HDR reconstruction for alternating gain (ISO) sensor readout

Saghi Hajisharif Joel Kronander Jonas Unger

Department Of Science and Technology
Linköping University, Sweden

Abstract

Modern image sensors are becoming more and more flexible in the way an image is captured. In this paper, we focus on sensors that allow the per pixel gain to be varied over the sensor and develop a new technique for efficient and accurate reconstruction of high dynamic range *(HDR) images based on such input data. Our method estimates the radiant power at each output pixel using a sampling operation which performs color interpolation, re-sampling, noise reduction and HDR-reconstruction in a single step. The reconstruction filter uses a sensor noise model to weight the input pixel samples according to their variances. Our algorithm works in only a small spatial neighbourhood around each pixel and lends itself to efficient implementation in hardware. To demonstrate the utility of our approach we show example HDR-images reconstructed from raw sensor data captured using off-the shelf consumer hardware which allows for two different gain settings for different rows in the same image. To analyse the accuracy of the algorithm, we also use synthetic images from a camera simulation software.*

1. Introduction

For more than a decade High dynamic range imaging, [DM97], has been a highly important tool in many applications areas. However, the users of HDR-imaging have until now mainly been researchers or other expert practitioners. The main reason for this is that traditionally HDR capture has been difficult as it requires either the capture of multiple exposures of the same scene or specialized and expensive camera systems. A key challenge is therefore to develop hardware, software and methodology that enables HDR-imaging to reach also everyday consumers. An enabling factor in solving this problem is the rapid development of cameras and sensors. Over the last years, we have seen a significant increase in the computational power onboard the cameras as well as a higher degree of control over the actual sensors. This produces a much higher freedom in the way we capture and reconstruct images; even with consumer cameras.

In this paper, we exploit this flexibility and present an algorithm for HDR-image reconstruction based on a single input image where the pixel gain is varied over the sensor, [UG07, a1e13]. The analog pixel gain is proportional to the *ISO* setting found on most cameras. The input to our algorithm is a multi-gain RAW sensor image where color is captured using a *color filter array* (CFA), e.g. a Bayer pattern. Figure 1 illustrates how two different distributions of per pixel gain settings are overlaid onto a raw CFA image. A low gain setting leads to a high saturation threshold but a lower signal to noise ratio compared to a high gain setting.

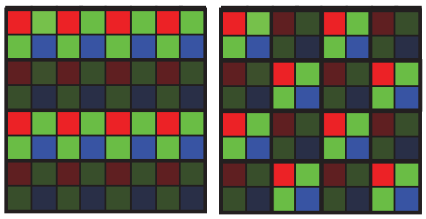

Figure 1: *The figure Illustrates two example gain distribution patterns: (**left**) every second pair of rows have a higher gain, and (**middle**) groups of* 2×2 *pixels have a higher gain..*

The goal of our algorithm is to, for each output pixel, fuse the information from the different gain settings available within a small neighbourhood of input pixels as illustrated in Figure 2. In this way, it is possible to maximize the saturation threshold while maintaining low sensor noise in dark parts of the image. To accurately take into account the varying noise properties in the image, we estimate the variance at each pixel using a camera noise model (see Section 2). The final HDR-estimate is carried out as a single sampling operation which simultaneously performs: *re-sampling*, *color interpolation*, *noise reduction*, and *HDR-fusion*.

In comparison to multiple exposure techniques [DM97, SKY*12, UGOJ04], our method has the advantage of operating on a single image instead of a time-multiplexed sequence. This means that the capture is easier and that ghosting artifacts from camera or scene motion are not a problem. Similarly to other techniques based on the idea of spatial multiplexing, e.g. using filters with different transmittance [NM00, NN02], there is a tradeoff between how the different gain settings are distributed over the sensor and the output resolution. We show, however, that our algorithm is capable of reconstructing full resolution HDR output with high accuracy. This is due to the noise aware adaptive filter kernel used in the reconstruction.

Our algorithm is inspired by [KGBU13] who used a similar sampling scheme to reconstruct HDR images captured using camera systems with multiple sensors (equipped with different natural density filters). Here, we extend this idea to instead operate on single CFA images with spatially varying gain. We also show how the technique can be used with consumer cameras instead of custom built multi-sensor systems. As our experimental platform we use *Canon 5D* Mark III cameras running the Magic Lantern firmware with the recent *dual ISO module* [a1e13]. This enables two different gain settings to be used simultaneously for different pixel rows. Using a camera simulation framework, we also evaluate other gain distributions possible with Bayer pattern CFAs. The result is a robust and efficient reconstruction algorithm that can be run in parallel in hardware.

2. Radiometric Calibration

The first step in our HDR-fusion is to transform each pixel value to a common radiometric space. In this space the pixel response, f_i, at location i corresponds to the number of photo-induced electrons collected per unit time. The pixel values are transformed using a radiometric model inspired by [ADGM13, KGBU13, KGB*13], where non-saturated pixel values are modeled as realizations of a random variable Y_i with distribution

$$Y \sim \mathcal{N}(g_i a_i t f_i + \mu_R, g_i^2 a_i t f_i + \sigma_R^2(g_i)) \quad (1)$$

where g_i is the pixel gain, a_i is the pixel non-uniformity, t is exposure time and μ_R and σ_R^2 are the readout noise mean and standard deviation. The readout noise variance generally depends on the gain setting used. Previous work [HDF10, KGBU13] have used simplistic parametric models to describe how the readout noise varies with different gain settings. However, we have found that such models cannot accurately describe this dependence for modern camera sensors. To handle sensors with varying gain we instead calibrate the readout noise standard deviation, $\sigma_R^2(g)$ for each gain/ISO setting individually.

A raw digital value, y_i, is transformed to an estimate of the irradiance at the pixel as

$$\hat{f}_i = \frac{y_i - \mu_R}{g_i a_i t} \quad (2)$$

The variance of $\hat{f}_i$ is in turn estimated as

$$\sigma_{fi}^2 = \frac{g_i^2 a_i t \hat{f}_i + \sigma_R^2(g_i)}{g_i^2 a_i^2 t^2} \quad (3)$$

For an in-depth overview of this radiometric camera model we refer the reader to [ADGM13, KGBU13, KGB*13].

The per pixel non-uniformity, a_i, can be estimated using a flat field image. For this paper, however, we have simply assumed the per pixel uniformity to be constant for all pixels. The mean and variance of the readout noise, μ_R, σ_R^2 can be estimated from a set of black images, captured so that no light reaches the sensor. The sensor gain, g_s, can be calibrated using the relation,

$$g_i = \frac{\text{Var}[y_i] - \sigma_R^2}{E[y_i] - \mu_R} \quad (4)$$

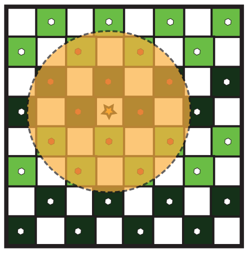

Figure 2: *Unified reconstruction exemplified for the green color channel. To reconstruct the HDR pixel (orange star), radiometrically calibrated samples in the neighborhood (red) are fitted to a local polynomial model taking into account the noise of the individual samples.*

3. Unified Reconstruction

To estimate an HDR pixel z_j at a location X_j, the observed samples $\hat{f}_i(X_i)$ in a local neighborhood around the pixel are fitted to a local polynomial model, see Figure 2. Assuming that the radiant power $f(x)$ is a smooth function in a local neighborhood around the output location X_j, for each color channel, an *M-th* order Taylor series expansion is used to predict the radiant power at a point X_i close to X_j, as:

$$\begin{aligned}\tilde{f}(X_i) &= C_0 + C_1(X_i - X_j) \\ &+ C_2 tril\{(X_i - X_j)(X_i - X_j)^T\} + \quad \ldots \end{aligned} \quad (5)$$

where *tril* lexicographically vectorizes the lower triangular part of a symmetric matrix and where

$$C_0 = f(X_j) \quad (6)$$

$$C_1 = \nabla f(X_j) = \left[\frac{\partial f(X_j)}{\partial x_1}, \frac{\partial f(X_j)}{\partial x_2}\right] \quad (7)$$

$$C_2 = \frac{1}{2}\left[\frac{\partial^2 f(X_j)}{\partial x_1^2}, 2\frac{\partial^2 f(X_j)}{\partial x_1 \partial x_2}, \frac{\partial^2 f(X_j)}{\partial x_2^2}\right] \quad (8)$$

Given the fitted polynomial coefficients, $C_{1:M}$, we can predict the radiant power at the output location X_j by $C_0 = f(X_j)$, and the first order gradients by C_1.

To estimate the coefficients we maximize a localized likelihood function defined using a Gaussian smoothing window centered around X_j:

$$\mathcal{W}_H(X_k) = \frac{1}{2\pi det(H)} exp\left\{ -(X_k - X_j)^T H^{-1}(X_k - X_j)\right\} \quad (9)$$

where H is a 2×2 smoothing matrix that determines the shape and size of the window. The polynomial coefficients, $\tilde{C}$, maximizing the localized likelihood function is found by the weighted least squares estimate

$$\begin{aligned}\tilde{C} &= \underset{C\in\mathcal{R}^M}{\operatorname{argmax}}(L(X_j,C)) \\ &= (\Phi^T W\Phi)^{-1}\Phi^T W\bar{f} \end{aligned} \tag{10}$$

where

$$\bar{f} = [\hat{f}_1(X_1),\hat{f}_2(X_2),...\hat{f}_N(X_N)]^T \quad X_i \in \operatorname{supp}\left(\mathcal{W}_H(X)\right)\}$$

$$W = diag[\frac{\mathcal{W}_H(X_1)}{\hat{\sigma}_{f_1}},\frac{\mathcal{W}_H(X_2)}{\hat{\sigma}_{f_2}},...,\frac{\mathcal{W}_H(X_K)}{\hat{\sigma}_{f_K}}]$$

$$\Phi = \begin{bmatrix} 1 & (X_1-X_j) & tril^T\{(X_1-X_j)(X_1-X_j)^T\} & ... \\ 1 & (X_2-X_j) & tril^T\{(X_2-X_j)(X_2-X_j)^T\} & ... \\ \vdots & \vdots & \vdots & \vdots \\ 1 & (X_K-X_j) & tril^T\{(X_K-X_j)(X_K-X_j)^T\} & ... \end{bmatrix}$$

The expected mean square error of the reconstructed image depends on a trade-off between bias and variance of the estimate. This trade-off is determined by: the order of the polynomial basis M, the window function $\mathcal{W}$, and the smoothing matrix H.

Using a piecewise constant polynomial, $M = 0$, the estimator corresponds to an ordinary locally weighted average of neighbouring pixels and is thus very fast to compute. Using $M = 0$ may, however, introduce asymmetries in the number of available samples around areas of sensor saturation and at image border. This bias may introduce artifacts in these locations. By instead fitting a linear polynomial (a plane, $M = 1$), the bias can be reduced significantly. Introducing higher order polynomials is possible, but may lead to increased variance in the estimates.

The shape of the Gaussian is determined by the smoothing matrix, H. Here, we use $H = hI$, where h is a global scale parameter and I is the identity matrix. This corresponds to an isotropic filter support. The choice of h depends on the sensor noise characteristics and the scene, and is therefore treated as a user parameter. A large scale parameter h will introduce a low-pass filtering effect and reduce noise and a small h will create sharper images. For Bayer patterns, we set h separately for the green and red/blue color channels, $h_G = \frac{h_{R,B}}{\sqrt{2}}$, to reflect the higher number of green samples per unit area.

4. Results and Discussion

As a result of our method, we show a set of reconstructed images for different input data as well as different gain/ISO patterns which has been discussed in section 1. The input data consists of the real data captured with Canon 5D Mark III running Magic Lantern's module "Dual-ISO" and also synthetic data that is generated using a camera simulation framework. The framework's input is a noise free HDR image and it generates raw sensor image for any camera, given the calibration data produced by using the radiometric model described in section 2. The noise free HDR-images are reconstructed using carefully calibrated exposure brackets captured 1 f-stop apart. Each exposure bracket is computed as the average of 100 images with the same exposure settings.

Real data - The data was captured using a ISO setting of 100 and 800 for each second row, as shown in the left patten of Figure 1. In Figure 4 we compare our reconstructed result to the original "Dual-ISO" reconstruction software [a1e13]. The figure to the left displays the original image reconstructed from alternating ISO100 and ISO1600 for every second pair of pixel rows. The cutouts shows the raw data, our method using input data alternating ISO100 and ISO800, our method using input data alternating ISO100 and ISO1600, and a comparison to the original Dual-ISO method. Using ISO100 and ISO800 extends the dynamic range of the camera with 3 *f*-stops and using ISO100 and ISO1600 extends the dynamic range of the camera with 4 *f*-stops. Taking into account that the noise level of the 14bit sensor lies around 3 bits, our method extends the working dynamic range of the camera up to 14-15 *f*-stops using these settings. The original Dual-ISO method includes several filtering steps in a very heuristic way making it difficult to control. Some of its problems can be seen in the over saturation on the red car.

Simulated data - We have also considered other gain patterns using simulated data generated from a camera simulation framework. Figure 4 displays a comparison between the two patterns shown in Figure 1. The left image in 4 corresponds to the left pattern in 1 and vice versa. As the result demonstrates both patterns produce quite similar images, however the diffrence can be found around edges and highlight regions. The zoom images of Figure 4 shows that the square pattern (right) tends to produce more accurate results.

5. Conclusion and Future work

This paper presented a novel technique for reconstruction of HDR-images from single raw CFA images with spatially varying per pixel gain. Using a radiometric camera model, the algorithm performs a sampling operation which includes color interpolation, noise reduction and HDR-fusion in a single unified step. Our experiments show that our method can be readily used with consumer cameras, and that it produces highly accurate results. However using isotropic filter kernel can introduce blur and color artifacts in the result. As a solution to this problem, we would like to incorporate anisotropic filter kernels for improved color interpolation and overall image quality which has been left for future works.

6. Acknowledgments

This project was funded by the Swedish Foundation for Strategic Research (SSF) through grant IIS11-0081,

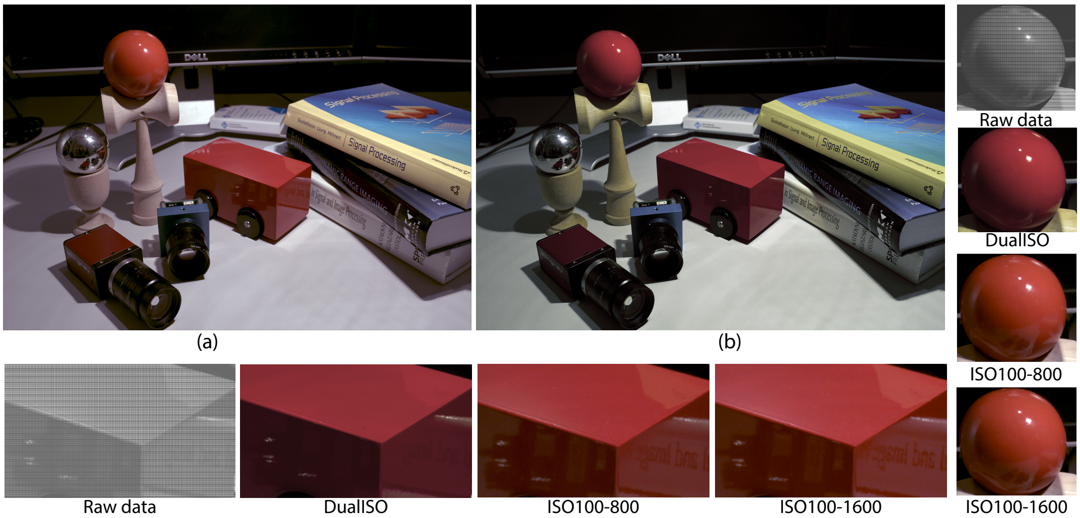

Figure 3: *Reconstructed HDR images based on real data from a Canon 5D Mark III running Magic Lantern alternating iso readout.* ***(a)*** *The full image reconstructed by our method using ISO100 alternating with ISO800 using the pattern in Figure 1 (left).* ***(b)*** *The full image reconstructed by DualISO method using ISO100 alternating with ISO800 using the pattern in Figure 1 (left). The* ***cutouts*** *shows (from left to right) the raw input, the DualISO method for ISO100-800, our method for ISO100-800, our method for ISO100-1600*

Figure 4: *The image displays the same raw data reconstructed using the two gains patterns displayed in Figure 1. The images are of full resolution and are best viewed zoomed in to display individual pixels.*

Linköping University Center for Industrial Information Technology (CENIIT), and the Swedish Research Council through the Linnaeus Environment CADICS.

References

[a1e13] A1EX: Dynamic range improvement for canon dslrs with 8-channel sensor readout by alternating iso during sensor readout. Technical documentation, url: http://acoutts.com/a1ex/dual_iso.pdf, July 2013.

[ADGM13] AGUERREBERE C., DELON J., GOUSSEAU Y., MUSÉ P.: Best algorithms for hdr image generation. a study of performance bounds. HAL preprint, 2013.

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (1997).

[HDF10] HASINOFF S., DURAND F., FREEMAN W.: Noise-optimal capture for high dynamic range photography. In *CVPR* (2010).

[KGB*13] KRONANDER J., GUSTAVSON S., BONNET G., YNNERMAN A., UNGER J.: A unified framework for multi-sensor hdr video reconstruction. *Signal Processing: Image Communication* (2013).

[KGBU13] KRONANDER J., GUSTAVSON S., BONNET G., UNGER J.: Unified hdr reconstruction from raw cfa data. *IEEE International Conference on Computational Photography (ICCP)* (2013).

[NM00] NAYAR S., MITSUNAGA T.: High Dynamic Range Imaging: Spatially Varying Pixel Exposures. In *CVPR* (Jun 2000), vol. 1, pp. 472–479.

[NN02] NAYAR S., NARASIMHAN S.: Assorted Pixels: Multi-Sampled Imaging With Structural Models. In *ECCV* (May 2002), vol. IV, pp. 636–652.

[SKY*12] SEN P., KALANTARI N. K., YAESOUBI M., DARABI S., GOLDMAN D. B., SHECHTMAN E.: Robust patch-based hdr reconstruction of dynamic scenes. *ACM Trans. Graph. 31*, 6 (2012), 203.

[UG07] UNGER J., GUSTAVSON S.: High-dynamic-range video for photometric measurement of illumination. *Proc. of SPIE 6501* (2007), 65010E.

[UGOJ04] UNGER J., GUSTAVSON S., OLLILA M., JOHANNESSON M.: A real time light probe. In *In Proceedings of the 25th Eurographics Annual Conference* (2004), vol. Short Papers and Interactive Demos, pp. 17–21.