FAST EXPOSURE FUSION USING EXPOSEDNESS FUNCTION

Mansour Nejati[1], Maryam Karimi[1], S.M. Reza Soroushmehr[2], Nader Karimi[1], Shadrokh Samavi[1,3], Kayvan Najarian[2,3]

[1]Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran
[2]Department of Emergency Medicine, University of Michigan, Ann Arbor, 48109, U.S.A.
[3]Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, 48109, U.S.A.

ABSTRACT

We propose a fast and effective method for multi-exposure image fusion. Our method blends multiple exposures under a base-detail decomposition of input images. Construction of blending weights in the proposed method is performed based on an exposedness function using luminance component of the input images. The fused base layer and detail layer are integrated into the final fused image which its detail strength is simply controlled through the integration process. Experimental results demonstrate that the proposed exposure fusion method is much faster than competing methods and can achieve state-of-the-art performance objectively and perceptually.

Index Terms— Multi-exposure fusion, image enhancement, exposedness function, guided filter, quality measure.

1. INTRODUCTION

When the radiation across a scene changes seriously, digital cameras usually take images that suffer from lack of details in the under-exposed and over-exposed regions because of their limited dynamic range. Multi-exposure image fusion (MEF) is an effective technique to solve this problem by taking several images of the same scene under different exposure levels and fusing them together. This technique is more efficient than the traditional high dynamic range (HDR) imaging approaches [1-3] which involve an intermediate HDR image construction process and may need to calibration of the camera response curve. Thus, MEF methods are more preferable for consumer electronic applications because of saving the computational cost and minimizing manual interactions.

In recent years, many MEF algorithms have been developed [4-16]. These methods mainly provide a weighted summation of input exposure images as the final fused image. The weight factor is spatially different proportional to structural details and visual importance. Different techniques were adopted by various MEF methods to find the weight maps. Most of them are based on some quality measures defined mainly in terms of local image characteristics. Mertens *et al.* [4] proposed a multi-scale fusion scheme in which the contrast, color saturation, and well-exposure measures are used for computing the fusion weights. Strength of image details extracted based on bilateral filter is employed in [5] as weights to guide the fusion process. The MEF method presented in [6] used Laplacian filtering for construction of the weight maps which are then refined using guided filter [17]. A relevant work in [7] applied edge-preserving recursive filter to generate the refined and edge aligned weight maps.

The gradient information of images in some cases is employed to compute the contribution of different input sources into final reconstructed image. In this regard, Gu *et al.* [8] used the gradient fields derived from structure tensor of input images and modified it iteratively to compute the final weights. A similar algorithm in [9] combined the input image sequence with the guidance of a gradient-based quality measure. Recently, Liu and Wang [10] applied dense scale invariant feature transform (SIFT) [18] for obtaining the contrast and spatial consistency weights based on local gradient information. Another recently proposed MEF method [11] decomposes each image patch into mean intensity, signal strength and signal structure. The final components are determined based on patch strength and exposedness measures. A sparse representation based exposure fusion method is also proposed in [12] where the fusion rule is applied on the sparse coefficients. The fused coefficients finally reconstruct the fused image.

In some of the existing works, MEF is incorporated into an optimization framework. These methods estimate the weight maps by solving an energy function. A generalized random walk framework proposed in [13] finds an optimal balance between local contrast and color constancy while combining the scene details. The fusion scheme in [14] employed a probabilistic model to preserve the maximum visible contrasts and the gradient consistency in the synthesized final image. MEF algorithm in [15] derived the optimal fusion weights by maximum-a-posteriori (MAP) estimation in a hierarchical multivariate Gaussian conditional random field model based on perceived local contrast and color saturation. In [16], fine details of input images are extracted by solving a quadratic optimization problem and then added to an intermediate reconstructed image to make it sharper.

In this paper, a simple yet effective multi-exposure image fusion method is proposed. This method exploits a fast two-scale image decomposition approach to separate each input image into base and detail layers. We employ

luminance component of input source images to extract the exposedness features which are mapped to blending weights for base layers and detail layers by an exposedness function. The final image is synthesized by combining the fused base layer and the fused detail layer. Our experiments demonstrate that the proposed method leads to compelling fusion results in both visual and objective quality. Furthermore, the proposed method is quite fast which makes it suitable for consumer cameras.

The rest of this paper is organized as follows. Section 2 presents the proposed fusion method in detail. Experimental results and comparisons are presented in Section 3. Finally, conclusions are given in Section 4.

2. PROPOSED FUSION METHOD

Let input multi-exposure images be denoted as I_k, $k = 1, \dots, K$. For each input image I_k, we first compute its luminance component L_k by weighted sum of red, green and blue channels of I_k. The luminance values are then used for obtaining the fusion weights. In our method, the input images are decomposed into two-scale representations where fusion rules are applied. The aim of this decomposition step is to separate each input image into a smooth base layer containing the large-scale intensity variations and a detail layer having the small scale details.

The base layer can be obtained by applying an average filter to the luminance component. However, this can lead to halo artifacts in the final reconstructed image due to smoothing of edges along with the details. To reduce this problem, we employ guided filter [16] as an efficient edge-preserving smoothing filter. Let $G_{r,\epsilon}(P, Q)$ denote the guided filtering operator where r is radius of the filter and ϵ controls its blurring degree. Also, P and Q indicate the input image and guidance image, respectively. Using this filter, the base layer of each input image is obtained as:

$$B_k = G_{r,\epsilon}(L_k, L_k) \tag{1}$$

where L_k serves as both input and guidance image. Given the base layer, the detail layer is easily computed as:

$$D_k = I_k - B_k, \;\; k = 1, \dots, K \tag{2}$$

where D_k contains the detail layer for red, green and blue channels of I_k. The exposure fusion process is guided by weight maps for the base and detail layer of each input image. In the proposed method, the weights are calculated using a general exposedness function f defined as:

$$f(\varphi, \sigma) = \exp\left(-\frac{(\varphi - C_e)^2}{2\sigma^2}\right) \tag{3}$$

where φ denotes exposedness feature and σ controls the spread of the Gaussian. Also, C_e indicates well-exposedness constant which is normally set to middle of the intensity range. The calculated weight based on the exposedness function in (3) shows how the exposedness feature is close to its desirable value C_e.

For the detail layer, we compute the exposedness feature at each pixel position as average luminance in a small local neighborhood. Indeed, the exposure quality of image details is assessed based on mean level of local intensity variations. Accordingly, for each location (x, y) in the detail layer of k^{th} input image, the weight $W_k^D(x, y)$ is obtained as:

$$W_k^D(x, y) = \exp\left(-\frac{(\varphi_k^D(x, y) - 0.5)^2}{2\sigma_D^2}\right) \tag{4}$$

where φ_k^D is simply computed by convolving the luminance component L_k with a 7×7 average filter.

The base layer, produced based on (1), mainly conveys large-scale structural information of image luminance. To obtain blending weights of this layer, we consider the exposure quality of local and global luminance. Local mean value of image luminance at each pixel position can be used as an exposedness feature for local exposure quality assessment. However, in order to have structural consistency between the base layer and its weight map, the base layer itself is used as the exposedness feature. Thus, for the local exposure weight $W_k^{B,l}(x, y)$ of the k^{th} input image we have:

$$W_k^{B,l}(x, y) = \exp\left(-\frac{(B_k(x, y) - 0.5)^2}{2\sigma_l^2}\right) \tag{5}$$

The global exposure weight $W_k^{B,g}(x, y)$ of the k^{th} input image at pixel position (x, y) is calculated in a similar way, except that the mean luminance value of entire image, $\bar{L}_k$, is used in exposure quality assessment:

$$W_k^{B,g}(x, y) = \exp\left(-\frac{(\bar{L}_k - 0.5)^2}{2\sigma_g^2}\right) \tag{6}$$

Then, the weights $W_k^{B,l}$ and $W_k^{B,g}$ at each pixel position (x, y) are combined together as

$$W_k^B(x, y) = W_k^{B,l}(x, y).W_k^{B,g}(x, y) \tag{7}$$

where W_k^B denotes the weight map for the base layer of k^{th} input image. After construction of the weight maps for all input images, they are normalized to have unit sum at each pixel position. Finally, weighted average of the base layers and the detail layers of input images are combined into the fused image F as follows:

$$F = \sum_{k=1}^{K} W_k^B B_k + \alpha \sum_{k=1}^{K} W_k^D D_k \tag{8}$$

where $\alpha \geq 1$ controls the detail strength and thus local contrast of the resulting fused image F.

3. EXPERIMENTAL RESULTS

We tested the proposed method on several multi-exposure image sequences with different numbers of exposure levels. In our experiments, the parameters of the proposed method are set to $r = 12$, $\epsilon = 0.25$, $\sigma_g = 0.2$, $\sigma_l = 0.5$, $\sigma_D = 0.12$, and $\alpha = 1.1$. These parameter values were chosen based on

Table 1. Performance evaluation of the proposed method against existing MEF methods in terms of the objective IQA model in [19]. Top two results for each source sequence are shown in bold.

Source Sequence	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[14]	[16]	[4]	Proposed
Balloons	0.768	0.948	0.945	0.913	0.945	0.927	0.963	0.918	0.941	**0.969**	**0.971**
Belgium House	0.810	0.964	0.947	0.896	0.948	0.930	0.965	0.873	0.954	**0.971**	**0.973**
Cadik Lamp	0.729	0.929	0.931	0.875	0.948	0.903	0.945	0.943	0.945	**0.969**	**0.961**
Candle	0.842	0.869	0.938	0.899	**0.965**	0.937	0.953	0.960	0.965	**0.971**	0.964
Cave	0.694	0.978	0.961	0.934	0.952	0.948	**0.980**	0.918	0.923	0.974	**0.979**
Chinese garden	0.911	0.984	0.982	0.927	0.978	0.972	0.988	0.967	0.951	**0.989**	**0.991**
Farmhouse	0.877	**0.985**	0.977	0.932	0.972	0.967	0.983	0.947	0.959	0.981	**0.983**
House	0.770	0.957	0.921	0.876	**0.959**	0.927	0.945	0.948	0.935	**0.965**	0.950
Kluki	0.902	0.968	0.965	0.922	0.969	0.959	0.961	0.939	0.949	**0.980**	**0.972**
Lamp	0.864	0.934	0.937	0.871	0.932	0.926	0.945	0.829	0.933	**0.948**	**0.949**
Landscape	0.954	0.942	0.972	0.941	0.984	0.978	**0.991**	0.944	0.948	0.976	**0.992**
Lighthouse	0.938	0.950	0.953	0.934	0.972	0.965	0.969	0.971	0.968	**0.980**	**0.975**
MadisonCapitol	0.763	0.968	0.918	0.864	0.952	0.928	0.974	0.945	0.949	**0.977**	**0.979**
Memorial	0.617	0.965	0.945	0.871	0.935	0.901	0.940	0.964	0.947	**0.967**	**0.966**
Office	0.907	0.967	0.972	0.900	0.968	0.967	**0.986**	0.961	0.954	0.984	**0.988**
Tower	0.895	**0.986**	0.984	0.932	0.974	0.976	0.981	0.939	0.950	**0.986**	**0.986**
Venice	0.892	0.954	0.952	0.889	0.962	0.954	**0.978**	0.942	0.937	0.966	**0.976**
Average	0.831	0.956	0.953	0.904	0.960	0.945	0.967	0.936	0.948	**0.974**	**0.974**

the fusion performance. We empirically found that our method is robust to variations of the guided filter parameters r and ϵ, although the increase in r leads to an increase in local contrast of the fused image.

For objective performance evaluations, a recently proposed MEF quality assessment model [19] is adopted. The results in [19] show that this quality measure outperforms all the existing MEF quality assessment measures in terms of correlation with subjective judgments. Using this quality measure, we compared the fusion performance of the proposed method with ten existing MEF algorithms including Mertens09 [4], Raman09 [5], Li13 [6], Li12 [7], Gu12 [8], Zhang12 [9], Liu15 [10], Ma15 [11], Song12 [14], and GuoLi12 [16]. The comparison results on a set of 17 multi-exposure image sequences[1] are listed in Table 1. The quality scores are in the range of [0,1] where a higher value is indicative of a better quality. It can be observed that the proposed method achieves the best or second best performance in most sequences in terms of the quality measure of [19]. On the average, our method and Mertens09 [4] have the best performance.

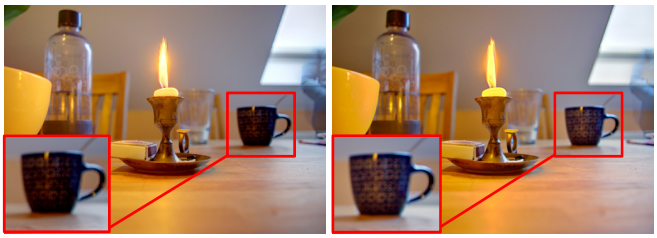
Visual comparison of the fusion results produced by the proposed method and Mertens09 [4] on the "Candle" and "Kluki" sequences is shown in Fig. 1. Although, the result of Mertens09 [4] on both of these sequences has better objective quality compared to result of our method in Table 1, our method achieves better visual appearance with more detail preservation and vivid colors. As can be observed in Fig. 1(b), our method better preserves the texture details on the cup and provides more natural colors on the table and around the candle. Moreover, in Fig. 1 (d) more structural details are visible on the cottage and trees. The bright and vivid green color in the meadow regions and the leaves of trees is also a significant advantage.

In Fig. 2, we compare Li12 [7] with the proposed method on the "MadisonCapitol" sequence. The weakness of the method Li12 [7] is clearly visible in Fig. 2(a). This fusion algorithm cannot properly make use of color information resulting in dreary color appearance in the fusion output. Moreover, the global luminance strongly changes in the final result where the center part of the image is clearly darker than the bottom and right parts. By contrast, the color information and details are well-preserved in the result of our method which leads to better visual appearance.

Fig. 3 compares Zhang12 [9] with the proposed method on the "Balloons" sequence. We observe that the former method fails to produce a well-exposed composite image and the result suffers from low brightness. On the contrary, the proposed method does a better job resulting in higher contrast and more natural and vivid color appearance on the sky, balloons, and the meadow regions. Fig. 4 shows the fused images obtained by Liu15 [10] and our method on the "Belgium House" sequence. The result of Liu15 [10] shows dark halo in the wall regions and door frames which degrades its visual quality. Compared to Liu15 [10], our method gives a higher contrast image with sharper details (for example on the floor) and the overall appearance of the fused image is more appealing. In Fig. 5, our method is compared with Song12 [14] on the "Memorial" sequence. It can be seen in Fig. 5(a) that the fused image produced by Song12 [14] exhibits over-exposure in the window regions resulting in detail loss. In comparison, the proposed method reveals more details especially in the window and ceiling areas and offers more contrast to the texture of the walls.

In addition to the fusion quality, since MEF algorithms are expected to be embedded in digital cameras, the fusion speed is a very important parameter. Therefore, we compare the fusion speed of our proposed method with that of the best MEF algorithms with available source codes. Table 2 gives the average computational time of the proposed method over 17 multi-exposure image sequences compared with the methods of [4], [6], [7], [10] and [11]. All the compared methods have MATLAB implementations and all experiments are performed on a laptop with 2.6 GHz Intel

[1] Available at: http://ivc.uwaterloo.ca/database/MEF/



(a) Mertens09 [4] (b) Proposed

(c) Mertens09 [4] (d) Proposed

Fig. 1. Comparison of the proposed method with Mertens09 [4] on the sequences: (a),(b) "Candle", and (c),(d) "Kluki".

(a) Li12 [7] (b) Proposed

Fig. 2. Comparison of the proposed method with Li12 [7]. Source image sequence courtesy of Chaman Singh Verma.

(a) Zhang12 [9] (b) Proposed

Fig. 3. Comparison of the proposed method with Zhang12 [9]. Source image sequence courtesy of Erik Reinhard.

(a) Liu15 [10] (b) Proposed

Fig. 4. Comparison of the proposed method with Liu15 [10]. Source image sequence courtesy of Dani Lischinski.

(a) Song12 [14] (b) Proposed

Fig. 5. Comparison of the proposed method with Song12 [14]. Source image sequence courtesy of P. Debevec.

Core i7 CPU[2]. From Table 2, it can be seen that the proposed method is much more computationally efficient than the other ones. The closest competitor to our proposed method is Mertens09 [4] which is more than 3 times slower. Overall, our method not only achieves state-of-the-art fusion performance according to both visual comparison and objective assessment but also is quite fast. Thus, it is a reliable MEF method for real applications.

Table 2. Execution time of different MEF algorithms averaged over a set of 17 multi-exposure sequences.

Method	[4]	[6]	[7]	[10]	[11]	Proposed
Average time (sec)	0.676	1.165	2.315	2.232	2.685	**0.218**

4. CONCLUSION

In this paper, we presented a new method for fast and effective fusion of multi-exposure sequences into high quality images. The proposed method decomposes input images into base and detail representations for which the blending weights are efficiently computed using an exposedness function. Without any post-processing step, the computed weights are directly applied to guide the fusion process. Our method was compared against several recently proposed multi-exposure fusion techniques. Objective evaluations based on a state-of-the-art quality assessment model demonstrated the effectiveness of the proposed method. Also, visual comparisons showed the advantages of our method over the other ones in terms of detail preservation, output contrast and vivid color appearance. Furthermore, our method is computationally efficient and fuses images much faster than the competing methods making it qualified for real applications.

[2] Source code of our algorithm can be accessed at: http://mansournejati.ece.iut.ac.ir/content/exposure-fusion

5. REFERENCES

[1] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 3, pp. 267–276, 2002.

[2] J. Kuang, G. M. Johnson, and M. D. Fairchild, "iCAM06: A refined image appearance model for HDR image rendering," *Journal of Visual Communication and Image Representation*, vol. 18, no. 5, pp. 406–414, 2007.

[3] L. Meylan, and S. Susstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Transactions on image processing*, vol. 15, no. 9, pp. 2820–2830, 2006.

[4] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," in *Computer Graphics Forum,* vol. 28, no. 1, pp. 161–171*,* Blackwell Publishing Ltd, 2009.

[5] S. Raman, and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proc. Eurographics*, March 2009.

[6] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864–2875, 2013.

[7] S. Li, and X. Kang, "Fast multi-exposure image fusion with median filter and recursive filter," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 626–632, 2012.

[8] B. Gu, W. Li, J. Wong, M. Zhu, and M. Wang, "Gradient field multi-exposure images fusion for high dynamic range image visualization," *Journal of Visual Communication and Image Representation*, vol. 23, no. 4, pp. 604–610, 2012.

[9] W. Zhang, and W. K. Cham, "Gradient-directed multiexposure composition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2318–2323, 2012.

[10] Y. Liu and Z. Wang, "Dense SIFT for ghost-free multi-exposure fusion," *Journal of Visual Communication and Image Representation*, vol. 31, pp.208–224, 2015.

[11] K. Ma and Z. Wang, "Multi-exposure image fusion: A patch-wise approach," in *IEEE International Conference on Image Processing (ICIP)*, 2015*,* pp. 1717–1721.

[12] J. Wang, H. Liu, and N. He, "Exposure fusion based on sparse representation using approximate K-SVD," *Neurocomputing*, vol. 135, pp.145–154, 2014.

[13] R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized random walks for fusion of multi-exposure images," *IEEE Transactions on Image Processing*, vol. 20, no. 12, pp.3634–3646, 2011.

[14] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, and C. Zhang, "Probabilistic exposure fusion," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 341–357, 2012.

[15] R. Shen, I. Cheng, and A. Basu, "QoE-based multi-exposure fusion in hierarchical multivariate Gaussian CRF," *IEEE Transactions on Image Processing*, vol 22, no. 6, pp.2469–2478, 2013.

[16] Z. G. Li, J. H. Zheng, and S. Rahardja, "Detail-enhanced exposure fusion," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4672–4676, 2012.

[17] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. European Conference on Computer Vision (ECCV)*, Greece, Sep. 2010, pp. 1–14.

[18] C. Liu, J. Yuen, and A. Torralba, "Sift flow: dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 978–994, 2011.

[19] K. Ma, K. Zeng and Z. Wang, "Perceptual Quality Assessment for Multi-Exposure Image Fusion," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345–3356, 2015.