IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 16, 2017, accepted April 3, 2017, date of publication April 13, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2694038

A Method for Fast Multi-Exposure Image Fusion

SEUNGCHEOL CHOI, OH-JIN KWON, AND JINHEE LEE
Department of Electrical Engineering, Sejong University, Seoul 05006, South Korea

Corresponding author: Oh-Jin Kwon (ojkwon@sejong.ac.kr)

This work was supported by the Standard Technology Development and Spread Program of KATS/KEIT [10069128, Developing Standard Platform and Library for JPEG Privacy-Security].

ABSTRACT This paper proposes a method for fusing multi-exposed images that can operate on digital cameras or smartphones. The proposed method consists of an automatic exposure bracketing algorithm that determines which exposures to capture and a newly proposed multi-exposure image fusion algorithm. This fusion algorithm attempts to improve the fusion performance on the basis of the recently proposed no-reference image-quality metrics, noting that the exposure change affects the change in the local luminance details, contrast, and colorfulness of a pixel. Experimental results of various sample image sequences show the superiority of the proposed fusion algorithm in terms of both objective and visual evaluations. By using the proposed method, users can capture high-dynamic range images directly on digital cameras or smartphones, without using offline image-processing software.

INDEX TERMS Multi-exposure image fusion, exposure bracketing, no-reference image quality metric, digital camera, smartphone.

I. INTRODUCTION

When photographers take a picture that comprises shadows or highlighted areas, they are faced with the challenge of setting the appropriate exposure. Fortunately, modern digital cameras and smartphones provide a high-dynamic range (HDR) mode to resolve this problem by fusing multi-exposed images of the same scene.

Figure 1 shows a typical operation of HDR mode in modern digital cameras and smartphones [1]. Modern digital cameras and smartphones provide their video-rate view-finding mode for taking pictures. When the shutter is on for HDR imaging, the HDR mode starts automatic exposure bracketing (or HDR metering) operation to capture images using predefined exposure settings. The exposure of an image may be controlled by varying the camera's exposure time and/or lens aperture. Since varying the aperture affects the depth of field between different captures, they normally perform exposure bracketing by varying exposure time via the camera's shutter speed [1]. For fast processing, reduced-size images are used for this operation. After the exposure bracketing operation, full-size images corresponding to the selected exposure settings are captured. Camera motion is aligned for the captured images. Then, reduced-size multi-exposure image fusion is performed for fast previewing. When the previewed fusion result needs to be changed by the user's decision, the aperture is optionally adjusted by the user and the exposure bracketing is repeated with a different exposure range. Full-size multi-exposure image fusion is performed when the user finally decides to save the fusion result.

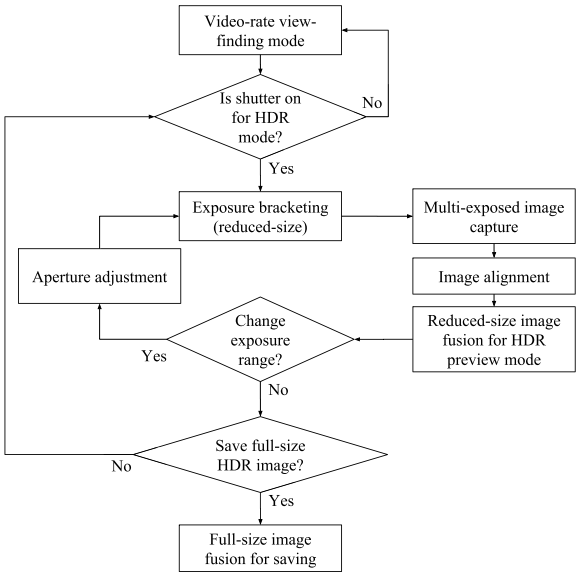

FIGURE 1. Typical operation of HDR mode in modern digital cameras and smartphones.

Many effective methods for fusing multi-exposed images have been proposed. For example, Debevec and Malik [2] developed an empirical model for determining the aggregate

mapping from scene radiance to pixel values. The authors fuse a set of differently exposed images and recover the HDR image whose pixel values are proportional to the true radiance values. Shen *et al*. [3] reproduced the HDR scene radiance in the way similar to the response of human visual system (HVS) by an HDR image generation method based on retinex theory. Robertson *et al*. [4] proposed a probabilistic formulation of the response function of cameras. The authors give the maximum likelihood solution of the HDR image for a known camera response and discuss how the response function can be estimated from the unknown camera response. Burt and Kolczynski [5] proposed a multi-purpose pyramid-based image fusion, which uses the gradient pyramid transform that can be used for fusing multi-exposed images. Jinno and Okuda [6] estimated occlusion and saturation regions by using maximum posteriori estimation and propose an algorithm of the HDR estimation based on the Markov random field model. Even though these methods have shown satisfactory fusion results, because of their computing cost, their usage without optimization for fast processing is inhibitive for the implementation on state-of-the-art digital cameras and smartphones.

Faster methods have been proposed in the literature on the subject. Mertens *et al*. [7] developed a flexible pyramid-based method by using a Laplacian decomposition of the multi-exposed images and a Gaussian pyramid of the weight maps representing measures of contrast, saturation, and well-exposedness. Goshtasby [8] proposed a block-based method for fusing multi-exposure images. His method partitions the image into uniform blocks, selects the image showing the most information for each block, and blends selected images together using monotonically decreasing blending functions centered at the blocks. Raman and Chaudhuri [9] designed a fast edge-preserving bilateral filter for compositing of a scene from images obtained through various exposure photographs. Shen *et al*. [10] used a generalized random walks framework for probabilistic model based fusion of multi-exposed images achieving an optimal balance between local contrast and color consistency. Gu *et al*. [11] noted that HVS is sensitive to contrasts between pixel intensities, not the absolute values. The authors propose the method fusing multi-exposure images in the gradient field where the gradient values of every pixel point are generated from maximizing the structure tensor. Kao *et al*. [12] proposed an empirical fusing method for multi-exposed images by detecting unstable pixels. The authors have also employed the sum of absolute difference of macro block for the camera motion compensation. Jacobs *et al*. [13] designed a method constructing the HDR image by calculating a weighted average of the irradiance values. The authors use the edge-based camera alignment. Zhang and Cham [14] proposed a method to handle the composition of multi-exposed images with the guidance of visibility and consistency measures derived from the gradient information. Li and Kang [15] proposed a weighed sum based multi-exposure image fusion method whose weight maps are initially measured by local contrast, luminance, and color dissimilarity, and then refined by the recursive edge-preserving smoothing filter.

In this paper, we propose a method for providing a novel HDR mode on digital cameras or smartphones. The proposed method attempts to include the following features:

1) The proposed method includes a new well-exposedness metric employed for fast processing. This new metric is designed on the basis of the usage of the recently proposed no-reference (NR) image quality assessment (IQA) metrics. NR-IQA provides the absolute score for the processed image itself without any reference. Therefore, NR-IQA metrics are very useful in practical applications where the reference image is not available, such as in the case of image fusion [16], [17]. Whereas the colorfulness attribute of image quality has received much less attention in previously proposed methods [2]–[15], we show that the exposure change affects the change in the colorfulness of a pixel and involve this attribute for the design of proposed metric.
2) The proposed method includes an automatic exposure bracketing procedure. Although HDR imaging has a long history, the literature on the subject [2]–[15] has rarely addressed the automatic exposure bracketing algorithm [1]. Compared with the well-studied problem of auto-focusing to find the best-focusing lens position, the problem of exposure bracketing to determine the priority of exposure settings for selecting input images for the fusion has received much less attention. Based on the knowledge that three types of standard search algorithms have been adopted for auto-focusing: global search, binary search, and rule-based search [18], [19], we design an automatic exposure bracketing algorithm using the proposed well-exposedness metric.

This paper is organized as follows. The proposed method including the proposed well-exposedness metric and an automatic exposure bracketing algorithm is described in Sect. II. The experimental results are presented in Sect. III. Finally, Sect. IV concludes the paper.

II. PROPOSED METHOD

A. PROPOSED WELL-EXPOSEDNESS METRIC

A previous study on human perception on image quality proposes a number of attributes for IQA. These attributes include overall luminance, contrast, sharpness, details, naturalness, and colorfulness [16], [17], [20]–[24]. Many researchers have proposed NR-IQA metrics on the basis of these attributes and investigated the effect of exposure change on these attributes. Their results show that the exposure change affects the change in the local luminance details, contrast, and colorfulness of a pixel [2]–[15], [25]. Therefore, in this paper, simple metrics corresponding to these attributes are proposed on the basis of the research results of previously proposed NR-IQA metrics. For the metrics measuring each of the local luminance details, contrast, and colorfulness attributes of a pixel, the number of zero-crossings (*ZC*) [20], sum-modified

Laplacian (*SML*) [26], and the colorfulness metric $\hat{M}^{(3)}$ proposed by Hasler and Suesstrunk [22] are chosen, respectively. All the proposed metrics are chosen on the basis of their effectiveness and possibility of fast calculation. The definitions of the proposed metrics with notations are as follows:

Notations:

- $X = \{x(m,n)\}, (m,n) \in \Omega = \{1 \le m \le M, 1 \le n \le N\}$:
  An image where M and N denote the vertical and the horizontal image sizes
- $x_r(m,n)$, $x_g(m,n)$, $x_b(m,n)$, and $x_y(m,n)$:
  The red, green, blue, and luminance component of $x(m,n)$, respectively
- $B_{m,n}$:
  The block centered at (m,n) whose size is 5×5

Definitions:

- Number of zero-crossings:

$$\begin{aligned} &ZC_X(m,n) \\ &= \text{Number of vertical zero-crossings of } x_y(k,l) \text{ in } B_{m,n} \\ &+ \text{Number of horizontal zero-crossings of } x_y(k,l) \text{ in } B_{m,n} \end{aligned}$$

- Sum-modified Laplacian:

$$\begin{aligned} &SML_X(m,n) \\ &= \sum_{(k,l)\in B_{m,n}} \{\, |\, 2x_y(k,l) - x_y(k-1,l) - x_y(k+1,l)\, | \\ &\qquad + |\, 2x_y(k,l) - x_y(k,l-1) - x_y(k,l+1)\, | \,\} \end{aligned}$$

- Colorfulness:

$$\hat{M}_X^{(3)}(m,n) = \sqrt{\sigma_{rg}^2 + \sigma_{yb}^2} + 0.3\sqrt{\mu_{rg}^2 + \mu_{yb}^2}$$

  where μ_{rg}, μ_{yb}, σ_{rg}^2, and σ_{yb}^2 denote the mean of $x_{rg}(k,l)$, the mean of $x_{yb}(k,l)$, the variance of $x_{rg}(k,l)$, and the variance of $x_{yb}(k,l)$, respectively, for $(k,l) \in B_{m,n}$, with $x_{rg}(k,l) = x_r(k,l) - x_g(k,l)$ and $x_{yb}(k,l) = \frac{1}{2}[x_r(k,l) + x_g(k,l)] - x_b(k,l)$.

It is noted that *ZC* and *SML* are normally computed using the luminance component. However, in some digital cameras, RGB color coordinates are the default, so luminance must be computed from the RGB colors. If this extra processing becomes a critical burden for fast processing, using just the green component, which is the largest contributor to luminance, provides sufficient information.

To demonstrate the effectiveness of proposed metrics, relative measurements of each metric for two differently exposed images of the same scene are exemplified. Sample images are shown in Fig. 2a, b. For each metric, measurements were obtained for two images and scaled by the same scaling factor for display. They are shown in Fig. 2c–h. It is shown that each metric not only reflects the well-exposed regions reasonably but also shows its own demerits. For example, *ZC* shows large overlapped results of similar measurement in the boundary area between a well-exposed region and a badly-exposed region. *SML* is sensitive on edges and does not reflect the well-exposedness in homogeneous regions. $\hat{M}_X^{(3)}$ measures the well-exposedness on the basis of colorfulness, therefore, its usage is prohibited in a grey area. In conclusion, no single metric among the proposed metrics can reflect the well-exposedness effectively.

This paper presents two simple and effective well-exposedness metrics developed using a combination of the aforementioned metrics. The first metric called the relative well-exposedness (RWE) metric measures the relative well-exposedness between two images. The second metric called the absolute well-exposedness (AWE) metric measures the well-exposedness of an image. When two images, namely image 1 and image 2, are denoted by $X^{(1)} = \{x^{(1)}(m,n)\}$ and $X^{(2)} = \{x^{(2)}(m,n)\}$, respectively, the proposed RWE metric of image 1 to image 2 at (m,n), denoted by $RWE_{1:2}(m,n)$, is calculated as follows:

1) Three metrics: *ZC*, *SML*, and $\hat{M}^{(3)}$, are calculated for image $X^{(1)}$ and $X^{(2)}$, respectively.
2) For each metric, the maximum and minimum values are obtained for a set of two images.
3) By using these maximum and minimum values, the measurement of each metric for image $X^{(1)}$ is normalized.
4) $RWE_{1:2}(m,n)$ is obtained by summing the results of three metrics.

It is noted that $RWE_{2:1}(m,n)$, may be obtained by exchanging $X^{(1)}$ and $X^{(2)}$ in the above procedure. Figures 2i and 2j show $RWE_{1:2}(m,n)$ and $RWE_{2:1}(m,n)$, respectively. Compared to the individual metrics shown in Fig. 2c–h, $RWE_{1:2}(m,n)$ and $RWE_{2:1}(m,n)$ show considerably improved well-exposedness. Similarly, the proposed AWE metric of image 1 at (m,n), denoted by $AWE_1(m,n)$, is calculated as follows:

1) Three metrics: *ZC*, *SML*, and $\hat{M}^{(3)}$, are calculated for image $X^{(1)}$.
2) For each metric, the maximum and minimum values are obtained for image $X^{(1)}$.
3) By using these maximum and minimum values, the measurement of each metric for image $X^{(1)}$ is normalized.
4) $AWE_1(m,n)$ is obtained by summing the results of three metrics.

B. PROPOSED AUTOMATIC EXPOSURE BRACKETING ALGORITHM

The goal of automatic exposure bracketing process is to select the optimal input images for fusion among a sequence of images obtained by increasing or decreasing the exposure step (ES). This process needs to determine how many input images and what values of ES settings are optimal for the best fusion result. When a larger number of input images are used, a better fusion result is expected but a higher computing cost is demanded. Therefore, the number of input images is chosen as the maximum number of image satisfying the limitation of the user's allowable fusion time on a dedicated hardware platform.

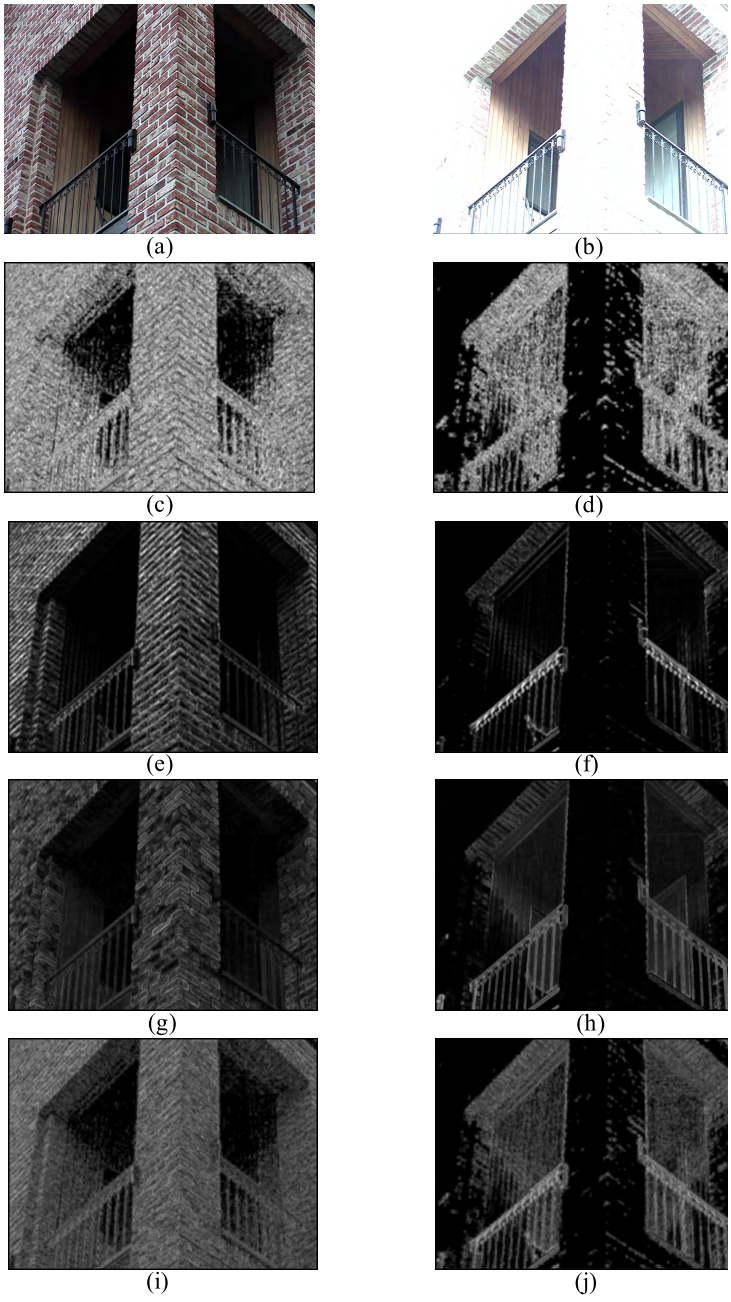

FIGURE 2. Relative measurements of proposed metrics: (a) sample image $X^{(1)} = \{x^{(1)}(m,n)\}$, (b) sample image $X^{(2)} = \{x^{(2)}(m,n)\}$, (c) $ZC_{X^{(1)}}(m,n)$, (d) $ZC_{X^{(2)}}(m,n)$, (e) $SML_{X^{(1)}}(m,n)$, (f) $SML_{X^{(2)}}(m,n)$, (g) $\hat{M}^{(3)}_{X^{(1)}}(m,n)$, (h) $\hat{M}^{(3)}_{X^{(2)}}(m,n)$, (i) $RWE_{1:2}(m,n)$, and (j) $RWE_{2:1}(m,n)$.

When the number of input images is fixed, the optimal ESs for capturing input images are determined on the basis of the AWE. The i^{th} exposed image is denoted $X^{(i)} = \{x^{(i)}(m,n)\}$, $i = 1, 2, \ldots, I$, and the well-exposedness change (WEC) of $X^{(i)}$ is calculated as follow:

$$WEC(i) = \sum_{(m,n)\in\Omega} [AWE_{(i-1)}(m,n)AWE_i(m,n)AWE_{(i+1)}(m,n)]. \quad (1)$$

Figure 3a shows a graph of WEC for a sample sequence of multi-exposed images. The images corresponding to several critical peaks are shown in Fig. 3b. It is shown that the WEC generates peaks reasonably.

Three types of standard search algorithms, namely global search (GS), binary search (BS), and rule-based search (RS) [18], [19], have been proposed for finding peaks in a graph such as WEC graphs. GS scans the entire range by eliminating any possibility of falsely obtaining a local peak but demands the highest computing cost. BS uses a divide

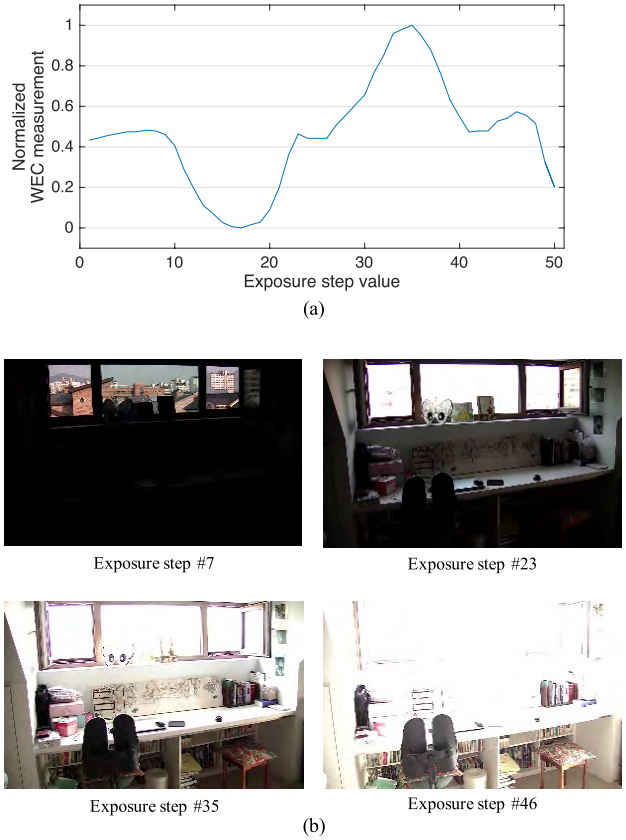

FIGURE 3. Measurements of WEC for a sequence of multi-exposed images: (a) measurements and (b) images corresponding to peaks.

and conquer algorithm that is faster than GS but requires frequent changes in the search direction to reach the local peak positions. RS scans the search range in a sequential manner in one direction by classifying the search range into four types, namely initial, fine, mid, and coarse, and changing the step size increments on the basis of the type of search range. RS is adopted in this paper for finding peaks because RS is known to reduce the number of iteration to about 40 % of the GS and be faster than BS for graphs similar to WEC graphs [19], [27].

Once all the local peaks are found, a predefined number of exposures are selected for fusion. The selection rule with notations is as follow:

Notations:

- k: Index for the k^{th} ES searched during the RS process
- p: Index for the p^{th} peak found during the RS process
- Num_p: Total number of peaks as the result of RS
- MAX_EV: Predefined maximum number of allowable selected ESs

Rule:

- Case 1 : $Num\_p \geq MAX\_EV$
  1) The global peak is chosen as the first selected ES.
  2) The next ranked ES is selected by the following score:

$$Score(p) = \frac{WEC(p)}{GHeight} + \frac{StepD(p)}{ERange}$$

where *Score(p), WEC(p), GHeight, StepD(p)*, and *ERange* are the score of the p^{th} peak, the height of the p^{th} peak, the height of the global peak, the ES distance from the p^{th} peak to the nearest pre-selected ES, and the total ES range, respectively.

  3) This selecting procedure continues until the number of selected ESs reaches MAX_EV.

- Case 2: $Num\_p < MAX\_EV$
  1) All the local peaks are chosen as the selected ES.
  2) The next ranked ES is selected by the following score:

$$Score(k) = \frac{WEC(k)}{GHeight} + \frac{StepD(k)}{ERange}$$

where *Score(k), WEC(k)*, and *StepD(k)* are the score of the k^{th} ES, the height of the k^{th} ES, and the ES distance from the k^{th} ES to the nearest pre-selected ES, respectively.

  3) This selecting procedure continues until the number of selected ESs reaches MAX_EV.

C. PROPOSED MULTI-EXPOSURE IMAGE FUSION RULE

Two types of standard fusion rules, namely weighted-sum-based rule and selection-based rule, have been adopted for image fusion [25]. However, previous research shows that rules have their own demerits. The former rule loses details and sharpness in homogeneous regions and transition regions between two differently exposed regions, respectively. The latter rule is highly sensitive to noise and results in discontinuities between two differently exposed regions [25], [28]. Furthermore, in case of fusing two differently exposed images with a high luminance difference, both rules typically show overall luminance discontinuities annoying human eyes between differently weighted or selected pixels.

To resolve these problems, this paper proposes a fusion rule which has the combined form of two standard rules with luminance compensation. The fusion is performed in the YCbCr color coordinate. Denoting the Y, Cb, and Cr component of an image $x(m,n)$ by $x_y(m,n)$, $x_{Cb}(m,n)$, and $x_{Cr}(m,n)$, respectively, and the resulting image of fusing two images $X^{(1)}$ and $X^{(2)}$ by $X^{(1\&2)} = \left\{x^{(1\&2)}(m,n)\right\}$, the following fusion rule is applied for compensating luminance discontinuities, as shown at the top of the next page, where λ denotes the threshold value that determines the balance between the selection-based and the weighted-sum-based fusion rules. It is noted that we use the weighted-sum-based fusion rule for all Y, Cb, and Cr components when the difference between $RWE_{1:2}(m,n)$ and $RWE_{2:1}(m,n)$ is small. We use the selection-based fusion rule for Cb and Cr components when the difference between $RWE_{1:2}(m,n)$ and $RWE_{2:1}(m,n)$ is large. In this case, we also use the

$$1)\ x_y^{(1\&2)}(m,n) = \frac{RWE_{1:2}(m,n)x_y^{(1)}(m,n) + RWE_{2:1}(m,n)x_y^{(2)}(m,n)}{RWE_{1:2}(m,n) + RWE_{2:1}(m,n)},$$

$$x_{Cb}^{(1\&2)}(m,n) = \frac{RWE_{1:2}(m,n)x_{Cb}^{(1)}(m,n) + RWE_{2:1}(m,n)x_{Cb}^{(2)}(m,n)}{RWE_{1:2}(m,n) + RWE_{2:1}(m,n)}, \text{ and}$$

$$x_{Cr}^{(1\&2)}(m,n) = \frac{RWE_{1:2}(m,n)x_{Cr}^{(1)}(m,n) + RWE_{2:1}(m,n)x_{Cr}^{(2)}(m,n)}{RWE_{1:2}(m,n) + RWE_{2:1}(m,n)},$$

if $RWE_{1:2}(m,n)/RWE_{2:1}(m,n) < \lambda$ and $RWE_{2:1}(m,n)/RWE_{1:2}(m,n) < \lambda$;

$$2)\ x_y^{(1\&2)}(m,n) = \frac{log(1+RWE_{1:2}(m,n))x_y^{(1)}(m,n) + log(1+RWE_{2:1}(m,n))x_y^{(2)}(m,n)}{log(1+RWE_{1:2}(m,n)) + log(1+RWE_{2:1}(m,n))},$$

if $RWE_{1:2}(m,n)/RWE_{2:1}(m,n) > \lambda$ or $RWE_{2:1}(m,n)/RWE_{1:2}(m,n) > \lambda$;

3) $x_{Cb}^{(1\&2)}(m,n) = x_{Cb}^{(1)}(m,n)$ and $x_{Cr}^{(1\&2)}(m,n) = x_{Cr}^{(1)}(m,n)$,
if $RWE_{1:2}(m,n)/RWE_{2:1}(m,n) > \lambda$;

4) $X_{Cb}^{(1\&2)}(m,n) = x_{Cb}^{(2)}(m,n)$ and $x_{Cr}^{(1\&2)}(m,n) = x_{Cr}^{(2)}(m,n)$,
if $RWE_{2:1}(m,n)/RWE_{1:2}(m,n) > \lambda$

logarithmic-scaled weighted-sum-based fusion rule for Y component to compensate luminance discontinuities.

When the number of multi-exposed images is greater than two, this fusion rule is applied recursively, as follows:

1) Given multi-exposed images $X^{(l)} = \{x^{(l)}(m,n)\}$, $l = 1, 2, \ldots, L$, set $s = 1$.
2) Using the above fusion rule, obtain $X^{(s\&(s+1))}$ and save the result as $X^{(s+1)}$.
3) Obtain $X^{(L-s+1)\&(L-s))}$ and save the result as $X^{(L-s)}$.
4) Repeat steps 2) and 3) by increasing s until all the input images are involved.
5) Declare the last resulting image as the final fusion result.

III. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

The performance of the proposed method was evaluated on a digital camera platform powered by a quad-core 2.1-GHz CPU and 3-GB RAM, which provides a software development kit. The camera module installed in this platform consists of 1/2.8-inch CMOS image sensor and a 20× zoom lens. For fast processing of exposure bracketing and previewing the fused image, video graphics array (VGA) resolution (640 × 480) images were used. For fusing full-size images, full high-definition (FHD) resolution (1920 × 1080) images were used.

The time for exposure bracketing is affected by the number of ESs selected during the RS process, the shutter speed corresponding to each selected ES, the lag time to change exposure from one ES to another ES, and the time for calculating WEC. The platform used in the experiment provides a shutter speed ranging from 0.3×10^{-4}s to 1s and an f-number from 1.6 to 25.6, to set the exposure time and the lens aperture, respectively. The maximum lag time of the used camera to change exposure is 50 ms. For this experiment, the f-number was set to 2.3 and the range of shutter speed was set from 0.5×10^{-1}s to 0.25×10^{-2}s. The computing time of Eq. (1) for the WEC used for bracketing was almost negligible as compared to the lag time. Therefore, the number of ESs to be searched within the search range was set on the basis of the lag time, shutter speed, average number of ESs skipped by the RS process, and user-defined allowable time for bracketing. The number of ESs to be searched was chosen to limit bracketing times to less than 1 s. In the platform used, this corresponded to 50 ESs as shown in Fig. 3a.

The time for fusing multi-exposed VGA and FHD images for previewing and saving, respectively, is dependent on the number of input images captured with exposure settings resulting from the bracketing procedure. The predefined number of selected peaks in the exposure-bracketing algorithm was chosen to be four in order to satisfy the requirement that the fusion time for VGA images and FHD images be less than 1 s and 4 s, respectively.

Image alignment by enhanced correlation coefficient (ECC) maximization [29] was performed following the exposure-bracketing procedure. The ECC algorithm was chosen because it is fast, well-known, and widely used in a wide range of applications, such as image registration, object tracking, super-resolution, and visual surveillance by moving cameras.

The size of block $B_{m,n}$ was set to 5 × 5. Larger sizes may give a better result, but at the expense of increased complexity. Experiments on various test image sequences showed that larger block sizes negligibly improved image quality. The threshold value λ for the fusion rule was experimentally chosen as 1.4 to achieve the best result on both objective and visual tests for multi-exposed sample image sequences used in the experiment.

Five sample image sequences were used to test the proposed method's performance. They include indoor and outdoor multi-exposed image sequences including at least two critical exposures. For presentation purpose, the VGA

FIGURE 4. Images selected for fusion from sample image sequences.

images selected by the proposed exposure bracketing algorithm are shown for each sequence in Fig. 4.

B. FUSION RESULTS

Based on a thorough review of previously proposed multi-exposure image fusion methods, two commercialized methods, namely "P" [30] and "H" [31], and Li and Kang's method [15] abbreviated as "LK" in this paper, were selected for performance comparison. "P" and "H" were selected because they are well-known offline software in the industrial field of multi-exposure image fusion. For the implementation of "P" and "H", the most recently released programs were used. "P" and "H" provide several options for adjusting the fused image quality on the basis of the user preferences. In this paper's experiment, default values were set for these options. "LK" was selected because it shows the best sets of objective results among the methods that are sufficiently fast to be implemented on state-of-the-art digital cameras. For the implementation of "LK", the authors' providing program was used. Unfortunately, "P", "H", and "LK" do not provide an exposure bracketing operation. Therefore, for comparisons of image quality, four images selected by the exposure bracketing algorithm proposed in this paper were used as the input images for the fusion of compared methods.

The mutual information-based quality metric MF^{XY} [32] proposed by Hossny *et al.* was used for the objective test because it is the metric that can be used for the case of fusing more than two input images and is the most widely used metric in the recently published papers. MF^{XY}, which was originally proposed for the case of fusing two input images, is extended for the case of fusing more than two input images, $X^{(l)}$, $l = 1, 2, \ldots, L$, as follow:

$$M_F^{X(L)} = \sum_{l=1}^{L} \frac{2I(F, X^{(l)})}{H(F) + H(X^{(l)})}, \tag{2}$$

where F denotes the fused image; $I(F, X^{(l)})$ represents the mutual information between F and $X^{(l)}$; $H(F)$ and $H(X^{(l)})$ indicates the entropy of F and $X^{(l)}$, respectively; and L denotes the number of input images. The values of $M_F^{X(L)}$ for both VGA and FHD images are listed in Table 1.

TABLE 1. $M_F^{X(L)}$ comparisons with other methods.

Resolution	Image Sequence	"P"	"H"	"LK"	Proposed
VGA	"P1"	1.1753	1.0267	0.8331	***2.0030***
	"P2"	0.9773	0.9855	0.8117	***1.6662***
	"P3"	1.3101	1.2564	1.2324	***2.2131***
	"P4"	0.9184	0.9767	0.8815	***1.9747***
	"P5"	1.1075	1.0813	0.8013	***1.9818***
Average		1.0977	1.0653	0.9120	***1.9678***
FHD	"P1"	1.1858	0.9863	0.7813	***1.9841***
	"P2"	1.0032	0.9407	0.7570	***1.6803***
	"P3"	1.3066	1.2240	1.2256	***2.2337***
	"P4"	0.9562	0.9677	0.8189	***2.0301***
	"P5"	1.1446	1.0260	0.7580	***1.9867***
Average		1.1192	1.0289	0.8682	***1.9830***

FIGURE 5. Fused VGA images using "P," "H," "LK," and the proposed method.

A larger value corresponds to better performance. Therefore, the proposed method can be considered superior based on this metric.

For a visual test of the global aspect of fused images, VGA images fused by using the compared methods are shown in Fig. 5. Each method shows a slightly different overall luminance depending on the sample image sequences. It is observed that the most distinct characteristic of the proposed method is colorfulness. This result may be attributed to the fact that all the other methods perform their fusions mainly on the basis of the luminance component only, whereas the proposed method considers colorfulness as one of the key features for its fusion.

Previous research shows that performance differences are typically clarified in textured regions, homogeneous regions, edge regions, and adjacent regions between underexposed and overexposed regions. Thus, for a visual test of the local aspect of fused images, portions of each fused FHD image including these types of regions were selected, enlarged, and shown in Fig. 6. Further, visual difference in textured regions or homogeneous regions located inside a foreground or a background is hardly distinguishable for the compared methods. However, the most distinguishing result is observed near the strong edge regions between a foreground and a background as indicated by arrows in Fig. 6. The proposed method alse shows the least distortion.

TABLE 2. Comparison of execution times (s).

Image Sequence	"P"	"H"	"LK"	Proposed
VGA	0.9686	2.8234	0.7826	0.6840
FHD	3.5320	5.6160	4.9933	3.6111

Besides the quality of fused images, the computing efficiency is an important factor for implementing a fusion algorithm. Table 2 shows the average execution times for both VGA and FHD image fusion on the digital camera platform used in the experiment. As shown in the table, the execution time of the proposed method was lower than that of other methods for VGA images. However, in the case of FHD images, "P" was faster than the proposed method. The execution time of the proposed method is approximately proportional to the size of the input images. Thus, the computing efficiency of "P" increases to be higher than that of the proposed method as the size of input images increases.

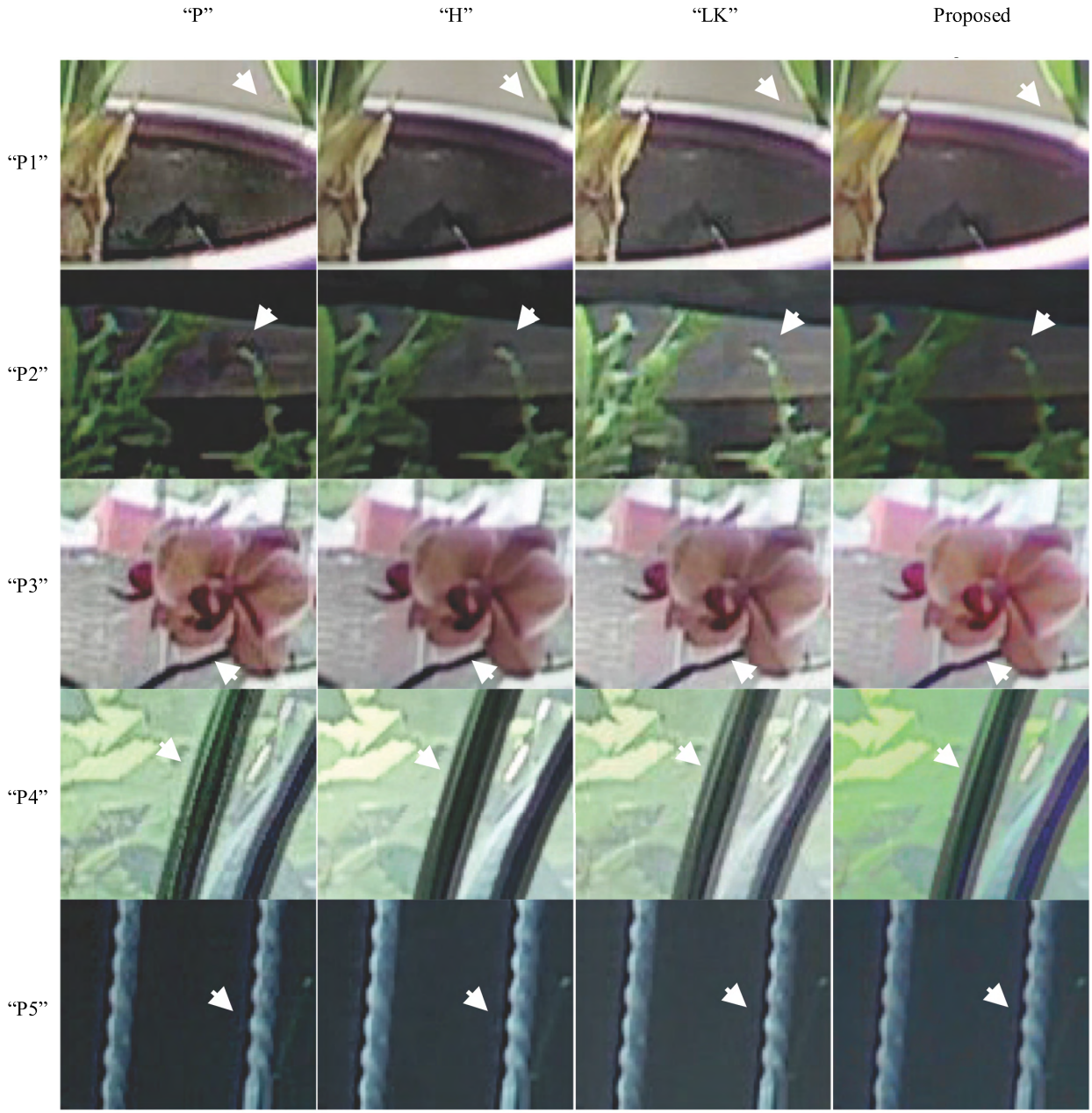

FIGURE 6. Enlarged parts of fused FHD images using "P," "H," "LK," and the proposed method.

IV. CONCLUSION

In this paper, we proposed a fast method for fusing multi-exposed images that can be implemented in digital cameras and smartphones. Two new well-exposedness metrics, namely RWE and AWE, were proposed. The exposure bracketing algorithm using AWE was implemented for VGA-size image sequences. Further, multi-exposure image fusion method using RWE was proposed. The proposed fusion method outperformed previously published methods for constructing VGA and FHD HDR images for several critical image sequences in both visual and objective tests of quality.

The proposed method works for still images. Future research will be focused on extending this method for multi-exposure video fusion.

REFERENCES

[1] A. Adams *et al.*, "The frankencamera: An experimental platform for computational photography," *ACM Trans. Graph.*, vol. 29, no. 4, p. 29, 2010.

[2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH Classes*. 2008, p. 31.

[3] F. Shen, Y. Zhao, X. Jiang, and M. Suwa, "Recovering high dynamic range by multi-exposure retinex," *J. Vis. Commun. Image Represent.*, vol. 20, no. 8, pp. 521–531, 2009.

[4] M. A. Robertson, S. Borman, and R. L. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," *J. Electron. Imag.*, vol. 12, no. 2, pp. 219–228, 2003.
[5] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," in *Proc. 4th Int. Conf. Comput. Vis.*, May 1993, pp. 173–182.
[6] T. Jinno and M. Okuda, "Multiple exposure fusion for high dynamic range image acquisition," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 358–365, Jan. 2012.
[7] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, 2009.
[8] A. A. Goshtasby, "Fusion of multi-exposure images," *Image Vis. Comput.*, vol. 23, no. 6, pp. 611–618, 2005.
[9] S. Raman and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proc. Eurograph. (Short Papers)*, 2009, pp. 1–4.
[10] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3634–3646, Dec. 2011.
[11] B. Gu, W. Li, J. Wong, M. Zhu, and M. Wang, "Gradient field multi-exposure images fusion for high dynamic range image visualization," *J. Vis. Commun. Image Represent.*, vol. 23, no. 4, pp. 604–610, 2012.
[12] W.-C. Kao, C.-C. Hsu, L.-Y. Chen, C.-C. Kao, and S.-H. Chen, "Integrating image fusion and motion stabilization for capturing still images in high dynamic range scenes," *IEEE Trans. Consum. Electron.*, vol. 52, no. 3, pp. 735–741, Aug. 2006.
[13] K. Jacobs, C. Loscos, and G. Ward, "Automatic high-dynamic range image generation for dynamic scenes," *IEEE Comput. Graph. Appl.*, vol. 28, no. 2, pp. 84–93, Mar./Apr. 2008.
[14] W. Zhang and W.-K. Cham, "Gradient-directed multiexposure composition," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2318–2323, Apr. 2012.
[15] S. Li and X. Kang, "Fast multi-exposure image fusion with median filter and recursive filter," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 626–632, May 2012.
[16] R. Soundararajan and A. C. Bovik, "Survey of information theory in visual quality assessment," *Signal, Image, Video Process.*, vol. 7, no. 3, pp. 391–401, May 2013.
[17] E. Cohen and Y. Yitzhaky, "No-reference assessment of blur and noise impacts on image quality," *Signal, Image Video Process.*, vol. 4, no. 3, pp. 289–302, 2010.
[18] V. Peddigari, M. Gamadia, and N. Kehtarnavaz, "Real-time implementation issues in passive automatic focusing for digital still cameras," *J. Imag. Sci. Technol.*, vol. 49, no. 2, pp. 114–123, 2005.
[19] N. Kehtarnavaz and H.-J. Oh, "Development and real-time implementation of a rule-based auto-focus algorithm," *Real-Time Imag.*, vol. 9, no. 3, pp. 197–203, Jun. 2003.
[20] J. Zhang and T. M. Le, "A new no-reference quality metric for jpeg2000 images," *IEEE Trans. Consum. Electron.*, vol. 56, no. 2, pp. 743–750, May 2010.
[21] K. Panetta, C. Gao, and S. Agaian, "No reference color image contrast and quality measures," *IEEE Trans. Consum. Electron.*, vol. 59, no. 3, pp. 643–651, Aug. 2013.
[22] D. Hasler and S. E. Suesstrunk, "Measuring colorfulness in natural images," in *Electronic Imaging* (International Soc. Opt. Photon.) 2003, pp. 87–95.
[23] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013.
[24] M. Narwaria, M. P. Da Silva, P. Le Callet, and R. Pepion, "Tone mapping based hdr compression: Does it affect visual experience?" *Signal Process. Image Commun.*, vol. 29, no. 2, pp. 257–273, 2014.
[25] P. Shah, T. Srikanth, S. N. Merchant, and U. B. Desai, "Multimodal image/video fusion rule using generalized pixel significance based on statistical properties of the neighborhood," *Signal, Image Video Process.*, vol. 8, no. 4, pp. 723–738, 2014.
[26] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 824–831, Aug. 1994.
[27] O.-J. Kwon, S. Choi, D. Jang, and H.-S. Pang, "All-in-focus imaging using average filter-based relative focus measure," *Digital Signal Process.*, vol. 60, pp. 200–210, Jan. 2017.
[28] S. Li, X. Kang, J. Hu, and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes," *Inf. Fusion*, vol. 14, no. 2, pp. 147–162, 2013.
[29] G. D. Evangelidis and E. Z. Psarakis, "Parametric image alignment using enhanced correlation coefficient maximization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1858–1865, Oct. 2008.
[30] HDRsoft, (2017). [Online]. Available: http://www.hdrsoft.com
[31] easyHDR, (2017). [Online]. Available: http://www.easyhdr.com
[32] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on 'information measure for performance of image fusion,'" *Electron. Lett.*, vol. 44, no. 18, pp. 1066–1067, Aug. 2008.

SEUNGCHEOL CHOI received the B.S. and M.S. degrees in computer science from Sejong University, Seoul, South Korea, in 1998 and 2001, respectively, where he is currently pursuing the Ph.D. degree in electrical engineering. From 2001 to 2013, he was a Researcher with Galaxia Communications. His research interests include image and video coding, high-dynamic range imaging, image processing, image fusion, and JPEG.

OH-JIN KWON received the M.S. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 1991, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 1994. From 1984 to 1989, he was a Researcher with the Agency for Defense Development of Korea. From 1995 to 1999, he was the Head of the Media Lab, Samsung SDS Co., Ltd., Seoul, South Korea. Since 1999, he has been a Faculty Member with Sejong University, Seoul, South Korea, where he is currently an Associate Professor. His research interests include image and video fusion, coding, watermarking, analyzing, and processing.

JINHEE LEE is currently pursuing the M.S. degree in electrical engineering from Sejong University, Seoul, South Korea. Her research interests include image processing and fusion.