

# Detail-Enhanced Multi-Scale Exposure Fusion

Zhengguo Li, *Senior Member IEEE*, Zhe Wei, Changyun Wen, *Fellow IEEE*, and Jinghong Zheng, *Member IEEE*

**Abstract**—Multi-scale exposure fusion is an effective image enhancement technique for a high dynamic range (HDR) scene. In this paper, a new multi-scale exposure fusion algorithm is proposed to merge differently exposed low dynamic range (LDR) images by using the weighted guided image filter (WGIF) to smooth the Gaussian pyramids of weight maps for all the LDR images. Details in the brightest and darkest regions of the HDR scene are preserved better by the proposed algorithm without relative brightness change in the fused image. In addition, a new weighted structure tensor is introduced to the differently exposed images and it is adopted to design a detail extraction component for the proposed fusion algorithm such that users are allowed to manipulate fine details in the enhanced image according to their preference. The proposed multi-scale exposure fusion algorithm is also applied to design a simple single image brightening algorithm for both low-light imaging and back-light imaging.

**Index Terms**—Exposure fusion, Differently exposed images, Weighted guided image filter, Low-light imaging, Weighted structure tensor,

## I. INTRODUCTION

Merging multiple differently exposed low dynamic range (LDR) images into a high dynamic range (HDR) image is an efficient way to overcome the limited dynamic ranges of cameras and to reduce noise in photographs [1]. Such an imaging technique is called HDR imaging. Due to possible camera movement and moving objects, all the LDR images are first aligned [2], [3] and all the moving objects are synchronized according to a pre-defined reference image [4]–[7]. An HDR image is then synthesized from the corrected images to include details of all the LDR images. The HDR image is finally converted into an LDR image by using tone mapping algorithms [8]–[11] so as to visualize the HDR scene by a conventional display device.

Besides the HDR imaging, another popular technique is called exposure fusion. A more informative and perceptually appealing LDR image is directly synthesized from all the LDR images without the generation of an intermediate HDR image as in the HDR imaging. Many exposure fusion algorithms were proposed, ranging from simple weighted averaging to sophisticated methods based on advanced statistical image models. Three quality measures of proper exposure, good contrast, and high saturation were used in [12] to determine how much a given pixel will contribute to the final synthesized image. All the LDR images were scaled into several down-sampled layers by using the Laplacian pyramid [13].

Z. Li and J. Zheng are with signal processing department, Institute for Infocomm Research, 1 Fusionopolis Way, 21-01 Connexis, Singapore 138632. ({ezgli,jzheng}@i2r.a-star.edu.sg). Z. Wei and C. Wen are with school of Electronic and Electric Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798. (WEIZ0008@e.ntu.edu.sg, ecy-wen@ntu.edu.sg).

The multi-resolution blending preserves the global contrast and ensures that transitions between regions where different images contribute more heavily are difficult to perceive. The exposure fusion scheme in [14] was based on an observation that gradient magnitude becomes larger when a pixel gets a state of better exposed and it decreases gradually as the pixel approaches over/under-exposure. Same as the algorithm in [12], all the LDR images were blended at multiple scales using the pyramidal image decomposition. A guided image filter (GIF) based exposure fusion algorithm was proposed in [15]. An average filter was first utilized to get the two-scale representation. The base and detail layers are fused through using a guided filtering based weighted average method. Two single-scale exposure fusion algorithms were proposed in [16], [17]. Different optimization problems are formulated in [16], [17] with optimal variables as pixel values of the final synthesized LDR image or weights of the LDR images. The final synthesized image can be obtained by solving these optimization problems. The exposure fusion neither requires lighting conditions of all the LDR images to be the same nor requires knowledge of exposure times as required by the HDR imaging. However, there is no fine detail extraction component in the above exposure fusion algorithms while fine details can be manipulated by the existing tone mapping algorithms [8]–[10]. Based on such an observation, a detail extraction component was proposed in [18] to enhance the existing exposure fusion algorithms. The component is formulated as a quadratic optimization problem and an iterative method was provided in [18] to solve the optimization problem. Since exposure fusion is an active research area, many new exposure fusion algorithms were proposed in [20]–[24] and their cons and pros of these algorithms were provided in [19], [25].

Recently, an interesting subjective user study was conducted in [19] to evaluate the quality of images generated by the above exposure fusion algorithms. It was found that no single state-of-the-art exposure fusion algorithm produces the best quality for all test images. The algorithm in [12] achieves the best performance on average, and the algorithm in [18] is the second best on average. Unfortunately, neither the algorithm in [12] nor the algorithm in [18] can always preserve details and color in the brightest and/or darkest regions of an HDR scene well [19]. Even though this problem is overcome by the algorithm in [26], the relative brightness among regions is changed if certain regions are not well-exposed in any of the captured LDR images. Thus, it is desired to design a new exposure fusion algorithm so as to preserve details and color in the brightest and/or darkest regions of the HDR scene well. It was also shown in [19] that the detail extraction component in [18] is useful to create perceptually appealing results on some images, but also creates unwanted artifacts in some other images. The overall performance gain is not always

guaranteed by the detail extraction component in [18]. One more issue is the speed of the detail extraction component in [18]. The quadratic optimization problem is solved via an iterative method in [18] and its computational cost is high, especially for mobile devices with limited computational resources. Therefore, it is also desired to design a new fast detail extraction component for an exposure fusion algorithm such that users can manipulate the fine details according to their preference.

In this paper, a new multi-scale exposure fusion algorithm is first introduced to merge all the LDR images by using the weighted GIF (WGIF) [23] to smooth the Gaussian pyramids of weighted maps for all the LDR images. The Gaussian pyramids of the luminance components serve as the guidance images. Fine details in the brightest and darkest regions of an HDR scene are preserved better by the proposed algorithm than the algorithm in [12] and the relative brightness among regions is preserved better by the proposed algorithm than the single-scale exposure fusion algorithm in [26]. However, the proposed fusion algorithm does not provide freedom for users to manipulate fine details according to their preference. Furthermore, fine details could be lost by the multi-resolution blending [18]. Thus, a properly designed detail extraction component is desired and useful for the proposed exposure fusion algorithm [19]. Based on this observation, a new detail extraction component is then proposed by extending the powerful Di Zenzo structure tensor [27] to the differently exposed images. Same as the detail extraction component in [18], the proposed one is also formulated as a quadratic optimization problem with the input as a vector field. It was pointed out in [18] that the vector field plays a crucial role in the detail extraction component. Intuitively, the set of the LDR images is a natural tensor, and the structure tensors in [28], [29] could be applied to generate the vector field. Unfortunately, noise will be amplified in the final image if the structure tensors in [28], [29] are directly adopted. This is because that differently exposed images include different level of noise. A weighted structure tensor is proposed to address the problem and the vector field is constructed using the weighted structure tensor. Fine details are extracted from the vector field through using a fast separate approach [30], [31] to solve the quadratic optimization problem. The extracted details are added to an intermediate image which is fused by the multi-scale exposure fusion algorithm to produce the final image. Experimental results show that only fine details are extracted by the new component while large scale components could be extracted by the one in [18]. As a result, the quality of the enhanced image is improved.

Due to small sensors and low-cost optics, cameras embedded in mobile devices have a reduced capacity to capture light. Setting longer exposure times only has a potential to help relatively static scenes. A possible solution is to capture multiple differently exposed images and to fuse them together to produce an image. However, both camera movement and moving objects are issues for the fusion of differently exposed images [7]. Here, a new method is introduced to capture images in low-lighting and back-lighting conditions, i.e., an image is

captured via a pair of a low ISO value and a short exposure time [32]. This is different from a conventional method which captures an image using a pair of a large ISO value and a low exposure time or a pair of a small ISO value and a long exposure time. There is negligible motion blur in the captured image due to the short exposure time. Furthermore, there are negligible or even no saturated pixels in the captured image because the product of the ISO value and the exposure time is small. On the other hand, the captured image is darker than an image that is captured via the conventional method. The proposed multi-scale exposure fusion algorithm is applied to design a simple algorithm for the brightening of such a dark image. Two virtual differently exposed LDR images are produced from the captured image and all the three LDR images are merged together using the proposed multi-scale exposure fusion algorithm to brighten the captured image. Experimental results demonstrated that the proposed algorithm can be applied to brighten dark objects in images that are captured in both low-lighting and back-lighting conditions. Overall, four major contributions of this paper are 1) a multi-scale exposure fusion algorithm using the WGIF. Fine details in brightest and darkest regions are preserved better using the new fusion algorithm; 2) a new concept of weighted structure tensor for the differently exposed images. Amplification of noise could be an issue for the existing structure tensors in [28], [29] while it is avoided by the proposed one; 3) a faster detail extraction component for the differently exposed images. Only fine details are extracted by the new method while large scale components could be extracted by the method in [18]. The speed is increased by more than four times; and 4) a simple single image brightening algorithm. The algorithm is useful for back-lit imaging and low-lighting imaging.

The rest of paper is organized as follows. A multi-scale exposure fusion algorithm is proposed in Section II to merge all the LDR images, and a new detail extraction component is provided to enhance the proposed algorithm. The proposed algorithm is applied to design a single image brightening algorithm in Section III. Finally, concluding remarks are given in Section IV.

## II. DETAIL ENHANCED FUSION OF DIFFERENTLY EXPOSED LDR IMAGES

In this section, a new multi-scale exposure fusion algorithm is first proposed using the WGIF in [23]. A fast detail extraction algorithm is then proposed to enhance the proposed multi-scale exposure fusion algorithm. For simplicity, the 8-bit images are focused in this section and the algorithm can be easily extended to other types of input images.

### A. A WGIF Based Multi-Scale Exposure Fusion

Let  $Z_i (1 \leq i \leq N)$  be a set of differently exposed LDR images with  $N$  be the number of the LDR images.  $Y_i$  is the luminance component of the image  $Z_i$ . Let  $p$  be a pixel position. There are three quality measures in [12],  $C_i(p)$ ,

$S_i(p)$ , and  $E_i(p)$  measure contrast, color saturation, and well-exposedness of pixel  $Z_i(p)$ , respectively.  $C_i(p)$  is obtained by applying a Laplacian filter to the gray-scale version of each image.  $S_i(p)$  is computed as the standard deviation within the R, G and B channel.  $E_i(p)$  is yielded by applying a Gauss curve to each channel separately and multiplying the results. Their product is denoted as  $\tilde{W}_i(p)$ . The weight map is then constructed as

$$W_i(p) = \frac{\tilde{W}_i(p)}{\sum_{j=1}^N \tilde{W}_j(p)}. \quad (1)$$

$G\{Y_i\}^{(l)}$  and  $G\{W_i\}^{(l)}$  are the Gaussian pyramids of luminance component  $Y_i$  and weight map  $W_i$ , respectively. The total number of layers is  $\kappa$  which is defined as

$$\kappa = \lfloor \log_2 \min(w, h) \rfloor - 2, \quad (2)$$

where  $\lfloor a \rfloor$  is the largest integer which is not greater than  $a$ .  $w$  and  $h$  are the width and height of the input image. Instead of using the Gaussian pyramid  $G\{W_i\}^{(l)}$  to fuse the differently exposed images, a WGIF based pyramid is introduced to fuse them.  $G\{W_i\}^{(l)}$  is the pyramid to be smoothed and  $G\{Y_i\}^{(l)}$  is the guidance pyramid. The proposed pyramid is based on an observation, the WGIF can transfer the structure of  $G\{Y_i\}^{(l)}$  to  $G\{W_i\}^{(l)}$ . Let the coefficients of the WGIF be denoted as  $\{\bar{a}_i\}^{(l)}$  and  $\{\bar{b}_i\}^{(l)}$ . To reduce the computational cost, only  $\{\bar{a}_i\}^{(\kappa)}$  and  $\{\bar{b}_i\}^{(\kappa)}$  as well as  $\{\bar{a}_i\}^{(4)}$  and  $\{\bar{b}_i\}^{(4)}$  are computed using the method in [23] with the radius  $\zeta$  and regularization parameter  $\lambda$  being fixed at 4 and 1/1024, respectively.  $\{\bar{a}_i\}^{(l)}$  and  $\{\bar{b}_i\}^{(l)}$  are interpolated from  $\{\bar{a}_i\}^{(\kappa)}$  and  $\{\bar{b}_i\}^{(\kappa)}$  if the value of  $l$  is larger than 4 and otherwise from  $\{\bar{a}_i\}^{(4)}$  and  $\{\bar{b}_i\}^{(4)}$ . The WGIF based pyramid is then given as

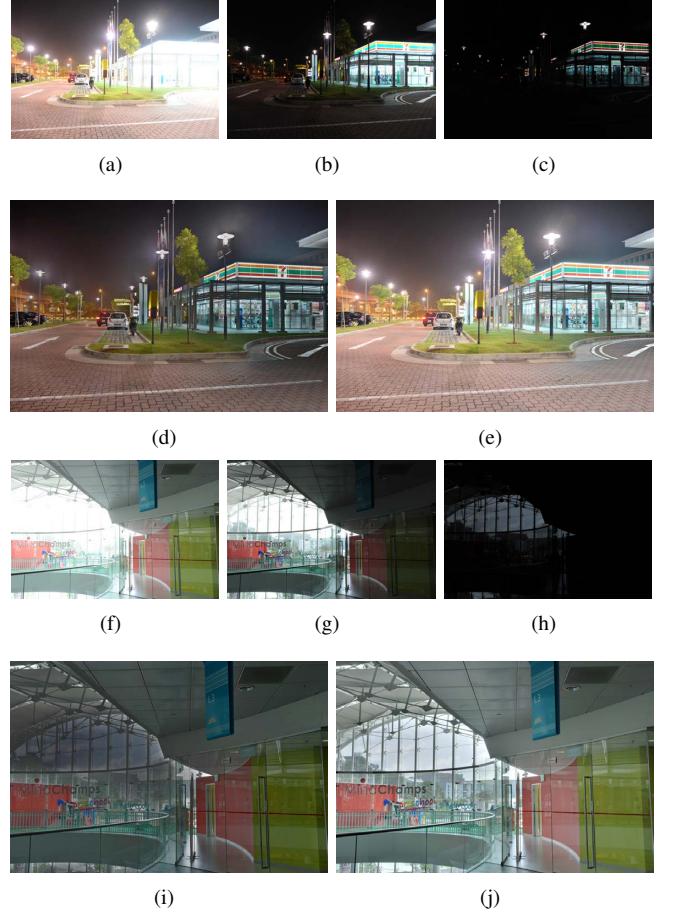
$$\{\hat{W}_i\}^{(l)} = \{\bar{a}_i\}^{(l)}\{Y_i\}^{(l)} + \{\bar{b}_i\}^{(l)}; 1 \leq l \leq \kappa. \quad (3)$$

Let  $L\{Z_{i,j}\}^{(l)}$  be the Laplacian pyramid of component  $Z_{i,j}$ . All the images  $Z_i (1 \leq i \leq N)$  at the different pyramid levels are blended as

$$L\{Z_j^{int}\}^{(l)} = \sum_{i=1}^N [L\{Z_{i,j}\}^{(l)} \{\hat{W}_i\}^{(l)}]; j \in \{r, g, b\}, \quad (4)$$

and the Laplacian pyramid  $L\{Z_j^{int}\}^{(l)}$  is collapsed to produce the intermediate image  $Z_j^{int}$ .

The proposed WGIF based multi-scale exposure fusion algorithm is compared with the single-scale exposure fusion algorithm in [26] which also preserves fine details in both the brightest and darkest regions of HDR scenes well. They are used to fuse two sets of differently exposed images. It is demonstrated in Fig. 1 that the relative brightness is changed by the algorithm in [26] if certain regions of an HDR scene are not well exposed in any of the LDR images. The problem is overcome by the proposed algorithm. However, the algorithm in [26] is simpler than the proposed exposure fusion algorithm. The complexity of the proposed exposure fusion algorithm is  $O(NM)$  for  $N$  differently exposed images with  $M$  pixels.



**Fig. 1:** Comparison of the proposed WGIF based multi-scale exposure fusion algorithm with the single-scale exposure fusion algorithm in [26]. (a)-(c) and (f)-(h) input images; (d,i) fused images by the algorithm in [26]; (e,j) fused images by the proposed algorithm. The relative brightness is changed by the algorithm in [26] while the relative brightness is preserved by the proposed algorithm.

### B. Detail-Enhanced Exposure Fusion

It was shown in [19] that a detail extraction component can be applied to improve the quality of images fused by an exposure fusion algorithm. The component also provides users freedom to manipulate fine details according to their preference [18]. A new detail extraction component is proposed in this subsection by introducing a new weighted structure tensor to the differently exposed images.

The set of the LDR images  $Z_i (1 \leq i \leq N)$  is a natural tensor. Intuitively, the structure tensors in [28], [29] can be applied to construct the vector field which contains the fine details of multiple differently exposed images. However, as shown in Fig. 2, noise could also be amplified if the structure tensors in [28], [29] are directly applied to the LDR images  $Z_i (1 \leq i \leq N)$ . To address the problem, a weighted structure tensor is introduced, and the new concept is based on an observation that the differently exposed images include different level of noise [1]. The luminance components of the

LDR images  $Z_i (1 \leq i \leq N)$  and the intermediate image  $Z^{int}$  in log domain are computed as

$$\hat{Y}_i(p) = \log(Y_i(p) + 1); \quad \hat{Y}^{int}(p) = \log(Y^{int}(p) + 1). \quad (5)$$

Define two weights  $\hat{W}_{x,i}(p)$  and  $\hat{W}_{y,i}(p)$  as

$$\hat{W}_{x,i}(p) = \frac{f(Y_i(p))f(Y_i(p_r))}{\max_j\{f(Y_j(p))f(Y_j(p_r))\}}, \quad (6)$$

$$\hat{W}_{y,i}(p) = \frac{f(Y_i(p))f(Y_i(p_b))}{\max_j\{f(Y_j(p))f(Y_j(p_b))\}}, \quad (7)$$

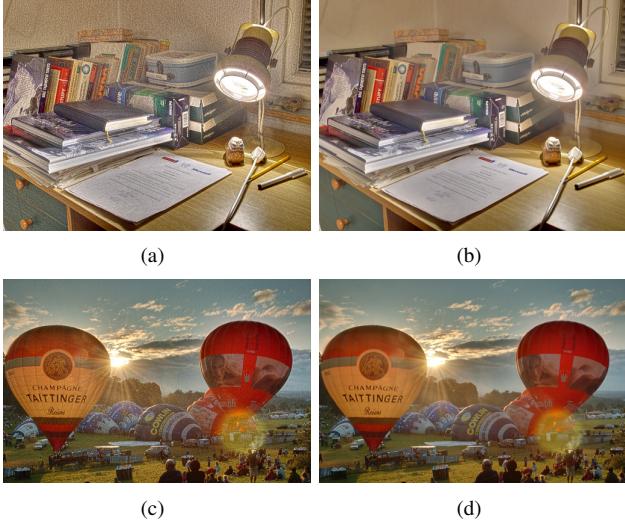
where  $p_r$  and  $p_b$  are the right and bottom pixels of the pixel  $p$ , and  $f(z)$  is defined as

$$f(z) = \begin{cases} z + 1; & \text{if } z \leq 128 \\ 257 - z; & \text{otherwise} \end{cases}. \quad (8)$$

The weighted gradients of all the LDR images  $Z_i (1 \leq i \leq N)$  at the pixel  $p$  are then given by the following matrix:

$$\nabla \hat{Y}(p) = \begin{bmatrix} \hat{W}_{x,1}(p) \frac{\partial \hat{Y}_1(p)}{\partial x} & \hat{W}_{y,1}(p) \frac{\partial \hat{Y}_1(p)}{\partial y} \\ \vdots & \vdots \\ \hat{W}_{x,N}(p) \frac{\partial \hat{Y}_N(p)}{\partial x} & \hat{W}_{y,N}(p) \frac{\partial \hat{Y}_N(p)}{\partial y} \end{bmatrix}. \quad (9)$$

Clearly, a large weight is assigned to the gradient  $\frac{\partial \hat{Y}_i(p)}{\partial x}$  ( $\frac{\partial \hat{Y}_i(p)}{\partial y}$ ) if the two pixels  $Y_i(p)$  and  $Y_i(p_r)$  ( $Y_i(p)$  and  $Y_i(p_b)$ ) are well exposed. A small weight is assigned otherwise. As such, the matrix  $\nabla \hat{Y}(p)$  includes less noise. The matrix

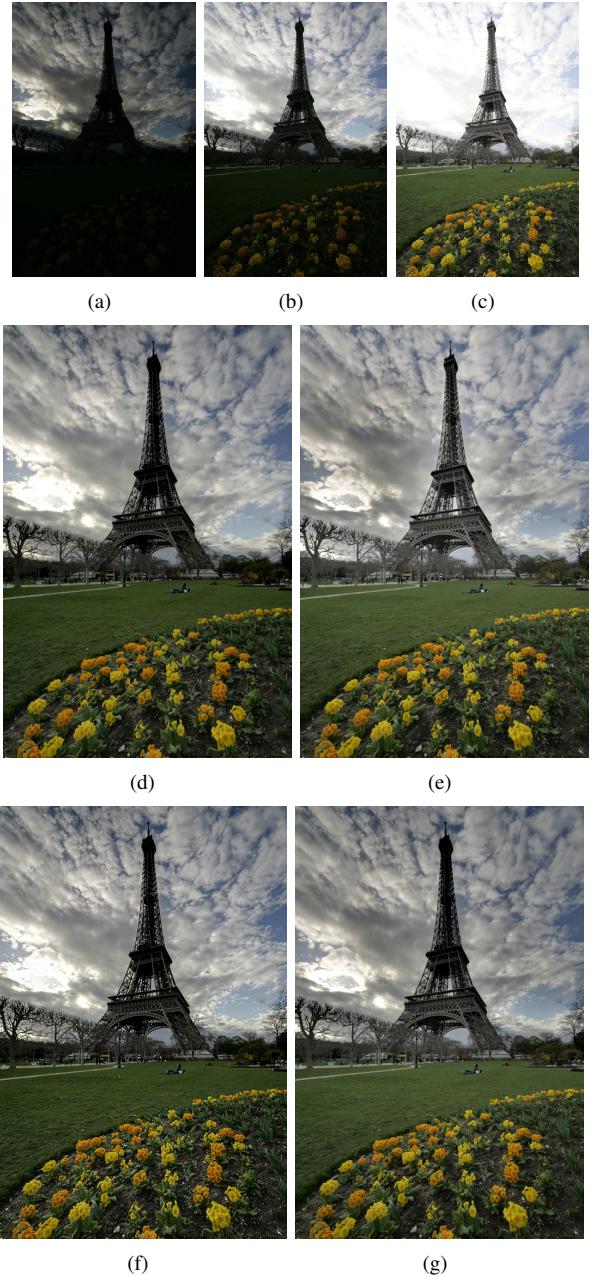


**Fig. 2:** Comparison of different structure tensors. (a) and (c) enhanced images by the structure tensor in [28]; (b) and (d) enhanced images by the proposed one. Noise is amplified by the structure tensor in [28]. Image courtesy of Martin Cadik and Erik Reinhard.

$\nabla \hat{Y}(p)$  is decomposed via the singular value decomposition as

$$\nabla \hat{Y}(p) = U_h(p) \Lambda_h(p) V_h^T(p), \quad (10)$$

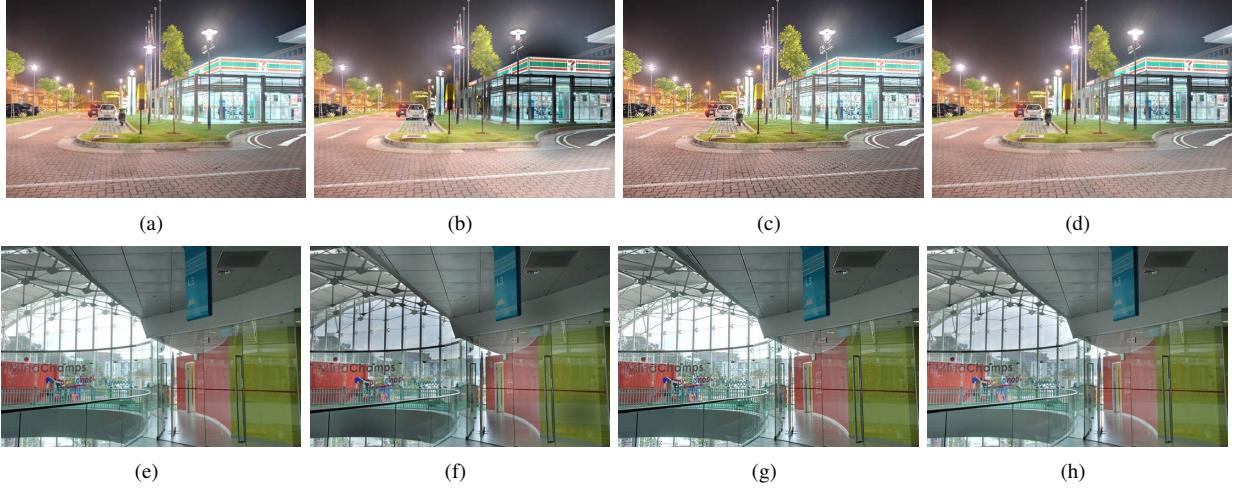
where  $U_h(p)$  is an  $N \times 2$  matrix,  $V_h(p) = [V_{h,1}(p), V_{h,2}(p)]$  is a  $2 \times 2$  matrix, and the diagonal matrix  $\Lambda_h =$



**Fig. 4:** Comparison of the proposed detail-enhanced exposure fusion algorithm with the exposure fusion algorithms in [12], [15], [18]. (a)-(c): input images; (d)-(g) images fused by the algorithms in [12], [15], [18], and the proposed one, respectively. Image courtesy of Jacques Joffre.

$\text{diag}\{\lambda_1(p), \lambda_2(p)\}$  contains the magnitude information for each of the primary directions.

The first fundamental form  $2 \times 2$  matrix  $\Phi_{\hat{Y}} = (\nabla \hat{Y})^T \nabla \hat{Y}$  was introduced in the image processing literature by Di Zenzo [27] as the structure tensor. With the weighting functions in the Equations (6)-(8), the matrix  $\Phi_{\hat{Y}}$  is a new weighted structure tensor. The eigenvalues of the structure tensor provide a rich and discriminative description of the local geometry of the image.  $\lambda_1^2(p)$  and  $\lambda_2^2(p)$  are the large and small eigenvalues



**Fig. 3:** Comparison of the proposed detail-enhanced exposure fusion algorithm with the exposure fusion algorithms in [12], [15], [18] by testing the two set of differently exposed images in Fig. 1. (a) and (e) results by the algorithm in [12]; (b) and (f) results by the algorithm in [15]; (c) and (g) results by the algorithm in [18]; (d) and (h) results by the proposed algorithm.

of the structure tensor  $\Phi_{\hat{Y}}$  at the pixel  $p$ , respectively. When both of them are small, the pixel  $p$  is located in a homogeneous region. When  $\lambda_1^2(p)$  is large while  $\lambda_2^2(p)$  is small, the pixel  $p$  is located close to an image edge. When both of them are large, the pixel  $p$  is close to an image corner. The vector field in [28] is constructed as  $\lambda_1(p)V_{h,1}^T(p)$ . It is shown in Fig. 2 that amplified noise becomes visible in homogeneous regions of the enhanced image. To overcome the problem, the desired vector field  $V$  is computed as

$$V(p) = \frac{\Gamma_1(p)\lambda_1(p)V_{h,1}^T(p) + \Gamma_2(p)\lambda_2(p)V_{h,2}^T(p)}{\sqrt{(\frac{\partial \hat{Y}^{int}(p)}{\partial x})^2 + (\frac{\partial \hat{Y}^{int}(p)}{\partial y})^2}}, \quad (11)$$

where  $\Gamma_1(p)$  and  $\Gamma_2(p)$  are defined as

$$\Gamma_1(p) = \begin{cases} \frac{\partial \hat{Y}^{int}(p)}{\partial x}; & \text{if } |\frac{\partial \hat{Y}^{int}(p)}{\partial x}| > |\frac{\partial \hat{Y}^{int}(p)}{\partial y}| \\ \frac{\partial \hat{Y}^{int}(p)}{\partial y}; & \text{otherwise} \end{cases}, \quad (12)$$

$$\Gamma_2(p) = \begin{cases} \frac{\partial \hat{Y}^{int}(p)}{\partial y}; & \text{if } |\frac{\partial \hat{Y}^{int}(p)}{\partial x}| > |\frac{\partial \hat{Y}^{int}(p)}{\partial y}| \\ \frac{\partial \hat{Y}^{int}(p)}{\partial x}; & \text{otherwise} \end{cases}. \quad (13)$$

The proposed vector field  $V$  in Equation (11) is a weighted average of  $\lambda_1(p)V_{h,1}^T(p)$  and  $\lambda_2(p)V_{h,2}^T(p)$  with the weight being determined by the gradient field of the image  $\hat{Y}^{int}$ . The proposed vector field is different from the vector field in [28] for pixels in homogeneous regions and close to image corners. Since the eigenspace is a line without a preassigned orientation, the direction of the vector field  $V$  is adjusted as follows:  $V_1(p)$  is multiplied by the sign of the product  $V_1(p)\frac{\partial \hat{Y}^{int}(p)}{\partial x}$ . Similarly,  $V_2(p)$  is multiplied by the sign of the product  $V_2(p)\frac{\partial \hat{Y}^{int}(p)}{\partial y}$ . The objective is to guarantee that the intermediate image  $Z^{int}$  is enhanced by the fine details in all the LDR images  $Z_i (1 \leq i \leq N)$ .

Same as the algorithm in [18], a quadratic optimization problem is formulated on the extraction of fine details from

the vector field  $(V_1, V_2)$  as follows

$$\min_{L_d} \left[ \|L_d\|_2^2 + \lambda \left( \left\| \frac{V_1 - \frac{\partial L_d}{\partial \mathbf{x}}}{\psi(V_1)} \right\|_2^2 + \left\| \frac{V_2 - \frac{\partial L_d}{\partial \mathbf{y}}}{\psi(V_2)} \right\|_2^2 \right) \right], \quad (14)$$

where  $\|\cdot\|_2$  is the  $l_2$  norm.  $L_d$  represents fine details to be extracted. The function  $\psi(z)$  is selected as [9]

$$\psi(z) = \sqrt{|z|^\gamma + \epsilon}, \quad (15)$$

the default values of  $\gamma$ ,  $\epsilon$ , and  $\lambda$  are 0.75,  $10^{-4}$  and 0.03125, respectively.

The solution to the optimization problem (14) is obtained by solving the following linear equation:

$$(I + \lambda(D_x^T A(V_1)D_x + D_y^T A(V_2)D_y))L_d = \lambda(D_x^T A(V_1)V_1 + D_y^T A(V_2)V_2) \quad (16)$$

where  $I$  is the identity matrix,  $D_x$  and  $D_y$  are discrete differentiation operators, and  $A(V_1)$  and  $A(V_2)$  are  $diag(\frac{1}{\psi^2(V_1(p))})$  and  $diag(\frac{1}{\psi^2(V_2(p))})$ , respectively.

Instead of using an iterative method to solve the above linear equation as in [18], a fast method is provided to solve the above linear equation and the fast method is inspired by the separate methods in [30], [31]. The matrix  $(I + \lambda(D_x^T A(V_1)D_x + D_y^T A(V_2)D_y))^{-1}$  can be approximated by applying sequential 1D edge-preserving smoother a multiple number of iterations [31]:

$$(I + \lambda(D_x^T A(V_1)D_x + D_y^T A(V_2)D_y))^{-1} \approx \prod_{k=1}^T (I + \lambda_k(D_x^T A(V_1)D_x))^{-1} (I + \lambda_k(D_y^T A(V_2)D_y))^{-1} \quad (17)$$

where  $T$  is a constant, and  $\lambda_k$  is  $\frac{3 \times 4^{T-k}}{2(4^T - 1)} \lambda$  [31]. This implies that the optimization problem (14) can be decomposed into two 1-D optimization problems. For each row of the image, the



**Fig. 5:** Comparison of the proposed detail-enhanced exposure fusion algorithm with the exposure fusion algorithms in [12], [15], [18]. (a)-(g) and (l)-(p): input images; (h) and (q): images fused by the algorithm in [12]; (i) and (r) images fused by the algorithm in [15]; (j) and (s): images fused by the algorithm in [18]; (k) and (t): image fused by the proposed one. Image courtesy of Laurance Meylan and Dani Lischinski.

corresponding cost function along the  $x$  dimension is defined with a 1D guidance vector field  $V^h(x)$  ( $x = 0, \dots, w-1$ ) as

$$\min_{L_d^h} \sum_{x=0}^{w-1} [(L_d^h(x))^2 + \lambda_k \left( \frac{dL_d^h(x)}{dx} - V_1(x) \right)^2]. \quad (18)$$

The solution to the above optimization problem is given in the Appendix. Similarly, the optimization problem along the  $y$  direction can be formulated and solved in a fast way. Clearly, this approximation can greatly accelerate the processing to solve the optimization problem (14). However a single iteration often leads to “streaking artifacts”, the 1D solvers are applied for  $T$  times. The value of  $T$  is selected as 3 as suggested in [31]. The speed of the proposed algorithm is almost the same as that of the fast algorithm in [31] which is comparable to those of local edge-preserving filters as shown in [31]. The complexity of the proposed detail extraction component is  $O(NM)$  for the  $N$  differently exposed images with  $M$  pixels.

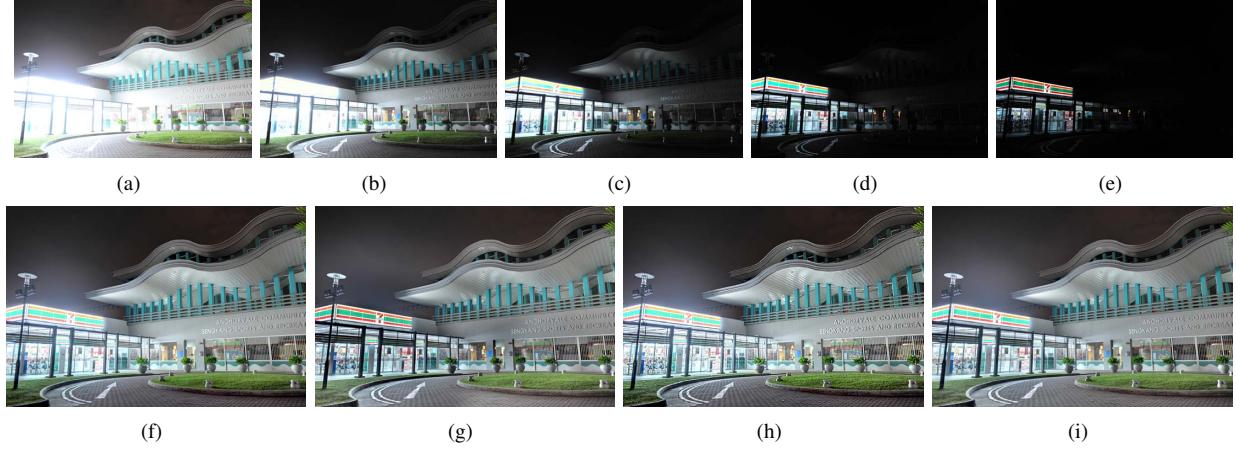
Once the detail layer  $L_d$  is extracted from the vector field  $V$ , it is added to the intermediate image. Similar to the tone mapping algorithms in [8]–[10], the final image is synthesized by

integrating the intermediate image  $Z^{int}(p)$  with the extracted details  $L_d(p)$  as

$$Z_f(p) = Z^{int}(p) \exp^{\theta L_d(p)}, \quad (19)$$

where  $\theta (\geq 0)$  is a constant and the default value of  $\theta$  is selected as 1 in this paper. It should be pointed out that the value of  $\theta$  can be selected by a user according to her/his preference.

The proposed detail-enhanced exposure fusion algorithm is compared with three multi-scale exposure fusion algorithms in [12], [15], [18] by testing six sets of differently exposed LDR images. The algorithm in [12] is the best exposure fusion algorithm, the algorithm in [18] is the second best and it also includes a detail enhancement component, and the algorithm in [15] is also based on the edge-preserving smoothing technique. It is shown in Figs. 3–6 that the algorithms in [12], [18] cannot preserve details in brightest and darkest regions well. Even though the problem is overcome by the algorithm in [15], visible halo artifacts are produced by it. The proposed algorithm can preserve details in brightest and darkest regions well and avoid halo artifacts from appearing in fused images.



**Fig. 6:** Comparison of the proposed detail-enhanced exposure fusion algorithm with the exposure fusion algorithms in [12], [15], [18]. (a)-(e): input images; (f)-(i): images fused by the algorithms in [12], [15], [18], and the proposed one, respectively.



**Fig. 7:** Comparison of detail layers extracted by the component in [18] and the proposed component. (a) detail layer extracted by the algorithm in [18]; (b) detail layer extracted by the proposed algorithm. Only fine details are extracted by the proposed component while large scale components are also extracted by the component in [18].

Due to the detail extraction components, the images fused by the algorithm in [18] and the proposed algorithm look sharper than those produced by the algorithms in [12], [15]. However, the images fused by the algorithm in [18] are over-sharpened. The problem is overcome by the proposed algorithm. This is due to the new method for the generation of the vector field. The four exposure fusion algorithms are also compared objectively using the quality metric in [19]. As shown in the Table I, the proposed exposure fusion algorithm outperforms the algorithms in [12], [15], [18] from the MEF-SSIM point of view. Clearly, the result of objective evaluation is consistent with that of the subjective evaluation.

**TABLE I:** MEF-SSIM of four different algorithms.

Set	1	2	3	4	5	6	Average	Rank
[15]	0.955	0.98	0.986	0.956	0.956	0.967	0.967	3
[12]	0.967	<b>0.988</b>	0.986	0.95	0.974	0.968	0.972	2
[18]	0.934	0.97	0.972	0.925	0.952	0.944	0.95	4
Ours	<b>0.971</b>	<b>0.988</b>	<b>0.989</b>	<b>0.957</b>	<b>0.977</b>	<b>0.978</b>	<b>0.977</b>	<b>1</b>

Since both the algorithm in [18] and the proposed algorithm include a component to simultaneously extract fine details from a set of differently exposed images. The extracted details

layers are compared in Fig. 7. It can be shown that the detail layer extracted by the proposed component only include fine details which is owe to the structure tensor while there are large scale components in the detail layer extracted by the component in [18]. In addition, the running times of the implementation in [18] are reduced by more than four times by the proposed faster implementation in this paper.

### III. SINGLE IMAGE BRIGHTENING

To capture well-posed images under various lighting conditions, the three key parameters, which are the camera aperture, the exposure time and the ISO value, are carefully tuned by professional photographers. Among the three available parameters, the latter two are widely adjusted to capture images in low-lighting and back-lighting conditions. An image captured via a pair of a long exposure time and a small ISO value is clean but blurred [33]. To reduce/avoid the blurring, a pair of a short exposure time and a large ISO value is usually chosen instead. Unfortunately, with the new combination of parameters, the captured image is sharp but noisy [33]. Meanwhile, both above methods have a limitation to capture HDR scenes because of the possible saturation [34]. An attractive method for capturing an image for a low-lighting or back-lighting scene, especially for an HDR scene is to use a pair of a short exposure time and a small ISO value. There is negligible motion blur due to the short exposure time. Furthermore, there are negligible saturated pixels because of the small ISO value. However, the captured image is darker than an image that is captured by the conventional method. The proposed exposure fusion algorithm in Section II is applied to design a simple algorithm for the brightening of the captured image. It is worth noting that several interesting single image brightening algorithms were proposed in [35]–[38].

Three virtual differently exposed LDR images are generated from the image  $Z$ . The first image  $\hat{Z}_1$  is generated by using a non-decreasing function to brighten the under-exposed regions

of the image  $Z$ . The image  $\hat{Z}_1$  is produced as [32]

$$\hat{Z}_{1,c}(p) = \Phi(Y(p))Z_c(p), \quad (20)$$

where  $c$  is the color channel and  $Y(p)$  is the luminance component of the pixel  $Z(p)$ . The function  $\Phi(\cdot)$  is derived from the conventional S-curve of automatic exposure control [39] and is empirically defined as

$$\Phi(z) = 1 + \exp^{-14(\frac{z}{255})^{1.6}}. \quad (21)$$

The function can be used to brighten dark regions of the input image without affecting bright regions of the input image. The image  $\hat{Z}_1$  can be further brightened by using a global factor to increase the brightness of the whole image  $\hat{Z}_1$  with negligible increment on the brightness of the brightest areas.

The second and the third images are generated by multiplying  $\hat{Z}_1$  by two constants  $\frac{5v}{8}$  and  $v(>\frac{8}{5})$  as follows [32]:

$$\hat{Z}_{2,c}(p) = \frac{5v}{8} \hat{Z}_{1,c}(p); \quad \hat{Z}_{3,c}(p) = v \hat{Z}_{1,c}(p), \quad (22)$$

where the value of  $v$  is empirically given as  $\frac{\sqrt{256-\bar{y}}}{4}$ ,  $\bar{y}$  is the average value of the luminance components of all under-exposed and well-exposed pixels in the image  $Z$ . The value of  $\bar{y}$  is usually less than 128. The value of  $v$  is thus usually larger than 2.828 but less than 4. Subsequently, the virtual exposure value (EV) gap between the images  $\hat{Z}_1$  and  $\hat{Z}_3$  is about 3-4 EV if the gamma value of the virtual camera is set at about 0.5. The second and third images are generated using the golden ratio.

The proposed brightening algorithm is compared with the brightening algorithm in [35], the brightening algorithm in [36], and the brightening algorithm in the Photoshop. The brightness of images is adjusted to the maximal level by Photoshop CS5. It is illustrated in Fig. 8 that 1) the brightness of all brightened images is increased; 2) the global contrast is preserved well by Photoshop CS5 but the brightest regions of the images produced by Photoshop CS5 are saturated even though the images are not as bright as those image produced by the other algorithms; and 3) saturation of the brightest regions is overcome by the algorithm in [36] while the global contrast is not preserved as well as the other algorithms. As a result, the brightened images by the algorithm in [36] look a bit flattened. As shown in Fig. 8, the brightened images by the algorithm in [36] are darker than the brightened images by the proposed method. Color is slightly distorted by the algorithms in [36], for example the color of the white cloth in Fig. 8(f) is changed. The algorithm in [35] generally works well except for slight color distortion and loss of fine details. The brightening algorithm in the Photoshop is the simplest, and followed by the proposed algorithm, the brightening algorithm in [36], and the brightening algorithm in [35]. It should be pointed out that noise in the under-exposed regions is amplified by all the brightening algorithms. This is an limitation of the proposed brightening algorithm. This problem could be addressed by using the solution provided in [38] if the computational cost is not an issue.

Despite that modern cameras are equipped with sophisticated metering techniques, it remains a challenge for users to take

well-exposed images, especially in presence of backlighting. Dark objects could appear when photos are taken for human subjects in a scene with the backlit. The proposed brightening algorithm is applied to brighten dark objects in two images that are captured at day time. It is shown in Fig. 9 that with the proposed algorithm, the darkest regions are brightened while the brightness of the brightest regions almost remains the same. As a result, the exposedness level of the dark human subjects is significantly improved in the brightened images.

#### IV. CONCLUSION REMARKS

A multi-scale exposure fusion algorithm has been proposed to merge differently exposed low dynamic range (LDR) images of a high dynamic range (HDR) scene into a more informative and perceptually appealing LDR image using the weighted guided image filter. In addition, a new concept of weighted structure tensor is introduced to the differently exposed images and it is used to design a detail extraction component to enhance the proposed algorithm. The proposed fusion algorithm is applied to design a single image brightening algorithm for both low-light imaging and back-light imaging. Experimental results show that the proposed algorithms can be applied to produce images with better visual quality.

One interesting problem is to determine the value of  $\theta$  in Equation (19) by using the quality metrics in [19]. Another interesting problem is how to avoid amplifying noise when an image is brightened. Both problems will be studied in our future research.

#### APPENDIX: SOLUTION TO THE OPTIMIZATION PROBLEM (18)

Defining  $\frac{dL_d^h(x)}{dx}$  as  $(L_d^h(x+1) - L_d^h(x))$ , it can be derived that the optimal solution is obtained by solving the following linear equation:

$$A * L_d^h = D, \quad (23)$$

where the matrices  $A$ ,  $L_d^h$  and  $D$  are given as

$$A = \begin{bmatrix} b_0 & c_0 & 0 & \cdots & 0 \\ a_0 & b_1 & c_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_{w-3} & b_{w-2} & c_{w-2} \\ 0 & \cdots & 0 & a_{w-2} & b_{w-1} \end{bmatrix}, \quad (24)$$

$$L_d^h = [ L_d^h(0) \quad L_d^h(1) \quad L_d^h(2) \quad \cdots \quad L_d^h(w-1) ]^T, \quad (25)$$

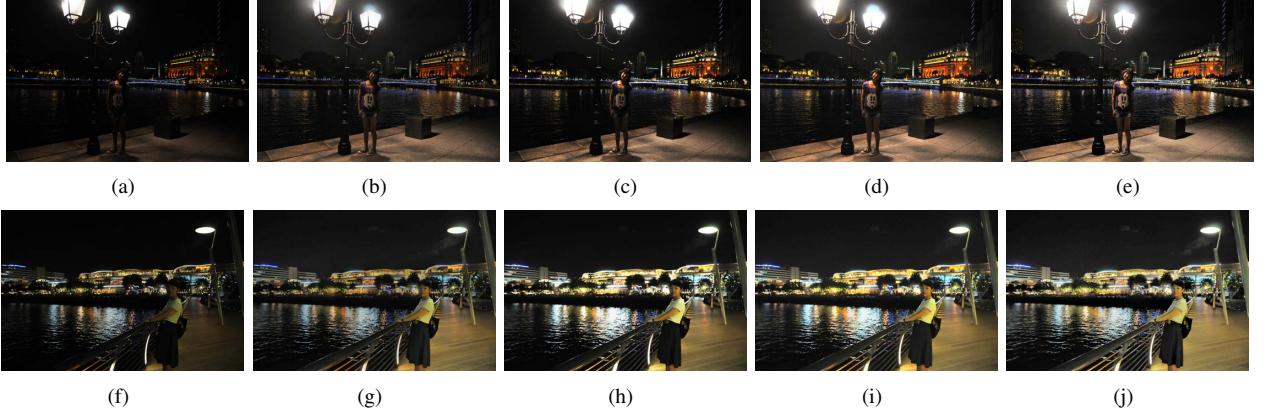
$$D = [ d_0 \quad d_1 \quad d_2 \quad \cdots \quad d_{w-1} ]^T, \quad (26)$$

$a_i$ ,  $b_i$ ,  $c_i$  and  $d_i$  are computed as

$$a_i = c_i = -\frac{\lambda_k}{\psi^2(V_1(i))}; \quad 0 \leq i \leq w-2,$$

$$b_0 = 1 + \frac{\lambda_k}{\psi^2(V_1(0))},$$

$$b_{w-1} = 1 + \frac{\lambda_k}{\psi^2(V_1(w-2))},$$



**Fig. 8:** Comparison of different image enhancement algorithms. (a, f) low-lighting images; (b, g) brightened images by the brightening algorithm [36]; (c, h) brightened images by Photoshop CS5; (d, i) brightened images by the algorithm in [35]; (e, j) brightened images by the proposed algorithm. Fine details in the darkest regions are enhanced better by the proposed algorithm without causing over-saturation to the brightest regions.



**Fig. 9:** Images with dark objects and their enhanced images. (a, c) images with dark objects; (b, d) brightened images by the proposed algorithm. The under-exposed human subjects become well-exposed in the brightened images.

$$\begin{aligned} b_i &= 1 + \frac{\lambda_k}{\psi^2(V_1(i))} + \frac{\lambda_k}{\psi^2(V_1(i-1))}; \quad 1 \leq i \leq w-2, \\ d_0 &= -\lambda_k \frac{V_1(0)}{\psi^2(V_1(0))}, \\ d_{w-1} &= \lambda_k \frac{V_1(w-2)}{\psi^2(V_1(w-2))}, \\ d_i &= \lambda_k \left( \frac{V^h(i-1)}{\psi^2(V_1(i-1))} - \frac{V^h(i)}{\psi^2(V_1(i))} \right); \quad 1 \leq i \leq w-2. \end{aligned}$$

For simplicity, define two matrices  $B$  and  $C$  as

$$B = \begin{bmatrix} \tilde{b}_0 & 0 & 0 & \cdots & 0 \\ a_0 & \tilde{b}_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & a_{w-3} & \tilde{b}_{w-2} & 0 \\ 0 & \cdots & 0 & a_{w-2} & \tilde{b}_{w-1} \end{bmatrix}, \quad (27)$$

$$C = \begin{bmatrix} 1 & \frac{c_0}{\tilde{b}_0} & 0 & \cdots & 0 \\ 0 & 1 & \frac{c_1}{\tilde{b}_1} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \frac{c_{w-2}}{\tilde{b}_{w-2}} \\ 0 & \cdots & 0 & 0 & 1 \end{bmatrix}, \quad (28)$$

where  $\tilde{b}_i (1 \leq i \leq w-1)$  is

$$\tilde{b}_i = b_i - \frac{a_{i-1}c_{i-1}}{\tilde{b}_{i-1}}; \quad 1 \leq i \leq w-1, \quad (29)$$

and the value of  $\tilde{b}_0$  is  $b_0$ . The optimal solution is given as

$$L_d^{h,*} = C^{-1} * B^{-1} * D, \quad (30)$$

where  $C^{-1}$  and  $B^{-1}$  are the inverse matrices of  $C$  and  $B$ , respectively.

## REFERENCES

- [1] P. E. Debevec and J. Malik, “Rendering high dynamic range radiance maps from photographs,” In *Proceedings of SIGGRAPH*, pp. 369-378, USA, May 1997.
- [2] K. Jacobs, C. Loscos, and G. Ward, “Automatic high dynamic range image generation for dynamic scenes,” *IEEE Computer Graphics and Applications*, vol. 28, no. 2, pp. 84-93, Apr. 2008.
- [3] S. Q. Wu, Z. G. Li, J. H. Zheng, and Z. J. Zhu, “Exposure robust method for aligning differently exposed images,” *IEEE Signal Processing Letter*, vol. 21, no. 7, pp. 885-889, Jul. 2014.
- [4] J. Hu, O. Gallo, and K. Pulli, “Exposure stacks of live scene with hand-held cameras,” In *Proceedings of the European Conference of Computer Vision*, pp.499-512, Italy, Oct. 2012.
- [5] P. Sen, N. Khademi Kalantari, M. Yaesoubi, S. Darabi, D. Goldman, and E. Shechtman, “Robust patch-based HDR reconstruction of dynamic scenes,” *ACM Trans. on Graphics*, vol. 31, no. 6, Art. 203, Nov. 2012.
- [6] J. H. Zheng, Z. G. Li, Z. J. Zhu, S. Q. Wu, and S. Rahardja, “Hybrid patching for a sequence of differently exposed images with moving objects,” *IEEE Trans. on Image Processing*, vol. 22, no. 12, pp.5190-5201, Dec. 2013.

- [7] Z. G. Li, J. H. Zheng, Z. J. Zhu, and S. Q. Wu, "Selectively detail-enhanced fusion of differently exposed images with moving objects," *IEEE Trans. on Image Processing*, vol. 23, no. 10, pp. 4372-4382, Oct. 2014.
- [8] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," In *SIGGRAPH*, pp. 257-266, USA, Aug. 2002.
- [9] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and details manipulation," *ACM Trans. on Graphics*, vol. 27, no. 3, pp. 249-256, Aug. 2008.
- [10] Z. G. Li and J. H. Zheng, "Visual salience based tone mapping for high dynamic range images," *IEEE Trans. on Industrial Electronics*, vol. 61, no. 12, pp. 7076-7082, Dec. 2014.
- [11] F. Kou, W. H. Chen, C. Y. Wen, and Z. G. Li, "Gradient domain guided image filtering," *IEEE Trans. on Image Processing*, vol. 24, no. 11, pp. 4528-4539, Nov. 2015.
- [12] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: a simple and practical alternative to high dynamic range photography," *Computer Graphics Forum*, Vol. 28, pp. 161-171, 2009.
- [13] P. Burt and T. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532-540, Apr. 1983.
- [14] W. Zhang and W. K. Cham, "Gradient-directed multiexposure composition," *IEEE Trans. on Image Processing*, vol. 21, no. 4, pp. 2318-2323, Apr. 2012.
- [15] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Processing*, vol. 22, no. 7, pp. 2864-2875, Jul. 2013.
- [16] 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.
- [17] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, and C. Zhang, "Probabilistic exposure fusion," *IEEE Trans. Image Process.*, vol. 21, no. 1, pp. 341-357, Jan. 2012.
- [18] Z. G. Li, J. H. Zheng, and S. Rahardja, "Detail-enhanced exposure fusion," *IEEE Trans. on Image Processing*, vol. 21, no. 11, pp. 4672-4676, Nov. 2012.
- [19] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3345-3356, Nov. 2015.
- [20] R. Shen, I. Cheng, and A. Basu, "QoE-based multi-exposure fusion in hierarchical multivariate gaussian CRF", *IEEE Trans. on Image Processing*, vol. 22, no. 6, pp. 2469-2478, Jun. 2013.
- [21] M. Bertalmio and S. Levine, "Variational approach for the fusion of exposure bracketed pairs", *IEEE Trans. on Image Processing*, vol. 22, no. 2, pp. 712-723, Feb 2013.
- [22] K. Hara, K. Inoue, and K. Urahama, "A differentiable approximation approach to contrast-aware image fusion", *IEEE Signal Processing Letters*, vol. 21, no. 6, pp. 742-745, Jun. 2014.
- [23] Z. G. Li, J. H. Zheng, Z. J. Zhu, W. Yao, and S. Q. Wu, "Weighted guided image filtering," *IEEE Trans. on Image Processing*, vol. 24, no. 1, pp. 120-129, Jan. 2015.
- [24] K. Ma and Z. Wang, "Multi-exposure image fusion: a patch-wise approach", In *IEEE International Conference on Image Processing*, pp. 1717-1721, Canada, Sept. 2015.
- [25] K. Zeng, K. Ma, R. Hassen and Z. Wang, "Perceptual evaluation of multi-exposure image fusion algorithms", In *The 6th International Workshop on Quality of Multimedia Experience (QoMEX)*, pp. 7-12, Singapore, Sept. 2014.
- [26] V. Vonikakis, O. Bouzos, and I. Andreadis, I, "Multi-exposure image fusion based on illumination estimation," In *LASTED SIPA 2011*, pp.135-142, 2011, Greece.
- [27] S. Di Zenzo, "A note on the gradient of a multi-image," *Comp. Vision, Graphics, and Image Proc.*, vol. 33, pp. 116125, 1986.
- [28] D. Socolinsky and L. Wolff, "Multispectral image visualization through first-order fusion," *IEEE Trans. Image Processing*, vol. 11, no. 8, pp. 923-931, Aug. 2002.
- [29] D. Connah, M. S. Drew, and G. D. Finlayson, "Spectral edge image fusion: Theory and applications," in *Euro.Conf. on Comp. Vis.*, pp. 65-80, 2014.
- [30] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing", *ACM Trans. on Graph.*, vol. 30, no. 4, Jul. 2011, Art. ID 69
- [31] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. Do, "Fast global image smoothing based on weighted least squares", *IEEE Trans. on Image Processing*, vol. 23, no. 12, pp. 5638-5653, Dec. 2014.
- [32] Z. G. Li and J. H. Zheng, "Single image brightening via exposure fusion," In *International Conference on Acoustics, Speech, and Signal Processing*, pp. 1756-1760, China, Mar. 2016.
- [33] L. Zhang, A. Deshpande, and X. Chen, "Denoising vs. deblurring: HDR imaging techniques using moving cameras," In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 522-529, Jun. 2010, USA.
- [34] S. W. Hasinoff, F. Durand, and W. T. Freeman, "Noise-optimal capture for high dynamic range photography," In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 553-560, Jun. 2010, USA.
- [35] T. H. Wang, C. W. Chiu, W. C. Wu, J. W. Wang, C. Y. Lin, C. T. Chiu, and J. J. Liou, "Pseudo-multiple-exposure-based tone fusion with local region adjustment," *IEEE Trans. on Multimedia*, vol. 17, no. 4, pp. 470-484, Apr. 2015.
- [36] X. Dong, G. Wang, Y. Pang, W. X. Li, J. T. Wen, W. Meng, and Y. Lu, "Fast efficient algorithm for enhancement of low lighting video," In *2011 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6, Jul. 2011, Spain.
- [37] J. W. Seo and S. D. Kim, "Contrast enhancement of back-light images via a regional rank-1 constraint", In *IEEE International Conference on Image Processing*, pp. 2955-2959, Canada, Sept. 2015.
- [38] L. Li, R. G Wang, W. M Wang, and W. Gao, "A low-light image enhancement method for both denoising and contrast enlarging", In *IEEE International Conference on Image Processing*, pp. 3730-3734, Canada, Sept. 2015.
- [39] L. Yuan and J. Sun, "Automatic exposure correction of consumer photographs", In *European Conference on Computer Vision*, pp.771-785, Italy, Oct. 2012.