High Dynamic Range Imaging Using Multiple Exposures

Xinglin Hou 1,2,3, Haibo Luo 1,3,4, Peipei Zhou 1,2,3,4 and Wei Zhou 5

1 Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016
2 University of Chinese Academy of Sciences, Beijing, 100049
3 Key Laboratory of Opto-Electronic Information Processing, CAS, Shenyang 110016
4 The Key Lab of Image Understanding and Computer Vision, Liaoning Province, Shenyang, 110016
5 AVIC Jiangxi HONGDU Aviation Industry Group LTD, Nanchang, China

Abstract. It is challenging to capture a high-dynamic range (HDR) scene using a low-dynamic range (LDR) camera. This paper presents an approach for improving the dynamic range of cameras by using multiple exposure images of same scene taken under different exposure times. First, the camera response function (CRF) is recovered by solving a high-order polynomial in which only the ratios of the exposures are used. Then, the HDR radiance image is reconstructed by weighted summation of the each radiance maps. After that, a novel local tone mapping (TM) operator is proposed for the display of the HDR radiance image. By solving the high-order polynomial, the CRF can be recovered quickly and easily. Taken the local image feature and characteristic of histogram statics into consideration, the proposed TM operator could preserve the local details efficiently. Experimental result demonstrates the effectiveness of our method. By comparison, the method outperforms other methods in terms of imaging quality.

1. Introduction
The dynamic range of a natural scene can be as high as the order of $10^8$. However, the widely used digital camera can capture a range of order from $10^2$ to $10^4$. A high dynamic range (HDR) scene cannot be well-expressed through a single image captured by a low dynamic range (LDR) camera. It is of great significance to develop an HDR imaging method for LDR cameras.

In recent years, many HDR prototype have been designed by redesign of pixel [1]. However, because of their high price, these cameras are unacceptable for ordinary consumers. Instead the expensive hardware, an HDR image can be reconstructed by using a set of multiple exposures.

During the process of HDR reconstruction (HDDR). The recovering of the CRF is the key which determine the assignment of the scene radiance to the image gray value [2-4]. After the obtaining of the HDR image, it should be mapped to a displayable gray range using a tone mapping operator. There are some advantages of the HDDR+TM method: there is no specialized hardware required; various operations can be performed on the HDR images, such as virtual exposure; and user interactions are allowed in the TM phase to generate a tone-mapped image with the desired appearance.

The propose of HDDR is to reconstruct the HDR radiance image in which the CRF is necessary. Large number of scholars have proposed method for recovering the CRF. Debevec [5] is the pioneer who recover the CRF by minimizing a quadratic objective function. After that, a weight function is used to construct HDR radiance image. In Robertson et al.‘s method [6], another object function of radiance, exposure time and the amount of light is constructed. A form of Gauss–Seidel relaxation is used to
minimize the object function and CRF is constructed with a Gaussian weight function. Grossberg [7] recover the CRF from intensity mapping functions, which are determined by image histograms. In Lu et al. ’s work [8], CRF is constructed through an expectation maximization like optimization algorithm using a maximum-likelihood approach based on a Bayesian framework-defined probabilistic model. Mitsunaga [9] describe the CRF using a high-order polynomial. Construction of the CRF is converted into the solution of the coefficients. The CRF is recovered by formulating an error function. In this work, the CRF is also described use a high-order polynomial. By adding a constraint, the polynomial can be solved easily using singular value decomposition.

For an HDR image, TM addresses the problem of dynamic range reduction from the scene radiance to the displayable range while preserving the image details. TM is usually applied in the luminance channel. There are two kinds of TM, local TM and global TM. Local TM methods [10-12] are based on the local spatial information; while in the global TM method [13], the same nonlinear transformation curve is used for all pixels. Reinhard et al.[10] present a local TM method based on the operation in film printing. A global TM function is initially used for the whole pixel luminance. Then a local process based on “dodging and burning” is adopted to enhance the image details. Shan et al.[11] propose a method for HDR reconstruction by using overlapping window-based linear functions. A guidance map based on local constraints is used in TM and it suppresses local high contrast and effectively preserves details. Larson et al.[13] present a global tone reproduction operator, which incorporates models for human contrast sensitivity, glare, spatial acuity, and color sensitivity to match the subjective experience. Local TM operator brightens areas around dark objects, which could produce some characteristic effects such as halos. While global TM enhances the global contrast, the local contrast may be compressed. For the two methods mentioned above, local TM provides a better result according to the perceptual evaluation introduced by Yoshida [14]. In our work, a local TM method is proposed based on the perceptual evaluation introduced by Yoshida [14]. In our work, a local TM method is proposed based on the perceptual evaluation introduced by Yoshida [14].

2. Camera Response Function Recovering

CRF is the assignment of scene radiance to the image gray value. CRF is the patent of a company and vary appreciably from one imaging system to the next. When analyze the CRF results obtained from imaging systems, we find that all CRFs are distributed like a high-order polynomial which is shown in Figure 1. We also conclude that the CRF is monotonic or at least semi-monotonic. Any curve can be fitted by a high order polynomial. Here, we modeling the CRF with a high order polynomial:

\[ E = f(M) = \sum_{n=0}^{N} c_n M^n \]  

where, \( M \) denotes pixel grey value, \( E \) denotes the exposure and \( n \) is the order of the polynomial. The minimum order of the polynomial required clearly depends on the fitting error.

Considering \( f_q \) and \( f_{q+1} \) are two images of a scene taken with different exposure time \( t_q \) and \( t_{q+1} \). The radiance radio of a pixel can be written as:

\[ \frac{E_{p,q}}{E_{p,q+1}} = \frac{I_{p,q}}{I_{p,q+1}} = \frac{t_q}{t_{q+1}} = R_{q,q+1} \]  

(2)

where, \( q \) is the index of image, \( p \) is the index of pixel and \( I \) denotes the scene radiance. Then, we can get:

\[ \frac{\sum_{n=0}^{N} c_n M_{p,q}^n}{\sum_{n=0}^{N} c_n M_{p,q+1}^n} = \frac{I_{p,q}}{I_{p,q+1}} = \frac{t_q}{t_{q+1}} = R_{q,q+1} \]  

(3)
\[
\sum_{n=0}^{N} c_n M_{p,q}^n - R_{q+1} \sum_{n=0}^{N} c_n M_{p,q+1}^n = 0
\]  

(4)

In the above equation (4), \( M \) and \( R \) are known while \( c_0 \ldots c_n \) are unknown for us. At least \( n \) equations are needed for the solution of \( c_0 \ldots c_n \). However, the homogeneous linear equations have no solutions when the coefficient matrix is full rank. When normalized the all gray value \( M \) to the range of \([0,1]\) and suppose the largest exposure is 1, we can add another constraint to the linear equations.

\[
\sum_{n=0}^{N} c_n = 1
\]

(5)

Then, we can get the polynomial coefficients by solving the nonhomogeneous equations. A total of \( n + 1 \) images are needed in the process. In fact, the sampled values of a given pixel in the images can be used for the solution. However, the distribution of the sampled value in the gray range will affect the fitting result. When the sampled values locate at the right part of the gray range, the left part of fitted curve will not precise. Therefore, a screen will be implemented before fitting and removing the uneven distributed value sets. For a set of values, if more than two of them are 255 or 0, the set should be removed. After screening, all available sample set can be applied for solving the equations. We can get multiple sets of \( c_0 \ldots c_n \). To make the fitting result more accurate and reduce the impact of noise, the mean of all \( c_0 \) is regarded as the final value. Figure 2 depicts the CRF of camera Canon EOS 60D.

3. Constructing the High Dynamic Range Radiance Map

Once the response curve \( f \) is recovered, it can be used to convert the gray values to relative radiance values. Assuming the exposure time is \( t_q \), the radiance map can be obtained by equation (6)

\[
I_q = \frac{\sum_{n=0}^{N} c_n M_{q}^n}{t_q}
\]

(6)

For precise and robustness, all available exposure images should be used to recover high dynamic range radiance values. Here, the weighting function is used to give higher weight to exposures in which the pixel’s value is closer to the middle of the response function [5]:

\[
I = \frac{\sum_{q=0}^{N} \omega(I_q) I_q}{\sum_{q=0}^{N} \omega(I_q)}
\]

(7)

where, \( q \) is the index of pixel and \( n \) denotes different radiance maps. \( \omega \) is a simple hat function:
\[ \omega(I) = \begin{cases} 
I - I_{\text{min}} & \text{for } I \leq \frac{1}{2}(I_{\text{min}} + I_{\text{max}}) \\
I_{\text{max}} - I & \text{for } I > \frac{1}{2}(I_{\text{min}} + I_{\text{max}}) 
\end{cases} \tag{8} \]

A set of multiple exposure images are shown in figure 3(a), the HDR radiance map is given in figure 3(b).

![Figure 3](image_url)

(a) LDR image sequence

(b) HDR radiance map

**Figure 3.** A set of multiple exposure images (a) and the HDR radiance map (b)

### 4. Local Tone Mapping

For a RGB image, if the tone mapping is implemented in each of the RGB space, color deviation may be occurred in the result image. Here, the RGB image is firstly translated into HSV space. Then, tone mapping is implemented in V space. At last, the image will be translated back to RGB space.

Image scale space is defined as the descriptions for the image in the full scale domain absolutely. If an image is processed in different scale domain, image features in different scales will be obtained. Here, to preserve the image details in different scales during the tone mapping process, we decompose the HDR image into detail layer and base layer. Then, local tone mapping is implemented in each of the layers.

#### 4.1. Multi-scale Decomposition

Tone mapping algorithm compress the HDR radiance map into a displayable LDR image. Here, a multiscale decomposition and histogram analysis-based method is proposed for tone mapping. A HDR scene contains both high frequency and low frequency information. Therefore, an input image is firstly decomposed into a base layer, containing only large-scale variations, and a detail layer. Then, tone mapping processing is applied to all the layers. At last, the result image is reconstructed by fusing the compressed images.

Bilateral filter is a widely used non-linear filter, where each pixel is weighted by the product of a Gaussian filtering in the spatial domain and another Gaussian filtering in the intensity domain. The bilateral filter is very effective in preserving the strong edge in the smooth intensity domain. However, it can not meet the demands of extracting image details of arbitrary scale. It can not be applied in multiscale detail extraction. In our method, an edge-preserving approach [15] based on the weighted least square (WLS) optimization framework is used for multiscale decomposition. Given an input image \( g \), another image \( u \) can be sought by minimum the formula:

\[
\sum_p \left( (u_p - g_p)^2 + \lambda \left( \alpha_x(u)(\frac{\partial u}{\partial x})_p + \alpha_y(u)(\frac{\partial u}{\partial y})_p \right) \right) \tag{9}
\]

where the subscript \( P \) denotes the spatial location of a pixel. The goal of the data term \( (u_p - g_p)^2 \) is to minimize the distance between \( u \) and \( g \), while the second term strives to achieve smoothness by
minimizing the partial derivatives of $u$. $\alpha$, and $\alpha_i$ are weights to enforce the smoothness of $u$. $\lambda$ is the balance factor and the increase of $\lambda$ will result in progressively smoother images $u$.

Let $u^l, \ldots, u^k$ denote progressively coarser version of $g$. The coarsest of these version, $u^k$ will serve as the base layer, with the $k$ detail layer defined as:

$$d^i = u^{i-1} - u^i$$

where $i = 1, \ldots, k$ and $u^0 = g$.

In figure 4, the source image is shown in the left of figure 4(a). Then, the filtered results by WLS is given with $\lambda=1, 10, 20$, respectively. As we can see, with the increase of $\lambda$, the result image is more and more smooth. The last image is regarded as base image which is shown in the left of figure 4(b). The rest part of figure 4(b) are detail layers.

![Figure 4](image)

(a) Source image and filtered images  
(b) Base layer and detail layers

**Figure 4.** (a) From left to right: source image, filtered images with $\lambda=1, 10, 20$ and (b) From left to right: base layer, detail layers of (a).

### 4.2. Local Tone Mapping

For the obtained base layer and detail layers, most image details are included in the detail layers and the base layer reflects the overall brightness of the scene area. To preserve details of the image while compressing the dynamic range, nonlinear compression is only implemented for the detail layers and linear compression is applied for base layer.

![Figure 5](image)

(a) Histogram of an HDR radiance map  
(b) Tone mapping curve

**Figure 5.** Histogram of an HDR radiance map (a) and tone mapping curve (b) of image in (a)

As shown in figure 5(a), for a detail layer, the main domain of the histogram locates at the left part of the radiance value range. Image details are mainly expressed by these radiance values. The higher radiance value in the right part of the histogram make less contribution to image detail. Therefore, the left part of the histogram should be slightly compressed and the right part should be seriously compressed. The compression radio should be related to the quantity of the radiance value. The larger the quantity of the value, the greater the compression ratio. Here, a cumulative distribution function (CDF) based mapping method is applied. As shown in figure 6, the quantity of the value in radiance map can be expressed by the increment in CDF of two adjacent point. Due to the fact that details of regions corresponding to radiance values with large increment are abundant, we set the compression radio inversely proportional to the increment. After linear compression of the ordinate of the CDF, we make it the tone mapping curve. For the histogram in figure 5(a), the mapping curve is shown in figure 5(b).
After the tone mapping, the result image $g$ can be easily recovered by simply adding up the base and detail layers:

$$g = b + \sum_{i=1}^{k} d^i$$

(11)

where $b$ and $d^i$ are compressed base layer and detail layers.

5. Experimental Results and Analysis

5.1. Local Tone Mapping

In order to demonstrate the effectiveness of the proposed method, many image sequences are captured using Canon EOS 60D camera for high dynamic range imaging. In figure 7(a), a set of multiple exposure images are captured with 1s, 1.6s, 2.5s, 4s, 5s, 6s, 8s and 10s respectively. As is shown, with the increase of the exposure time, image brightness increases. However, no image else could perfectly express the scene detail. Figure 7(b) is the linear compressed image of reconstructed HDR map. Because the dynamic range of the HDR map is much larger than 256 and the image is dark in general, there is a few details in the bright region when linear compressed. Figure 7(c) is the tone mapping result of the proposed method. Both bright and dark regions are abundant in detail. The image could perfectly express the high dynamic scene.

For simulations, we compare the proposed method with previous methods [5], [16], [17] perceptually. The fundamental method [5] is chosen as baseline. Others are state-of-the-art fusion methods with the ghost removal [16], [17]. Those source codes are available on the authors’ website. Free parameters of
the proposed method are experimentally determined. The input image sequence of ‘Desk’ is shown in figure 8(a). Shutter speed settings are simulated 1/500, 1/125, 1/30, 1/8 in seconds. The HDR imaging results of each methods are shown in figure 8(b).

![Input image sequence of ‘Desk’](image)

![HDR imaging results with methods [5], [16], [17] and proposed method](image)

**Figure 8.** HDR images of different HDR imaging methods

To better illustrate the validity of this method, an objective comparison is implemented with other state-of-the-art HDR imaging methods. HDRVDP-2[18] is used for measurement. HDR-VDP-2 scores the similarity between true and fused HDR images and a perceptual quality of fused one based on the statistics is analyzed. Table 1 describes the scores of each method which indicates us that the proposed method outperforms other methods obviously.

|                | Ref. [5] | Ref. [16] | Ref. [17] | Proposed |
|----------------|----------|-----------|-----------|-----------|
| Cathedral      | 59.94    | 54.24     | 69.18     | 70.21     |
| Desk           | 58.84    | 53.61     | 68.26     | 70.33     |
| Memorial       | 61.62    | 55.08     | 67.6      | 69.92     |
| Night atrium   | 67.11    | 57.09     | 68.59     | 70.38     |
| Tree           | 57.45    | 51.58     | 66.79     | 66.80     |

**Table 1.** HDRVDP-2 scores of each method

5.2. Analysis of the order of polynomial

In the proposed method, the CRF is modeled with a high-order polynomial. To recover the CRF precisely, an appropriate order should be chosen for polynomial. Here, the order \( n \) is determined by analyzing the error of the polynomial. Suppose the CRF is \( n \)-order differentiable on \([0,1]\) and \((n+1)\)-order differentiable on \((0,1)\), then for \( x \in [0,1] \), \( f(x) \) an be expressed as:

\[
f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + R_n(x)
\]

(12)

This is the famous Taylor formula. \( R_n(x) \) is the error term which can be expressed by Lagrange remainder.
Here $\theta \in [0,1]$. If $R_n(x)$ is less than an acceptable threshold, the fitted curve is accurate enough for our experiments. Here, $n = 3$ is selected for analysis. Four images are captured for CRF fitting. Supposing the sampling points are uniformly distributed, the maximum distance between any point and sampling point is 0.1. For any point $x \in [0,1]$, $f(x) < 1$. From the analysis above, we know that $R_n(x)$ is less than 0.00001. Therefore, $n = 3$ is precise enough for CRF fitting. In all our experiments, the polynomial order is set 3.

6. Conclusion
A high efficient HDR imaging method is proposed in this paper. Firstly, the CRF is modelled with a high-order polynomial. The recovery of CRF can be transformed into the solution of the polynomial coefficients. Then, HDR image can be obtained by weight summation of the relative radiance maps. After that, a local tone mapping algorithm is proposed based on the analysis of the histogram. In the result image, the local contrast of the image is preserved while compressing the dynamic range. Experimental results demonstrate that the proposed method outperforms other method obviously.

References
[1] Tocci M D, Kiser C, Tocci N and Sen P 2011 A versatile HDR video production system ACM Trans. Graphics 30 p 41
[2] Lin D, Huayan S, Tinghua Z and Shuai W 2016 Calibration camera response function model algorithm based on principal component analysis Infrared and Laser Engineering 10 pp 251-9
[3] Huo Y Q and Zhang X D 2016 Single image-based HDR imaging with CRF estimation 2016 Int. Conf. On Communication Problem-Solving (ICCP) IEEE pp 1-3
[4] Lin D, Huayan S, Tinghua Z and shuai W 2016 Camera Response Function Calibration Algorithm Based on Single Frame Image Acta Optica Sinica 07 pp 114-22
[5]Debevec P E and Malik J 2008 Recovering high dynamic range radiance maps from photographs in ACM SIGGRAPH 2008 Classes ACM p 31
[6]Robertson M A, Borman S, and Stevenson R L 1999 Dynamic range improvement through multiple exposures in Proc. Int. Conf. on Image Processing (ICIP’99) IEEE 3 pp 159–63
[7]Grossberg M D and Nayar S K 2003 Determining the camera response from images: what is knowable? IEEE Trans. Pattern Anal. Mach. Intell. 25(11) pp 1455–67
[8]Lu P Y, Huang T H, Wu M S, Cheng Y T and Chuang Y Y 2009 High dynamic range image reconstruction from hand-held cameras IEEE Conf. on Computer Vision and Pattern Recognition IEEE pp 509–516
[9]Mitsunaga T and Nayar S K 1999 Radiometric self calibration IEEE Computer Society Conf. on Computer Vision and Pattern Recognition IEEE 1 pp 374-80
[10]Reinhard E, Stark M, Shirley P and Ferwerda J 2002 Photographic tone reproduction for digital images ACM Trans. Graphics 21(3) pp 267–76
[11]Shan Q, Jia J and Brown M S 2010 Globally optimized linear windowed tone mapping IEEE Trans. Visual Comput. Graphics 16(4) pp 663–75
[12]Kim B K, Park R H and Chang S K 2016 Tone mapping with contrast preservation and lightness correction in high dynamic range imaging Signal, Image and Video Processing 10(8) pp 1425-32
[13]Larson G W, Rushmeier H and Piatto C 1997 A visibility matching tone reproduction operator for high dynamic range scenes IEEE Trans. Visual Comput. Graphics 3(4) pp 291–306
[14]Yoshida A, Blanz V, Myszkowski K and Seidel H P 2005 Perceptual evaluation of tone mapping operators with real-world scenes Int. Electronic Imaging Society for Optics and Photonics pp 192-203
[15] ARBMAN Z, FATTAL R, LISCHINSKI D and Szeliski R 2008 Edge-preserving decompositions for multi-scale tone and detail manipulation *ACM Transactions on Graphics* 27(3) pp 1-10

[16] Sen P, Kalantari N K, Yaeouabi M, Darabi S, Goldman D B, and Shechtman E 2012 Robust patch-based HDR reconstruction of dynamic scenes *ACM Trans. Graph.* 31(6) p 203

[17] Kimura D, Kondo A, Yoshida T and Iwahashi M 2016 High dynamic range imaging based on camera response estimation and multi-gradients fusion 2016 *Signal and Information Processing Association Annual Summit and Conference (APSIPA)* Asia-Pacific pp 1-4

[18] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, “HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions,” *ACM Trans. Graph.*, vol. 30, no. 4, pp. 40:1–40:14, 2011.