HDR Image Noise Estimation for Denoising Tone Mapped Images

Miguel Granados* MPI for Informatics
Tunç Ozan Aydın Disney Research
J. Rafael Tena Disney Research
Jean-François Lalonde Laval University
Christian Theobalt MPI for Informatics

ABSTRACT

Tone mapping operators are designed to compress the dynamic range of high dynamic range (HDR) images while preserving the perceived image brightness, but they often enhance image noise in the process, specially in low-light conditions. We propose a method for reducing noise in images created by any tone mapping operator. Our approach leverages the noise distribution of the HDR image to guide the range kernel of a cross bilateral filter that is used to denoise the tone mapped image. When the noise distribution is unknown, we use a new method to automatically estimate it assuming that the HDR image was produced as an average of multiple exposures taken in RAW or JPEG compressed format. Our method performs quantitatively better than existing denoising methods applied on either the original HDR or the tone-mapped images directly, and a user study confirms that it produces visually preferable results.

Keywords

High Dynamic Range Imaging, Noise Estimation, Bilateral Filter

1. INTRODUCTION

High Dynamic Range (HDR) images need to be tone mapped for visualization on low dynamic range displays. Tone mapping operators (TMOs) reduce the dynamic range while trying to preserve the original contrast, in particular true color gradients [26].

Any existing TMO is challenged by the presence of noise in the HDR images, as it is hard to distinguish real light variations from gradients caused by *camera noise* (a detailed description of camera noise is given in Sec. 2). While this is less of a concern for HDR images taken with sufficient light, it becomes critical for images taken under low light conditions. For example, outdoor shots taken at dawn with poorly illuminated backgrounds, and indoor shots where large parts of a scene are scarcely lit, often exhibit significant noise. In such cases, existing TMOs create objectionable artifacts by amplifying the noise (see fig. 3-(a)).

*Miguel Granados is also with Disney Research.

To appear in the 12th International Conference on Visual Media Production (CVMP 2015), 24-25 November, 2015, London, UK
Author's manuscript copy

Let us assume for a moment that the noise distribution in the HDR image is known. Given this distribution, the first option to try would be to apply a denoising algorithm [23] directly on the HDR image, prior to tone mapping. However, any noise left after this process could be amplified by the TMO. The user may thus frequently need to alternate between denoising with more aggressive parameters and tone mapping before obtaining a satisfactory result.

The second, less trivial, option to try would be to apply the tone mapping transformation to the original (noisy) HDR image, recover the noise distribution in the resulting image, and give this information to a denoising algorithm applied on the tone mapped image. Unfortunately, this is not feasible because the tone mapping transformation is highly non-linear and depends upon the particular TMO employed and its parameters. In addition, standard denoising algorithms could fail to remove noise since the assumption of additive Gaussian noise (even signal-dependent Gaussian or Poisson noise) is not satisfied for images that have been tone-mapped.

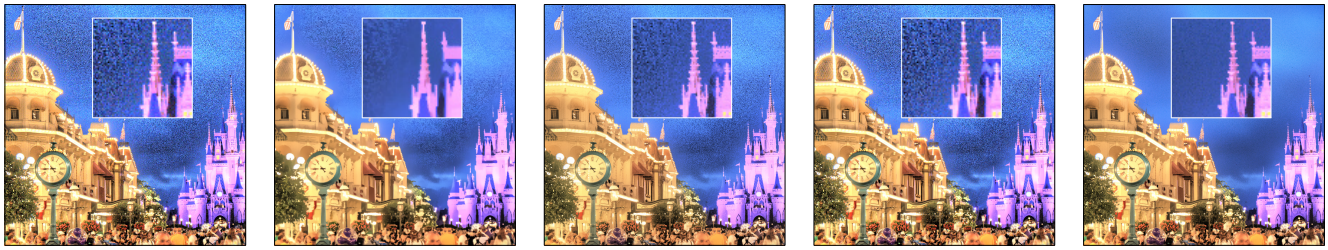
Our first key contribution is to introduce a novel method which uses knowledge of the HDR noise distribution to remove noise in the tone mapped image, *without knowledge of the tone mapping transformation*. In fact, our method is completely agnostic to the TMO algorithm itself. We propose a new "HDR cross bilateral filter", which assumes that the original HDR image noise distribution is known (sec. 3).

To estimate the noise distribution in HDR images, one typically needs to calibrate the camera directly [12], or have access to the RAW exposure stack from which the HDR image was created [14]. Our second key contribution is to estimate the noise distribution from the HDR image alone, *without requiring the original exposures*. Our method takes into account the camera noise model to predict the noise distributions of HDR images reconstructed from RAW or JPEG exposure stacks (sec. 4).

Our approach can be easily used by non-professional users and it is applicable for HDR images reconstructed from maximum-likelihood averages of RAW or JPEG exposure stacks. We demonstrate the high performance of our method by quantitative and qualitative comparisons to state-of-the-art filtering methods (sec. 5).

2. RELATED WORK

The tone mapping problem has been thoroughly investigated by the computer graphics community. An excellent summary of the numerous tone mapping techniques can be found in [26]. Modeling common photographic practices of film development in digital



(a) Tone mapped image (b) Non-local means (c) Bilateral filtering (d) BM3D (e) Our approach

Figure 1: Existing tone mapping operators often over-emphasize the noise of HDR images, specially for low-light scenes (a). Image denoising methods, such as (b) non-local means [11], (c) bilateral filter [28], and (d) BM3D [5] are not effective due to the non-linear noise distribution in tone mapped images. Our method predicts the noise of the original HDR image to effectively remove it (e).

media has been shown to be an effective tone mapping strategy. A popular example in that direction is the photographic tone mapping operator [25] that utilizes a global sigmoid-shaped tone curve in combination with local dodging and burning operators. Another popular approach has been *local* tone mapping where the HDR image is decomposed into base and detail layers using shift variant filtering, such as the bilateral filter [6]. Dynamic range is reduced through tonal compression of only the base layer, without modifying the detail layer. This way one can achieve significantly higher local contrast compared to *global* tone mapping. Each tone mapping type produces images with a distinct visual style, and thus using a specific operator is an artistic choice. Previous work on tone mapping has also focused in understanding and utilizing how the properties of the human visual system change for extremely low luminance levels [9, 24, 16, 18]. These methods mainly focus on faithfully modeling the loss of color and luminance contrast acuity in dark scenes, but do not account for camera noise.

Camera noise is a well studied topic in the fields of optics and photonics (c.f. [17] for an in-depth analysis of noise in CCD/CMOS sensors). Noise factors can be classified into two categories: spatially-dependent noise and spatially-independent noise. The first category includes factors that depend on the particular pixel such as its surface area (inducing *photo-response non-uniformity* noise) and the level of heat-dependent distortions (inducing *dark current non-uniformity* noise). The second category includes factors that affect every pixel equally such as *shot noise* (associated to the Poisson process of light emission) and *readout noise* (an umbrella term for several sources that affect the acquisition process). In this paper, we apply a simplified noise model (see sec. 4.1) that takes into account the spatially-independent sources (shot and readout noise) and ignores other spatially dependent sources. The parameters of the simplified model (i.e. a gain factor and the variance of the readout noise distribution) can be either recovered from a set of calibration images [17] or directly from photographs of a given scene [20, 10, 14]. In sec. 4.2 we propose a method to recover the parameters from HDR images directly, which we will then use to guide the filtering of noisy tone mapped images.

Image denoising has been a very active field in recent years (see [23] for a recent overview). However, although this is considered a very mature field [4], it has been mainly targeted at standard images (i.e. *low dynamic range* (LDR) images). Existing state-of-the-art denoising strategies include non-local means filtering [2] in combination with efficient high dimensional filters [11], and strategies based on collaborative filtering [5]. In contrast to existing work on LDR images, we take advantage of noise models calibrated specifically for HDR images to improve the denoising quality in tone mapped images. This strategy significantly improves on the naive application of the state-of-the-art denoising methods to HDR and tone mapped images. Similar noise models have been applied in [13] to the selection of suitable tone mapping parameters for noisy HDR images.

3. HDR CROSS BILATERAL FILTERING

To function properly, a denoising algorithm needs an estimate of the noise distribution of the image. However, such a distribution is difficult to recover in a tone mapped image since it was created by transforming an HDR image using (potentially) highly non-linear processes. To address this difficulty, we observe that in the original HDR domain, noise is linearly proportional to the light in the scene, so its distribution is easier to model. We then propose to use the pixel noise distribution of the original HDR image, and *transfer* it to the tone mapped image by using a cross (or joint) bilateral filtering approach [7]. Let us briefly summarize the general concept of this filter, and see how it can be adapted to our setting.

Given a color image I, standard bilateral filtering [28] replaces the color I_p of a pixel p by a weighted average of the color of the surrounding pixels q. A domain kernel f gives higher importance to closer pixels (e.g. a Gaussian kernel with standard deviation σ_f), and a range kernel g gives higher importance to color values closer to I_p (e.g. Gaussian with standard deviation σ_g). The filtered image J_p is given by

$$J_p = \frac{\sum_{q\in\Omega} W(p,q) I_q}{\sum_{q\in\Omega} W(p,q)}, \text{ with weights} \quad (1)$$

$$W(p,q) = f(\|p-q\|;\sigma_f)\, g(\|I_p - I_q\|;\sigma_g), \quad (2)$$

where Ω is the image domain. In cross bilateral filtering [7] the range kernel is guided using a different image $\tilde{I}$.

We remove noisy gradients from a tone mapped image using cross bilateral filtering with a range kernel guided by the noise distribution of the original HDR image H. Let H_p be the irradiance at pixel p in the HDR image. The noise distribution at every pixel is assumed to be a Gaussian with mean H_p and standard deviation σ_{H_p} [27]; in sec. 4 we provide a method for estimating the σ_H automatically. Using this distribution we compute the z-scores for each pixel q as $z_q = |\frac{H_p - H_q}{\sigma_{H_p}}|$, and produce the range kernel $g(z_q)$, where g is a Gaussian kernel with unitary standard deviation. To prevent color bleeding in color images, for every averaged pixel q we take

the minimum score z_q among the color channels.[1] The final filtering is computed by replacing the weights in (1) by

$$W(p,q) = f(\|p-q\|;\sigma_f)g\left(\min_{c\in\{R,G,B\}}\left\{\left|\frac{H_p^c - H_q^c}{k\sigma_{H_p}^c}\right|\right\}\right), \quad (3)$$

where parameter k is a user-defined parameter which controls the desired level of smoothing. Our HDR cross bilateral filter correctly averages color values that are dissimilar in the tone mapped image but which correspond to similar irradiance values in the HDR image (see fig. 3-(d), insets), while preserving even faint edges in the tone mapped result that are present in the HDR image.

4. AUTOMATIC ESTIMATION OF HDR IMAGE NOISE

Our approach requires an estimate of noise distribution of the original HDR image. Methods for estimating the noise from single (LDR) image exist, for instance, by segmenting a RAW image into level sets to obtain an intensity mean-variance scatterplot from which the parameters are estimated[10]. Using a similar scatterplot, Liu et al. [20] infer the noise distribution of JPEG images by fitting the parameters of an image formation model for CCD cameras, and then predicting the noise level of the image values. Similarly, Heide et al. [15] model each step of the image formation pipeline such as demosaicking, denoising, and deconvolution to reconstruct sharp images from noisy input. However, in order to achieve a similar effect in tone mapped images, one would need to include the tone mapping pipeline in the model for every existing tone mapper. This is highly impractical, given the large number of existing TMOs, and the number of parameters might become too large to obtain a reliable estimate from a single image. Instead, we advocate for performing the noise estimation in the original HDR image (or the corresponding exposure stack) where a general image formation model can be used.

There are methods to predict the noise distribution of HDR images under some conditions, for instance when averaging exposure stacks captured with a pre-calibrated camera [12], or by calibrating the camera from the exposure stack of the HDR image [14]. However, to the best of our knowledge, there are no methods to estimate the noise distribution of HDR images when the original image stack is not available. We thus introduce a novel, practical method that handles the general case when only the tone mapped image and the source HDR image are given.

To constrain the noise prediction, we assume that: *i)* the HDR image was reconstructed as a weighted average of multiple photographs [21], *ii)* a maximum likelihood weighting scheme [27] was employed (a reasonable assumption since most HDR images are created this way), and *iii)* every well-exposed pixel contributes to the average (e.g., no motion compensation). The noise prediction performance will depend on how much the input HDR satisfies these assumptions. Similar to [20, 14], we perform a super-pixel tessellation of the HDR image into regions with as-uniform-as-possible irradiance. The tesselation is obtained via an energy-minimization framework to ensure super-pixels are compact, and do not cross image boundaries [29]. The irradiance mean and variance of these regions define a scatter-plot (fig. 2). The lower envelope of this plot is a tight upper bound for the variance parameter of the irradiance distribution [20] because any color variation within a region with uniform irradiance can only be attributed to camera noise. However, this upper bound might overestimate the variance parameter due to a lack of uniform super-pixels for every irradiance range. Therefore, we use the upper bound only as guidance for fitting a noise model that assumes that the HDR image was constructed as the maximum likelihood irradiance estimate from multiple photographs. The model and the fitting process are described next.

[1] Ideally, color distances should be measured in a perceptually uniform color space [28], but the transformation of the corresponding image noise distribution to such a space is not trival, and it does not necessarily correlate with noise perception.

4.1 Noise model for HDR images

Assuming maximum-likelihood weighting when averaging the n photographs (*ii*), our noise model assumes that the irradiance measurement at pixel p given by the HDR image is sampled from a Gaussian distrubtion [27]. The parameters of the distribution, i.e. the expected irradiance X_p and the variance $\sigma^2(x_p)$ for each p, can be estimated as

$$X_p = \frac{\sum_{i=1}^n w_p^{(i)} \frac{X_p^{(i)}}{\sigma^2\left(x_p^{(i)}\right)}}{\sum_{i=1}^n w_p^{(i)} \frac{1}{\sigma^2\left(x_p^{(i)}\right)}}, \text{ and} \quad (4)$$

$$\sigma^2(x_p) = \frac{1}{\sum_{i=1}^n w_p^{(i)} \frac{1}{\sigma^2\left(x_p^{(i)}\right)}}, \quad (5)$$

where $X_p^{(i)}, \sigma^2\left(x_p^{(i)}\right)$ are the parameter estimates obtained from each input photograph with index i. The weight $w_p^{(i)} = [X_p < X_{\max}^{(i)}]$ excludes saturated values from the average, where the value $X_{\max}^{(i)}$ represent the maximum irradiance measurable on each photograph i before saturation occurs.

Following the camera noise model in [14], $\sigma^2\left(x_p^{(i)}\right)$ of irradiance measurement $X_p^{(i)}$ at pixel p is approximated by

$$\sigma^2\left(x_p^{(i)}\right) \approx \frac{X_p^{(i)} + \frac{C^{(i)}}{t^{(i)}}}{t^{(i)}}, \quad (6)$$

where $t^{(i)}$ is the exposure time, and $C^{(i)} = \frac{\sigma_R^{2(i)}}{(g^{(i)})^2}$ is a constant that depends on the camera gain $g^{(i)}$ and readout noise $\sigma_R^{2(i)}$ parameters (these parameters vary between photographs depending on the ISO setting).

Therefore, in order to estimate the variance parameter $\sigma^2(X)$ for each pixel in a given HDR image X, we need to estimate $3n+1$ parameters: $t^{(i)}$, the saturation point $X_{\max}^{(i)}$, the constant $C^{(i)}$ of each photograph used to reconstruct X, and the number of photographs n. Our *model-based self-calibration* procedure to estimate these parameters is described next.

4.2 Estimation of the Noise Model Parameters

The *model-based self-calibration* proceeds in four steps:

1) Super-pixel tesselation of the HDR image The HDR image is tessellated into *super-pixels*, i.e. into a subsets $\{S_j\}$ of spatially adjacent pixels, where each subset $\{S_j\}$ has as low color variation as possible [29]. For each super-pixel $\{S_j\}$, the irradiance sample mean and sample variance $(\bar{S}_j, \sigma_j^2)$ of the corresponding pixels are computed independently for each color channel (fig. 2, blue dots).

2) Selection of super-pixels with low irradiance variation The irradiance range x is divided into n_r intervals of equal length in log-scale ($n_r = 100$ in practice). For each interval x_r, the set of super-pixels whose mean falls within the interval are selected, and the one with minimum variance among them is singled out. This yields the set $U = \{(\bar{S}_r, \sigma_r^2)\}_{r=1}^{n_r}$, where each element $(\bar{S}_r, \sigma_r^2)$ is the super-pixel of minimal spatial irradiance variance for the corresponding interval. The sample variance of each super-pixel in U provides an upper bound for the variance parameter of the irradiance distribution in the corresponding interval.

3) Estimation of upper bound of the variance parameter From the set U, a piece-wise-linear function $\sigma_{\text{NP}}^2 : x \to \mathbf{R}$ is constructed to approximate the points in U using constrained least-squares linear regression. In the following, we refer to σ_{NP}^2 as the *non-parametric self-calibrated variance* (fig. 2, cyan curve). Given the linear relation between irradiance and its variance (eq. 6), the curve is constrained to be positive and monotonically-increasing.

4) Estimation of noise parameters as energy minimization Since σ_{NP}^2 might over-estimate the variance parameter of the irradiance distribution, this curve cannot be used directly to guide the filtering of tone mapped images derived from X. Nevertheless, this upper bound can be used to guide the estimation of the parameters of the noise model for HDR images (sec. 4.1). Taking σ_{NP}^2 as a noisy observation of the noise model, the parameters $P_n = \{t^{(i)}, X_{\max}^{(i)}, C^{(i)}\}_{i=1}^{n}$ are estimated as the minimum of an energy function that penalizes deviations from the noise model on each interval. This energy function is defined as

$$\mathcal{E}(P_n) = \sum_{r=1}^{n_r} \left(\frac{\log \sigma_{\text{NP}}^2(x_r) - \log \sigma_{P_n}^2(x_r)}{\log \sigma_{\text{NP}}^2(x_r)} \right)^2, \qquad (7)$$

where $\sigma_{P_n}^2(X)$ is the estimate for the variance parameter obtained by replacing the parameters P_n in (5) and (6). In the following, we refer to $\sigma_{P_n}^2$ as the *model-based self-calibrated variance*. Logarithms are used in eq. 7 to better deal with the non-linear relationship between the parameters and the variance (i.e. $\log \sigma^{2(x^i)} \approx \log[X + \frac{C^{(i)}}{t^{(i)}}] + \log t^{(i)}$ is linear in $\log t^{(i)}$ as $C \to 0$). The denominator serves as a normalization factor to dampen the effect of large over-estimates in σ_{NP}^2 (e.g. at upper irradiance levels in Fig. 2-b,c).

We minimize the energy in (7) following an iterative approach. Starting with $n = 1$, $C^{(1)} = 0$, and $X_{\max}^{(1)} = \infty$, we jointly estimate $t^{(1)}$ and $C^{(1)}$ using the Quasi-Newton method ($t^{(i)}, C^{(i)} > 0$). Then, we add a second photograph ($n = 2$) and select the saturation point $X_{\max}^{(2)} \in \{x_r\}_{r=1}^{n_r}$ that minimizes the energy (7); the exposure times and the constant are re-estimated for each candidate x_r. We repeat the same procedure for each additional photograph until the target number of photographs n is reached, and select the parameter configuration with the minimum energy. The number of photographs n is considered as a hyper-parameter. It was set to $n = 3$ in our experiments, however, it would be possible to let the algorithm automatically select a value for n, e.g. by adding an additional error term to penalize photographs with similar parameters [1].

Discussion Fig. 2 illustrates the result of the self-calibration process on three images of different types: 1) the simulated HDR image of fig. 4; 2) the HDR image reconstructed from RAW photographs from fig. 1; and 3) the HDR reconstructed from JPEG photographs from fig. 3. The plots corresponding to RAW and JPEG images indicate that the rate of increase of the irradiance variance parameter with respect to irradiance mean parameter is well approximated, but the absolute values of the variance are lower than the grounth truth (e.g. by a factor k, applied in Eq. 3). Two reasons explain this behavior. First, the model does not take into account the distribution of the sample variance for super-pixels. Therefore, the resulting upper bound σ_{NP}^2 is lower than the expected value by an amount proportional to the expected sample variance. Second, in the case of JPEG images, the variance of super-pixel irradiances is drastically reduced by in-camera processing such as denoising, and luminance and color compression.

5. EXPERIMENTAL VALIDATION

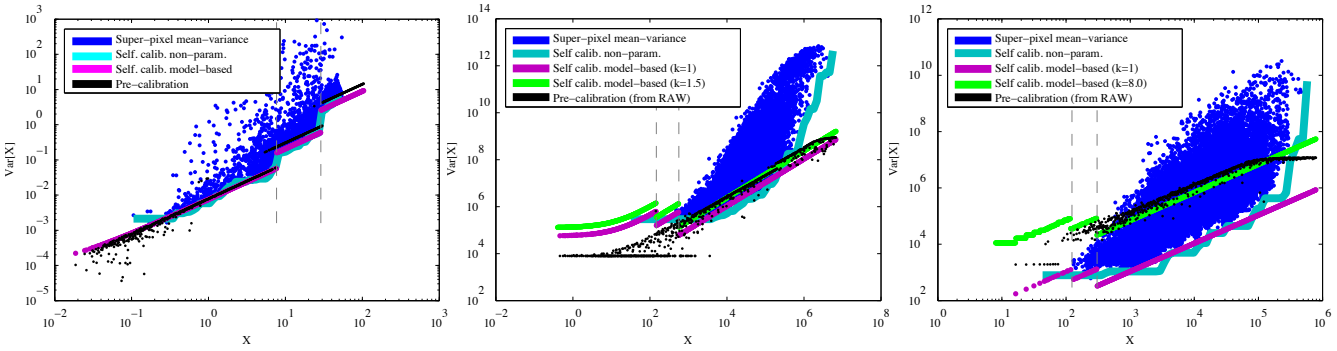
In this section, we experimentally show that our HDR cross bilateral filter improves the quality of tone mapped images, irrespective of the TMO employed. First, we performed an empirical evaluation of the denoising quality (sec. 5.1); second, we conducted a quantitative evaluation of the correlation between tone mapped images obtained by applying different denoising strategies and the original HDR image (sec. 5.2); and third, a qualitative validation through a user study (available in the supplementary material[2]). These evaluations indicate that our approach produces high quality results outperforming existing alternatives, and that the HDR noise distribution estimate can be used during filtering to improve the quality of tone mapped images.

5.1 Empirical evaluation

We captured eight night scenes using a Canon EOS 5D Mark III, four examples are shown in figs. 1 and 3 (all eight photos are available in the supplementary material). Each scene was captured as a bracketed-exposure sequence (-3, 0, +3EV) in aperture priority mode and automatic ISO setting. Note that our method handles the variation on ISO settings by estimating a gain factor per exposure (eq. 6). The photographs were saved both in RAW and JPEG format. We reconstructed two HDR images per scene (one from RAW and one from JPEG) using the method of [19] with pre-calibrated camera parameters [12]. For JPEG sequences, the inverse camera response was estimated from a RAW-JPEG image pair: first, we applied black level subtraction and white balancing to the RAW image (using the values reported in the JPEG file), and second, we fitted a smooth, monotonically increasing curve to the corresponding RAW-JPEG image values. We also use the *Still Life* scene (courtesy of Industrial Light and Magic, see bottom row of fig. 3) [31]. Each HDR image was tone mapped using the operator of Fattal et al. [8] using settings that emphasize gradients ($\beta = 0.8$, fig. 3-(a)). The corresponding tone mapping is aesthetically interesting but contains severe noise.

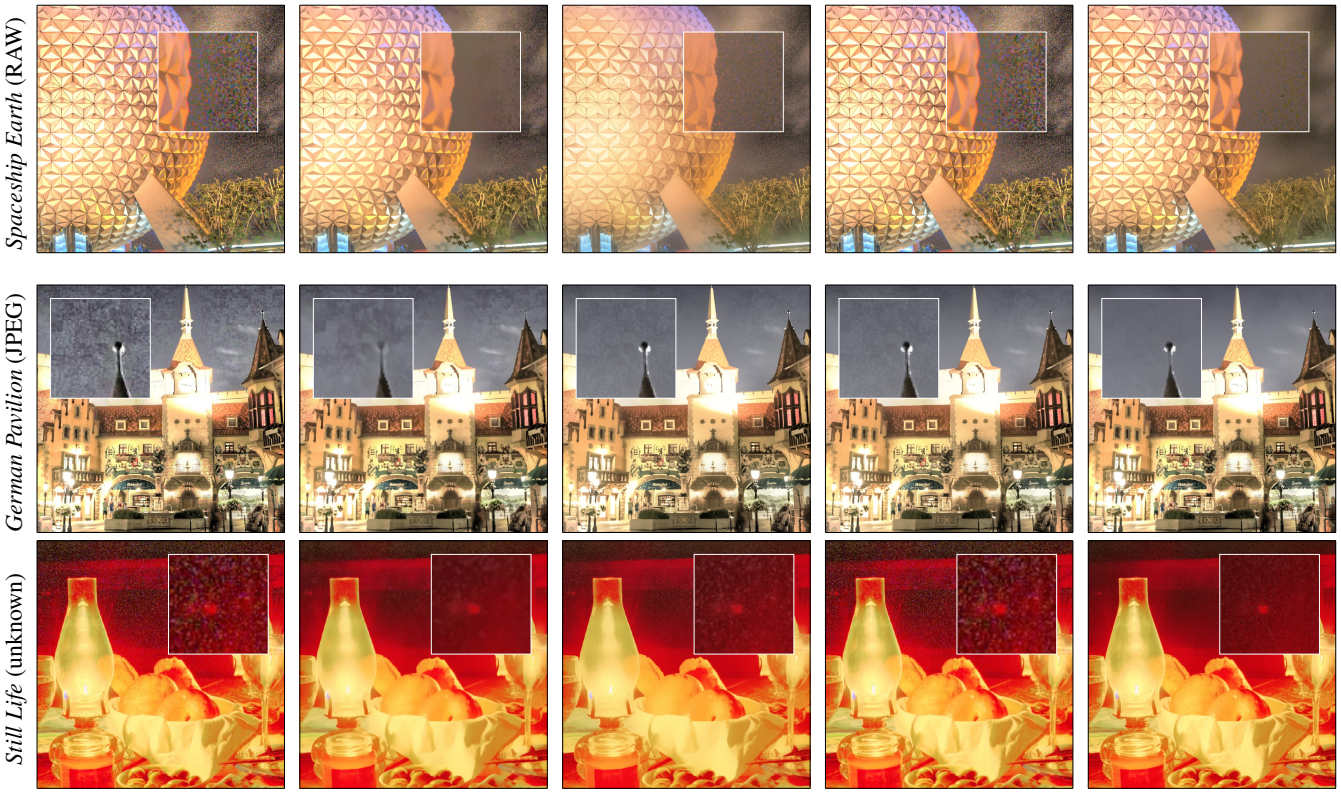
For each scene, we denoised the tone mapped image using non-local means [11], standard bilateral filtering [28], and BM3D [5] (fig. 3-(b)–(d)), using the optimal parameters obtained from the quantitative evaluation in sec. 5.2. No method can perform effective denoising in all cases. This result is expected as the assumptions made by image denoising algorithms (e.g., additive Gaussian noise) are not satisfied in tone mapped images. In contrast, after estimating the noise distribution of the original HDR image, our method performs better denoising while preserving the image structure (fig. 3-(e)). The supplementary material shows more results with additional tone mapping operators [25, 22].

[2] http://www.disneyresearch.com/publication/hdr-image-noise-estimation/



(a) Simulated image (b) *Cinderella Castle* (RAW) (c) *German Pavilion* (JPEG)

Figure 2: Noise estimation from HDR images. First, we extract the mean irradiance and variance of the super-pixels in an HDR image (blue); the lower envelope of this mean-variance scatter plot (cyan) defines an upper bound for the image variance. Given this upper bound, we fit a HDR image noise model to predict the image variance (magenta); our prediction lays within a factor ($k = 2$ for RAW, $k = 8$ for JPEG) of the true variance (yellow) (see sec. 4.2). Our noise model accounts for HDR images produced as the average of multiple photographs (vertical gray lines represent the boundary of the photograph's contribution).

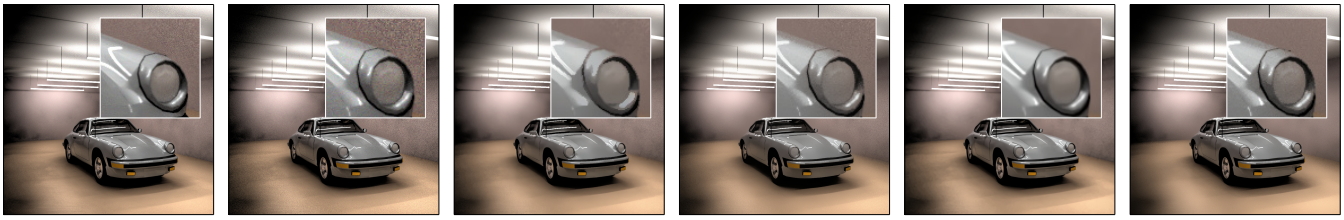


(a) Input (b) Non-local means (c) Bilateral filtering (d) BM3D (e) Ours

Figure 3: Comparison of image denoising algorithms. (a) Input tone mapped image created using [8]); (b) [11]; (c) [28]; (d) [5]; (e) Our method. *Still Life* courtesy of Industrial Light and Magic. Please see the supplementary material for additional results.

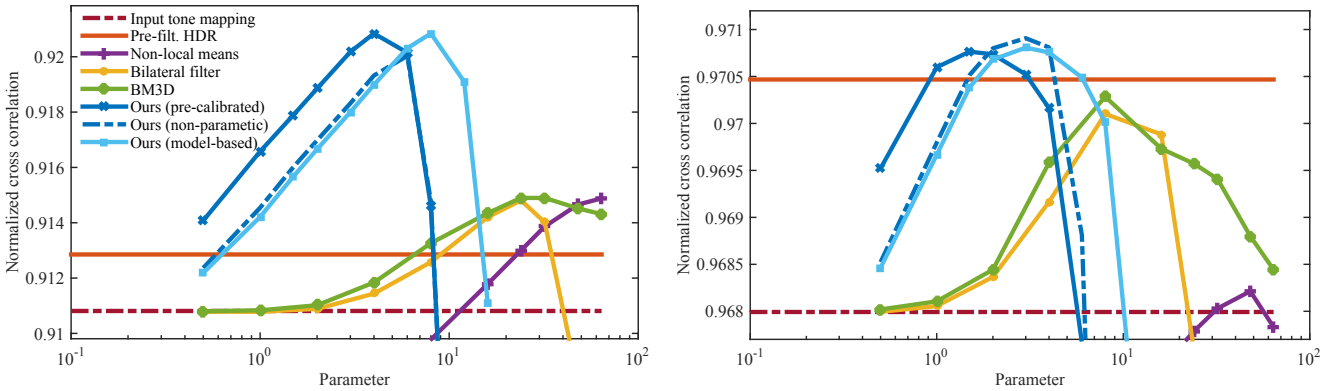
5.2 Quantitative evaluation

For quantitative evaluation, we used a total of 12 rendered HDR images from Ward's image collection [31]. Here, we report results for the file *rend13_o7B0.hdr*. Results for the other 11 images, provided in the supplementary material, show similar behavior. These synthetic images are virtually noise-free (except possible rendering noise, e.g. Monte Carlo noise), so they can be used as a clean plate to simulate noisy images with arbitrary distributions. First, we created a noise-free HDR image by merging the three noise-free exposures (a tone mapped version is shown in fig. 4-(a) for display purposes). Then, we generated a noisy HDR image by capturing three simulated noisy photographs (using the EOS 5D noise parameters estimated in sec. 5.1). We applied the same HDR reconstruction pipeline and tone mapping algorithm described in sec. 5.1, and obtained a noisy HDR-tone mapped image pair (fig. 4-(b)). Taking the noise-free HDR image as ground truth, we compared our method quantitatively with the result of applying a denoising algorithm di-



(a) Noise-free image (b) Noisy image (c) Non-local means (d) Bilateral filtering (e) BM3D (f) Our result ($\sigma^2_{P_n}$)

Figure 4: Rendered scene used to quantitatively evaluate the quality of our method with respect to image denoising strategies. (a) tone mapped version of the noise-free HDR image; (b) input noisy tone mapping; (c)-(f) best denoising result for each method according to normalized cross-correlation with the noise-free HDR image (fig. 5). Results better visualized in electronic version. Image courtesy of Greg Ward. Please see the supplementary material for qualitative results on 11 additional ground truth test images from [31].



(a) Tone mapping of Fattal et al. [8] (b) Tone mapping of Reinhard et al. [25]

Figure 5: Quantitative comparison with standard image denoising methods applied before and after tone mapping, for different values of each denoising method's varied parameter (see text). For two different TMOs, the plots show the normalized cross-correlation (NCC) obtained by non-local means [11], bilateral filter [28], BM3D [5], and our HDR cross bilateral filter. In addition, we also show the NCC obtained by pre-denoising the HDR image using a noise-aware filter [12] followed by tone mapping (solid horizontal line). The NCC of our method is higher than the alternative HDR and tone mapping denoising methods. The legend is the same for both plots. Please see NCC curves on 11 additional ground truth test images from [31] in the supplementary material.

rectly on *i)* the final tone mapped image, and *ii)* the HDR image.

5.2.1 *Comparison with denoising the tone mapped image*

To evaluate the quality of our denoising method, we compared it with three image denoising strategies: non-local means [11], standard bilateral filtering [28], and BM3D [5]. We evaluated our HDR cross bilateral filter using three different noise estimates: *i)* the ground truth noise distribution, *ii)* the non-parametric self-calibration σ^2_{NP}, and *iii)* the model-based self calibration $\sigma^2_{P_n}$. The resulting denoised images are shown in fig. 4.

As mentioned earlier, we seek to compare the resulting denoised images with the original, noise-free HDR image. Inspired by the structure similarity component of the well-known SSIM quality metric [30], we used the normalized cross-correlation because it captures similarity to the original HDR, while normalizing for the expected changes in brightness and contrast due to the tone mapping process. A similar approach has recently been used successfully in other HDR-related applications [3].

In particular, we estimated the normalized cross-correlation (NCC) between the result of: *i)* tone mapping the noise-free HDR image directly and *ii)* denoising the tone mapping of the noise-corrupted HDR image. For each denoising method, we estimated the NCC for a wide range of values of its intrinsic parameter: For non-local means and bilateral filtering, the intrinsic parameter is the standard deviation of the range kernel (σ_g in (1)), for BM3D it is the standard deviation σ of the additive Gaussian noise that is assumed to corrupt the image, and for our method, the parameter is the multiplier of the estimated standard deviation (k in (3)).

Figure 5 shows the resulting NCC for each method and parameter setting, and for two TMOs [8, 25]. The NCC is highest for the three settings of our method (see sec. 4 for a reminder on each setting; "pre-calibrated" denotes the noise model estimated using [12]), indicating their superior ability to reduce noise while preserving the structure of the original HDR image. All three of the other approaches (non-local means, standard bilateral filtering and BM3D) achieved lower NCC values, showing the benefit of exploiting HDR image noise information during denoising.

5.2.2 *Comparison with denoising the HDR image*

We also tested the effect of denoising the HDR image before tone mapping. For this test, we denoised the HDR image using the signal-dependent denoising method of Granados et al. [12] which takes into account the predicted noise distribution of the HDR image. Fig. 5 shows that the NCC of the pre-filtered HDR image was

lower than the one obtained with our method. Moreover, even setting aside NCC evidence, we argue that HDR images can never be completely noise-free because their numerical values are quantized. This quantization noise can always be amplified by TMOs, especially gradient-domain tone mappers. Therefore, we argue that post-filtering is a more principled approach for denoising tone mapped images than pre-filtering the source HDR images.

5.3 Limitations and Future Work

Our approach is subject to the following limitations. First, our method estimates the HDR image noise distribution by taking the lower bound of the color mean-variance scatter plot of a superpixel tessellation of the image (sec. 4.2). However, this lower bound is affected by the uncertainty of the sample standard deviation. Such uncertainty is not accounted for in our model, and therefore, the actual lower bound could be higher than the one estimated by our method, yielding suboptimal results such as the ones shown in fig. 6. Second, the algorithm expects as input the (assumed or known) number of images averaged during HDR reconstruction. This parameter could also be estimated automatically by testing different numbers and selecting the lowest one that properly explains the noise distribution. And third, the current strategy (bilateral filtering) only considers local neighborhoods around each pixel. Other strategies, such as non-local means, could be extended to use the noise distribution of the HDR image as a reference space for comparing local neighborhoods.

6. CONCLUSION

Tone mapping operators aim to preserve perceived luminance differences in high dynamic range (HDR) images, but they sometimes strongly emphasize image noise, particularly in low-light scenarios. We introduce a new HDR cross bilateral filtering method to remove noise in HDR images that were tone mapped with any operator by using the noise distribution of the original HDR image as guide. We also contribute with the first self-calibration method to estimate the noise distribution of HDR images, generated by maximum-likelihood average of RAW or JPEG exposure stacks, when the images in the stack are not available. Quantitative and qualitative evaluation show that the proposed method is more effective than signal-dependent HDR denoising, and image-based denoising strategies that do not consider the noise in the original HDR image.

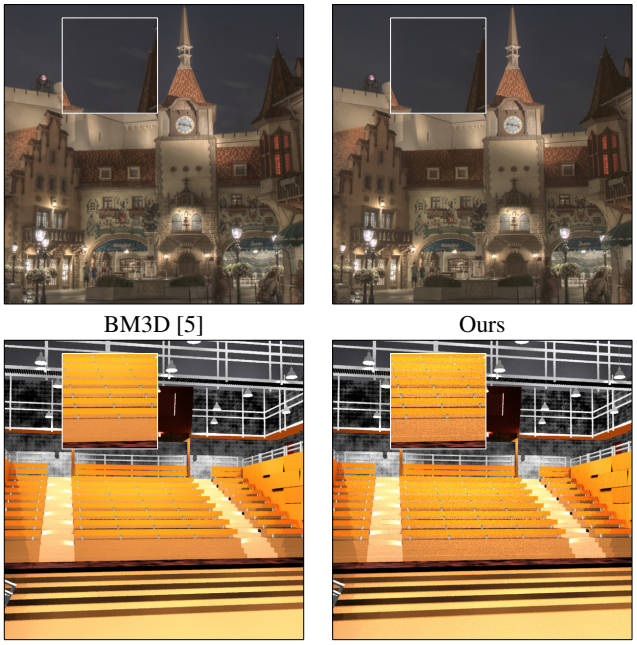


Figure 6: The main reason for failure is inaccurate self-calibration. This yields two main consequences: competing methods may perform better (top row), and our self-calibration procedure may be too conservative, leading to suboptimal results (bottom row). Bottom row: *Candelstick Point State Park Community Theater lighting design*, courtesy of Charles Ehrlich

7. ACKNOWLEDGMENTS

We would like to thank Gregory Ward, Charles Ehrlich, Steve Walker (Arup U. K.), Martin Moeck, and Industrial Light and Magic for providing some of the images used in this paper and the supplementary material. These images were sourced from [31].

8. REFERENCES

[1] D. Batra, P. Yadollahpour, A. Guzmán-Rivera, and G. Shakhnarovich. Diverse m-best solutions in markov random fields. In *Proc. ECCV*, pages 1–16, 2012.

[2] A. Buades, B. Coll, and J. Morel. A non-local algorithm for image denoising. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), 20-26 June 2005, San Diego, CA, USA*, pages 60–65. IEEE Computer Society, 2005.

[3] M. Cadik, R. Herzog, R. Mantiuk, K. Myszkowski, and H.-P. Seidel. New measurements reveal weaknesses of image quality metrics in evaluating graphics artifacts. *ACM Transactions on Graphics (SIGGRAPH Asia)*, 31:1–10, 2012.

[4] P. Chatterjee and P. Milanfar. Is denoising dead? *IEEE Transactions on Image Processing*, 19(4):895–911, 2010.

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.

[6] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-dynamic-range images. *ACM TOG*, 21(3):257–266, July 2002.

[7] E. Eisemann and F. Durand. Flash photography enhancement via intrinsic relighting. *ACM TOG*, 23(3):673–678, 2004.

[8] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. *ACM TOG*, 21(3):249–256, July 2002.

[9] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg. A model of visual adaptation for realistic image synthesis. In *Proc. SIGGRAPH*, pages 249–258, 1996.

[10] A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE Transactions on Image Processing*, 17(10):1737–1754, 2008.

[11] E. S. L. Gastal and M. M. Oliveira. Adaptive manifolds for real-time high-dimensional filtering. *ACM TOG*, 31(4):33, 2012.

[12] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. P. A. Lensch. Optimal HDR reconstruction with linear digital cameras. In *Proc. CVPR*, pages 215–222, 2010.

[13] M. Granados, T. O. Aydin, J. R. Tena, J.-F. Lalonde, and C. Theobalt. Contrast use metrics for tone mapping images. In *International Conference on Computational Photography*, 2015.

[14] M. Granados, K. I. Kim, J. Tompkin, and C. Theobalt. Automatic noise modeling for ghost-free HDR reconstruction. *ACM TOG*, 32(6):201, 2013.

[15] F. Heide, M. Steinberger, Y.-T. Tsai, M. Rouf, D. PajÄĚk, D. Reddy, O. Gallo, J. Liu, W. Heidrich, K. Egiazarian, J. Kautz, and K. Pulli. Flexisp: A flexible camera image processing framework. *ACM Transactions on Graphics*, 33(6), December 2014.

[16] P. Irawan, J. A. Ferwerda, and S. R. Marschner. Perceptually based tone mapping of high dynamic range image streams. In *Proc. EGSR*, pages 231–242, 2005.

[17] J. Janesick. *Scientific charge-coupled devices*. SPIE Press, 2001.

[18] A. G. Kirk and J. F. O'Brien. Perceptually based tone mapping for low-light conditions. *ACM TOG*, 30(4):42:1–42:10, July 2011.

[19] K. Kirk and H. J. Andersen. Noise characterization of weighting schemes for combination of multiple exposures. In *Proc. BMVC*, volume 3, pages 1129–1138, 2006.

[20] C. Liu, R. Szeliski, S. B. Kang, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Trans. PAMI*, 30(2):299–314, 2008.

[21] S. Mann and R. Picard. Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *IS&T 46th annual conference*, pages 422–428, 1995.

[22] R. Mantiuk, K. Myszkowski, and H.-P. Seidel. A perceptual framework for contrast processing of high dynamic range images. *TAP*, 3(3):286–308, 2006.

[23] P. Milanfar. A tour of modern image filtering: New insights and methods, both practical and theoretical. *IEEE Signal Process. Mag.*, 30(1):106–128, 2013.

[24] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Proc. SIGGRAPH*, pages 47–54, 2000.

[25] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images. *ACM TOG*, 21(3):267–276, July 2002.

[26] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski. *HDR Imaging - Acquisition, Display, and Image-Based Lighting, Second Edition*, chapter 6–7. Morgan Kaufmann, 2010.

[27] M. Robertson, S. Borman, and R. Stevenson. Estimation-theoretic approach to dynamic range improvement using multiple exposures. *J. Elec. Imag.*, 12(2):219–228, 2003.

[28] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. ICCV*, 1998.

[29] O. Veksler, Y. Boykov, and P. Mehrani. Superpixels and supervoxels in an energy optimization framework. In *Proc. ECCV*, pages 211–224, 2010.

[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[31] G. Ward. High dynamic range image examples. `http://www.anyhere.com/gward/hdrenc/pages/originals.html`, December 2003. Accessed on May 2014.