

Multi Exposure Image Fusion

Ali Alsam

*Sør-Trøndelag University College, Faculty of Informatics and e-Learning (AITeL) Trondheim
ali.alsam@gmail.com*

Abstract

High dynamic range images are images that have a ratio between their maximum and minimum intensities greater than that of display or capture devices. One method to create high dynamic range images is to combine a set of differently exposed shoots into a single image that has to be compressed before visualization. Compressing high dynamic range images is thus based on a so called high dynamic range map (HDR). In this paper we show that a solution to the problem can be obtained without the need for the calculation of HDR maps. To achieve this we use an image fusion method in the gradient domain. We start by calculating a gradient field which encompasses the details found in the exposure image series. This gradient field is then integrated to result in the final within dynamic range image. The resulting algorithm achieves high image contrast due to a non-linear thresholding of gradients.

1 Introduction

A camera is a light sensitive device designed to measure the intensity of the light reflected off the different parts of the imaged scene. Similar to other measuring devices, a camera has a range that is defined by the minimum and maximum intensities that it can, faithfully, record. The ratio: Maximum to minimum intensities defines the dynamic range of the imaging device. This ratio is fixed as both the minimum light energy level needed to evoke a reliable measurement and that, level, that would cause a saturated measurement are fixed. It is, however, possible to shift this range by controlling the amount of light allowed to reach the light sensitive part of the camera. Shifting the range can be achieved by controlling the exposure time, shutter speed, or by adjusting the camera's aperture.

The dynamic range of natural sunlit scenes, i.e. the ratio between direct sunlight and the darkest shadows is much higher than that of today's cameras [6]. As an example, it is estimated that the dynamic range of sunlit scenes is 100 000 000:1 candelas per square metre (cd/m^2) while a 12 bit digital camera has a range of 4096 : 1 (cd/m^2). The discrepancy between the dynamic range of the imaged scene and that of the imaging device cannot be resolved: When imaging a scene with a high dynamic range parts of the resultant image will be under-exposed, over-exposed or both.

This paper was presented at the NIK-2010 conference; see <http://www.nik.no/>.

In applications where loss of detail, due to the scenes high dynamic range, is not acceptable, photographers resort to imaging the same scene with different exposure levels. Consequently, the higher the scene's dynamic range the more images are needed to capture it. Representing a single scene with a large number of images is, however, not desirable. To resolve this problem, Debevec and Malik [6] presented an algorithm to construct a high dynamic range (HDR) map based on a few images with different exposures or a panoramic video scan [1]. By definition, a high dynamic range map is an image that has a higher, or much higher, dynamic range than the input device. Unfortunately, HDR maps, present us with a new challenge: Given that the dynamic range of output devices is limited, how can we visualize high dynamic range maps?

The past decade has witnessed increasing interest in efficient algorithms to display HDR maps on low dynamic range devices. These algorithms can generally be grouped in two categories: In the first, we have global tone mapping algorithms [5, 10, 11, 13, 16, 17] and in the second we have local or spatial operators [4, 7, 8, 9, 12]. Global operators assume that a non-linear transformation of the HDR map is sufficient to compress the data into the range of the output device. Spatial operators, on the other hand, assume that the compression has to be based on the ratios between the image pixels. In other words, the pixel intensity is defined based on its location in the image plane.

We make the assertion that for most users it is the final tone mapped image that is of interest. In other words, most imaging users are not interested in a HDR map but rather an image that has a dynamic range no greater than their output device. Indeed it is the interest in an image with a dynamic range that matches the output device that is the drive behind all tone mapping algorithms.

In this paper, we take a novel approach to the problem of constructing an image from multiple exposures. We show that a solution to the problem can be obtained without the need to construct a HDR map. Using recent advances in image fusion [2, 15, 14], specifically the algorithm developed by Alsam and Drew [2, 3], we show that images with different exposures can be fused to result in a single image with the correct dynamic range. Further, by exploring the state of the art algorithms in image fusion we show that constructing a tone-mapped image from an exposure series can be treated as a problem of converting multi-spectral image data to grayscale [3].

The first step in the proposed algorithm is the calculation of a gradient field for each image in the exposure series. These gradient fields are then combined into a unified representative field that is integrated to result in the final image. As proposed by Alsam and Drew [2] we combine the gradient fields based on the maximum gradient at the horizontal and vertical directions. The main contribution of this article is a new method to calculate a tone-mapped representation of an image series with varying exposures without the calculation of a HDR map.

Calculation of a gradient field from a series of exposures

It is theoretically possible to treat the colour channels separately and thus for an RGB image with m exposures, independently, calculate a combined red, green and blue channels that are within the dynamic range of the output device. Such an approach leads, however, to a change in the colour values that needs to be corrected. Instead we work on the luminance channel of each image with the assumption that the colour values can be recovered from an the average of all the images.

Thus we calculate the gradient (z_{j_x}, z_{j_y}) for each grayscale exposure z_j separately. For each pixel (x, y) we choose the gradient from the exposure which has the maximum absolute value over

all exposures.

There are of course other ways to construct a gradient field [18]; we have chosen the maximum absolute value method because it is fast and has been shown to results in improved visualization [2].

2 Numerical Implementation:

Relaxation Method

Knowing that our assumption is that the best grayscale image is one having gradients as close as possible to a one dimensional representation of the combined multi-exposure gradient, we can state that we wish to minimize:

$$\min_g \int \int_{\Omega} \|\nabla \mathbf{Z} - \nabla g\|^2 dx dy \quad (1)$$

where $\nabla \mathbf{Z}$ is the combined maximum gradient over all exposures and ∇g is gradient of the resultant grayscale, luminance, image. Eq. 1 can be solved using Jacobi iteration, with homogeneous Neumann (so-called “natural”) boundary conditions to ensure zero derivative at the image boundaries. To handle boundary points, one inserts so-called “ghost” points outside the actual image, with values given by setting the appropriate derivative at each boundary to zero.

The iteration steps are expressed as:

$$g_{x,y}^{t+1} = \frac{1}{4} [g_{x+1,y}^t + g_{x-1,y}^t + g_{x,y+1}^t + g_{x,y-1}^t + z_{x+1,y}^{max} + z_{x-1,y}^{max} + z_{x,y+1}^{max} + z_{x,y-1}^{max}] \quad (2)$$

This is a generalization of the standard solution of the Poisson equation corresponding to Eq. 1.

Equation 2 can be solved with an arbitrary guess image g . In our implementation we use the average luminance image over all the exposures as starting point g^0 . As iteration proceeds, information is passed to neighboring pixels (relaxation) and the closer we get to an image that satisfies the equality of output gradient to input maximum-gradient.

Further, to control the dynamic range of the solution image, we introduced a reset step. Thus, at every iteration level we impose the constraint that the resultant pixel values, $g_{x,y}^{t+1}$ are between zero and m where m is the maximum intensity level that can be displayed by the output device. Thus if $g_{x,y}^{t+1}$ is less than zero we clip the value to zero; and if it is greater than m we clip it to m . Therefore, to Eq. 2 we add the conditions

$$\begin{aligned} \text{if } (g_{x,y}^{t+1} < 0) \quad g_{x,y}^{t+1} &= 0, \\ \text{if } (g_{x,y}^{t+1} > m) \quad g_{x,y}^{t+1} &= m \end{aligned} \quad (3)$$

Without this reset step, we found that image appearance had less contrast.



Figure 1. Seven different exposures of the Goldstein Synagogue: The Hebrew University of Jerusalem. In our approach these images were fused without the creation of a HDR map.

3 Results

In this section, we present an example image results. We compare our findings with those presented by Fattal *et al* [8] . We chose the Fattal algorithm for a number of reasons: Firstly, the algorithm has over the years, since its publication, been considered the method with best and most consistent results. Secondly, knowing that like our method the Fattal approach is in the gradient domain we believe that a comparison with this algorithm would offer the reader the chance to compare two algorithms that are conceptually similar. The main difference between the two algorithms is that in the case of our algorithm we don't require the calculation of a HDR map. Thus the main objective with this section is to show that the calculation of a tone-mapped version of a high dynamic scene that is represented by a series of image exposures does not necessitate the calculation of a HDR map. Finally, the calculations were based on the HSV colour space where the images from each exposure level were transformed to HSV and the calculation was based on the value channel only. The hue and saturation of the final image were assumed to correspond to those of the average over all exposures.

Goldstein Synagogue, The Hebrew University of Jerusalem

For this image, seven different exposures were used. These are shown in Figure 1. In our approach, these were combined without the calculation of a HDR map. Further, the results for the Fattal method are based on a high dynamic map that was created using the Debevec and Malik [6] algorithm. To achieve an impartial comparison we obtained the Fattal's results from the author's web-site.

Figure 2 shows a comparison between the colour output of the two algorithms. We note that both algorithms result in visibly pleasing outputs. We refrain from comparing the colour fidelity as we know of no objective method to measure it. Here we would like to point out that the Fattal's algorithm includes a sharpening step which we didn't employ.

4 Conclusions and future work

In this paper we have presented a mathematical derivation of a fusion method that we employed to combine a series of images with different exposures. The results show that the resulting



Figure 2. Shows a comparison between the luminance output of both algorithms. New output, top, and Fattal's bottom. We note that in our approach all the variations seen in the different exposures are represented in the final image.

algorithm achieves high image contrast while retaining the details in the different exposures. Comparison with the Fattal algorithm also show that the use of image fusion represent a method to optimize the calculation of an in-dynamic-range image by bypassing the calculation of a HDR map. We believe that this approach is better suited for an in-camera processing of an exposure series.

References

- [1] M. Aggarwal and N. Ahuja. Split aperture imaging for high dynamic range. *International Journal of Computer Vision*, 58(1):7–17, June 2004.
- [2] A. Alsam and M. S. Drew. Fast colour2grey. In *16th Color Imaging Conference: Color, Science, Systems and Applications.*, pages 342–346. Society for Imaging Science & Technology (IS&T)/Society for Information Display (SID) joint conference, 2008.
- [3] A. Alsam and M. S. Drew. Fast multispectral2gray. *The Journal of imaging science and technology*, 53(6):060401.1–060401.10, 2009.
- [4] K. Chiu, M. Herf, P. Shirley, S. Swamy, C. Wang, and K. Zimmerman. Spatially nonuniform scaling functions for high contrast images. In *Graphics Interface Toronto, Ontario, Canada: Canadian Information Processing Society*, pages 245–253, May 1993.
- [5] J. Cohen, C. Tchou, T. Hawkins, and P. Debevec. Real-time high dynamic range texture mapping. In *12th Eurographics Workshop on Rendering*, pages 313–320, June 2002.
- [6] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH*, pages 369–378, 1997.
- [7] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high dynamic range image. *SIGGRAPH Conference Graphics Proceedings, Annual Conference Series*, pages 257–265, 2002.
- [8] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. In *Proceedings of ACM SIGGRAPH*, 2002.
- [9] D. J. Jobson, Z. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, 6(7):965–976, July 1997.
- [10] G. W. Larson, H. Rushmeier, and C. Piatko. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics*, 3(4):219–306, October - December 1997.
- [11] N. Miller, P. Ngai, and D. Miller. The application of computer graphics in lighting design. *Journal of the IES*, 14:6–26, 1984.
- [12] A. Oppenheim, R. Schafer, and T. Stockham. Nonlinear filtering of multiplied and convolved signals. In *Proceedings of the IEEE*, 56(7):1264–1291, August 1968.
- [13] A. Scheel, M. Stamminger, and H. Seidel. Tone reproduction for interactive walkthroughs. *Computer Graphics Forum*, 19(3):301–312, August 2000.
- [14] D. Socolinsky and L. Wolff. A new visualization paradigm for multispectral imagery and data fusion. In *CVPR*, pages I:319–324, 1999.
- [15] D. Socolinsky and L. Wolff. Multispectral image visualization through first-order fusion. *IEEE Trans. Im. Proc.*, 11:923–931, 2002.
- [16] J. Tumblin, J. K. Hodgins, and B. K. Guenter. Two methods for display of high contrast images. *ACM Transactions on Graphics*, 18(1):56–94, January 1999.
- [17] J. Tumblin and H. E. Rushmeier. Tone reproduction for realistic images. *IEEE Computer Graphics & Applications*, 13(6):42–48, November 1993.
- [18] S. D. Zeno. A note on the gradient of a multi-image. *Comp. Vision, Graphics, and Image Proc.*, 33:116–125, 1986.