RECONSTRUCTION OF HIGH DYNAMIC RANGE IMAGES WITH POISSON NOISE MODELING AND INTEGRATED DENOISING

Bart Goossens, Hiêp Luong, Jan Aelterman, Aleksandra Pižurica and Wilfried Philips

Ghent University - TELIN - IPI - IBBT
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
bart.goossens@telin.ugent.be

ABSTRACT

In this paper, we present a new method for High Dynamic Range (HDR) reconstruction based on a set of multiple photographs with different exposure times. While most existing techniques take a deterministic approach by assuming that the acquired low dynamic range (LDR) images are noise-free, we explicitly model the photon arrival process by assuming sensor data corrupted by Poisson noise. Taking the noise characteristics of the sensor data into account leads to a more robust way to estimate the non-parametric camera response function (CRF) compared to existing techniques. To further improve the HDR reconstruction, we adopt the split-Bregman framework and use Total Variation for regularization. Experimental results on real camera images and ground-truth data show the effectiveness of the proposed approach.

Index Terms—High dynamic range imaging, denoising

I. INTRODUCTION

Digital photographs are acquired by exposing the camera sensor to incident light (also called radiance) for a specified period of time, called the exposure time. During the exposure, the electrical charge of the sensor builds up and finally, the charge is converted to a digital number using an analog-to-digital converter (ADC). However, because the sensor has a limited working range, it is not always possible to capture the full dynamic range of a natural scene: underexposed regions in the image will be mapped to the lowest intensity value while overexposed regions are mapped to the highest intensity value. Low dynamic range (LDR) images are then obtained.

The goal of HDR reconstruction is to compute the radiance map of a real scene, based on several LDR images taken at different exposure times. In the past decade, several methods have been proposed for this task: the method of Debevec and Malik [1] jointly recovers a *non*-parametric camera response curve from image pixels and the corresponding radiance map. Because every pixel corresponds to an equation in a linear system, in practice, only a small number of pixels are used (e.g. 256) to compute the actual response curve and this curve is then used to directly compute the radiance map. The method of Mann and Picard [2] uses a parameteric function for photometric calibration of the different images. One difficulty is choosing a proper parametric function for a given camera, because a simple model is often too restrictive. Mitsunaga and Nayar [3] further improve this

work by proposing a more realistic polynomial model, for which linear regression can be used. The method of Robertson et al. [4] iteratively updates the estimation of the camera response function (CRF) and the irradiance values, until convergence. Their technique assumes an additive sensor noise model. Finally, (joint) histogram-based methods have been proposed for estimating the CRF [5], [6], these techniques can then for example be used in combination with the reconstruction from Debevec and Malik [1]. In [7], we proposed a different HDR weighting function to yield higher reconstruction SNR in presence of camera noise. Moreover, we presented a solution to the algorithmic complexity issue of [1] such that *all* image pixels can be taken into account, resulting in better estimation of the CRF and a better overall reconstruction quality.

A problem with many of these techniques is that they do not explicitly take sensor noise into account, and as a consequence these methods reconstruct images *with* noise. In this paper, will integrate denoising in the HDR reconstruction. First, we will show that estimation for a Poisson sensor model leads to a different but more effective weighting scheme than in [1], [7]. Then, we also reformulate the HDR reconstruction itself based on this noise model. Our approach is then integrated in the Split-Bregman [8] or SALSA [9] optimization framework (both frameworks are equivalent in our case) to include spatial regularization (and hence denoising). In this paper, we will use Total Variation [10] for regularization as a proof of concept.

The remainder of this paper is as follows: in Section II we present the image and noise model that we will use throughout this paper. The estimation of the CRF is explained in Section III. In Section IV we describe the HDR reconstruction method (i.e. the estimation of the radiance map). Results and a discussion are given in Section V. Finally, Section VI concludes this paper.

II. IMAGE AND NOISE MODELING

We start from a set of P digital photographs of a static scene taken from a fixed camera position, and we assume that illumination changes can be ignored. Let $z_{ij} \in [0, 1, ..., 255]$ denote the captured pixel intensities of each LDR image, where $i = 1, ..., N$ is a one-dimensional position index in the image (e.g., using raster scanning) and where $j = 1, ..., P$ is the photograph index. The exposure time for shot j is given by Δt_j and is assumed to be known. We are interested in reconstructing the radiance map E_i of the scene based on the measurements z_{ij}. As in [1], we start from the following "*ideal*" relationship between E_i and z_{ij}:

$$z_{ij} = f\left(\Delta t_j E_i\right), \tag{1}$$

where E_i is first integrated over a time period Δt_j and then subjected to the CRF f. The CRF models various nonlinear operations in the digital camera such as gamma correction, ISO setting, gain control, white balancing, contrast enhancement and is camera/manufacturer dependent. We further assume that f is monotonic on its domain and smooth (such that small relative changes in E_i result in small relative changes with the same sign in z_{ij}).

Due to the stochastic nature of the photon arrival process, (1) is never exact in practice. Therefore, let x_{ij} denote the measured signal intensity for exposure j at position i, then x_{ij} is Poisson distributed:

$$x_{ij} \sim \mathcal{P}\left(\Delta t_j E_i\right) \quad \text{and} \quad z_{ij} = f\left(x_{ij}\right). \tag{2}$$

Because working with densities of nonlinear functions of Poisson distributed random variables is highly complicated and often not analytically tractable, we will seek for good probability density function (PDF) approximations. In the following, we will consider $\log x_{ij}$, which has the following statistical moments:

$$\mathrm{E}\left[\log x | E\right] = e^{-E} \sum_{k=1}^{+\infty} \frac{\log\left(k/E\right) E^k}{k!} + \log E \approx \log E,$$

$$\mathrm{Var}\left[\log x | E\right] = e^{-E} \sum_{k=1}^{+\infty} \frac{\log^2\left(k/E\right) E^k}{k!} - e^{-2E}\left(\sum_{k=1}^{+\infty} \frac{\log\left(k/E\right) E^k}{k!}\right)^2$$
$$\approx E^{-1}, \tag{3}$$

where we used $E = \Delta t_j E_i$ to shorten the notation. The accuracy of the approximations in (3) improves when E becomes larger (i.e., for higher SNRs). Consequently, by defining $g(z) = \log f^{-1}(z)$ as the logarithm of the inverse CRF (see [1]), we have:

$$\mathrm{E}\left[g\left(z_{ij}\right)|E_i\right] \approx \log \Delta t_j + \log E_i$$
$$\mathrm{Var}\left[g\left(z_{ij}\right)|E_i\right] \approx \left(\Delta t_j E_i\right)^{-1}. \tag{4}$$

This gives the following approximative Gaussian (but signal-dependent, i.e., the noise variance depends on the signal intensity) model for $g\left(z_{ij}\right)$:

$$g\left(z_{ij}\right) \sim \mathrm{N}\left(\log \Delta t_j + \log E_i, \left(\Delta t_j E_i\right)^{-1}\right). \tag{5}$$

Hence, the higher the exposure time, the lower the noise variance becomes. This Gaussian model has the great advantage that the estimation of the CRF and radiance map becomes considerably easier compared to exact PDF models, as we will show next.

III. ESTIMATING THE CAMERA RESPONSE FUNCTION

Because the CRF f (or alternatively, g) is often not known in advance (this would require deep knowledge of the internal processing in the camera and sensors being used), we wish to estimate the CRF from the set of LDR images as well. Fortunately, the CRF can easily be estimated through joint histograms of z_{ij} and $z_{ij'}$ for $j \neq j'$. Therefore, relying on the statistical independence of z_{ij} and $z_{ij'}$, the difference $g\left(z_{ij}\right) - g\left(z_{ij'}\right)$ has the following PDF:

$$g\left(z_{ij}\right) - g\left(z_{ij'}\right) \sim \mathrm{N}\left(\log \frac{\Delta t_j}{\Delta t_{j'}}, \frac{1}{E_i} \frac{\Delta t_j + \Delta t_{j'}}{\Delta t_j \Delta t_{j'}}\right).$$

Next, maximizing the likelihood function would amount to minimizing:

$$\min_g \sum_{\substack{i,j,j' \\ j \neq j'}} E_i \frac{\Delta t_j \Delta t_{j'}}{\Delta t_j + \Delta t_{j'}}\left(g\left(z_{ij}\right) - g\left(z_{ij'}\right) - \log \frac{\Delta t_j}{\Delta t_{j'}}\right)^2, \tag{6}$$

with respect to the unknown parameters. However, there are some issues with this approach which need to be solved first:

1) The estimation of $g(z), z = 0, ..., 255$, which is a discrete function. To estimate the values of this function, we treat $g(0), ..., g(255)$ as unknown variables (as in [1]).
2) The parameter E_i in (6) is unknown. We may choose to replace E_i by $\exp g\left(z_{ij}\right)/\Delta t_j$ and then solve iteratively for $g(z)$, as in iteratively reweighted least squares approaches. However, in our experiments we noted that a simpler and faster technique (which consists of dropping the pre-factor E_i in (6)) yields results that are already very good.
3) $g(0)$ and $g(255)$ are undefined, due to clipping of the intensity values. Moreover, the clipped LDR images z_{ij} are often stored in a compressed format (e.g. JPEG) on the camera memory. Therefore, as in [1], [7], we will use a weighting function $w(z)$ that determines the importance of the samples $g(z)$ of a given intensity.
4) To deal with scenarios with a low number of samples for certain value z (in this case the histogram of z_{ij} contains zero-bins), we also include the smoothness regularization term from [1].
5) Due to the assumption of noise presence, including a non-smoothness penalty does not guarantee that the estimated $g(z)$ is monotonic. Therefore, we include the conditions $g(z) > g(z-1), z = 1, ..., 255$ as constraints to the problem, yielding a quadratic programming (QP) problem.

Taking all these factors into account, the problem can be stated as:

$$\min_g \sum_{\substack{i,j,j' \\ j \neq j'}} \frac{\Delta t_j \Delta t_{j'}}{\Delta t_j + \Delta t_{j'}} w\left(z_{ij}\right) w\left(z_{ij'}\right)\left(g\left(z_{ij}\right) - g\left(z_{ij'}\right) - \log \frac{\Delta t_j}{\Delta t_{j'}}\right)^2$$

$$+ \nu\left|w(z)g''(z)\right|^2 \text{ subject to } g(z) \geq g(z-1) + \epsilon, z = 1, ..., 255,$$

where ϵ is a very small positive constant (e.g., $\epsilon = 10^{-8}$), ν is chosen in $[0, 1]$ and where a numerical second derivative is used to compute $g''(z)$. The problem can then be efficiently solved using standard QP solvers [11] (`quadprog` in MATLAB).

One point that we still need to address is the choice of the weighting function. On the one hand, the weighting function $w(z)$ should be maximally flat, to minimize the statistical errors due to noise in the estimation [7]. On the other hand, $w(z)$ should allow the algorithm to be resilient to clipping and JPEG artifacts near the boundaries of the dynamic range of z. Experimentally, in presence of noise we found that a good trade-off is given by:

$$w(z) = \exp\left(-\left[1.1\left|\frac{z - 127.5}{127.5}\right|\right]^{20}\right). \tag{7}$$

We remark that the particular choice of the numerical constants in (7) is not of big importance, we want to ensure that $w(z) \approx 0$ for $z = 0$ and $z = 255$ and $w(z) = 1$ for a large part of the dynamic range.

IV. ESTIMATING THE RADIANCE MAP

To estimate the radiance map E_i, we can select again a ML type of approach (based on (5)):

$$\min_{E_i > 0} \sum_j E_i \Delta t_j w(z_{ij})\left(g\left(z_{ij}\right) - \log E_i - \log \Delta t_j\right)^2, \tag{8}$$

Algorithm 1 The proposed split-Bregman (or SALSA) HDR reconstruction algorithm, solving (11).

---

initialize $\mathbf{d}^{(0)} = \mathbf{0}$, $\mathbf{b}^{(0)} = \mathbf{0}$, $\mathbf{E}^{'(0)} = \hat{\mathbf{E}}'_{\mathrm{WLS}}$, $n = 0$
repeat
$\mathbf{E}^{'(n+1)} = \arg\min_{\mathbf{E}'} \frac{\lambda}{2} H(\mathbf{z}, \mathbf{E}') + \frac{\mu}{2} \left\| \mathbf{E}' \text{-} \mathbf{D}^H \left(\mathbf{d}^{(n)} \text{-} \mathbf{b}^{(n)} \right) \right\|^2$
$\mathbf{d}^{(n+1)} = \arg\min_{\mathbf{d}} |\mathbf{d}|_1 + \frac{\mu}{2} \left\| \mathbf{d} - \mathbf{D}\mathbf{E}^{'(n+1)} - \mathbf{b}^{(n)} \right\|^2$
$\mathbf{b}^{(n+1)} = \mathbf{b}^{(n)} + \left(\mathbf{D}\mathbf{E}^{'(n)} - \mathbf{d}^{(n)} \right)$
$n \leftarrow n + 1$
until $\left\| \mathbf{E}^{'(n)} - \mathbf{E}^{'(n-1)} \right\|_2 < \text{tolerance}$

---

where, for similar reasons as in Section III, we also included the weighting function $w(z)$. Noting that E_i in (8) is a constant factor that can be dropped, we obtain the weighted mean estimate in $E_i' = \log E_i$:

$$\hat{E}'_{i,\mathrm{WLS}} = \frac{\sum_j \Delta t_j w(z_{ij}) \left(g(z_{ij}) - \log \Delta t_j \right)}{\sum_j \Delta t_j w(z_{ij})}. \qquad (9)$$

This is a pointwise estimate (i.e., it is applied to every position i of the image without taking neighbors into account). Compared to the reconstruction in [1], the exposure times Δt_j play a more important role and now also determine the weights. However, there are two scenarios in which (9) may not yield the desired result: 1) in case the number of LDR images, P, is small and the exposure times Δt_j are also relatively small, 2) when the weights $w(z_{ij})$ for a given i are all small. In both cases, the denominator becomes very small resulting in a noisy or even unstable reconstruction. This problem can be solved by including some prior knowledge about the ideal radiance map E_i. We will explain this in the next Subsections.

IV-A. Reconstruction using split-Bregman

We will use some vector notation: let $\mathbf{z}$ denote all intensities z_{ij} stacked into a row vector, and let $\mathbf{E}'$ represent a row vector containing the logarithms of the radiance map values $E_i' = \log E_i$. We consider the following data fitting function:

$$H(\mathbf{z}, \mathbf{E}') = \sum_{i,j} E_i \Delta t_j w(z_{ij}) \left(g(z_{ij}) \text{-} E_i' \text{-} \log \Delta t_j \right)^2. \qquad (10)$$

Then the regularized reconstruction can be expressed as an unconstrained problem:

$$\hat{\mathbf{E}}' = \arg\min_{\mathbf{E}'} \left| \mathbf{D}\mathbf{E}' \right|_1 + \frac{\lambda}{2} H(\mathbf{z}, \mathbf{E}'), \qquad (11)$$

where $\left| \mathbf{D}\mathbf{E}' \right|_1$ is the Total Variation (TV) norm [10]. In this notation, $\mathbf{D}$ is a block matrix that consists of the horizontal and vertical discrete derivative operators. For generality, to allow different matrices $\mathbf{D}$ to be used as well (e.g. curvelets, shearlets) and later to solve the problem jointly with demosaicing, we use a more general solver than the gradient descent used in TV minimization. Examples of such a solver are the split-Bregman [8] method and the SALSA method [9]. In our application, both formulations give the same algorithm, shown in Algorithm 1. All steps of this algorithm amount to simple and fast point-wise operations, such as weighted averaging and soft-thresholding. To have a quadratic data fitting function, we use the radiance map value E_i from previous iteration, $E_i^{(n)}$, as weight in (10).

Table I. Comparison of the MSE of the inverse CRFs compared to the ground-truth.

	(noise-free)	(simulated Poisson noise)		
Exposure times Δt_j	$1s$, $2.5s$, $5s$, $10s$	$1s$, $2.5s$, $5s$, $10s$	$2s$, $5s$, $10s$, $20s$	$3s$, $7.5s$, $15s$, $30s$
Debevec [1]	**8.75**	39.48	40.30	39.58
De Neve [7]	9.94	17.89	17.89	17.87
Proposed	10.15	**17.43**	**17.53**	**17.41**

IV-B. Dealing with color images

To reconstruct color images, we apply the CRF estimation and HDR reconstruction in RGB color space to each color channel *individually*. However, because the CRF can be estimated up to a constant scale factor, it is necessary to calibrate the scale factor for the different color channels. A common assumption [1] is that the RGB value $(255/2, 255/2, 255/2)$ has equal radiance values for R, G, B, such that the pixel is achromatic. This directly yields the different scale factors.

We found that, next to this, a much more faithful reconstruction with less color artifacts can be obtained by fixing the values of the *weighting function* across color channels. Instead of computing the weights for each channel individually, we compute the weights based on the luminance intensities $0.2126z_{ij,R} + 0.7152z_{ij,G} + 0.0722z_{ij,B}$.

V. RESULTS AND DISCUSSION

To evaluate our technique, we acquired real camera images, containing real CCD sensor noise. We captured two sets of LDR images: the first set in bright conditions (Figure 1) and a second set in a more dark environment (Figure 3). Consequently, the images in set 1 do not suffer from the influence of noise, while the images in set 2 have a low SNR, especially in the dark regions in the image. HDR reconstruction results are given in Figure 2 and Figure 4. We compare to Debevec and Malik [1], for their method we use 2048 sampling points, randomly selected in the image. For our method, the parameters μ and λ are chosen experimentally as $\mu = 300$ and $\lambda = 1$. In Figure 2 it can be seen that the color artifacts in the sun region has been remedied, this is partly by fixing the weights across color channels as explained in Subsection IV-B. In Figure 4, we observe that for our method, noise is well suppressed, while numerical instabilities (due to a denominator that becomes zero) are completely avoided.

To have an indication of the objective quality of our method, we also perform an experiment using the ground-truth image from [7]. First, the HDR image is multiplied by pre-specified exposure times and subsequently corrupted with Poisson noise. Next a fictious camera response function $f(x) \sim \sqrt{x}$ and clipping of the dynamic range are applied to the resulting image. In Table I, we compare the mean square error (MSE) of the estimated $g(z)$ compared to the ground-truth $g(z) = \log f^{-1}(z) = 2\log z$, for different settings of exposure times. We use $\nu = 0.5$ as smoothing parameter for the CRF estimation for all methods. It can be seen that in the Poisson noise case (this is the scenario we are interested in, the noise-free case is shown as reference), the proposed CRF estimation method performs the best, due to the modified weighting scheme that takes the exposure times into account.

Fig. 1. A first set of LDR images used of the experiments, with corresponding exposure times. Images were captured using a Nikon D90 camera at ISO 200.
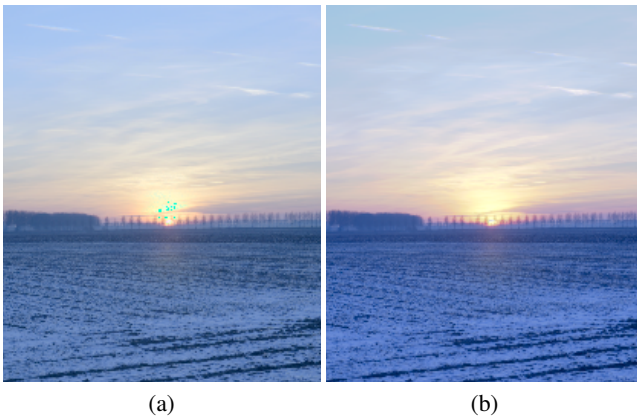

(a)   (b)

Fig. 2. HDR synthesis results (after the tone mapping from [12]): (a) Debevec and Malik [1], (b) Proposed method. Images are cropped due to space limitations.
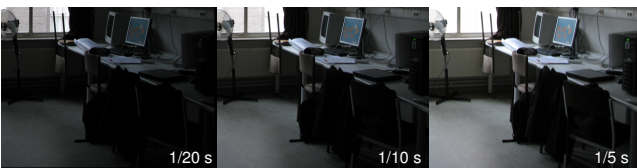

Fig. 3. A second set of LDR images used of the experiments, with corresponding exposure times. Images were captured using a Canon Powershot S5 camera at ISO 80.

VI. CONCLUSION

We presented a new CRF estimation and HDR reconstruction method that is based on a Poisson noise model for the sensor data. We showed how the CRF estimation needs to be modified in order to take the noise characteristics into account and have a better estimation. Furthermore, we extended the HDR reconstruction by integrating it into a more general split-Bregman framework for image restoration, offering denoising and solving some instability problems of non-regularized reconstruction methods. In our future work, we wish to combine this approach with shearlet-based regularization and demosaicing, and this can all be done within the presented framework.
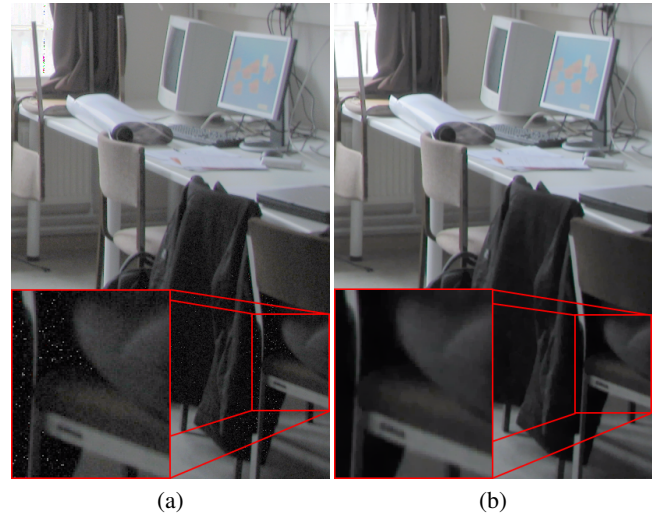

(a)   (b)

Fig. 4. HDR synthesis results (after the tone mapping from [12]): (a) Debevec and Malik [1], (b) Proposed method. Images are cropped due to space limitations.

VII. REFERENCES

[1] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of SIGGRAPH97, Computer Graphics Proceedings*, pp. 369–378, 1997.

[2] S. Mann and R. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *In Proceedings of IS&T 48th annual conference (Cambridge, Massachusetts, May 1995)*, pp. 422–428, 1994.

[3] T. Mitsunaga and S. Nayar, "Radiometric self calibration," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 374–380, 1999.

[4] M. Robertson, S. Borman, and R. Stevenson, "Dynamic range improvement through multiple exposures," in *IEEE Int. Conf. Image Process.*, vol. 3, pp. 159–163, 1999.

[5] M. D. Grossberg and S. K. Nayar, "Determining the camera response from images: what is knowable?," *IEEE Trans. Pat. Analysis and Mach. Intell.*, vol. 25, no. 11, pp. 1455–1467, 2003.

[6] M. D. Grossberg and S. K. Nayar, "Modeling the space of camera response functions," *IEEE Trans. Pat. Analysis and Mach. Intell.*, vol. 26, no. 10, pp. 1272–1282, 2004.

[7] S. De Neve, B. Goossens, H. Luong, and W. Philips, "An Improved HDR Image Synthesis Algorithm," in *IEEE Int. Conf. Image Processing (ICIP2009)*, (Cairo, Egypt), pp. 1545–1548, nov 2009.

[8] T. Goldstein and S. Osher, "The split Bregman method for L1 regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[9] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, 2010.

[10] L. Rudin and S. Osher, "Total variation based image restoration with free local constraints," in *Proc. of IEEE International Conference on Image Processing (ICIP)*, vol. 1, pp. 31–35, nov 1994.

[11] P. Gill, W. Murray, and M. H. Wright, *Practical Optimization*. London, UK: Acedemic Press, 1981.

[12] G. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. on Vis. and Comp. Graphics*, vol. 3, pp. 291–306, 1997.