Reconstruction of Super Resolution High Dynamic Range Image from Multiple-Exposure Images

Tae-Hyoung Lee, Ho-Gun Ha, and Yeong-Ho Ha
School of Electrics Engineering, Kyungpook National University, Taegu, South Korea

Abstract

Recent research efforts have focused on combining high dynamic range (HDR) imaging with super-resolution (SR) reconstruction to enhance both the intensity range and resolution of images beyond the apparent limits of the sensors that capture them. The processes developed to date start with a set of multiple-exposure input images with low dynamic range (LDR) and low resolution (LR), and require several procedural steps: conversion from LDR to HDR, SR reconstruction, and tone mapping. Input images captured with irregular exposure steps have an impact on the quality of the output images from this process. In this paper, we present a simplified framework to replace the separate procedures of previous methods that is also robust to different sets of input images. The proposed method first calculates weight maps to determine the best visible parts of the input images. The weight maps are then applied directly to SR reconstruction, and the best visible parts for the dark and highlighted areas of each input image are preserved without LDR-to-HDR conversion, resulting in high dynamic range. A new luminance control factor (LCF) is used during SR reconstruction to adjust the luminance of input images captured during irregular exposure steps and ensure acceptable luminance of the resulting output images. Experimental results show that the proposed method produces SR images of HDR quality with luminance compensation.

Introduction

Many efforts have been made to enhance the image quality of digital still cameras by improving the physical performance of their image sensors. Even so, digital cameras still suffer from limited dynamic range and resolution that is less than that encountered in the real world.

High dynamic range (HDR) imaging algorithms have been developed to overcome the problem of underexposed or overexposed images caused by the narrow dynamic range of cameras. This involves assembling multiple-exposure low dynamic range (LDR) images from a normal camera to obtain a full dynamic range image [1]. However, HDR imaging techniques require that the camera response curve (CRC) for each camera is capable of recovering the intensity of an actual scene [2]. Additionally, many common displays have a limited dynamic range and cannot display HDR images directly; such displays require a tone-mapping process to compresses the dynamic range of the image to fit their dynamic ranges [2]. Many tone-mapping algorithms have been proposed, but each is tailored to a specific purpose and thus has its own advantages and disadvantages. Thus, the quality of HDR images is not always preserved.

The limited resolution of normal digital cameras has been overcome with SR reconstruction, which increases the spatial resolution by exploiting the correlation of several sequential input images obtained by the camera under identical conditions. Because SR reconstruction requires the use of multiple images with the same exposure time, combining it with HDR imaging is difficult because the later requires multiple-exposure images. Even so, recent studies have addressed the challenge of combining HDR images and SR reconstruction to obtain high-quality, high-resolution images with high dynamic range. Gunturk sought to obtain HDR-SR images, proposing a new imaging model during SR reconstruction that included dynamic range and spatial domain effects [3]. Figure 1 and the following equation describe that imaging model:

$$\mathbf{z}_k = f(t_k \mathbf{H}_k q + v_k) + \mathbf{w}_k \tag{1}$$

where z_k is a low-resolution observation for the k^{th} LDR-LR image with Y channel from YCbCr values, $f(\cdot)$ is the nonlinear camera response function, t_k is the exposure time, q is a high-resolution input signal, and $\mathbf{H}_k$ is the linear mapping that incorporates motion, the point spread function, and down-sampling. Although the resulting images contain HDR information, presenting them on normal displays requires tone mapping, which may not give satisfactory results depending on the tone mapping method used.

Schubert also suggested a framework for combining HDR and SR [4]. He used a similar imaging model that included photometric camera calibration data obtained using Debevec's method and tone mapping [2].

Both methods require estimating the CRC function to obtain the dynamic range of the actual scene; however, tone mapping is also required and produces results that are not always satisfactory. Indeed, estimation of the real-world CRC function is not a simple task. Furthermore, each camera requires its own CRC function, and tone-mapping algorithms have a very significant influence on the quality of the resulting images.

Here, we propose a new framework for obtaining high-quality images that appear to have high dynamic range and super resolution. Because it blends multiple-exposure images, the framework is robust in the face of input images from irregular exposure steps. The proposed method first calculates weight maps,

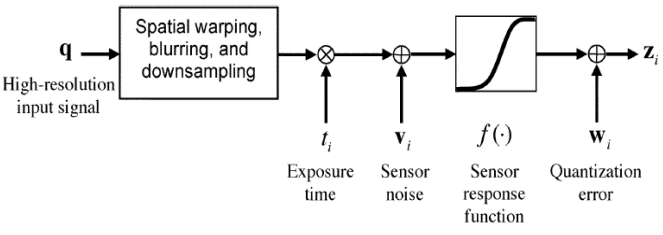


Figure 1. *Imaging model by Gunturk.*

which it then applies directly to SR reconstruction with pyramid merging [5, 6]. These weight maps are obtained by retaining only

the best visible parts in each multiple-exposure input image to preserve details of dark and highlighted regions during SR reconstruction and produce SR output images that have HDR quality without LDR-to-HDR conversion [5]. However, the use of weight maps alone does not result in suitable luminance in output images because of the set of input images from irregular exposure steps. Thus, a new luminance control factor (LCF) is proposed for application during SR reconstruction to correct this. The resulting images have suitable resolution in both light and dark sections.

SR reconstruction based on weight maps

The overall objective of this work was to obtain an SR image that has fine details in light and dark sections. Figure 2 is a flowchart for the proposed method assuming four input images with proper exposure steps. The specific procedure has two parts: color reproduction and SR reconstruction from multiple-exposure images. During color reproduction, exposure fusion is performed on a series of LDR-LR input images to produce a single HDR-LR image [5]. Then, only CbCr values are extracted and resized for use in the last conversion from YCbCr to digital RGB values. Next, SR reconstruction takes place assuming perfect image registration. During this process, weight maps are first calculated to determine the necessary areas in each input image that contain the "best" intensities. These weight maps are then applied to SR reconstruction with pyramid merging to prevent intensity distortion from unwanted iterations. Simultaneously, LCFs (η_k) are applied to each input image to adjust the luminance so that the overall luminance of the output image is suitable.

Weight map construction

The exposure fusion algorithm generates high-quality images such as those that result from HDR imaging [5]. That is, the resulting images show fine details in dark and lighted areas. This algorithm selects the "best" parts of images in a multi-exposure image sequence. These "best" parts are defined as a weighted map based on a combination of quality measures, such as contrast, saturation, and good exposure [5], and this map is used to blend the input images.

The contrast measure C is calculated as the absolute value of the Laplacian-filtered image for each gray-scaled image. The saturation measure S is obtained as the standard deviation within the R, G, and B channels at each pixel. The measure of good exposure E indicates how well a pixel is exposed, and it is used as a weight of the intensity, based on how close it is to 0.5 using a Gaussian curve defined as $e^{-(i-0.5)2/2\sigma 2}$. This is applied to each channel. The final weight maps for each k^{th} (k = 1, .., N) image are defined as follows:

$$W_{ij,k} = C_{ij,k} \times S_{ij,k} \times E_{ij,k} \tag{2}$$

where i and j are the location of pixels, and k is the index of the input sequence images.

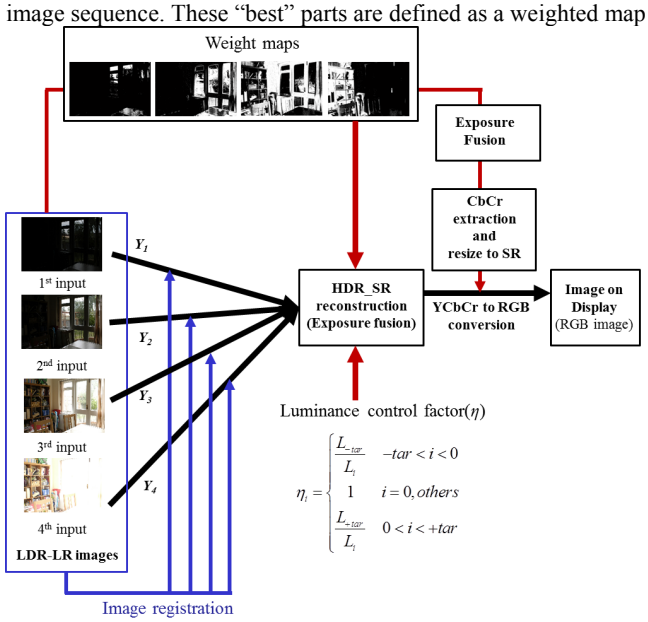


Figure 2. *Flow of the proposed method.*

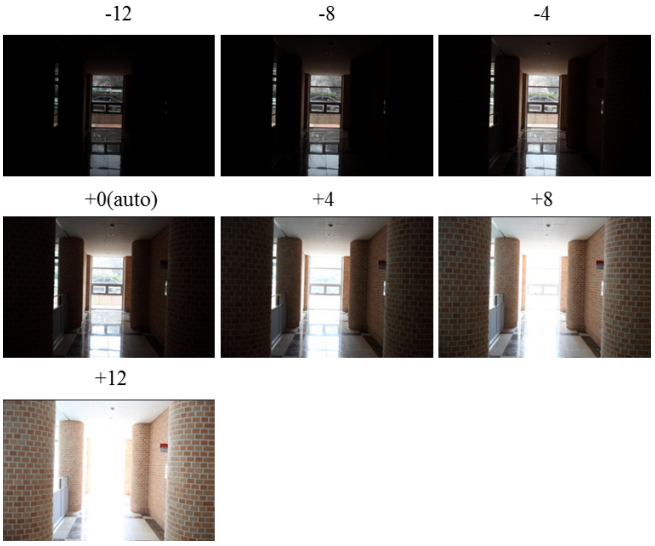


Figure 3. *A set of test input images.*

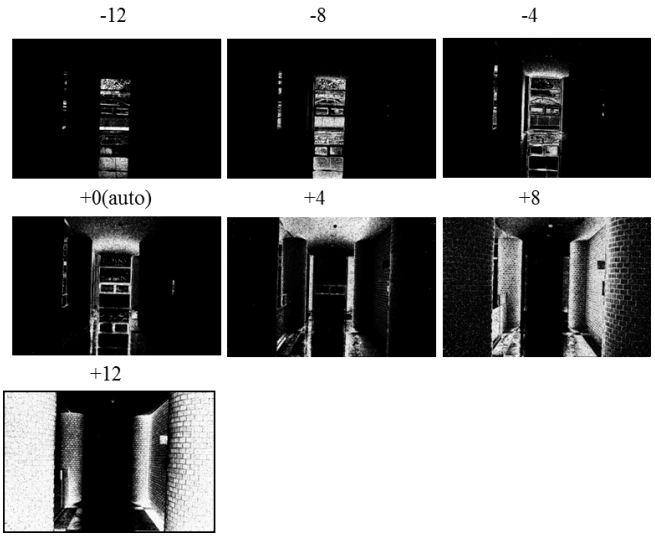


Figure 4. *Weight maps for each input image.*

Figure 3 shows a test set of multiple-exposure image sequences with RGB channels with Tiff format. Because the resulting image is affected by the set of input images, seven images were prepared from the auto-exposed image using a Canon 5D Mark 2 camera with fixed aperture value of 5.0 for use a as reference image.

Figure 4 shows the results of weight maps for each input image. Dark and highlighted areas have low values on the weight maps; high values indicate the "best" areas of the input images.

Proposed SR reconstruction

We first assume the imaging model to be the following:

$$z_k = \mathbf{H}_k q + v_k \tag{3}$$

where z_k, $\mathbf{H}_k$, q, v_k, and k are the same from equation (1). If the image registration is perfect, then $\mathbf{H}_k$ is composed of spatial warping, blurring, and down-sampling. In the deterministic approach of SR reconstruction, the inverse imaging model can be solved by choosing q to minimize the following cost function:

$$\left[\sum_{k=1}^{N} \| \mathbf{z}_k - \mathbf{H}_k q \|^2 + \alpha \| Cq \|^2\right] \tag{4}$$

where $\| \bullet \|$ is l_2-norm, C is a high-pass filter, and α is a regularization parameter. Gradient descent techniques are used with the equation above to estimate q. Iterative updates are shown for the m^{th} iteration:

$$q^{m+1} = q^m + \beta \nabla E(q^m) \tag{5}$$

where ∇E is a negative gradient calculated as follows:

$$\nabla E(q) = \sum_{k=1}^{N} \mathbf{H}_k^T (\mathbf{z}_k - \mathbf{H}_k q) - \alpha \mathbf{C}^T \mathbf{C} q \tag{6}$$

In the proposed method, the weight map is applied to gradient-descent terms. That is, the gradient descent is performed with weight maps to apply the "best" intensity during SR reconstruction. Thus, the gradient descent is modified with the weight maps as follows:

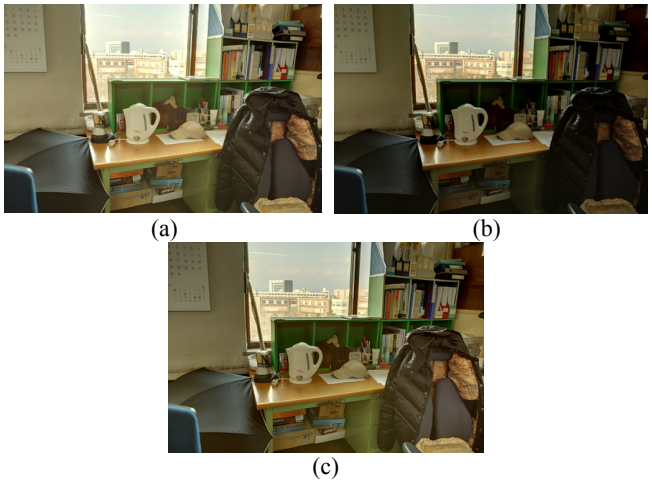
$$\nabla E(q) = -\sum_{k=1}^{N} \mathbf{H}_k^T W_k (\mathbf{z}_k - \mathbf{H}_k q) - \alpha \mathbf{C}^T \mathbf{C} q \tag{7}$$

However, as in the case of exposure fusion described above, the simple application of weight maps causes seams and contour artifacts. In particular, during SR reconstruction, this process also induces unstable ∇E, causing image distortion due to over-iteration. Thus, Eqs. (7) and (8) can be expressed as

$$L\{q^{m+1}\} = L\{q^m\} - \alpha \nabla E(L\{q^m\}) \tag{8}$$

$$\nabla E(L\{q^m\}^l) = -\sum_{k=1}^{N} \mathbf{H}_k^T G\{W_k\}^l (L\{\mathbf{z}_k\}^l - L\{\mathbf{H}_k q^m\}^l) - \alpha \mathbf{C}^T \mathbf{C} q^m \tag{9}$$

Figure 5. *Resulting images from the proposed method by using 7 images.*



(a) (b)

(c)

Figure 6. (*a*) *bi-cubic interpolated image from the result of exposure fusion,* (*b*) the proposed method using 4 images, (c) the proposed method using 7 images.

where $G\{\bullet\}$ and $L\{\bullet\}$ are Gaussian and Laplacian pyramids, respectively, and l is the level of pyramid. Here, the SR reconstruction is performed for each level of pyramid by multiplying by the weight maps for each input image. The final resulting image is reconstructed using the inverse Laplacian pyramid after SR reconstruction for each level.

Figure 5 is the image obtained using the proposed method. The dark and highlighted areas are well represented. However, evaluation of the proposed method is difficult because reference images exist. Figure 6 compares our resulting images by using different sets of input images in Figures 6(b) and 6(c) with the bi-cubic interpolated image from the exposure fusion algorithm in Figure 6(a). Although the images in Figures 6(b) and 6(c) have better resolution, the overall luminance in Figure 6(b) is poorer than in either Figure 6(a) or 6(c). That is, the quality of resulting images depends on the set of input images.

Luminance compensation

The resized image resulting from exposure fusion in Figure 6(a) is not suitable for experimental comparison because it

was produced using a totally different process. Here, we suggest a base image for evaluation, then introduce the new LCF to adjust the luminance of the resulting image depending on the set of input images.

Images resulting from different sets of input images were analyzed to determine the base image. Test sets are distinguished according to the composition of different exposure steps. Here, we assume the reference image for SR reconstruction is obtained with auto exposure (+0), because normal still cameras usually capture photos using auto exposure. Then, images were obtained in four exposure steps, the intervals of which were decided empirically.

Figure 7(a) is reconstructed from seven input images with exposures of –12, –8, –4, +0, 4, 8, and 12 while for Figures 7(b) and 7(c), the number of input images is decreased by one level of exposure interval. The resulting image in Figure 7(a) looks better than the other images, which use input images with closer exposure steps, as in LDR images; these poorer images do not have enough information in the dark and highlighted areas. Our experimental results show that using more than seven images does not improve the quality. Additionally, the image in Figure 6(c) has HDR quality that is similar to that obtained with exposure fusion. The other test set produced almost the same result using seven images with exposures of –12, –8, –4, +0, 4, 8, and 12 as the base image.

A new LCF η_k is introduced to adjust the luminance of the image resulting from an irregular set of input images, such as those shown in Figures 7(b) and 7(c). That is, the LCF makes the images in Figures 7(b) and 7(c) similar to that of 7(a). This LCF is applied to the input images to adjust their overall luminance, making low-exposure areas darker and high-exposure areas lighter. This is required because the proposed method requires input images with a large exposure interval. Different LCFs should be applied to each input image because the luminances of the input images are not the same. Thus, we modified Eq. (9) to include the LCF as follows:

$$\nabla E(L\{q^n\}^l) = -\sum_{k=1}^{N} \mathbf{H}_k^T G\{W_k\}^l (\eta_k L\{\mathbf{z}_k\}^l - L\{\mathbf{H}_k q^n\}^l) - \alpha \mathbf{C}^T \mathbf{C} q^n \quad (10)$$

where η_k is the LCF for each input image.

Indeed, if the set of input images has enough information for dark and highlighted areas from the reference image, the resulting images appear similar to the base image, as shown in Figure 8. The images in Figures 8(b) and 8(c) were reconstructed with a set of three and five input images, respectively, which include the requisite +12 and –12 exposure steps. The appearance and average luminance of all three resulting images in Figure 8 are almost the same. Thus, the set of input images should include enough dark and light images with proper exposure steps, defined here as +target and –target exposure steps.

If we assume that the +target and –target exposure steps can be determined ahead of time, it is possible to make η_k a function to adjust the average luminance of the input images. That is, the role of η_k is to ensure that the average luminance of input images farthest from the reference image is as close to +target and –target exposure steps as possible.

Determining η_k as a function requires first determining the +target and –target exposure steps for sufficiently dark and light images. In experimental simulation, +12 and –12 steps from the

(a) (b)

(c)

Figure 7. *(a) 7 images with -12, -8, -4, +0, +4, +8, and +12 exposure steps, (b) 5 images with -8, -4, +0, +4, and +8 exposure steps,* (c) 3 *images with -4, +0, and +4 exposure steps.*

(a) (b)

(c)

Figure 8. *(a) 5 images with -12, -4, 0, +4, and +12 exposure steps, (b) 5 images with -12, -8, +0, +8, and +12 exposure steps,* (c) 3 *images with -12, +0, and +12 exposure steps.*

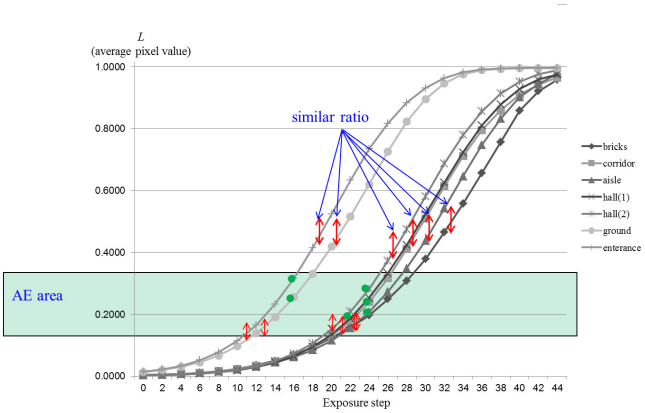


Figure 9. *Average luminance for each test images according to the exposure steps.*

reference (+0) were appropriate targets because more distant steps did not have any geometrical information. Determination of these target exposure steps is a subject for future study.

The average luminance values according to the exposure steps were investigated (Fig. 9) to determine the specific value of η_k. For each test image, even though the absolute average values were different, the percentage change for the corresponding exposure steps was almost the same. A similar concept applies during
construction of a camera response function. That is, the ratio between exposure times is the same as that between radiances [1]:

$$\frac{I_{p,q}}{I_{p,q+1}} = \frac{t_q}{t_{q+1}} = R_{q,q+1} \tag{11}$$

where I is the radiance, t is the exposure time, p is a pixel, and q is the sequence of input images.

Consequently, η_k is the ratio between the current and the target average luminance values:

$$\eta_k = \begin{cases} \frac{L_{-tar}}{L_i} & -tar < i < 0 \\ 1 & i = 0, others \\ \frac{L_{+tar}}{L_i} & 0 < i < +tar \end{cases} \tag{12}$$

where L is the average luminance, i is the exposure step, and k is the sequence of input images.

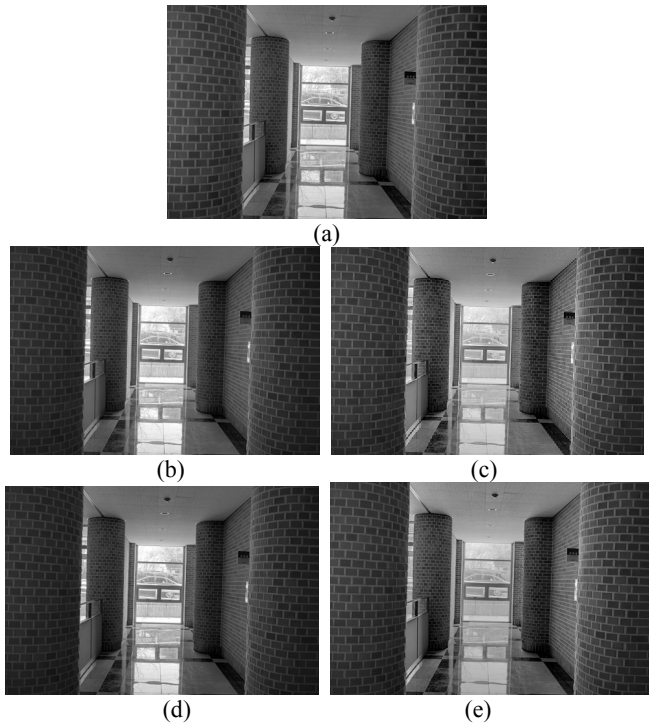


Figure 10. *Resulting images by adopting □k.(a) using 7 images (b) using 5 images without □k, (c) using 5 images with □k, (d) using 3 images without □k, and (e) using 3 images with □k*

Figure 11. *Color image version from figure 10(c) and (e).*

Experimental results

Figure 10 shows the images produced using η_k along with gray images for comparison of the luminance only. The resulting images with η_k in Figures 10(c) and 10(d) appear similar to the image in Figure 7(a). However, the color versions of Figures 10(c) and 10(e) in Figure 11 are clearly darker due to low saturation because the input images had already lost their color information. Color compensation is another area of future research.

The last comparison is performed with other methods. Narasimhan as method_1 suggested a HDR-SR image by sequential performance of HDR imaging, SR reconstruction, then tone mapping[7]. Hardie as method_2 also has the similar results with SR reconstruction, HDR imaging, and tone mapping algorithms[8]. Their results are shown in figure 12, presenting almost the same, because the same conversion algorithms are used during their process. However they are quite different from our

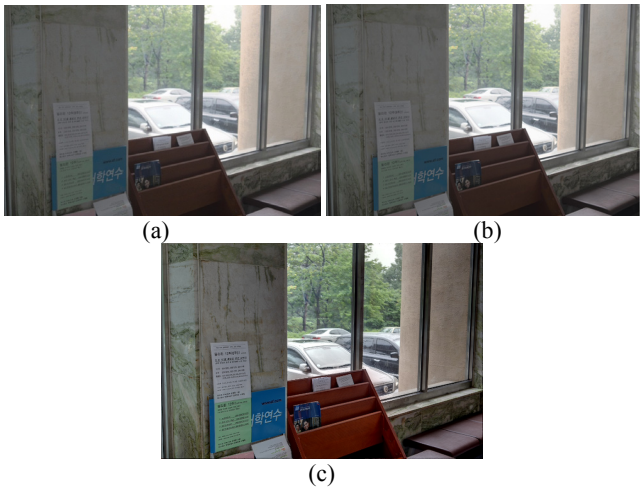


Figure 12. *Resulting images from other method. (a)method_1, and (b) method_2, and (b) proposed method.*

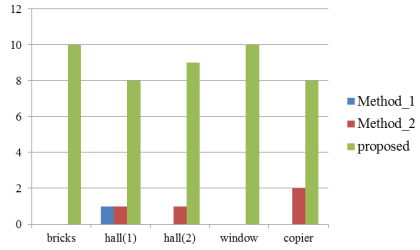


Figure 13. *Subjective evaluation with other methods.*

results because we don't use any conversion. As follows objective evaluation such as PSNR and SSIM is difficult. Finally, subjective evaluations are performed with 5 resulting images, and results are shown in figure 13.

Conclusions and discussions

We have proposed a simplified method for generating SR images using multiple-exposure images to obtain HDR information about dark and light areas. The use of weight maps permits the elimination of HDR conversion using the CRC function and LDR conversion using tone-mapping algorithms, both of which can affect the quality of the resulting image. The dynamic range is extended by determining the best visible areas from the weight maps. Use of the LCF function also makes preservation of suitable luminance of the resulting images possible based on different sets of input images.

Acknowledgement

This work was supported by Mid-career Researcher Program through NRF grant funded by the MEST (No.2011-0000152).

References

[1] K. H. Park, D. G. Park, and Y. H. Ha, "High Dynamic Range Image Acquisition from Multiple Low Dynamic Range Images Based on Estimation of Scene Dynamic Range," *Journal of Imaging Science and Technology*, **53**, 2, 020505-1 - 020505-12, Mar./Apr. (2009).

[2] E. Reinhard, G. Ward, S. Pattanik, and P. Debevec*, High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting, Morgan Kaufmann Publishers*, Dec. (2005).

[3] B. K. Gunturk and M. Gevrekci, "High-resolution image reconstruction from multiple differently exposed images"*, IEEE signal processing letteres*, **13**, 4, Apr. (2006).

[4] F. Schubert, K. Schertler, and K. Mikolajczyk, "A Hands-on Approach to High-Dynamic-Range and Superresolution Fusion," *IEEE Workshop on Application of Computer Vision,*1-8, (2009).

[5] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure Fusion," *IEEE 15th Pacific Conference on Computer Graphics and Application*, 382-390, Dec. (2007).

[6] S. C. Park, M. K. Park, and M. G. Kang, "Super-Resolution Image reconstruction: a Technical Overview," *IEEE Signal Processing Magazine*, **20**, 3, May (2003).

[7] Narasimhan, S.G. and Nayar, S.K., "Enhancing resolution along multiple imaging dimensions using assorted pixels," *IEEE Trans. Pattern Anal. Mach. Intell.*, 27, 518–530, (2005)

[8] Hardie, R.C., Barnard, K.J., Bognar, J.G., Armstrong, E.E. and Watson, E.A., "High-resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system," *Opt. Eng.*, 37, 247–260, (1998)

Author Biography

Tae-Hyoung Lee received his BS and MS in Electronic Engineering from Kyungpook National University, Taegu, Korea, in 2005 and 2007, respectively. Now he is Ph. D. candidate in Kyungpook National University. His research interests include color management, image quality evaluation, auto exposure control, super resolution, and high dynamic range imaging.

Ho-Gun Ha received his B.S. degrees in Electrical Engineering & Computer Science from Kyungpook National University, Korea in 2007. He has worked M.S. student at Color & Imaging Laboratory at Kyungpook National University since 2007. His research interests are image registration and color image enhancement.

Yeong-Ho Ha received the B. S. and M. S. degrees in Electronic Engineering from Kyungpook National University, Taegu, Korea, in 1976 and 1978, respectively, and Ph. D. degree in Electrical and Computer Engineering from the University of Texas at Austin, Texas, 1985. In March 1986, he joined the Department of Electronics Engineering of Kyungpook National University and is currently a professor. He served as TPC chair, committee member, and organizing committee chair of many international conferences held in IEEE, SPIE, and IS&T and domestic conferences. He served as president and vice president in Korea Society for Imaging Science and Technology (KSIST), and vice president of the Institute of Electronics Engineering of Korea (IEEK). He is a senior member of IEEE, a member of Pattern Recognition Society and Society for IS&T and SPIE, and a fellow of IS&T. His main research interests are in color image processing, computer vision, and digital signal and image processing.