ORIGINAL ARTICLE

Fast exposure fusion of detail enhancement for brightest and darkest regions

Chunmeng Wang[1] · **Chen He**[2] · **Minfeng Xu**[3]

Abstract
Multi-exposure fusion is the common approach to generate high dynamic range (HDR) images that combines multi-exposure images captured for the same scene, but the traditional multi-exposure fusion algorithms lose details in the brightest and darkest regions of the scene. Therefore, many detail enhancement-based exposure fusion algorithms have been proposed to extract these details. However, these algorithms have low efficiency because of the complexity of detail enhancement mechanism, and most of them excessively enhance all the pixels besides of the necessary brightest and darkest pixels. We propose a local detail enhancement mechanism to enhance only the details of brightest and darkest regions by using fast local Laplacian filtering (FLLF). A large number of experiments show that the proposed algorithm has much more high efficiency than the current detail enhancement-based exposure fusion algorithms, and the brightest and darkest details in the high dynamic range scene are preserved well.

Keywords High dynamic range image · Exposure fusion · Detail enhancement · Fast local Laplacian filtering

1 Introduction

The dynamic range of the real scene is extremely large while the common camera cannot capture all the information and details in the bright or dark regions of high dynamic range (HDR) scene because camera sensors are unable to recover so much luminance and color information [10]. The general practice of generating HDR images is to directly fuse multi-exposure images of the same scene into one HDR image, which is called exposure fusion. The traditional exposure fusion methods preserved details in most regions but they lost useful details in the brightest and darkest regions.

Recently, many detail enhancement-based exposure fusion methods have been proposed to enhance the details in the brightest and darkest regions. Most of these methods are based on multiresolution Laplacian pyramids and introduce one detail extraction mechanism to enhance details. However, these methods are mostly much slower than traditional fusion methods because of two reasons. Firstly, most of them not only enhance the brightest and darkest pixels, but also enhance all the other pixels of all input LDR images. This indistinguishable processing way will lead to two drawbacks. One is that it increases the algorithm complexity, and the second is that it leads to excessive enhancement of details because traditional fusion methods have preserved enough details in the most regions without need of over enhancement. Secondly, the detail enhancement mechanisms of them are very time-consuming, such as gradient-domain guided image filter (GGIF) in [13], weighted least square (WLS) optimization in [5] and so on.

In order to improve the efficiency of detail enhancement-based exposure fusion methods, we introduce a local enhancement strategy by determining the region of enhancement (RoE) for each input image to enhance details only in the brightest and darkest regions, and our detail enhancement mechanism is based on fast local Laplacian filtering (FLLF) which is faster than the previous detail extraction mechanisms. The local enhancement strategy and the fast detail enhancement mechanism not only improve the efficiency greatly, but also avoid excessive detail enhancement

*   Chunmeng Wang
     wchm87@jit.edu.cn

1   School of Computer Engineering, Jinling Institute
     of Technology, Nanjing 211169, Jiangsu, China

2   Media and Communication College, Weifang University,
     Weifang 261061, Shandong, China

3   School of Computer Science and Technology, Shandong
     University of Finance and Economics, Jinan 250014,
     Shandong, China

of none-RoE regions. The experiments show that the proposed algorithm is much faster than previous detail enhancement based exposure fusion algorithms and generates visually pleasing and detail-preserving fusion image in the brightest and darkest regions.

2 Related work

Many kinds of exposure fusion methods have been proposed to preserve details in the overexposed and underexposed regions of the high dynamic range scene. Wei and Cham [24] fused multiple images with different exposures in both static and dynamic scenes with the gradient-based quality measures because the gradient magnitude decreases gradually as the pixel tends to overexposure or underexposure. Mertens et al. [16] proposed a multi-resolution based method to decompose each input image into Laplacian pyramid and build the weight map Gaussian pyramid for each image determined by contrast, saturation and well-exposedness values and then collapse the weighted averaged Laplacian pyramid for each pyramid level. The multiscale method fused multi-exposure images well for most scenes but lost useful details in the brightest and darkest regions. Our method is also based on multiresolution pyramid frame, but we introduce a detail enhancement mechanism in the brightest and darkest areas. Rui et al. [21] proposed a generalized random walks framework to achieve a global optimization probabilistic model taking neighborhood information into account with local contrast and color consistency. Mingli et al. [17] proposed a probabilistic model with both local adaptation and gradient consistency optimization. Mansour et al. [15] proposed a two-scale decomposition approach, and the blending weight is based on an exposedness function with the input luminance component. These traditional exposure fusion methods were mostly of high efficiency but the details and color visibility in the brightest and darkest regions were preserved not so well. Moreover, the deep learning-based method [19] was proposed to fuse static multi-exposure images, but the training process was less flexible because it only learned to fuse a fixed number of images, and the scale of multi-exposure image dataset was too small to cover all real scenes.

Recently, multi-exposure fusion method was applied to generate one HDR image with detail enhancement from one underexposed input image in [22] or one low-contrast input image in [8]. On the contrary, detail enhancement methods can also be introduced to multi-exposure fusion. Many detail enhancement-based exposure fusion methods have been proposed to extract the detail information in the brightest and darkest regions. Most of them introduced a detail extraction mechanism based on edge-preserving filters that could avoid artifacts with strong edges not be blurred in the

decomposition process, such as bilateral filter [2], weighted least squares (WLS) [3], guided filter [9] and other improved filters. For example, Raman and Chaudhuri [20] generated the weight value for each pixel with the difference between the original pixel value and the pixel value after bilateral filtering. The method preserved strong edges but lost global contrast and color visibility. Shutao et al. [23] proposed a two-scale weighted average fusion method based on guided filtering by controlling the two measures of pixel saliency and spatial consistency. Details were abstracted from the gradient vector in [25] by using improved weighted least square optimization and then added to the fused image, but the iterative processing for solving the optimization problem is very slow. Jianbing et al. proposed a detail-preserving fusion method-based subband architecture and the gain control strategy in [7] and a method-based boosting Laplacian pyramid guided by the local exposure weight and the JND-based saliency weight in [6]. The method preserved both color appearance and texture structure, but the details in the fused image were overenhanced. Li et al. [13] also used an weighted guided image filter (WGIF) to smooth the Gaussian pyramids and adopted one fast separate approach to solve the optimization problem to accelerate the method [25], but the computational cost is still high while generating the vector fields. Ma et al. [14] proposed one structural patch decomposition based MEF (SPD-MEF) that fused three independent components: signal strength, signal structure and mean intensity based on image patches, but its structural patch decomposition operation is time-consuming. Li et al. [12] proposed a fast multiscale method to accelerate SPD-MEF without normalization process, which has achieved much faster speed than SPD-MEF. Kou et al. used the fast weighted least square (FWLS) optimization in [5] and introduced edge-preserving smoothing pyramid in [4] to extract the details in the brightest and darkest regions efficiently for exposure fusion. These detail enhancement-based exposure fusion methods mostly extracted details for all the pixels of all input LDR images instead of only dealing with the brightest and darkest pixels. This processing way not only increases the algorithm complexity, but also leads to detail overenhancement for the normal pixels.

Unlike previous detail enhancement-based fusion methods, we propose a local enhancement strategy to only deal with the brightest and darkest regions determined by the average brightness and keep the other regions unchanged. Although the method [5] only extracted the fine details in the brightest regions from the most underexposed image and the darkest regions from the most overexposed image instead of processing all input images, these details will not be extracted when the most overexposed or overexposed image contains very few useful details.

Furthermore, because the previous detail enhancement mechanisms were mostly slow, we adopt one faster detail

enhancement mechanism inspired by fast local Laplacian filters (FLLF) proposed by Aubry et al. [1]. The FLLF is the accelerated version based on local Laplacian filters (LLF) proposed by Paris et al. [18]. The local Laplacian filters use standard Laplacian pyramids to achieve edge-preserving smoothing or small-scale details enhancement. The approach is very simple for relying only on simple point-wise nonlinearities and small Gaussian convolutions instead of complex optimization process. Our method proposes an improved fast local Laplacian filters to use in the Laplacian pyramid-based fusion method that yields speedup by several times.

3 Our approach

3.1 Overview of our approach

Suppose that the N input exposure images are sorted from the most underexposed I_1 to the most overexposed I_N and have been strictly aligned for pixels at the same location (x, y), if we use the multiresolution strategy [16] to fuse the multi-exposure images, it will lose fine details in the brightest or darkest regions though it preserved enough details in the most normal regions. This is because that the Laplacian pyramid is unable to represent edges or preserve details well and ill-suited for edge-preserving smoothing. The over-averaging of weights

based on spatially invariant Gaussian kernels makes the weight values which contain fine details in brightest or darkest regions lose dominance. The weakness is more obvious as the pyramid level d is larger.

In this paper, we firstly determine the region of enhancement (RoE) for each image in the brightest and darkest regions by computing the average brightness for each location (x, y). For the pixels in the RoE, we use the improved fast local Laplacian filtering as our detail extraction mechanism to extract darkest or brightest details, and finally use the multiresolution method to generate the fused image. Figure 1 shows the pipeline of our algorithm.

3.2 Region of enhancement (RoE)

The value of average brightness of N input exposure images for each location (x, y) means the luminance of this location within the HDR scene, so we determine the darkest region $D_{x,y}$ and the brightest region $B_{x,y}$ according to the average brightness by

$$\bar{l}_{x,y} = \frac{\sum_{k=1}^{N} l_{x,y,k}}{N} \tag{1}$$

where $l_{x,y,k}$ is the brightness value of (x, y) of the kth image. We regard pixels to be located in the normal region if
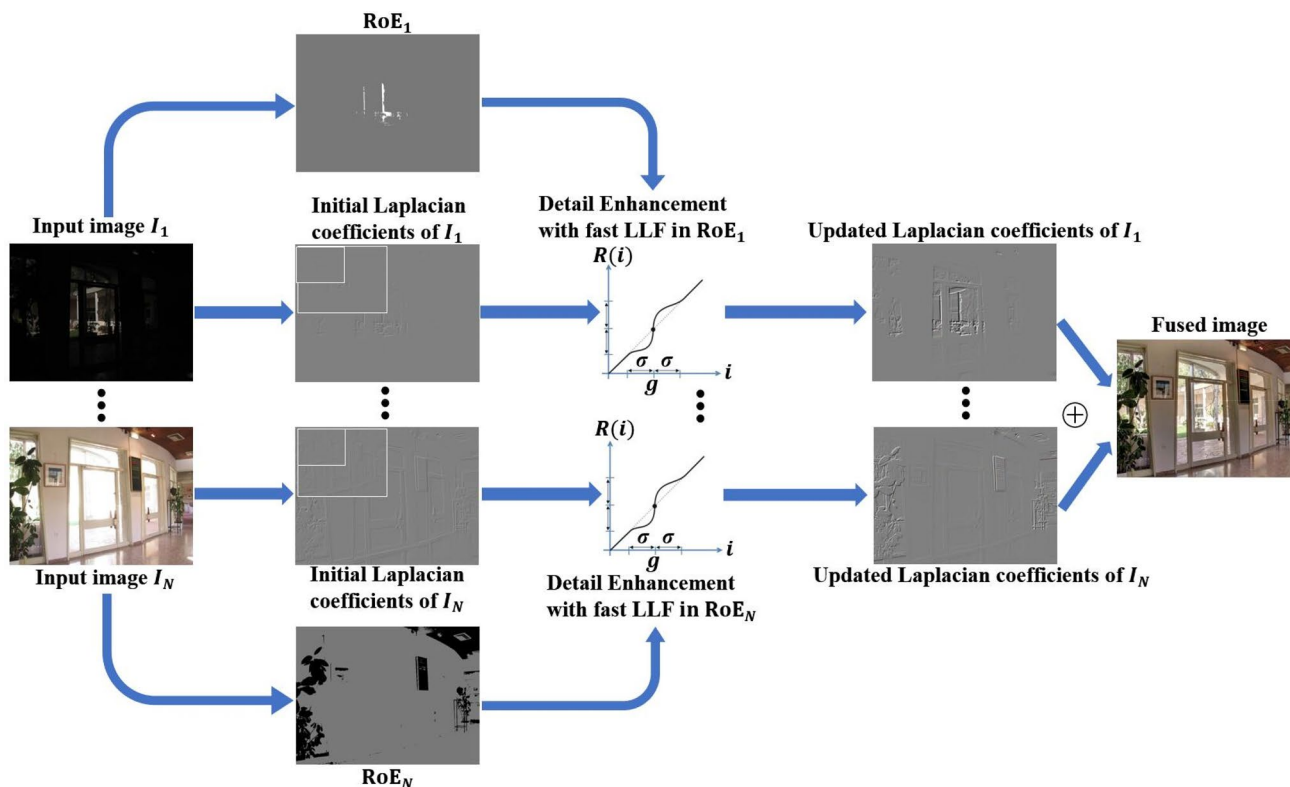

Fig. 1 Pipeline of our method

$\lambda \leq \bar{l}_{x,y} \leq 1 - \lambda$. Parameter λ is variable from 0 to 0.5. In this paper, we set $\lambda = 0.1$ that means the pixel (x, y) is in the darkest region $D_{x,y}$ if $\bar{l}_{x,y} < 0.1$ and in the brightest region $B_{x,y}$ if $\bar{l}_{x,y} > 0.9$, as shown in Fig. 2b. We use the detail enhancement mechanism to enhance the details only in $D_{x,y}$ or $B_{x,y}$ and remain the normal regions unchanged.

However, the pixels of $D_{x,y}$ in the underexposed images are always invisible with little details, so it is completely unnecessary to enhance them for reducing the computation time (Fig. 2c), and the pixels of $B_{x,y}$ in the overexposed images are always the same situation (Fig. 2d). Therefore, in order to further reduce the pixels of detail enhancement, we determine the region of enhancement (RoE) for each image (Fig. 2e) by excluding the darkest region in the underexposed images and the brightest region in the overexposed images as the following way:

$$\text{RoE}_{x,y,k} = \begin{cases} B_{x,y}, & \text{if } k \leq \left\lfloor \frac{N}{2} \right\rfloor \\ \text{none}, & \text{if } \left\lfloor \frac{N}{2} \right\rfloor < k \leq \left\lfloor \frac{N}{2} \right\rfloor \\ D_{x,y}, & \text{if } k > \left\lfloor \frac{N}{2} \right\rfloor \end{cases} \quad (2)$$

The RoE includes a set of pixels located in the brightest and darkest regions whose details need to be enhance. To reduce computation time and avoid overenhancement, we only manipulate the Laplacian pyramid coefficients in the RoE to enhance details in the next section and keep other Laplacian pyramid coefficients of input images unchanged. Compared with previous detail extraction mechanisms of other methods, our method processes much fewer pixels because the details of pixels only in the RoE are extracted.

3.3 Detail enhancement mechanism with local Laplacian filters in the RoE

For each pixel (x, y) of I_k, we first use the traditional Laplacian pyramid generation method to construct the initial Laplacian

(a) Input images

(b) $D_{x,y}$ and $B_{x,y}$ **(c)** $D_{x,y}$ in I_1 **(d)** $B_{x,y}$ in I_N
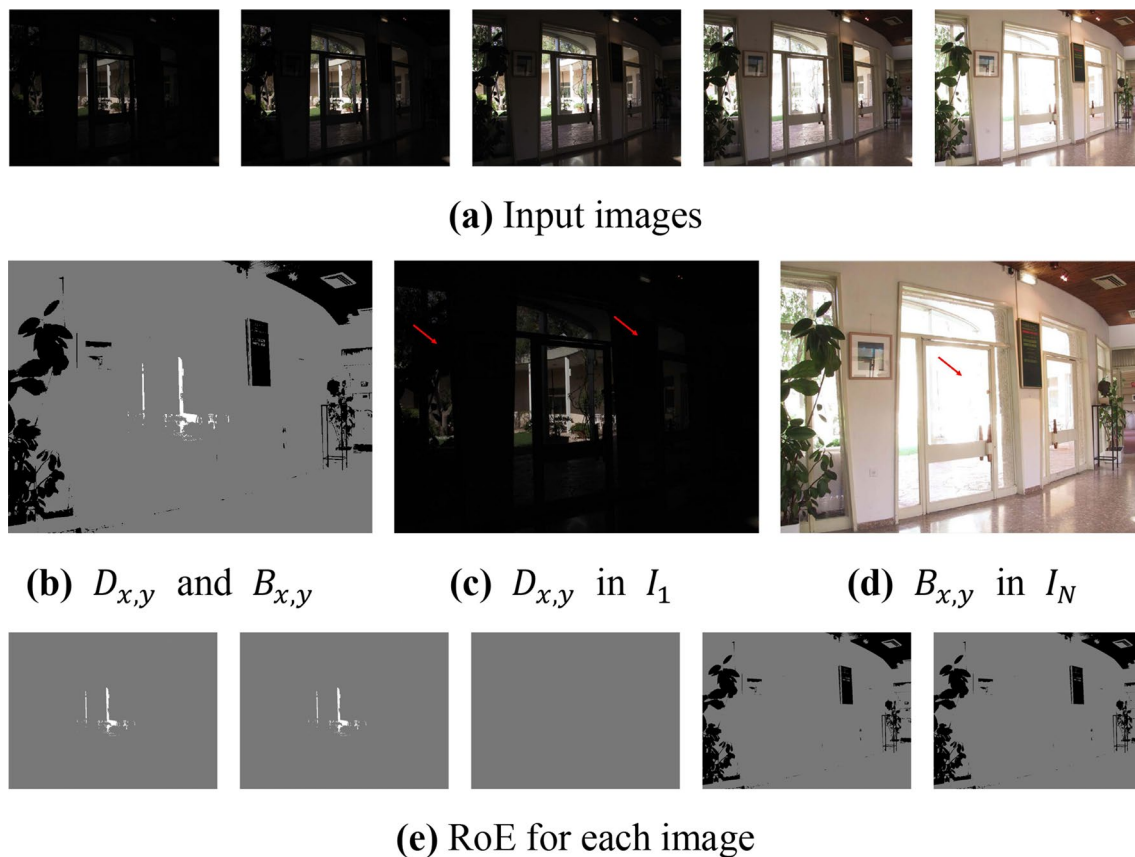

(e) RoE for each image

Fig. 2 **a** Shows input images of the set BeligiumHouse courtesy of Dani Lischinski. **b** Shows the darkest and brightest regions. **c** Shows the darkest region in the first image and **d** Shows the brightest region in the last image. **e** Shows the RoE pixels for each image. The black color means darkest pixels, the white color means brightest pixels and the gray color means the none-RoE pixels

pyramid $L\{I\}^d_{x,y,k}$ for the dth level with the range $0 \le d < L$ as follows

$$L\{I\}^d_{x,y,k} = G\{I\}^d_{x,y,k} - \text{upsample}\left\{G\{I\}^{d+1}_{x,y,k}\right\} \qquad (3)$$

where $G\{I\}^d_{x,y,k}$ is the dth level Gaussian pyramid. Suppose H and W are the height and width of the input image, respectively, then the largest possible pyramid level L is calculated as follows

$$L = \log_2 \min(H, W) \qquad (4)$$

For detail enhancement in the RoE, we need to update the Laplacian coefficients that are located in the RoE. For each coefficient (x, y) of the level d in the RoE of $I_k (k = 1 \cdots N)$, inspired by the local Laplacian filters (LLF) method [18], we generate an local intermediate image I' that is a portion of the input image with size of $K \times K$, where K is related to the current pyramid level d by $K = 2^{d+3} + 1$. We remap each coefficient from the intermediate image I' to I'_R by a point-wise remapping function $R(I')_{g,\sigma}$ and then compute the Laplacian subpyramid of the remapped image $L\{I'_R\}^d_{x,y,k}$, finally copy the corresponding pyramid coefficient to the output pyramid by

$$L\{\text{RoE}\}^d_{x,y,k} = L\{I'_R\}^d_{x,y,k} \qquad (5)$$

where $L\{\text{RoE}\}^d_{x,y,k}$ is the updated Laplacian pyramid of RoE of I_k on the dth level.

This remapping function depends on the local image value from the Gaussian pyramid $g = G_0(x, y)$ for one channel, and the parameter σ is used to distinguish edges from details. Therefore, the parameter σ controls what magnitude of variations should be considered edges that are preserved constant. Intensity variations smaller than σ should be considered fine-scale details and larger variations are edges. That is, if $|i - g| < \sigma$, it is regarded as details, where i is an intensity value. We aim to manipulate multilevel details and leave the edges not treated. The parameter σ is set to 0.5 in our method to extract details in the RoE. Larger value allows the filter to manipulate more details. The remapping function modifies the fine-scale details by altering the amplitude around the value g defined as

$$R(i)_{g,\sigma} = g + \sigma \frac{i - g}{|i - g|} f\left(\frac{|i - g|}{\sigma}\right) \qquad (6)$$

where f is a point-wise smooth function mapping from [0,1] to [0,1] for the local details manipulation by $f(\Delta) = \Delta^\alpha$ to modify the RoE details of I_k, where $0 < \alpha < 1$ is a user parameter to control the detail enhancement amount, and the amplitudes of details are enhanced more as α is smaller. The parameter α is set to 0.25 in our method. We also use the same smooth strategy in [18] to reduce noise and artifacts.

(a) Mertens09 [16]   **(b)** Li17 [13]   **(c)** Kou18 [5]   **(d)** Shutao13 [23]

(e) Rui11 [21]   **(f)** Ma17 [14]   **(g)** Li20 [12]   **(h)** Ours

(i) Zoomed result of (a), (b), (c), (d), (e), (f), (g) and (h) from left to right, respectively.

Fig. 3 Comparison with other exposure fusion methods [5, 12–14, 16, 21, 23] on image set BeligiumHouse. Image courtesy of Dani Lischinski

(a) Input images

(b) Mertens09 [16]   **(c)** Li17 [13]   **(d)** Kou18 [5]   **(e)** Shutao13 [23]

(f) Rui11 [21]   **(g)** Ma17 [14]   **(h)** Li20 [12]   **(i)** Ours

(j) Zoomed result of (b), (c), (d), (e), (f), (g), (h) and (i) from left to right, respectively.
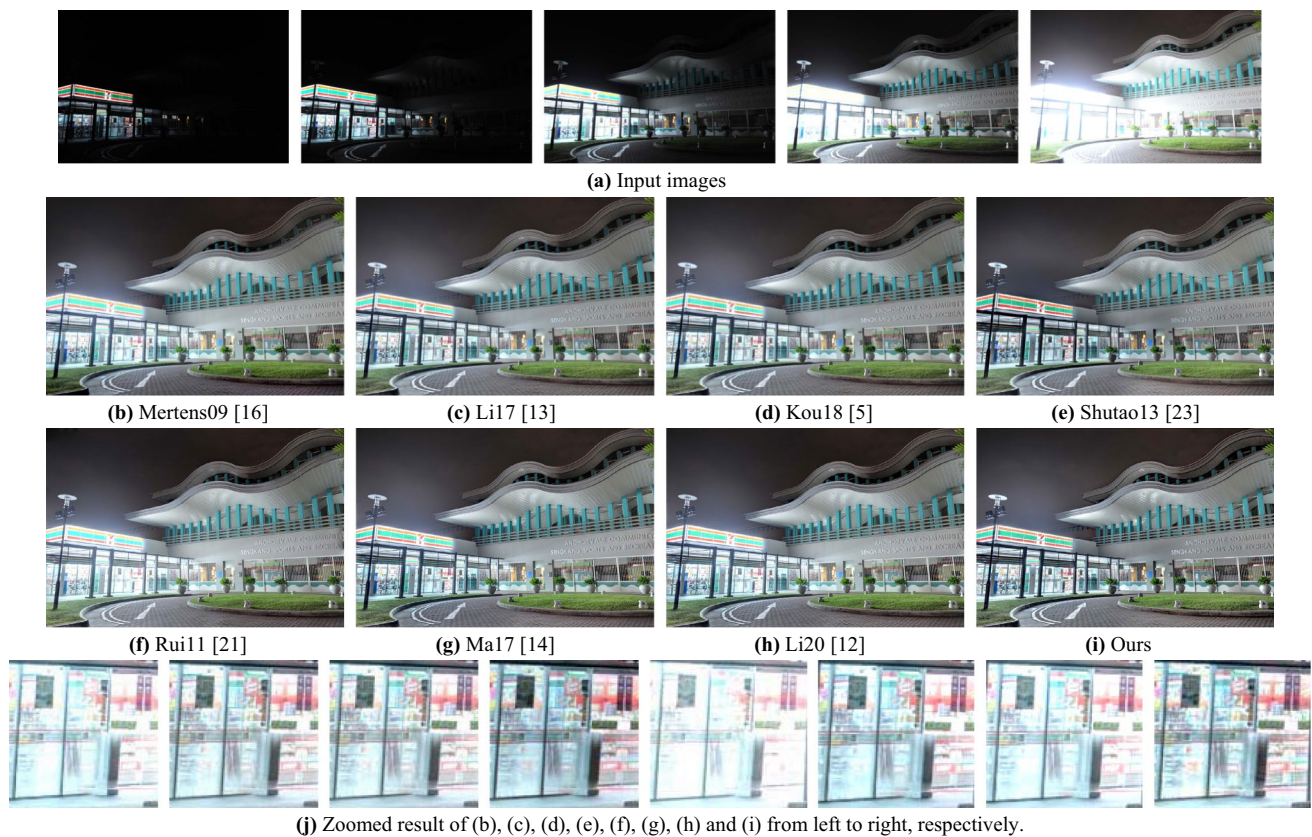
Fig. 4 Comparison with other exposure fusion methods [5, 12–14, 16, 21, 23] on image set SportsCentre. Image courtesy of Laurence Meylan

For the color images, it is possible to treat three independent channels, e.g., the RGB channels by Eq. (6) to achieve similar enhancement results, but in order to improve the efficiency, we extend the remapping function to 3D color channels by directly dealing with RGB vectors. The advantage of adopting RGB vectors instead of independent channels is shown in Sect. 4.3. The intensity remapping function $R(i)_{g,\sigma}$ is altered to the vector version $R(\mathbf{i})_{\mathbf{g},\sigma}$ that modifies details by altering the amplitude around the local vector $\mathbf{g}$ defined as

$$R(\mathbf{i})_{\mathbf{g},\sigma} = \mathbf{g} + \sigma \frac{\mathbf{i} - \mathbf{g}}{\|\mathbf{i} - \mathbf{g}\|} f\left(\frac{\|\mathbf{i} - \mathbf{g}\|}{\sigma} \right) \tag{7}$$

where $\mathbf{i}$ is a three-dimensional vector for the RGB channels.

The process of computing Laplacian pyramid of the intermediate image $L\left\{ R(I')_{g,\sigma} \right\}^{d}_{x,y,k}$ involves the time-consuming power function for each (x, y). In order to reduce the complexity of this process, we introduce the acceleration mechanism in the next section to speed up computing of the intermediate Laplacian pyramid.

3.4 Acceleration scheme

Since each pyramid coefficient is constructed by a point-by-point mapping function in a detail-enhanced intermediate image I', which is time-consuming. Inspired by the fast local Laplacian filters (FLLF) [1], one acceleration scheme is proposed based on discrete sampling and interpolation instead of directly calculating each Laplacian pyramid coefficient. Specifically, taking a single channel as an example, we calculate the Gaussian pyramid $g = G(x, y)$ of each remapped image and then regularly sample the value range of g with a set of γ_j values from small to large ($\gamma_1 < \gamma_2 \cdots < \gamma_{S-1} < \gamma_S$), where $1 \le j \le S$ and S is the number of sampling. For one sampling value γ_j of g in the RoE, a set of local Laplacian pyramid coefficient values $L\left\{ R(I')_{\gamma_j,\sigma} \right\}$ are calculated and stored according to Eq. (5) and Eq. (6), where $R(I')_{\gamma_j,\sigma}$ is $g = \gamma_j$ mapping function; this is also a predetermined sampling point, the mapping function image $L\left\{ R(I')_{\gamma_j,\sigma} \right\}$ corresponding to each sample value γ_j and its Laplacian pyramid are calculated in advance. When we calculate the pyramid coefficient (x, y) of a specific g for the dth level value and it is located in the interval $[\gamma_j, \gamma_{j+1}]$, that is $\gamma_j \le g < \gamma_{j+1}$, the interpolation parameters λ and γ_j are

(a) Input images

(b) Mertens09 [16]   **(c)** Li17 [13]   **(d)** Kou18 [5]   **(e)** Shutao13 [23]

(f) Rui11 [21]   **(g)** Ma17 [14]   **(h)** Li20 [12]   **(i)** Ours

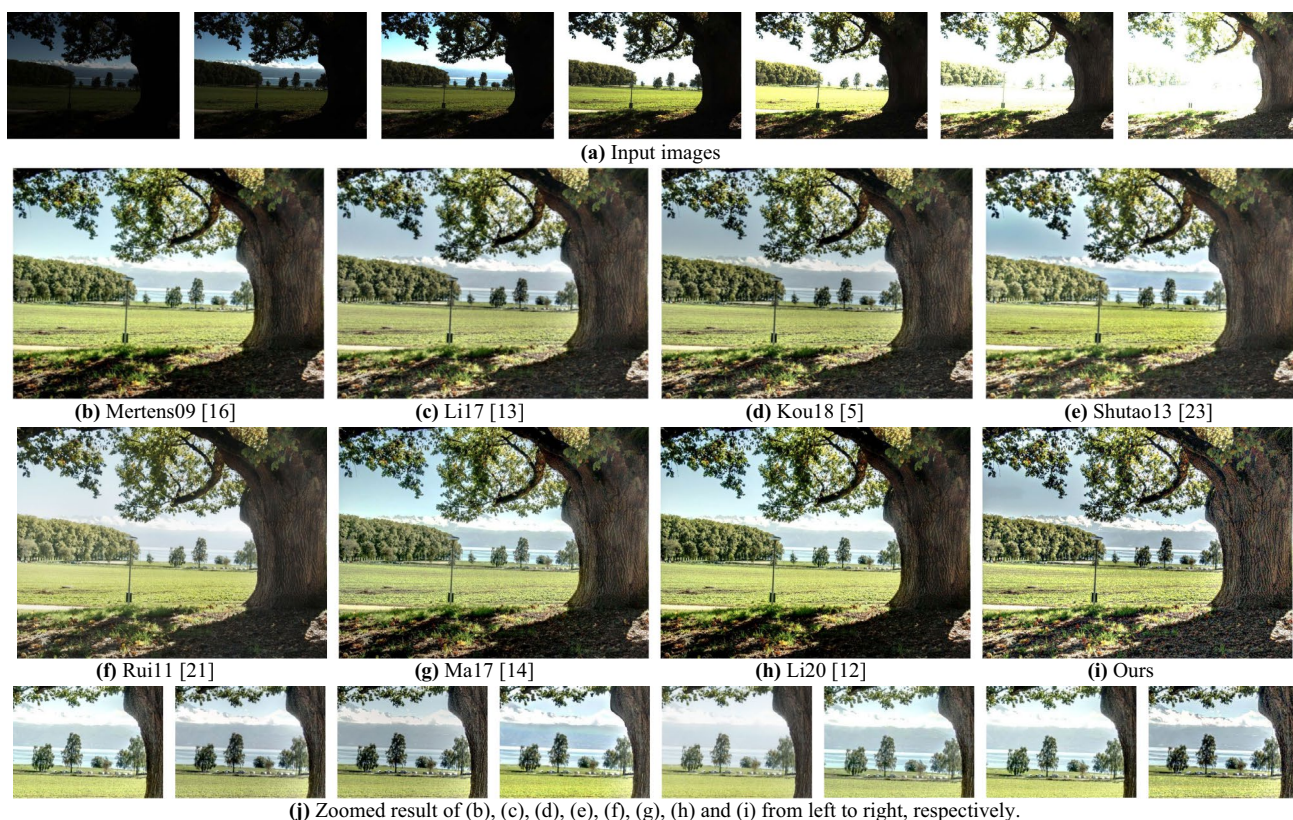**(j)** Zoomed result of (b), (c), (d), (e), (f), (g), (h) and (i) from left to right, respectively.

Fig. 5 Comparison with other exposure fusion methods [5, 12–14, 16, 21, 23] on image set Treeunil. Image courtesy of Laurence Meylan

calculated such that $g = (1 - \lambda)\gamma_j + \lambda\gamma_{j+1}$. After obtaining the values of λ and γ_j, we use the linear interpolation algorithm to calculate the corresponding pyramid coefficient in the RoE of an image by

$$L\{\text{RoE}\}_{x,y}^d = (1 - \lambda)L\left\{R(I')_{\gamma_j,\sigma}\right\}_{x,y}^d + \lambda L\left\{R(I')_{\gamma_{j+1},\sigma}\right\}_{x,y}^d \tag{8}$$

The sampling interval is recorded as q which is related to is the sampling number S. In order to reduce the calculation and increase the speed, we try to make the sampling interval q as sparse as possible without losing quality. We estimate q as same as the method [1] according to the Nyquist sampling theorem, which is half the minimum wavelength present in the signal. In practice, for our detail enhancement, the sampling interval is set as $q = 0.5$ and $S = (\max(g) - \min(g))/q$.

The experiments verified that the peak signal-to-noise ratio PSNR of the image after acceleration is higher than 30 dB, and the differences are almost invisible. With the combination of sampling and linear interpolation strategy, only the modified Laplacian pyramid coefficients in the intermediate image of the γ_j sampling values need to be computed with LLF, and for the other values in the interval $[\gamma_j, \gamma_{j+1}]$, the Laplacian pyramid coefficients are computed with simple linear interpolation. Therefore, the complexity of computing the modified

Laplacian pyramid coefficients of the intermediate image is significantly reduced to $O(S)$ because the sampling interval q is sparse enough.

3.5 Multiscale fusion

After obtaining the Laplacian pyramid of each image, we use the multiscale strategy [16] to fuse the multi-exposure images. With the same as the method [16], we also compute the initial weights $W_{x,y,k}$ for each input image by weighted multiplication of three quality measures including contrast C, saturation S and well-exposedness E as follows

$$W_{x,y,k} = \left(C_{x,y,k}\right)^{w_C} \times \left(S_{x,y,k}\right)^{w_S} \times \left(E_{x,y,k}\right)^{w_E} \tag{9}$$

where $C_{x,y,k}$, $S_{x,y,k}$ and $E_{x,y,k}$ are the contrast, saturation and well-exposedness value for the pixel (x, y) in the kth image with the corresponding weights as w_C, w_S and w_E, respectively. We use the default weight values as $w_C = w_S = w_E = 1$. Then, the initial weights $W_{x,y,k}$ are normalized to $\hat{W}_{x,y,k}$ as follows

$$\hat{W}_{x,y,k} = \frac{\sum_{k'=1}^{N} W_{x,y,k'}}{W_{x,y,k}} \tag{10}$$

(a) Input images

(b) Mertens09 [16]   **(c)** Li17 [13]   **(d)** Kou18 [5]   **(e)** Shutao13 [24]

(f) Rui11 [23]   **(g)** Ma17 [14]   **(h)** Li20 [12]   **(i)** Ours

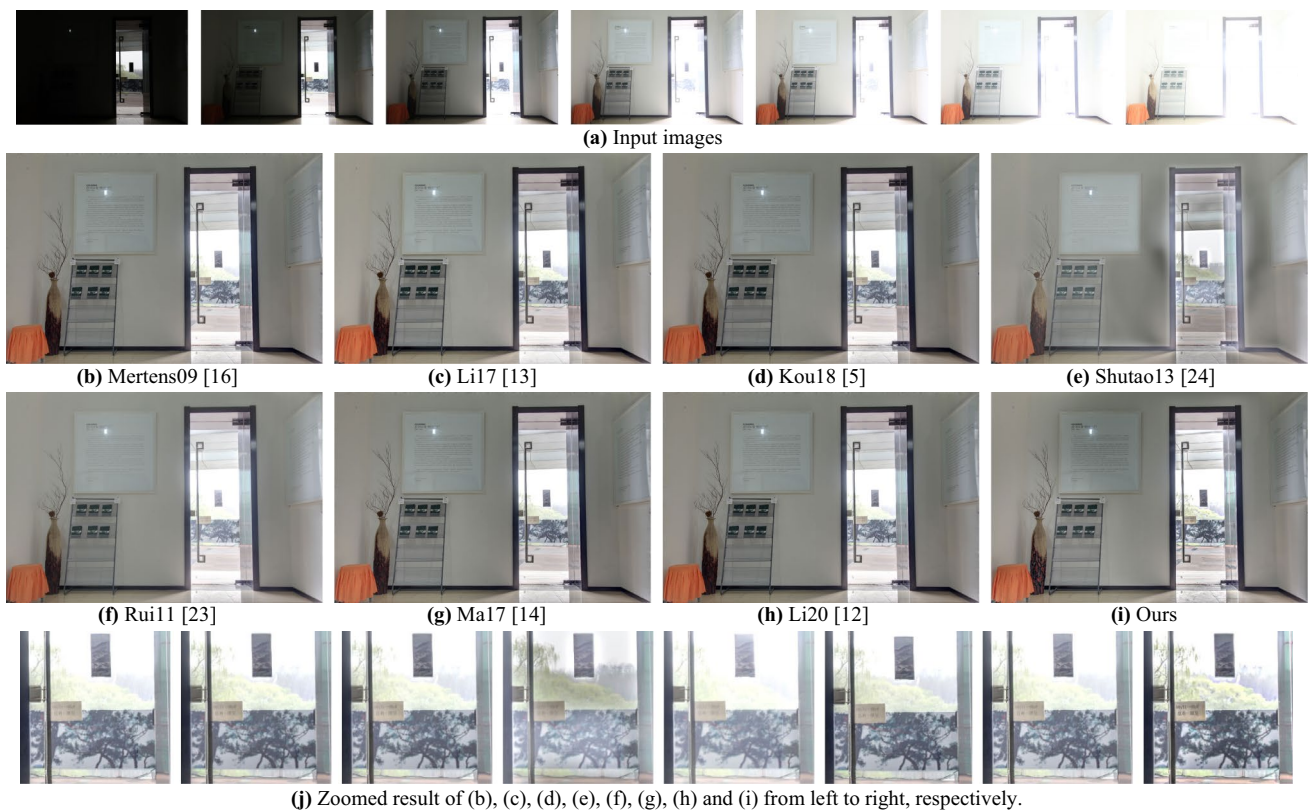**(j)** Zoomed result of (b), (c), (d), (e), (f), (g), (h) and (i) from left to right, respectively.

Fig. 6 Comparison with other exposure fusion methods [5, 12–14, 16, 21, 23] on image set ArtGallery

For multiscale fusion, we have to compute the Gaussian pyramid of the corresponding weighting map $\hat{W}_{x,y,k}$. The first level, that is the 0th Gaussian pyramid of each weighting map is noted as $G\{\hat{W}\}_{x,y,k}^{0}$ which is equal to $\hat{W}_{x,y,k}$. Then, we construct the dth ($1 \leq d \leq L$) level Gaussian pyramid of the weighting map with a down-sampled Gaussian smoothed version of the previous level $G\{\hat{W}\}_{x,y,k}^{d-1}$. And then, we fuse all the Laplacian pyramid by weighted averaging for each level as follows:

$$L\{F\}_{x,y}^{d} = \sum_{k=1}^{N} L\{I\}_{x,y,k}^{d} G\{\hat{W}\}_{x,y,k}^{d} \tag{11}$$

where $L\{I\}_{x,y,k}^{d}$ is the Laplacian pyramid of dth level and kth input image and $L\{F\}_{x,y}^{d}$ is the fused Laplacian pyramid. Finally, the fused Laplacian pyramid is collapsed to generate the final fused image F.

4 Experimental results and comparisons

We evaluate the quality performance of our algorithm by testing ten sets of differently exposed images that involve a variety of real-life scenes. All the experiments were run on

a PC with a 3.5 GHz Intel core i7-4770 k CPU and 16 GB memory. We evaluate seven state-of-the-art exposure fusion methods including [5, 12–14, 16, 21, 23] and ours by both subjective observation and objective quality measure comparison. We compare the proposed method with the other methods to prove that our method could preserve more useful details in the brightest and darkest regions. Besides, the running time comparison shows that our method is much faster than previous detail enhancement-based exposure fusion methods [5, 13, 14, 23].

4.1 Subjective observation comparison

The method [21] can fuse multi-exposure images without producing halo, but cannot preserve global contrast well and lose lots of details and colors. The algorithm [23] preserves details in the brightest and darkest regions well but it also suffers from halo artifacts, such as the wall around the drawing board in Fig. 3 and the wall around the drawing door in Fig. 6. The weighted guided image filter (WGIF)-based method [13] and the fast weighted least square (FWLS) method [5] adopted the multiscale strategy with the detail enhancement mechanism, but they did not focus on extracting details in the brightest and darkest pixels that leads to

Table 1 Objective metric MEF-SSIM values comparisons with the other exposure fusion methods on different image sets

Set name	$H \times W \times N$	[16]	[13]	[5]	[23]	[21]	[14]	[12]	Ours
Kluki	$512 \times 341 \times 3$	0.960	0.967	0.962	0.966	0.913	0.963	0.966	**0.970**
Farmhouse	$512 \times 341 \times 3$	0.975	0.978	0.979	0.973	0.926	0.984	**0.986**	0.982
Tower	$530 \times 793 \times 3$	0.986	0.990	**0.992**	0.986	0.954	0.986	0.988	0.990
Treeunil	$808 \times 600 \times 7$	0.950	0.963	0.966	0.956	0.904	0.959	0.962	**0.967**
BeligiumHouse	$1024 \times 768 \times 5$	0.974	0.977	0.972	0.956	0.945	0.973	0.977	**0.979**
SportsCentre	$2128 \times 1416 \times 5$	0.971	0.975	0.976	0.971	0.960	0.972	0.975	**0.978**
Memorial	$502 \times 716 \times 9$	0.976	0.978	0.977	**0.980**	0.958	0.977	**0.980**	0.979
ArtGallery	$1384 \times 920 \times 7$	0.971	0.969	0.971	0.958	0.923	0.971	0.974	**0.980**
Laptop	$1188 \times 792 \times 7$	0.966	0.970	0.972	0.968	0.944	0.970	0.972	**0.976**
McDonaldNight	$2816 \times 2112 \times 3$	0.983	0.985	0.986	0.973	0.964	0.984	0.987	**0.988**
Average		0.971	0.975	0.975	0.969	0.939	0.974	0.977	**0.979**

The maximum MEF-SSIM values for each set are shown in bold

Table 2 Computation time (seconds) comparisons with the other exposure fusion methods on different image sets

Set Name	$H \times W \times N$	[16]	[13]	[5]	[23]	[21]	[14]	[12]	Ours
Kluki	$512 \times 341 \times 3$	1.45	5.21	14.3	2.91	0.118	15.2	1.32	1.47
Farmhouse	$512 \times 341 \times 3$	1.42	5.19	14.2	2.89	0.104	15.6	1.30	1.44
Tower	$530 \times 793 \times 3$	3.03	9.97	28.9	6.01	0.279	29.8	2.98	3.45
Treeunil	$808 \times 600 \times 7$	10.1	23.3	58.6	20.3	0.801	60.7	9.79	10.6
BeligiumHouse	$1024 \times 768 \times 5$	8.24	23.4	96.6	19.9	0.904	102	8.11	8.54
SportsCentre	$2128 \times 1416 \times 5$	33.4	90.9	296	76.8	0.904	338	32.0	37.6
Memorial	$502 \times 716 \times 9$	7.54	24.5	68.7	18.6	0.768	89	7.32	8.05
ArtGallery	$1384 \times 920 \times 7$	21.1	66.2	180	50.7	2.82	207	20.7	25.8
Laptop	$1188 \times 792 \times 7$	15.9	47.7	121	37.1	1.60	131	15.5	17.2
McDonaldNight	$2816 \times 2112 \times 3$	40.3	116	358	104	5.12	386	38.5	48.9

detail lost in these regions and detail overenhancement for other regions. The method SPD-MEF [14] achieved more natural color visibility by treating RGB color channels of an image patch jointly, but it suffered from some halo artifacts. The method [12] used the multiscale approach to improve SPD-MEF and produce more natural result that is similar to the method [21] with fewer halos, but it still did not preserve enough details in the brightest and darkest regions.

Our method can preserve more details in the brightest and darkest regions than the other methods, such as the drawing boards on the wall in Fig. 3, the drawing boards on the wall in Fig. 4, the cloud in the sky and the grass and tree in Fig. 5 and faraway trees outside the art gallery door in Fig. 6. Our result also looks more natural without excessive detail enhancement of other regions.

4.2 Objective quality measure comparison

We use the objective metric MEF-SSIM [11] based on the principle of the structural similarity and patch structural consistency to evaluate the quality of the fused images, with results summarized in Table 1. From the table, it can be found that our fused image has the highest average structural similarity (SSIM) score of total ten sets and has slightly lower scores than [5, 23] or [12] for only three testing sets. The results of objective quality measure evaluation are consistent with those of the subjective observation evaluation.

4.3 Running time comparison

Besides comparing the visual quality of images fused by different algorithms, the complexity of different algorithms is also compared. For fair comparisons, all the algorithms are implemented in MATLAB without any special optimization on a single CPU core. The computational time of the proposed method and the comparison methods is summarized in Table 2. The multiscale strategy is based on Laplacian pyramid decomposition [16], and there are $O(\log M)$ levels in the pyramid, so the complexity of computing all the Laplacian pyramid coefficients of all the input images is $O(NM \log M)$, where M is the number of pixels in an input image and N is the number of images in the input sequence. This is the lowest complexity for the multiscale-based fusion methods because no detail enhancement mechanism is introduced. The methods in [5, 13] also adopt the multiscale

strategy, and they introduce detail enhancement mechanism. For example, method [13] uses a weighted guided image filter (WGIF) and method [5] uses the fast weighted least square (FWLS). The complexity of the two methods [5, 13] including two components, one is the computing of all the Laplacian pyramid coefficients $O(NM \log M)$, and the other is detail extraction component with the complexity $O(NM)$, and the two detail extraction components including WGIF and FWLS are very time-consuming. The method [23] is also based on edge-preserving smoothing technique but it is based on two-scale strategy. The method SPD-MEF [14] is based on structural patch decomposition that is time-consuming with the complexity $O(KNM)$, where K is the patch size. The method [12] uses the multiscale strategy and the

Fig. 7 Comparison of the fused results between using RGB vectors (**a**) and using RGB channels separately (**b**) on image set Beigium-House. Their MEF-SSIM values are both 0.979

unnormalized approximation to reduce the complexity of SPD-MEF to about $O(NM \log M)$, which is the same with [16].

Our method also uses the multiscale strategy, but the complexity of our detail enhancement mechanism is only $O(M')$, where M' is the number of pixels in the RoE, which is much fewer than M. Besides, since only a small portion of coefficients need to be processed by the FLLF based on discrete sampling, the total complexity of our method is reduced close to the complexity of [12, 16]. Our exposure fusion method is much faster than these detail enhancement-based methods [5, 13, 23] and the patch-based method [14]. The single-scale-based method [21] is much faster than our method, but the global contrast and details are not preserved well.

As referred in Sect. 3.3, we can treat RGB channels separately to get similar results and the same MEF-SSIM values with using RGB vectors as shown in Fig. 7. However, for each pyramid coefficient, the smooth function f in Eq. (6) with the time-consuming exponential operation needs to be performed three times for three channels, respectively, while

with RGB vectors, the smooth function needs to be performed only once. The execution time with RGB vectors is about 8.5 s, while the execution time with three channels is up to 12 s. Therefore, all our results are obtained by directly dealing with RGB vectors so as to improve the efficiency.

In summary, our method achieves a good trade-off between the algorithm complexity and the quality of fused images.

5 Discussion and future work

In this paper, we proposed a fast exposure fusion method with detail enhancement in the brightest and darkest regions. We defined the Region of Enhancement (RoE) for each image to reduce the pixels of detail enhancement. Besides, we adopted local Laplacian filters (LLF) to enhance details in the RoE and proposed one acceleration scheme based on discrete sampling and interpolation. Our local enhancement strategy and the fast detail enhancement mechanism improve the efficiency greatly and much faster than previous detail enhancement based exposure fusion algorithms. Besides, the proposed method could preserve more useful details in the brightest and darkest regions without excessive enhancement.

We would like to deal with the detail enhancement-based fusion process with GPU implementation to further improve the efficiency. We also would like to explore more effective detail enhancement strategy to achieve more better quality.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Aubry, M., Paris, S., Hasinoff, S.W., Kautz, J., Durand, F.: Fast local Laplacian filters: theory and applications. ACM Trans. Graph. **33**(5), 1–14 (2014)
2. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. ACM Trans. Graph. **21**(3), 257–266 (2002)
3. Farbman, Z., Fattal, R., Lischinski, D.: Edge-preservingdecompositions for multi-scale tone and detail manipulation. ACM Trans. Graph. **27**(3), 67-1-67–10 (2008)
4. Fei, K., Li, Z., Wen, C., Chen, W.: Edge-preserving smoothing pyramid based multi-scale exposure fusion. J. Vis. Commun. Image Represent. **53**, 235–244 (2018)
5. Fei, K., Zhe, W., Chen, W., Wu, X., Li, Z.: Intelligent detail enhancement for exposure fusion. IEEE Trans. Multimed. **20**(1), 484–495 (2018)
6. Jianbing, S., Ying, Z., Shuicheng, Y.: Exposure fusion using boosting laplacian pyramid. IEEE Trans. Cybern. **44**(9), 1579–1590 (2014)
7. Jianbing, S., Ying, Z., Ying, H.: Detail-preserving exposure fusion using subband architecture. Vis. Comput. **28**(5), 463–473 (2012)
8. Jianrui, C., Shuhang, G., Lei, Z.: Learning a deep single image contrast enhancer from multi-exposure images. IEEE Trans. Image Process. **27**(4), 2049–2062 (2018)
9. Kaiming, H., Jian, S., Xiaoou, T.: Guided image filtering. IEEE Trans. Pattern Anal. Mach. Intell. **35**(6), 1397–1409 (2013)
10. Karr, B.A., Debattista, K., Chalmers, A.G.: Optical effects on HDR calibration via a multiple exposure noise-based workflow. Vis. Comput. **17**, 1–16 (2020)
11. Kede, M., Kai, Z., Zhou, W.: Perceptual quality assessment for multi-exposure image fusion. IEEE Trans. Image Process. **24**(11), 3345–3356 (2015)
12. Li, H., Ma, K., Yong, H., Zhang, L.: Fast multi-scale structural patch decomposition for multi-exposure image fusion. IEEE Trans. Image Process. **29**, 5805–5816 (2020)
13. Li, Z., Wei, Z., Wen, C., Zheng, J.: Detail-enhanced multi-scale exposure fusion. IEEE Trans. Image Process. **26**(3), 1243–1252 (2017)
14. Ma, K., Li, H., Yong, H., Wang, Z., Meng, D., Zhang, L.: Robust multi-exposure image fusion: a structural patch decomposition approach. IEEE Trans. Image Process. **26**(5), 2519–2532 (2017)
15. Mansour, N., Maryam, K., S.M., Reza, S., Nader, K.: Fast exposure fusion using exposedness function. In: IEEE International Conference on Image Processing (ICIP), pp. 2234–2238 (2017)
16. Mertens, T., Kautz, J., Reeth, F.V.: Exposure fusion: a simple and practical alternative to high dynamic range photography. Comput. Graph. Forum **28**(1), 161–171 (2009)
17. Mingli, S., Dacheng, T., Chun, C., Jiajun, B., Jiebo, L., Chengqi, Z.: Probabilistic exposure fusion. IEEE Trans. Image Process. **21**(1), 341–357 (2012)
18. Paris, S., Hasinoff, S.W., Kautz, J.: Local Laplacian filters: edge-aware image processing with a Laplacian pyramid. In: International Conference on Computer Graphics and Interactive Techniques, pp. 1–12 (2011)
19. Prabhakar, K.R., Srikar, V.S., Babu, R.V.: Deepfuse: a deep unsupervised approach for exposure fusion with extreme exposure image pairs. In: IEEE International Conference on Computer Vision (2017)
20. Raman, S., Chaudhuri, S.: Bilateral filter based compositing for variable exposure photography. In: Proceedings of Eurographics Short Papers (2009)
21. Rui, S., Irene, C., Jianbo, S., Anup, B.: Generalized random walks for fusion of multi-exposure images. IEEE Trans. Image Process. **20**(12), 3634–3646 (2011)
22. Shiguang, L., Yu, Z.: Detail-preserving underexposed image enhancement via optimal weighted multi-exposure fusion. IEEE Trans. Consum. Electron. **65**(3), 303–311 (2019)
23. Shutao, L., Xudong, K., Jianwen, H.: Image fusion with guided filtering. IEEE Trans. Image Process. **22**(7), 2864–2875 (2013)
24. Wei, Z., Cham, W.K.: Gradient-directed composition of multi-exposure images. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 530–536 (2010)
25. Zhengguo, L., Jinghong, Z., Susanto, R.: Detail-enhanced exposure fusion. IEEE Trans. Image Process. **21**(11), 4672–4676 (2012)
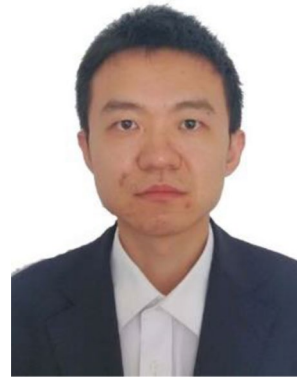
Chunmeng Wang is currently a lecturer at School of Computer Engineering, Jinling Institute of Technology, Nanjing, China. He received his Ph.D. degree from the School of Computer Science and Technology of Shandong University in 2015. His research interests include computer graphics, image processing, high dynamic range image.

Chen He is currently an associate professor at Media and Communication College, Weifang University, China. He received his Ph.D. degree in digital media technology from Shandong University in 2014. His research interests include computer graphics, digital image processing, digital geometry processing and data compression.

Minfeng Xu is a lecturer in the School of Computer Science and Technology at Shandong University of Finance and Economics. He received his Ph.D. degree from the School of Computer Science and Technology of Shandong University in 2019. His research interests include computer graphics and digital geometry processing.