

Optimal HDR Reconstruction with Linear Digital Cameras

Miguel Granados
MPI Informatik

Boris Ajdin
Ulm University

Michael Wand
MPI Informatik

Christian Theobalt
MPI Informatik

Hans-Peter Seidel
MPI Informatik

Hendrik P. A. Lensch
Ulm University

Abstract

Given a multi-exposure sequence of a scene, our aim is to recover the absolute irradiance falling onto a linear camera sensor. The established approach is to perform a weighted average of the scaled input exposures. However, there is no clear consensus on the appropriate weighting to use. We propose a weighting function that produces statistically optimal estimates under the assumption of compound-Gaussian noise. Our weighting is based on a calibrated camera model that accounts for all noise sources. This model also allows us to simultaneously estimate the irradiance and its uncertainty. We evaluate our method on simulated and real world photographs, and show that we consistently improve the signal-to-noise ratio over previous approaches. Finally, we show the effectiveness of our model for optimal exposure sequence selection and HDR image denoising.

1. Introduction

Most real world scenes span a radiance range that exceeds the capabilities of standard film and digital cameras. However, many applications, such as image-based lighting [4], and BRDF measurement [9], require access to the whole dynamic range of a scene. Although high dynamic range (HDR) cameras are available [15], cost and performance issues still prevent their widespread use.

Using standard cameras, one can recover HDR images from a set of photographs with different exposure times by first linearizing the images using the inverse of the camera response function, and then averaging them into a single HDR image [12]. A weighted average is used to account for the reliability of every pixel measurement. However, weighting functions proposed in the literature (Sec. 2) do not appropriately consider the individual noise sources of the acquisition process (except [18], [8]). This problem becomes increasingly relevant when the HDR images are not only meant for visualization but as accurate measurements

for the physical scene irradiance.

Taking into account all noise sources, i.e., temporal (photon and dark current shot noise, readout noise) and spatial (photo-response and dark current non-uniformity), we establish an optimal compound-Gaussian model that faithfully estimates the variance in the measurement depending on exposure and exposure time. Based on this model, we derive an optimal weighting function for linear sensors (Sec. 4). As the noise recursively depends on the incident irradiance, we propose an iterative optimization for the maximum likelihood estimate of the irradiance and its uncertainty (Sec. 5).

The parameters of the Gaussian HDR model for a real camera are obtained per pixel to account for spatial noise (Sec. 6). The performance of our HDR reconstruction is compared to the signal-to-noise ratio achieved by previous approaches in Sec. 7.

Based on the novel noise model we can further analyze optimal solutions to two additional applications: The optimal sequence of exposure times can be found by maximizing the SNR of the resulting HDR image [2], or by explicitly setting the desired sampling density of the radiance range [5]. Following the former approach our estimate for the SNR benefits from the advanced noise model (Sec. 8). Finally, the per pixel uncertainty obtained as a byproduct of our HDR reconstruction can be used to drive a spatial smoothing process for HDR denoising (Sec. 9, Fig. 1).

2. Previous work

Several algorithms have been proposed to combine a set of low dynamic range (LDR) images into a single HDR image [3, 11, 12, 13, 15, 16, 18]. In general, these methods estimate both an HDR image that roughly corresponds to the real-world irradiance X (up to scale), and the inverse of the camera response function f^{-1} that maps a digital output v to its inducing exposure Xt , where t is the exposure time. The irradiance can be estimated from the weighted average

$$\hat{\mu}_X = \frac{\sum_i w(v_i) \hat{x}_i}{\sum_i w(v_i)}, \text{ where } \hat{x}_i = \frac{f^{-1}(v_i)}{t_i}. \quad (1)$$

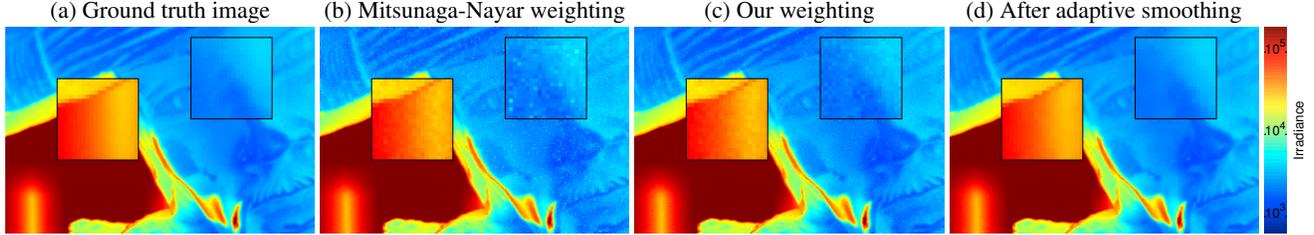


Figure 1. HDR image reconstruction and denoising. (a) Ground truth HDR image, recovered from 180 images (color-coded). (b) Reconstruction using Mitsunaga-Nayar weighting, which is suboptimal at lower intensities since readout noise is unaccounted for. (c) Reconstruction using our weighting, which accounts for readout and dark current noise. (d) Denoised reconstruction using the predicted pixel uncertainty to control the smoothing bandwidth parameter. Blue colors correspond to the lowest irradiance in the scene; reds to the highest.

| | Type | Formula |
|------------------------|----------|---|
| Mann & Picard [12] | Quant. | $\frac{1}{\frac{d}{dv}(\log g(v))}$ |
| Debevec & Malik [3] | Hat | $\min(v - v_{min}, v_{max} - v)$ |
| Mitsunaga & Nayar [13] | SNR | $\frac{g(v)}{g'(v)}$ |
| Reinhard et al. [15] | SNR·Hat | $\frac{g(v)}{g'(v)} \left[1 - \left(\frac{v}{v_{mid}} - 1 \right)^{12} \right]$ |
| Robertson et al. [16] | Variance | $\frac{t^2}{\frac{d}{dv}(\log g(v))}$ |
| Tsin et al. [18] | St. dev. | $\frac{t}{\hat{\sigma}_{g(v)}}$ |
| Kirk & Andersen [8] | Variance | $\frac{t^2}{g'(v)^2 \sigma_v^2}$ |

Table 1. Weighting functions for HDR reconstruction. Here, v is the camera digital output, t the exposure time, and $g(v) \equiv f^{-1}(v)$ is the inverse of the camera response function.

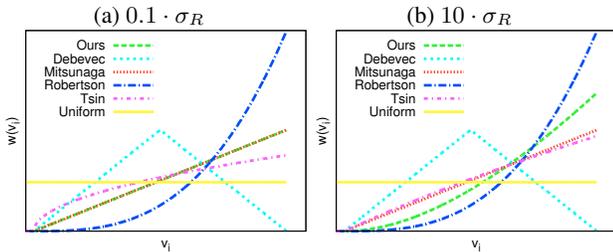


Figure 2. Weighting functions at different readout noise levels σ_R . Weights are given to output values v_i , which measure a constant irradiance on increasing exposure times. Note that Mitsunaga-Nayar converges to our optimal weighting as the readout noise goes to zero. In linear sensors, saturation can be handled by estimating the threshold and excluding the affected values (Sec. 6). Simulations performed on *camera-B* (Table 2).

Each method proposes a different weighting function $w(v)$. These are listed in Table 1, and compared in Fig. 2.

In their seminal paper, Mann and Picard [12] assign weights to output values according to the derivative of the inverse camera response, in order to avoid coarsely quantized values. Derivatives are computed in a logarithmic scale to make the quantization error perceptually uniform.

Debevec and Malik [3] propose a hat function that assigns higher weights to mid-gray values as they are farthest

from both under-exposed and saturated outputs.

Mitsunaga and Nayar [13] suggest a weighting function based on the signal-to-noise ratio (SNR). However, since the camera noise behavior is deemed unknown, they assume the uncertainty is constant across the output range. Although the noise structure is signal dependent (Sec. 3), their weighting is still optimal under specific conditions (Sec. 7). Reinhard et al. [15] extend the Mitsunaga-Nayar weighting by applying a broad hat function that suppresses under-exposed and saturated values.

Assuming compound-Gaussian noise, Robertson et al. [16] propose a probabilistic approach to derive an optimal weighting function. The resulting weighting decreases linearly with the digital output variance, and increases quadratically with the exposure time. Analogously to Mitsunaga and Nayar, they replace the unknown variance with the Mann-Picard weighting, thus accounting only for quantization noise.

Tsin et al. [18] are the first to exploit the camera noise model. They propose a weighting function based on the output standard deviation, which they estimate directly from the input images. However, we show that the standard deviation is still a suboptimal weighting under the assumption of compound-Gaussian noise.

Similarly to Robertson et al., Kirk and Andersen [8] derive a variance penalizer as the optimal weighting function. For the case of linear sensors, they provide variance estimates based on a simplified camera noise model. However, their variance estimates are derived directly from the camera output, which transfers the measurement uncertainty into the weighting function. This condition is shared by all previous methods.

In this paper, we extend Kirk-Andersen by adopting a more rigorous camera noise model that accounts not only for temporal but also spatial noise sources. We calibrate the sensor parameters beforehand, which leaves us room to iteratively estimate the irradiance and its uncertainty, in a way that noisy measurements do not perturb the weighting function. Lastly, we take advantage of the predicted uncertainty to denoise the resulting irradiance map in an optimal way.

3. Camera noise

As we discuss in Sec. 4, the optimal weighting function requires us to provide a model for the camera noise behavior at different exposure settings. Several methods exist for characterizing noise in digital cameras [6, 7, 14]. In Sec. 6, we present a pipeline for noise characterization tailored to HDRI reconstruction. We need to deal with two noise classes: temporal sources and spatial sources. Noises that manifest in pixel value differences between exposures after all acquisition parameters are left untouched are regarded as *temporal noise*. Measurement discrepancies that occur between pixels exposed to the same light intensity are referred to as *spatial noise*. We briefly discuss them next.

3.1. Temporal noise

Photon Shot Noise (PSN). The number of photons arriving to the camera sensor in a given time slack follows a Poisson distribution, whose uncertainty is called *photon shot noise*. For instance, long exposures and bright objects suffer from larger shot noise. However, since the objective is to estimate the irradiance X , the exposure Xt is divided by the exposure time t . The resulting uncertainty is $\sqrt{Xt^{-1}}$. Therefore, long exposures are preferred as they provide better SNRs.

Dark Current Shot Noise (DCSN). In a perfect sensor, a single arriving photon frees a constant amount of electrons (just one for the visible spectrum). Unfortunately, due to thermal energy, some electrons are freed without any interacting photon. These are called *dark current*. The amount collected depends on the exposure time and the sensor's temperature. Similar to PSN, the generation of thermo-electrons follows a Poisson distribution, whose uncertainty is called *dark current shot noise*.

Readout noise. Several noise sources disturb the conversion from charge to digital values. These include *reset noise*, occurring during charge-to-voltage transfer; *white noise* and *flicker noise* during voltage amplification; and *quantization noise* during analog-to-digital conversion. Please refer to [7] for an in-depth review. Since readout noise is a combination of several independent noise sources, it can be described by a Gaussian distribution.

3.2. Spatial noise

Photo-response non-uniformity (PRNU). Despite high quality manufacturing processes, differences persist between the photosensitive area of different pixels. These differences cause that two pixels under the same light intensity produce consistently different readings. The resulting discrepancy pattern is called *photo-response non-uniformity*, and can be understood as a per pixel gain factor. The offset caused by PRNU increases with the signal, and thus it is more evident in brighter image regions.

Dark current non-uniformity (DCNU). Due to temperature differences, the amount of dark current varies between pixels. This variation is referred as *fixed pattern noise (FPN)* or *dark current non-uniformity (DCNU)*, and it can be understood as a per pixel bias. We can correct DCNU by subtracting from each photograph a *dark frame*, i.e. an image acquired with the lens covered but otherwise equal settings (including the integration time and sensor temperature).

4. Stochastic model

By exploiting the camera noise model, our goal is to obtain the best possible irradiance estimate, i.e. that of minimum variance, from a set of N measurements $\{(v_{ij}, b_{ij}, t_i)\}_{i=1\dots N}$, where v_{ij} is the image intensity on pixel $j \in \Omega$, b_{ij} is the dark frame intensity, and t_i is the exposure time. Only the exposure time is allowed to vary, whereas all other camera settings (ISO value, aperture size, focal length) are left fixed. In order to analytically solve for a minimum variance irradiance estimate, we assume *raw* pixel values, i.e. measurements before any in-camera processing (e.g. dark frame subtraction, demosaicing, denoising, white balancing, and compression).

We start by relating the irradiance on the imaging sensor to the digital output value. Let X_j be the amount of photo-induced electrons collected at the pixel capacitor per unit time (scaled irradiance¹). Let D_j be the amount induced by dark current. Let a_j be the pixel gain factor induced by the PRNU. During the exposure, the pixel capacitor will collect $E_{ij} \equiv t_i (a_j X_j + D_j)$ electrons. The output digital value is given by

$$V_{ij} = f(E_{ij}) = [g \cdot E_{ij} + N_R], \quad (2)$$

where g is the overall camera gain factor, N_R is the readout noise with mean μ_R and variance σ_R^2 , and $[\cdot]$ is the round-off operator corresponding to quantization. Assuming the gain factors g , a_j , and the exposure time t_i to be known, the variance of V_{ij} is

$$\sigma_{V_{ij}}^2 = g^2 \sigma_{E_{ij}}^2 + \sigma_R^2, \quad (3)$$

where $\sigma_{E_{ij}}^2 = E[E_{ij}]$ is the shot noise (both PSN and DCSN), and σ_R^2 is the readout noise including the quantization error. Similarly, the dark frame B_{ij} and its variance $\sigma_{B_{ij}}^2$ are given given by

$$B_{ij} = [g \cdot t_i D_j + N_R], \text{ and} \quad (4)$$

$$\sigma_{B_{ij}}^2 = g^2 \sigma_{D_{ij}}^2 + \sigma_R^2, \quad (5)$$

where $\sigma_{D_{ij}}^2 = E[t_i D_j]$ corresponds to the DCSN only.

¹Absolute irradiance values can be derived if the pixel area and the quantum efficiency at the current wavelength are known.

For easing the notation, let us assume a fixed location and, henceforth, omit the location index j . For a given exposure i , we can now derive X_i and its uncertainty σ_{X_i} (Eq. 2, 4), obtaining

$$X_i \approx \frac{V_i - B_i}{t_i \cdot g \cdot a}, \text{ with } \sigma_{X_i}^2 = \frac{\sigma_{V_i}^2 + \sigma_{B_i}^2}{t_i^2 g^2 a^2}. \quad (6)$$

Note that quantization forces the approximation. Analogously, we derive D_i and obtain

$$D_i \approx \frac{B_i - \mu_R}{t_i \cdot g}, \text{ with } \sigma_{D_i}^2 = \frac{\sigma_{B_i}^2 + \sigma_R^2}{t_i^2 g^2}. \quad (7)$$

In the following, we assume that X_i has a Gaussian distribution with mean μ_X (equal for all exposures) and variance $\sigma_{X_i}^2$ (different for every exposure). The suitability of this assumption is discussed in Sec. 7.1.

Due to saturation, which occurs when the sensor capacitor cannot accumulate more charge, Eq. 6 is only valid for values $v < v_{sat}$, where v_{sat} is the saturation limit (see Sec. 6). Naturally, noise introduces uncertainty on the classification saturated values. Therefore, we introduce a probability mass $P^{(nsat)}(v)$ that an observed pixel is not saturated. Let $P(x_i|\mu_X, \sigma_{X_i}^2)$ be the conditional probability density of an observation x_i . We can describe this function by *blending* the unclipped probability density (under no saturation) with a uniform probability, where the blending factor is given by the saturation probability:

$$P(x_i|\mu_X, \sigma_{X_i}^2) = \left(1 - P^{(nsat)}(v)\right) P^{(unif)} + P^{(nsat)}(v) \cdot P^{(unclipped)}(x_i|\mu_X, \sigma_{X_i}^2). \quad (8)$$

5. Optimal Reconstruction

Since our goal is to reconstruct the mean radiance μ_X with the lowest variance from a set $\{x_i\}_{i=1\dots N}$ of independent measurements, we compute the conditional probability of X as the joint probability

$$P(x_i|\mu_X, \sigma_{X_1}^2, \dots, \sigma_{X_N}^2) = \prod_{i=1}^N P(x_i|\mu_X, \sigma_{X_i}^2). \quad (9)$$

The maximum likelihood estimate for X is given by

$$\hat{\mu}_X = \arg \max_{\mu_X} \prod_{i=1}^N P(x_i|\mu_X, \sigma_{X_i}^2). \quad (10)$$

As this p.d.f. is typically convex, we can approximate $\hat{\mu}_X$ iteratively from N images by Newton estimation, starting from an initial averaged estimate $\bar{x} = (\sum_i S(v_i)x_i)/(\sum_i S(v_i))^{-1}$, where $S(v_i)$ indicates pixel non-saturation.

We can further derive an analytic solution of Eq. 10 by *ignoring* all pixels that are close or beyond the saturation limit by setting $P^{(nsat)}(v) = 1$ for $v < v_{sat} - \epsilon$, and zero otherwise. In this setting, Eq. 10 simplifies to

$$\hat{\mu}_X = \arg \max_{\mu_X} \prod_{i \in S_j} P^{(unclipped)}(x_i|\mu_X, \sigma_{X_i}^2), \quad (11)$$

where $S_j \subseteq \{1, \dots, N\}$ is the set of non-saturated exposures for pixel j . From this expression, we obtain the maximum likelihood estimate

$$\hat{\mu}_X = \frac{\sum_{i \in S_j} \frac{1}{\hat{\sigma}_{X_i}^2} \hat{x}_i}{\sum_{i \in S_j} \frac{1}{\hat{\sigma}_{X_i}^2}}, \text{ with } \hat{\sigma}_{\mu_X}^2 = \frac{1}{\sum_{i \in S_j} \frac{1}{\hat{\sigma}_{X_i}^2}}. \quad (12)$$

Analogously, from Eq. 7 we can draw estimates for the dark current parameters $\mu_D, \sigma_{\mu_D}^2$.

From Eq. 1,12, it follows that the optimal weighting function for HDR reconstruction is $w_{opt}(v_i) = (\sigma_{X_i}^2)^{-1}$. After plugging in Eq. 3, 5, 6, we obtain

$$w_{opt}(v_i) = \frac{1}{\sigma_{X_i}^2} = \frac{t_i^2 g^2 a_j^2}{g^2 t_i (a_j \mu_X + 2\mu_D) + 2\sigma_R^2}. \quad (13)$$

We can see that shot noise introduces a circular dependency between the estimate of μ_X and the variances $\{\sigma_{X_i}^2\}_{i=1\dots N}$. The same applies for the dark current. For this reason, we iteratively solve for both. Assuming initial constant variances, we solve for the $\hat{\mu}_X, \hat{\mu}_D$ estimates (Eq. 6, 7, 12). Using the latter, we estimate the variances $\hat{\sigma}_{X_i}^2$ (Eq. 13), and iterate until convergence. The complete reconstruction pipeline is presented in Fig. 3.

6. Parameter estimation

In order to optimally recover μ_X , we provide estimates for the readout noise parameters μ_R, σ_R^2 , the saturation limit v_{sat} , the per-pixel gain factors a_j , and the overall camera gain g , based on the work by Janesik [7].

Readout noise. We estimate the readout noise from a *bias frame*, i.e. an image acquired with zero integration time. Assuming the readout noise is decoupled from the sensor location, we can consider each pixel as an independent random variable of the same distribution, and thus estimate the noise parameters μ_R, σ_R^2 from the frame spatial mean and variance.

Saturation limit. We estimate the saturation limit v_{sat} as the spatial mean of a *saturation frame*, where the sensor is exposed long enough so that every pixel reaches full well capacity. One would expect that all pixels values be at the maximum $2^{bpp} - 1$, where bpp is the number of bits of the digital-to-analog converter. However, this is can be prevented by readout noise. In order to ensure that no saturated values be averaged, we set the limit to $v_{sat} - \epsilon$, where

ϵ corresponds to three times the saturation frame's (spatial) standard deviation.

Per pixel gain (PRNU). Per pixel gain factors a_j are expected to follow a normal distribution with unit mean and a small standard deviation (1%). The factors can be recovered from a *flat field*, which is acquired by illuminating the sensor with a spatially uniform, narrow band light source (e.g. using a diffuser and bandpass wavelength filters). In order to maximize the flat field signal-to-noise ratio, the exposure time and illumination intensity should be set such that the output values are close to saturation. In principle, gain factors can be derived by simply dividing each flat field pixel value ff_j by the spatial frame average \overline{ff} . However, in order to deal with bias and temporal noise, we need to first correct for dark current, and then frame average several corrected flat fields, leading to the estimates

$$a_j = \frac{E[ff_j] - E[b_j]}{\frac{1}{|\Omega|} \sum_j (E[ff_j] - E[b_j])}. \quad (14)$$

Camera gain. The camera gain is the factor converting charge stored at pixel capacitors, i.e. the exposure, to digital values. Due to the sensor's quantum efficiency, this factor is wavelength dependent. In practice, we can ignore this dependency. According to Janesick [7], from Eq. 2 we derive

$$g = \frac{E[V_{ij}] - \mu_R}{E[E_{ij}]}. \quad (15)$$

The expectation over the output value can be approximated by the spatial average of a flat field image, i.e. $E[V_{ij}] \approx \text{avg}_\Omega[ff_j]$. Since $E[E_{ij}] = \sigma_{E_{ij}}^2$, we could also approximate this term by the spatial variance $\text{var}_\Omega[ff_j]$. However, this variance carries readout and PRNU noise as well. By taking the difference of two flat fields, we can virtually eliminate PRNU induced variance, and derive the estimate

$$\hat{g} = \frac{1}{k} \cdot \frac{\frac{1}{2} \text{var}_\Omega[ff_j^{(1)} - ff_j^{(2)}] - \sigma_R^2}{\text{avg}_\Omega[ff_j] - \mu_R}, \text{ where} \quad (16)$$

$$k = \frac{\mu_X(1 + \sigma_a^2) + \mu_D}{\mu_X + \mu_D}, \quad (17)$$

and σ_a^2 is the PRNU spatial variance. Since σ_a^2 is usually very low ($\sigma_a \approx 1\%$), the factor k is usually omitted. To handle readout noise, several gain estimates obtained from distinct flat frame differences should be averaged.

7. Evaluation

In this section, we compare the performance of our weighting function with previous approaches. We tested our method using two digital cameras: a high end Canon EOS 5D, 12-bit DAC (named *camera-A*), and a consumer Canon PowerShot S5, 10-bit DAC (named *camera-B*), both

Calibration (Sec. 6)

1. From bias frame; estimate readout noise μ_R, σ_R^2
2. From saturation frame; estimate saturation value v_{sat}
3. From n flat fields; estimate gain per pixel a_j (Eq. 14)
4. Using two+ flat fields, estimate camera gain g (Eq. 16)

HDRI reconstruction (Sec. 4)

1. Acquire LDR images v_i and dark frames b_i
2. Assume constant variances $\hat{\sigma}_{X_i}^{2(0)}$
3. Estimate $\hat{\mu}_X^{(i)}$ assuming $\hat{\sigma}_{X_i}^{2(i-1)}$ (Eq. 12)
4. Estimate $\hat{\sigma}_{X_i}^{2(i)}$ assuming $\hat{\mu}_X^{(i-1)}$ (Eq. 13)
5. Iterate 3 & 4 until convergence
6. Smooth final $\hat{\mu}_X$ according to $\hat{\sigma}_{\mu_X}^2$ (Sec. 9)

Figure 3. Pipeline for optimal HDRI reconstruction.

Table 2. Estimated sensor parameters for our two test cameras

| # | Model | ISO | \hat{g} | $\hat{\mu}_R$ | $\hat{\sigma}_R^2$ | \hat{v}_{sat} |
|---|--------------------|-----|-----------|---------------|--------------------|-----------------|
| A | Canon EOS 5D | 400 | 0.23 | 128 | 6.5 | 3709 |
| B | Canon PowerShot S5 | 400 | 0.92 | 32 | 18 | 1023 |

set to ISO400 sensitivity, and with noise removal features disabled. We estimated the camera parameters from one bias, one saturation, and 36 flat field frames, as described in Sec. 6. The resulting parameters are presented in Table 2. In order to assess the reliability of our camera model, we run the experiments both on real world images and on simulated images. The simulation was performed using the ground truth irradiance, acquisition sequence, and calibrated camera parameters.

Ground truth acquisition. For each camera, we setup a scene with an intensity range of roughly four orders of magnitude, which we sampled using six exposure times. In order to provide a reference HDR image, we acquired 36 frames and dark frames per exposure time. We averaged them to obtain a six-fold noise reduction. We projected the sample variance to the irradiance domain to derive $\sigma_{X_i}^{2(opt)}$ (Eq. 6). From the camera parameters, image averages, and uncertainties, we derived the ground truth HDR image (Fig. 1-a), and the reconstruction's optimal signal-to-noise ratio (Eq. 12).

Note that, in order to avoid the uncertainty introduced by shutter speed variability, we normalized the images such that the spatial average is constant for every sample of each exposure time. Additionally, due to clamping, the sample variance becomes unreliable as the saturation limit approaches. Therefore, we excluded the affected output values from the reconstruction.

Performance comparison. The quality of a weighting function $w_i(v_{ij})$ depends on how well it emphasizes low variance samples, without discarding the information carried in the high variance ones. Given a single sequence $\{(v_{ij}, t_i)\}_i$, this can be measured by the variance of the weighted mean $\sigma_{\mu_X}^{2(w)} = (\sum_i w_i^2 \sigma_{X_i}^{2(opt)}) (\sum_i w_i)^{-2}$.

From $\sigma^{(w)}$, the ground truth HDR image μ_X , and the bias $(\mu_X - \hat{\mu}_X)$, we compute the signal-to-noise ratio achieved by each weighting. In Fig. 4, we present the resulting ratios on our two test cameras. Note that the SNR for the upper irradiance range is virtually equal for all methods as there are fewer (or just one) images to average, and that discontinuities in the SNR occur at locations where the longest currently contributing exposure becomes saturated. The SNR curves were smoothed to allow for better visual distinction of the performance differences.

In all test cases, our weighting function closely follows the optimal reference, and it consistently achieves higher SNR than previous approaches. The uniform weighting (derived to be optimal in [1]) achieves the lowest SNR, since reliable and unreliable measurements contribute equally. The Kirk-Andersen weighting [8] properly accounts for temporal noise sources (except DCSN), but lacks the spatial ones, hence obtains the optimal SNR minus the bias error. The Tsin et al. weighting [18] penalizes samples according to their standard deviation, and thus, it overemphasizes values close to the noise floor, and downplays those close to saturation, which have the highest SNR. The Debevec-Malik hat [3] also down-weights values close to saturation, which explains the sudden SNR drop on each exposure segment.

As previous weightings are derived directly from the digital output, the measurement uncertainty is carried over into the weighting function. This affects negatively all previous approaches, but becomes evident on Debevec-Malik and Mitsunaga-Nayar [13] on low irradiance ranges (Fig. 1-b). The remaining methods are less susceptible as t^2 factors reduce the influence of short exposures.

Please note the performance difference between Mitsunaga-Nayar and Tsin et al., even though both assign weights according to the output value SNR (Tsin implicitly). Given the unknown noise distribution, Mitsunaga and Nayar assume it to be constant for all output values. As previously discussed, this assumption is invalidated by shot noise. Nevertheless, it can be shown (Appendix A) that this approximation leads to the maximum likelihood weighting when the readout noise approaches zero. This is evident in Fig. 5-a, where the readout noise is suppressed. There, Debevec-Malik also reaches the optimal on the first half of the output range, where weights increase linearly.

Lastly, Robertson et al. [16] performs consistently across the irradiance range since its t^2 factor discards noisy values from short exposures. Recall that the output uncertainty is approximated as derivative of response function in log-scale. The resulting weighting can be shown to be $\sim t^3$. Therefore, longer exposures receive too high importance, and the optimal SNR is no longer achieved. A special case occurs on sensors with high readout noise where all other noise sources become negligible (Fig. 5-b).

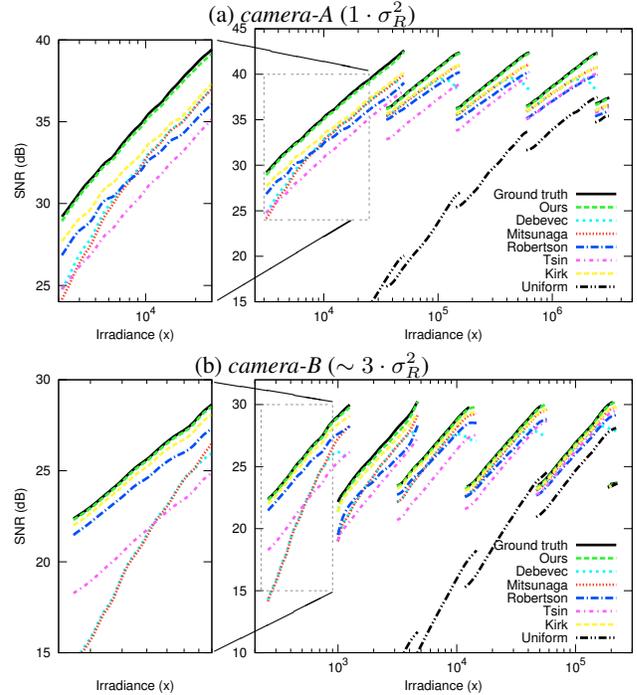


Figure 4. Signal-to-noise ratio achieved by different weighting functions. (a) SNR on *camera-A*. (b) SNR on *camera-B*, whose higher readout noise makes Mitsunaga-Nayar and Debevec-Malik perform worse than on *camera-A*. The SNR curves were smoothed for better visual distinction of the performance. Colors available in the on-line version.

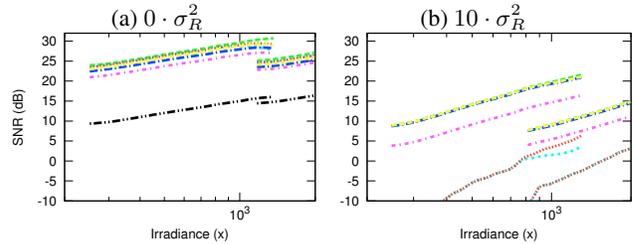


Figure 5. Signal-to-noise ratio in relation to readout noise. (a) Under zero readout noise, the Mitsunaga-Nayar reaches the optimal, and Debevec-Malik does it as well on half of the output range. (b) After a 10-fold increase on readout noise, Robertson and Kirk-Andersen approach the optimal SNR since the high readout noise shadows all other noise sources. Simulations performed with *camera-B* parameters. Legend as in Fig. 4.

7.1. Gaussian noise assumption

We now test the validity of a compound-Gaussian noise assumption. We chose *camera-B* since it has higher readout noise. Fixing the scene and camera parameters, we draw a sample of 36 images, and select three representative pixel locations according to their expected value: right above the noise floor, middle range, and right before saturation. In Fig. 6, we show the corresponding normal Q-Q plots, which indicate that the noise distribution matches that of a Gaus-

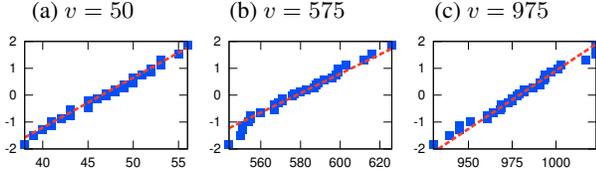


Figure 6. Gaussian assumption validation. Normal Q-Q plots for a sample of 36 pixel outputs at different intensity levels: (a) right above noise floor, (b) middle of the output range, and (c) right before saturation.

sian. This comes at no surprise since the sum of shot noise (Poisson) and readout noise (normal) is a normal distribution for measurable exposure levels.

8. Optimal exposure time selection

If we assume that the irradiance distribution μ_X and camera parameters p are known, we can apply our noise model to compute an exposure sequence \mathbf{t} that provides irradiance estimates whose SNR is above a minimum limit at every pixel. Instead of optimizing the minimum SNR directly, which is highly local, we approximate the sum of penalized ratios

$$\mathbf{t}^{opt} = \arg \min_{\mathbf{t}} \sum_{j \in \Omega} Q(\text{snr}_j(\mathbf{t} : \mu_{X_j}, p)), \quad (18)$$

where $\text{snr} = \hat{\mu}_X \hat{\sigma}_X^{-1}$ is the reconstruction’s SNR, and $Q(\cdot) = \exp^{-1}(\cdot)$ is a heavy penalizer for higher ratios. Note that arbitrary many exposures could be acquired to achieve the desired minimum SNR. Therefore, we follow a greedy algorithm that computes optimal sequences of increasing size until the target SNR is reached.

In Fig. 7, we show the resulting exposure sequences for the irradiance distribution of Fig. 1-a, using *camera-B* parameters. We set the target SNRs to 15, 25, and 30dB. The computed sequences include 3, 8, and 15 exposures, respectively. The first image in the sequence always corresponds to the longest exposure before saturation as the SNR increases with the signal. Subsequent exposures resample the irradiance range with the lowest ratio. Our method slightly favors shorter exposures since they improve wider intensity ranges and produce fewer saturated pixels.

9. HDR image smoothing

Liu et al. [10] propose a method for LDR image denoising, where the noise is estimated locally and then filtered out. Since we simultaneously estimate the irradiance and its uncertainty, we can extend this filtering concept to the HDR domain. In Fig. 8, we present the denoised HDR image using bilateral filtering [17], where the range bandwidth is spatially variant according to the predicted uncertainty. We

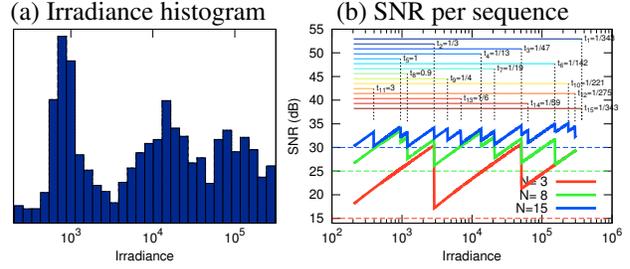


Figure 7. Optimal exposure sequences. (a) Irradiance distribution. (b) Resulting exposure sequences including 3, 8, and 15 images, for a target reconstruction SNR of 15, 25, and 30dB, respectively. Reconstruction using *camera-B* parameters. Please zoom-in in the on-line version for time labels.

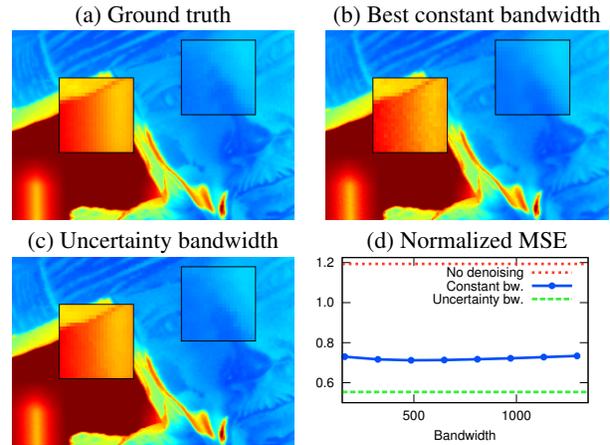


Figure 8. Optimal bandwidth for HDR image denoising. (a) Ground truth HDR image, low irradiances shown in blue, high in red. (b) Smoothing using the constant bandwidth with minimum normalized MSE. (c) Smoothing according to our predicted image uncertainty. (d) The normalized MSE shows that our adaptive method achieves lower error than any given constant bandwidth.

computed the mean normalized squared error between the denoised image and the ground truth image, and compared it against different constant bandwidths. The resulting error plot (Fig. 8-d) indicates that the reconstruction error of the uncertainty-based denoising is lower than that of any constant bandwidth.

10. Conclusions and future work

In this paper, we have presented a novel framework for the creation of HDR images from a set of differently exposed LDR raw images that takes into account the camera noise characteristics. The HDR pipeline is built around a statistical noise model that incorporates both temporal and spatial noise sources. This approach allows for easy theoretical insights into camera behavior with regard to the exposure time, scene intensity, and the interfering noise sources. As a result we have derived an optimal weighting function

for the estimation of HDR values from a series of LDR raw images, based on a compound-Gaussian noise assumption. Our model further allows us to compute optimal exposure sequences, and to perform high quality HDRI denoising.

The next logical step is to characterize the noise behavior of cameras where raw data is not available. There, one needs to account for the effect of non-linear transformations occurring during in-camera processing. This may prove difficult as the exact pipeline is rarely provided by manufacturers, unlike with the imaging sensors.

11. Acknowledgments

We would like to thank the reviewers for their helpful remarks. Also, our thanks go to Thorsten Grosch, Matthias Hullin, and Tongbo Chen for their valuable input. This work has been partially funded by the DFG Emmy Noether fellowship (Le 1341/1-1).

References

[1] A. A. Bell, C. Seiler, J. N. Kaftan, and T. Aach. Noise in high dynamic range imaging. In *Proc. IEEE Int'l Conf. Image Proc.*, pages 561–564, 2008. 6

[2] T. Chen and A. E. Gamal. Optimal scheduling of capture times in a multiple-capture imaging system. In *Proc. SPIE*, volume 4669, pages 288–296, 2002. 1

[3] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proc. ACM SIGGRAPH*, pages 369–378, 1997. 1, 2, 6

[4] P. E. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Proc. ACM SIGGRAPH*, pages 189–198, 1998. 1

[5] M. Grossberg and S. Nayar. High dynamic range from multiple images: Which exposures to combine? In *Proc. ICCV Workshop on Color and Photometric Methods in Computer Vision*, 2003. 1

[6] G. Healey and R. Kondepudy. Radiometric CCD camera calibration and noise estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(3):267–276, 1994. 3

[7] J. Janesick. *Scientific charge-coupled devices*. SPIE Press, 2001. 3, 4, 5

[8] K. Kirk and H. J. Andersen. Noise characterization of weighting schemes for combination of multiple exposures. In *Proc. British Mach. Vis. Conf.*, volume 3, pages 1129–1138, 2006. 1, 2, 6

[9] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Trans. Graph.*, 22(2):234–257, 2003. 1

[10] 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. Pattern Anal. Mach. Intell.*, 30(2):299–314, 2008. 7

[11] S. Mann and R. Mann. Quantigraphic imaging: Estimating the camera response and exposures from differently exposed images. In *Proc. IEEE Conf. Comp. Vis. Patt. Recog.*, pages 842–849, 2001. 1

[12] S. Mann and R. W. Picard. Extending dynamic range by combining different exposed pictures. In *Proc. IS&T Ann. Conf.*, pages 442–448, 1995. 1, 2

[13] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. In *Proc. IEEE Comp. Vis. Patt. Recog.*, pages 1374–1380, 1999. 1, 2, 6

[14] Y. Reibel, M. Jung, M. Bouhifd, B. Cunin, and C. Draman. CCD or CMOS camera noise characterization. *Eur. Phys. J.*, 21(21):75–80, 2003. 3

[15] E. Reinhard, G. Ward, S. Pattanaik, and P. Debevec. *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, 2005. 1, 2

[16] 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. 1, 2, 6

[17] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *Proc. IEEE Int'l Conf. Comp. Vis.*, pages 839–846, 1998. 7

[18] Y. Tsin, V. Ramesh, and T. Kanade. Statistical calibration of the CCD imaging process. In *Proc. IEEE Int'l Conf. Comp. Vis.*, pages 480–487, 2001. 1, 2, 6

A. Mitsunaga-Nayar approximation

Mitsunaga and Nayar propose that values with higher signal-to-noise ratio receive higher weight, as they carry more information. Given the exposure $E \equiv \mu_X t_i$ inducing a pixel value v_i , their weighting function is given by

$$w_{\text{mitsunaga}}(v_i) \equiv \frac{E}{\sigma_E} \approx \frac{E}{\frac{\partial E}{\partial v_i} \sigma_{v_i}} \approx \frac{E}{\frac{\partial E}{\partial v_i}}, \quad (19)$$

where σ_{v_i} is assumed to be constant across different exposure times, disregarding shot noise. For linear cameras, we get $w_{\text{mitsunaga}}(v_i) = \mu_X t_i$. Since the irradiance μ_X is constant across exposures, we obtain $w_{\text{mitsunaga}}(v_i) = t_i$.

Now, let us compare this result the optimal weighting in Eq. 13. Only if $\mu_X \gg \mu_D$ and $g^2 t_i \mu_X \gg \sigma_R^2$ holds true, is $w_{\text{mitsunaga}} \approx w_{\text{opt}}$. Finally, if σ_{v_i} is not neglected in Eq. 19, then $w_{\text{mitsunaga}} = \sqrt{t_i}$, which is different from the optimal.