Multi-exposure image fusion based on recursive filter

Ting Li, Kai Xie, Tong Li, Yubing Yan, Ling Huang

Department of Information Engineering, Beijing Institute of Graphic Communication, Beijing, 102600, China

2531424229@qq.com
xiekai@bigc.edu.cn

Abstract. Multi-exposure image fusion technology is to fuse images of the same scene with different exposures into a high-quality image that contains more scene information. This paper proposes a multi-exposure image fusion method based on a recursive filter. The method first calculates three characteristic indicators of the multi-exposure image sequence: image contrast, entropy, and exposure information. And use three characteristic indicators to estimate the initial image fusion. In order to obtain a more accurate weight map for image fusion, a recursive filter is used to refine the weight map. Finally, the weighted sum method is used to construct the fused image. The experimental results show that the fusion results obtained by the algorithm in this paper have high image clarity, better subjective visual fusion effects and objective performance indicators, and maintain a high computational efficiency.

1. Introduction
The dynamic range of brightness in real life scenes can reach 8 to 10 orders of magnitude, but the dynamic range of ordinary digital imaging equipment is 2 to 3 orders of magnitude, and the actual image dynamic range exceeds all the information of the real scene that the equipment can capture [1]. Due to the limitation of dynamic range, the captured single image often has underexposed and overexposed areas, the edge information and texture details are seriously lost, the image quality is obviously reduced, and the visual perception is poor. The current methods for solving high dynamic range image (HDR) are mainly divided into the following two categories: tone mapping method based on inverse camera response curve and direct fusion method of low dynamic range image (LDR) image. This article mainly studies the second method.

The method of direct fusion of LDR images is to use LDR images of the same scene with different exposures to directly fusion generate HDR images according to certain rules. This method is simpler and more efficient than the first type of fusion technology. Burt and Adelson proposed the pyramid model for the first time, and applied it to the image fusion algorithm, and achieved good results [2]. Mertens et al. proposed a multi-exposure fusion method to generate HDR images based on the Laplacian pyramid model [3]. This method selects the weighted fusion of image contrast, saturation, and exposure information to generate HDR images. The result of the fusion is high in definition, but the image loses more information in the darker and lighter areas. Wu et al. proposed an HDR image generation method based on optimal block fusion [4]. First, the multi-exposure image is divided into blocks, and then the light gradient factor, saturation factor and appropriate exposure factor are used to select the optimal exposure block. Finally, the HDR image is generated according to certain fusion rules. The resulting image of the fusion has a blocky effect, which needs to be eliminated later, which increases the complexity and running time of the algorithm. Zhang et al. proposed a method to
generate HDR images based on the gradient method [5]. This algorithm can eliminate the "ghost" of the image, but this algorithm is only suitable for dynamic images. Shen et al. proposed a multi-exposure image fusion algorithm based on random walk [6], which calculates the image contrast and color constancy as probability, and obtains the fused image through the global optimal solution.

This paper proposes a simple and effective method for multi-exposure image fusion. According to the three characteristic indicators of image contrast, entropy and good exposure, the initial fusion weight of the image is calculated, and then the recursive filter [7] is used to modify and refine the initial weight map to obtain an accurate fusion weight map, and finally, use the weighted sum method to fuse the image. This method is only suitable for image fusion of static scenes, and further research is needed for image fusion of dynamic scenes. Experiments show that this method is superior in both subjective and objective aspects.

2. Algorithm introduction

The algorithm in this paper is mainly divided into four steps: image feature index calculation, weight map estimation, weight map correction, and weighted fusion. The algorithm flow chart is shown in Figure 1:

![Algorithm flow chart](image)

Fig 1. The algorithm flow chart of this article

2.1. Local contrast

Contrast is the ratio of black to white in the picture, which reflects the details of the image and the Grayscale and the degree of color performance. The better the color performance of the image with relatively large contrast, the clearer the image, the more natural the visual effect of the image. The local contrast of each pixel [8] is calculated as follows:

\[
C_n(x, y) = \frac{Y_n(x, y)}{h(x, y)}
\]  

(1)

In the formula: \( Y_n \) represents the input image; \( n \) is the number of input images; \( h \) is the convolution operation.; \( Y_n \) is defined as follows:

\[
Y_n = 0.299 \times I_n^r + 0.587 \times I_n^g + 0.114 \times I_n^b
\]  

(2)

In the formula: \( I_n^r \), \( I_n^g \), \( I_n^b \) denote the intensity of R, G, and B channels respectively. \( h \) represents a high-pass filter, and the specific formula is as follows:

\[
h = \begin{bmatrix}
0 & -1 & 0 \\
1 & 4 & -1 \\
0 & -1 & 0
\end{bmatrix}
\]  

(3)

Since local differences indicate contrast, the higher the contrast, the more details the corresponding area contains. If two pixels look very different and share the same grayscale value, the surrounding pixels are different, so they get different values through the high-pass filter. In the multi-exposure fusion, the difference between the multi-exposure images is caused by the exposure time and has a great influence on the grayscale value in the image. Therefore, in the same area, the pixel patch will not be different. The colors share the same grayscale value, and this measurement can detect good
details from multiple exposure images. When obtaining the local contrast of each pixel, it can be represented by the following arrangement:

$$\hat{C}_n(x,y) = \begin{cases} 1, & C_a(x,y) = \max \{C_n(x,y), n = 1,2,\cdots,N\} \\ 0, & \text{otherwise} \end{cases}$$

(4)

In the formula: $N$ is the number of source images; $\hat{C}_n(x,y)$ is the generated local contrast information, whose purpose is to maintain image details.

2.2. Image entropy

In information theory, it can be known that the entropy value of information can reflect the size of its information volume, so the entropy value of the image is calculated as an indicator to evaluate the size of the image information volume. The calculation formula is as follows:

$$S = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$$

(5)

In the formula, $p(z)$ is the histogram data of the image; $L$ is the gray level of the image.

2.3. Exposure information

The exposure factor is introduced based on the human visual characteristics and the spatial frequency characteristics of the image. A moderately exposed image has a more even distribution of pixels and a wider range of pixel values. In the case of over-exposure or under-exposure, the pixel points will be more distributed in extreme points. In order to get a better result, the image should be Choose well-exposed pixels for fusion. The calculation formula is as follows:

$$E_n(x,y) = \exp \left( \frac{-(I_{(x,y)} - 0.5)^2}{2\sigma^2} \right)$$

(6)

In the formula, $I_{(x,y)}$ is the normalized pixel value of the image; $\sigma$ is the standard deviation of the distribution of the control weight mapping function, which 0.02 is taken in the experiment.

2.4. Weight map estimation

In order to retain more scene details and rich color information, and eliminate the effects of underexposed and overexposed pixels, the algorithm in this paper multiplies the three characteristic indicators of the exposure sequence: contrast, entropy, and good exposure information to obtain the initial weight map. The formula is as follows:

$$W_n = C_n \times S_n \times E_n$$

(7)

The weight map is normalized to ensure that the sum of the weights of each pixel is 1. The formula is as follows:

$$\hat{W}_n(x,y) = \left[ \sum_{m=1}^{N} W_m(x,y) \right]^{-1} W_n(x,y)$$

(8)

2.5. Weight refinement and weighted fusion

Because the weight map is easy to introduce noise in the processing process, in order to eliminate the influence of noise, a recursive filter is used to filter the initial weight map to obtain an accurate weight map. Recursive filter is a real-time edge-preserving filter, which is widely used in image and video processing. The processing formula is as follows:

$$\hat{W}^t_n = RF(\hat{W}_n(x,y), I_n(x,y))$$

(9)
In the formula, \( RF(\cdot) \) represents the recursive filtering operation. The weight map refined by the recursive filtering is used to obtain the final fused image through weighted fusion. The formula is as follows:

\[
F(x, y) = \sum_{n=1}^{N} I_{n}(x, y) \times W_{n}^{\wedge}(x, y) \quad (10)
\]

3. Experimental results and analysis

All experiments are done using MATLAB (R2018a) programming on a PC platform with Intel i7 processor (3.5GHz, 64-bit) and 16G memory. This article selects several sets of classic static multi-exposure image fusion sequences for testing, and compare and analyze algorithm of this paper with the algorithm of Mertens et al and FMMR.

During the experiment, the relevant parameters of the algorithm were set, and the spatial parameters and smoothing coefficients of the recursive filter were taken as \( \sigma_{r} = 100, \sigma_{r} = 4, r_{i} = 3, r_{z} = 30 \).

In addition to subjective evaluation, this article uses three commonly used image quality evaluation indicators: image clarity, information entropy and standard deviation. Quantitatively evaluate algorithm performance through the above three indicators.

3.1. Average gradient

The definition of the image is described by the average gradient of the image. The larger the value, the richer the minute details contained in the image and the clearer the result image of the fusion. The calculation formula is:

\[
\tilde{E} = \frac{1}{(R-1)(C-1)} \sum_{r=1}^{R-1} \sum_{c=1}^{C-1} \left[ \frac{(Z_{r,c} - Z_{r+1,c})^2 + (Z_{r,c} - Z_{r,c+1})^2}{2} \right] \quad (11)
\]

In the formula, \( Z_{r,c} \) represents the pixel grayscale value, \( r \) and \( c \) represent the row and column coordinates of the pixel, respectively; \( R \) and \( C \) represent the number of pixel rows and columns of the test image.

3.2. Information entropy

Information entropy is a measurement method that reflects the richness of image information. The larger the entropy value, the richer the information contained in the image. Information entropy is closely related to the image quality perceived by human subjectively, and the influence of different types and distorted images can be predicted and found. The calculation formula is as follows:

\[
S = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (12)
\]

3.3. Standard deviation

The standard deviation reflects the degree of dispersion between the image pixel value and the mean value. The larger the standard deviation, the better the image quality.

3.4. Experimental results

This part conducts an experimental analysis on the proposed method, and compares the performance of the three algorithms through average gradient, information entropy and standard deviation. This experiment uses six "garage" images with different exposures as the source images for fusion display.
Figure 2 shows a set of "garage" image sequences with different exposure levels and the resulting images of three different fusion methods. Table 1 shows the evaluation results of three evaluation indicators of different algorithms. From this group of images and table data, it can be concluded that the classic pyramid algorithm has a good effect on fusing images, but it is obvious from the figure that there is a problem of loss of detailed information in the underexposed areas, and the image is too bright. The result of FMMR algorithm fusion results in serious loss of image detail information, and many details of the car cannot be displayed well. Although the algorithm in this paper is not as good as the overall image fusion effect of the classic algorithm, the algorithm in this paper has better local contrast and rich color information, and the result of image fusion is more natural in terms of visual effects.

4. Conclusion
This paper proposes a multi-exposure HDR image fusion method based on a recursive filter, without HDR construction and tone mapping, and directly uses multiple LDR images with different exposures for fusion. The algorithm in this paper obtains the initial weight map of image fusion through three image feature indicators, uses recursive filters to modify the weight map to obtain a more accurate
weight map, and finally uses the weighted sum fusion rule to obtain the final result image. It is proved that the method is simple and the fused image has rich details, natural color and high image quality.

Acknowledgments
This work was financially supported by the practical training project of "Multi-exposure Image Fusion Based on Convolutional Neural Network" (03150120001/014).

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. Burt and T. Adelson. The Laplacian Pyramid as a Compact Image Code. IEEE Transactions on Communication, COM-31:532–540, 1983.
[3] 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.
[4] WU X J, SONG Z, YU G. A Novel Multiple Exposure Merging Method for High Dynamic Range Image Generation. Proceedings of the IEEE International Conference on signal Processing Systems, Chengdu, China, 2010: V174-V177.
[5] ZHANG W, Cham W k. Gradient-Directed Multiexposure Composition. IEEE Transactions on Image Processing.2010,21(4):2318-2323.
[6] SHEN R, CHENG I, SHI J, et al. generalized Random Walks for Fusion of Multi-Exposure Images[J]. IEEE Transactions on Image Processing Apublication of the IEEE Signal Processing Society,2011,20(20):3634-3646.
[7] E.S.L.Gastaland 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.
[8] YAN M, HU Y L, HU Y SH. Dynamic range compression with local contrast preservation[J]. Journal of Shanghai University (Natural Science Edition),2017,23(5):682-689.