

Glare Encoding of High Dynamic Range Images

Supplemental Material

Mushfiqur Rouf¹ Rafał Mantiuk^{1,2} Wolfgang Heidrich¹
 Matthew Trentacoste¹ Cheryl Lau¹

¹The University of British Columbia ²Bangor University

1. Introduction

In this document, we present some additional background material for our work. The sections are organized as follows: Section 2 expands on parts of derivation in the main paper. Section 3 gives a few synthetic results in addition to the real results presented in the paper. And finally, Section 4 gives an analysis of effects of noise on our algorithm.

2. Additional derivations

In this section, we include more details on some aspects of the derivation described in the technical paper.

2.1. MLE of Laplace distribution mean

Here we derive the maximum likelihood estimator of the mean of a Laplace distribution. If a random variable X has a Laplace distribution,

$$X \sim \text{Laplace}(\mu, b), \quad (1)$$

where μ is the location parameter and $b > 0$ is the scale parameter, then the probability density function is given by

$$f(x|\mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}}. \quad (2)$$

The maximum likelihood estimator of μ is given by

$$\hat{\mu} = \arg \max_{\mu} \left(\log \prod_i \frac{1}{2b} e^{-\frac{|x_i-\mu|}{b}} \right) = \arg \min_{\mu} \sum_i |x_i - \mu|. \quad (3)$$

This result shows that the maximum likelihood estimator of mean of a Laplace distribution can be computed using L_1 norm.

2.2. Gradient properties of the residual glare, r_S

In our paper, we introduced a low energy, high frequency, shift variant component $\hat{\rho}$ of the light transport

function H , that captures cross-screen filter diffraction effects r_S present in the captured image. In this section, we derive statistical properties of this function r_S which we use in our paper (Equation 9 in the paper). First, we present empirical properties of derivatives of $\hat{\rho}$, and then relate them to derivatives of r_S .

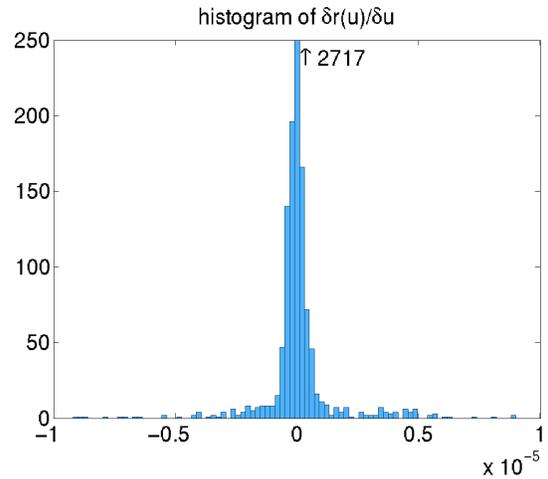


Figure 1: Histogram of $\frac{\partial r}{\partial v}$ has a highly peaked structure, which can be closely approximated with a Laplace distribution.

The r_S function accounts for wavelength-dependent variation is the PSF. There is no need to exactly model its spectral characteristics in our application, since we only need to estimate the resulting color distortions. Removing them is important for glare removal, but is not essential for estimating the saturated pixel.

ρ has the same star-shape as the shift-invariant part, although the exact energy distribution within the star is dependent on image location, i.e.:

$$\rho(\mathbf{x}, \mathbf{y}) = \begin{cases} \sum_{i=1}^{p/2} \hat{\rho}_i(\mathbf{x}, \mathbf{y}) & \text{when } \mathbf{v}_i \cdot (\mathbf{y} - \mathbf{x}) = 0, \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

where, \mathbf{u}_i and \mathbf{v}_i form an orthogonal coordinate system aligned along the i^{th} glare direction. Now we can apply the same separability argument again, and assume that the rays are independent. For each individual glare ray, we show that $\hat{\rho}_i$ gradients are distributed as a Laplace distribution (Figure 1). By definition of r ,

$$\frac{\partial}{\partial \mathbf{u}_i} r_{\mathbb{S}}(\mathbf{x}) = \sum_{\mathbf{y}} f_{\mathbb{S}}(\mathbf{y}) \frac{\partial}{\partial \mathbf{u}_i} \hat{\rho}_i(\mathbf{x}, \mathbf{y}) \quad (5)$$

$$\sim \text{Laplace}\left(0, \sum f_{\mathbb{S}}(\mathbf{y})\right), \quad (6)$$

and

$$\frac{\partial}{\partial \mathbf{v}_i} r_{\mathbb{S}}(\mathbf{x}) = \sum_{\mathbf{y}} \left(f_{\mathbb{S}}(\mathbf{y}) \frac{\partial}{\partial \mathbf{v}_i} \hat{\rho}_i(\mathbf{x}, \mathbf{y}) + \hat{\rho}_i(\mathbf{x}, \mathbf{y}) \frac{\partial}{\partial \mathbf{v}_i} f_{\mathbb{S}}(\mathbf{y}) \right) \quad (7)$$

$$\sim \text{Laplace}\left(0, \sum (f_{\mathbb{S}}(\mathbf{y}) + \hat{\rho}_i(\mathbf{x}, \mathbf{y}))\right), \quad (8)$$

since $\hat{\rho}_i$ is directly related to pixel color and hence vertical gradients follow a $\text{Laplace}(0, b)$ distribution similar to image gradients. Our experiments have confirmed this assumption.

Although the standard deviations in the above expressions contain unknown quantities, since our algorithm runs row-by-row, they can be considered constant and hence can be ignored while solving the linear system.

2.3. Modeling constraints on residual glare

We can combine the statistical properties of derivatives of residual glare r to build the constraint term R :

$$R = \sum_{i=1}^{p/2} \left(\lambda_1 \|r_i\|_2 + \lambda_2 \left\| \frac{\partial}{\partial \mathbf{v}_i} r_i \right\|_1 + \lambda_3 \left\| \frac{\partial}{\partial \mathbf{u}_i} r_i \right\|_1 \right). \quad (9)$$

The λ_1 term ensures that r has a zero-mean normal distribution. The λ_2 term ensures that the function is consistent along \mathbf{v} axis. The rationale behind this consistency is that the light in the saturated region has a consistent spectral composition and thus produce similar diffraction patterns. This was also confirmed in our experiments. The λ_3 term ensures that the r function is also smooth along glare rays. We found that the color variations due to diffraction are much smoother than the image content and thus is also a strong prior that let us separate glare from the latent image.

2.4. Arguments against deconvolution

Although our derivation so far as Equation 9 in the paper demonstrates that image restoration fundamentally involves a deconvolution problem, we found during extensive experiments that standard deconvolution methods are impractical for a number of reasons.

First, the support of the exponential glare PSF K typically exceeds the width of the image, making robust deconvolution methods prohibitively expensive. This property of our glare also makes every pixel in the image a ‘‘boundary’’ case, which most standard deconvolution methods do not handle gracefully.

Secondly, standard deconvolution methods do not exploit the characteristic of the cross-screen filter. The star shape of light transport function lets us ignore a large number of inter-dependencies that are unimportant for the solution but increases the complexity of the problem.

Third, the non-Dirac components act as a low-pass filter along the glare direction, which implies that estimated glare only gives 1D ‘‘projections’’ of the highlights, which obviates the use of a tomography approach. It also results in suppressed frequencies, in which case conventional deconvolution algorithms enhance the noise in measurement and estimation, and produce noisy and inaccurate highlight estimates.

Finally, the problem is heavily under-determined, and can be solved only with natural image priors, well-selected regularization and robust solvers. Using these methods in the context of the full 2D deconvolution problem is very problematic, computationally expensive, and, according to our experiments, impractical.

3. Synthetic results

Figure 2 shows a number of results using input images with *synthetic* glare. In the left column, we show the LDR input image, which has been generated by convolving an HDR image with the filter PSF, and adding noise and quantization. The center two columns show different exposures of our result, while the right column shows a short exposure rendition of the original HDR image. Row (a) shows an image with number of small specular highlights, which are reconstructed faithfully by our approach. The second row tests a difficult case for our glare estimation approach, since the glare rays are aligned with a strong image edge (the horizon). Sparse gradient prior for gradient estimation is not valid in this case, resulting in a misestimation of the horizontal glare (circled region in the center-left image), and hence a lower-quality reconstruction of the saturated region (zoomed-in in the center-right image). These artifacts are best seen in the electronic version of the paper. The example in row (c) shows that these artifacts can be solved by rotating the cross-screen filter such that glare rays do not align with strong image edges.

4. Analysis

In this section we analyze the trade-offs between glare detection and increased camera noise.

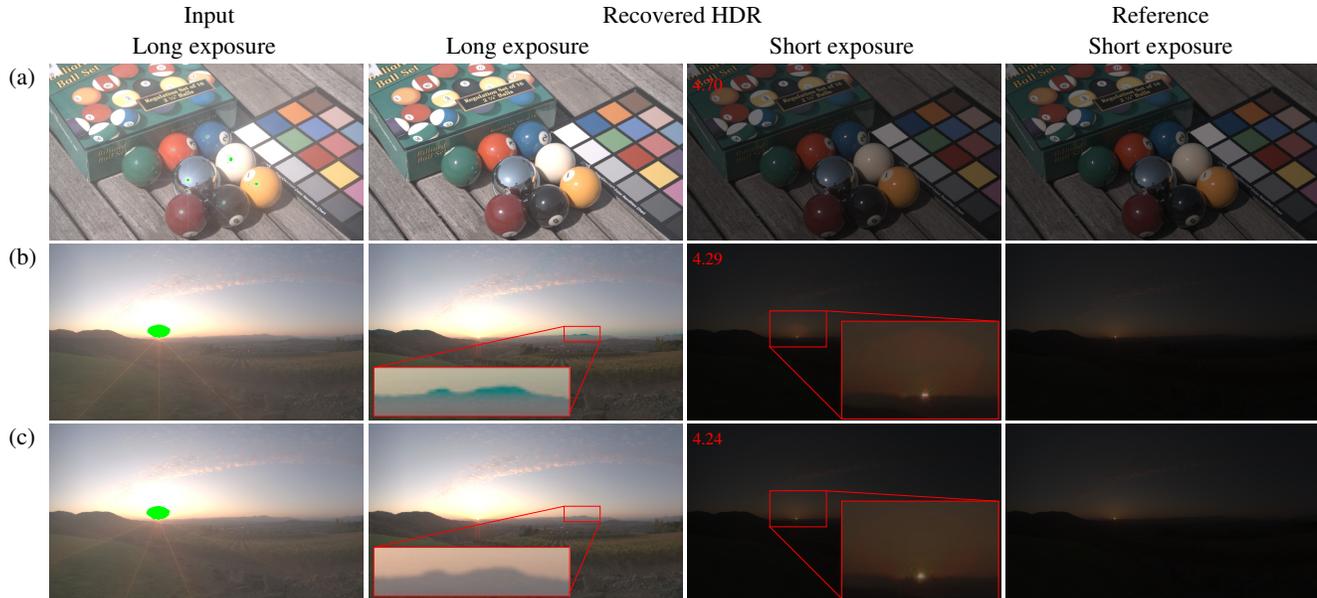


Figure 2: Reconstruction results for simulated glare from an 8-point cross-screen filter. The first column shows the synthetically blurred input image with saturated pixels marked green. (b) and (c) show the same image but in case of (c) the filter was rotated by 22.5 degrees to avoid aligning glare patterns with the sky gradient and thus better glare estimation was obtained. The third column gives the dynamic range increase at the top-left corner of each image. See text for a full discussion.

4.1. Glare detection

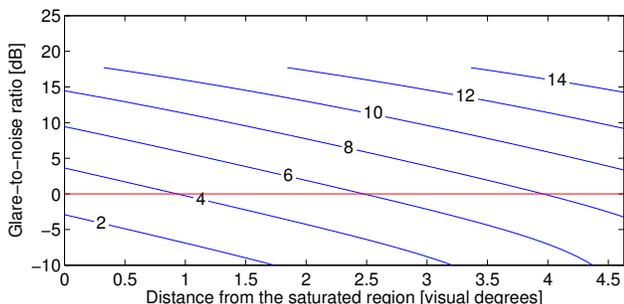


Figure 3: Glare-to-noise ratio that characterizes the detectability of a glare signal. The numbers next to lines indicate how much brighter (in f-stops) the glare source is relative to the sensor clipping level.

The cross-screen filter can effectively encode information about most, but not all clipped pixels. This is because, in order to be detected, a glare pattern must be at least a few times stronger than the camera noise level without saturating glare pixels. Figure 3 shows the glare-to-noise ratio in dB for a source of glare that has a width of 0.5 visual degrees, is perfectly flat, and is from 2^2 to 2^{16} times brighter than the sensor clipping level (numbers on the lines). The glare-to-noise values are given for the pixels that are located x visual degrees from the source of glare (x -axis, 100mm lens). The image region receiving the glare signal is uni-

form, and its pixel value is 2 f-stops below the sensor clipping level. The exponential model of the 8-point PSF was used to create the plot. We used a simple camera noise model that consists of normally distributed static noise with the standard deviation $\sigma_s = 0.0002$ (for a maximum sensor value equal to 1) and signal-dependent noise with the standard deviation $\sigma_d = 0.013$. The glare-to-noise ratio was computed as

$$GNR = 10 \log_{10} \left(\frac{g - f}{\sqrt{\sigma_d^2 g + \sigma_s^2}} \right), \quad (10)$$

where f is the original pixel value without glare, and g is the pixel value with glare. We used the noise parameters to approximate the characteristics of our Canon D40 (200 ISO, 5.6f, standard post-processing settings), although these can vary with aperture, ISO settings, sensor temperature and other factors. The parameters we found by least-square fitting of the model to the noise found in a gray card photographed with varying illumination levels.

Figure 3 shows the trade-off between clipping glare pixels and capturing glare that is too weak to be detected. A 0.5 visual-degree segment of clipped pixels must be at least 5–6 f-stops (32–64 times) brighter than the clipping level to produce glare that is detectable. A brighter or larger source of glare produces a higher glare-to-noise ratio, but if it is much brighter, pixels can get saturated, and thus lose encoded information. This is shown on the plot as clipping

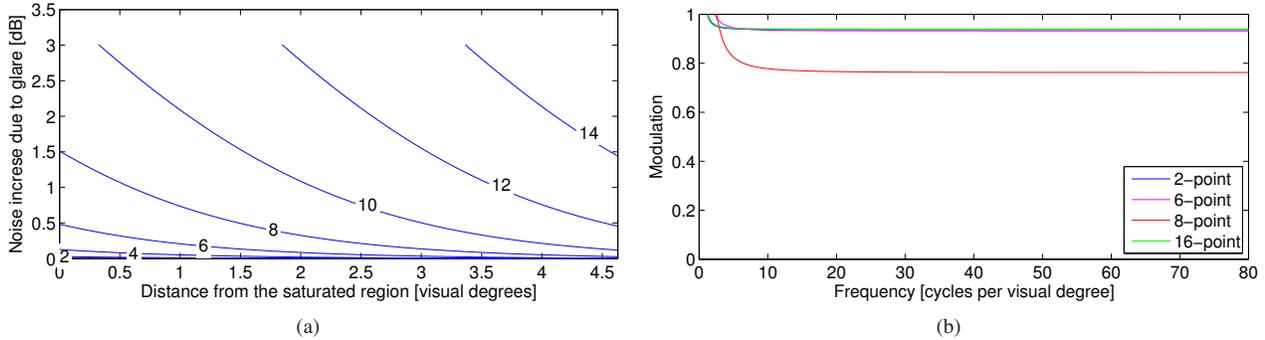


Figure 4: (a) Noise increase due to glare removal. The curves are generated for the same conditions as in Figure 3. The higher noise is caused by higher shot noise for the pixels captured with glare than for the same pixels captured without glare. (b) Modulation transfer functions (MTFs) of the cross-sections filters. The filters and projections are the same as in Figure 2(c) in the paper. The exponential models of the PSFs were used to compute the MTFs in order to remove the MTF of the lens system.

of the lines above 20 dB. To avoid saturation, the exposure time needs to be shortened, but this increases noise in an image [1]. Saturation of glare pixels can be also avoided if a cross-screen filter that produces weaker glare (*i.e.*, smaller β) is used. Such a filter, however, results in a smaller glare-to-noise ratio, making the glare difficult to detect and estimate. The best results are achieved if the cross-screen filter is selected to produce just detectable glare with possibly large exposure time, while avoiding saturation of glare pixels. The specific values on the plot apply to the specific setting outlined above, but the qualitative analysis applies equally to other cameras and glare sources.

4.2. Noise analysis

Encoding additional information in unsaturated pixels has one drawback: it increases the noise level. Fortunately the cross-screen filter has a relatively small impact on noise. Since shot noise is proportional to the square root of the signal, pixels affected by glare due saturated pixels ($\mathbb{S} \rightarrow \mathbb{U}$) have higher shot noise than if the same pixels were captured without glare. In Figure 4(a) this effect is simulated for the same parameters as used in the glare analysis (Figure 3).

Since the cross-screen filter spreads light in discrete directions, this noise increase is much smaller than for typical veiling glare in lenses [1], and affects only a small percentage of pixels. The noise is also increased due to the deconvolution that we perform when removing glare due to unsaturated pixels. This noise increase is also moderate, about 0.31 dB for all our filters except the 8-point one, which can boost noise up to 1.17 dB. The numbers are explained by the modulation transfer functions (MTFs) of cross-screen filters, which have very high values for all frequencies, as shown in Figure 4(b). If the deconvolution is performed in the Fourier domain, the frequency components are multiplied by the inverse of the MTF. Since this multiplication boosts the contrast of both image details and noise, the noise increase can be approximated by the inverse of the MTF values.

References

- [1] E. Talvala, A. Adams, M. Horowitz, and M. Levoy. Veiling glare in high dynamic range imaging. *ACM Trans. Graph.*, 26(3):37, 2007. 4