

Glare Encoding of High Dynamic Range Images

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

¹The University of British Columbia ²Bangor University

Abstract

Without specialized sensor technology or custom, multi-chip cameras, high dynamic range imaging typically involves time-sequential capture of multiple photographs. The obvious downside to this approach is that it cannot easily be applied to images with moving objects, especially if the motions are complex.

In this paper, we take a novel view of HDR capture, which is based on a computational photography approach. We propose to first optically encode both the low dynamic range portion of the scene and highlight information into a low dynamic range image that can be captured with a conventional image sensor. This step is achieved using a cross-screen, or star filter. Second, we decode, in software, both the low dynamic range image and the highlight information. Lastly, these two portions can be combined to form an image of a higher dynamic range than the regular sensor dynamic range.

1. Introduction

Camera sensors can capture a certain maximum number of photons before they start to saturate and no longer register additional light. Although it is possible to increase the saturation point by increasing the capacity of the sensor electron well, producing large sensors is excessively expensive and reduces sensor resolution. Such sensors are also hard to justify for general imaging applications because, on average, only a small portion of a scene contains very bright spots and thus needs high capacity sensors.

The human visual system has developed a clever mechanism to cope with highly saturated scene regions, such as highlights or light sources. Like camera sensors, the photoreceptors in the human retina are also prone to saturate. However, the visual system is able to infer higher brightness of those saturated regions from glare, which is produced by the light that is scattered in the ocular fluid and spread over the retina. The glare surrounding bright areas boosts their perceived brightness, giving additional information to the

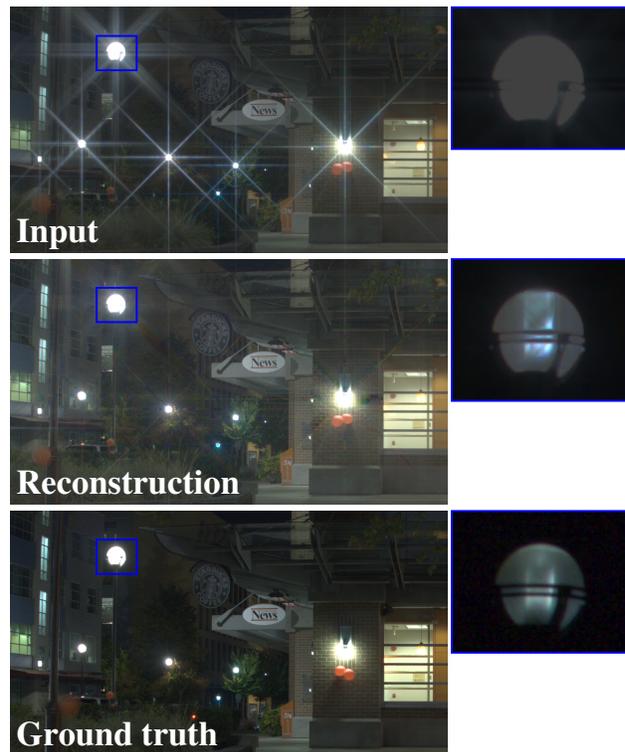


Figure 1: Capturing a high dynamic range (HDR) image with a cross-screen filter. Insets show bright regions at a shorter virtual exposure. The dynamic range increase is 9.21 f-stops. The ground truth has been constructed from a series of 16 exposures and hence has a lower noise level compared to our single exposure result.

brain that this part of a scene is much brighter than the photoreceptor saturation point [25].

In this paper we propose to use a similar approach to improve camera dynamic range without resorting to custom sensors, multi-sensor cameras, or time-sequential imaging. Unlike the eye, we are not limited to specific optics. Instead, we can choose to modify the optical system in order to increase the information that is encoded for the saturated areas. Our goal can thus be more ambitious than simply to estimate the overall brightness of the saturated image

regions. Instead, we would like to reconstruct spatial detail for the pixels in those areas. Specifically, we propose a computational photography approach comprised of the following steps:

- **Encoding.** Details of bright image regions in a high dynamic range (HDR) image, such as highlights and directly visible light sources, are encoded into specially shaped glare patterns optically added to the image.
- **Capture.** The encoded image is captured using a standard image sensor. Bright regions in the captured image are saturated due to limited sensor dynamic range.
- **Decoding.** In software, we separate the glare pattern from the low dynamic range version of the image. The glare pattern can be used to infer the radiometric intensity distributions in the saturated image regions.

We have experimented with a number of specific optical encodings to implement this general principle. Some obvious candidates are regular lens glare and defocus blur to spread out energy from saturated image regions to other pixels. However, to provide enough information of the highlight regions for detailed reconstruction, energy spread must be significantly larger than standard lens flare. Likewise, a defocus blur implementation would have to use very large blur radii on the order of dozens of pixels. For such large blur, even the most recent deconvolution algorithms in combination with coded apertures fail to reconstruct high quality images [24].

In this paper, we therefore focus on the optical encoding that we found most successful: a glare pattern that scatters light in a fixed set of discrete directions. Such patterns are produced by inexpensive photographic cross-screen filters (also known as star filters), which are mounted in front of a camera lens. The scattering pattern of these filters is most salient for very bright scene features since the star filters concentrate most energy in a Dirac peak rather than the glare rays. Star filters spread the light in discrete directions, and therefore one dimensional techniques can be applied instead of more expensive and less stable 2D techniques. These properties let us estimate the amount of light spread from bright image features into several discrete directions (from 2 to 16), and then reconstruct clipped pixels using a tomographic reconstruction technique.

2. Related work

Multi-exposure HDR capture: Blending multiple exposures [4] is the most accurate method for acquiring high dynamic range images with conventional cameras. However, this approach is limited by ghosting and misalignment problems [9], which are still largely unsolved for difficult cases such as moving tree leaves or waves on the water. There are ways of obtaining multiple simultaneous exposures (e.g. [15, 16]), and to design sensors that directly sup-

port multi-exposure capture (e.g. [6]), but such cameras and sensors are not currently widely available.

LDR to HDR enhancement: Reconstructing an HDR image from a single exposure with clipped values is a challenging problem that yields only approximate solutions. Several techniques have been developed (e.g. [3, 21, 14, 5]), however these are merely heuristics that are used to plausibly guess content that has ultimately not been captured.

Clipped signal restoration: For band limited 1D signals, reconstruction algorithms have been proposed for situations where the number of clipped samples is low [1], or where a statistical model of an undistorted signal is known [17]. However, neither of these approaches can be trivially extended to images, because natural image statistics are too weak to restore detailed texture in clipped regions. Therefore, only special cases have been successfully solved in the image domain, for example images where only a subset of the color channels is clipped [26], or noisy images with pixel values just above the clipping threshold [8].

Deconvolution: A large body of recent work has focused on the development of new deconvolution algorithms, as well as special, frequency-preserving convolution kernels for both motion blur (e.g. [19, 13]) and depth-of-field blur (e.g. [12, 27]). In principle, both motion blur and depth-of-field blur could be used to spread energy of bright pixels in a fashion similar to what we propose in this paper. However, a sufficiently large energy spread can only be achieved with very large blur kernels. In our experiments, we found that even the combination of state-of-the-art deconvolution methods with special kernel shapes fails to recover a high quality, sharp image for these large radii. This is consistent with recently published results [24]. Another problem with using convolution methods is that most recent deconvolution algorithms cannot reconstruct clipped pixels.

Our approach using a cross-screen filter avoids these problems, since the filter produces a collection of 1D streaks that can be detected and removed reliably, while encoding enough information of the saturated regions to allow for detailed reconstruction of clipped pixel values.

Glare removal: Over the years, a number of approaches have been proposed for removing lens glare. Since we rely on strong glare for obtaining information about clipped image regions, the methods that optically suppress glare (e.g. [23, 20, 11]) are not applicable in our setting. On the other hand, deconvolution methods that remove the glare after the fact (e.g. [22, 7]) suffer from the same shortcomings as the other deconvolution methods discussed above.

3. Image formation model

In the following, we outline the image formation process for cameras with a cross-screen filter before we go into the details of our approach.

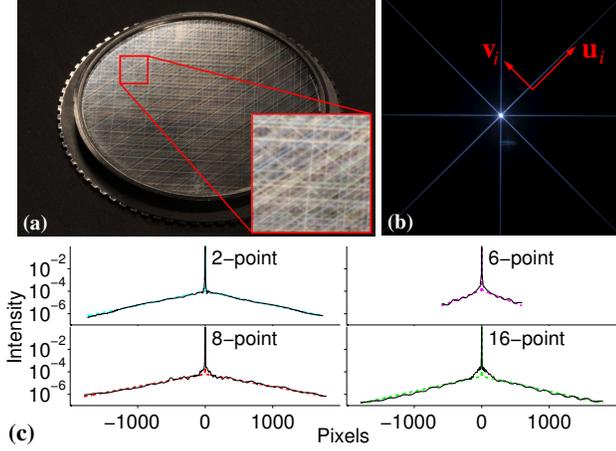


Figure 2: (a) An 8-point cross-screen filter. (b) A point light source seen through it. (c) Measured point-spread functions (PSF) for the 1D slices along glare lines for different cross-screen filters, taken from images like (b). Exponential approximations are shown as dashed lines.

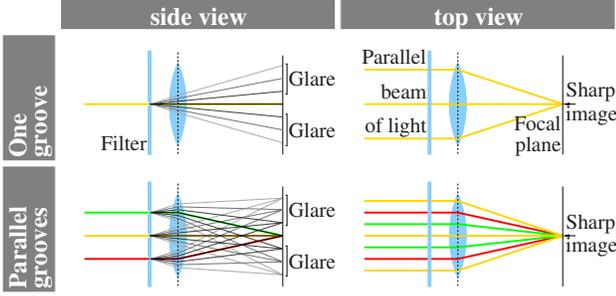


Figure 3: The working principle of a 2-point cross-screen filter. **Top row:** A single horizontal groove systematically spreads out incoming light only vertically, yet keeps it focused horizontally. **Bottom row:** Multiple parallel grooves makes the effect stronger.

A cross-screen filter is a transparent photographic filter with parallel scratch marks or grooves on its surface (Figure 2(a)). When mounted in front of a camera lens, the grooves disperse and diffract the light, creating a star-shaped glare – linear streaks (Figure 3) in a number of discrete directions. This glare is very faint and hence star-shaped glare patterns are usually noticeable only around very bright areas.

A captured image g can be expressed as a result of applying a light transport operator H describing the glare to the *latent* image f , and then clipping the result to the maximum sensor value:

$$g(\mathbf{x}) = \min(1, \sum_{\mathbf{y}} f(\mathbf{y})H(\mathbf{x}, \mathbf{y}) + n). \quad (1)$$

Here, \mathbf{x} and \mathbf{y} refer to two dimensional image coordinates, and n represents noise. For simplicity, we ignore noise n

in the rest of the derivation and discuss its influence on results in the supplemental material. H can be modeled as a combination of following components (Figure 4),

- a Dirac peak representing the light that does not hit one of the scratches on the cross-screen filter,
- a glare function K which has been empirically found to be both shift- and depth-invariant¹, and
- a zero-mean residual waviness in glare, ρ , that is *not* shift-invariant, but several orders of magnitude weaker in intensity.

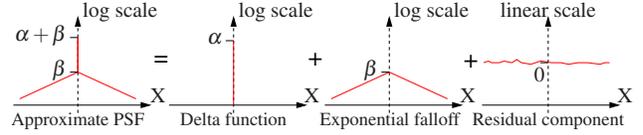


Figure 4: The kernel can be approximated by a sum of a Dirac delta function and an exponential falloff. The residual component accounts for a shift-variant wavelength-dependent response.

Thus,

$$H(\mathbf{x}, \mathbf{y}) = \alpha \delta(\mathbf{y} - \mathbf{x}) + \beta K(\mathbf{y} - \mathbf{x}) + \gamma \rho(\mathbf{x}, \mathbf{y}). \quad (2)$$

For the filters we used, $\alpha \approx 1$, $\beta \approx 10^{-4}$ and $\gamma \approx 10^{-7}$. The glare function K is itself composed of 1D streaks,

$$K(\mathbf{x} - \mathbf{y}) = \begin{cases} \sum_{i=1}^{p/2} k_i(\mathbf{u}_i \cdot (\mathbf{x} - \mathbf{y})) & \text{when } \mathbf{v}_i \cdot (\mathbf{x} - \mathbf{y}) = 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

with an exponential falloff $k_i(d) = e^{-m|d|}$. Here, \mathbf{u}_i and \mathbf{v}_i form an orthogonal coordinate system aligned along the i^{th} glare direction (see Figure 2(b)). Note that the parameters α , β , γ and m can be measured for each cross-screen filter by capturing an (almost) point light source and measuring these statistics. In our experiments, we have observed that these quantities are independent of focal depth and position.

The scene dependent residual waviness function ρ is primarily a function of the (unknown) spectral composition of the scattered light. Although this function is shift-variant, it too only distributes energy along radial lines, like K .

Figure 2(c) shows cross-sections along glare streaks of 2D PSFs for several cross-screen filters we obtained. These measurements show that an exponential falloff model fits the overall shape of the glare quite well. In our application, this exponential model is sufficient for glare estimation with sufficient precision for saturated pixel reconstruction. The high-frequency variations captured in ρ are, however, important for removing glare from low dynamic range portion

¹The glare streaks are, however, created by focusing the glare pattern through the camera lens, and hence are subject to radial lens distortion. In our discussion, we assume that radial distortion has been removed.

of the image. The overall image formation model is then given as

$$g = \min(1, \alpha f + \beta K * f + \gamma r), \quad (4)$$

where $r(\mathbf{x}) = \sum_{\mathbf{y}} \rho(\mathbf{x}, \mathbf{y}) f(\mathbf{y})$ is the result of a ‘‘convolution’’ of the intrinsic image with the shift-variant residual waviness pattern.

In summary, our image formation model consists of a Dirac part and a combination of $p/2$ 1D functions describing both an exponential falloff and a residual waviness. In the following, we can therefore consider the glare removal problem as a set of independent 1D problems.

4. Decoding method

We now describe our proposed method for decoding both the low dynamic range image and the highlight details from a glare photograph taken with a cross-screen filter.

Considering the light transport (Equation 4), we can see that it is not possible to directly solve for the glare-free latent image, due to sensor saturation. Instead, we split the problem by separately considering the saturated and the unsaturated pixels in the observed image g . We define g_{U} to be the unsaturated pixels of g , with the values of all saturated pixels set to 0. We also define $g_{\text{S}} = g - g_{\text{U}}$ to be a mask that is 1 for saturated pixels and 0 for unsaturated ones. Similarly, we define $f_{\text{S}} = f \cdot g_{\text{S}}$ and $f_{\text{U}} = f \cdot (1 - g_{\text{S}})$. Finally, we define r_{S} (r_{U}) as only that part of the residual from Equation 4, which is due to scattering of light from saturated (unsaturated) pixels.

With these definitions, we can rewrite the unsaturated component of Equation 4 as follows:

$$\begin{aligned} g_{\text{U}} &= \alpha (f_{\text{U}} + f_{\text{S}}) + \beta K * (f_{\text{U}} + f_{\text{S}}) + \gamma (r_{\text{U}} + r_{\text{S}}) \quad (5) \\ &= \alpha f_{\text{U}} + (\beta K * f_{\text{U}} + \gamma r_{\text{U}}) + (\beta K * f_{\text{S}} + \gamma r_{\text{S}}), \quad (6) \end{aligned}$$

since $f_{\text{S}} = 0$ for unsaturated pixels. As a result, we can now obtain the latent image by estimating and removing several kinds of glare:

- Glare generated by *unsaturated* pixels that affects other *unsaturated* pixels — first bracketed term of Equation 6. This type of glare is fairly weak and does not contain high spatial frequencies (Section 4.1). We can further simplify this term, since r_{U} is so small as to be negligible.
- Glare generated by *saturated* pixels that affects *unsaturated* pixels can be estimated and removed through the use of image priors (second bracketed term in the equation above, Section 4.2). The estimated glare also provides information about the saturated regions from which it emerges, and can therefore be used to reconstruct spatial detail within those regions (Section 4.3).

- Glare that contributes to already saturated pixels — either originating from unsaturated or saturated pixels — is not measured in the captured image and therefore does not need to be modeled.

While in essence we do perform a 2D deconvolution, to make the solution possible and robust, we decompose it into an ‘easy’ 2D deconvolution (a series of 1D problems) and finally a tomographic reconstruction. The supplemental material contains further discussion about the relationship to deconvolution.

4.1. Glare due to unsaturated pixels

Because the Dirac peak dominates the PSF of the cross-screen filters, the glare due to unsaturated pixels is very weak. As mentioned above, we can further simplify the situation by neglecting the shift-variant residual r_{U} , which is several orders of magnitude weaker than the shift-invariant part of the PSF. With these observations, we can remove the glare due to unsaturated pixels using a deconvolution approach similar to [23]:

$$g'(\mathbf{x}) = g(\mathbf{x}) - \beta (K * f_{\text{U}})(\mathbf{x}) \quad \text{for } \mathbf{x} \in \text{U} \quad (7)$$

$$= \left(\sum_{t=0}^{\infty} \left(-\frac{\beta}{\alpha} K \right)^t * g \right) (\mathbf{x}) \quad \text{from (6),} \quad (8)$$

where g' is the image with the unsaturated pixel glare removed, and the operator \cdot^t denotes t -times convolution.

4.2. Glare due to saturated pixels

The next step is to estimate and remove glare due to saturated pixels. This glare component will also be used for reconstructing saturated pixel values in Section 4.3. As mentioned in Section 3, we can factor this step into a number of 1D problems along directions \mathbf{u}_i , where $\mathbf{u}_i, \mathbf{v}_i$ form a coordinate frame aligned with the i^{th} glare ray (see Figure 5). In the following, we consider each glare direction separately, and thus omit the i subscript for notational convenience.

Image priors. Knowing both which pixels are saturated in the observed image, as well as the direction of the 1D glare rays, we can determine which image pixels exhibit a glare contribution from saturated pixels. In order to separate the latent image information from the glare in these pixels, we employ results from natural image statistics, specifically a sparse gradient prior [12, 18]. We model the distribution of gradients in the latent image using a Laplace distribution, which is the best approximation of the heavy-tailed distribution that still leads to a convex problem [2].

Glare rays cause the largest distortion of the image gradients in the direction *orthogonal* to the glare rays. According to the sparse gradient prior, we obtain $\frac{\partial f}{\partial v} \sim \text{Laplace}(0, b)$. Any deviations from a zero mean in the observed image g are attributed to glare. In the supplemental material we show that the Maximum Likelihood (ML) estimator for the

mean of a Laplace distribution is obtained by minimizing the L_1 norm. Therefore, we can solve for the intrinsic image as follows:

$$\arg \min_f \left\| \frac{\partial}{\partial v} (g' - \beta K * f_S - \gamma r_S) \right\|_1 + R, \quad (9)$$

where R gives constraints on r_S (see supplemental material, Section 2.2 for details):

$$R = \left(\lambda_1 \|r_S\|_2 + \lambda_2 \left\| \frac{\partial}{\partial v} r_S \right\|_1 + \lambda_3 \left\| \frac{\partial}{\partial u} r_S \right\|_1 \right). \quad (10)$$

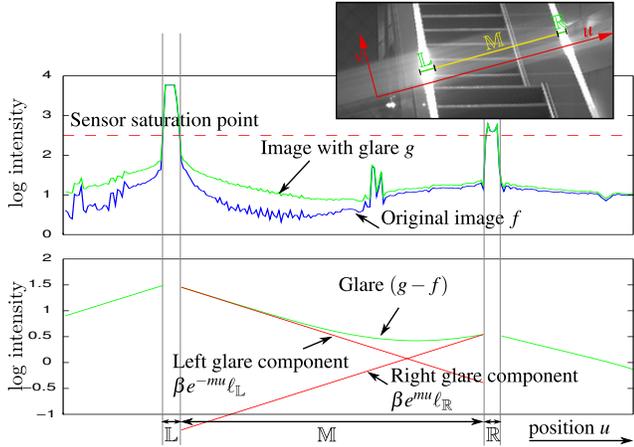


Figure 5: Cross-section of the glare due to saturated pixels, forming between a pair of saturated regions \mathbb{L} and \mathbb{R} . The cross-section is extracted from a single line of pixels along the glare direction (\mathbf{u} -axis in the inset). The glare can be split into two components with an exponential slope in the opposite directions.

Optimization. To actually apply the image prior in the glare estimation, we consider a single continuous segment \mathbb{M} of unsaturated pixels along a glare direction \mathbf{u} . \mathbb{M} is bounded by two sets of saturated pixels \mathbb{L} and \mathbb{R} on the left and on the right, as shown in the inset of Figure 5.

We can now use the exponential nature of the 1D glare streaks (from Equation 3), and expand the convolution operator in Equation 9:

$$(k * f_{S_v})(u, v) = \ell_{\mathbb{L}} e^{-mu} + \ell_{\mathbb{R}} e^{mu} \quad \text{for } u \in \mathbb{M}, \quad (11)$$

where

$$\ell_{\mathbb{L}} = \sum_{i \in \mathbb{L}} e^{mi} f(i, v), \quad \ell_{\mathbb{R}} = \sum_{i \in \mathbb{R}} e^{-mi} f(i, v). \quad (12)$$

Note that $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$ have the same value for all unsaturated pixels $u \in \mathbb{M}$, and therefore all pixels in \mathbb{M} can be used to robustly estimate these two quantities.

Also note that $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$ represent the amount of energy present in the glare from the saturated pixels to the left and to the right of \mathbb{M} . These quantities, which we refer to as

line integrals will be useful for reconstructing detail in the saturated regions in Section 4.3.

Now we can reformulate the glare estimator in terms of line integrals $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$ rather than saturated pixel values. From Equations 9, 11 and 12 we obtain

$$\arg \min_{\ell_{\mathbb{L}}, \ell_{\mathbb{R}}, r} \sum_{u \in \mathbb{M}} \left\| \frac{\partial}{\partial v} g'(u) - \beta \frac{\partial}{\partial v} (\ell_{\mathbb{L}} e^{-mu} + \ell_{\mathbb{R}} e^{mu}) - \gamma \frac{\partial}{\partial v} r_S(u) \right\|_1 + R. \quad (13)$$

This equation allows us to efficiently optimize on each segment \mathbb{M} independently. However, to solve for $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$, the partial derivatives $\frac{\partial \ell_{\mathbb{L}}}{\partial v}$ and $\frac{\partial \ell_{\mathbb{R}}}{\partial v}$ must be found for all segments and then integrated. To solve the minimization problem efficiently, we use several EM iterations. We initially set $\gamma \frac{\partial r}{\partial v} = 0$. Since $\gamma \ll \beta$, this provides a reasonable initial estimate of exponential glare component, but enhances color artifacts when this monochromatic glare is removed. In the E-step, we solve for $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$, and in the M-step we refine the estimate of r . Minimizing Equation 13 is sufficient to remove most of the glare (Figure 6).

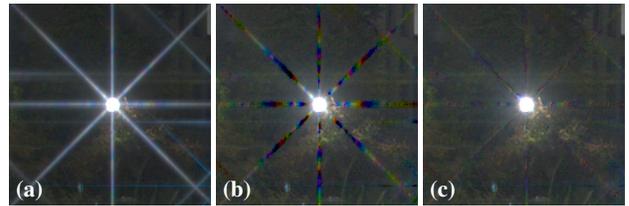


Figure 6: The glare left by a cross-screen filter is not monochromatic due to diffraction and dispersion effects. (a) Although the color artifacts seem to be very faint in captured images, (b) they are strongly enhanced after removing achromatic exponential glare because it boosts chromatic contrast. (c) Estimating wavelength-dependent variations can remove most of the color artifacts.

Finally, we prepare line integral estimates for the energy contributed by individual, *continuous* regions of saturated pixels, which will be used in the next section. Each value $\ell_{\mathbb{L}}$ and $\ell_{\mathbb{R}}$ can contain contributions from multiple saturated segments on the left and right of \mathbb{M} (not shown in Figure 5). However, isolating glare due to each saturated region is trivial since there are exactly as many line integrals as there are regions \mathbb{M} along a glare line, and therefore the contributions for each region can be found with a simple linear system. For convenience, we shift the origin of (u, v) to the leftmost or rightmost pixel of each segment \mathbb{M} to get isolated line integrals $\hat{\ell}_{\mathbb{L}}$ and $\hat{\ell}_{\mathbb{R}}$.

4.3. Reconstruction of saturated pixels

So far, we have decoded the values of the intrinsic image f for the previously unsaturated pixels only; the values of

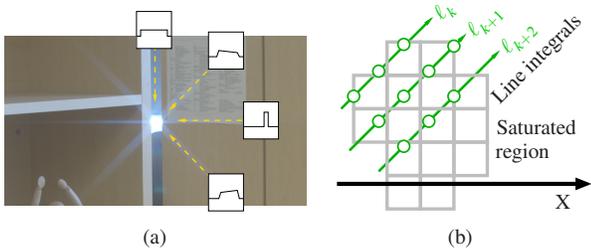


Figure 7: (a) Glare along discrete directions give different “projections” of the saturated region. (b) Bilinear sampling along these directions relates line integrals to saturated pixels in (14) and (15).

the saturated pixels are still unknown. However, glare removal procedure from Section 4.2 also yields line integrals along p discrete directions, as shown in Figure 7(a). In the final step of the decoding procedure, we use this information to reconstruct the saturated region. To this end, we need to find saturated pixel values that can produce the line integrals matching the observations. This requires solving a standard tomographic reconstruction problem [10].

Unlike the glare estimation, the tomographic reconstruction is inherently a 2D problem. We gather the estimated line integrals along all p directions in a linear system that describes the relationship between line integrals and saturated pixels f . We therefore use a one-index representation for all line integrals contributing to a given region: $\hat{\ell}_i$. This relationship is then expressed as

$$\hat{\ell}_i = \sum_j w_{ij} f_j, \quad (14)$$

where the weight term w_{ij} for line integral i and an unknown pixel j is the product of exponential falloff and a bilinear resampling weight a_{ij} , as shown in Figure 7(b):

$$w_{ij} = a_{ij} e^{-m|u_i - u_j|}. \quad (15)$$

Here, u_i is the reference location used while computing $\hat{\ell}_i$. The absolute value consolidates different signs for glare falloffs to the left and right.

We solve this tomography problem using Simultaneous Iterative Reconstruction [10, pp 284]. We start with an initial guess, $f^{(0)} = 0$. Then, in each iteration t , the residual error in the current estimate of line integrals,

$$\Delta \hat{\ell}_i = \hat{\ell}_i - \sum_j w_{ij} f_j^{(t)}, \quad (16)$$

is backprojected over the participating unknown pixels regardless of distance from the reference location, *i.e.*, energy distribution is proportional to resampling weight (a) only,

$$f_j^{(t+1)} = f_j^{(t)} - \Delta \hat{\ell}_i \frac{a_{ij}}{\sum_k a_{ik}}. \quad (17)$$

Using a uniform distribution for the backprojected residual, independent of any falloff, is a standard procedure in tomography. One should think of this as a (weak) prior on the intensity distribution within the unknown region. We employ a simple two-scale approach which solves the problem for a low resolution image first. Since we know that actual values at the saturated pixels are larger than the saturation threshold for the camera, we enforce this simple constraint during backprojection.

5. Results

Figure 8 shows a number of examples of HDR images, decoded from single images captured as RAW images with a Canon 40D DSLR camera using 8- and 16-point cross-screen filters, and Canon lenses ranging from 50mm to 100mm. In this figure, the first two columns represent two exposures of the 12-bit input image, while the right two columns represent two virtual exposures of our reconstructions. Saturated regions are reconstructed, and glare produced by the filter is removed. For color images, we run our algorithm separately and independently on each color channel. Radial lens distortion was removed in a preprocessing step. Insets in the right column show ground-truth comparisons for some of the results, *i.e.* short exposure images taken without the filter, using the same camera and lens. Note that the geometric and photometric alignment may not be perfect due to the changes in the acquisition setup. These results demonstrate a number of points:

Glare estimation: Accurate estimation of glare is necessary not only to correctly reconstruct saturated regions, but also to remove glare. Our sparse-gradient prior was robust enough to estimate glare both for a multitude of small light sources (Figure 8a), as well as relatively large saturated areas (Figure 8c). The main requirement for successful glare estimation is that saturated regions be both bright and large enough (*i.e.* sufficient cumulative energy) to produce glare above the camera noise level.

Highlight reconstruction: Given only 8–16 directional line integrals, tomographic reconstruction is a challenging task. Even so, the results demonstrate that our method can estimate the total energy of the saturated regions as well as the approximate values of the saturated pixels. This is in contrast to the previous single-image methods, which could achieve neither of these two goals. Our method can also easily distinguish between very bright light sources and diffuse surfaces that are just above the clipping level, thus making complicated classification methods for the LDR-to-HDR enhancement unnecessary [5].

Figure 8(a) also demonstrates that the multi-exposure HDR can exhibit some artifacts due to alignment issues, particularly at the outline of the light sources. Ours being a single exposure method, does not show any such artifacts.

5.1. Limitations

Our method is not suitable for scenes with large saturated regions, such as a sky, because large saturated regions do not leave enough unsaturated pixels to register glare patterns, and the gradient distortions too hard to detect. The method can conceptually handle scenes with light sources outside the image frame, but we found that the accuracy of glare estimation is often not sufficient in such cases.

Finally, our method is also likely to fail if a scene contains color gradients oriented the same way as the glare patterns. It breaks the assumption of a zero-mean gradient distribution, and results in mis-estimates of the glare. It is usually possible to avoid such problems by rotating the filter out of alignment with image gradients. A synthetic example of this scenario is analyzed in the supplemental material.

6. Conclusion and future work

The distinctive feature of our proposed single-image HDR capture method is that the information lost in clipped pixels is encoded in the remaining portions of an image. This approach is very different from existing HDR capture methods, which attempt to register HDR information within each pixel or a group of closely located pixels. Unlike the LDR to HDR methods that only enhance clipped pixels, the proposed method can restore a close approximation of their original values. Our method does all that without requiring specialized sensor or invasive camera modifications, as it needs only a cross-screen filter mounted on top of a lens.

Our reconstruction method contains several technical contributions, including the use of natural image priors to separate encoded information (glare) from image content. We also propose a novel application of tomographic reconstruction.

In the future, we would like to design cross-screen filters that produce easily detectable patterns benefitting from redundancy and sparsity of information in natural images. A promising application of our method is HDR video capture, which could be further improved by combining information from several frames to better reconstruct clipped pixels.

Acknowledgments

This work was supported by Dolby under the Dolby Research Chair in Computer Science at UBC.

References

- [1] J. Abel and J. Smith III. Restoring a clipped signal. In *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing, 1991*, pages 1745–1748, 1991. 290
- [2] U. Ascher, E. Haber, and H. Huang. On effective methods for implicit piecewise smooth surface recovery. *SIAM Journal on Scientific Computing*, 28(1):339, 2006. 292
- [3] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers. Inverse tone mapping. In *Proc. GRAPHITE '06*, pages 349–356, 2006. 290
- [4] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH '97*, pages 369–378, 1997. 290
- [5] P. Didyk, R. Mantiuk, M. Hein, and H. Seidel. Enhancement of bright video features for HDR displays. *Computer Graphics Forum*, 27(4):1265–1274, 2008. 290, 294
- [6] A. El Gamal. High dynamic range image sensors. In *Tutorial at International Solid-State Circuits Conference*, 2002. 290
- [7] K. Faulkner, C. Kotre, and M. Louka. Veiling glare deconvolution of images produced by x-ray image intensifiers. In *Conf. on Image Processing and its Applications*, pages 669–673, 1989. 290
- [8] A. Foi. Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, 89(12):2609–2629, 2009. 290
- [9] O. Gallo, N. Gelfand, W. Chen, M. Tico, and K. Pulli. Artifact-free high dynamic range imaging. In *Proc. ICCP*, 2009. 290
- [10] A. Kak and M. Slaney. *Principles of computerized tomographic imaging*. SIAM, 2001. 294
- [11] F. Koreban and Y. Schechner. Geometry by deflaring. In *Proc. ICCP*, 2009. 290
- [12] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.*, 26(3):70, 2007. 290, 292
- [13] A. Levin, P. Sand, T. Cho, F. Durand, and W. Freeman. Motion-invariant photography. *ACM Trans. Graph.*, 27(3):71, 2008. 290
- [14] L. Meylan, S. Daly, and S. Süsstrunk. The reproduction of specular highlights on high dynamic range displays. In *IS&T/SID 14th Color Imaging Conference*, 2006. 290
- [15] S. Nayar and T. Mitsunaga. High dynamic range imaging: Spatially varying pixel exposures. In *Proc. CVPR*, pages 472–479, 2000. 290
- [16] S. Nayar and S. Narasimhan. Assorted pixels: Multi-sampled imaging with structural models. In *ACM SIGGRAPH'05 Courses*, 2005. 290
- [17] T. Olofsson. Deconvolution and model-based restoration of clipped ultrasonic signals. *IEEE Trans. Instrumentation and Measurement*, 54(3):1235–1240, 2005. 290
- [18] B. Olshausen and D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996. 292
- [19] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: Motion deblurring using fluttered shutter. *ACM Trans. Graph.*, 25(3):795–804, 2006. 290
- [20] R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan. Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses. *ACM Trans. Graph.*, 27(3), 2008. 290
- [21] A. Rempel, M. Trentacoste, H. Seetzen, H. Young, W. Heidrich, L. Whitehead, and G. Ward. Ldr2Hdr: on-the-fly reverse tone mapping of legacy video and photographs. *ACM Trans. Graph.*, 26(3):39, 2007. 290

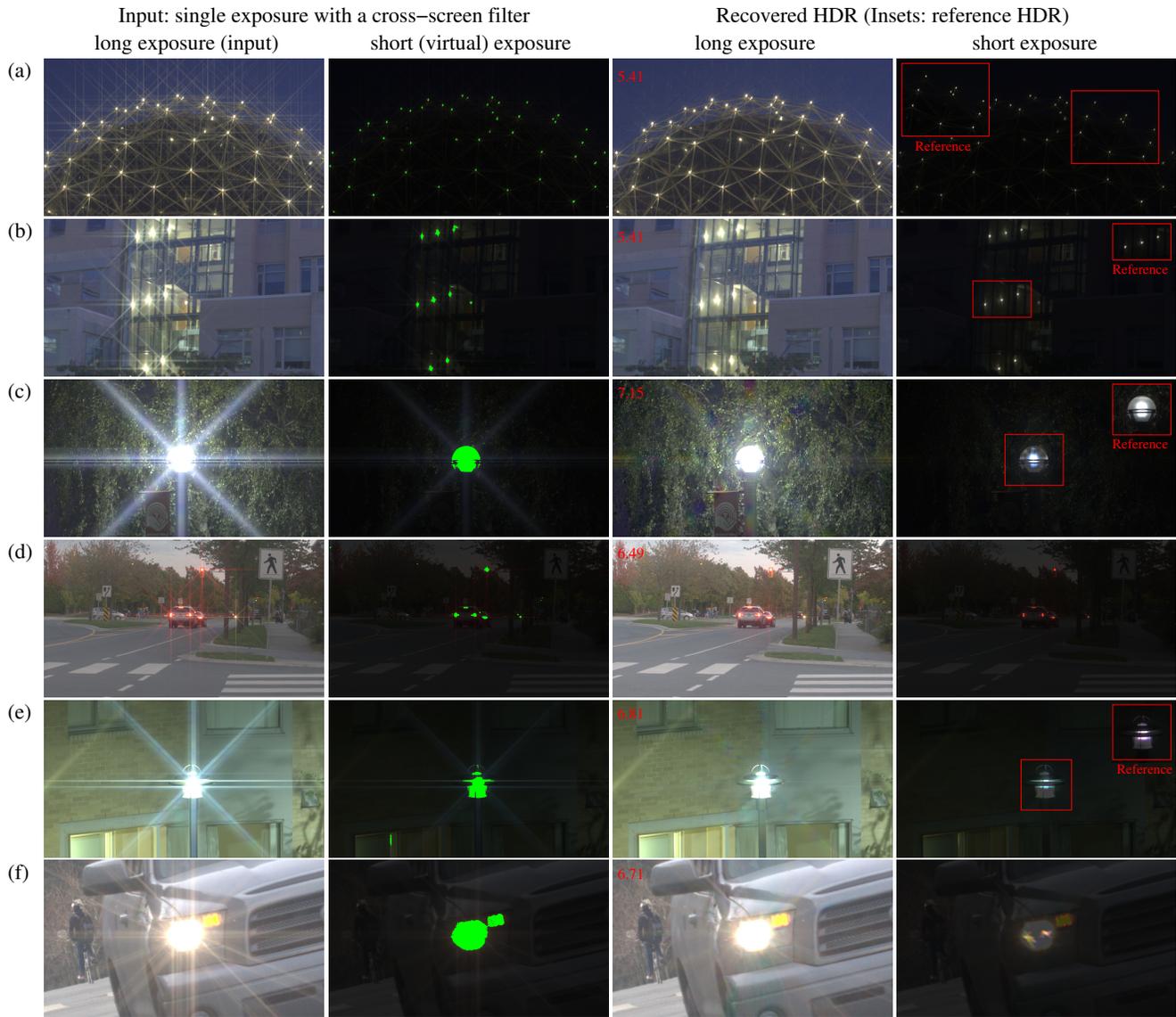


Figure 8: Results of reconstruction from images captured by a camera with an 8-point cross-screen filter. The first column shows the single-exposure images captured with the filter (input to the proposed algorithm). The second column shows the saturated pixels by marking them in green and artificially shortening the exposure. The third column shows a long virtual exposure of the recovered HDR image. Note that most of the glare present in the first column was removed. The numbers in the top-left corner indicate the dynamic range increase in f-stops. The last column shows the short exposure of the same reconstructed image. The insets show the reference image captured with the multi-exposure method and the corresponding regions are marked with a red frame. We were not able to capture multi-exposure reference images for rows (d) and (f) because of the moving objects.

- [22] J. A. Seibert, O. Nalcioglu, and W. Roeck. Removal of image intensifier veiling glare by mathematical deconvolution techniques. *Medical Physics*, 12(3):281–288, 1985. 290
- [23] E. Talvala, A. Adams, M. Horowitz, and M. Levoy. Veiling glare in high dynamic range imaging. *ACM Trans. Graph.*, 26(3):37, 2007. 290, 292
- [24] M. Trentacoste, C. Lau, M. Rouf, R. Mantiuk, and W. Heidrich. Defocus techniques for camera dynamic range expansion. In *SPIE Electronic Imaging*, 2010. 290
- [25] A. Yoshida, M. Ihrke, R. Mantiuk, and H. Seidel. Brightness of the glare illusion. In *Proc. APGV*, pages 83–90, 2008. 289
- [26] X. Zhang and D. Brainard. Estimation of saturated pixel values in digital color imaging. *JOSA A*, 21(12):2301–2310, 2004. 290
- [27] C. Zhou and S. K. Nayar. What are good apertures for defocus deblurring? In *Proc. ICCP*, 2009. 290