Manifold Training Technique to Reconstruct High Dynamic Range Image

Cheng-Yuan Liou* and Wei-Chen Cheng

Department of Computer Science and Information Engineering
National Taiwan University
Republic of China
Supported by National Science Council
correspondence: cyliou@csie.ntu.edu.tw

Abstract. This paper presents two manifold training techniques to reconstruct high dynamic range images from a set of low dynamic range images which have different exposure times. It provides the performance on noisy images.
Keyword–SIR algorithm, SOM, HDR image, high dynamic range

1 Introduction

The ordinary digital camera is a low dynamic range device. The intensity of the environment scene may have a very wide dynamic range. It may exceed the camera's range limit, 255. Those intensity values which exceed the limit will be set to 0 or 255. Many efforts have been done to recover the high dynamic range (HDR) images with varying degrees of success. Many camera systems transform the sensor exposure value of the CCD (Charge-Coupled Device) through a non-linear function which is called the camera response function (CRF) and record the transformed value as the restored scene intensity. Since this function may not be available from the manufacturer, the key to obtain the HDR image is to recover the CRF. With this CRF one can produce the 'real' time-invariant irradiance of the scene. The method in [1] shows how to reconstruct the CRF from a series of images which are taken from the same scene with different exposures. It develops a parametric model for the CRF. The method in [2] uses a series of digital pictures and solves a set of linear equations to estimate the inverse of the CRF. Those pictures are taken with a fixed aperture and different known shutter speeds. Debevec's method [2] is not a parametric model. It assumes the inverse of the CRF is smooth. Mitsunaga [3] proposed an iterative method to adjust the coefficients of a high-order polynomial to fit the CRF. In this work, we devise two manifold training techniques to obtain the HDR without any irradiance information. One is based on SIR method [4][5], the second is a relaxation method similar to SOM [6]. The technique based on SIR will not use the continuous polynomial [3] and not use the smooth assumption [2]. Without matrix decomposition, the SIR and SOM are relatively easy in implementation.

2 Camera Model

Suppose there are N pictures with different exposures taken from the same scene. We assume these images are aligned. Therefore the same pixel location of all images should correspond to the same point of the scene. Each image has P pixels. For an 800×600 image, P is equal to 480000. The difference among images is the shutter speed setting. All images are taken with the same aperture setting. The different exposure time, Δt, can be obtained by varying the shutter speed. Let Δt_j denote the exposure time of the jth image. The sensor exposure value, X_{ij}, of the ith pixel in the jth image can be modeled as

$$X_{ij} = E_i \Delta t_j, \; i \in \{1, \ldots, P\}, \; j \in \{1, \ldots, N\}. \tag{1}$$

The E_i is the sensor irradiance of the ith pixel. The X_{ij} is the output of the ith CCD unit during the jth photo image. The unit of X_{ij} is Jm^{-2}. After cutting out all large and small intensity values of the outputs of the CCD that exceed the range limits, the rest values are passed through the CRF and are digitalized (quantized). A function f is used for representing the whole quantization process,

$$Z_{ij} = f\left(X_{ij}\right) = f\left(E_i \Delta t_j\right), \; Z_{ij} \in \{0, \ldots, 255\}. \tag{2}$$

The Z_{ij} is the intensity value which is finally stored in the storage device. We can rewrite (2) with an inverse function and take log of both sides,

$$\ln f^{-1}\left(Z_{ij}\right) = \ln X_{ij} = \ln E_i + \ln \Delta t_j. \tag{3}$$

Defining a $g = \ln f^{-1}$, (3) can be written as

$$g\left(Z_{ij}\right) = \ln E_i + \ln \Delta t_j. \tag{4}$$

3 Manifold training

3.1 SIR method

We use the SIR method [4][5] to solve the function $g = \ln f^{-1}$. For the ith pixel, the SIR energy function is

$$O_i = \frac{1}{4} \sum_{k=1}^{N} \sum_{r=1}^{N} \left(\left(g\left(Z_{ik}\right) - g\left(Z_{ir}\right)\right)^2 - \left(\ln \Delta t_k - \ln \Delta t_r\right)^2 \right)^2, \tag{5}$$

where $Z_{ik}, Z_{ir} \in \{0, \ldots, 255\}$. In this energy, O_i, we assume that the image pixels at the same location have the same or similar irradiance value. Then we have the idea form

$$\begin{aligned}
& g\left(Z_{ik}\right) - g\left(Z_{ir}\right) \\
&= \ln E_i + \ln \Delta t_k - \ln E_i - \ln \Delta t_r \\
&= \ln \Delta t_k - \ln \Delta t_r, \tag{6}
\end{aligned}$$

where $k, r \in \{1, \ldots, N\}$. We plan to utilize the energy O_i to seek a solution $g(Z_{ij})$ that satisfies the idea form. One way to minimize (5) is to adjust $g(Z_{ik})$ and $g(Z_{ir})$ toward the gradient descent direction. Differentiate O_i with respect to $g(Z_{ik})$,

$$\frac{\partial O_i}{\partial g(Z_{ik})} = \sum_{r=1}^{N} \left(\left(g(Z_{ik}) - g(Z_{ir}) \right)^2 - \left(\ln \Delta t_k - \ln \Delta t_r \right)^2 \right) \left(g(Z_{ik}) - g(Z_{ir}) \right).$$
(7)

Differentiate O_i with respect to $g(Z_{ir})$,

$$\frac{\partial O_i}{\partial g(Z_{ir})} = -\sum_{k=1}^{N} \left(\left(g(Z_{ik}) - g(Z_{ir}) \right)^2 - \left(\ln \Delta t_k - \ln \Delta t_r \right)^2 \right) \left(g(Z_{ik}) - g(Z_{ir}) \right).$$
(8)

The SIR method is briefly described as follows:

1. Randomly initialize the function g with a discrete form.
2. Randomly select a pixel i from $\{1, \ldots, P\}$.
3. Update $g^{T+1}(Z_{ik}) = g^T(Z_{ik}) - \eta \frac{\partial O_i}{\partial g(Z_{ik})}$ and $g^{T+1}(Z_{ir}) = g^T(Z_{ir}) - \eta \frac{\partial O_i}{\partial g(Z_{ir})}$, for a pair (k, r) selected from the N images, $k, r \in \{1, \ldots, N\}$. η is the training rate and T is the number of training epoch.
4. Gradually decrease η and repeat Step 2-4.

Note that the discrete form g is much more flexible to operate with by the SIR method than the continuous high-order polynomial and smooth functions used in other methods. We expect that g will approximate the 'real' $\ln f^{-1}$ when the training time is long enough. The concept is illustrated in Figure 1(a). Once the whole discrete form function $g(Z_{ij})$ is determined, we can calculate the irradiance, E_i, for every pixel by using the formula in [2],

$$\ln E_i = \frac{\sum_{j=1}^{N} w(Z_{ij}) \left(g(Z_{ij}) - \ln \Delta t_j \right)}{\sum_{j=1}^{N} w(Z_{ij})}, \ i \in \{1, \ldots, P\}.$$
(9)

The w is a weighting function. We set $w(x) = e^{\frac{-1}{80} \|x - 127\|}$ in this paper. The recovered HDR image will include the irradiance maps, $\{E_i, i = 1, \ldots, P\}$. We will use the tone mapping [7] to display the HDR images in all experiments.

3.2 Relaxation using the self-organization method

The CRF can also be obtained by a relaxation method similar to the self-organizing map (SOM) [6]. Suppose there are 256 cells regularly aligned on a straight line, the marked horizontal axis in Figure 1(b). Each cell has a single weight. The mth weight value is $g(m)$. The neighborhood function h in SOM is set as

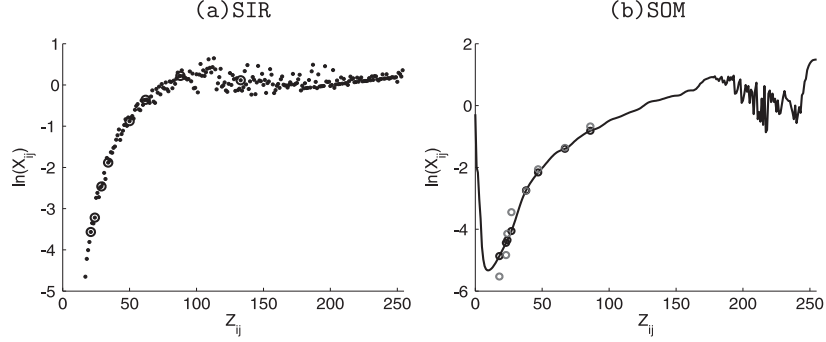$$h(u, v) = e^{-\left(\frac{u - v}{\sigma} \right)^2},$$
(10)

Fig. 1. (a) The concept of SIR training. The 256 points are obtained and updated during each training epoch. The eight black circle dots are those points for the ith pixel, $\{(Z_{ij}, g^{50}(Z_{ij})), j \in \{1, 2, \ldots, 8\}\}$. (b) The concept of SOM training. The grey circles are the values $\{(\ln E_i)^{50} + \ln \Delta t_j, j \in \{1, \ldots, N\}\}$.

where $u, v \in \{0, \ldots, 255\}$ and $\sigma \in R$. The σ is a parameter which controls the size of neighborhood. We suppose each pixel i has it own irradiance, E_i. In each epoch T, the current estimate of $\ln E_i$ is

$$(\ln E_i)^T = \frac{1}{N} \sum_j \left(g^T(Z_{ij}) - \ln \Delta t_j\right). \tag{11}$$

Based on this estimate, the g can be updated by

$$\Delta g^T(m) = h(m, Z_{ij}) \left[\left((\ln E_i)^T + \ln \Delta t_j\right) - g^T(m)\right], \tag{12}$$

and

$$g^{T+1}(m) = g^T(m) + \eta \Delta g^T(m), \ m \in \{0, \ldots, 255\}, \tag{13}$$

where η is the training rate. Figure 1(b) shows an example of the self-organizing CRF curve during the 50th training epoch. The eight black circles denote the pairs, $\{(Z_{ij}, g^{50}(Z_{ij})), j = 1, \ldots, 8\}$ for a specific pixel i. The eight gray circles are the values $\{(\ln E_i)^{50} + \ln \Delta t_j, j \in \{1, \ldots, N\}\}$. The SOM method randomly selects a pixel i from the jth image, then uses the Z_{ij} to update the function g by (12) and (13). The training epoches are repeated until the curve g is converged. The irradiance maps, $\{E_i, i = 1, \ldots, P\}$, are then calculated using the formula (9).

4 Experiments

We have two sets of images. One is a scene of buildings and the other is of natural scenery. Figure 2 and Figure 3 plot the inverse CRFs using the building images and the natural scenery images. The red, green and blue lines (points) represent the three inverse CRFs of the RGB channel, respectively. The vertical

axis, $\ln X_{ij}$, is defined in (3). Figure 2 and Figure 3 also show the HDR images obtained by the two methods. The result of Debevec's algorithm is presented for comparison. We randomly sample 300 pixels to solve the linear equations and the parameter λ is set to 15, $\lambda = 15$, in Debevec's algorithm [2]. Note that the SIR and SOM use all pixels to solve the CRF.

We show that the SIR method can recover the inverse CRF when the images are corrupted. The noisy images contain normal distributed noises whose variance is $\sigma = 0.089$, see Figure 4(a). Figure 4(b) is the CRF trained by the SIR method using the images in Figure 4(a). Figure 4(c) is obtained by using the method in [2]. The parameter λ in [2] is set to 15, $\lambda = 15$. Three hundred selected pixels are used in solving the linear equations. Figures 4(d,e) show the HDR images constructed by using the noisy images in Figure 4(a) and the CRFs in Figure 4(b,c). Figures 4(f,g) show the HDR images constructed from the images without noise, $\sigma = 0$, and the CRF in Figure 4(b,c). Figure 4(f) shows better image on the right-top dark corner. We also used the software [8] to solve [3] the noise images, it can not recover the three CRFs from noisy images.

In summary, this paper proposes two manifold techniques to reconstruct HDR images. The trained CRF can be used in estimating the irradiance value from a series of photos with different exposures. Furthermore, we test the performance of the SIR method using images with heavy noise. The experimental results show that the SIR method can recover the CRF from noisy images. The reconstructed HDR image has many potential applications, such as film, astronomy image, and medical imaging.

References

1. Mann, S., Picard, R.: On being 'undigital' with digital cameras: extending dynamic range by combining differently exposed pictures. In: Proc. IS&T 46th Ann. Conf. (1995) 422–428
2. Debevec, P., Malik, J.: Recovering high dynamic range radiance maps from photographs. In: Proceedings of the 24th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co. New York, NY, USA (1997) 369–378
3. Mitsunaga, T., Nayar, S.: Radiometric self calibration. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Volume 1. (Jun 1999) 374–380
4. Liou, C.Y., Chen, H.T., Huang, J.C.: Separation of internal representations of the hidden layer. In: Proceedings of the International Computer Symposium, ICS, Workshop on Artificial Intelligence. (2000) 26–34
5. Liou, C.Y., Cheng, W.C.: Manifold construction by local neighborhood preservation. In: International Conference on Neural Information Processing, LNCS 4985. (2008) 683–692
6. Kohonen, T.: Self-organized formation of topologically correct feature maps. Biological Cybernetics **43**(1) (1982) 59–69
7. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. In: Proceedings of the 29th annual conference on Computer graphics and interactive techniques. (2002) 267–276
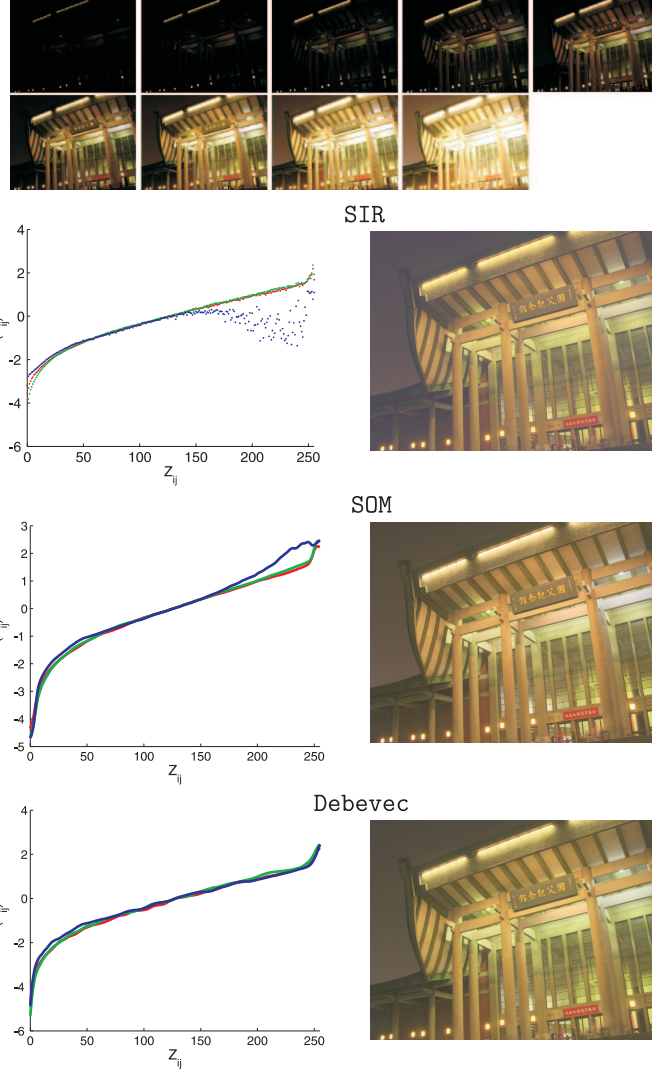8. http://www1.cs.columbia.edu/CAVE/software/rascal/rrhome.php

Fig. 2. These CRFs are obtained by using the same image set and the same camera settings. Three inverse CRFs obtained by SIR, SOM and Debevec's method. The nine small size images on top are taken with different exposure times in the night. The three HDR images on right are obtained by the three methods.
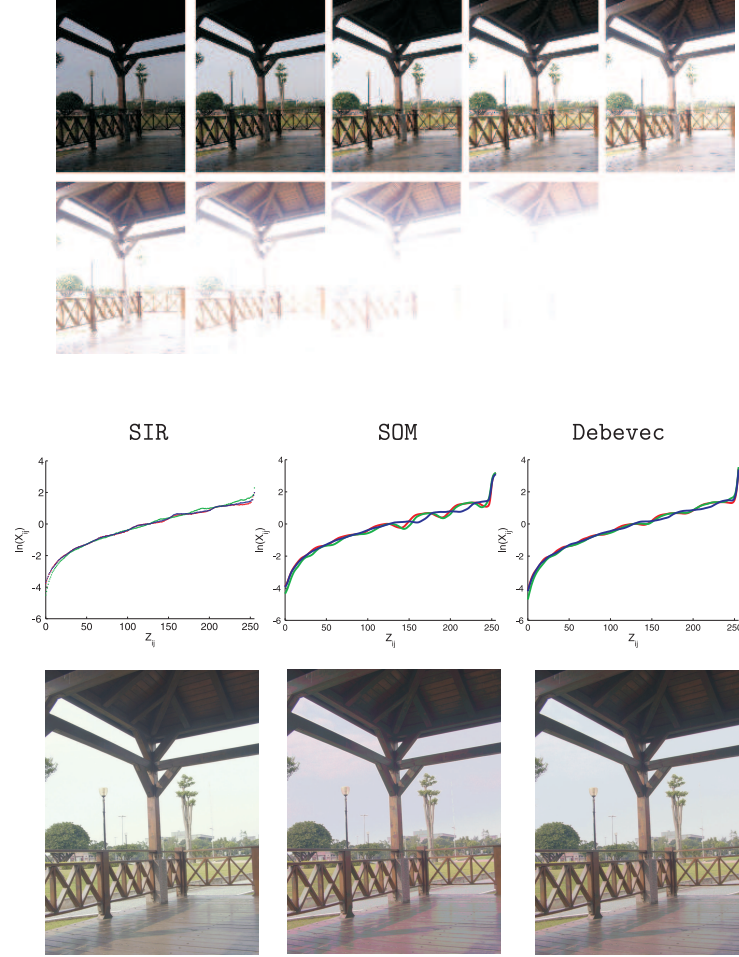
Fig. 3. The images are taken from the nature scenery. The nine small size images on top are the scenery images during sunlight. Three HDR images are reconstructed. The color of the sky in Debevec's image and in SOM image tends to be bluer than the SIR image.
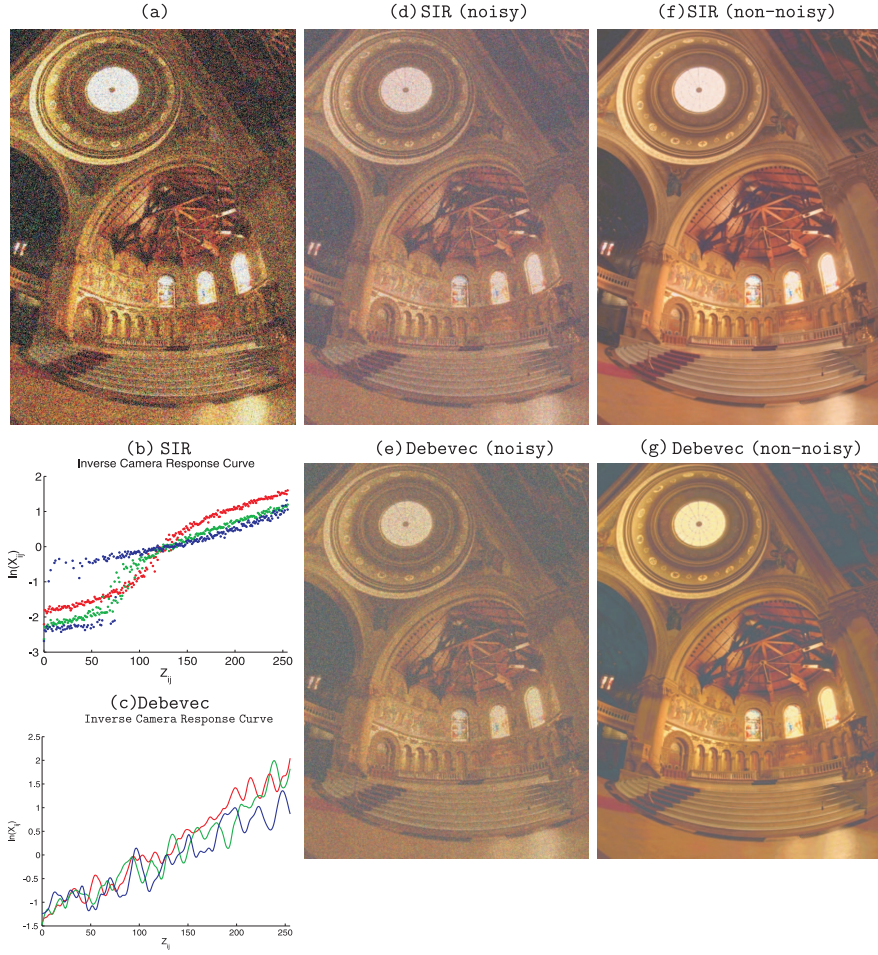
Fig. 4. (a) One noisy image in a series of photos. **(b)** The CRF reconstructed by SIR. **(c)** The CRF recovered by Debevec's method. **(d)** The HDR image by SIR. **(e)** The HDR image by Debevec's method. **(f)** The HDR image using the clean image, $\sigma = 0$, and the CRF in (b). **(g)** The HDR image using the clean image, $\sigma = 0$, and the CRF in (c).