A Layered-Based Exposure Fusion Algorithm

Journal:	*IET Image Processing*
Manuscript ID:	IPR-2012-0494
Manuscript Type:	Research Paper
Date Submitted by the Author:	04-Sep-2012
Complete List of Authors:	Li, Xiaoguang Li, FengHui Zhuo, Li Feng, Dagan; University of Sydney, School of Information Technologies;
Keyword:	IMAGE PROCESSING, IMAGE REPRESENTATION, IMAGE ENHANCEMENT

A Layered-Based Exposure Fusion Algorithm

Xiaoguang Li[1], Fenghui Li[1], Li Zhuo[1] and David Dagan Feng[2]

[1] Signal & Information Processing Lab., Beijing University of Technology, Beijing, China

lxg@bjut.edu.cn, lifenghui@emails.bjut.edu.cn, zhuoli@bjut.edu.cn

[2] School of Information Technologies, The University of Sydney

feng@it.usyd.edu.au

ABSTRACT

Due to the limitation of dynamic range, a single still image is usually insufficient to describe a high contrast scene. Fusing multi-exposure images of the same scene can produce a resulting image with details both in bright and dark regions. However, they may be sensitive to the exposure parameters of the input images. To improve the robustness of the method, a novel layered-based exposure fusion algorithm is proposed in this paper. In our algorithm, a global-layer is introduced to improve the robustness of the fusion method. The global-layer is employed to preserve the overall luminance of a real scene and avoid possible luminance reversion. Then details are recovered in gradient domain by a Poisson solver. Experimental results show the superior performance of our approach in terms of robustness and color consistency.

Index Terms: Image fusion; Multi-exposure High dynamic range; Exposure fusion; Poisson edit

1. INTRODUCTION

Real-world scenes contain a vast range of luminance, which the magnitude spans over 9 orders (10^{-4} to 10^{5} cd/m^2). The human visual system can perceive about 5 orders of luminance magnitude simultaneously, and is capable of adapting to these lighting conditions. However, the current standard image capture and display devices cannot directly reproduce such high contrasts. The High Dynamic Range (HDR) imaging techniques are developed to capture and display the full visible dynamic luminance in real world scenes. HDR image can be synthesized from a bracketed exposure sequence. While, tone mapping operators should be applied in order to display them in standard displayers.

Exposure fusion [1] is proposed to generate a high quality low dynamic range (LDR) image with HDR details.

A. Vavilin et al. [2] proposed a simple and fast exposure fusion approach based on exposure blending. Using three differently exposed images as input, pixels in the output image are computed as a weighted averaging of the input pixels. The fusion weights are formulated as a simple exponential function. The variable of the function depends on local feathers of each pixel in the input images. There are three feature extraction methods viz. pixel-based method, local contrast-based method and bilateral-based method are taken into account. There are several arguments should be carefully adjusted, or halo artifacts will appear near strong edges of the output image.

A. R. Varkonyi-Koczy et al. [3] developed a multiple exposure image synthesization technique, where the red, green and blue color components of the pixels are processed independently. In their method, the input images are tiled into blocks. For each block, the summation of gradient magnitudes is calculated as a character of details. Then, the resulting image is composed to the blocks with the maximum detailed feature. To remove the sharp transitions along borders of the region, a Gaussian blending function is employed. The results of this method are impressive. However, the fusion process with Gaussian blending function is very time consuming.

K.H. Jo et al [4] segment the target resulting image into different arbitrary shaped regions. That is achieved by an intensity clustering and region merging scheme applied in the input images. Then, gradients (details) and entropies are computed to select the best input exposure image for each region. Therefore, for each input image, there will be a binary weight mask to indentify that whether a pixel will appear in the resulting image. To deal with the halo artifacts appearing near the border of regions, the weight masks are filtered by a bilateral filter. Compared with the block-based methods, using regions of arbitrary shape allows more precise detail recovering on the borders of the regions with high contrast changing. However, this algorithm may be sensitive to the differences of the exposure values.

Recently, several probabilistic models for exposure fusion have been proposed. R. Shen et al [5] consider the exposure fusion as a probabilistic composition process. A probabilistic model is proposed in their paper. They aim to achieve an optimal balance between local contrast and color consistency. A generalized random framework is introduced to solve the optimal problem. M Song et al [6] formulate the exposure fusion problem with a maximum a posteriori (MAP) estimation problem. In their work,

local adaptive detail and scene gradient are extracted and utilized in the specification of the probabilistic terms in a MAP framework. One is to make sure the most visible details can be preserved in the resulting image, and another is to correct the gradient reversion. The probabilistic models for exposure fusion are easy to incorporate with many priors about the ideal target HDR scene. However, they usually suffer from expensive time complexity to solve the optimal problems.

The above-mentioned works have made significant contributions to the way we now exploit layered-based exposure fusion algorithm. However, most of these existing algorithms only considered local information in the estimation of the fusion weights. Therefore, these algorithms will limit to deal with input images which exposure differences are relatively small. In other words, the results of local-based algorithms may be sensitive to the value of exposure differences.

In this paper, Exposure Fusion refers to the algorithms that fuse multi-exposure images to a LDR image. The resulting LDR images can depict the HDR scene more efficiently, like the result of a tone mapping operator. The input multiple exposure images of a static scene are supposed to be exactly pixel alimented. As mentioned above, most existing approaches calculate the resulting pixels as a weighted average of the pixels those locate at the same position of the input images. The main difference between them is weights estimation model. Feather extraction is usually involved in the weights estimators. The basic idea is that the most important or salient features should be preserved in the output image. Therefore, the pixels with salient features will be weighted significantly.

We present a novel layered-based exposure fusion algorithm. A global layer of the target HDR scene is considered in our fusion procedure. The global layer which is an estimation of the overall luminance of the high dynamic range scene will help to avoid the reversion of the bright regions and the relative less bright regions in the output image. Inspired by gradient domain image cloning and Poisson image edit techniques [7,8,9], a scheme for detailed-layer fusion is introduced in gradient domain.

This paper makes the following contributions: (1) The global information is introduced into the exposure fusion. It will improve the robustness of the fusion results with a great exposure differences. (2) A gradient domain fusion scheme is introduced to fuse the detailed layer and the global layer efficiently.

2. THE PROPOSED ALGORTIHM

2.1. Overview

Fig. 1 shows the flowchart of the proposed layered-based exposure fusion algorithm. Input images are used to estimate a global layer. The global layer is an estimation of the overall luminance of a HDR scene. In the global layer, the dark and bright regions are represented as a smoothing map. Then, the global layer is clustered and segmented to different regions. Those pixels with similar intensities and located within a neighborhood will be high likely to fall into the same region. This step is aim to segment the background with different luminance levels. For each region or luminance level, there will be a best exposure image to describe its details. The best exposure selector will return an index for each region, which denotes the best exposure input image for the current region. Therefore, an exposure index map will be generated in this step. A detailed layer is computed from the input images with the guide of the exposure index map. Finally, the global layer and the detailed layers are fused in gradient domain. Finally, the resulting image will be obtained by solving a Poisson problem.

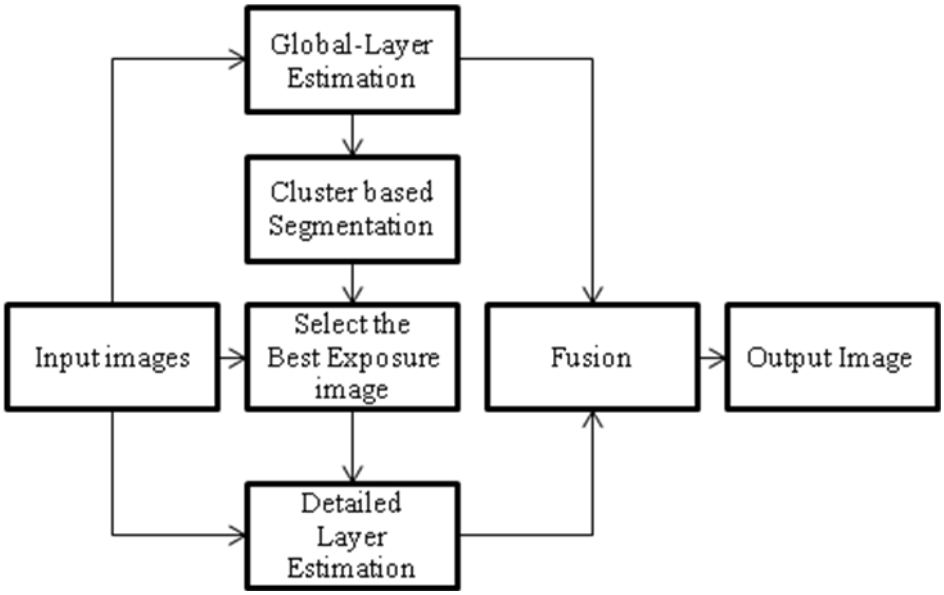


Fig. 1 The flowchart of the proposed algorithm

2.2. Global-Layer Estimation

The global layer is calculated to estimate the overall luminance levels of a HDR scene. It is used to keep the consistency of bright, dark and normal luminance levels between the target output image and the real scene. Given a single exposure input image, its luminance level should not reverse to the real scene, although some gradient details may be lost in saturation regions due to inappropriate exposure. Based on this observation, the global layer is estimated by average of all the input images.

Then, a bilateral [10] filter is employed to the averaging image. Compared to a Gaussian filter, the bilateral filter can preserve the strong edges in the output image by averaging those pixels having a similar intensity and located within a neighborhood of the pixel under consideration. Therefore, those regions near the strong edges will not be over-smoothed. This bilateral filtering can smooth regions with different luminance more accurately.

Let I_i ($i = -m,\ldots,-1, 0, 1,\ldots, m$) denotes the ith input image. As shown in Fig. 2, each image is separated by 1 f-stop. I_0 is a normal exposure image. Therefore, the global layer can be computed as Eq. (1) and (2):

$$I_{avg}(x,y) = \frac{1}{2m+1}\sum_{i=-m}^{m} I_i(x,y) \quad , \tag{1}$$

$$I_g = F_b(I_{avg}), \tag{2}$$

where, x and y are accordance of image pixels, I_{avg} is the averaging image, F_b denotes a bilateral filter and I_g is the global layer.

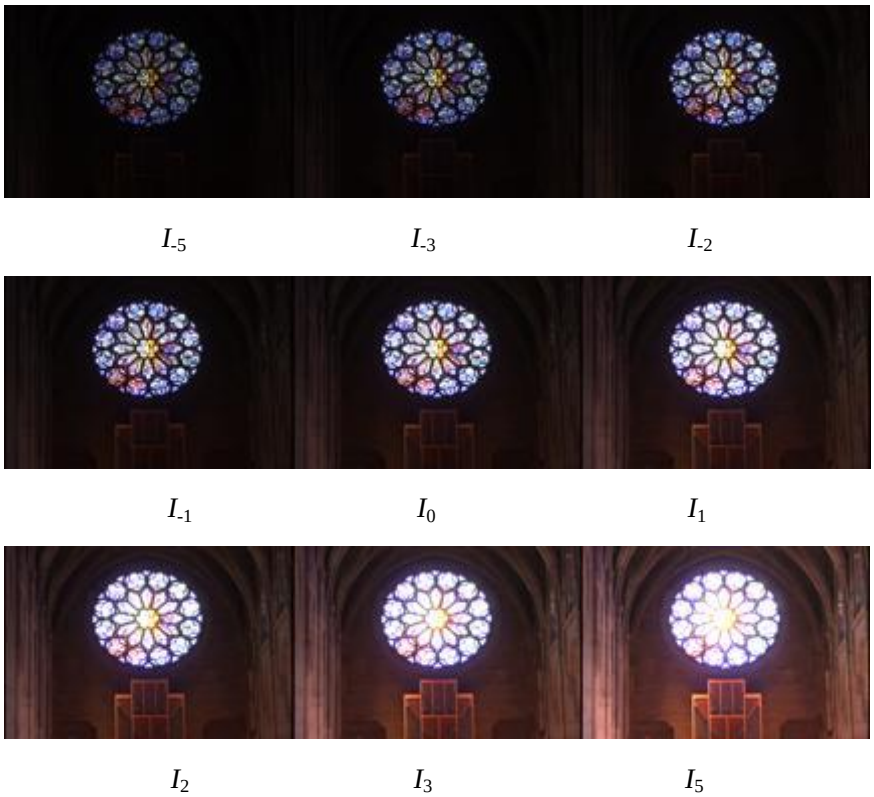


Fig. 2 Example exposure sequence (courtesy of Paul Debevec [11])

The bilateral filter is a non-linear filter whose output is a weighted average of the input. The weight of a pixel depends not only on a spatial kernel f, but also on a function g in the intensity domain. The output of the bilateral filter for a pixel at position q with an intensity of I_q is given as follows:

$$F_b(I_q) = \frac{1}{k(q)} \sum_{p \in \Omega} f(p-q) g(I_p - I_q) I_p, \quad (3)$$

where $k(q) = \sum_{p \in \Omega} f(p-q) g(I_p - I_q)$ is a normalization term, Ω denotes the image space, I_p and I_q represent the pixel intensities at positions p and q, respectively. In practice, both of these two weighting functions f and g are Gaussian. Suppose that σ_f and σ_g, respectively, are the variances of f and g. The fast bilateral filter [10] is employed in our paper. In our experiments, we set σ_f=15 and σ_g =60. The values of parameters are set according to the concept of the bilateral filter and experiments.

2.3 Luminance segmentation

To segment the global-layer into different luminance levels, many segmentation algorithms can be considered. For example, in [4] an adaptive threshold based intensity clustering is applied to each of the input image. Then, an iterative merging process is employed to estimate the luminance level of the target image. The performance of this scheme is well. However, the iterative procedure leads to an expensive time consuming.

We find that a simple K-mean clustering applied to our global layer works well. To keep simple and efficient of the proposed algorithm, a histogram based K-mean cluster is implemented in our paper.

Inspired by [12], K-mean clustering for image can be efficiently implemented via histogram. Given an image, the K-mean based intensity clustering can be considered as a problem of range cutting or intensity division. In other word, it is to find some thresholds that cut the intensity range into K sub-ranges so that pixels fall in a sub-range to form a cluster.

Firstly, given an image I, we compute its histogram $H[i]$, accumulative histogram $C[i]$ and a weighted accumulative histogram $W[i]$ as follows:

$$H(i) = \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \delta(I(x,y) - i), \quad i = 0,1,...,255, \quad (4)$$

where, $\delta(\ )$ is Dirac delta function, W and H denote the width and height of the Image, $I(x,y)$ is the intensity value of the pixel in (x, y) and $H[i]$ is the number of pixels in I whose intensity is equal to i.

$$C(i) = \sum_{k=0}^{i} H(k), \quad i = 0,1,...,255, \quad (5)$$

$$W(i) = \sum_{k=0}^{i} kH(k), \quad i = 0,1,...,255. \quad (6)$$

Actually, accumulative histogram is the integration of a histogram. We can calculate the number of pixels within a given range with $C(i)$ more efficiently. For example, the number of pixels within range $(t_1,t_2]$ is: $N_r=C(t_2)-C(t_1)$. Similarly, we can figure out the sum of these pixels: $S_r=W(t_2)-W(t_1)$. Then, the mean intensity of the pixels within a given range can be calculated as:

$$U_r=S_r/N_r = (W(t_2)-W(t_1))/(C(t_2)-C(t_1)) , \quad (7)$$

Therefore, the K-mean clustering will be implemented more efficiently with the help of $C(i)$ and $W(i)$. This scheme is illustrated in Fig. 3. The curve shown in the graph is an accumulative histogram. There are 3 clusters will be segmented. The initial thresholds are set so that the number of pixels for each cluster is equal to each other. Then, the mean of each cluster are calculated, and used to update the range cutting thresholds.

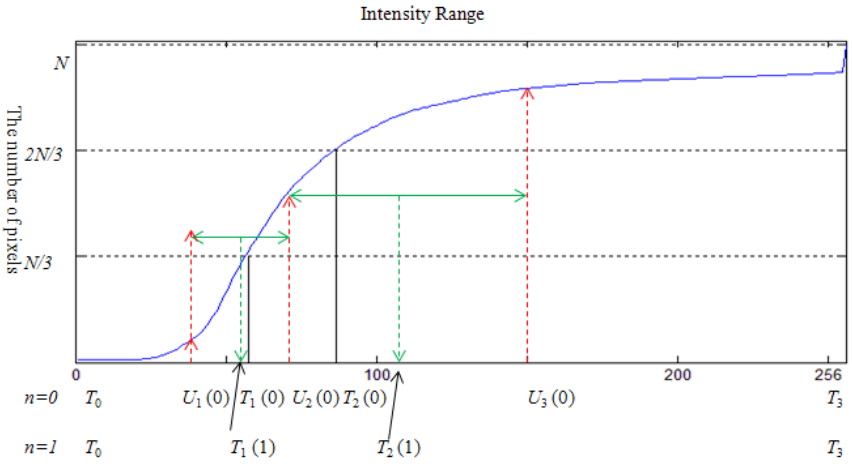


Fig. 3 Histogram-based K-mean clustering algorithm.

The histogram-based K-mean clustering algorithm can be summarized as follows:

Step 1) Initial the number of clusters K and the initial thresholds $T_0(n)$, $T_1(n)$, …, $T_K(n)$,$T_{K+1}(0)$, where, n=0 is the iterative index.

Step 2) Calculate the mean intensity of pixels using (7) for each sub-range: U_k, k=1, 2, … , K.

Step 3) Update the thresholds as follows: $T_0(n+1)=T_0(0)$; $T_K(n+1)=T_K(0)$.

$$T_k(n+1) = (U_k+U_{k+1})/2, \quad k=1, 2, \ldots, K-1. \quad (8)$$

Step 4) If any threshold T_k $(n+1)$ has updated its value in step (3), then set $n=n+1$ and go to step 2); Otherwise, go to step 5).

Step 5) End.

Finally, a set of thresholds are obtained:

$$T=\{T_k \mid k=0,1,2,\ldots,K\}. \quad (9)$$

The global layer is segmented with the K-mean clustering algorithm described above. The number of clusters (regions) should be set to no less than the number of input images.

2.4 Best Exposure Selection

To reproduce details of each luminance level, a single image should be taken with carefully selected exposure. Therefore, given a segmented region of the global-layer, its details in real HDR scene can be depicted appropriately in one of the input images. In this section, we will select the best input image for each segmented region.

For a region with complex textures, the contrast of textures in an image with appropriately exposure will relatively greater than that images with inappropriately exposure. Thus, there are many metrics can be employed to select the most suitable image for a textural region, such as gradients, several kinds of contrasts, entropy, visible luminance, salient features and so on. However, for flat regions, almost all the above mentioned metrics will be similar in different exposures. The above-mentioned schemes will depend on the order of the input images. It will lead to unreasonable results in most of the existing local-based exposure fusion algorithms.

In our algorithm, the overall luminance and details layer are processed separately. Details in flat regions are similar in any input images, while the global layer will take charge in the brightness of background. Therefore, the problem of order dependent will not lead to unreasonable results in our algorithm.

As for the implementation of this step, we employed the scheme proposed in [4]. A normalized level of details (gradients) and entropy are selected as metrics. For convenience, the scheme is summarized as follows.

Suppose that there are $2M+1$ input images and K segmented regions, for each segmented region k, a normalized details level $\overline{D_{k,i}}$ and a normalized entropy $\overline{P_{k,i}}$ are calculated in each input image I_i. Then, a utility function $U(i,k)$ which reflects the effectiveness of the using exposure i for region k is defined as follows:

$$U(i,k) = (\overline{D_{k,i}} + \overline{P_{k,i}})/2, \quad (10)$$

The best exposure image for region k can be selected:

$$R(k) = \arg\max_{i}\{U(i,k)\}, \quad i=-M,\ldots,0,1,\ldots,M. \tag{11}$$

To calculate the details level of a given region *k*, the gradient of each input image is computed firstly:

$$G_i = [G_{x,i}, G_{y,i}], \tag{12}$$

$$G_{x,i} = |I_i(x+1,y) - I_i(x,y)|, \tag{13}$$

$$G_{y,i} = |I_i(x,y+1) - I_i(x,y)|, \tag{14}$$

Then, the level of details defined as:

$$D_{i,k} = \sum_{(x,y)\in\Omega_k} \max(G_{x,i}(x,y), G_{y,i}(x,y)) \tag{15}$$

where, Ω_k denotes region *k*. The normalized details level for region *k* is defined as eq. (16):

$$\overline{Dk,i} = D_{k,i} / \max_{j=-M,\ldots M}\{D_{k,j}\} \tag{16}$$

To calculate the normalized level of entropy for region *k*, a normalized histogram $H_{i,k}$ for each input image is computed:

$$H_{i,k}(j) = \left(\sum_{(x,y)\in\Omega k} \delta(I_i(x,y) - j) \right) / N_k \quad , j=0,\ldots,255, \tag{17}$$

where, N_k is the number of pixels belong to the cluster *k*. The level of entropy for region *k* can be defined as:

$$P_{i,k} = -\sum_{j=0}^{255} H_{i,k}(j)\log_2(H_{i,k}(j)+\varepsilon), \tag{18}$$

where, ε is a small number to avoid the variable of logarithmic function from becoming zero. (We set $\varepsilon = 10^{-25}$ in our experiments.) Then, the normalized level of entropy for region *k* can be calculated as:

$$\overline{Pk,i} = P_{k,i} / \max_{j=-M,\ldots M}\{P_{k,j}\} \tag{19}$$

Until now, a best input image for each region is selected. We can construct an index map for details selection. In this map, each segmented region is denoted with a number to indicate an input image. The details of this region will be selected from the input image with the best exposure.

2.5 Details Fusion in Gradient Domain

In the proposed algorithm, the resulting image is produced by fusing the global layer and the selected details from input images. The detailed layer of an input image can be estimated by several

approaches. For example, the difference of an input image and its bilateral filtered version can be considered as an estimation of the detailed layer [13, 15]. However, the bilateral filter should be applied to each input image in this scheme. That will result in expensive computational complexity.

Inspired by seamless cloning technique in Poisson image editing [9], we considered the details fusion as a problem of a region based image cloning. We would like to clone the details in the selected best exposure to a given region. That can be implemented more efficiently in gradient domain. In the last section, the gradient map for each input image has already been computed in (12). Let $G_g = \nabla I_g$, where $\nabla . = [\frac{\partial .}{\partial x}, \frac{\partial .}{\partial y}]$ is the gradient operator, G_g is gradient of the global-layer. Details fusion in gradient domain can be described as follows:

$$G(x,y) = \begin{cases} \nabla I_{R(k)}(x,y), & \text{if } (x,y) \in \Omega_k \text{ and } |\nabla I_{R(k)}(x,y)| >= |G_g(x,y)| \\ G_g(x,y), & \text{otherwise.} \end{cases} \quad (20)$$

where $G(x,y)$ is the fused gradient of the output image. Ω_k denotes the segmented region k, which the corresponding best exposure input image is $I_{R(k)}$. $R(k)$ is obtained in (11).

The problem of exposure fusion can be formulated as follow minimization problem:

$$I = \min_I \iint_{(x,y)\in\Omega} |\nabla I - G|^2 \, dxdy \,, \quad (21)$$

where Ω is the spatial domain of image. According to the variational principle, the solution of (21) is the unique solution of the following Poisson equation:

$$\Delta I = div(G) \,, \quad (22)$$

where $\Delta . = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is the Laplacian operator. The derivatives around boundaries are set to 0, namely the Neumann boundary condition is employed in our paper. There are a number of ways to solve the Poisson PDE (Partial Differential Equation) with Neumann boundary as (22), such as analytical solution or multi-grid method[14]. A. Agrawal's [7, 8] procedure is utilized in our implementation.

2.6 Implementation

Details of the implementation of our layered-based exposure fusion algorithm are summarized as follows.

Step 1) Convert all the input RGB color images into gray image. The proposed fusion process is

implemented in gray domain.

Step 2) Estimate the global-layer according to equation (1) and (2).

Step 3) Segment the global-layer with the histogram based K-mean cluster which is outlined in section C.

Step 4) Select a best exposure image for each segmented cluster in step 3), obtain (11). In this step, the gradients of all the input images are saved for the next step.

Step 5) Fuse the global layer and details in gradient domain according to (20).

Step 6) Solve equation (22) to obtain the target gray image I.

Step 7) Recover the color of the output image using (23) [15].

$$C_{out}(x,y) = I(x,y) * \left(\sum_{k=-M}^{M} C_{in,k}(x,y) \Big/ \sum_{k=-M}^{M} I_k(x,y) \right)^{S}, \quad (23)$$

where C_{out}=[R_{out}, G_{out}, B_{out}] stand for the RGB components of the resulting image, $C_{in,k}$ =[$R_{in,k}$, $G_{in,k}$, $B_{in,k}$] are the RGB components of the kth input image, I is the result of step 6) and I_k is the luminance intensity of the kth input image, (x,y) is a coordinate in a image and S is the saturation parameter. We set S=0.8 in our experiments.

3. EXPERIMENTAL RESULTS AND DISCUSSION

To justify the effectiveness of the proposed layered-based exposure fusion scheme, we have done several set of experiments. In our experiments, we use those HDR images that are widely used in the literature to generate a set of different exposure standard dynamic range images (SDR). A shot of HDR images used in this paper are list in Fig. 4. For each HDR image, there are 11 standard dynamic range (SDR) images with different exposure level are manually collected using HDRShop [16]. As images shown in Fig. 2, they are named as I_{-5},…, I_0, I_1,…, I_5. There are 3 images selected as input image in most experiments unless otherwise mentioned. Exposure Difference (ED) denotes the images selected. For example, DE=1 means I_{-1}, I_0 and I_1 are selected, while DE=3 means I_{-3}, I_0 and I_3 are selected.

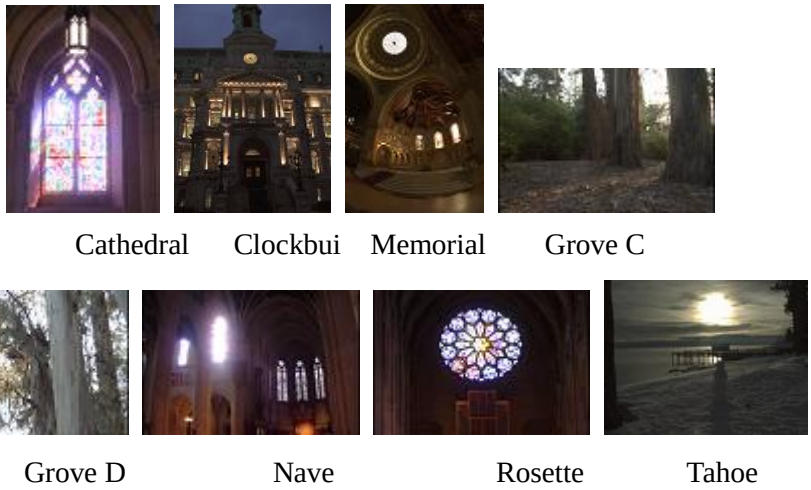


Fig. 4 Images used in our experiments

3.1 Comparison with other algorithms

Six algorithms are considered: the pixel-based fast fusion [2], the block-based fast fusion [2], the cluster fusion [4], the gradient directed fusion [17], the gradient synthesis fusion [3] and our algorithm. To compare the subjective performances of different methods, results of these 6 methods are shown in Fig. 5(a)-(f). There are 3 images in each row. From left to right they are fused with ED=1, ED=3 and ED=5 respectively.

As shown in Fig. 5(a), the pixel-based fast fusion method works well with ED=1, both the details and the overall brightness can be persevered in the output image. The middle image of Fig. 5(a) (ED=3) is with an acceptable subjective quality. However, compared with the result of ED=1, the region of window looks dim. When ED=5 is selected, form the right image of Fig. 5(a) we can see that the background of window is incorrect. The brightness of this region is dark than it should be in a real scene, and it even darker than the region of wall. We refer to this problem as luminance reversion.

From images shown in Fig. 5 (b)-(d), we can see that most of the tested methods suffer from the luminance reversion problem. With increasing of the exposure difference, the luminance reversion effects will appear with a high likelihood. Images shown in Fig. 5 (d) and (e) seems not so sensitive to the exposure difference of the input images.

To further analysis on the luminance reversion problem, a single line across the window of Rosette is taken into account, as shown in Fig. 6. The curve shown in Fig. 6(a) is a real HDR scene form the HDR Rosette image. Curves in Fig. 6 (b) are the same line in the input images I_{-5}, I_0 and I_5 respectively. It is obviously that pixels from 200 to 500 are located in the region of window. Other pixels are located in the wall. It is well known that a short exposure time will be more suitable for

capturing details in a brighter region. Therefore, details in I_{-5} (the blue curve) will be more likely to be weighted more significantly within the window region (namely pixels form 200 to 500) for fusion image. While, details in I_5 (the red curve) are more suitable for the dark regions in real scene. However, from Fig. 6 (b) we can see that the overall intensity level of the blue curve in the region of window can be no greater than the overall intensity of the red curve in the region of wall. It results in luminance reversion problem in the fused image. For example, it is difficult to tell which region is brighter than others in Fig. 6 (c). Furthermore, from the right image shown in Fig. 5 (a) we can see that the luminance of window is more similar with image I_{-5} in Fig. 2.

Most of the existing exposure fusion methods estimate the weights of input pixels only based on local features. Pixel-based fast fusion method is a simple and fast algorithm. Only pixels in the same location of the input images are used to estimate their weights for fusion. Block-based fast fusion takes a local-block region to estimate the contrast of details. Compared with the pixel-based fusion, it is slightly less sensitive to the value of exposure difference. From the result image of ED=3 shown in Fig. 5 (b), we can see that the region of window is brighter than the middle image of Fig. 5 (a). Both of the cluster fusion and gradient-directed methods are local-based fusion algorithms. They cannot avoid the luminance reversion problem.

The gradient-synthesis fusion is also a local-based fusion algorithm. It selects the based exposure input images based on blocks. To remove the artifacts on the borders of blocks, a Gaussian blending function is applied to estimate the weights of each input pixels. The greater spatial variance of the Gaussian blending function makes the fusion result more robust respect to the value of exposure difference. However, the great spatial variance leads to very expensive computational complexity.

ED=1 ED=3 ED=5

(a) Pixel-based fast Fusion

(b) Block-based fast fusion

(c) Cluster Fusion

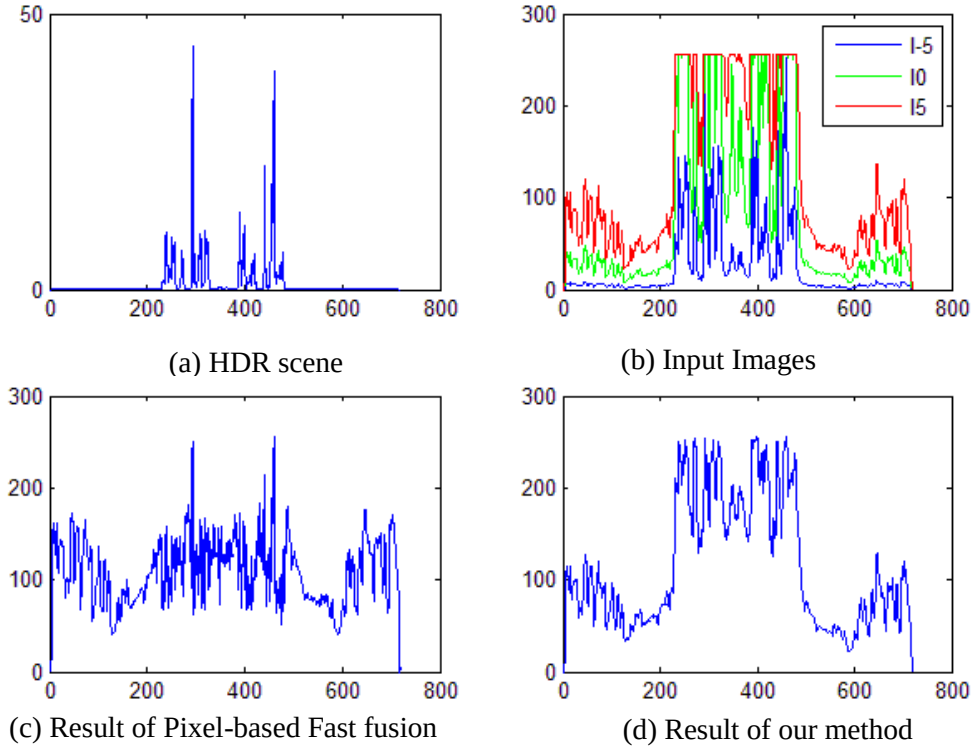
(d) Gradient Directed

(e) Gradient Synthesis

(f) The proposed algorithm

Fig. 5 Comparison of different methods (Rosette)

The global-layer is considered in our algorithm, which helps to avoid the luminance reversion problem efficiently. Fig. 5 (f) and Fig. 6 (d) show that the proposed algorithm is more robust to the exposure difference of the input images.



(a) HDR scene (b) Input Images

(c) Result of Pixel-based Fast fusion (d) Result of our method

Fig. 6 A single line across the window in Rosette.

3.2 Analysis on detail preservation

Until now, there has been no standard objective evaluation method available for measuring the quality of exposure fused images. Most exiting methods evaluate their own results by means of subjective evaluation. To further analysis the performances of detail preservation in the fused images, PSNR is employed in our experiments. For each tested HDR scene, the HDR image is tone mapped to a target LDR image using Photomatix Pro 4.2.3 with default method and parameters. In our experiments, we find that for all the tested HDR images, the tone mapped image with default method and parameters shows superior subjective quality than other methods that included in Photomatix Pro. Then, PSNR between the fused image and the target LDR image is computed in gray domain.

Tab. 1 shows the objective performances of different algorithms, which ED is 3. The average PSNRs for each algorithm are listed in the last row. Fig. 7 shows that the average PNSRs for each algorithm respect to exposure differences of the input images. From the Fig. 7, we can see that almost all the compared algorithms can preserve HDR details well in their results when the exposure differences of the input images are 1 or 2. While, with the increasing of ED, PSNRs falls steeply. However, the proposed algorithm performs more steadily.

We also can observe that for ED is set to 1 or 2, the pixel-based or block based fast fusion algorithms can solve the problem more efficiently. However, in such cases, the input images will be more similar to each other, which may be difficult to contain the details in very bright and dark regions.

Tab. 1 The objective performances of different algorithms (PSNR: dB, ED=3)

	Pixel-based fast Fusion	Block-based fast fusion	Cluster Fusion	Gradient Directed	Gradient Synthesis	Our algorithm
cathedral	22.25	23.39	17.85	22.52	21.99	22.06
groveC	22.47	19.63	23.17	18.50	16.45	23.24
memorial	15.13	13.44	10.71	14.44	11.47	15.40
rosette	20.42	17.93	15.98	22.06	16.84	24.66
tahoe	19.81	23.15	19.32	19.50	20.56	22.21
groveD	19.10	22.98	22.45	18.85	19.69	21.66
clockbui	24.78	22.66	16.70	21.55	16.11	21.75
nave	21.64	18.90	17.15	21.96	16.89	26.44
Average	**20.70**	**20.26**	**17.91**	**19.90**	**17.50**	**22.17**

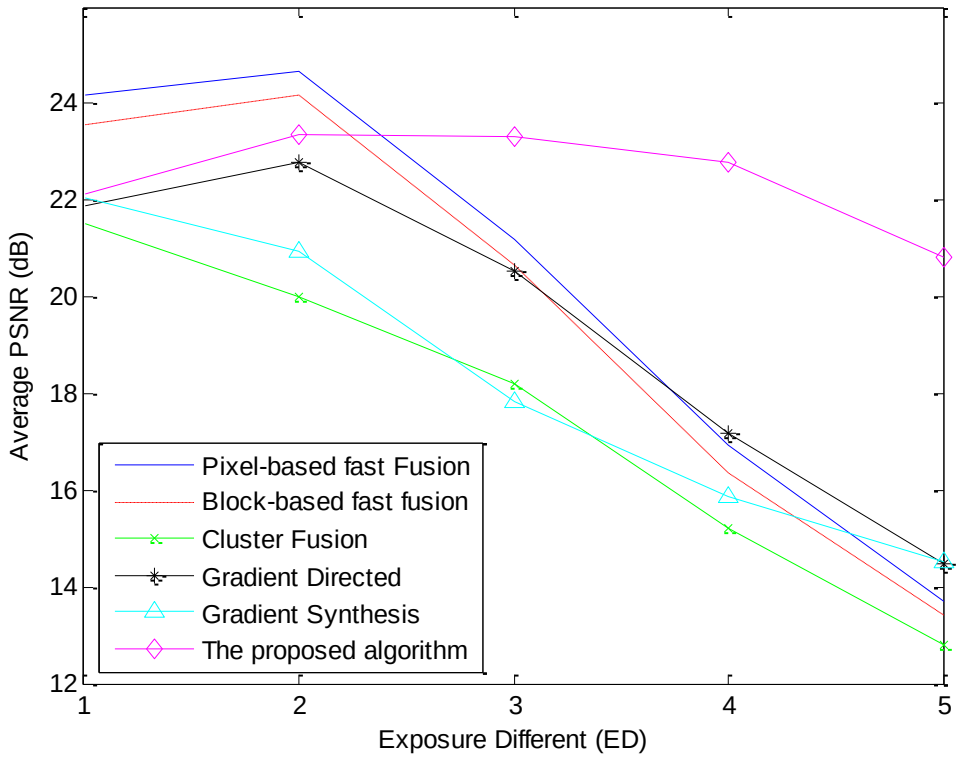


Fig. 7 Average PSNR performances respect to exposure differences of the input images.

3.3 Analysis on color consistency

Besides the above mentioned luminance reversion and detail preservation problem, color distortion also tends to appear in the results of many local-based exposure fusion methods. The color appearance of the fused image may sensitive to the exposure difference values of the input image sequence. However, most of the existing algorithms haven't considered the color consistency problem in their paper.

As images shown in Fig. 5, with the increasing of ED, color of the results tend to gray out (in Fig. 5 (a) and (b)) or become brighter (in Fig. 5 (c), (d) and (e)). While, the color appeared in the fused image should be as similar as the real scene. The results of our algorithm are less sensitive to the ED of the input images. In other word, it can preserve the color consistency in the resulting images more efficiently than other methods.

To compare the color consistency of the proposed method and the gradient synthesis method [3] in detail, fusion results of Cathedral using these two methods are shown in Fig. 8 In the gradient synthesis method, which processes the RGB channels separately, the color of the resulting images will become brighter with the increasing of ED. We also can see that the color in the window frame is unnatural. The gradient synthesis method is a local-based fusion algorithm in nature. Therefore, the fusion result of the brighter regions in a real scene will be affected by the input image with a shorter exposure time more significantly. In Fig. 8 (b), the results of the proposed algorithm can preserve the color consistency respect to the different input images. While, with the increasing of ED, the color will become slightly gray out.

To further improve the color consistency of the fused images, more suitable color space and color recover scheme should be investigated in our future work. Some research on color transform methods [18,19,20,21] may be helpful.

Some other results of the proposed algorithm are shown in Fig. 9. We can see that there are no halo artifacts and luminance reversion problem.

ED=1 ED=3 ED=5

(a) Gradient Synthesis

(b) The proposed algorithm

Fig. 8 Compare on color consistency (Cathedral)

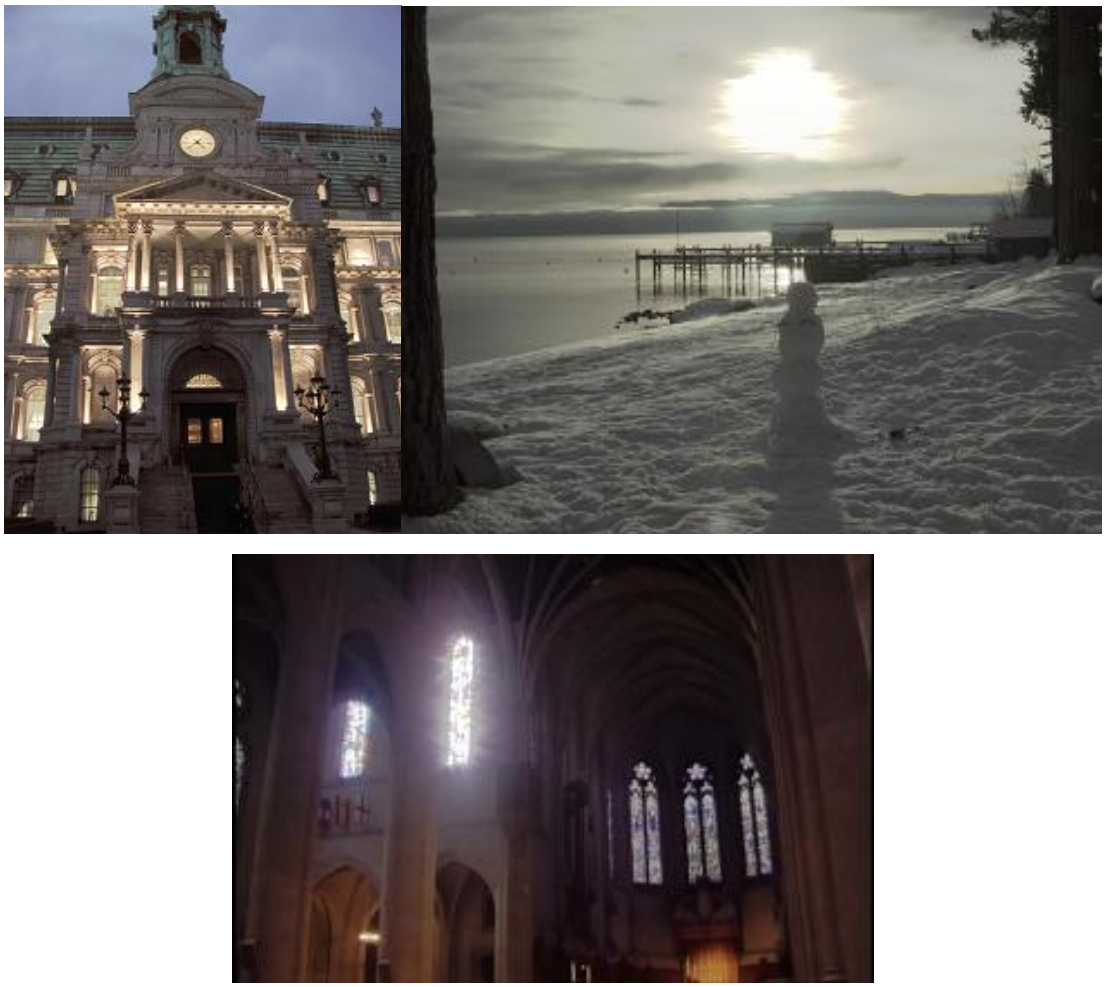


Fig. 9 Other fusion results

4. CONCLUSIONS

For different high contrast scenes, the input images for exposure fusion may be captured with different exposure parameters. Most existing exposure fusion algorithms are sensitive to the exposure difference between the input images. A layered-based exposure fusion algorithm is presented in this paper. The global information can improve the robustness of the exposure fusion algorithm efficiently. The local gradients are helpful to preserve details. The proposed algorithm can produce a rational result for a HDR scene without halo artifacts and luminance reversion problem.

5. ACKNOWLEDGMENTS

The work in this paper is supported by the National Natural Science Foundation of China (No 61003289, No.61100212), the Natural Science Foundation of Beijing (No. 4102008), the Excellent Science Program for the Returned Overseas Chinese Scholars of Ministry of Human Resources and

Social Security of China, and the Scientific Research Foundation for the Returned Overseas Chinese Scholars of MOE.

References

1 Mertens, T., J, Kautz., F, Van, Reeth.: 'Exposure Fusion', *IEEE Computer Society*, Washington, DC, USA, July 2007, pp. 382-390

2 Vavilin, A., K, Deb., K, Jo.: 'Fast HDR image generation technique based on exposure blending', *In IEA/AIE'10. Part III. LNAI6098.*, Berlin, Heidelberg, Springer-Verlag, 2010, pp. 379-388

3 Varkonyi-Koczy, A, R., A, Rovid., T, Hashimoto.: 'Gradient-Based Synthesized Multiple Exposure Time Color HDR Image', *IEEE Transactions on Instrumentation and Measurement*., 2008, **57**, (8), pp. 1779 -1785

4 Jo, K., A, Vavilin.: 'HDR Image Generation based on Intensity Clustering and Local Feature Analysis', *Comput. Hum. Behav.*, 2011. **27**, (5), pp. 1507-1511

5 Shen, R.: 'Generalized Random Walks for Fusion of Multi-Exposure Images', *IEEE Transactions on Image Processing*., 2011, **20**,(12), pp. 3634 -3646

6 M, Song.: 'Probabilistic Exposure Fusion', *IEEE Transactions on Image Processing*., 2012, **21**, (1), pp. 341-357

7 A, Agrawal.: 'An algebraic approach to surface reconstruction from gradient fields', 2005, pp. 174-181

8 A, Agrawal.: 'What is the Range of Surface Reconstructions from a Gradient Field, Book What is the Range of Surface Reconstructions from a Gradient Field, Series What is the Range of Surface Reconstructions from a Gradient Field', *Editor ed*., 2006, pp. 578-591

9 P, Perez.: 'Poisson Image Editing', Proc. ACM SIGGRAPH., 2003, pp. 313-318

10 F, Durand., J, Dorsey.: 'Fast bilateral filtering for the display of high-dynamic-range images', *ACM Trans. Graphics*., 2002, **21**, (3), pp. 257-266

11 E, Reinhard., G, Ward., S, Pattanaik., P, Debevec.: 'High dynamic range imaging: acquisition, display and image-based lighting', *Elsevier Inc*., 2006

12 J, Duan.: 'Tone-mapping high dynamic range images by novel histogram adjustment', *Pattern Recognition*., 2010, **43**, (5), pp. 1847-1862

13 F, Durand., J, Dorsey.: 'Fast bilateral filtering for the display of high-dynamic-range images', Proc. SIGGRAPH '02. ACM., 2002, pp. 257-266

14 W, H, Press., S, A, Teukolsk., W, T, Vetterling., B, P, Flannery.: 'Numerical recipes in C: the art of scientific computing', Cambridge, U.K., 1992

15 X, Li.: 'An adaptive algorithm for the display of high-dynamic range images', J. Vis. Comun., Image Represent, 2007, **18**, (5), pp. 397-405

16 HDRShop V1: http://www.hdrshop.com/

17 W, Zhang., W, Cham.: 'Gradient-Directed Multiexposure Composition', *IEEE Transactions on Image Processing*., 2012, **21**, (4), pp. 2318-2323

18 E, Reinhard.: 'Color transfer between images', *IEEE* Comput *Graph.*, 2001, **21**, (5), pp. 34-41

19 E, Reinhard., T, Pouli.: ‘Colour spaces for colour transfer’, Proc. CCIW'11., Springer-Verlag, 2011, pp. 1-15

20 A, Maslennikova., V, Vezhnevets., ‘Interactive Local Color Transfer Between Images, Book Interactive Local Color Transfer Between Images, Series Interactive Local Color Transfer Between Images’ , 2007

21 Y, Xiang.: ‘Selective color transfer with multi-source images’, *Pattern Recogn Lett*., 2009, **30**, (7), pp. 682-689