OPTIMAL EXPOSURE CONTROL FOR HIGH DYNAMIC RANGE IMAGING

Keigo Hirakawa
University of Dayton (ECE)
300 College Park, Dayton, OH 45469 USA
k.hirakawa@notes.udayton.edu

Patrick J. Wolfe
Harvard University (SEAS)
33 Oxford Ave, Cambridge, MA 02138 USA
patrick@seas.harvard.edu

ABSTRACT
A common technique used to acquire high dynamic range image data is that of exposure bracketing—short exposure times are required to capture bright regions of the image without saturation, whereas long exposure times are needed to capture darker image regions effectively. This article describes how to take into account the statistics of the photon arrival process to derive optimal exposure control for maximizing signal recoverability in high dynamic range imaging.

Index Terms— Censored data, dynamic range, exposure bracketing, Poisson process, saturation

1. INTRODUCTION
Suppose for a moment that image data were completely deterministic. Then acquisition and reconstruction of high dynamic range (HDR) image data would pose no problems, as contrast in an underexposed image could be “stretched” indefinitely to enhance low-light regions. In practice noise and HDR imaging are tightly coupled because the photon arrival process is stochastic. Exposure bracketing is commonly used to overcome this limitation—as a long-exposure image is needed to process dark regions, yet a short exposure is needed to prevent bright parts of the images from saturating. The ability to choose a finite set of camera exposures that maximizes the recoverability of HDR images is critically important.

Prior work in this area has focused on post-capture reconstruction of HDR images by synthesizing multiple exposures [1, 2]. In this paradigm, a calibration step is required to estimate camera response function (a map between light intensity and output value), and reconstruction is typically based on heuristic objective functions. Another active research area is that of tone maps for purposes of displaying a HDR image on a limited range display device [3, 4]. A compressive tone mapping is designed to preserve the local regularity and contrast, and combines human visual system models to highlight image details despite hardware limitations.

In contrast, this article is concerned with identifying a set of exposures that maximizes the HDR recoverability. Unlike the aforementioned problems, our understanding of the influence that a set of exposures has on the HDR image reconstruction is limited. For example, the work of [5] derives optimal exposure set based on maximum and minimum scene irradiances, ignoring the heteroscedastic noise properties entirely. The goal of this work is to bridge this gap by providing a statistical analysis of exposure controls.

We begin with the basic assumptions that the quantum efficiency of the image sensor is linear, and that the predominant sources of variability are the photon arrival process itself and the so-called “shot noise” caused by the random behavior of electrons. Noise of these types are well-modeled by Poisson distribution [6]. We consider various practical scenarios under which multiple exposure techniques operate; and based on the distribution of pixels, we make inference on the reconstruction properties of the images yet to be taken.

2. OBSERVATION AND SIGNAL MODELS
Let $x(i)$ be the latent light intensity variable we are interested in measuring. An image sensor is an integrating detector, meaning it accumulates photons over an integration period $k$. Thus, the number of photons that reach the surface of the $i$th pixel sensor over $k$ seconds is stochastic with Poisson distribution: $y(i)|x(i) \sim \mathcal{P}(k|x(i))$. Here we further assume that the distribution of $x(i)$ is Gamma: $x(i) \sim \text{Gamma}(\alpha, \beta)$. This prior model is employed for two reasons. First, Gamma distribution is a conjugate prior for Poisson distribution, making the subsequent inferences computationally tractable. Second, the distribution of pixel values has a heavy positive skew. Multimodality can subsequently be captured by a finite Gamma mixture model (see Section 3.3). Where understood, the location index $i$ is omitted from text.

Owing to the fact that electrons generated by the photocurrent are stored in a capacitor of limited size $\tau$, the image sensor observation is saturated: $z = \min(y, \tau)$. This phenomenon is often referred to as “right censoring” in statistics. Our main challenge is to quantify the impact of the right censoring in the recovery of HDR images.

3. EXPOSURE CONTROL FOR HDR IMAGING
Extending the signal observation models to the exposure bracketing setting, we are interested in optimally capturing $N$ images $z_1, \ldots, z_N$, where $z_n = \min(y_n, \tau)$, $y_n \sim \mathcal{P}(k_n, x)$. 

Poisson distribution: $y_i \sim \text{Poisson}(\lambda)$. The number of photons that reach the surface of the $i$th pixel sensor over $k$ seconds is stochastic with Poisson distribution: $y(i)|x(i) \sim \mathcal{P}(k|x(i))$. Here we further assume that the distribution of $x(i)$ is Gamma: $x(i) \sim \text{Gamma}(\alpha, \beta)$. This prior model is employed for two reasons. First, Gamma distribution is a conjugate prior for Poisson distribution, making the subsequent inferences computationally tractable. Second, the distribution of pixel values has a heavy positive skew. Multimodality can subsequently be captured by a finite Gamma mixture model (see Section 3.3). Where understood, the location index $i$ is omitted from text.

Owing to the fact that electrons generated by the photocurrent are stored in a capacitor of limited size $\tau$, the image sensor observation is saturated: $z = \min(y, \tau)$. This phenomenon is often referred to as “right censoring” in statistics. Our main challenge is to quantify the impact of the right censoring in the recovery of HDR images.
We consider here two practical scenarios for HDR imaging using multiple exposure techniques: (i) the entire exposure set \( k_1, \ldots, k_N \) is predetermined; or (ii) the \( n \)th exposure \( k_n \) is dependent on \( z_1, \ldots, z_{n-1} \) (i.e., the images that preceded it). Under the first scenario, the empirical histograms of pixel values are used to determine model parameters \((\alpha, \beta)\), and \( k_1, \ldots, k_N \) that maximize recoverability is determined. After acquisition, a standard reconstruction method may be used to produce the final irradiance map from \( N \) images. Under the second scenario, we envision an image acquisition system that operates iteratively by computing the optimal exposure \( k_n \) and updating parameter estimates \((\alpha, \beta)\) after each image capture; after \( N \) acquisitions, a standard reconstruction method is used to produce the final irradiance map.

The mathematics and the methodologies described here apply to both scenarios under consideration (and to the case of single exposure imaging), and the presentation below will remain agnostic to them. To simplify notation, let \( z_{\text{obs}} \) be the set of previously acquired images and \( z_{\text{mis}} \) be the images yet to be taken (i.e., “missing”); \( k_{\text{obs}} \) and \( k_{\text{mis}} \) are defined analogously; and \( \mathbf{z} := \{z_{\text{obs}}, z_{\text{mis}}\} \), \( k := \{k_{\text{obs}}, k_{\text{mis}}\} \). When the entire exposure set must be determined prior to image capture, we let \( z_{\text{obs}} = z_1 \) be a proxy for the empirical histogram of pixel values and \( z_{\text{mis}} = (z_2, \ldots, z_N) \) be the “actual” image capture. When iteratively updating exposure, \( z_{\text{obs}} = (z_1, \ldots, z_{M-1}) \) is the set of previously captured images, and we make inference on \( k_{\text{mis}} = k_M \) and \( z_{\text{mis}} = z_M \) during the \( M \)th iteration.

### 3.1. Parameter Training

Based on \( z_{\text{obs}} \), we would like to train the parameters for the priors distribution, \( \alpha \) and \( \beta \). Although the maximal likelihood estimate (MLE) of \( \beta \) in Gamma distribution is computable, \( \alpha \) is not. Moreover, the right-censoring in \( z_n \) prevents the use of other popular techniques (e.g., moment matching). To overcome this shortcoming, we derive a robust rank-order matching estimate (ROE) of \( \alpha \) based on the mode of the empirical pixel histogram \( p(z_n|z_n < \tau) \). If this mode falls below the capacitor limit \( \tau \), then it is identical to the mode of \( p(y_n) \) which is a known function of \( \alpha \) and \( \beta \).

Below, we detail algorithms for computing the MLE of \( \beta \) when \( \alpha \) is known, and an ROE of \( \alpha \) when \( \beta \) is known. A fixed-point scheme is then used to iterate until convergence.

#### Estimation of \( \beta \)

Suppose \( \alpha \) is known. Recall that the marginal distribution of \( y \) is the negative binomial (or Poisson-Gamma) distribution:

\[
y_n \sim \text{NegBin}(\alpha, \rho_n) \quad \quad p(Y_n = y; \alpha, \beta) = \frac{\Gamma(y + \alpha)}{y!\Gamma(\alpha)} (1 - \rho_n)^{\alpha} \rho_n^y
\]

where \( \rho_n = \frac{\beta}{k_n + \beta} \). Owing to right censoring in \( z_n \), the likelihood based on a pixel \( z_n \) is [7]:

\[
p(z_n; \alpha, \beta) = \begin{cases} p(Y_n = z_n; \alpha, \beta) & \text{if } z_n < \tau \\ p(Y_n \geq z_n; \alpha, \beta) & \text{if } z_n = \tau \end{cases}
\]

Hence the log likelihood function is:

\[
\ln p(z_n; \alpha, \beta) = \delta(z_n < \tau) \ln p(Y_n = z_n; \alpha, \beta) + (1 - \delta(z_n < \tau)) \ln p(Y_n \geq z_n; \alpha, \beta)
\]

\[
= \delta(z_n < \tau) \ln \left( \frac{(z_n + \alpha - 1)!}{z_n!(\alpha - 1)!} + \alpha \ln (1 - \rho_n) + z_n \ln \rho_n \right) + (1 - \delta(z_n < \tau)) \ln B\beta(\rho_n; \tau, \alpha) - \ln B(\tau, \alpha)
\]

where \( B(\rho_n; \tau, \alpha) \) is the incomplete beta function, and \( B(\tau, \alpha) \) is the beta function. To maximize log likelihood, we take a derivative with respect to \( \rho_n \):

\[
\frac{\partial \ln p(z_n; \alpha, \beta)}{\partial \rho_n} = \delta(z_n < \tau) \left[ \frac{\alpha}{\rho_n - 1} + \frac{z_n}{\rho_n} \right] + (1 - \delta(z_n < \tau)) \frac{\rho_n^{\tau-1}(1 - \rho_n)^{\alpha-1}}{B(\rho_n; \tau, \alpha)},
\]

where we applied fundamental theorem of calculus to the incomplete beta function to obtain this result. Define

\[
\mathcal{L}(\beta|z_{\text{obs}}) := \prod_{n=1}^{M-1} \prod_{i=1}^{\#\{z_n(i) = \tau\}} p(z_n(i); \alpha, \beta).
\]

Then setting its derivative with respect to \( \rho_n \) to zero,

\[
\frac{\partial \mathcal{L}(\beta|z_{\text{obs}})}{\partial \rho_n} = \sum_{n=1}^{M-1} \frac{\#\{z_n(i) = \tau\} \alpha}{\rho_n - 1} + \sum_{\tau = 0}^{\#\{z_n(i) = z\}} \frac{\beta^{\tau-1}(1 - \rho_n)^{\alpha-1}}{B(\rho_n; \tau, \alpha)}
\]

where \( \#\{\cdot\} \) is the number of coefficients that satisfy the set. A standard technique (e.g., Newton’s method) is used to solve for the MLE of \( \rho_n \) (and solve for \( \beta = k_n \frac{\rho_n}{\beta} \)).

#### Estimation of \( \alpha \)

Suppose \( \beta \) is known. If the mode of the empirical histogram of \( z_n \) falls below the threshold \( \tau \), then it is identical to the mode of \( p(y_n) \) and \( \alpha \) is identical to the mode of \( p(y_n; \alpha, \beta) \). Recalling that \( p(y_n; \alpha, \beta) \) is a negative binomial distribution,

\[
\text{mode} \{p(y_n; \alpha, \beta)\} = \left( \frac{\alpha}{\alpha - 1} \left( \frac{p_n}{1 - p_n} \right) \right) = \left( \alpha \left( \frac{1}{\beta} \right) \frac{\beta}{k_n} \right).
\]

Ignoring the flooring function, we have

\[
\alpha \approx 1 + \beta \text{ mode} \left\{ p \left( \frac{z_n}{k_n}; \alpha, \beta \right) \right\}.
\]

Important feature of this method is that \( z_n \) is normalized by \( k_n \), and the mode of \( p(z_n/k) \) should coincide for all \( n \)’s. Hence an empirical histogram of \( z_n \) (conditioned on \( z_n < \tau \)) aggregated over \( M - 1 \) multiple exposure images will have the same mode, and each successive image acquisition should improve the performance of this estimator.
3.2. Optimal Exposure

Supposing images $z_{\text{obs}}$ have been acquired already with exposures $k_{\text{obs}}$, we want $k_{\text{mis}}$ that maximizes the recoverability when images $z_{\text{mis}}$ are taken. For a set of exposures $k = \{k_{\text{obs}}, k_{\text{mis}}\}$, first define the reconstruction risk of exposure $k_{\text{mis}}$ as the expected mean squared error based on model parameters $(\alpha, \beta)$, observed images $z_{\text{obs}}$, and estimate $\hat{x}(Z; k)$:

$$\text{risk} := E[||X - \hat{x}(Z; k)||^2 | z_{\text{obs}}].$$

Naturally, we define recoverability as the opposite of risk infimum: $R(k_{\text{mis}}) := -\inf_{\hat{x}} \text{risk}$. By orthogonality,

$$R(k_{\text{mis}}) = -\min_{\hat{x}} E[||X - \hat{x}(Z; k)||^2 | z_{\text{obs}}] = -E[X^2 | z_{\text{obs}}] + E[E[X|Z]^2 | z_{\text{obs}}]. \quad (3)$$

Note $z_{\text{mis}}$ is marginalized out from the predictive risk, as it is yet to be acquired. We have the following results.

**Lemma 3.1** Let $z_n = \min(y_n, \tau)$, $y_n|x \sim \mathcal{P}(k_n x)$, $x \sim \Gamma(\alpha, \beta)$. Then

$$E[X^m | z] = \Gamma(\alpha + m) p(z; \alpha + m, \beta) \frac{\Gamma(\alpha^m)}{\Gamma(\alpha)}.$$

**Proof** The Gamma distribution has the following property:

$$x^m p(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha - 1 + m} e^{-\beta x} = \frac{\Gamma(\alpha + m)}{\Gamma(\alpha)} p(x; \alpha + m, \beta)$$

Substituting this to the following proves the theorem:

$$E[X^m | z] = \int x^m \frac{\sum y p(z; y)p(y|x)p(x; \alpha, \beta)}{p(z; \alpha, \beta)} dx.$$

**Theorem 2** Let $z_n = \min(y_n, \tau)$, $y_n|x \sim \mathcal{P}(k_n x)$, $x \sim \Gamma(\alpha, \beta)$. Then for all $n \in \{1, \ldots, N\}$

$$y_n | z_{\text{obs}} \sim \text{NegBin}(\alpha + \sum \hat{z}_{\text{obs}}, \hat{\rho}_n)$$

$$\hat{\rho}_n = \beta + \sum (k_{\text{obs}}).$$

**Proof** By conditional independence $p(y_n | x, z_{\text{obs}}) = p(y_n | x)$,

$$p(y_n | z_{\text{obs}}; \alpha, \beta) = \int p(y_n | x) p(x|z_{\text{obs}}) dx.$$
acquisition with exposure $k_n$ is simulated using the observation model in Section 2 from the RIT/MSCL HDR database [3]. HDR reconstruction is performed using the off-the-shelf method of `makehdr` in Matlab [8]. Images are displayed using Matlab `tonemap` function (logarithm in CIELAB space).

HDR images are reconstructed from 10 multiple exposure images using optimal exposure values and compared to the reconstruction from 10 images taken with linear exposure over the same range—their tonemapped versions are shown in Figure 1. The HDR recovery with proposed exposure control results in a slightly higher level of noise, but not seriously so. However, the recovery from proposed exposure control preserves details in the regions of high dynamic range more faithfully without introducing contouring—an evidence that exposure resolution in this range is adequate.

5. CONCLUSION AND DISCUSSION

We presented a new exposure optimization method for high dynamic range imaging based on the underlying distribution of pixel values. We introduced the notion of recoverability, defined as the negative of the infimum over reconstruction risk, and showed how to solve for a set of exposure values that maximizes it. We note that the techniques presented here are meant to augment, not replace, current research approaches to HDR image reconstruction. Our model currently ignores correlations across pixels and remain agnostic to the choice of HDR reconstruction method—natural next steps for improvements in the future. With proper modifications to (3), we can incorporate aperture and ±EV controls as well.

The proposed work has other applications. The recoverability metric can predict the minimum number of exposures required to achieve a desired image quality. Spatially multiplexed panchromatic and chromatic pixels in mixed-exposure image sensors [9] reportedly have two separate gain controls.

6. REFERENCES


