High Dynamic Range Imaging using Quanta Image Sensors

Abhiram Gnanasambandam[1], Jiaju Ma[2] and Stanley H. Chan[1]

[1] School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907

[2] Gigajot Technology Inc., Pasadena, CA 91107, USA

Abstract

An algorithmic solution is proposed for reconstructing high dynamic range (HDR) images from single-bit and multi-bit Quanta Image Sensor (QIS). Given a space-time cubicle of the QIS data, the algorithm partitions the frames into groups of different exposures. After summation and denoising, the resulting frames are combined to form the HDR image. The combination weights are determined according to a new theoretical result showing how signal to noise changes with the exposure. The new method is compared with conventional CMOS-based HDR image reconstruction methods on both synthetic and real QIS data.

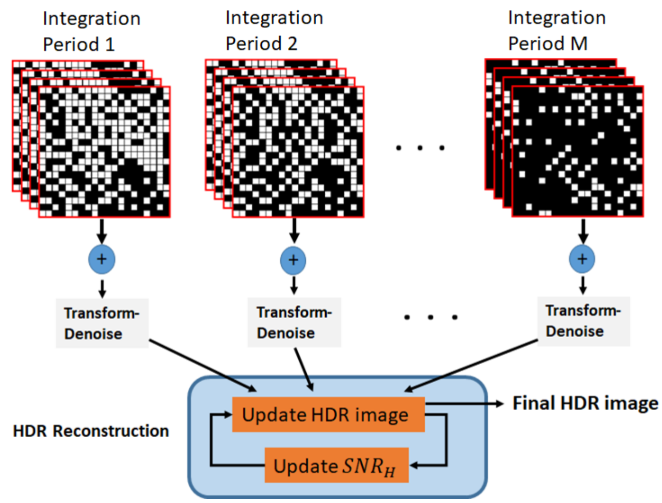

Figure 1. HDR Reconstruction pipeline. The raw QIS frames are summed and denoised. Then the denoised images are combined according to the weights proportional to SNR_H. (See Section 3 for details.)

1. Introduction

Quanta Image Sensor (QIS) is a single photon image sensor with extremely small full well capacity. Since its introduction in 2005, the feasibility of QIS has been reported in many occasions, e.g., [1, 2]. The latest QIS prototype in [2] can achieve a read noise below $0.25e^-$ rms at room temperature, and frame rate beyond 1000 frames per second. This high level of photon sensitivity, low read noise, and high speed has made QIS an ideal sensor for low-light applications.

The subject of this paper is to use QIS for high dynamic range imaging. Compared to conventional CMOS image sensors (CIS) which acquire multiple frames and fuse them using linear averaging, QIS acquires a space-time data cubicle. To reconstruct the high dynamic range image, the algorithm has to first sum the frames, denoise, and then form an average. However, because of the quantized Poisson statistics of QIS, reconstruction methods for conventional CIS are not applicable. In particular, weighted averages based on estimating the pixel variance are extremely sensitive to noise which are not suitable for QIS.

The contribution of this paper is to propose a high dynamic range image reconstruction algorithm for QIS. The paper consists of two parts. First, we theoretically derive the dynamic range offered by QIS and compare it with CIS. This provides a foundation of how much dynamic range we can expect from QIS. Second, we present a new reconstruction pipeline as illustrated in Figure 1. The reconstruction uses the theoretical results to predict the optimal combination weights. We compare the new algorithm with other state-of-the-art image reconstruction methods.

2. Dynamic Range - QIS vs. CIS

We start by reviewing the concept of dynamic range. Dynamic range of a sensor is the range of exposures that the signal-to-noise ratio is sustained before it drops below 1 (i.e., signal = noise). The signal-to-noise ratio considered in this paper follows from [3, 4, 5], which is known as the exposure referred signal-to-noise ratio SNR_H. Our goal is to analytically derive SNR_H, and use the results to design the reconstruction algorithm.

The mathematical model of the sensor is as follows. Denote c as the exposure (photons per second) of a particular pixel in the scene, and let τ be the integration time. The number of photons X reaching the sensor is a Poisson random variable $X \sim \mathrm{Poisson}(\tau c)$. The full-well capacity is assumed to be ℓ. Thus, the observed signal is a random variable B such that $B = X$, if $X < \ell$, and $B = \ell$, if $X \geq \ell$.

The exposure referred signal-to-noise SNR_H (for

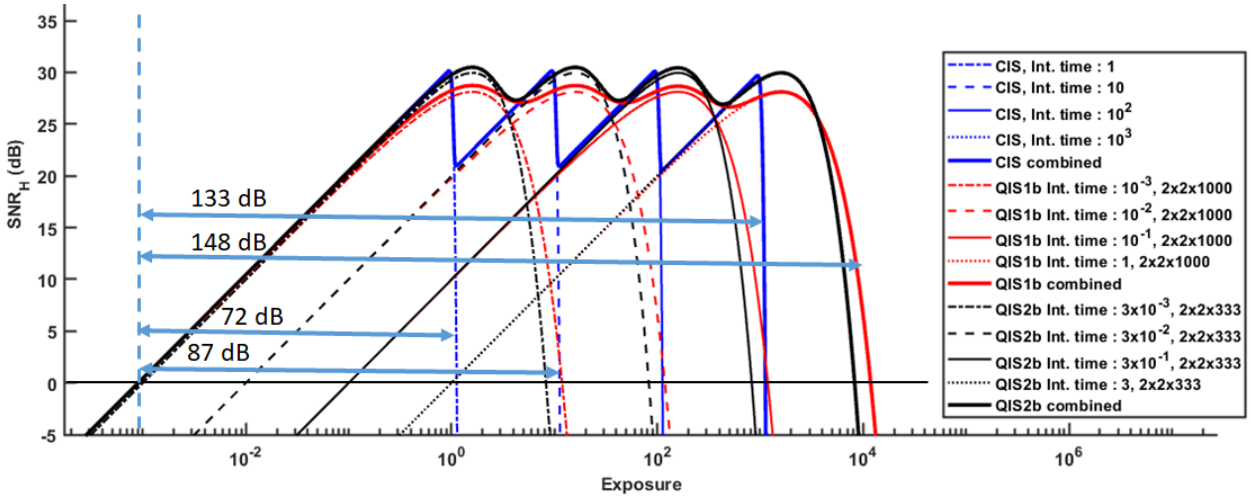

Figure 2. **Comparison of Exposure referred signal-to-noise (SNR$_H$) for CIS and QIS**. (QIS1b: Single-bit QIS, QIS2b: Two-bit QIS). Ideal QIS and CIS are considered, i.e. the read noise is assumed to be zero in both cases. CIS is assumed to have a full well capacity of 4000 electrons. For single-bit QIS, binning of 1000 frames with 2×2 spatial oversampling is assumed. For 2-bit QIS, $T = 333$ frames with 2×2 spatial oversampling is assumed. QIS has a larger dynamic range than CIS for each exposure, and has a more consistent SNR over the entire range when the low dynamic range images are combined to get a single high-dynamic range image.

both CIS and QIS) is defined as

$$\mathrm{SNR}_H \stackrel{\text{def}}{=} \frac{\theta \frac{\partial \mu_B}{\partial \theta}}{\sigma_B}, \tag{1}$$

where $\mu_B = \mathbb{E}[B]$ is the mean, $\sigma_B^2 = \mathrm{Var}(B)$ is the variance, and $\theta \stackrel{\text{def}}{=} \tau c$ is the average number of photons seen in time τ. The exact expressions of the quantities in (1) are given by the theorem below, where the proof is skipped due to space limit.

Theorem 1 *The quantities μ_B and σ_B are*

$$\mu_B = \theta(\Psi_{\ell-1}(\theta)) + \ell(1 - \Psi_\ell(\theta)), \tag{2}$$

$$\sigma_B^2 = \ell^2 - \sum_{q=0}^{\ell-1}((2q+1)\Psi_{q+1}(\theta)) - \mu_B^2, \tag{3}$$

where $\Psi_q(\theta) = \sum_{k=0}^{q-1} \frac{\theta^k e^{-\theta}}{k!}$ the incomplete gamma function. Consequently, the partial derivative $\partial\mu_B/\partial\theta$ is

$$\frac{\partial \mu_B}{\partial \theta} = \begin{cases} \Psi_{\ell-1}(\theta) - \theta \frac{e^{-\theta}\theta^{\ell-2}}{(\ell-2)!} + \ell \frac{e^{-\theta}\theta^{\ell-1}}{(\ell-1)!}, & \ell > 1 \\ e^{-\theta}, & \ell \leq 1. \end{cases} \tag{4}$$

This exact analytic expression for arbitrary ℓ is a new result compared to previous work, e.g., [3, 4]. Using the expressions in the theorem, we can plot SNR$_H$ as a function of the exposure. Figure 2 illustrates the behavior of SNR$_H$ for CIS and QIS. For CIS, we set $\ell = 4000$, and for QIS we set ℓ to either $\ell = 1$ or $\ell = 3$, which are typical values of the sensors. The results in Figure 2 show that a single-exposure of QIS (87dB) is already larger than that of CIS (72dB). The combined-exposure is 148dB compared to 133dB. Therefore, regardless if we use single or multiple integration times, QIS offers a greater dynamic range. The result also indicates that 1-bit QIS has a larger dynamic range than a 2-bit QIS. This is a direct consequence of the decrease in over-exposure latitude as the number of bits increases [3].

To visually compare the images captured by a CIS and a QIS, we show in Figure 3 a simulated experiment. The images are simulated by setting the maximum illumination to 6×10^6 photons per pixel per second, and using CIS and QIS to measure the photons. For evaluation, we report the peak-signal-to-noise ratio (PSNR) which is proportional to the negative log of the mean squared error between the observed and the ground truth. The result again shows that CIS has a smaller contrast due to the low dynamic range.

3. HDR Reconstruction for QIS

Using the analytic expressions derived in Theorem 1, we now present a new HDR reconstruction algorithm for QIS. Consider a static scene captured by a QIS usings N frames of 1-bit or multi-bit measurements. To construct a HDR image, the N frames are grouped into M groups where each group corresponds to a different integration time. Within each group, the frames are summed to generate a single output, thus giving $\widetilde{S}_m$, $m = 1, \ldots, M$. Since each $\widetilde{S}_m$ is a simple sum, there will be Poisson noise. To denoise, we follow [4] by using

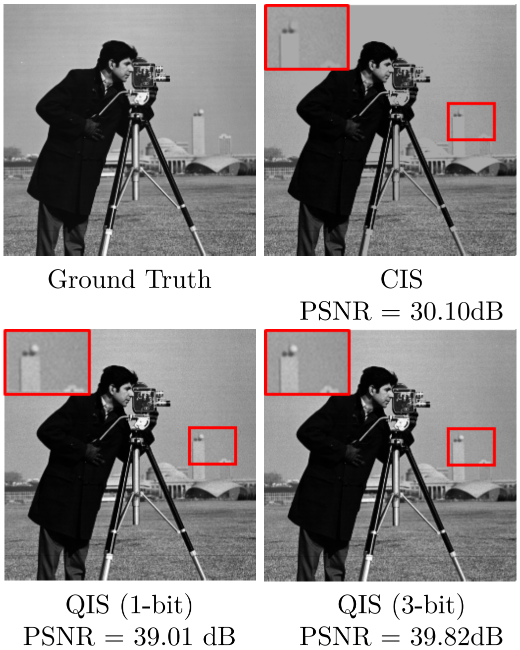

Figure 3. Dynamic range of QIS and CIS. The image is setup in such a way that the maximum illumination of the image is 6×10^6 photons per pixel per second. We assume that CIS can count upto 4000 electrons, single bit QIS - 1 electon and 3 bit QIS - 7 electrons. The exposure times are: CIS - 1ms, single bit QIS - $0.25\mu s$, and 3 bit QIS - $1.75\mu s$. We use 1 CIS frame, 4000 frames for single bit QIS and 571 frames for 3 bit QIS to ensure the equivalent total photon counts are the same for different imaging devices

a Transform-Denoise procedure. Denote the denoised frames as $S_m = \mathsf{denoise}(\widetilde{S}_m)$.

Given the low dynamic range images $\{S_1, \ldots, S_M\}$, we construct the high dynamic range image $\widehat{c}$ by formulating a weighted average:

$$\widehat{c}(i,j) = \sum_{m=1}^{M} \frac{1}{\tau_m} w_m(i,j) S_m(i,j). \tag{5}$$

Here, τ_m is the equivalent exposure time used when constructing S_m, and $w_m(i,j)$ is the weight of the m-th exposure image. The running index (i,j) denotes the pixel. Thus, (5) is a per-pixel weighted average.

Readers familiar with the HDR literature will notice that (5) is just the classical weighted averaging. However, in the classical CIS based HDR reconstruction, the combination weight is inversely proportional to the local image variances [6]. For QIS, such variance-based weight can cause serious problems, because in bright regions the noise is squeezed and so the variance is close to zero. The following new theorem gives the optimal combination weights for QIS.

Theorem 2 *The optimal weights $w_m(i,j)$ (for pixel (i,j)) which maximize the SNR_H of the signal $\widehat{c}(i,j)$ in (5) is given by*

$$w_m(i,j) = \frac{(SNR_{H,m}(i,j))^2}{\sum_{m=1}^{M} (SNR_{H,m}(i,j))^2}, \tag{6}$$

where $SNR_{H,m}$ is m-th SNR_H corresponding to the m-th exposure (curve) in Figure 2.

The reason why Theorem 2 resolves the vanishing variance problem is that it optimizes the exposure-referred SNR whereas the classical algorithms optimize the output SNR. The latter does not apply to QIS. The algorithm of the reconstruction is shown in Algorithm 1. It is an iterative algorithm. We need to update the image and estimate the SNR simultaneously because calculating the SNR requires the latent image c. The latent image c needs to be estimated.

Algorithm 1 Reconstruction Algorithm

1. Capture N frames at M different integration times.
2. Compute summed LDR images $\widetilde{S}_m$, $m = 1, \ldots M$.
3. $S_m = \mathsf{denoise}(\widetilde{S}_m)$, $m = 1, \ldots M$.
4. Initialize $\text{SNR}_{H,m}(i,j) = 1$
5. Update the weights $w_m(i,j)$ according to (6).
6. Estimate the HDR image $\widehat{c}(i,j)$ according to (5).
7. Update $\text{SNR}_{H,m}(i,j)$ using $\widehat{c}(i,j)$ and (1).
8. Repeat 5, 6, 7 until convergence.

To evaluate the effectiveness of the new algorithm we compare it with two existing CIS-based image reconstruction methods: A built-in function in MATLAB, and linear sum used in [7]. Figure 4 shows the simulation result on an existing synthetic HDR dataset by Stanford. Since the inputs are simulated according to QIS statistics, both existing methods fail to reconstruct the image. In Figure 5, we show a set of real images captured using the Gigajot QIS pathfinder camera, at 1-bit mode and 3-bit mode. The result shows that the method can effectively recover the dynamic range while smoothing out the noise.

4. Conclusion

A new image reconstruction method is proposed for high dynamic range imaging using QIS. The algorithm is based on a linear weighted averaging method, where the weights are determined according to a new theoretical study on how the exposure referred signal to noise ratio changes with the exposure. Experimental results demonstrate the effectiveness of the new method on both synthetic and real QIS datasets.

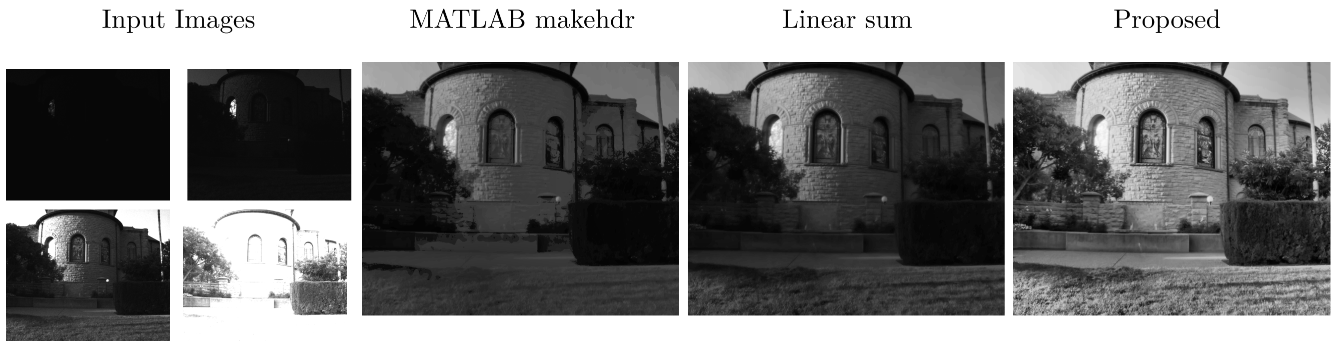

Figure 4. Simulated Data. One thousand frames of simulated single bit data were obtained at 4 different exposures each. These images are combined using 3 different methods to get HDR images. It is clearly visible that the proposed method outperforms the other two methods. MATLAB's tonemap is used to display the image.

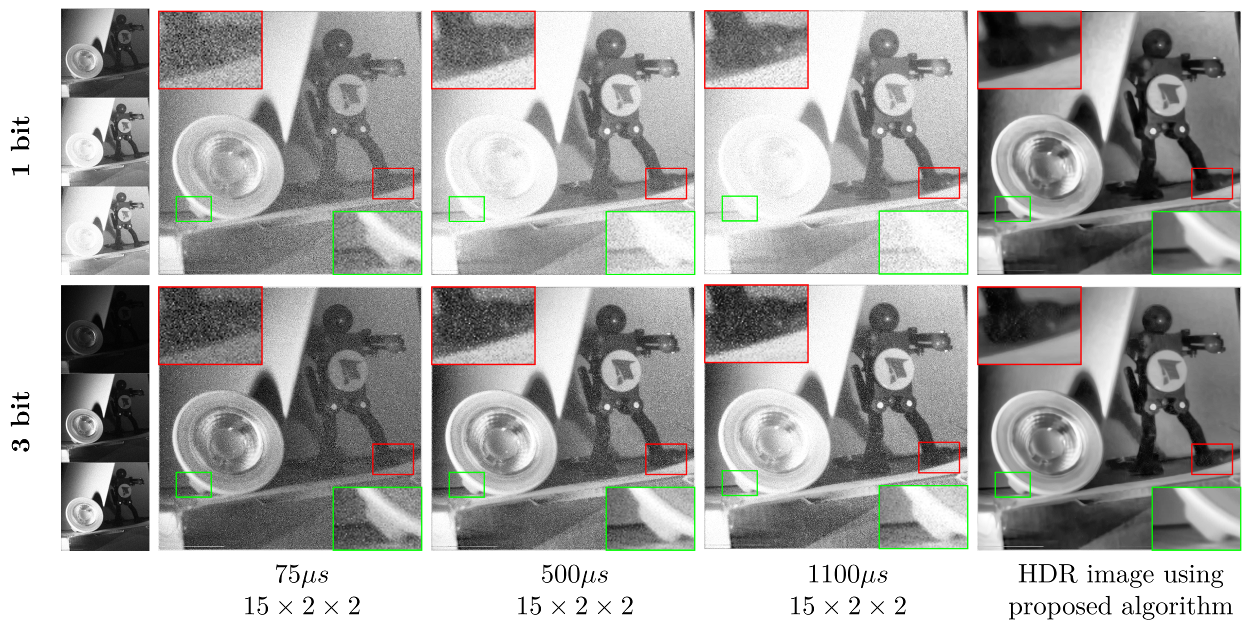

Figure 5. Real Data. In this experiment, we obtain 15 frames each at 3 different exposures - $75\mu s$, $500\mu s$, and $1100\mu s$ in 1 bit and 3 bit modes using a 1 Mpixel QIS image sensor. The average number of photons per pixel per frame are $0.5, 2.1$ and 3.3 for the 3 exposures respectively. Spatial oversampling of 2×2 is used. The proposed HDR reconstruction algorithm is used to obtain the final HDR image. MATLAB's tonemap is used to display the images. The first column are images before tone-map.

References

[1] E. R. Fossum, "Some thoughts on future digital still cameras," *Image sensors and signal processing for digital still cameras*, p. 305, 2006.

[2] J. Ma, S. Masoodian, D. A. Starkey, and E. R. Fossum, "Photon-number-resolving megapixel image sensor at room temperature without avalanche gain," *OSA Optica*, vol. 4, pp. 1474–1481, Dec 2017.

[3] E. R. Fossum, "Modeling the performance of single-bit and multi-bit quanta image sensors," *IEEE Journal of the Electron Devices Society*, vol. 1, no. 9, pp. 166–174, 2013.

[4] S. H. Chan, O. A. Elgendy, and X. Wang, "Images from bits: Non-iterative image reconstruction for quanta image sensors," *Sensors*, vol. 16, no. 11, p. 1961, 2016.

[5] E. R. Fossum, "Multi-bit quanta image sensors," in *Proc. Int. Image Sens. Workshop*, pp. 292–295, 2015.

[6] M. Granados, B. Ajdin, M. Wand, C. Theobalt, H.-P. Seidel, and H. P. Lensch, "Optimal hdr reconstruction with linear digital cameras," in *CVPR*, pp. 215–222, IEEE, 2010.

[7] N. Dutton, T. Al Abbas, I. Gyongy, F. Mattioli Della Rocca, and R. Henderson, "High dynamic range imaging at the quantum limit with single photon avalanche diode-based image sensors," *Sensors*, vol. 18, no. 4, p. 1166, 2018.