

# One-Shot Optimal Exposure Control

David Ilstrup and Roberto Manduchi

University of California, Santa Cruz

**Abstract.** We introduce an algorithm to estimate the optimal exposure parameters from the analysis of a single, possibly under- or over-exposed, image. This algorithm relies on a new quantitative measure of exposure quality, based on the average rendering error, that is, the difference between the original irradiance and its reconstructed value after processing and quantization. In order to estimate the exposure quality in the presence of saturated pixels, we fit a log-normal distribution to the brightness data, computed from the unsaturated pixels. Experimental results are presented comparing the estimated vs. “ground truth” optimal exposure parameters under various illumination conditions.

**Keywords:** Exposure control.

## 1 Introduction

Correct image exposure is critical for virtually any computer vision application. If the image is under- or over-exposed, features or texture are lost, colors are washed out, and the overall perceptual quality of the image is decreased. Correct exposure means that the best possible use is made of the quantization levels provided by the digitization system – in other words, that the *rendering error* due to the non-ideal imaging system is minimized, where the rendering error is the difference between the true irradiance at a pixel and what can be reconstructed based on the measured brightness.

In this paper we propose a quantitative measure for the quality of exposure, along with an algorithm to estimate the optimal exposure based on single, possibly under- or over-exposed, image. By using only one image (rather than several images taken at different exposures) our algorithm enables a fast mechanism for exposure control, a useful characteristic in many contexts. For example, vision system mounted on mobile robots need to adapt quickly to new scenes imaged as the robots moves around. Surveillance systems require prompt response to sudden changes in illumination, such as a light turned on or off. Likewise, through-the-lens (TTL) digital cameras systems for the consumer or professional market may benefit from fast and accurate exposure control.

Our definition of exposure quality requires estimation of the rendering error and of its expected behavior with varying exposure parameters. Unfortunately, the rendering error can only be computed if the original, unprocessed irradiance data is available - a luxury that is not usually available. In particular, if some of the pixels are saturated, their value and thus their rendering error is simply

unknown. We note in passing that, in general, a correctly exposed image contains a certain amount of saturated pixels: an exposure control strategy that simply avoids saturation is usually sub-optimal. We propose a procedure to estimate the rendering error for saturated pixels based on a prior statistical model of the image brightness. Basically, we fit a parametric distribution model to the unsaturated data; the “tail” of this distribution tells us what to expect beyond the saturation point. Computing this model boils down to a problem of parameter estimation from right-censored data, a well-studied statistical technique. Combined with the brightness histogram of the unsaturated data, the model-based distribution for the saturated data allows us to predict how the rendering error changes as one increases or decreases the exposure time, and thus to estimate the optimal exposure, as the one that minimizes the rendering error.

This paper is organized as follows. After presenting related work in Sec. 2, we introduce our quantitative definition of exposure quality in Sec. 3. Next Sec. 4 shows how the exposure quality can be evaluated from a single image, and introduces our parametric statistical model for the unobserved (saturated) pixels. This concept is brought forward in Sec. 5, where we describe how to estimate the rendering error for various exposures from observation of an image at a fixed exposure, enabling a mechanism for estimating the optimal exposure. Quantitative experiments are described in Sec. 6.

## 2 Related Work

Much of the existing literature for automatic exposure control appears as patents (e.g. [1,2,3]). A common theme in all these works is the use of some *scene evaluation* heuristics. Scene evaluation can range from relatively simple goals such identifying back-lit and front-lit scenes [4] to the complex task of face detection [5]. Once the most important areas of the scene are determined, exposure control is adjusted so that some statistic of these pixels, such as the mean, reaches a desired value, often near the middle of the pixel range (e.g. 128 for an 8-bit image). Adjustment is normally achieved via dynamic control algorithms [6,7,8].

A per-pixel control algorithm where the objective function is based on a model of the camera’s response function is given in [9]. The goal of this system is to modify the exposure of each pixel (or, in this case, the transmittance of a coupled spatial light modulator) so that the irradiant energy is just below saturation. If the pixel is unsaturated, then the next exposure is computed trivially. If the pixel is saturated, then the exposure is decreased by a large constant fraction.

Schulz *et al.* [10] measure the goodness of exposure by the integral of the brightness histogram within the bounds of the camera’s dynamic range (from the minimum brightness above noise level to the maximum brightness before saturation). Although this measure of goodness may resemble the one proposed in this paper, it lacks a sound theoretical justification, and may give very different results from ours.

Recent work on high-dynamic range (HDR) imaging has addressed the issue of how to efficiently combine low-dynamic range(LDR) images into an HDR stack

(see *e.g.* [11]). The goal is to find a *minimal image-bracketing* set that covers all of the scene dynamics. In order to minimize the acquisition time, one needs an efficient strategy for selecting the exposure of the next LDR image to take. Barakat *et al.* [12] propose three different approaches: Blind acquisition; Clairvoyant acquisition; and Blind acquisition with feedback. Under this terminology, our proposed approach can be defined as a blind acquisition system that tries to best capture the scene dynamics after observation of just one previous image.

### 3 Exposure Quality: A Quantitative Definition

A pixel in a camera’s focal plane converts irradiant energy into a number (*brightness*). For a given *exposure time* (or, concisely, *exposure*)  $T$ , the irradiant energy  $I_T$  is a function of the irradiant power integrated over the pixel’s surface<sup>1</sup>,  $I$ :

$$I_T = I \cdot T \quad (1)$$

Note that the irradiant power  $I$  is approximately a linear function of the iris aperture area, especially for pixels near the center of the image, which adds one multiplicative factor in (1). We will assume constant iris aperture in this paper.

Conversion from irradiant energy  $I_T$  to brightness  $B_T$  normally comprises two steps: (a) transformation of  $I_T$  into electrical charge; (b) quantization of a voltage signal that is a function of this charge. For the sake of simplicity, subsequent operations on the digital data (such as white balancing, gamma correction, or sub-quantization) are neglected in this work. We note that, at least for cameras in the higher market segments, these operations can usually be overridden by proper configuration setting.

Formally, this conversion can be represented as follows:

$$B_T = Q(f(I_T)) \quad (2)$$

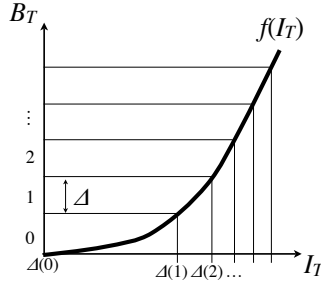
The *sensor’s characteristic*  $f$  can usually be modeled as an invertible noisy function, and can be estimated using standard methods (see *e.g.* [13,14]). The inverse function of  $f$  will be denoted by  $g$ :  $g(f(I_T)) = I_T$ . Note that embedded in the sensor’s characteristic  $f$  is also the variable gain amplification, which can be also used as an exposure parameter.

The non-invertible quantization operator  $Q$  maps values of  $f(I_T)$  into numbers between 0 and  $2^N - 1$ , where  $N$  is the number of bits. Using a mid-tread model [15], the quantization operation can be formalized as follows:

$$Q(x) = \begin{cases} \text{round}(x/\Delta) & , x < (2^N - 1)\Delta \\ 2^N - 1 & , x \geq (2^N - 1)\Delta \end{cases} \quad (3)$$

where  $\Delta$  is the quantization step. In practice, values of  $I_T$  within an *equivalent bin*  $[\Delta(i), \Delta(i + 1)]$ , where  $\Delta(i) = g(i\Delta)$ , are mapped to  $B_T = i$  (see Fig. 1).

<sup>1</sup> Without loss of generality, it will be assumed that the pixel has unit area in an appropriate scale.



**Fig. 1.** Conversion of irradiant energy  $I_T$  into brightness  $B_T$

Values of  $I_T$  above  $g((2^N - 1)\Delta)$  are saturated to  $2^N - 1$ . Note that in the case of linear sensor characteristic ( $f(x) = ax$ ), increasing the exposure by a factor of  $k$  ( $T \rightarrow kT$ ) is completely equivalent to reducing the quantization step by the same factor ( $\Delta \rightarrow \Delta/k$ ).

We define by *rendering error*  $e_T$  at a pixel the difference between the true irradiant power,  $I$ , and the best reconstruction from the brightness  $B_T$ :

$$e_T = I - g(B_T \Delta) / T \tag{4}$$

The irradiant power  $I$  is independent of the exposure setting (for constant iris aperture) and thus represents a more natural domain for the definition of rendering error  $e_T$  than the radiant energy  $I_T$ . Note that the dependence of  $e_T$  on  $T$  as we analyze it is only due to the presence of the quantizer (but see Appendix B). When  $I_T < g((2^N - 1)\Delta)$ , the signal is said to be in the *granular region*.

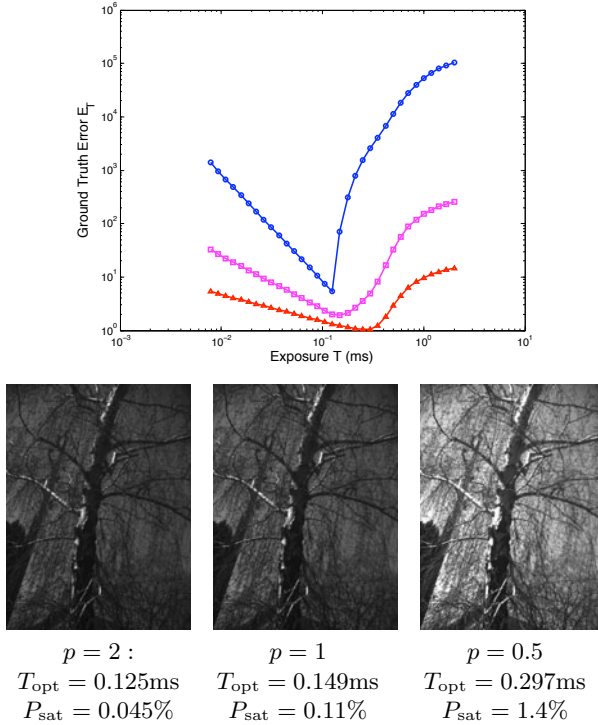
If the equivalent quantization bins are small enough that the sensor’s characteristic  $f(I_T)$  has constant slope within each individual bin, then one easily sees that, when  $I_T$  is within the  $i$ -th equivalent bin, the error  $e_T$  is confined between  $-\alpha(i)\Delta/2$  and  $\alpha(i)\Delta/2$ , where  $\alpha(i) = g'((i + 1/2)\Delta)$ . When  $I_T > g((2^N - 1)\Delta)$ , the signal is said to be in the *overload region*, generating an unbounded error (meaning that the pixel is saturated).

In order to assess the effect of quantization, we can define a positive measure of the rendering error  $L(e_T(m))$  at each pixel  $m$ , and average it over the whole image:

$$E_T = \sum_{m=1}^N L(e_T(m)) / M \tag{5}$$

where  $M$  is number of pixels in the image. The *optimal exposure* for a particular scene is the value of the exposure  $T$  that minimizes the associated error  $E_T$ . The goal of exposure control is thus one of finding the optimal exposure, given the observations (images) available. In this paper, we describe an algorithm that attempts to find the optimal exposure from analysis of a single image, taken with a known (and presumably suboptimal) exposure  $T_0$ .

Our definition of exposure quality promotes a “good” balance between the overload error due to saturation and the granular error for unsaturated pixels. The optimal exposure depends on the chosen error measure  $L$ . One may choose,



**Fig. 2.** The error  $E_T$  as a function of exposure  $T$  for an 8-bit system with  $L(e_T) = |e_T|^p$ . Blue circles:  $p = 2$ . Magenta squares:  $p = 1$ . Red triangles:  $p = 0.5$ . The minimizer of each curve represents the optimal exposure  $T_{\text{opt}}$  for the corresponding measure. The optimally exposed image for the each measures are also shown, along with the percentage of saturated pixels  $P_{\text{sat}}$ .

for example,  $L(e_T) = |e_T|^p$  for an appropriate value of the exponent  $p$ . Larger values of  $p$  penalize the overload error more (since it can grow unbounded). For example, in Fig. 2 we show the error  $E_T$  for  $p = 0.5, 1$  and  $2$  as a function of  $T$  using 8-bit pixel depth for a particular scene. (For this and other experiments we synthesized 8-bit images from a 12-bit image as discussed in Appendix A, and used data from the 12-bit image as “ground truth”). Optimally exposed images for the three measures chosen (corresponding to the minimizers of the curves) are also shown in the image. Note that using  $p = 0.5$ , a brighter image with more saturated pixels (1.4% of the image) is obtained, while  $p = 2$  allows for much fewer saturated pixels only. Other error measures (*e.g.* robust measures such as Tukey’s biweight function) could also be considered. For all experiments in this paper, we used the measure  $L(e_T) = |e_T|$ .

## 4 Evaluating Exposure Quality

Computation of (5) is only feasible if the irradiance  $I$  is known for each pixel, an unrealistic assumption in any practical situation. Instead, one may estimate

$E_T$  by means of the expected error measure over a suitable probability density. More precisely, we model the values of irradiant power at the pixels as samples independently drawn from a density  $p(I)$ . Thus, the expected value  $E_T$  of  $L(e_T)$  can be written as:

$$E_T = \int_0^\infty L(e_T)p(I) dI = E_T^g + E_T^o \quad (6)$$

$$E_T^g = \sum_{i=0}^{2^N-2} \int_{\Delta(i)/T}^{\Delta(i+1)/T} L(e_T)p(I) dI ; E_T^o = \int_{g((2^N-1)\Delta)/T}^\infty L(e_T)p(I) dI \quad (7)$$

In the following analysis we only consider the effect of quantization noise. While the overall level and variance of photon noise can be significant, in Appendix B we argue that this has little effect on the optimum exposure value  $T_{\text{opt}}$ , especially compared to the effect of changing  $L$  in (5) or changes in the distribution of irradiance at the sensor.

If the density  $p(I)$  can be considered constant within each equivalent bin (“high rate” assumption [15]), and still assuming that the sensor characteristic has linear slope within each equivalent bin, the granular error is uniformly distributed within  $-\alpha(i)\Delta/2T$  and  $\alpha(i)\Delta/2T$ . This enables easy computation of the granular error  $E_T^g$ . The dependence of  $E_T^g$  on  $T$  is normally complex, except when the sensor has a linear characteristic  $f(I_T)$ , in which case the following identity holds:

$$E_T^g = \Phi_T \cdot \text{Prob}(I < (2^N - 1)\Delta/T) \quad (8)$$

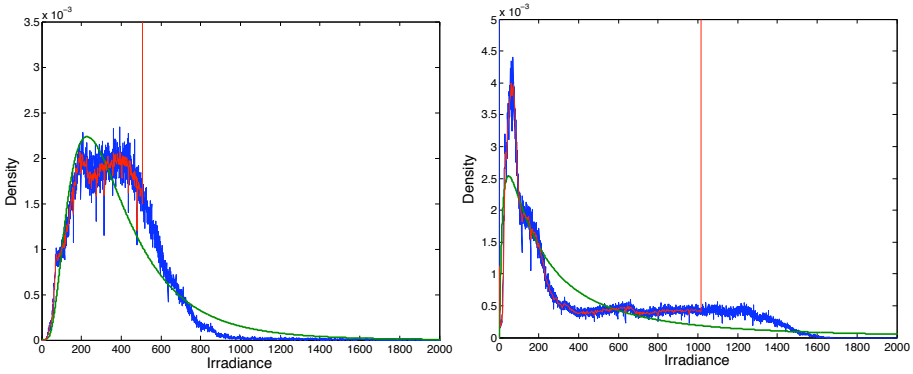
where  $\Phi_T$  is a quantity that decreases with  $T$  but *does not* depend on the density  $p(I)$ . For example, if  $L(e_T) = |e_T|^p$ , then  $\Phi_T = (\Delta/T)^2/12$  for  $p = 2$ ,  $\Phi_T = \Delta/4T$  for  $p = 1$ , and  $\Phi_T = \sqrt{2}\Delta/T/3$  for  $p = 0.5$ .

Eq. (8) formalizes a very intuitive concept, termed “Expose to the right” (ETTR) in the photography community [16]: increasing the exposure time improves the rendering quality for the non-saturated pixels. At the same time, increasing the exposure leads to more saturated pixels as well as to higher over-load error for the saturated pixels.

#### 4.1 Modeling the Irradiance Distribution

What is a good model for the density  $p(I)$ ? Suppose for a moment that all pixels in the image, taken at exposure  $T$ , are unsaturated. Let us define the “continuous domain” histogram as the piecewise constant function  $h_T(x)$  representing the proportion of pixels with  $B_T = \text{round}(x)$ . Note that  $h_T(2^N - 1)$  is the proportion of saturated pixels in the image. The continuous domain histogram  $h_T(x)$  can be used to model the density  $p(I)$  by means of the auxiliary function  $\bar{h}_T(I)$ , defined by (9) where  $f'$  is the derivative of  $f$ .

$$\bar{h}_T(I) = h_T(f(IT)/\Delta) \cdot f'(IT) \cdot T/\Delta \quad (9)$$



**Fig. 3.** The histogram function  $\bar{h}_T(I)$  for the “ground truth” 12-bit image (blue) and for a derived synthetic 8-bit image (red) are shown along with the lognormal density  $\bar{q}_T(I)$  fitted to the right-censored data from the 8-bit image for two different scenes. Note that the 8-bit images saturates for  $I = g((2^8 - 1)\Delta)/T$ .

But what if the image has saturated pixels? The brightness value of these pixels is not observed, and thus the histogram provides only partial information. For these case, we propose to model  $p(I)$  by means of a parametric function, with parameters learned from the unsaturated pixels. Parameter estimation from “right-censored” data is a well studied methodology, and standard methods exist [17,18]. In our experiments, we used the Matlab function `mle.m` which performs ML parameter estimation with right-censored data for a variety of parametric distributions.

We decided to use the lognormal parametric function for representing the marginal probability density function (pdf) of the irradiance data. This choice was suggested by the theoretical and experimental analysis of Richards [19] and Ruderman [20]. In particular, Richards [19] observed that random variables modeling distributions of important contributors to scene brightness, such as illumination sources and angles, surface reflectance, and the viewing angle for non-Lambertian surfaces, affect recorded brightness in a multiplicative fashion. Thus, the logarithm of brightness should be distributed as a sum of random variables, which the central limit theorem approximates as a normal distribution. It should be clear that any choice for a prior distribution of the brightness data is bound to fail in certain instances. For example, the presence of a strong illuminator, or even of the sky in an image, generates a peak in the brightness histogram that cannot be easily accounted for by a parametric distribution, especially if these peaks belong to the saturated region. Still, we believe that the chosen fit provides a simple and, in most cases, realistic estimation of the behavior of the irradiance even for the pixels that are saturated. An example of parametric fit is shown in Fig. 3 for two different scenes. Note that in both cases the 8-bit image saturates; the irradiance values for the saturated pixels are modeled by the lognormal fit.

Let  $q_T(B)$  be the parametric model learned from the right-censored brightness data taken with exposure  $T$ . Similarly to (9), a model  $\bar{q}_T(I)$  for  $p(I)$  based on  $q_T(B)$  can be defined as by:

$$\bar{q}_T(I) = q_T(f(IT)/\Delta) \cdot f'(IT)T/\Delta \quad (10)$$

At this point, we have two different representations for  $p(I)$ : the histogram-based function  $\bar{h}_T(I)$ , which is the best model for the unsaturated data; and the parametric density function  $\bar{q}_T(I)$ , which models the saturated and thus unobservable data. We propose a “composite” model  $\bar{p}_T(I)$  for  $p(I)$  that combines the two models above:

$$p(I) \approx \bar{p}_T(I) = \begin{cases} \bar{h}_T(I) & , I < g((2^N - 1)\Delta)/T \\ \bar{q}_T(I) / K_T & , I \geq g((2^N - 1)\Delta)/T \end{cases} \quad (11)$$

where  $K_T$  is a normalization constant:

$$K_T = h_T(2^N - 1) / \int_{g((2^N - 1)\Delta)/T}^{\infty} \bar{q}_T(I) dI \quad (12)$$

where we used the fact that  $h_T(2^N - 1)$  is the proportion of saturated pixels in the image. Basically, the image histogram is used to model  $p(I)$  for values of the radiant power  $I$  that do not generate saturation. For larger values (the “tail” part), the parametric model is used. Note that if all pixels are unsaturated, then the tail part of the density vanishes because  $K_T = 0$ . Note that, ideally,  $\bar{p}_T(I)$  should not change with  $T$ . The dependence of  $\bar{p}_T(I)$  on  $T$  is due to the fact that both histogram and fitting distribution are computed from a single image taken at exposure  $T$ .

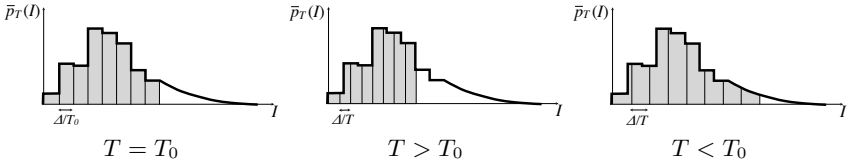
Using the density  $\bar{p}_T(I)$  as an approximation to  $p(I)$ , one may compute the expected error  $E_T$  for a given image, taken at exposure  $T$ , as by (6). Note that, in the case of linear sensor characteristic, term  $\Phi_T$  in the expression (8) of the granular error component  $E_T^g$  can be pre-computed, as it does not depend on the data. The term  $\text{Prob}(I < 2^N \Delta/T)$  in (8) simply represents the portion of non-saturated pixels, and can be easily computed from the histogram. The overload error can be computed by integration of the parametric function  $q_T(I)$  via numerical or Monte Carlo methods.

## 5 Predicting the Optimal Exposure

In the previous section we showed how to estimate the expected rendering error for a given image. Now we extend our theory to the *prediction* of the expected error when  $T$  varies. Formally, we will try to predict the exposure error  $E_T$  at any value of  $T$  based on the observation of just one image taken with (known) exposure  $T_0$ . We will do so by modeling  $p(I)$  with our composite model  $\bar{p}_{T_0}(I)$  in (11). Then, the expected error at any value of exposure  $T$  can be estimated via (6).



Here are some details about our prediction algorithm (see also Fig. 4). We begin by considering values of  $T$  larger than  $T_0$ . The granular component  $E_T^g$  is easily computed from (7) or (8). The overload component  $E_T^o$  is equal to the sum of two terms. The first term represents the “projection” of the histogram  $\bar{h}_{T_0}$  into the overload area, that is, for  $I$  between  $g(2^N \Delta)/T$  and  $g(2^N \Delta)/T_0$ . Integration of  $L(e_T)\bar{p}_T(I)$  over this segment amounts to a sum using histogram values. The second term is obtained by integration of the error weighed by the parametric density  $\bar{q}_T(I)$  for values of  $I$  above  $g(2^N \Delta)/T_0$ . This term can be computed offline and stored in a look-up table for various parameters of the parametric function used.



**Fig. 4.** A representation of the composite density function  $\bar{p}_T(I)$ , under three different exposures. The shaded are represents the granular region. The area of the density within the shaded areas represents  $\text{Prob}(I < (2^N - 1)\Delta/T)$ .

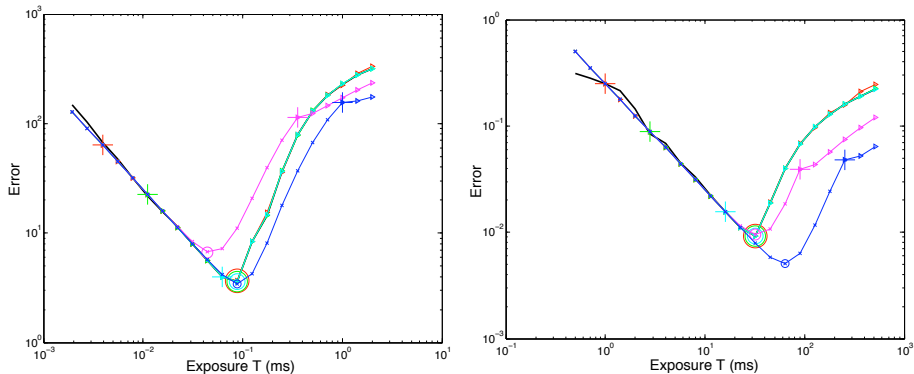
The predicted values for the granular and overload error components,  $\bar{E}_T^g$  and  $\bar{E}_T^o$ , can be expressed in a relatively simple form if the sensor’s characteristic  $f(I_T)$  is linear. In this case, the following identities hold:

$$\begin{aligned}
 T < T_0 : \quad & \bar{E}_T^g = \left[ (1 - h_{T_0}(2^N - 1)) + K_T \int_{(2^N - 1)/T_0}^{(2^N - 1)/T} \bar{q}_{T_0}(I) dI \right] \Phi_T \\
 & \bar{E}_T^o = K_T \int_{(2^N - 1)/T}^{\infty} L(I - (2^N - 1)/T) \bar{q}_{T_0}(I) dI \\
 T > T_0 : \quad & \bar{E}_T^g = \left[ \sum_{m=0}^{\text{floor}((2^N - 1)T_0/T)} h_{T_0}(m) \right] \Phi_T \\
 & \bar{E}_T^o = \sum_{m=\text{ceil}((2^N - 1)T_0/T)}^{2^N - 2} L(m/T_0 - (2^N - 1)/T) h_{T_0}(m) \\
 & \quad + K_T \int_{(2^N - 1)/T_0}^{\infty} L(I - (2^N - 1)/T) \bar{q}_{T_0}(I) dI
 \end{aligned} \tag{13}$$

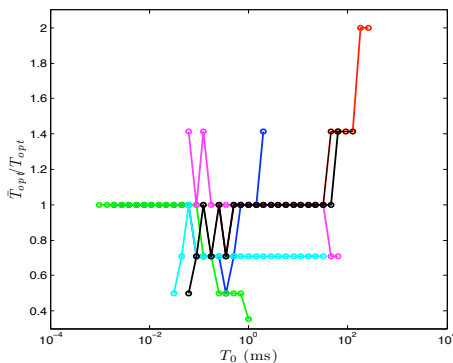
At this point, one may sample the estimated error  $\bar{E}_T = \bar{E}_T^g + \bar{E}_T^o$  for various values of  $T$  in order to find the estimated optimal exposure  $T_{\text{opt}}$ .

## 6 Experiments

We have used synthetically generated 8-bit images from a “ground truth” 12-bit image as discussed in the Appendix. The 12-bit images were taken with a Dragonfly 2 camera from Point Grey that has a very linear sensor characteristic  $f(I_T)$  [13,14,21]. The ground-truth 12-bit image is used for the computation of the ground-truth error  $E_T$  and of the optimal exposure  $T_{\text{opt}}$  that minimizes  $E_T$ .



**Fig. 5.** The ground-truth error  $E_T$  (black thick line) and the estimated errors  $\bar{E}_T$  starting from different values of  $T_0$  (thin colored lines, one line per choice of  $T_0$ ) for two different scenes. For each  $\bar{E}_T$  plot, the large '+' signs is placed at  $T_0$ : the whole plot is built from the analysis of the image at  $T_0$ . The large circles within each line represent the minimum of the plot, corresponding to the optimal exposure.



**Fig. 6.** Experiments with the proposed algorithm for estimating the optimal exposure  $T_{\text{opt}}$  from a single image. Each color represents a different scene. For each scene, the image exposed at  $T_0$  was used to find the estimate  $\bar{T}_{\text{opt}}$  using the algorithm in (13). The ratio  $\bar{T}_{\text{opt}}/T_{\text{opt}}$  is shown for each image with varying  $T_0$ . A value of  $\bar{T}_{\text{opt}}/T_{\text{opt}}$  equal to 1 means that the algorithm found the optimal exposure correctly.

Fig. 5 shows a number of estimated error plots  $\bar{E}_T$  as a function of exposure  $T$ . Each plot corresponds to a different starting point  $T_0$ . The thick black line is the “ground-truth” error  $E_T$ . Note that the left part of  $E_T$  has linear 45° slope in log-log space. This is because, for our choice of  $L(e_T) = |e_T|$ , the expected granular error is equal to  $\Delta/4T$  as mentioned in Sec. 4. However, for very small values of  $T$ , the granular error characteristic is not linear anymore, due to the fact that the “high rate” assumption does not hold true in these cases. The estimated error curves  $\bar{E}_T$  are generally good when the starting point  $T_0$  is in a location with few saturated pixels. The more challenging (and interesting)

situations are for larger  $T_0$ , chosen when a considerable portion of the image is saturated. In these cases, the estimated  $\bar{E}_T$  may fail to represent  $E_T$  in some areas, possibly leading to errors in the estimation of  $T_{\text{opt}}$ .

Results showing the quality of estimation of the optimal exposure from an image taken at exposure  $T_0$  for different values of the “start” exposure  $T_0$  are shown in Fig. 6 for various scenes. The optimal exposure  $T_{\text{opt}}$  for each scene was computed as discussed in Sec. 3. The plots in Fig. 6 show the ratio  $\bar{T}_{\text{opt}}/T_{\text{opt}}$ , which is indicative of the quality of the algorithm (values equal to 1 indicate correct estimation). Note that the different scenes had different optimal exposures  $T_{\text{opt}}$ . In most situations, our algorithm predicts the optimal exposure with good accuracy. However, when  $T_0$  is much smaller or higher than  $T_{\text{opt}}$ , the estimate may become incorrect. Small values of  $T_0$  mean that the histogram has little information due to high quantization step. Large values of  $T_0$  mean that the “start” image had a considerable number of saturated pixels.

## 7 Conclusion

We have presented a technique to estimate the optimal exposure from analysis of a single image. This approach relies on a definition of exposure quality based on the expected rendering error. Predicting the exposure quality for varying exposure times requires accessing the saturated (and thus unobservable) pixels. We proposed the use of a parametric distribution that fits the observable data, and allows reasoning about the saturated data. Our experiments show that this model enables accurate one-shot estimation of the correct exposure as long as the image being analyzed does not contain too many saturated pixels, or is not too under-exposed.

One main limitation of our approach is that we do not consider sensor noise and the use of gain as an exposure parameter. Future work will address both these issues, along with the possibility of using more accurate models for the distribution of irradiance in the image. Eventually, our algorithm will be integrated in a dynamic loop for real-time exposure control in video applications.

## Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. BES-0529435.

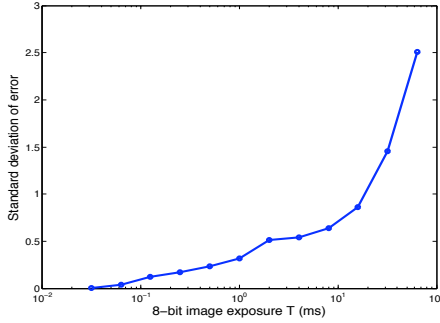
## References

1. Muramatsu, M.: Photometry device for a camera (1997)
2. Johnson, B.K.: Photographic exposure control system and method (1997)
3. Kremens, R., Sampat, N., Venkataraman, S., Yeh, T.: System implications of implementing auto-exposure on consumer digital cameras. In: Proceedings of the SPIE Electronic Imaging Conference, vol. 3650 (1999)

4. Shimizu, S., Kondo, T., Kohashi, T., Tsuruta, M., Komuro, T.: A new algorithm for exposure control based on fuzzy logic for video cameras. *IEEE Transactions on Consumer Electronics* 38, 617–623 (1992)
5. Yang, M., Crenshaw, J., Augustine, B., Mareachen, R., Wu, Y.: Face detection for automatic exposure control in handheld camera. In: *IEEE International Conference on Computer Vision Systems* (2006)
6. Nuske, S., Roberts, J., Wyeth, G.: Extending the range of robotic vision. In: *IEEE International Conference on Robotics and Automation* (2006)
7. Nourani-Vatani, N., Roberts, J.: Automatic exposure control. In: *Australasian Conference on Robotics and Automation* (2007)
8. Kuno, T., Matoba, N.: A new automatic exposure system for digital still cameras. *IEEE Transactions on Consumer Electronics* 44, 192–199 (1998)
9. Nayar, S., Branzoi, V.: Adaptive dynamic range imaging: optical control of pixel exposures over space and time. In: *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, pp. 1168–1175 (2003)
10. Schulz, S., Grimm, M., Grigat, R.R.: Optimum auto exposure based on high-dynamic-range histogram. In: *Proceedings of the 6th WSEAS International Conference on Signal Processing, Robotics and Automation (ISPRA'07)*, Stevens Point, Wisconsin, USA, pp. 269–274. World Scientific and Engineering Academy and Society, WSEAS (2007)
11. Grossberg, M., Nayar, S.: High dynamic range from multiple images: Which exposures to combine? In: *Proceedings of the ICCV Workshop on Color and Photometric Methods in Computer Vision, CPMCV* (2003)
12. Barakat, N., Hone, A., Darcie, T.: Minimal-bracketing sets for high-dynamic-range image capture. *IEEE Transactions on Image Processing* 17, 1864–1875 (2008)
13. Mitsunaga, T., Nayar, S.K.: Radiometric self calibration. *Proceedings of the IEEE Computer Vision and Pattern Recognition* 1, 1374 (1999)
14. Matsushita, Y., Lin, S.: Radiometric calibration from noise distributions. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR '07* (2007)
15. Gersho, A., Gray, R.: *Vector quantization and signal compression*. Kluwer Academic Publishers, Norwell (1991)
16. Anon, E., Grey, T.: *Photoshop for Nature Photographers: A Workshop in a Book*. Wiley, Chichester (2005)
17. Miller, R., Gong, G., Muñoz, A.: *Survival analysis*. Wiley, New York (1981)
18. Gross, Shulamith, T., LaiTze, L.: Nonparametric estimation and regression analysis with left-truncated and right-censored data. *Journal of the American Statistical Association* 91, 1166–1180 (1996)
19. Richards, W.: Lightness scale from image intensity distributions. *Applied Optics* 21, 2569–2582 (1982)
20. Ruderman, D.: The statistics of natural images. *Network: computation in neural systems* 5, 517–548 (1994)
21. Point Grey Research, Inc. *Point Grey Dragonfly2 Technical Specification* (2007)
22. Chen, T., El Gamal, A.: Optimal scheduling of capture times in a multiple capture imaging system. In: *Proc. SPIE*, vol. 4669, pp. 288–296 (2002)

## Appendix A

This Appendix describes the process used to generate synthetic 8-bit images at different exposure  $T$  starting from a 12-bit image. We used a Dragonfly 2 camera



**Fig. 7.** The standard deviation of the error  $B_{T,8} - B_{T,12}$  between the synthetic and the real 8-bit images as a function of the exposure  $T$

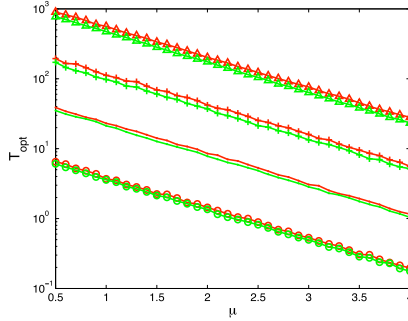
from Point Grey that has a very linear sensor characteristic  $f(I_T)$  [13,14,21] and provides images both at 12-bit and 8-bit pixel depth. Images were taken at 12 bits, carefully choosing the exposure  $T_0$  so as to best exploit the camera's dynamic range while avoiding saturation. Images with more than 0.1% pixels saturated were discarded. The brightness data  $B_{T_0,12}$  was dithered by adding white noise with uniform distribution between 0.5 and 0.5, then divided by  $2^{12-8} = 16$ . This quantity is multiplied by  $T_0/T$  and then quantized with  $\Delta = 1$  in order to obtain the equivalent 8-bit image for exposure  $T$ , named  $B_{T,12}$ . In this way, multiple 8-bit synthetic images can be obtained for different exposure value  $T$ .

In order to assess the error that should be expected with this processing, we took a number of real 8-bit images ( $B_{T,8}$ ) of a static scene with various exposures  $T$ , and then compared them with their synthetic counterparts obtained by synthesis from a 12-bit image of the same scene. The results, in terms of standard deviation of the error  $B_{T,8} - B_{T,12}$ , are plotted in Fig. 7. As expected, the error increases with increasing exposure  $T$  (remember that the dithered 12-bit image is multiplied by  $T/T_0$ ). Note that for most of the exposure, the standard deviation stays below 1 (PSNR = 48 dB), and it reaches a maximum of about 2.5 (PSNR = 40 dB).

## Appendix B

In this Appendix we consider the effect of photon noise in the determination of the optimal exposure. For a given value of irradiant power  $I$  and of exposure  $T$ , the variance of the rendering error due to photon noise is equal to  $\sigma_{\text{pht}}^2 = qI/T$ , where  $q$  is the electrical charge of an electron, and  $I$  is measured in terms of photoelectronic current [22]. Let  $N_{\text{sat}}$  be the full well capacity of the sensor. It is reasonable to assume that (in the absence of amplification gain), the quantizer saturates when the sensor saturates, that is,  $\Delta(2^N - 1) = qN_{\text{sat}}$ .

When computing the optimal exposure, both quantization and photon noise should be considered. Unfortunately, the resulting rendering error depends on



**Fig. 8.** Monte Carlo simulation of  $T_{\text{opt}}$  for  $L(e_T) = |e_T|$ , red: photon noise and quantization noise considered, green: quantization noise only, triangles:  $\sigma = 0.5$ , plus marks:  $\sigma = 1$ , dots:  $\sigma = 1.5$ , circles:  $\sigma = 2.0$

the characteristics of the irradiance distribution. For example, one can easily derive the expression of the quadratic norm of the granular error under the assumption of linear sensor characteristic:

$$E_T^g = \frac{q^2 N_{\text{sat}}}{T^2} \left( \frac{N_{\text{sat}}}{12(2^N - 1)^2} + \frac{TE[I]}{qN_{\text{sat}}} \right) \quad (14)$$

where  $E[I]$  is the average value of the irradiant power. The second term within the parenthesis is a number that represents the “average degree of saturation”. In particular, when no pixel is saturated, then  $TE[I]/qN_{\text{sat}} < 1$ . Note that for (14), the *relative* effect of the term due to photon noise is increased as  $N_{\text{sat}}$  *decreases*. Unfortunately, computation of the average error under different metrics (in particular,  $L(e_T) = |e_T|$ , which is the metric considered in this paper) requires knowledge of the probability density function of the irradiance  $I$ .

Fig. 8 shows results of a Monte Carlo simulation to find  $T_{\text{opt}}$  for  $L(e_T) = |e_T|$ , assuming a log-normal distribution with parameters  $\mu$  and  $\sigma$ . Fixed parameters in the simulation are  $N_{\text{sat}} = 6000$  (representing a sensor with a relatively small well capacity) and bit depth  $N = 8$ . Two million points are sampled to generate error values for each  $T$  used in the search for  $T_{\text{opt}}$ . Results shown in Fig. 8 suggest that the ratio of  $T_{\text{opt}}$  with photon noise considered, relative to  $T_{\text{opt}}$  where it is not, is a small positive value.  $T_{\text{opt}}$  appears more sensitive to a choice of  $L$  or change in the irradiance distribution at the sensor than to the consideration of photon noise.