MULTIPLE EXPOSURE INTEGRATION FOR RESTORING ALL IN-FOCUS IMAGES

Ryo Matsuoka, Haruki Ishibashi, Tatsuya Baba, and Masahiro Okuda

Univ. of Kitakyushu Faculty of Environmental Engineering Kitakyushu, Japan

ABSTRACT

When taking a photograph of a high dynamic range scene, saturation of pixel values (that is under/over-exposure) often occurs due to the narrower dynamic ranges of commercial cameras. It is possible to solve the problem of the under/over-exposure by generating High Dynamic Range (HDR) Images. The HDR image is generated by integration of multiple exposure images that are taken with different exposure settings. When taking multiple exposure images for the scene that contains dark areas, one often copes it by adjusting one of the three settings: (1) camera's sensitivity, (2) exposure time, (3) lens' aperture, in order to compensate shortage of the quantity of light. Each of the three methods have a drawback. The high camera sensitivity enhances noises as well as signals. The long exposure causes motion blur. The lens with wide aperture yields out-of-focus images. The first two problems can be solved by denoising and deblurring, respectively, and there are many methods for solving the problems. Our aim is to address the third problem. We generate an all in-focus image without the under/over-exposure and out-of-focus areas due to the wide aperture by integrating the multiple exposure images. We introduce a new technique for image integration which simultaneously addresses the problems of the under/over-exposure and defocus. The validity of the proposal technique is shown by comparing it with several conventional methods for real scenes.

Index Terms— Multiple Exposure Images, All In-focus Image, Image Restoration

1. INTRODUCTION

The Human Visual System (HVS) can acquire a large dynamic range of light in a natural scene. The dynamic ranges of many commercial camera devices are narrower, and thus pixel saturation (i.e. under/over-exposure) often occurs in an image. Many methods have been proposed to integrate multiple exposure images to enhance the dynamic range of an image without the under/over-exposure [1], [2], [3], [4], [5]. In [1], [2], a design method for the ICRC (Inverse Camera Response Curve) that cancels the effect of a nonlinear camera response called CRC (Camera Response Curve) is proposed. After linearizing the camera response effects in multiple exposure images, the high dynamic range (HDR) image is obtained by the weighted sum of the images, which restores a natural scene with few under/over-exposures areas. The intensity of the HDR image is transformed to a low dynamic range image by so called tone mapping when displaying it in devices with limited dynamic ranges [6], [7], [8]. Some methods such as [3] and [9] skip the HDR generation step and directly create the low dynamic range image without the under/over-exposure. The procedure is called exposure fusion. In [3], contrast enhancement is realized by integrating multiple exposure images in the multiresolution pyramid with Gaussian weights calculated based on the color saturation, exposure, and contrast of the images. One of desirable features in the exposure fusion is that it can generate an integrated image with high contrast, but it can skip the procedure of camera response curve linearization, the information on exposure time, and tone mapping for displaying on devices with limited range. However the method does not handle images with out-of-focus areas.

We propose a new integration method for multiple exposure images which realizes simultaneous restoration of the under/over-exposure and out-of-focus areas. In the proposed method we generate two images by a weighted image integration technique. The two images are generated with the weights based on the color saturation/exposure and contrast. The two images are further combined by an optimization based image integration to obtain a final image that is free from under/over-exposure and out-of-focus artifacts.

In the paper, first we explain conventional work related to our method, and then propose a new integration algorithm to realize the restoration for the under/over-exposure and our-of-focus artifacts in Section 3. In Section 4, we show the validity of our algorithm by some experimental results, and then we conclude the work in Section5.

2. CONVENTIONAL METHOD

The exposure fusion in [3] utilizes the following three weights derived from the color saturation, exposure, and contrast of the input image to perform the weighted image integration.

This work was supported in part by JSPS Grants-in-Aid (24560473) and KDDI foundation.

Fig. 1. Example of input multiple exposure images with different focal lengths.

$$\text{Color saturation: } w_s = \sqrt{\frac{(I_R-\mu)^2+(I_G-\mu)^2+(I_B-\mu)^2}{3}}, \tag{1}$$

where I_r, I_g and I_b are the RGB channels of an image I, respectively. μ is its mean value over the three channels.

$$\text{Exposure:} \quad w_e = \exp(-\frac{(I-0.5)^2}{2\sigma^2}), \tag{2}$$

where σ is a parameter for the standard deviation of the Gaussian-like function.

$$\text{Contrast:} \quad w_c = |h_l * I_y|, \tag{3}$$

where h_l is a Laplacian filter and , I_y is the intensity of the image, and $*$ indicates a convolution.

In [3], they use multiresolution integration with Gaussian and Laplacian pyramids for the image fusion. Letting the t-th layer of an image $\mathbf{u} \in \mathbb{R}^N$ (where N is the number of pixels) in the Gaussian and Laplacian pyramids be $g_t(\mathbf{u}) \in \mathbb{R}^{N/4^t}$ and $l_t(\mathbf{u}) \in \mathbb{R}^{N/4^t}$, respectively, then the fusion in the t-th layer is given by

$$l_t(\mathbf{r}) = \sum_{k=1}^{K} g_t(\mathbf{w}_f^{(k)}) \otimes l_t(\mathbf{u}^{(k)}), \tag{4}$$

where (k) indicates the k-th exposure, and $\mathbf{w}_f^{(k)} \in \mathbb{R}^N$ is a weight map calculated with a k-th exposure image. $g_t(\mathbf{w}_f^{(k)})$ is the t-th layer in Gaussian pyramid of the k-th weight map, and $l_t(\mathbf{u}^{(k)})$ is t-th layer in Laplacian pyramid of the k-th weight map. The operator $\otimes$ is a pixel-wise multiplication. The image integration using the pyramid is performed by the weighted mean (4) for each layer, followed by pyramid synthesis for $l_t(\mathbf{r})$ to obtain a composed result.

3. PROPOSED METHOD

In our method, the multiple exposure images are taken with wide aperture (i.e. low f-stop). The use of the wide aperture can take a larger amount of light and relieve motion blur, but it causes out-of-focus regions. Our main aim is to combine multiple exposure images which are taken with different focal lengths (see Fig.1) and to simultaneously restore a single all in-focus image with a high dynamic range.

To construct our algorithm we assume that a multiple exposure image set shares a rough structure with sharp edges, while the presence of detailed texture depends on exposure level, i.e. the texture in highlights only appears in a high exposure image. Many of the tone mapping methods treat the structure and texture components separately for flexible handling of edges. Our method also starts with the structure/texture decomposition. We use Karacan et al.'s method [10] for the decomposition. The structure and texture components of the input images are further processed by the method described in the Sec.3.2 and 3.3, respectively. The procedure is illustrated in Fig.2.

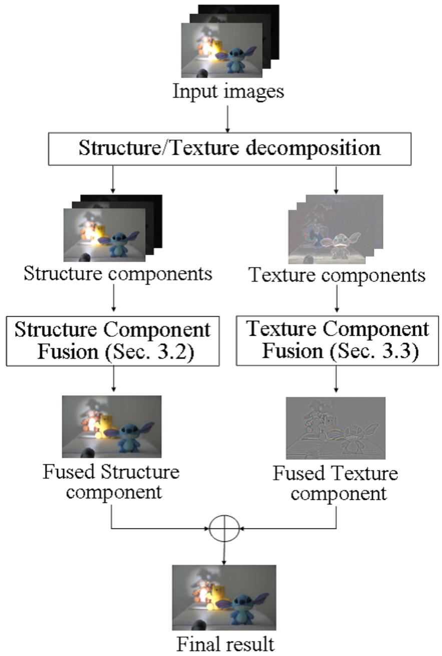


Fig. 2. Proposed exposure fusion

3.1. Image Integration

We combine the structure components of the multiple exposure images before obtaining a final image by optimization. When input images have out-of-focus area, one should assign low weights for the area. The conventional method [3] sometimes yields blurring artifacts and regions with low contrast, since the weights are obtained by a simple multiplication of several images. In our method, we take a different approach to address the problem. First we generate two images $\mathbf{y}_{w_{s,e}} \in \mathbb{R}^N$ and $\mathbf{y}_{w_c} \in \mathbb{R}^N$ by combining the structure components of the multiple exposure images using (4). One image $\mathbf{y}_{w_{s,e}}$ is obtained by the weights $w_{s,e} = w_s \times w_e$ introduced in [3], which is calculated from the color saturation and exposure. The other image $\mathbf{y}_{w_c}$ is obtained by the weight derived from the contrast. In our method, we slightly modify the procedure to calculate w_c in (3), that is, we further apply Gaussian filter to the output $h_l * i_y$ in (3), and then we adopt its absolute value as the weight w_c , since simply applying Laplacian filter sometimes yields ghost due to the ringing artifacts caused by Laplacian filter in our case.

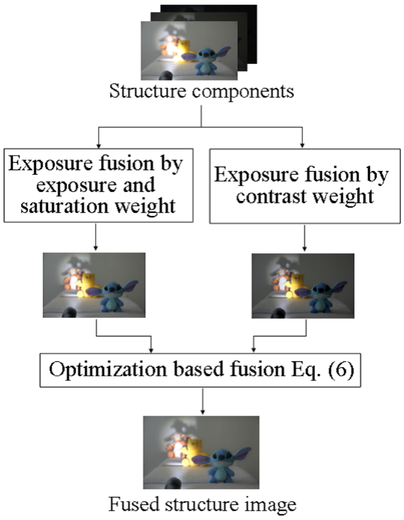


Fig. 3. The detail of structure component fusion.

3.2. Structure Image Integration with Optimization

Using the two images, $\mathbf{y}_{w_{s,e}}$ and $\mathbf{y}_{w_c}$, we generate a final image by the optimization based integration. The image $\mathbf{y}_{w_{s,e}}$ calculated with the weights of the color saturation and exposure has few under/over-exposed areas, while the other image $\mathbf{y}_{w_c}$ has clear contrast over the whole image. A next step is to utilize the features of the two images and integrate them to a final image. The procedure of this step is shown in Fig.3.

The integration problem in our method is stated as the minimization of the following cost function

$$\min_{\mathbf{s}} \|\mathbf{y}_{w_{s,e}} - \mathbf{s}\|_2^2 + \lambda \|\mathbf{L}(\mathbf{y}_{w_c} - \mathbf{s})\|_2^2, \tag{5}$$

where $\mathbf{L} \in \mathbb{R}^{N \times N}$ is a convolution matrix for the Laplacian filter, and $\lambda \in \mathbb{R}$ is a parameter to balance the two terms. The cost function is designed based on the concept that the latent image roughly resembles $\mathbf{y}_{w_s,e}$, while the detailed contrast is similar to $\mathbf{y}_{w_c}$. The first term is a fidelity function over the image $\mathbf{y}_{w_{s,e}}$ combined by the weights of the color saturation and exposure. The second term is introduced to approximate the image $\mathbf{y}_{w_c}$ in the gradient domain.

As Eq.(5) is a quadratic form w.r.t. $\mathbf{s}$, the solution is simply obtained by solving the normal equation:

$$\mathbf{s} = \left(\mathbf{I} + \lambda \mathbf{L}^T \mathbf{L}\right)^{-1} \left(\mathbf{y}_{w_{s,e}} + \lambda \mathbf{L}^T \mathbf{L} \mathbf{y}_{w_c}\right), \tag{6}$$

where $\mathbf{I} \in \mathbb{R}^{N \times N}$ is an identity matrix. Note that as the matrix to solve is a block circulant matrix with circulant blocks (BCCB), it is diagonalized by FFT, and thus the solution can be quickly calculated.

3.3. Texture Image Integration

The texture components are combined by taking the maximum value for each pixel. If the max values are taken in RGB channels independently, the color balance of the image is often damaged. Thus we perform the following procedure instead. First, we find the exposure image that has the maximum value of the l_2 norm of the RGB channels for each pixel,

$$k_i^* = \arg\max_k \sqrt{(t_{Ri}^{(k)})^2 + (t_{Gi}^{(k)})^2 + (t_{Bi}^{(k)})^2}, \tag{7}$$

where $t_{Ri}^{(k)}, t_{Gi}^{(k)}, t_{Bi}^{(k)}$ are the RGB components of i-th pixel of the texture components derived from the k-th exposure. Then all the three RGB values of k_i^*-th exposure image is stored at i-th pixel of the resultant texture image $\mathbf{t}$. This procedure is applied to all the pixels. Since the texture components have large energy around the in-focus area, one can obtain the texture with high contrast by taking the maximum values

Using the integrated structure component $\mathbf{s}$ in (6) and texture component $\mathbf{t}$, the final result is obtained by

$$\mathbf{y} = \mathbf{s} + \mathbf{t}. \tag{8}$$

4. EXPERIMENTAL RESULTS

We prepare several multiple exposure image sets for different scenes. We took the images by varying exposure time and focal length with other settings fixed. We integrate three exposure images and compare it with some conventional methods. Fig.4 shows the results of the three conventional methods: EF [3], DPEF[9], IFGF[11], and ours. In the results of DPEF and IFGF of Fig.4(a), halo artifacts appear around sharp edges, resulting in unnatural appearance (we indicate it by the blue circles in Fig.4(a)), while both of EF and our method preserve natural appearance without the halo artifacts. In 4(b), EF and DPEF do not restore high frequency components, and the out-of-focus artifact still remains, especially in the area with high contrast. The areas with low contrast are indicated by the red circles in Fig.4(b). On the contrary, IFGF and our method restores sharp edges in the whole images.

After all, the comparison can be summarized as follows. The methods EF and DPEF often fail in restoring contrasts especially in the out-of-focus areas. IFGF mostly restores high contrast in whole images, but it also yields severe halo artifacts. In contrast, our method successfully realizes the simultaneous restoration for under/over-exposure and out-of-focus areas without the halo artifacts and yields more natural images.

For quantitative comparison, we adopt Mean Absolute Laplacian (MAL), which is the mean absolute values of the Laplacian output, as a measure to evaluate the contrast. Table 1 shows the results, in which higher values indicate higher contrast. One can see from the table that our method outperforms the other methods in the scores.

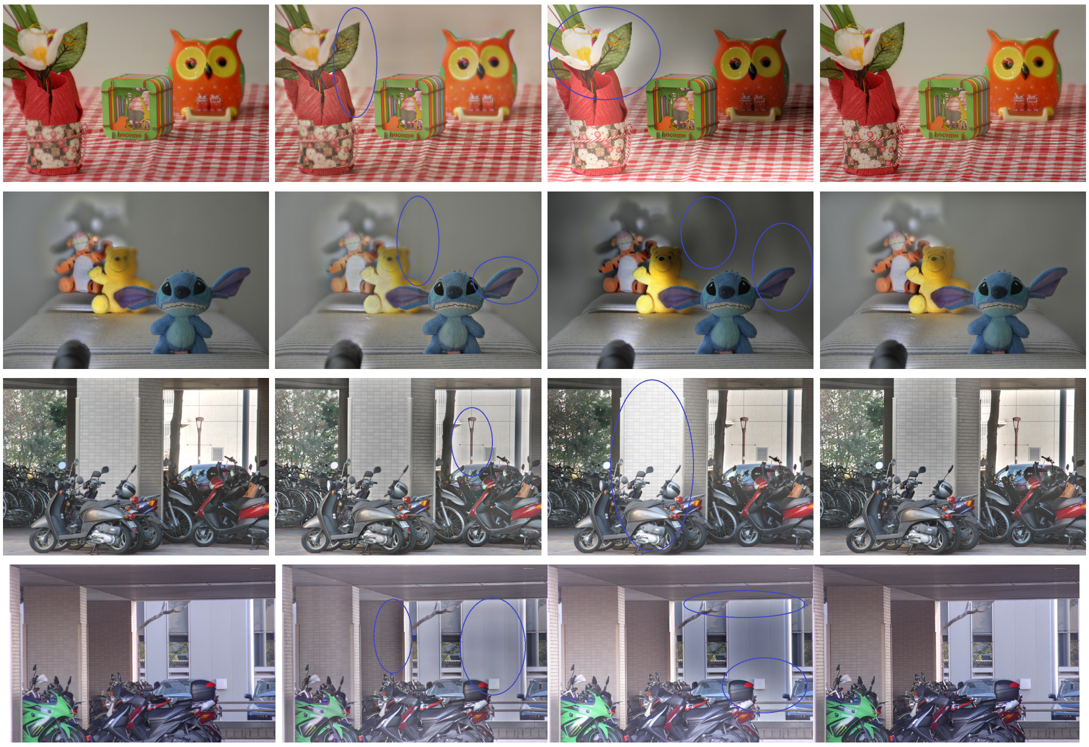
5. CONCLUSION

We propose a method to simultaneously restore under/over-exposure and out-of-focus areas by integrating multiple exposure images. Our method can generate all in-focus im-

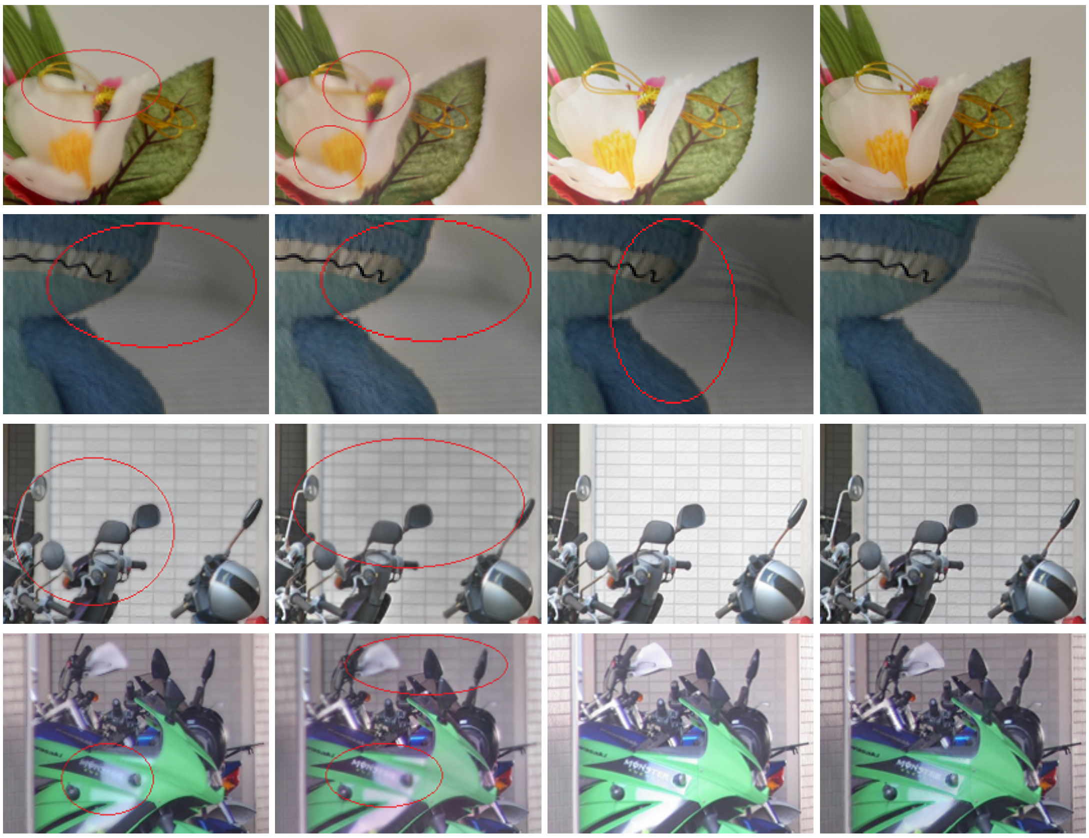
Table 1. Comparison of Mean Absolute Laplacian

	scene1	scene2	scene3	scene4
EF [3]	0.0258	0.0115	0.0502	0.0325
DPEF [9]	0.0244	0.0126	0.0527	0.0330
IFGF [11]	0.0314	0.0123	0.0545	0.0371
Our method	0.0348	0.0137	0.0579	0.0385

ages without any under/over-exposure. From the experimental comparison, we show the superiority of our method.



(a) Fused results of multiple exposure images



(b) Close up of fused results

Fig. 4. Fused results of multiple exposure images - (from top to bottom) scene1, scene2, scene3, and scene4: (from left to right) EF [3], DPEF [9], IFGF [11], and Our method.

6. REFERENCES

[1] T. Mitsunaga and S. Nayar, "Radiometric self calibration," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, vol. 1, 1999, pp. 374–380.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes*, ser. SIGGRAPH '08, 2008, pp. 31:1–31:10.

[3] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion," *Computer Graphics and Applications, Pacific Conference on*, vol. 0, pp. 382–390, 2007.

[4] R. Matsuoka, T. Baba, M. Okuda, and K. Shirai, "High dynamic range image acquisition using flash image." in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2013, pp. 1612–1616.

[5] R. Matsuoka, T. Yamauchi, T. Baba, and M. Okuda, "Weight optimization for multiple image integration," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, Sept 2013, pp. 795–799.

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

[7] N. Sugiyama, H. Kaida, X. Xue, T. Jinno, N. Adami, and M. Okuda, "Hdr image compression using optimized tone mapping model," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*. IEEE, 2009, pp. 1001–1004.

[8] T. Jinno, M. Okuda, and N. Adami, "New local tone mapping and two-layer coding for hdr images," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 765–768.

[9] J. Shen, Y. Zhao, and Y. He, "Detail-preserving exposure fusion using subband architecture." *The Visual Computer*, vol. 28, no. 5, pp. 463–473, 2012.

[10] L. Karacan, E. Erdem, and A. Erdem, "Structure-preserving image smoothing via region covariances," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 176:1–176:11, Nov. 2013.

[11] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, pp. 2864–2875, 2013.