

Fast Multi-exposure Image Fusion with Median Filter and Recursive Filter

Shutao Li, *Member*, IEEE and Xudong Kang

Abstract — This paper proposes a weighted sum based multi-exposure image fusion method which consists of two main steps: three image features composed of local contrast, brightness and color dissimilarity are first measured to estimate the weight maps refined by recursive filtering. Then, the fused image is constructed by weighted sum of source images. The main advantage of the proposed method lies in a recursive filter based weight map refinement step which is able to obtain accurate weight maps for image fusion. Another advantage is that a novel histogram equalization and median filter based motion detection method is proposed for fusing multi-exposure images in dynamic scenes which contain motion objects. Furthermore, the proposed method is quite fast and thus can be directly used for most consumer cameras. Experimental results demonstrate the superiority of the proposed method in terms of subjective and objective evaluation.¹

Index Terms — Multi-exposure, dynamic scene, median filter, recursive filter.

I. INTRODUCTION

Images taken by ordinary digital cameras usually suffer from a lack of details in the under-exposed and over-exposed areas if the camera has a low or high exposure setting. High dynamic range (HDR) imaging solves this problem by taking multiple images at different exposure levels and merging them together. This technique has been widely used in digital camera and mobile phone devices. Generally speaking, existing HDR imaging approaches can be divided into two categories: tone mapping based methods and image fusion based methods.

Tone mapping based methods consist of two main steps: HDR image construction and tone mapping. Multiple low dynamic range (LDR) photographs are first captured and combined together to construct a HDR image [2]. Then, through using tone mapping techniques [3], the overall contrast of the HDR image is reduced to facilitate display of HDR images on devices with lower dynamic range. This two-phase workflow i.e., HDR image construction and tone mapping, can generate a tone mapped image where all areas appear well exposed. Many effective tone mapping methods have been proposed [4]-[6]. For example, Kim *et al.* proposes a tone mapping method based on Retinex model [4]. Kuang *et*

al. propose a new image appearance model and use it for tone mapping [7]. However, this kind of methods is usually time consuming and thus not very qualified for ordinary digital cameras.

Different from tone mapping based methods, image fusion based methods i.e., multi-exposure image fusion, can bypass the HDR image construction process and directly yield a tone mapped-like fused image. Multi-exposure image fusion is preferred for consumer electronic applications since it does not require the HDR image construction process which increases some computing cost. Many multi-exposure image fusion methods have been proposed. For instance, Mertens *et al.* [8] propose a multi-scale image fusion framework. The source images are first decomposed as Laplacian pyramids and then blended together in each level to construct the fused image. However, the performance of this kind of methods may be unsatisfactory when the decomposition level is too large or too small. To solve this problem, this paper proposes a weighted sum based image fusion method without multi-scale analysis.

The basic assumption of most existing multi-exposure fusion methods is that the scene is static during different captures. However, while fusing images taken in dynamic scenes which contain camera movement or motion objects, the methods mentioned above may produce serious distortions. To remove the impacts of camera movement, many multi-exposure image alignment methods have been proposed [9]-[11]. Furthermore, this problem can be reduced by taking images using a tripod. Besides the camera movement, the more challenging problem is caused by motion objects which may appear in the fused image as ghost artifacts. Various solutions have been proposed such as variance based [1], gradient orientation based [12], and Maximum-a-posteriori (MAP) based methods [13]. However, these methods usually require a user assigned reference image for motion detection or are quite time consuming. To solve these problems, a histogram equalization and median filter based motion detection method for fusing images in dynamic scenes is proposed.

This paper proposes a simple yet effective multi-exposure image fusion method. A recursive filter [14] based weight map refinement method is adopted to obtain accurate weight maps for weighted sum based image fusion. Furthermore, the color dissimilarity between pixels of source images and pixels of the scene's static background is considered for fusing images taken in dynamic scenes. The proposed image fusion approach is very fast and thus is quite suitable for consumer

¹ S. Li and X. Kang are with the College of Electrical and Information Engineering, Hunan University, Changsha, China, 410082. (e-mail: shutao_li@yahoo.com.cn and xudong_kang@163.com)

cameras. Experiments on images captured at both static and dynamic scenes demonstrate the superiority of the proposed method in terms of objective and subjective evaluation. Furthermore, it is shown that the proposed method can be applied for other image fusion applications.

The rest of this paper is organized as follows. Section II explains the proposed algorithm in detail. Experimental results and comparisons are presented in Section III. Finally, conclusions are given in Section IV.

II. FUSION OF MULTI-EXPOSURE IMAGES WITH MEDIAN FILTER AND RECURSIVE FILTER

Fig. 1 shows the schematic diagram of the proposed weighted sum based image fusion method. The weights of pixels of different source images are first estimated, and then refined by recursive filtering with the corresponding source image serving as the reference image. Finally, the fused image is constructed by weighted sum of source images.

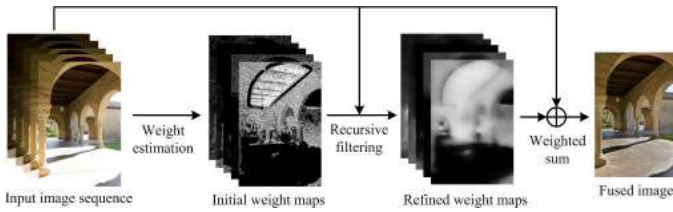


Fig. 1. Schematic diagram of the proposed multi-exposure image fusion method.

A. Local Contrast and Brightness

When fusing images in static scenes without motion objects, two image features i.e., local contrast and brightness should be considered for weight estimation. The local contrast of each pixel is calculated as follows:

$$A_n(x, y) = \hat{I}_n(x, y) * h(x, y), \quad (1)$$

where $*$ denotes convolution operation, \hat{I}_n is the gray image (calculated by weighted sum of red, green, and blue channel $\hat{I}_n = 0.299 \times I_n^r + 0.587 \times I_n^g + 0.114 \times I_n^b$ [15]) of the n th source image I_n , and h is a high-pass filter defined as:

$$h = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \quad (2)$$

Once the local contrast of each pixel is obtained, it is ranked in a winner-take-all manner as follows:

$$\hat{A}_n(x, y) = \begin{cases} 1 & \text{if } A_n(x, y) = \max \{A_n(x, y), n = 1, 2, \dots, N\} \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where N is the number of source images and \hat{A}_n is the resulting local contrast feature which aims at preserving image details.

The brightness of each pixel is used to decide whether a pixel is under-exposed or over-exposed. This feature ensures that the fused image will not be constructed by pixels from under-exposed, over-exposed areas. Based on the fact that

pixels with very weak or strong brightness are usually under-exposed or over-exposed, the under-exposed and over-exposed pixels can be detected as follows:

$$B_n(x, y) = \begin{cases} 1 & T < \hat{I}_n(x, y) < 255 - T \\ 0 & \text{otherwise} \end{cases}, \quad (4)$$

where T is a threshold value which is suggested to be an integral value between 10-30. This assumption has been widely used in many related publications [12], [13], [16] to remove the impacts of under-exposed and over-exposed pixels. For the proposed method, the variation of T in 10-30 has little influence on the fusion performance.

B. Color Dissimilarity

When fusing images in dynamic scenes which contain motion objects, as well as the local contrast and brightness feature, the influence of motion objects should also be considered for weight estimation. Through measuring the color dissimilarity between pixels of source images and pixels of the scene's static background, a novel histogram equalization and median filter based motion detection method is proposed. Fig. 2 shows the schematic diagram of the proposed motion detection method.

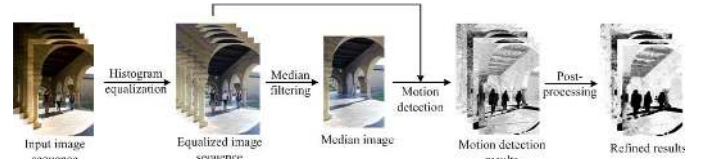


Fig. 2. Schematic diagram of the proposed median filter based motion detection.

Firstly, histogram equalization [15] is performed on each source image so that the color distributions of input images can be transformed into a similar color distribution (see Fig. 3).



Fig. 3. Median image generation with histogram equalization and median filtering. Top-left: the original "Arch" LDR image sequence. Bottom-left: the histogram equalized "Arch" LDR sequence. Right: the median image of histogram equalized LDR image sequence.

Then, median filtering is performed in the time domain to obtain the median image (see Fig. 3) of the histogram equalized image sequence as follows:

$$I^M(x, y) = \text{median} \{I_n^E(x, y)\} \quad n = 1, 2, \dots, N, \quad (5)$$

where I_n^E is the n th histogram equalized image and I^M is the median image i.e., the scene's static background image. The temporal median filter has been widely used for motion

detection in video applications [17]. It is based on the assumption that motion objects usually appear less times than pixels from the static background. Here, the temporal median filter is used to obtain the scene's static background image so that the motion objects of each image can be easily detected by calculating the color dissimilarity between each histogram equalized image and the scene's static background image as follows:

$$\tilde{C}_n(x, y) = \exp\left(\frac{(I_n^E(x, y) - I_n^M(x, y))^2}{\delta^2}\right), \quad (6)$$

where δ equals 0.1 controlling the curvature of the Gauss curve presented in (6).

Finally, \tilde{C}_n is refined by morphological operators (dilation followed by erosion) to remove noise estimation:

$$C_n = (\tilde{C}_n \oplus s_1) \ominus s_2, \quad (7)$$

where s_1 and s_2 are disk-like structure elements with radius r_1 and r_2 respectively, \oplus and \ominus denote dilation and erosion operation respectively, and C_n is the resulting color dissimilarity feature which indicates where is the motion objects. The main advantage of the proposed motion detection method is that the user does not need to assign a reference image for motion detection. As shown in Fig. 2, the walking people are accurately detected by the proposed method.

C. Weight Estimation

In order to preserve image details and remove influences of under-exposed pixels, over-exposed pixels, and pixels from motion objects, the three image features i.e., local contrast, brightness, and color dissimilarity should be combined together for weight estimation. The straightforward way to this objective is by multiplication. However, the pixels of the same location of different LDR images may be all labeled as under-exposed, over-exposed or motion objects, and this is unreasonable especially when these pixels appear in a large number. To solve this problem, the brightness feature and color dissimilarity feature are first combined together by multiplication:

$$D_n = B_n \times C_n. \quad (8)$$

Then, D_n are normalized such that they sum to one at each pixel (x, y) .

$$\tilde{D}_n(x, y) = \left[\sum_{m=1}^N D_m(x, y) \right]^{-1} D_n(x, y), \quad (9)$$

Next, only these pixels under average score i.e., $1/N$ are labeled as zero so that the pixels of the same location of different LDR images will not be all labeled as under-exposed, over-exposed or motion objects.

$$\hat{D}_n(x, y) = \begin{cases} 0 & \tilde{D}_n(x, y) < 1/N \\ 1 & \text{otherwise} \end{cases}, \quad (10)$$

where N is the number of source images and $\hat{D}_n(x, y) = 0$ means that the pixel (x, y) of the n th source image is under-

exposed, over-exposed, or from the motion objects. Finally, \hat{D}_n and the local contrast feature \hat{A}_n are combined together to estimate the weights:

$$\hat{W}_n = \hat{A}_n \times \hat{D}_n \quad (11)$$

It should be noticed that when it is known that the scene is static, the color dissimilarity feature C_n in (8) should not be considered.

D. Weight Refinement and Weighted Fusion

As shown in Fig. 1, the weights estimated above are noisy and hard (most weights are either 0 or 1). So the weight maps should be refined for weighted sum based image fusion. This is possible due to the recently proposed recursive filter [14], which is a real-time edge-preserving smoothing filter. Here, recursive filtering is performed on the weight maps \hat{W}_n with the corresponding source image I_n serving as the reference image.

$$W_n = R(\hat{W}_n, I_n), \quad (12)$$

where R denotes the recursive filtering operation. The weight map refinement step with recursive filter is based on two simple assumptions: first, pixels from the same objects which have similar image color should have similar weights; second, smooth weights are preferred since it will not introduce seam artifacts in the resulting fused image. The two assumptions can be easily satisfied with the recursive filter which is defined as follows:

$$J[k] = (1 - a^d)I(k) + a^d J[k-1], \quad (13)$$

where $a \in [0, 1]$ is a feedback coefficient, $I[k]$ is the value of the k th pixel of the input weight map, $J[k]$ is the k th pixel of the refined weight map, d is the distance between neighborhood pixels of the source LDR image [14]. As d increases, a^d goes to zero, stopping the propagation chain and thus pixels with similar colors tend to have similar smoothed weight. The detailed description of the real-time recursive filter can be found in Gatal and Oliveira's paper [14]. As shown in Figs. 1, 5(a), 6(a), 8(a), and 9(a), the proposed weight map refinement step can transform the hard and noise weight maps into accurate and smooth weight maps. These weight maps can well indicate where are well exposed and static in each LDR image.

Once the resulting weight maps W_n are obtained, the resulting fused image I^F can be directly calculated as follows:

$$I^F = \sum_{n=1}^N I_n \times W_n. \quad (14)$$

Furthermore, to reduce the computing and memory consumption, all the weight maps can be computed at a half resolution of the original image size and then up-sampled to the original size for image fusion. This acceleration scheme has little influence to the performance of the proposed method.

III. EXPERIMENTS AND DISCUSSIONS

A. Experimental Setup

Ten multi-exposure image sequences are used in the experiments. Three of them are taken in dynamic scenes and the other seven are taken in static scenes. Four HDR imaging methods i.e., Reinhard *et al.*'s method [3], Kuang *et al.*'s method named as iCAM06 [7], Photomatix [18], and Gallo *et al.*'s de-ghosting method [16] are used for comparison. For Reinhard *et al.*'s method, the source code is publically available (<http://user.cs.tu-berlin.de/~eitz/hdr/>). The results of iCAM06 are generated by the Matlab implementation downloaded from their homepage (<http://www.cis.rut.edu/Mcsl/icam/hdr/>). The results of Gallo *et al.*'s method are taken from its homepage (<http://users.soe.ucsc.edu/~orazio/deghost>). The software package of Photomatix is available online (<http://www.hdrsoft.com/>).

To assess the fusion performance of different methods objectively, three objective fusion quality metrics: Q_0 [19], $Q^{AB/F}$ [20], and visual information fidelity (VIF) [21] are adopted. The index Q_0 is designed through modeling image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. The index $Q^{AB/F}$ evaluates the success of edge information transferred from the source images to the fused image. The quality metric VIF is derived from a statistical model for natural scenes. It can accurately quantify the distortions and the improvements in visual quality. The larger the $Q^{AB/F}$, Q_0 , and VIF value are, the better the fusion results are.

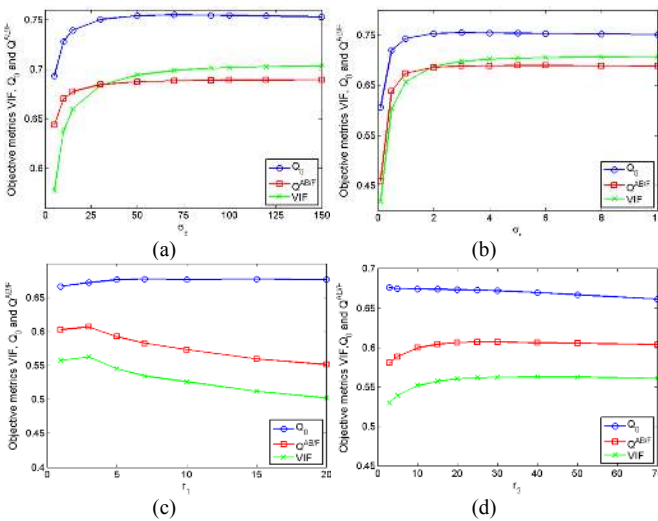


Fig. 4. The variation of the VIF, Q_0 , and $Q^{AB/F}$ on different σ_s (a), σ_r (b), r_1 (c), and r_2 (d).

B. Analysis of Free Parameters

This sub-section analyzes the influence of the parameters: σ_s , σ_r , r_1 , and r_2 which respectively control the space, range supports of the recursive filter, the radius of the structural elements s_1 and s_2 [see (7)]. The fusion performance is evaluated by the value of VIF, Q_0 , and $Q^{AB/F}$ (see Fig. 4).

σ_s is first analyzed with $\sigma_r = 4$, $r_1 = 3$, and $r_2 = 30$. Then, in the analysis of σ_r , we fix $\sigma_s = 100$, $r_1 = 3$, $r_2 = 30$. Next, r_1 and r_2 are analyzed in the same way. As shown in Fig. 4, the proposed method generates the best fusion performance when $\sigma_s = 100$, $\sigma_r = 4$, $r_1 = 3$, and $r_2 = 30$. So, $\sigma_s = 100$, $\sigma_r = 4$, $r_1 = 3$, and $r_2 = 30$ are set as the default parameters. This parameter setting can generate good subjective performance for most images.

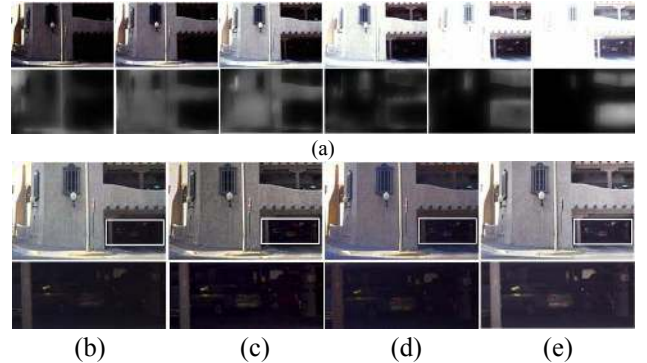


Fig. 5. The “Garage” LDR image sequence (up) and the resulting weight maps of the proposed method (bottom) (a), the results produced by Reinhard *et al.*'s method [3] (b), iCAM06 [7] (c), Photomatix [18] (d), and the proposed method (e). (Image courtesy of CAVE lab.)

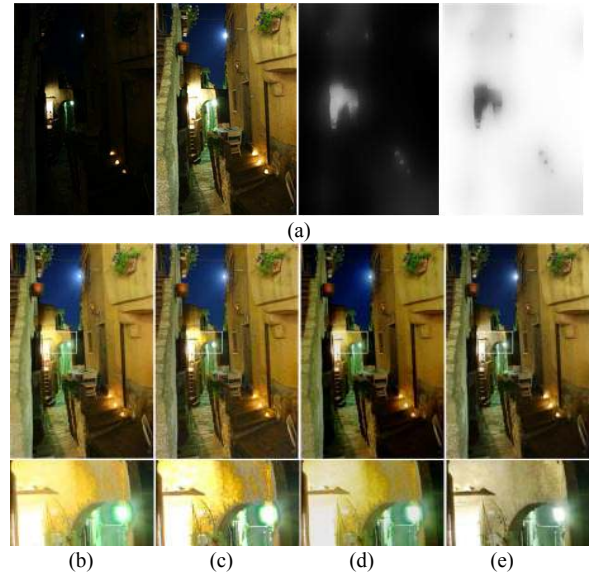


Fig. 6. The “Via” LDR image sequence (left) and the resulting weight maps of the proposed method (right) (a), the results produced by Reinhard *et al.*'s method [3] (b), iCAM06 [7] (c), Photomatix [18] (d), and the proposed method (e). (Image courtesy of easyHDR.)

C. Results

Figs. 5 and 6 show the comparison of the proposed method with three high dynamic imaging methods for fusing images in static scenes. Figs. 5(a) and 6(a) show the source LDR image sequences and the resulting weight maps estimated by the proposed method. As is shown, the weight maps accurately indicate where are well exposed in each LDR image. The last row of the two figures gives a close-up view of the garage region and the light region. It can be seen that

the results produced by Reinhard *et al.*'s method, iCAM06, and Photomatrix [see Figs. 5(b-d) and 6(b-d)] suffer different levels of under-exposure and over-exposure in the garage region and the light area respectively. By contrast, the result produced by the proposed method shows the best visual appearance in these regions [see Figs. 5(e) and 6(e)].

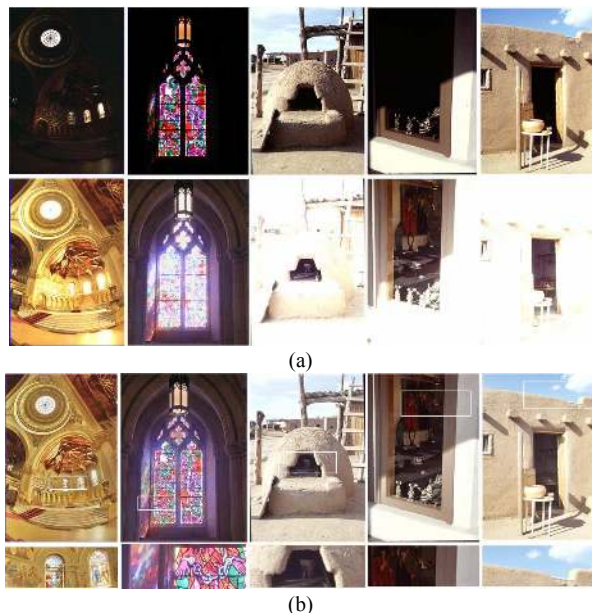


Fig. 7. From left to right: selected image pairs of the “Memorial church”, “National cathedral”, “Igloo”, “Window”, and “Door” image sequences (a). The fused images produced by the proposed method (b).

Fig. 7 presents more results of the proposed method on other classic LDR image sequences. Fig. 7(a) shows 5 image pairs of the “Memorial church” image sequence (the original sequence is of size $512 \times 768 \times 16$) provided by Paul Debevec, “National cathedral” image sequence ($1536 \times 2048 \times 2$) provided by Max Lyons, “Igloo”, “Window”, and “Door” image sequences ($348 \times 222 \times 6$, $226 \times 341 \times 5$, and $231 \times 338 \times 6$ respectively) which are all provided by the CAVE Lab. The proposed method is able to obtain high quality fused images for different LDR image sequences [see the first row of Fig. 8(b)]. Furthermore, from the magnified images in the second row of Fig. 8(b), it can be seen that the details in dark areas (see the inside area of the “igloo” and “window” images) and bright areas (see the window area of the “church” and “cathedral” images and the sky area of the “door” image) are all well-preserved in the fused images.

To further demonstrate the effectiveness of the proposed method, experiments are also performed on multi-exposure images captured in dynamic scenes. Fig. 8 shows an example of HDR imaging in dynamic scenes. As is shown, the results produced by Reinhard *et al.*'s method and iCAM06 which do not have de-ghosting function [see Fig. 8 (b) and (c)] will produce serious ghosting artifacts. Although Photomatrix has the de-ghosting function, it is not very effective in this example [see Fig. 8(c)]. By contrast, the proposed image fusion method can effectively remove the influence of pixels from the motion objects [see the bottom of Fig. 8(a)]. At the

same time, the details of source images are well preserved in the fused image [see Fig. 8(e)].

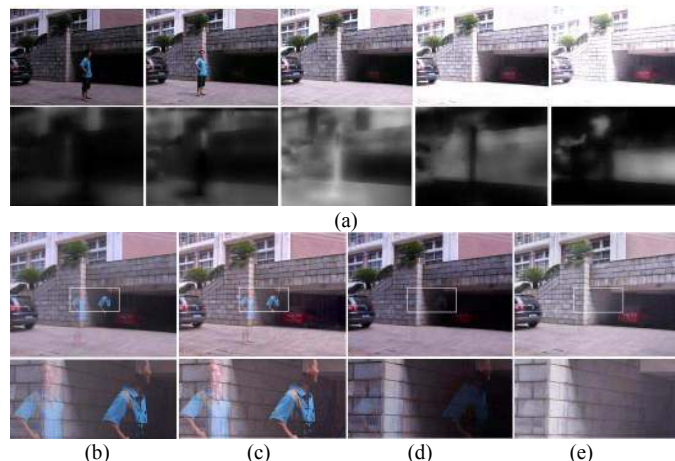


Fig. 8. The “Park” LDR image sequence(up) and the resulting weight maps of the proposed method (bottom) (a), the results produced by Reinhard *et al.*'s method [3] (b), iCAM06 [7] (c), Photomatrix [18] (d), and the proposed method (e).

The experiments are also performed on publically available dynamic image sequences, i.e., Figs. 9 and 3. The first row of Fig. 9(a) shows the “Forest” LDR image sequence which have moving people and waving branches. The first row of Fig. 3 shows the “Arch” LDR image sequence which contains moving people. The results produced by Reinhard *et al.*'s method, Photomatrix, Gallo *et al.*'s method, and the proposed method are presented in Figs. 9(b)–(e) and 10(a)–(d) respectively. It can be seen that although Photomatrix has the de-ghosting function, some ghosting artifacts is still visible [see Figs. 9(c) and 10(b)]. Gallo *et al.*'s method can effectively remove ghost artifacts while introducing blocking artifacts in the floor area [see Fig. 10(c)]. By contrast, the proposed method can effectively remove all ghosting artifacts and well preserve the details of different LDR images [see Figs. 9(d) and 10(d)].

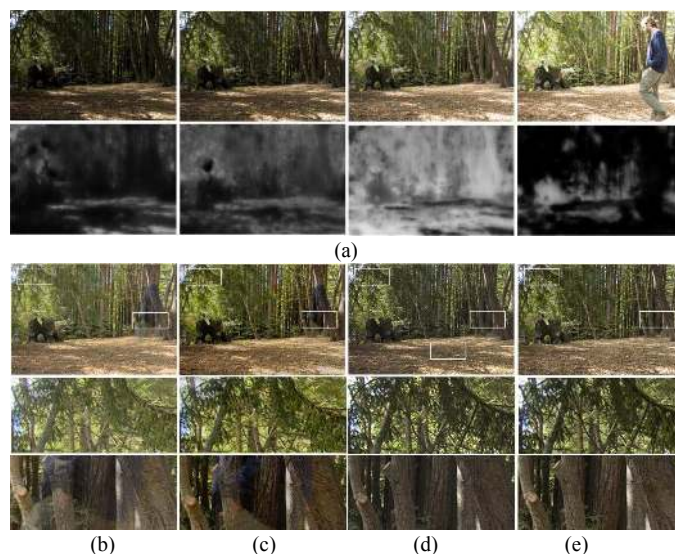


Fig. 9. The “Forest” LDR image sequence(up) and the resulting weight maps of the proposed method (bottom) (a), the results produced by Reinhard *et al.*'s method [3] (b), Photomatrix [18] (c), Gallo *et al.*'s method [16] (d), and the proposed method (e). (Image courtesy of Orazio Gallo [16].)

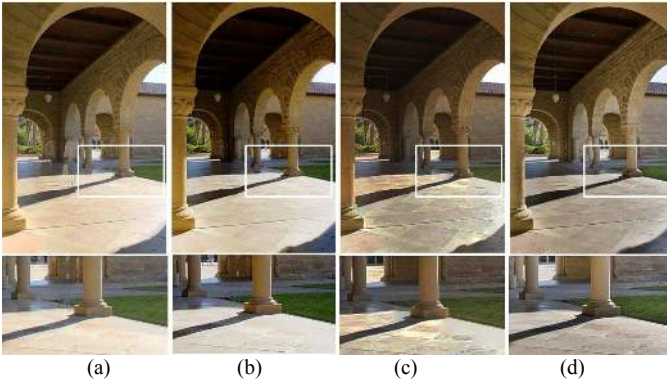


Fig. 10. The “Arch” example. The results produced by Reinhard *et al.*'s method [3] (a), Photomatix [18] (b), Gallo *et al.*'s method [16] (c), and the proposed method (d). (Image courtesy of Orazio Gallo [16].)

D. Objective Performance Comparison

In this sub-section, the objective performances of different methods are evaluated. Table I gives the values of Q_0 , $Q^{AB/F}$, and VIF for the fused results by Reinhard *et al.*'s method [3], Photomatix [18], iCAM06 [7], and the proposed method. From this table, it can be observed that the proposed image fusion method provides the best performance in terms of the largest Q_0 , $Q^{AB/F}$, and VIF. It can be concluded that the proposed image fusion method can better preserve the edge information of source images (see the performance of $Q^{AB/F}$) and introduces less color, contrast, and image quality distortions (see the performances of Q_0 and VIF).

TABLE I
QUANTITATIVE ASSESSMENT OF DIFFERENT HDR IMAGING METHODS.

Images	Indexes	Methods			
		Reinhard [3]	iCAM06 [7]	Photomatix [18]	Proposed
Garage	Q_0	0.555	0.576	0.469	0.675
	$Q^{AB/F}$	0.520	0.561	0.441	0.674
	VIF	0.467	0.443	0.377	0.522
Via	Q_0	0.805	0.696	0.643	0.942
	$Q^{AB/F}$	0.750	0.694	0.691	0.838
	VIF	0.987	0.979	0.911	1.105
Park	Q_0	0.577	0.664	0.586	0.723
	$Q^{AB/F}$	0.571	0.651	0.619	0.684
	VIF	0.592	0.642	0.624	0.676
Forest	Q_0	0.493	0.543	0.514	0.570
	$Q^{AB/F}$	0.384	0.407	0.374	0.519
	VIF	0.234	0.235	0.224	0.368
Arch	Q_0	0.615	0.602	0.558	0.722
	$Q^{AB/F}$	0.539	0.544	0.510	0.614
	VIF	0.540	0.581	0.538	0.641

TABLE II
CONSUMING TIME OF DIFFERENT METHODS FOR PROCESSING THE “MEMORIAL CHURCH” IMAGE SEQUENCE OF SIZE $512 \times 768 \times 16$ (s).

Methods	Reinhard [3]	iCAM06 [7]	Proposed (Matlab)	Proposed (C++)
Consuming Time	5.32	11.08	3.51	0.32

Besides the fusion performance, computing efficiency is also very important in real applications. Table II gives the comparison of different methods on computational efficiency.

The tone mapping algorithms and the proposed image fusion algorithm are all implemented using Matlab. The HDR images used for tone mapping algorithms are generated by using the Matlab implementation of Debevec *et al.*'s method [2] which takes about 4.03 seconds for the Memorial church image sequence. This time is included into the consuming time of Reinhard *et al.*'s method and iCAM06 for fair comparison. All experiments are simulated on a CPU 2.8 GHz PC with 2048 MB RAM. From Table II, compared with Reinhard *et al.*'s method and iCAM06, the proposed method is more computational efficient. Furthermore, by implementing the proposed method in C++, the computing time of the proposed method can be reduced to 0.32 seconds. A real-time implementation of the recursive filter [14], a speedup of ten times compared to our implementation, can be used for further acceleration. Thus, it is concluded that the proposed method is a time efficient multi-exposure image fusion method and thus it can be directly used for most consumer digital cameras.

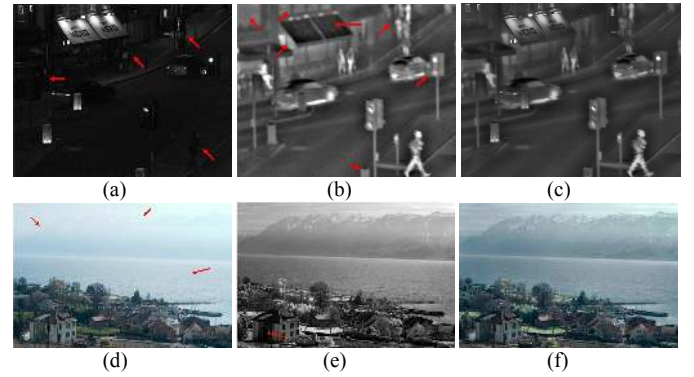


Fig. 11. Two examples of infrared and visual image fusion. The visual image (a), the infrared image (b), the fused image of (a) and (b) by the proposed method (c), the visual image (d), the near-infrared image (e) and the fused image of (d) and (e) by the proposed method (f). (Image courtesy of Lex Schaul [22].)

E. Other Applications

The proposed method can be also used for other image fusion applications such as infrared and visual image fusion. Fig. 11(a) and (b) shows a visual image and an infrared image. In the area directed by the red arrows, people are unclear in the visual image and background details are unclear in the infrared image. The fused image produced by the proposed method is shown in Fig. 11(c). It can be seen that the walking people and background details are all clear in the fused image. Furthermore, Fig. 11(d)–(e) shows the other example of infrared and visual image fusion for image de-hazing. Fig. 11(d) is captured through using ordinary digital camera. Fig. 11(e) is produced through using a near-infrared camera which is able to capture the details behind the haze. The fused image shown in Fig. 11(f) has a de-hazing effect.

IV. CONCLUSIONS

In this paper, a fast and effective multi-exposure image fusion approach is proposed. A novel histogram equalization and median filter based motion detection method is developed

for fusion of multi-exposure images in dynamic scenes. Furthermore, a recursive filter based weight map refinement method making full use of the color consistency between nearby pixels is adopted for weight refinement. Experiments demonstrate that proposed method can create high visual quality tone-mapped-like fused images in both dynamic scenes and static scenes. Furthermore, the effectiveness of the proposed method is demonstrated by using objective image fusion quality indexes. The proposed method has been applied for infrared and visual image fusion. In the future, whether the proposed method can be applied for fusion of multi-focus images in dynamic scenes can be further researched.

ACKNOWLEDGMENT

This paper is supported by the National Natural Science Foundation of China (No. 61172161).

REFERENCES

- [1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*, San Francisco, Elsevier, 2010.
- [2] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH*, Aug. 1997, pp. 369–378.
- [3] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," in *Proc. ACM SIGGRAPH*, Jul. 2002, pp. 267–276.
- [4] K. Kim, J. Bae, and J. Kim, "Natural HDR Image tone mapping based on retinex," *IEEE Trans. Consumer Electronics*, vol. 57, no. 4, pp. 1807–1814, Nov. 2011.
- [5] J. Lee, R. Park, and S. Chang, "Tone mapping using color correction function and image decomposition in high dynamic range imaging," *IEEE Trans. Consumer Electronics*, vol. 56, no. 4, pp. 2772–2780, May 2010.
- [6] Y. Monobe, H. Yamashita, T. Kurosawa, and H. Kotera, "Dynamic range compression preserving local image contrast for digital video camera," *IEEE Trans. Consumer Electronics*, vol. 51, no. 1, pp. 1–10, Feb. 2005.
- [7] J. Kuang, G. M. Johnson, and M. D. Fairchild, "iCAM06: A refined image appearance model for HDR image rendering," *J. Vis. Commun. Image Representation*, vol. 18, no. 5, pp. 406–414, June 2007.
- [8] T. Mertens, J. Kautz, and F. V. Reeth, "Exposure fusion: a simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, Mar. 2009.
- [9] J. Im, S. Lee, and J. Paik, "Improved elastic registration for ghost artifact free high dynamic range imaging," *IEEE Trans. Consumer Electronics*, vol. 57, no. 2, pp. 932–935, May 2011.
- [10] J. Im, J. Jeon, M. H. Hayes and J. Paik, "Single image-based ghost-free high dynamic range imaging using local histogram stretching and spatially-adaptive denoising," *IEEE Trans. Consumer Electronics*, vol. 57, no. 4, pp. 1478–1484, Nov. 2011.
- [11] W. Kao, C. Hsu, L. Chen, C. Kao, and S. Chen, "Integrating image fusion and motion stabilization for capturing still images in high dynamic range scenes," *IEEE Trans. Consumer Electronics*, vol. 52, no. 3, pp. 735–741, Aug. 2006.
- [12] W. Zhang and W. K. Cham, "Gradient-directed multi-exposure composition," *IEEE Trans. Image Process*, vol. 21, no. 4, pp. 2318–2323, Apr. 2012.
- [13] 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.
- [14] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge-aware image and video processing," *ACM Trans. Graphics*, vol. 30, no. 4, pp. 69:1–69:11, Jul. 2011.
- [15] R.C. Gonzalez, R.E. Woods, and S. Eddins, *Digital image processing using MATLAB*. New York, USA: Prentice Hall, 2004.
- [16] O. Gallo, N. Gelfand, W. Chen, M. Tico, and Kari Pulli, "Artifact-free high dynamic range imaging," in *Proc. IEEE Int. Conf. Computational Photography*, Apr. 2009, pp. 1–7.
- [17] R. Li, S. Yu, and X. Yang, "Efficient spatio-temporal segmentation for extracting moving objects in video sequences," *IEEE Trans. Consumer Electronics*, vol. 53, no. 3, pp. 1161–1167, Aug. 2007.
- [18] G. Joffre, W. puech, F. Comby, and J. Joffre, "High dynamic range images from digital cameras raw data," in *Proc. ACM SIGGRAPH*, Aug. 2005, pp. 72:1.
- [19] Z. Wang and A. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, Mar. 2002.
- [20] C.S. Xydeas and V.S. Petrović, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, Feb. 2000.
- [21] H.R. Sheikh and A.C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
- [22] L. Schaul, C. Fredembach, and S. Süsstrunk, "Color image dehazing using the near-infrared," in *Proc. IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 1629–1632.

BIOGRAPHIES



Shutao Li (M'07) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Hunan University, Changsha, China, in 1995, 1997, and 2001, respectively. In 2001, he joined the College of Electrical and Information Engineering, Hunan University. From May 2001 to October 2001, he was a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology, Kowloon, Hong Kong. From November 2002 to November 2003, he was a Postdoctoral Fellow with the Royal Holloway College, University of London, Egham, U.K., working with Prof. J.-S. Taylor. From April 2005 to June 2005, he was a Visiting Professor with the Department of Computer Science, Hong Kong University of Science and Technology. He is currently a Full Professor with the College of Electrical and Information Engineering, Hunan University. He has authored or coauthored more than 140 refereed papers. His professional interests are information fusion, pattern recognition, bioinformatics, and image processing. Dr. Li served as a member of the Neural Networks Technical Committee from 2007 to 2008. He has won two Second-Grade National Awards at the Science and Technology Progress of China in 2004 and 2006.



Xudong Kang received the B.S. degree in College of Information Science and Engineering from Northeast University, Shenyang, China, in 2007. He is currently pursuing the Ph.D. degree in electrical engineering at the Hunan University, Changsha, China. He is engaged in image fusion, image super-resolution and remote sensing image classification.