. We can get result images by fusion of two degraded images using the weight maps. The proposed method can be extended to high dynamic range imaging without using information of a camera response function or generating a radiance map. Experimental results with various sets of images show the effectiveness of the proposed method in enhancing details and removing ghost artifacts.

KEYWORDS

Image fusion, Blur estimation, Noise estimation, Motion detection, High dynamic range image, Exposure

1. INTRODUCTION

Digital camera users want to take good pictures without constraints of time and space, however, hand-held camera users are still suffering from many kinds of distortions such as blur and noise artifacts. Especially, images captured in low-light condition have dark regions, thus we have a difficult in figuring out what appears in them. To overcome the lack of lightness, a user can take an image with high International Organization for Standardization (ISO) setting, however, the image will contain lots of noise. Another method is to control the exposure time, however, blur degrades the image quality by camera shaking during taking a picture.

If a user takes pictures, then light comes into the camera through the lens. During sensing using image sensors such as charge-coupled device (CCD) or complementary metal-oxide semiconductor (CMOS), the amount of light determines the quality of an image, which is specified according to the shutter speed and the aperture [1].

In this paper, we use multiple images that are taken with different exposures to reduce the distortion in dim-light condition. During taking photos, the amount of light is a very important factor that determines the quality of photos, in which the amount of light is controlled, for example, according to the shutter speed. With long-exposure time, we get a blurry image by

camera shaking, however, brightness of an image is preserved. On the other hand, with short-exposure time, we get a dark and noisy image according to sensor characteristics or ISO setting, however, it is free from blurring. To get a crisp and clear image with adequate brightness and clear boundaries, we propose a fusion method of multi-exposure images that are degraded by blur and noise.

The rest of the paper is organized as follows. In Section 2, related image fusion methods are described. In Section 3, we propose a ghost-free image fusion method that uses blur and noise estimation. Experimental results and comparison of the proposed algorithm with previous work are shown in Section 4. Finally, conclusions are given in Section 5.

2. PREVIOUS WORK

Various image enhancement methods in low-light condition using different exposure images have been proposed. First, a single image is used to produce an acceptable image in low light. For reducing noise in a short-exposure image, total variation [2] is used as a regularization term in denoising, which is an ill-posed problem. Another denoising methods use a multi-resolution bilateral filter [3] and a motion-adaptive spatio-temporal filter [4]. However, these methods are not acceptable, because they do not have enough brightness and color information. Also for removing motion blur in a long-exposure image, many studies have been done [5]. First, they estimate the point spread function (PSF) and perform deconvolution to reconstruct the image [6-8]. Fergus *et al.*'s method [6] uses a maximum a priori (MAP) method and Shan *et al.*'s method [7] uses an ensemble learning method to get a clear output image. Yuan *et al.*'s method [8] uses two input images: long- and short-exposure images. It is good to reduce the ringing artifacts, which is one of the difficult problems in deblurring. However, the deblurring methods that consist of estimation of the PSF and deconvolution method require lots of iterations and thus have a drawback of heavy computation.

Second, fusion methods of multi-exposure images have been proposed to obtain the acceptable output image in dim light and they reduce the computation time with simple processing. These fusion methods are also used in high dynamic range (HDR) because they give better result without any information about the camera response function (CRF) and radiance map [9-13]. Raman and Chaudhuri [9] proposed a weighted sum of intensity differences between input images and bilateral filtered results. This fusion method of multi-exposure images shows good performance in preserving details and weak edges, however, noise in the short-exposure image is emphasized in the output image. Zhang and Cham [10] used gradient direction and magnitude as weights. Also, Martens *et al.* [11] proposed a weighted sum method using three quality measurements (contrast, saturation, and well-exposedness). Especially, Tico and Pulli [12] and Tico *et al.* [13] proposed fusion methods of two different exposure images: long- and short-exposure images. These methods perform fusion of luminance in the wavelet domain and then chrominance fusion is followed based on luminance fusion. These methods contain the weight that considers noise, thus they give better result in view of noise reduction. However, they do not reduce blurred edges in a long-exposure image, thus the fused image still suffers from blurring artifacts.

2.1. Bilateral filtering based compositing method

Raman and Chaudhuri's method [9] combines multi-exposure images with weights from bilateral filtering [14]. The bilateral filter is a non-linear filter which smoothes flat regions while preserving edge regions. It uses two Gaussian weights: one is a range weight considering the

intensity difference between neighboring pixels and the other is a domain weight considering the spatial distance. Weak edges and details can be detected by the difference between the bilateral filtered image and the original image. In their method [9], the difference with the bilateral filtered image is used as a weight to preserve both details and weak edges. Let $I_k(x, y)$ represent k^{th} input image and $I_k^{BS}(x, y)$ denote a bilateral filtered image, then the weight of k^{th} image at (x, y) is describe as

$$W_k(x, y) = \frac{C + |I_k(x, y) - I_k^{BS}(x, y)|}{\sum_{k=1}^K (C + |I_k(x, y) - I_k^{BS}(x, y)|)} \quad (1)$$

where C is a constant and K is the number of input images. The total sum of weights $W_k(x, y)$ is equal to one. The final output image using (1), $I_o(x, y)$, is described as

$$I_o(x, y) = \sum_{k=1}^K W_k(x, y) I_k(x, y). \quad (2)$$

An image with lots of details has a large weight value. The compositing method based bilateral filter [9] preserves edges well. This method is easy to implement, but it is difficult to separate signal and noise when an input image is noisy. Also more consideration on ghost artifacts is required.

2.2. Exposure fusion

Exposure fusion [11] can obtain HDR image using several input images of low dynamic range without generating a radiance map or information about camera setting. Quality measurements from each image are used to compute weights for image fusion. Quality measurements consist of three components: contrast, saturation, and well-exposedness. First, weight for contrast measurement has a high value in edge and texture regions that are detected by a Laplacian filter. Among the input images, there are images that include saturation region. To avoid artifacts from saturation, weight for the saturation is detected by the standard deviation in the RGB color space. At last, weight for well-exposedness is used by the value how close the normalized intensity of the input image is to 0.5. With three quality measurements, the weighted sum is used for fusion. In fusion, input images, which are taken with different exposure times, give some artifacts like seams. To remove them, a pyramid decomposition method [15] is used, in which Laplacian pyramid is used for input images whereas Gaussian pyramid is used for weights. Therefore, exposure fusion suffers from ghost artifacts because this method does not consider motion between input images. Also, it is not appropriate for input images that contain lots of noise, thus result of exposure fusion is degraded by noise that is contained in input images.

2.3. Image fusion

The goal of image fusion [12] is to get a proper result image by fusion of blurred and noisy images that are taken under low-light condition. First, image alignment is performed to remove the misaligned region between two images because of the finite time interval during image acquisition. Next, photometric calibration is performed to compensate for the brightness change between two input images. For photometric calibration, a brightness transfer function, which is also called compagram, is estimated using a joint histogram function between two input images.

Using this function, a short-exposure image is compensated in view of a long-exposure image. Then, two differently exposed images that are aligned and compensated with compagram [16] are fused with different weights in luminance and chrominance channels. In the CIE L*a*b* space, fusion of luminance channel is performed using the difference of long- and short-exposure images in the wavelet transform domain. The difference $D(k)$ is described as

$$D(k) = G_L(k) - G_S(k) \quad (3)$$

where k represents k^{th} coefficient, G_L denotes a long-exposure image, and G_S signifies a short-exposure image in the wavelet domain. The result wavelet coefficient of luminance fusion $\hat{F}(k)$ is described as

$$\hat{F}(k) = G_S(k) + W(k)D(k) \quad (4)$$

where W is a weight. The weight is calculated as

$$W(k) = \sigma^2(k) / E[D(k)^2] \quad (5)$$

where $E[D(k)^2]$ represents expectation of the squared difference $D(k)$ and σ^2 signifies the variance. This weight emphasizes the edge region of the short-exposure image and flat regions of the long-exposure image. On the other hand, weight for chrominance fusion is calculated with luminance value (0~255) from luminance fusion. The weight of the short-exposure image $I'_S(x, y)$, which is the brightness-compensated image, is described as

$$W_S(L) = \left(\frac{L}{255} \right)^6 \quad (6)$$

and the weight of the aligned long-exposure image $\tilde{I}_L(x, y)$ is expressed as

$$W_L(L) = 1 - W_S(L). \quad (7)$$

With an image fusion method [12], a user can get a blur- and noise-free image in low light without PSF estimation. However, blur of a long-exposure image still remains and affects the quality of the fused image. To reduce blur and ghost artifacts, a motion-blur-free exposure fusion method [13] was proposed.

Figure 1 shows performance comparison of existing methods described above. Figure 1(a) shows input images with different exposures, which were used in [11]. Figure 1(b) shows a result image obtained by a bilateral filter based method [9] and Figure 1(c) shows a result image by the exposure fusion method [11].

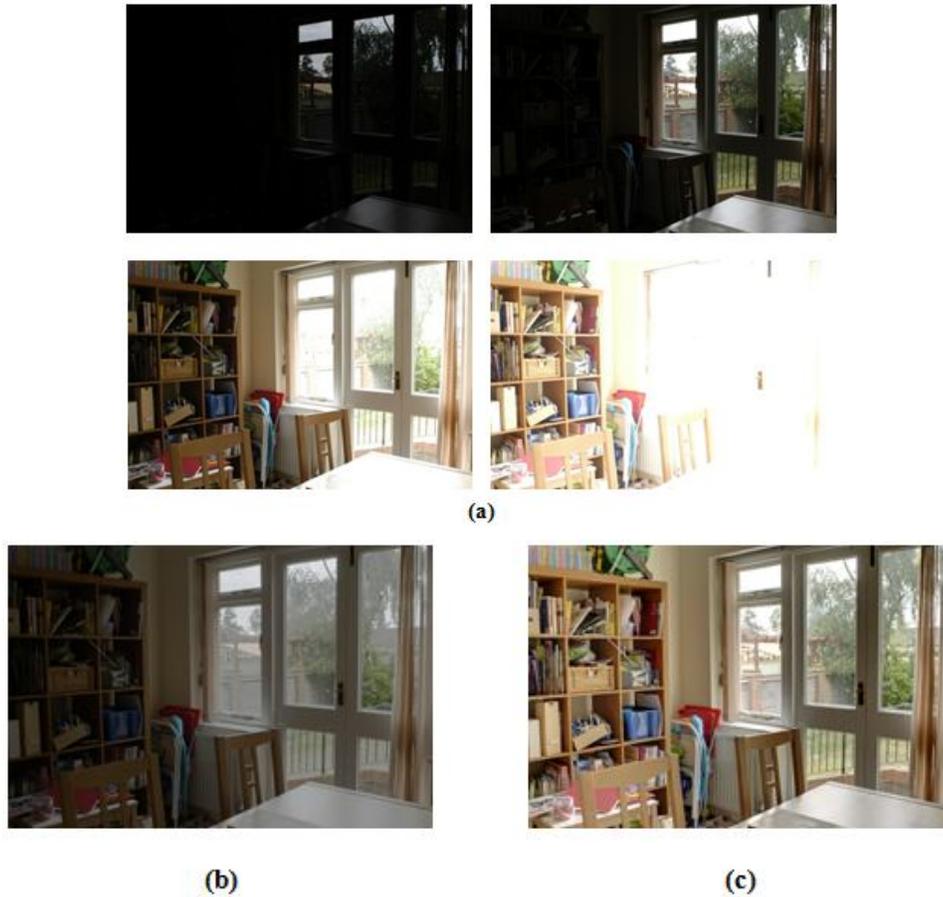


Figure 1. Comparison of existing methods. (a) Multi-exposure input images [11], (b) Bilateral filtering based compositing method [9], (c) Exposure fusion method [11].

3. PROPOSED IMAGE FUSION METHOD

The proposed method uses input images with different exposure images obtained in low- light condition. The characteristics of each exposure image are shown in Table 1. A long-exposure image is obtained with enough brightness and proper color, however distorted by blur artifacts. Whereas, a short-exposure image is very dark and contains lots of noise, with edges preserved. By the pixel-wise fusion method with these two images, blur and noise artifacts in input images degrade the fused image. To reduce degradation in the output image, pixel-wise weights that reflect the characteristics at pixels of input images are used in fusion. As mentioned in Section 2, existing methods [9-11] use multi-exposure images that do not contain lots of blur and noise, so they are not appropriate for low-light condition. If the same weights are used for degraded input images, the fused image suffers from blur and noise. Therefore, with long- and short-exposure images, weights that reflect the characteristics of input images are to be used.

Table 1. Characteristics of images with different exposures.

Exposure time	Artifacts	Brightness	Color	Detail
Long	Blur	Enough	Preserved	Blurred
Short	Noise	Lack	Loss	Preserved

3.1. Pre-processing

In this paper, we use two different exposure images that are taken successively. Small motion or camera shaking during time interval can produce misaligned regions between two images, which are caused by the global motion. Therefore, global registration is required for compensation of the global motion between two input images. In this paper, to estimate a single motion model and to compensate for global motion, the gradient method [17] is used. For fast computation, a fast gradient motion estimation method [18] is used. With the parametric coordinate transformation, two images can be represented as

$$\begin{aligned} x_i^{(2)} &= f(x_i^{(1)}, y_i^{(1)}, \mathbf{p}) \\ y_i^{(2)} &= g(x_i^{(1)}, y_i^{(1)}, \mathbf{p}) \end{aligned} \quad (8)$$

where $(x_i^{(1)}, y_i^{(1)})$ and $(x_i^{(2)}, y_i^{(2)})$ are corresponding pixel positions of $I_L(x, y)$ and $I_S(x, y)$ images, respectively, with $I_L(x, y)$ representing a long-exposure image and $I_S(x, y)$ denoting a short-exposure image. f and g are defining functions. The global motion method [12] estimates the global motion parameter vector \mathbf{p} that minimizes the difference between pixel values of two images. The global motion parameter vector is calculated as

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} \sum_{(x_i^{(1)}, y_i^{(1)}) \in S} (I_L(x_i^{(1)}, y_i^{(1)}) - I_S(x_i^{(2)}, y_i^{(2)}))^2 \quad (9)$$

where i represents position index of pixel. The relation between $I_L(x, y)$ and $I_S(x, y)$ can be described as

$$\begin{aligned} I_L(x_i^{(1)}, y_i^{(1)}) &= I_S(x_i^{(2)}(0, x_i^{(1)}, y_i^{(1)}), y_i^{(2)}(0, x_i^{(1)}, y_i^{(1)})) \\ &+ \sum_{p_j \in \mathbf{p}} \left. \frac{\partial I_S(x_i^{(2)}(\mathbf{p}, x_i^{(1)}, y_i^{(1)}), y_i^{(2)}(\mathbf{p}, x_i^{(1)}, y_i^{(1)}))}{\partial p_j} \right|_{\mathbf{p}=\mathbf{0}} p_j \end{aligned} \quad (10)$$

using the value of pixel-wise first-order Taylor expansion at $\mathbf{p} = \mathbf{0}$. (10) is a linearization of Gauss-Newton nonlinear minimization [19]. Let H be the partial differential equation of \mathbf{p} , with its elements $H_{i,j}$ expressed, by the derivative chain rule, as

$$\begin{aligned} H_{i,j} &= \frac{\partial I_S(x_i^{(2)}, y_i^{(2)})}{\partial p_i} \\ &= \frac{\partial I_S(x_i^{(2)}, y_i^{(2)})}{\partial x_i^{(2)}} \frac{\partial x_i^{(2)}(x_i^{(1)}, y_i^{(1)}, \mathbf{p})}{\partial p_i} + \frac{\partial I_S(x_i^{(2)}, y_i^{(2)})}{\partial y_i^{(2)}} \frac{\partial y_i^{(2)}(x_i^{(1)}, y_i^{(1)}, \mathbf{p})}{\partial p_i}. \end{aligned} \quad (11)$$

Then, motion parameter vector \mathbf{p} can be described using (10) and (11) as

$$\mathbf{p} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{I}_t \quad (12)$$

where \mathbf{I}_t is described as

$$\mathbf{I}_t = ((I_t)_1, \dots, (I_t)_n)^T. \quad (13)$$

\mathbf{I}_t is used as error vector and $(I_t)_i$ is represented as

$$(I_t)_i = I_L(x_i^{(1)}, y_i^{(1)}) - I_S(x_i^{(2)}(\mathbf{p}, x_i^{(1)}, y_i^{(1)}), y_i^{(2)}(\mathbf{p}, x_i^{(1)}, y_i^{(1)})). \quad (14)$$

As a result, aligned $\tilde{I}_L(x, y)$ can be calculated using \mathbf{p} . Moreover, the fast gradient method [18] uses multi-scale structure for better performance and the pixel subset selection [20] for speed. In the pixel subset selection method, motion is estimated by dividing the region into several sub-regions and by using the gradient value which is upper 10% in each region.

After alignment, the short-exposure image should be corrected to compensate for low light according to the long-exposure image. As a histogram matching method for color correction of the short-exposure image, a cumulative histogram [21] is used in this paper. The histogram of the long-exposure image is calculated as

$$h_L(v) = \frac{1}{W \cdot H} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} \delta(v, \tilde{I}_L(x, y)) \quad (15)$$

where W is width and H is depth of an image. The function δ is described as

$$\delta(a, b) = \begin{cases} 1, & \text{if } a = b \\ 0, & \text{otherwise} \end{cases} \quad (16)$$

where v is pixel intensity between 0 and 255. \tilde{I}_L is an aligned long-exposure image. The mapping function M between the short-exposure image I_S and the pixel value to be compensated is used in calibration for color correction. The cumulative histogram of C_L can be calculated as

$$C_L(v) = \sum_{i=0}^v h_L(i). \quad (17)$$

Also, the cumulative histogram of I_S , $C_S(v)$ is calculated in the same method. With the cumulative histogram, the mapping function is described as

$$M(v) = u \quad \text{with } C_L(u) < C_S(v) \leq C_L(u+1). \quad (18)$$

If the mapping function in (18) is applied to the image, luminance value of I_S smaller than threshold is clipped and mapped into the first active bin. The result image is unnatural, and the first and last bins in lower boundary and upper boundary, respectively, should be modified. The highest value of the reference image is applied to the clipped interval value. Lower interval is $(0, \dots, M(0))$ and upper interval is $(M(255), \dots, 255)$. Using each interval value, the center of mass is used by the mapping function. The center of mass value for lower interval s_l and that for upper interval s_u are described as

$$s_l = \frac{\sum_{i=0}^{M(0)} i \cdot h_L(i)}{\sum_{i=0}^{M(0)} h_L(i)} \quad (19)$$

$$s_u = \frac{\sum_{i=M(255)}^{255} i \cdot h_L(i)}{\sum_{i=M(255)}^{255} h_L(i)}. \quad (20)$$

With (19) and (20), the mapping function can be described as

$$M(0) = s_l \text{ and } M(255) = s_u. \quad (21)$$

Using the mapping function M , the color-corrected short-exposure image, $I'_s(x, y)$ can be calculated as

$$I'_s(x, y) = M(I_s(x, y)). \quad (22)$$

3.2. Proposed image fusion method

In this paper, the proposed method uses a weighted sum based fusion method and weights are obtained by estimating blurriness and noisiness of input images. Aligned long-exposure image \tilde{I}_L and color-corrected short-exposure image I'_s are used as input images. \tilde{I}_L contains blur artifacts around edges whereas I'_s contains lots of noise which is amplified after color correction. These artifacts from two input images affect the final fused image. Also, if there is motion during the time taking two input images, the fused image contains ghost artifacts. Therefore, proper weights are required to effectively reduce artifacts in input images and motion detection between two input images is needed to reduce ghost artifacts. To explain the proposed fusion method, first, image fusion is reviewed and then blur and noise estimation methods are described. At last, a motion detection method used in the proposed method is presented.

The final output of the proposed fusion method is obtained using two weights: blur weights W_L and noise weight W_S . W_L is applied to the long-exposure image whereas W_S to the short-exposure image. However, if motion occurs, the fused image is degraded with ghost artifacts. Therefore, the motion map is used to remove the ghost artifact. The output image $I_O(x, y)$ is calculated as

$$I_O = \begin{cases} I'_{S,D}, & \text{if } M_M = 1 \\ I_F, & \text{otherwise} \end{cases} \quad (23)$$

where M_M is the binary motion map obtained from moving object detection. When the value of the motion map is one, the denoised image is the output image without fusing processing, otherwise, the fused image I_F is used. When motion occurs, fusion processing of images degrades the output image, thus fusing is performed at pixels with no motion. The fused image, I_O is computed as

$$I_O(x, y) = \begin{cases} I'_{S,D}(x, y), & M'(x, y) = 1 \\ \tilde{I}_S(x, y), & S(x, y) = 1 \\ W_S(x, y)I'_s(x, y) + W_L(x, y)\tilde{I}_L(x, y), & N(x, y) < T_N \text{ and } B(x, y) < T_B \\ W_S(x, y)I'_{S,D}(x, y) + W_L(x, y)\tilde{I}_L(x, y), & \text{otherwise.} \end{cases} \quad (24)$$

where the subscript L represents a long-exposure image and the subscript S denotes a short-exposure image. $I'_s(x, y)$ is the color-corrected short-exposure image and $I'_{S,D}(x, y)$ is the denoised short-exposure image. Also, $\tilde{I}_L(x, y)$ is the aligned long-exposure image. B is blur measurement and N is noise measurement. T_B and T_N are threshold values for blur and noise estimation, respectively. In experiments, T_B and T_N are set to 0.1 and 0.2, respectively. $M_M(x, y)$ is the binary motion map to reduce the ghost artifacts between two images. If motion occurs, the value of M_M is 1, otherwise, 0. With the motion map, if motion occurs, result of image fusion uses only $I'_{S,D}(x, y)$, which is the denoised image with the existing denoising method [4]. When motion does not occur, the weighted sum of long- and short-exposure images is used. Note that they smooth not only noise but also high-frequency component, thus details are not preserved. To

preserve details, I'_S rather than $I'_{S,D}$ is used in fusion process, when blur and noise estimates are both smaller than threshold values. If either blur or noise estimate is larger than threshold value, I'_S produces a fused image with lots of noise remained. $W_S(x, y)$ has a smaller value in noisy region of I'_S to suppress the noise in the result image. Whereas $W_L(x, y)$ has a smaller value in the blur region of $\tilde{I}_L(x, y)$.

Each weight is determined depending on the amount of blur and noise in the images and has the value between 0 and 1. The normalized blur measurement is computed as

$$B'(x, y) = \frac{B(x, y) - B_{\min}}{B_{\max} - B_{\min}} \quad (25)$$

where B_{\min} represents the minimum value of $B(x, y)$ and B_{\max} denotes the maximum value. Also the normalized noise measurement is expressed as

$$N'(x, y) = \frac{N(x, y) - N_{\min}}{N_{\max} - N_{\min}} \quad (26)$$

where N_{\min} represents the minimum value of $N(x, y)$ and N_{\max} denotes the maximum value. With each normalized measurement of blur and noise, we can get the weight for image fusion. First of all, the weight of the long-exposure image $W_L(x, y)$ is calculated as

$$W_L(x, y) = \frac{\alpha N(x, y)}{\alpha N(x, y) + B(x, y)} \quad (27)$$

where $N(x, y)$ denotes the noise measurement whereas $B(x, y)$ represents the blur measurement. Also, α is used to scale the parameter value. Then, the weight for the short-exposure image is calculated as

$$W_S(x, y) = \frac{B(x, y)}{\alpha N(x, y) + B(x, y)}. \quad (28)$$

Figure 2 shows the block diagram of the proposed image fusion method. There are four steps for fusing two degraded images. The first step is global registration to compensate for global motion that affects all the pixels in the whole image with a single motion model applied to a long-exposure image. Then, color correction is performed to adjust brightness of the short-exposure image. Compensated input images by two pre-processing steps are fused using appropriate weights. These weights are based on two measurements: blur and noise estimation. Blur and noise measurements are computed by estimating the amount of blurriness and noisiness from the long-exposure image and short-exposure image, respectively. Also the motion map is used to remove ghost artifacts. Finally, computing the weighted sum of the long-exposure image, short-exposure image, and denoised short-exposure image, we can get the output image with acceptable brightness and details, with blur and noise artifacts significantly reduced.

The long-exposure image contains blur and produces blur artifacts in the fused image, especially around edges. To get the clear fused image, we estimate the amount of blur and apply this quantity to compute weights used for fusing a set of input images. A small blur weight is applied to $\tilde{I}_L(x, y)$ when pixel is degraded by blur. For the same reason, a large weight is applied to I'_S , therefore, pixel with blur is reduced in the result image.

Generally, blur artifacts occur around the boundary region of an object rather than in flat region [22]. Therefore, existing methods detect edges first, and then estimate the amount of blurriness at

detected edges [22,23]. In the long-exposure image, it is difficult to detect accurately edges by existing edge detection methods, in which edges are detected at wrong locations or represented as discontinuous line. Elder and Zucker [22] proposed an edge detection method with blurred images. Then, edges are detected by the reliability of the filter response that is obtained using threshold at each scale. For selecting the threshold value, prior knowledge such as noise characteristics of sensor is needed. Threshold values, c_1 and c_2 in the first- and second-derivative Gaussian filters, respectively, are calculated as

$$c_1(\sigma_1) = \frac{s_n \sqrt{-2 \ln \alpha_p}}{2\sqrt{2\pi}\alpha_1^2} \tag{29}$$

$$c_2(\sigma_2) = \frac{s_n \sqrt{2} \operatorname{erf}^{-1}(\alpha_p)}{4\sqrt{\pi/3}\alpha_2^3} \tag{30}$$

where s_n represents the noise level of sensor and α_p denotes the false positive tolerance.

The minimum reliable scale is decided when the filter response at each pixel is larger than the threshold value. Elder and Zucker’s method [22] detects edges where gradient has nonzero value at selected scale and the value of the second-derivative satisfies the zero crossing condition.

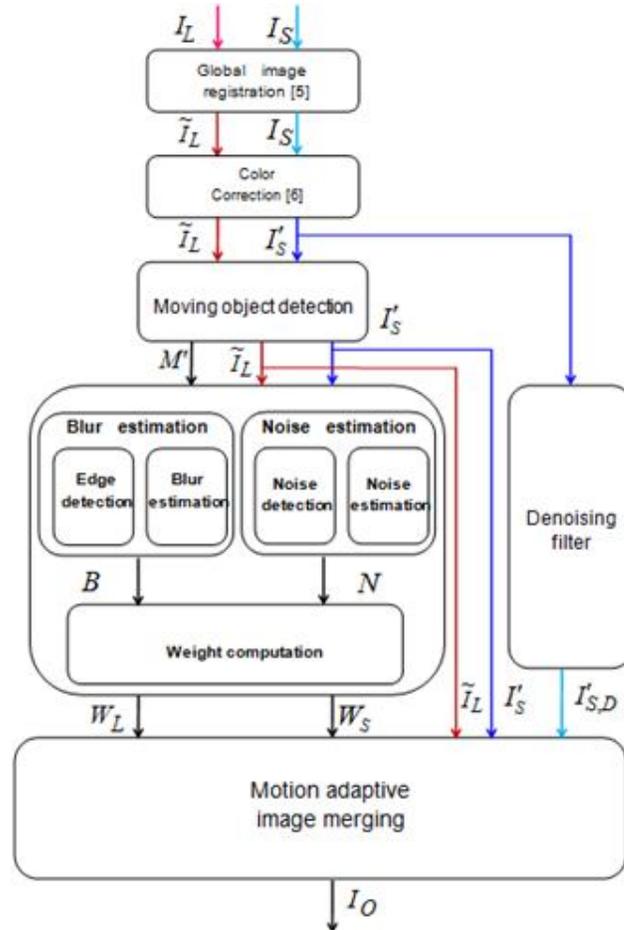


Figure 2. Block diagram of the proposed image fusion method.

In detected edges, the amount of blurriness is estimated using the difference of extreme values of plus and minus of the second-derivative. It can be described as

$$BM = \sqrt{(d/2)^2 - \sigma_2^2} \quad (31)$$

where d is calculated to minimize the least square fitting error at edges. The blur estimation methods mentioned previously need post-processing in computing the weights in our fusion method because estimated measure of blurriness is discontinuous. For post-processing, we use Bae and Durand's method [24], in which edges are detected and the amount of blur is estimated. Then, they used two steps. The first step is refinement, in which a cross bilateral filter [25] is used to remove outliers from estimated blur measurement. Then, propagation [26] is performed for extending blur measurement from edges to the flat region. This blur measurement is used to compute the blur weight in image fusion.

Under low-light condition, an image is degraded mostly by noise. Especially, the short-exposure image after color correction is degraded by magnified noise, which affects the quality of the fused image. In low-light condition, there are many kinds of noises such as photon noise and dark current noise. These noises can be modeled by Gaussian or Poisson distribution. In this paper, we classify noise into two types: Poisson counting noise (PCN) and false color noise (FCN) [5]. We first detect pixels that are distorted by noise and then estimate the amount of noise at detected locations. For detection and estimation of noise, we use a noise adaptive spatio temporal filter proposed by Lee *et al.* [5]. Also, two types of noises are decided as PCN and FCN according to motion that is detected by the sum of absolute differences (SAD). This method is for noise removal for video in low-light condition. Our proposed fusion method uses two images instead of video that consists of a sequence of temporal frames. Even though input type of the proposed method is different from that of Lee *et al.*'s method, noise characteristics in low-light condition are the same, so we can apply the noise detection algorithm to the compensated short-exposure image.

According to [5], two types of noises are detected. First, using local statistics PCN is detected in the short-exposure image that contains lots of low-light noise. Then, PCN noise measurement in I'_s , $N_p(x, y)$ is calculated as

$$N_p(x, y) = |I'_s(x, y) - S_N(x, y)| \quad (32)$$

where $S_N(x, y)$ represents local statistics. It is described as

$$S_N(x, y) = \text{median}\{|I'_s(x, y) - I'_s(r, s)|, (r, s) \in \Omega\} \quad (33)$$

where (x, y) is the coordinate of the center pixel and (r, s) represent the coordinates of neighboring pixels in widow Ω .

Meanwhile, FCN noise is independent of neighboring pixels and it can be detected using statistics between two sequential images. The local statistics of the short-exposure image is calculated as

$$S_s(x, y) = \left| (I'_s(x, y) - \bar{I}'_s(x, y))^2 - \sigma_s^2(x, y) \right| \quad (34)$$

where $\bar{I}'_s(x, y)$ represents the mean value and $\sigma_s^2(x, y)$ denotes the variance of the long-exposure image. With each local statistics, noise measurement of FCN is calculated as

$$N_F(x, y) = \begin{cases} |I'_S(x, y) - S_S(x, y)|, & S_L(x, y) > S_S(x, y) \\ |\tilde{I}_L(x, y) - S_L(x, y)|, & \text{otherwise.} \end{cases} \quad (35)$$

With $N_P(x, y)$ and $N_F(x, y)$, measurement in low light, N is computed as

$$N(x, y) = \begin{cases} N_P(x, y), & M(x, y) = 0 \\ N_F(x, y), & \text{otherwise} \end{cases} \quad (36)$$

where $M(x, y)$ is the motion map which is calculated in terms of the SAD. Noise measurement N is used as weight in image fusion.

The proposed method uses input images taken consecutively. Motion during time interval between two successive images reduces the quality of the fused image. Global motion is reduced by registration processing whereas local motion still produces artifacts in the fused image. There are many approaches to reducing ghost artifacts [27-28]. In this paper, to reduce ghost artifacts, the binary motion map is generated by detecting motion between two input images and by applying it to the fusion process.

To detect moving objects, the proposed method uses Markowski's method [27], which is a motion detection method for HDR. This method consists of three steps. First, the initial motion map is described as

$$M_M(x, y) = F(\tilde{I}_L(x, y), I'_S(x, y)) \quad (37)$$

where F is a de-ghosting comparison function. The function F is computed as

$$F(\tilde{I}_L(x, y), I'_S(x, y)) = \sum_{c=r,g,b} \frac{1}{1 + \exp\left(-\left(\max\left(\frac{\tilde{I}_L(x, y)_c}{I'_S(x, y)_c}, \frac{I'_S(x, y)_c}{\tilde{I}_L(x, y)_c}\right) - 1.35\right) \cdot 20\right)} \quad (38)$$

where c represents three color channels in the RGB color space. The comparison function is a sigmoid function that compares two images. Then, the normalized motion map is computed as

$$M'(x, y) = \left(\frac{M_M(x, y) - M_{\min}}{M_{\max} - M_{\min}}\right)^n \quad (39)$$

where M_{\min} (M_{\max}) represents minimum (maximum) value of the initial motion map. Artifacts such as seams near the detected motion region by the normalized $M'(x, y)$ can occur after images with different exposures are fused. For smoothing and reducing artifacts, three morphology operations are sequentially used: dilation, erosion, and conversion [27], after which a more accurate motion map is obtained.

Figure 3(a) shows blur measurement, Figure 3(b) shows noise measurement, and Figure 3(c) shows motion map. With blur and noise measurements, the weight of the short-exposure image, $W_S(x, y)$ is illustrated in Figure 3(d) and the weight of the long-exposure image, $W_L(x, y)$ is shown in Figure 3(e).

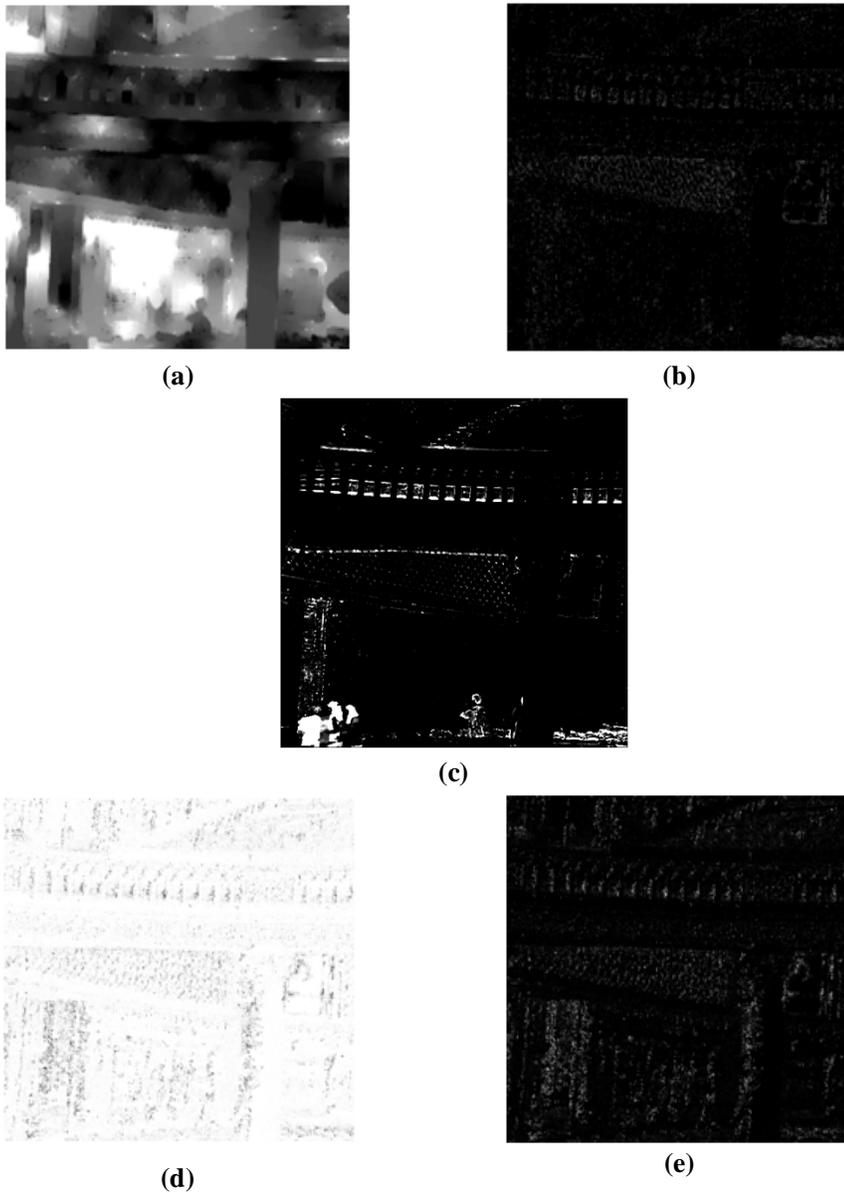


Figure 3. Results of each step of the proposed image fusion method. (a) Blur measurement ($B'(x,y)$), (b) Noise measurement ($N'(x,y)$), (c) Motion map ($M'(x,y)$), (d) Weight of the short-exposure image ($W_s(x,y)$), (e) Weight of the long-exposure image ($W_L(x,y)$).

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

Figure 4 shows two test image sets used in our experiments, which are used in the existing image fusion method [12] and their size is 1024×1024 . Figure 4(a) is a short-exposure image in low-light condition, which is degraded with lots of noise in flat region, however, preserves details. In contrast, Figure 4(b) is a long-exposure image which is degraded with blur artifacts in edges and saturated in window of the right side of an image. Also, in bottom left of Figure 4(b), there are

ghost artifacts because of motion of people. However, it has proper color with enough light. Figure 4(c) is a short-exposure image with noise artifacts in flat region whereas Figure 4(d) is a long-exposure image with blurred edges. Figure 4(c) has sharp edges and Figure 4(d) has proper color with enough light.

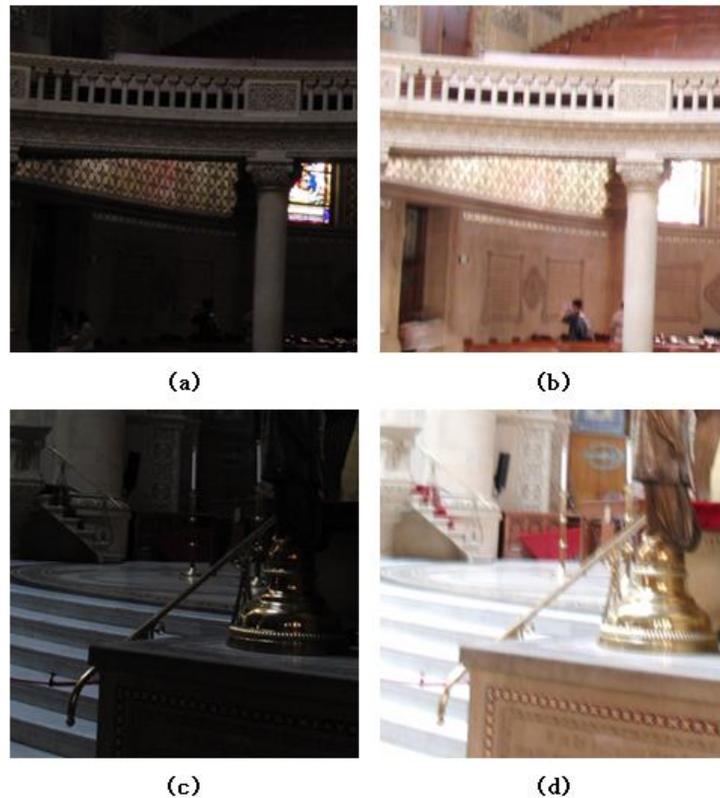


Figure 4. Test image set with different exposures. (a), (c) Short-exposure images, (b), (d) Long-exposure images.

In experiments, two methods are used to compare the performance of the proposed method. The exposure fusion method [11] explained in Section 2.1 and the image fusion method [12] described in Section 2.3 are used. Figures 5(a)-5(d) are result images of image enhancement methods in low-light condition. Figure 5(a) shows a color-corrected result of the short-exposure image based on color of the long-exposure image. The color-corrected result image in low light is noisy because noise is also boosted along with an image signal. Figure 5(b) is a result of the exposure fusion method [11] that uses multi-exposure images, in which blur and noise artifacts of input images affect the result image. Figure 5(c) is a result of the image fusion method [12], which uses two images: long- and short-exposure images. The image fusion method [12] computes weights

according to the amount of noise to give small weight in noisy region. However, blur artifacts from the long-exposure image still remain. Result of the proposed method is shown in Figure 5(d) and blur artifacts near edges and noise in flat region are reduced. Also, the proposed method reduces ghost artifacts. Figure 6 shows more details on result images.

Figure 6 shows enlarged subimages enclosed by the red box in Figure 5. The subimage enclosed by the red rectangle shows preserved details in the image. The subimage enclosed by the blue rectangle in Figure 6 shows enhanced performance of removing ghost artifacts caused by a moving object or people in the image. Figure 6(a) shows result images of color correction of the short-exposure image. The subimage enclosed by the red rectangle in Figure 6(a) shows a sharp image that is degraded with lots of noise because noise boosting of the short-exposure image

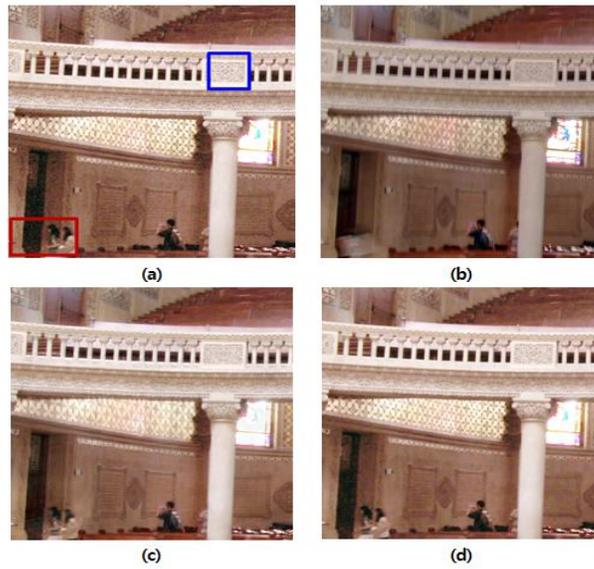


Figure 5. Result images with Figures 4(a) and 4(b). (a) Color-corrected short-exposure image, (b) Result image of the exposure fusion method [11], (c) Result image of the image fusion method [12], (d) Result image of the proposed method.



Figure 6. Enlarged images of Figures 5(a)-(d). (a) Color-corrected short-exposure image, (b) Result image of the exposure fusion method [11], (c) Result image of the image fusion method [12], (d) Result image of the proposed method.

appears with a gain in terms of color correction. Figure 6(b) shows the result of the exposure fusion method [11]. The subimage enclosed by the blue rectangle shows ghost artifacts because this method uses the weighted sum of multi-exposure images without considering ghost artifacts. The subimage enclosed by the red rectangle shows degraded details. It is because blur and noise artifacts in input images affect the result image directly. Figure 6(c) shows the result of the image fusion method [12] and ghost artifacts are almost removed in the subimage enclosed by the blue rectangle, but still blur artifacts in edges and noise in flat and low-light regions are shown in the subimage enclosed by the blue rectangle. Also, the subimage enclosed by the red rectangle shows blurred edges. Figure 6(d) shows the result of the proposed method. Ghost artifacts are removed in the subimage enclosed by the blue rectangle and details of people are preserved. Even though noise is still shown in the subimage enclosed by the blue rectangle, the proposed method shows better performance in view of details in edges and noise in flat region.

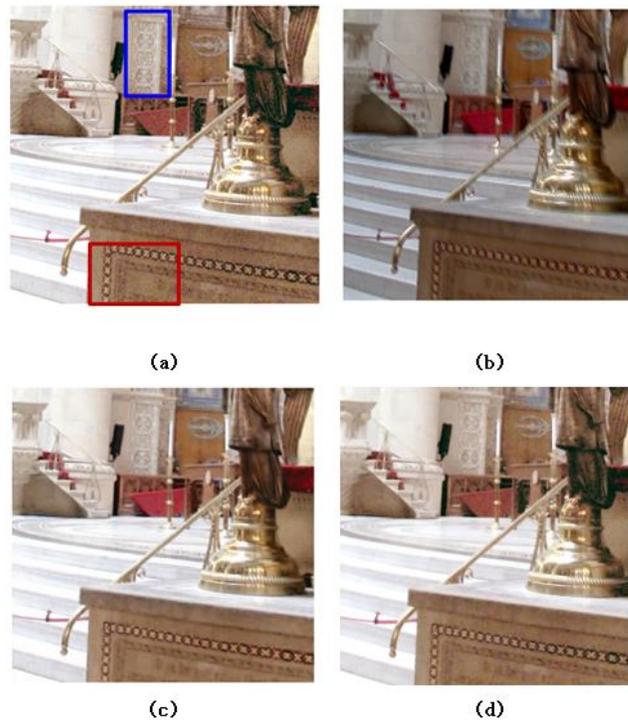


Figure 7. Results images with Figures 4(c) and 4(d). (a) Color-corrected short-exposure image, (b) Result image of the exposure fusion method [11], (c) Result image of the image fusion method [12], (d) Result image of the proposed method.

Figure 7 shows results of three enhancement methods in low-light condition and input images are from Figures 4(c) and 4(d). Figure 7(a) shows a color-corrected image which uses only a short-exposure image with details preserved, however, the image is distorted by the noise of low-light condition. Figure 7(b) shows better result of the exposure fusion method [11] than other method, with blur artifacts in low-light condition. Figure 7(c) shows the result of the image fusion method [12]. The image fusion method gives denoised result, however, blur artifacts are still remain. Figure 7(d) shows the result of the proposed method with the reduced noise and details preserved.

Figure 8 shows enlarged images of Figure 7. Figure 8(a) shows preserved detail, but distorted by lots of noise. Especially, flat region in the subimage enclosed by the red rectangle is distorted

with lots of noise. Figure 8(a) shows the result of the exposure fusion method [11]. In Figure 8(a), details are preserved with an image distorted in the subimages enclosed by the red and blue rectangles and low-light noise. In Figure 8(b), the subimages enclosed by the red and blue rectangles show blur artifacts in the exposure fusion method. In Figure 8(c), the subimage enclosed by the red rectangle shows blur artifacts and the subimage enclosed by the blue rectangle shows denoised image in flat region. The result image by the proposed method is shown in Figure 8(d). In Figure 8(d), the subimage enclosed by the red rectangle shows preserved edges and the subimage enclosed by the blue rectangle also shows denoised clear result and preserved edges. Flat region in the subimage enclosed by the blue rectangle shows especially denoised and preserved edges well.

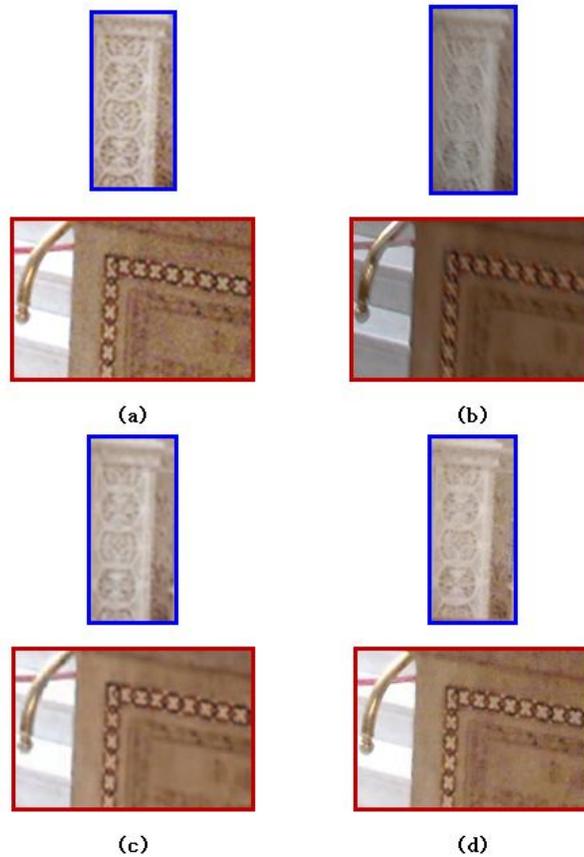


Figure 8. Enlarged image of Figures 7(a)-(d). (a) Color-corrected short-exposure image, (b) Result image of the exposure fusion method [11], (c) Result image of the image fusion method [12], (d) Result image of the proposed method.

5. CONCLUSIONS

In this paper, we propose an image fusion method using two differently exposed images in low-light condition. For fusion based on the weighted sum approach, weights are computed by estimating the amount of blurriness and noisiness. Blur is measured by detecting edges and estimating the amount of blurriness at detected edges in the compensated long-exposure image. Then, outliers of blur measurement are removed by a cross bilateral filter. Meanwhile, noise is estimated in the color-corrected short-exposure image according to the local statistics using the adaptive noise method. Next, the motion map is applied to fusion for removing ghost artifacts.

With weights and motion map, the final fused image is computed by the weighted sum of short- and long-exposure images and denoised short-exposure image. A multi-resolution bilateral filter [3] is used for denoising a short-exposure image. The denoised image is clear in flat region, however, with details lost. To preserve details in the fused image, a short-exposure image is used in region where both amounts of blur and noise are small, i.e., where the blur and noise estimates are smaller than the threshold values.

We simulate the exposure fusion and image fusion method for experimental comparison of the proposed method. Experimental results show that the proposed image fusion method using two different exposure images shows better performance for removing blur and noise artifacts than the existing methods. For future work, the proposed method is to be extended to deal with multiple exposure images as in HDR imaging.

ACKNOWLEDGEMENTS

This work was supported in part by Samsung Electronics, Co., Ltd.

REFERENCES

- [1] B. Long, *Complete Digital Photography*, Third ed. Hingham, MA, USA: Charles River Media, 2004.
- [2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D*, vol. 60, no. 1, pp. 259–268, Nov. 1992.
- [3] Z. Ming and B. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2324–2333, Dec. 2008.
- [4] H. Malm and E. Warrant, "Motion dependent spatiotemporal smoothing for noise reduction in very dim light image sequence," in *Proc. Int. Conf. Pattern Recognition*, pp. 135–145, Hong Kong, Aug. 2006.
- [5] S.-W. Lee, V. Maik, J. Jang, J. Shin, and J. Paik, "Noise-adaptive spatio-temporal filter for real-time noise removal in low light level images," *IEEE Trans. Consumer Electronics*, vol. 51, no. 2, pp. 648–653, May 2005.
- [6] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in *Proc. ACM SIGGRAPH*, pp. 787–794, Boston, MA, USA, July/Aug. 2006.
- [7] Q. Shan, J. Jia, and A. Agarwala, "High quality motion deblurring from a single image," *ACM Trans. Graphics*, vol. 27, no. 3, pp. 47–60, Aug. 2008.
- [8] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Image deblurring with blurred/noisy image pairs," in *Proc. ACM SIGGRAPH*, vol. 26, no. 3, pp. 1–10, San Diego, CA, USA, July 2007.
- [9] S. Raman and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proc. Eurographics Conf.*, Munich, Germany, Mar. 2009.
- [10] W. Zhang and W.-K. Cham, "Gradient-directed composition of multi-exposure images," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 530–536, San Francisco, CA, USA, Aug. 2010.
- [11] T. Martens, J. Kautz, and F. V. Reeth, "Exposure fusion," in *Proc. 15th Pacific Conf. Computer Graphics and Applications*, pp. 382–390, Maui, HI, USA, Oct./Nov. 2007.
- [12] M. Tico and K. Pulli, "Image enhancement method via blur and noisy image fusion," in *Proc. IEEE Int. Conf. Image Processing*, Cairo, Egypt, pp. 1521–1524, Nov. 2009.
- [13] M. Tico, N. Gelfand, and K. Pulli, "Motion-blur-free exposure fusion," in *Proc. 2010 IEEE Int. Conf. Image Processing*, pp. 3321–3324, Hong Kong, Sept. 2010.
- [14] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. IEEE Int. Conf. Computer Vision*, pp. 839–846, Bombay, India, Jan. 1998.
- [15] P. Burt and T. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Communications*, vol. 31, no. 4, pp. 532–540, Apr. 1983.
- [16] S. Mann, "Comparametric equations with practical applications in quantigraphic image processing," *IEEE Trans. Image Processing*, vol. 9, no. 8, pp. 1389–1406, Aug. 2000.

- [17] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *Int. J. Computer Vision*, vol. 12, no. 1, pp. 43–77, Aug. 1994.
- [18] Y. Keller and A. Averbuch, "Fast gradient methods based on global motion estimation for video compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 4, pp. 300–309, Apr. 2003.
- [19] P. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. New York, USA: Academic Press, 1981.
- [20] F. Dellaert and R. Collins, "Fast image-based tracking by selective pixel integration," in *Proc. ICCV 99 Workshop on Frame-Rate Vision*, pp. 1–22, Corfu, Greece, Sept. 1999.
- [21] U. Fecker, M. Barkowsky, and A. Kaup, "Improving the prediction efficiency for multi-view video coding using histogram matching," in *Proc. Picture Coding Symposium*, pp. 2–16, Beijing, China, Apr. 2006.
- [22] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and blur estimation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 20, no. 7, pp. 699–716, July 1998.
- [23] H. Hu and G. Haan, "Low cost robust blur estimator," in *Proc. 2006 IEEE Int. Conf. Image Processing*, pp. 617–620, Atlanta, GA, USA, Jan. 2007.
- [24] S. Bae and F. Durand, "Defocus magnification," in *Proc. Eurographics*, vol. 26, no. 3, pp. 571–579, Prague, Czech Republic, Sept. 2007.
- [25] E. Eisemann and F. Durand, "Flash photography enhancement via intrinsic relighting," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 673–678, Aug. 2004.
- [26] A. Levin, D. Lischinski, and Y. Welss, "Colorization using optimization," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 689–694, Aug. 2004.
- [27] M. Markowski, "Ghost removal in HDRI acquisition," in *Proc. Central European Seminar Computer Graphics*, Vienna, Austria, Apr. 2009.
- [28] T.-H. Min, R.-H. Park, and S. Chang, "Histogram based ghost removal in high dynamic range images," in *Proc. IEEE Int. Conf. Multimedia and Expo ICME 2009*, pp. 530–533, New York, USA, June/July 2009.

Authors

Ji-Hye Kim received the B.S. degree in electronic engineering from Myong-Ji University in 2005 and M.S. degree in electronic engineering from Sogang University in 2011. Her current research interests are image processing and image enhancement. Currently she works at Samsung Electronics Co., Ltd., Korea.

Rae-Hong Park received the B.S. and M.S. degrees in electronics engineering from Seoul National University, Seoul, Korea, in 1976 and 1979, respectively, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1981 and 1984, respectively. In 1984, he joined the faculty of the Department of Electronic Engineering, School of Engineering, Sogang University, Seoul, Korea, where he is currently a Professor. In 1990, he spent his sabbatical year as a Visiting Associate Professor with the Computer Vision Laboratory, Center for Automation Research, University of Maryland at College Park. In 2001 and 2004, he spent sabbatical semesters at Digital Media Research and Development Center (DTV image/video enhancement), Samsung Electronics Co., Ltd., Suwon, Korea. In 2012, he spent a sabbatical year in Digital Imaging Business (R&D Team) and Visual Display Business (R&D Office), Samsung Electronics Co., Ltd., Suwon, Korea. His current research interests are video communication, computer vision, and pattern recognition. He served as Editor for the Korea Institute of Telematics and Electronics (KITE) *Journal of Electronics Engineering* from 1995 to 1996. Dr. Park was the recipient of a 1990 Post-Doctoral Fellowship presented by the Korea Science and Engineering Foundation (KOSEF), the 1987 Academic Award presented by the KITE, the 2000 Haedong Paper Award presented by the Institute of Electronics Engineers of Korea (IEEK), the 1997 First Sogang Academic Award, and the 1999 Professor Achievement Excellence Award presented by Sogang University. He is a co-recipient of the Best Student Paper Awards of the IEEE Int. Symp. Multimedia (ISM 2006) and the IEEE Int. Symp. Consumer Electronics (ISCE 2011).

SoonKeun Chang received his B.S. degree in astronomy and space science from KyungHee University, Korea, in 2000 and M.S. degree in control engineering from Kanazawa University in 2003. He received a Ph.D. degree in control engineering from Tokyo Institute of Technology (TITech) in 2007. Now he works at Samsung Electronics Co., Ltd., Korea. His main research interests include computer vision and image processing.