

HIGH DYNAMIC RANGE MAP ESTIMATION VIA FULLY CONNECTED RANDOM FIELDS WITH STOCHASTIC CLIQUES

F. Y. Li, M. J. Shafiee, A. Chung, B. Chwyl, F. Kazemzadeh, A. Wong, and J. Zelek

Vision & Image Processing Lab, Systems Design Engineering Dept., University of Waterloo
{f27li, mjshafie, agchung, bchwyl, fkazemzadeh, a28wong, jzelek}@uwaterloo.ca

ABSTRACT

The reconstruction of high dynamic range (HDR) images via conventional camera systems and low dynamic range (LDR) images is a growing field of research in image acquisition. The radiance map associated with the HDR image of a scene is typically computed using multiple images of the same scene captured at different exposures (i.e., bracketed LDR images). This approach, though inexpensive, is sensitive to noise under high camera ISO. Each bracketed image is associated with a different level of noise due to the change in exposure time, and the noise is further amplified when tone-mapping the HDR image for display. A new framework is proposed to address the associated noise in the context of random fields. The estimation of the HDR image from a set of LDR images is formulated as a stochastically fully connected conditional random field where the spatial information is incorporated to compute the HDR value in combination with the LDR image values. Experimental results show that the proposed framework compensated the non-stationary ISO noise while preserving the boundaries in the estimated HDR images.

Index Terms— High Dynamic Range Imaging, Conditional Random Fields, Image Denoising, HDR Reconstruction, SFCRF

1. INTRODUCTION

High dynamic range (HDR) imaging has recently become a growing area of research. While the human visual system is able to interpret scenes with a high dynamic range of illumination, cameras are unable to properly capture these scenes due to the limited dynamic range of conventional sensor arrays found in digital cameras. As such, details within regions of very high illumination and/or very low illumination are lost. HDR imaging captures a wider range of illumination, allowing for the representation of additional detail in scenes with extreme illumination. The applications of HDR imaging are widespread and include remote sensing [1], computer



Fig. 1: Example of LDR images captured at different exposures (first row) and the associated HDR reconstructed toned-map image by Debevec framework [5] (second row), where the right image shows a zoomed view of the marked area in the left image. As seen the reconstructed image is highly contaminated by noise.

graphics [2], physically based rendering [3, 4], and various image processing algorithms [5].

HDR imaging can be achieved with hardware via high dynamic range sensors [6]; however, cameras with HDR sensors are relatively expensive, making them undesirable for practical application. As a result, several algorithms have been proposed for reconstructing HDR images from low dynamic range (LDR) images taken using conventional imaging equipment [5, 7, 8]. The standard approach to HDR reconstruction computes a radiance map of the scene using multiple LDR images of the same scene captured at different exposures. These images, commonly referred to as bracketed images, can then be combined using the camera response function to generate an HDR image.

Debevec and Malik [5] approximated the camera response function by use of pixel intensities across different exposures via least squares optimization. The final radiance map was obtained by applying the approximated response function and a triangular “hat-shaped” weighting function to lower contributions from over-exposed or under-exposed regions (pixel values near zero or 255). Robertson *et al.* [7] determined

This work was supported by the Natural Sciences and Engineering Research Council of Canada, Ontario Ministry of Research and Innovation, and the Canada Research Chairs Program.

the camera response function probabilistically and weighted images taken at higher exposure times more heavily. Lastly, Mitsunaga and Nayar [8] proposed a method to derive the response function without explicit knowledge of the exposure times via a parametric polynomial model for curve fitting.

While performance is generally good given bracketed images with minimal noise, these methods are sensitive to embedded noise. Often times fast shutter speeds and high ISO settings are used for scenes with changing light conditions, such as outdoor or dynamic scenes. However, using high ISO tends to cause noisy image captures. Various sources of noise affect digital images and may arise during image acquisition, transmission, or processing. Three primary sources of noise are present in digital cameras: photon shot noise, dark current noise, and read noise [9]. While dark current noise and read noise are dependent on digital camera design, photon shot noise is directly affected by ISO settings. For the remainder of this paper, any noise amplified through the use of high ISO will be referred to as 'ISO noise'. As shown in Figure 1 the constructed HDR image by conventional methods is highly sensitive to ISO noise. Due to the changing exposure times, a different level of ISO noise is present in each bracketed image and is subsequently amplified by the tone mapping process used to constrain pixels back to standard image values (i.e., between zero and 255). Since the noise level of each bracketed image is different, the associated noise of the HDR image is non-stationary. To better account for noise, methods have been proposed that performed weighted averaging on the bracketed images as a preprocessing step to the standard HDR reconstruction process [9, 10].

Methods that explicitly denoise bracketed images have also been proposed [11, 12, 13]. Rameshan *et al.* [11] used a Bayesian method and maximum a posteriori (MAP) formulation to perform denoising in the HDR domain. Goossens *et al.* [12] modeled the sensor noise with a Poisson distribution to denoise the bracketed images. Hasinoff *et al.* [13] proposed an imaging framework for acquiring a set of images to optimize worst case SNR. However, these methods assume a static level of ISO noise across all bracketed images, resulting in inconsistent noise levels in the HDR image.

The aforementioned methods utilized a denoising method as a pre-process or a post-process to HDR reconstruction. Here we present a novel framework for HDR reconstruction that simultaneously creates the HDR map while compensating for non-stationary ISO noise using a stochastically fully connected conditional random field (SFCRF). The SFCRF enforces a consistency constraint across pixels that are spatially compact and similar in intensity, enabling each bracketed image to be denoised dynamically while creating the HDR image. The clique connectivities are formed based on the stochastic clique formation framework. Thus, the appropriate long-range clique connectivities are formed in the SFCRF, improving the model accuracy while relaxing the computational complexity of high-order clique connectivity.

2. METHODS

To compensate for non-stationary ISO noise, we propose a novel framework for HDR reconstruction via a SFCRF. The creation of a HDR image is modeled as a conditional probability given a set of LDR images captured using different exposure times. LDR images are usually captured using common digital cameras that tend to be noisy under low light conditions and high sensitivity settings. Thus, the LDR images have varying levels of ISO noise, resulting in inconsistent noise levels across the HDR image.

We model the HDR estimation as a maximum a posteriori (MAP) optimization where the HDR map is estimated to maximize the conditional probability of the HDR map given LDR images. The proposed framework utilizes a SFCRF [14] to address the non-stationary ISO noise by incorporating long-range spatial information to create the HDR image. Given the LDR images, the conditional probability of the HDR map is formulated as

$$P(H|B) = \frac{1}{Z(B)} \exp(\psi(H, B)) \quad (1)$$

where H is the estimated HDR image, $B = \{B_1, B_2, \dots, B_m\}$ is the set of m LDR images with different exposure times, $Z(B)$ is the normalization constant. $\psi(\cdot)$ is the potential function encoding the relationship between pixels in the HDR map H :

$$\psi(H, B) = \sum_{i=1}^n \psi_u(h_i, B) + \sum_{c \in C} \psi_p(h_c, B) \quad (2)$$

where h_i represents pixel i in the HDR image containing $|H| = n$ pixels, and $\psi_u(h_i, B)$ is the unary potential representing the likelihood of each pixel i and its corresponding pixel intensities in the observed LDR images. The spatial relationships between pixels in the HDR image are formulated by $\psi_p(h_c, B)$, where h_c represents a set of pixels that construct a clique c in the set of all stochastic cliques C .

The unary potential encodes the likelihood of the same pixel across a set of differently exposed images to its associated HDR value. The Debevec and Malik [5] camera response approach is applied to formulate the unary potential in the random field:

$$\psi_u = \ln(h_i) = \frac{\sum_{j=1}^l w(b_{ij})(g(b_{ij}) - \ln \Delta t_j)}{\sum_{j=1}^l w(b_{ij})} \quad (3)$$

where b_{ij} is the i^{th} pixel in B_j , $g(\cdot)$ represents the inverse camera response function that maps the LDR value to a HDR value based on the exposure time, Δt_j encodes the exposure time of LDR image j , and $w(\cdot)$ represents a weighting function to lessen the contribution of pixels that are under-exposed or over-exposed (near zero or 255):

$$w(b) = \begin{cases} b - b_{min} & b \leq \frac{1}{2}(b_{min} + b_{max}) \\ b_{max} - b & b > \frac{1}{2}(b_{min} + b_{max}) \end{cases} \quad (4)$$

similar to Debevec and Malik [5], $w(\cdot)$ resembles a simple hat function centered between b_{min} and b_{max} , the minimum and maximum of LDR intensity values.

The spatial relationships between pixels in the HDR image are modeled by a fully connected conditional random field where the clique connectivities are constructed based on the stochastic clique framework proposed by [14]:

$$C = \{(i, j) | \mathbb{1}_{\{i, j\}} = 1\} \quad (5)$$

where C is the set of all pairwise cliques and $\mathbb{1}_{\{i, j\}}$ represents the stochastic clique indicator (SCI) function. The SCI encodes a stochastic function that determines if two nodes can construct a clique in the random field based on its underlying probability distribution. The underlying probability distribution of the SCI is based on the spatial similarity and color intensity similarity of two nodes in the random field:

$$\mathbb{1}_{\{i, j\}} = \begin{cases} 1 & \frac{\exp(-E_d(i, j)) \cdot \exp(-E_c(i, j))}{\gamma} \geq r \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where $E_d(\cdot)$ and $E_c(\cdot)$ represent spatial distance and color dissimilarity between two nodes, respectively; γ encodes the sparsity of the conditional random field and r is a random number selected from a uniform distribution over the unit interval.

The motivation to utilize the long-range clique connectivities is to address the smoothing problem associated with local random fields while compensating underlying noise of the HDR image by incorporating more information in the computation procedure. The proposed approach is a unified framework which reconstructs the HDR image while compensating the associated noise. As shown in Figure 2, long-range interaction connectivities are incorporated into the model by applying the stochastic clique formation. The proposed framework provides more appropriate clique interactions and reduces the computational complexity associated with long-range clique connectivity in conditional random fields via the sparse nature of stochastic cliques. The clique interactions are formed by considering the similarity between its associated nodes, allowing for the SFCRF framework to model the underlying noise of LDR images implicitly.

3. RESULTS

3.1. Experimental Setup

A Canon T3i DSLR camera was used to capture bracketed images with ISO 6400 to increase the amount of ISO noise seen in the LDR images. The proposed framework was evaluated under different situations; reported here as demonstration are three scenes including a Macbeth Colorchecker chart, an outdoor scene of a tree, and an indoor scene of a stack of books. The standard Colorchecker was utilized to compare methods quantitatively. The averaged of signal-to-noise (SNR) ratio

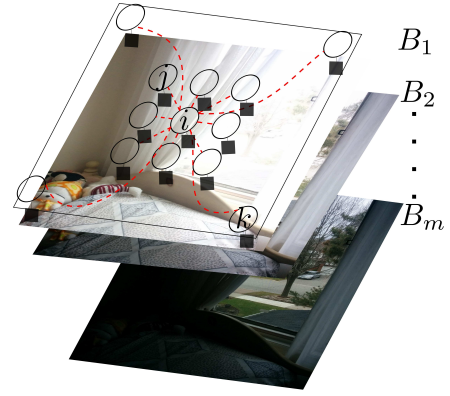


Fig. 2: The proposed SFCRF framework to estimate a HDR map. All LDR images, $\{l_1, \dots, l_m\}$, are considered as the measurements which the actual HDR map value for each pixel is estimated by use of corresponding LDR values while considering the pixel within its neighbors. Each node such as i can be connected to other nodes (i.e., i or k) in the random field by a chance based on their similarity.

of all blocks in the Colorchecker board is reported as quantitative analysis. The camera parameters for each scene is summarized in Table 1 and the sequence of bracketed images of the Colorchecker is shown in Figure 3. Since Debevec’s original HDR reconstruction algorithm [5] was applied as the unary potential this method was evaluated as the comparison method. The Debevec’s algorithm was applied by use of Banterle’s MATLAB implementation [15], where the camera response function is computed using the original MATLAB code from Debevec’s paper. The Mantiuk tonemap operator [16], implemented in the open source software ‘Luminance HDR’ [17], was used to tonemap the HDR image back to the LDR domain for display. The choice of tonemapping operator was simply for illustrative purposes.

3.2. Experimental Results

Figure 4 shows a zoomed in view of a patch of the Macbeth Colorchecker results after HDR reconstruction and tonemapping. As seen the proposed method estimated the pixel intensities much more homogeneously in the smooth regions compared to the standard method [5]. The average SNR across all colour patches and colour channels is calculated in the radiance domain after HDR reconstruction, where the proposed method shows a higher average SNR by 1.9 dB. Higher noise was observed in underexposed areas; in the blue patch the proposed SFCRF method shows a higher SNR of about 9.5 dB.

The reconstructed HDR images of an outdoor and indoor scene are shown in Figure 5. The reported results demonstrate that the proposed method is able to compute the correct HDR image while preserving boundary details and addressing the associated ISO noise of the image. Referring to the tree scene, a zoomed in view of a building corner is shown where the standard HDR reconstruction method shows noticeable noise in the sky, giving a ‘spotty’ look. The proposed method shows a much smoother view of the sky while maintaining the build-



Fig. 3: The bracket images (LDR) corresponding to the Colorchecker with different exposure times.

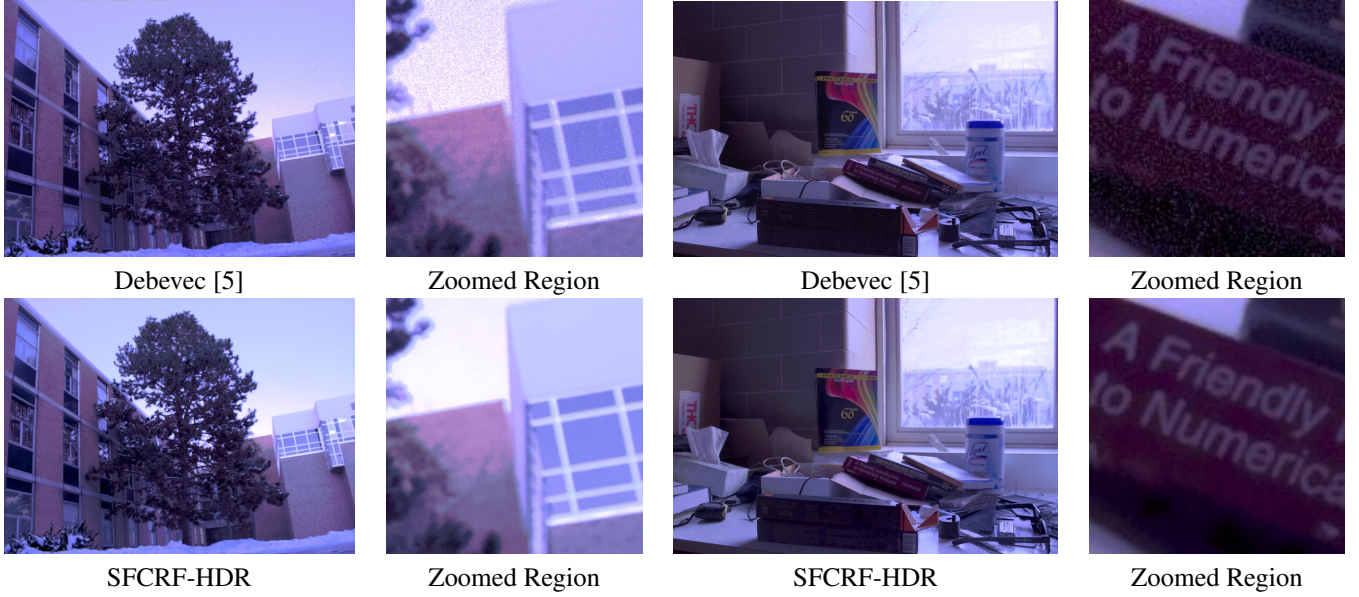


Fig. 5: Example of the estimated HDR images of natural scenes by the SFCRF-HDR compared to the Debevec & Malik method [5]. The first and third columns demonstrate the whole scene and the second and fourth columns shows the zoomed regions.

ing edge. The zoomed in view of the book scene illustrates the SFCRF HDR computation framework can preserve very fine boundary structures as well as compensate non-stationary noise in the image.

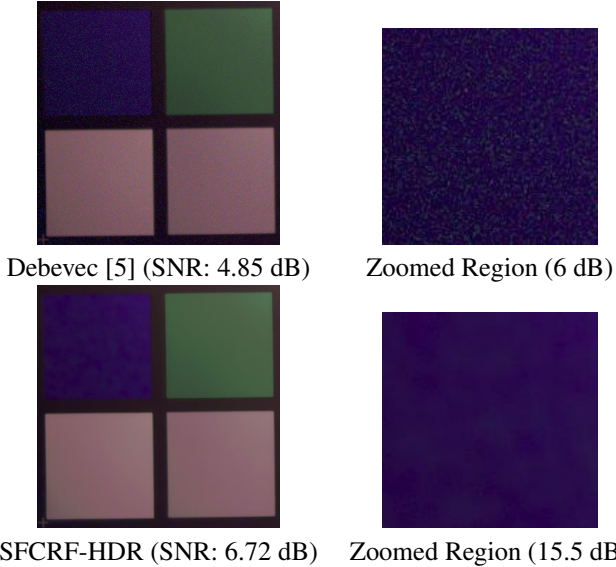


Fig. 4: The HDR estimated result of the proposed SFCRF-HDR framework (bottom) compared to Debevec & Malik [5] approach (top) on Colorchecker board. The reported SNR on the left is the averaged SNR of all colour patches in the ColorChecker (not just the four patches shown). Shown on the right is the blue patch with its corresponding SNR.

Table 1: Camera Settings for Bracketed Image Capture

Camera: Canon T3i ISO: 6400 Aperture Size: 5.0	
Scene	Exposure Times (seconds)
ColorChecker	1/4096, 1/2048, 1/512, 1/128, 1/64
Tree	1/1600, 1/400, 1/100
Books	1/4000, 1/1000, 1/250

4. DISCUSSION

In this paper we present a new HDR reconstruction framework that extends upon existing HDR reconstruction methods to reduce noise in HDR images given bracketed images succumbed to high ISO noise. We showed that the HDR radiance map can be inferred by using a SFCRF approach and modelling the LDR images as noisy observations. Results demonstrated that the proposed method is able to significantly reduce noise in the HDR images, especially in underexposed areas, while preserving edge boundaries. Our method allows photographs to be taken at higher ISO settings and faster shutter speeds with reduced noise. Future work include modelling external light sources in the conditional random field model and autonomously learning the noise characteristics depending on the exposure time.

5. REFERENCES

- [1] T. M Lillesand and R. Kiefer, "Remote sensing and image interpretation, 1994," .
- [2] J. Munkberg, P. Clarberg, J. Hasselgren, and T. Akenine-Möller, "Practical hdr texture compression," in *Computer Graphics Forum*. Wiley Online Library, 2008.
- [3] G. Ward, "The radiance lighting simulation and rendering system," in *Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. ACM, 1994.
- [4] G. Ward and M. Simmons, "Subband encoding of high dynamic range imagery," in *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*. ACM, 2004.
- [5] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes*. ACM, 2008.
- [6] D. Stoppa, A. Simoni, L. Gonzo, M. Gottardi, and G. Dalla Betta, "Novel cmos image sensor with a 132-db dynamic range," *IEEE Journal of Solid-State Circuits*, 2002.
- [7] M. A Robertson, S. Borman, and R. Stevenson, "Estimation-theoretic approach to dynamic range enhancement using multiple exposures," *Journal of Electronic Imaging*, 2003.
- [8] T. Mitsunaga and S. Nayar, "Radiometric self calibration," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 1999.
- [9] A. Akyüz and E. Reinhard, "Noise reduction in high dynamic range imaging," *Journal of Visual Communication and Image Representation*, 2007.
- [10] W. Yao, Z. Li, and S. Rahardja, "Noise reduction for differently exposed images," in *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2012.
- [11] R. Rameshan, S. Chaudhuri, and R. Velmurugan, "High dynamic range imaging under noisy observations," in *International Conference on Image Processing (ICIP)*. IEEE, 2011.
- [12] B. Goossens, H. Luong, J. Aelterman, A. Pizurica, and W. Philips, "Reconstruction of high dynamic range images with poisson noise modeling and integrated denoising," in *International Conference on Image Processing (ICIP)*. IEEE, 2011.
- [13] S. Hasinoff, F. Durand, and W. Freeman, "Noise-optimal capture for high dynamic range photography," in *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2010.
- [14] M. J. Shafiee, A. Wong, P. Siva, and P. Fieguth, "Efficient bayesian inference using fully connected conditional random fields with stochastic cliques," in *International Conference on Image Processing (ICIP)*. IEEE, 2014.
- [15] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*, AK Peters (CRC Press), 2011.
- [16] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," in *ACM Transactions on Graphics (TOG)*. ACM, 2008.
- [17] "Luminance HDR," <http://qtpfsgui.sourceforge.net/>, 2014.