ADAPTIVE EXPOSURE FUSION FOR HIGH DYNAMIC RANGE IMAGING

Sidhdharthkumar Patel, Dimitrios Androutsos, Matthew Kyan

Ryerson University
Department of Electrical and Computer Engineering
350 Victoria Street, Toronto, Ontario, Canada M5B 2K3

ABSTRACT

This paper proposes a novel exposure fusion algorithm that directly fuses exposure bracketed shots into a displayable image. Most techniques targeted for direct fusion do not have an effective exposure control mechanism that can compensate for the limitations in the human visual system (HVS). The proposed algorithm offers a novel approach that adaptively adjusts its parameter for the best viewing experience. Changing the parameter adaptively aims to mitigate the perceptual loss of details caused by Weber's effect, brightening the dark regions while darkening the bright regions. The proposed method yields a perceptually detailed image when compared against other methods of similar nature.

Index Terms— Dynamic fusion, HDR, perceptual

1. INTRODUCTION

The Human Visual System (HVS) has a remarkable ability to adapt to changes in lighting conditions in order to view objects and scenes more effectively. Sunlit sky vs. moonlit sky has approximately 10^9 order of difference in the amount of radiance incident on objects. Of course, this difference is not visible to the naked eye due to the adaptive nature of eye receptors and the expansion and contraction of the pupil.

The goal of High Dynamic Range (HDR) imaging is to capture the light as faithfully as possible in order to preserve maximum information. This can be done in two ways: 1) use a camera that can capture more than 256 levels of light intensity or 2) capture images using multiple exposures and then combine the most relevant parts of the images into one final image. The former approach is an expensive solution to HDR imaging and is currently inaccessible to the general public; furthermore most display devices can only display 8 bit (per channel) images so retaining information in captured HDR image is problematic. The second method however, can be employed by existing cameras. The issue of rendering the HDR image for either method requires that the image be converted back to 8 bits before displaying. Reducing the dynamic range for displaying the image is termed tone mapping and its goal is to retain perceptually relevant information while discarding the rest.

Many existing techniques for HDR image capture have difficulty retaining information in bright areas. The human visual system (HVS) cannot perceive changes in bright areas as much as dark areas (Weber's law) and therefore the areas in bright regions are not easily detected. The problems get amplified when the image is to be displayed on a larger screen monitor, due to interpolation used to resize the image. This paper proposes a novel algorithm to improve the visual perception of the resulting image by adaptively adjusting the exposure level. The proposed exposure adjustment algorithm adaptively increases the perceived loss of details in bright/dark regions by reducing the brightness of the bright regions and increasing the brightness in darker regions. This method also allows for the image to look more natural and leverages an adjustable parameter to adjust the tone of the image.

Rest of the paper is organized as follows; first the review of previous techniques is discussed, following by the proposed algorithm. The experimental setup and the results are presented next, followed by the conclusion.

2. PREVIOUS WORK

There are many techniques available for HDR image capture (using a standard camera). These techniques follow a traditional pipeline: 1) acquire multiple images at different exposures; 2) linearize the pixels by estimating the inverse of the camera response curve; 3) fuse and process the images, and; 4) convert the resulting image back to a standard or low dynamic range (LDR) via tone mapping. Recent HDR techniques fuse images directly into a tone mapped image skipping steps 2 and 4. This paper improves upon the latter, direct approaches.

2.1 Image fusion

Most of the fusion techniques separate fusion of a set of differently exposed image from the tone mapping stage. Furthermore, almost every fusion technique requires that the camera and the scene are perfectly still. Scenes with movement are first subjected to registration techniques in

Figure 1: Result of equation 3 after setting:(a) *k*=0;(b) *k*=0.2;(c) *k*=0.4;(d) *k*=0.6;(e) *k*=0.8;(f) *k*=1 (left to right respectively)

order to align the captured images. Pixel-wise image fusion, first proposed by Mann and Picard [1] assumes that the camera and the scene are both completely static and aligned. The proposed idea was to do simple weighted average of the corresponding pixels across the exposures.

$$I = \frac{\sum \frac{w(v) f^{-1}(v)}{t}}{\sum w(v)} \quad (1)$$

where:
v = captured pixel value at a given exposure
$f^{-1}(v)$ = inverse camera response function that maps from pixel value to irradiance
t = exposure time
w = weight

Equation 1 is the general framework for weighted average approaches, and requires the estimation of camera response function and a tone mapping stage after the fusion. Mann and Picard propose the weighing function as the derivative of the camera response function. Taking motivation from Mann and Picard's work, many other weighing functions have been proposed [2-4].

The problem of weighted average was solved by fusing images in a multi-resolution fashion. Multi-scale image fusion [6-8] attempts to fuse the sequence by first decomposing the images into different sub images with different resolutions (pyramid levels). For each location in the transformed image, the value of the pyramid level with highest saliency is selected. The fused Image is then constructed by taking an inverse transform of the combined images. [6] uses this approach to fuse the images in wavelet domain whereas [7] uses wavelet packet transformation for fusion. The technique used by [8] applies fusion in gradient domain which is similar to [6] and [7]. Finally, [9] uses dictionary learning to identify which components from across the different exposures should be used in the fusion process.

2.2 Exposure fusion

Image fusion is usually computed in radiance domain and therefore, a tone mapping operator (TMO) is required for display on standard monitor. There exist another class of HDR algorithms that attempt to directly fuse images into standard image domain skipping the camera response curve estimation and TMO steps of the HDR image acquisition pipeline. These algorithms are named exposure fusion (EF).

The most influential exposure fusion algorithm was introduced by Mertens et al [10]. Merten's algorithm begins by calculating a series of quality metrics such as *contrast*, *saturation* and *well exposedness*. The algorithm combines the product of these quality metrics with the original images in multi-resolution fashion.

A slightly different method of EF was proposed by Raman and Chaudhuri [15] who used the bi-lateral filter to obtain the weight map. A more recent method by Neil Bruce [14] fuses the images in log domain using local entropy as a weighing function.

3. PROPOSED METHOD

One issue with most of the EF techniques is that these algorithms do not have any mechanism to adjust the values of the exposure images. Mertens's algorithm was modified by [11] to work in L*a*b* colour space in order to make the fused image look more natural; however details in the darker regions seem to be attenuated. In order to address the issue of loss of details this paper proposes a novel exposure adjustment stage as a pre processing step and a novel contrast and saturation calculation algorithm, which adaptively darkens the brighter region and brightens the darker regions whilst retaining the naturalness of the scene.

The proposed algorithm begins by converting the colour space to L*a*b* space as in [11]. The next stage in the algorithm applies exposure adjustment to the L* channel of all input images. The aim of the exposure adjustment is to adaptively increase the details in dark/bright region of the image. Unlike the original algorithm of Mertens and [11], the exposure adjustment is applied to the image before any fusion has taken place. The exposure adjustment needs to be applied before the fusion because the image itself needs to be *corrected* before any fusion takes place. This change is important because combining (multiplying) the *well exposedness* metric with other metrics (as per original algorithm) has a chance of corrupting the weight map. For instance, effect of a high *contrast* value would be nullified by a low *well exposedness* value. In the proposed approach it is thus dealt with separately from other quality measures. The proposed algorithm then computes a novel *contrast* and *saturation* metric that can give pleasing results when combined with exposure adjustment algorithm.

3.1 Proposed algorithm

The fusion of images is performed in multi-resolution fashion as described in [10]. The proposed algorithm is summarized in pseudo-code below (with modifications to [10] highlighted in bold):

```
Exposure Fusion (LDR_images) {
    LDR_images = exposure_Adjust(LDR_images)
    C = scale(contrast(LDR_images);
    S = scale(saturation(LDR_image));
    Quality_measure = normalize(C*S);
    for i = number_LDR_images{
            W = gaussian_pyramid(quality measure)
            I = laplacian_pyramid(lab_im)
            for j = 1:number_levels
            R[j] =R[j]+ W[j]*I[j]
        }
    }
    out = reconstrunct_pyramid(R)
}
```

3.2 Contrast and Saturation

The contrast is calculated by filtering the image with a Laplacian of Gaussian (LoG) filter. The LoG filter was chosen based on the result of [11] which showed a variation of Difference of Gaussian (DoG) filter [12] can yield a better indication of a contrast in the image. The DoG filter is a close approximation of LoG filter and therefore a LoG filter was used in the present implementation. Saturation is defined as the measure of colourfulness of the picture. The saturation is calculated as the absolute difference of a* and b* channel. The rationale behind it is that for a region containing no colour, the a* and b* component of the L*a*b* colour space would have same value while an area with a colour would have different a* and b* component.

3.3 Exposure adjustment

The main contribution of this paper is to propose a novel exposure adjustment algorithm that can give more detailed results. Equation 2 shows the general equation for exposure adjustment:

$$I(x, y) = I(x, y) \exp\big(k - I_{ave}(x, y)\big) \qquad (2)$$

Where: k = brightness level
I = input image (assumed normalized to range 0-1)
I_{ave} = input image filtered with 3x3 averaging mask

Equation 2 applies a simple exponential weighing to the original image which is characterized by the local average information. The local average (of luminance) is a good measure of amount of brightness in a region surrounding a given pixel. The parameter (k) of equation 3 determines when a pixel should be amplified or attenuated and by how much. For example, any value of I_{ave} larger than k will attenuate the pixel while any value of I_{ave} smaller than k will amplify the input luminance; any I_{ave} value equal to k will not affect the luminance value. In cases where salt noise is present (typically in low light), there is a possibility that the resulting luminance value is greater than 1, therefore the output of the exposure adjustment needs to be scaled to a value between 0 and 1.

Figure 1 shows the result after setting k to 0, 0.2, 0.4...1. A note worthy aspect from figure 1 is that k value of equation 2 determines how dark or bright the resulting image will appear. For example setting k=0 will result in very dark image compared to setting k=1.

3.4 Adaptive *k*

Equation 2 is the general equation describing how the exposure of a given image should be adjusted. As discussed earlier, k determines the overall tone of the image and it can be left as a parameter that the user can set. To make the exposure adjustment algorithm more robust, we propose an automated scheme to set k adaptively, in order to more appropriately compensate for Weber's effect.

An image usually contains many types of regions but these regions can be classified into two major groups:1) light/dark regions and 2) detailed/homogeneous regions. In order to overcome the limitation of HVS, it is natural to darken the bright regions and similarly the dark regions need to be amplified, however, Weber's effect indicates that only regions containing information need to be adjusted while the information free regions of the image can remain un-altered.

Based on rationale presented above, the k value needs to satisfy the following:

- bright and high information region - low k value
- bright and low information region - high k value
- dark and high information region - high k value
- dark and low information region - low k value

The first 2 points will insure that only the informative regions of bright areas will darken (to compensate for Weber's effect). Similarly, the last 2 points insure that only informative regions in the dark area will get brighter. Thus we aim to selectively modify contrast when a neighborhood of fine details is present.

As discussed earlier, an image filtered by an averaging mask is a good indication of how bright or dark the surrounding region is. Edges in an image (high frequency regions) can be considered as highly informative. Therefore an image filtered by a Laplacian of Gaussian (LoG) mask can be used as an indicator of information content in the image. Using an image filtered by an averaging mask and a

LoG mask to determine the k value, can provide an image of localized k values instead of a constant over the entire scene. Equation 3 illustrates a simple scheme to calculate the k values satisfying all four points discussed above.

$$\boldsymbol{k} = (\boldsymbol{1} - \boldsymbol{L_{edge}}) * (\boldsymbol{L_{ave}}) + (\boldsymbol{L_{edge}}) * (\boldsymbol{1} - \boldsymbol{L_{ave}}) \qquad \textbf{(3)}$$

The first part of equation 3 satisfies the first 2 points while the second part of equation 3 satisfies the last 2 points. In order to remove the over amplification of dark regions, the k value is set to 0 when the input luminance lies below 0.1.Note: L_{edge} and L_{ave} are scaled between 0-1 range before using them in equation (3);

4. EXPERIMENTAL SETUP AND RESULTS

The proposed algorithm was tested on the dataset provided by Yeganeh and Wang [13]. The photographs in the dataset are a compilation of images used in many HDR research dating back as far as 1997 [1]. Figures 2(a)-(d) show the resulting image after applying proposed algorithm, [11], [14] and [15] respectively. The source code for the image in figure 2(c) and 2(d) was obtained from Banterle's site [16].

4.1 Discussion

Figures 2(a)-(d) show the fusion results using the most recent exposure fusion algorithms. The images in Figure 3 are based on algorithms that are designed specifically to combine images into tone-mapped images. The key difference between all four of the images is the amount of details that can be perceived, and the washed out effect in the dark, detailed (edge rich) areas of the image after fusion. The proposed algorithm aims to increase the details in both bright and dark regions, whilst retaining the naturalness of the image which other methods fail to show.

Figure 2(c) and 2(d) shows the result of using Bruce's and Raman's algorithm respectively [14][15]. The bright regions (sky) of the images are overly bright which have led to loss of details in the clouds. The region in the clouds in figure 2(c) and 2(d) shows where this loss of detail has occurred and the same region in figure 2(a) clearly show that the proposed algorithm has more details. Because figure 2(c) and 2(d) are bright, the regions in the dark area are visible but they look washed out. Figure 2(b) is based on a method that is most similar to proposed method. The cloud regions in figure 2(b) appear detailed however dark regions (e.g. enclosed by the lower rectangle), show that these areas shows a loss of details and appear "washed out"; the same regions in figure 2(a) show more visible structural details.

5. CONCLUSION

This paper proposes a novel algorithm to fuse images taken at different exposures using a multi-resolution exposure fusion approach. A novel adaptive exposure adjustment is proposed that can increase an image's perceived details and naturalness. The main objective of proposed exposure adjustment algorithm is to increase the perceived details in the bright/dark region of the image. This paper also proposes a novel saturation and contrast metric that can increase the naturalness of the image. The resulting image yielded a perceptually detailed image when compared against related techniques from the literature. The results were evaluated empirically due to lack of available evaluation tools that measure a human's "perceived" details. However, future research will be focused on more accurate and robust evaluation tools using perceptual based metrics.

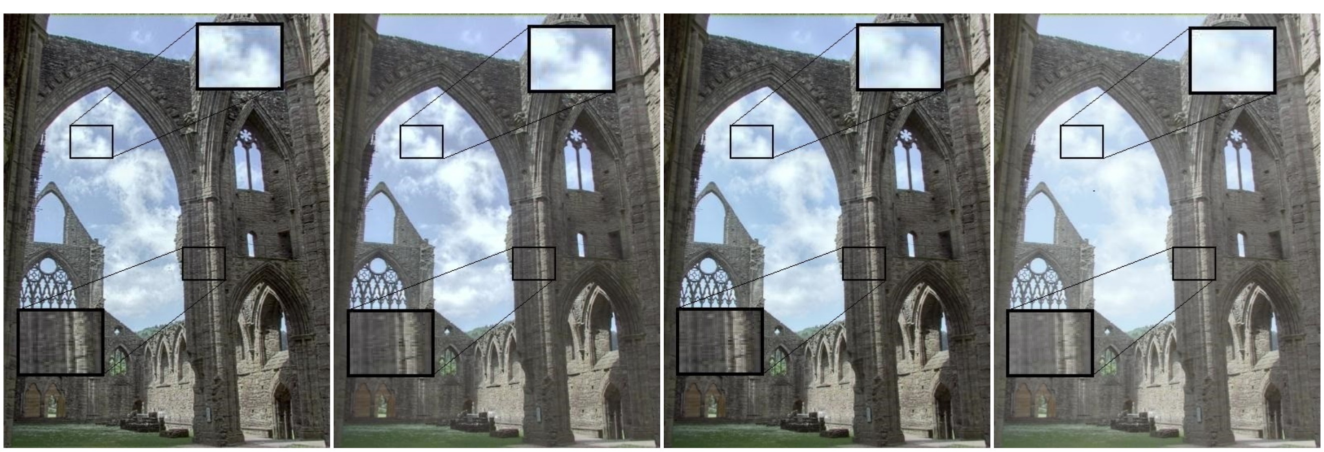

Figure 2: Comparison of proposed method with other similar methods that combine fusion with tone mapping: (left to write)(a) proposed algorithm;(b) Pablo Martínez-Cañada (2014) [11]; (c) Bruce(2014) [14]; (d) Raman and Chaudhuri (2007) [15]

6. REFERENCES

[1] S. Mann and R. W. Picard, "On Being `undigital' with Digital Cameras:Extending Dynamic Range By Combining Differently Exposed Pictures," in *Proceedings of IS&T*, Boston, Massachusetts, 1994.

[2] Paul E. Debevec and Jitendra Malik, "Recovering high dynamic range radiance maps from photographs," in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, New York, NY, 1997.

[3] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," in *IEEE Computer Society Conference on. , vol.1, no., pp.,380 Vol. 1*, 1999.

[4] ERIK REINHARD et al., *High Dynamic Range Imaging: Acquisition, Display*.: Morgan Kaufmann, 2010.

[5] T. Jinno and M. Okuda, "Multiple Exposure Fusion for High Dynamic Range Image Acquisition," in *mage Processing, IEEE Transactions on , vol.21, no.1, pp.358,365*, 2012.

[6] Xu Jianbo, Youjun Huang, and Jianli Wang, "Multi-exposure images of wavelet transform fusion," in *Fifth International Conference on Digital Image Processing* , 2013.

[7] Qi Wang, Zongxi Song, and Wei Gao, "A multi-exposure image fusion method based on wavelet packet transform," in *International Symposium on Photoelectronic Detection and Imaging 2013*, 2013.

[8] Bo Gu, Wujing Li, Jiangtao Wong, Minyun Zhu, and Minghui Wang, "Gradient field multi-exposure images fusion for high dynamic range image visualization," *Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. Pages 604–610, 2012.

[9] Jinhua Wang, Hongzhe Liu, and Ning He, "Exposure fusion based on sparse representation using approximate K-SVD," *Neurocomputing*, vol. 135, pp. 145–154, 2014.

[10] Tom Mertens, J. Kautz, and F. Van Reeth, "Exposure Fusion," *Computer Graphics and Applications*, pp. 382,390, 2007.

[11] Pablo Martínez-Cañada and Marius Pedersen, "Exposure Fusion Algorithm Based on Perceptual Contrast and Dynamic Adjustment of Well-Exposedness," *Lecture Notes in Computer Science*, vol. 8509, pp. 183-192, 2014.

[12] A, A. Rizzi, G. Simone, and R Cordone, "A modified algorithm for perceived contrast," in *Conference on Colour in Graphics, Imaging, and Vision*, 2008.

[13] Hojatollah Yeganeh and Zhou Wang. [Online]. https://ece.uwaterloo.ca/~z70wang/research/tmqi/

[14] Neil Bruce, "ExpoBlend: Information preserving exposure blending based on normalized log-domain entropy," *Computers & Graphics*, vol. 39, pp. 12-23, April 2014.

[15] S. Raman and S. Chaudhuri, "Bilateral Filter Based Compositing for Variable Exposure," January 2009.

[16] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*. Natick, USA: AK Peters (CRC Press)}, 2011.