SIMULTANEOUS IMAGE FORMATION AND MOTION BLUR RESTORATION VIA MULTIPLE CAPTURE

Xinqiao Liu and Abbas El Gamal

Information Systems Laboratory, Department of Electrical Engineering
Stanford University, Stanford CA 94305
chiao@stanford.edu, abbas@isl.stanford.edu

ABSTRACT

Advances in CMOS image sensors enable fast image capture, which makes it possible to capture multiple images within a normal exposure time. An algorithm that takes advantage of this capability by simultaneously constructing a high dynamic range image and performing motion blur restoration from multiple image captures is described. The algorithm comprises two main procedures – photocurrent estimation and motion/saturation detection. It operates completely locally – each pixel's final value is computed using only its captured values, and recursively, requiring the storage of only a constant number of values per pixel independent of the number of images captured. These modest computational and storage requirements make it feasible to integrate all needed memory and processing with the image sensor on a single CMOS chip. Simulation results demonstrate the enhanced SNR, dynamic range, and the motion blur restoration obtained using our algorithm.

1. INTRODUCTION

Blurring due to object or camera motion during image capture can cause substantial degradation in image quality. As a result, a great deal of research has been conducted on developing methods for restoring motion blurred images, *e.g.*, see [1]. These methods make certain assumptions on the blurring process, the ideal image, and the noise. Various image processing techniques are then used to identify the blur and restore the image. However, due to the lack of sufficient knowledge of the blurring process and the ideal image, the developed image blur restoration methods have limited applicability and their computational burden can be quite substantial.

Recent advances in CMOS image sensor technology enable digital high speed capture up to thousands of frames per second [2, 3]. This benefits traditional high speed imaging applications and enables new imaging enhancement capabilities such as multiple capture for increasing the sensor dynamic range [4]. In this scheme, multiple images are captured at different times within the normal exposure time. Shorter exposure time images capture brighter areas of the scene, while longer exposure time images capture darker areas of the scene. The images are then combined into a single high dynamic range image.

In this paper we propose to use this multiple capture capability to simultaneously form a high dynamic range image and reduce or eliminate motion blur. Our algorithm operates completely locally – each pixel's final value is computed using only its captured values. Moreover, our method can operate recursively, requiring the storage of only a constant number of values per pixel independent of the number of images captured. These modest computational and storage requirements make it feasible to integrate all the processing and memory needed with the image sensor on the same CMOS chip [5].

In the next section we briefly describe the image sensor pixel operation and statistical model. In Section 3 we present our algorithm. Simulation results are presented in Section 4.

2. PIXEL OPERATION AND MODEL

The area image sensor used in an analog or digital camera consists of a 2-D array of pixels. During capture each pixel converts incident light into photocurrent $i_{ph}(t)$, for $0 \leq t \leq T$, where T is the exposure time. This process is quite linear, and thus $i_{ph}(t)$ is a good measure of incident light intensity. Since the photocurrent is too small to measure directly, it is integrated onto a capacitor and the charge $Q(T)$ (or voltage) is read out at the end of exposure time T. Dark current i_{dc} and additive noise corrupt the output signal charge. The noise can be expressed as the sum of three independent components, (i) shot noise $U(T) \sim \mathcal{N}(0, q\int_0^T (i_{ph}(t) + i_{dc})dt)$, where q is the electron charge, (ii) readout circuit noise $V(T)$ (including quantization noise) with zero mean and variance σ_V^2, and (iii) reset noise $C \sim \mathcal{N}(0, \sigma_C^2)$ caused by resetting the capacitor prior to capture. Thus the output charge from a pixel can be expressed as

$$Q(T) = \int_0^T (i_{ph}(t) + i_{dc})dt + U(T) + V(T) + C, \quad (1)$$

provided $Q(T) \leq Q_{sat}$, the saturation charge, also referred to as *well capacity*. If photocurrent is constant over exposure time, SNR can be expressed as

$$\mathrm{SNR}(i_{ph}) = 10\log_{10}\frac{(i_{ph}T)^2}{q(i_{ph} + i_{dc})T + \sigma_V^2 + \sigma_C^2} \quad (2)$$

Note that SNR increases with i_{ph}, first at 20dB per decade when reset and readout noise variance dominates, and then at 10dB per decade when shot noise variance dominates. SNR also increases with T. Thus it is always preferred to have the longest possible exposure time. Saturation and change in photocurrent due to motion, however, makes it impractical to make exposure time too long.

We illustrate the effect of saturation and motion and how multiple capture may mitigate their effects via the examples in Figures 1 and 2. The first plot in Figure 1 represents the case of a constant low light, where photocurrent can be well estimated from $Q(T)$. The second plot represents the case of a constant high light, where $Q(T) = Q_{sat}$ and the photocurrent cannot be well estimated from $Q(T)$. The third plot is for the case when light changes during exposure time, *e.g.*, due to motion. In this case, photocurrent at the beginning of exposure time $i_{ph}(0)$ again cannot be well estimated from $Q(T)$.

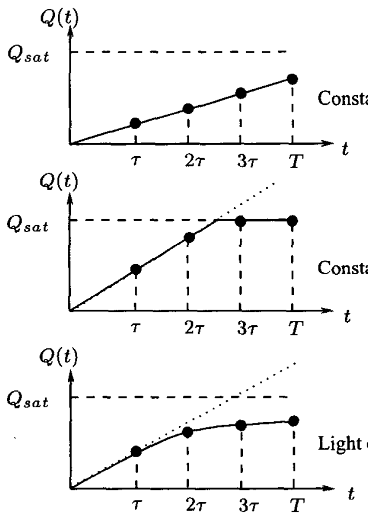

Figure 1: $Q(t)$ vs. t for three lighting conditions.

To avoid saturation and the change of $i_{ph}(t)$ due to motion, exposure time may be shortened, *e.g.*, to τ in Figure 1. Since in conventional sensor operation, exposure time is set globally for all pixels, this results in reduction of SNR, especially for pixels with low light. This point is further demonstrated by the images in Figure 2, where a bright square object moves diagonally across a dark background. If exposure time is set long to achieve high SNR, it results in significant motion blur as shown in image (b) of Figure 2. On the other hand if exposure time is set short, SNR deteriorates resulting in the noisy image (c) of Figure 2.

Recent advances in CMOS image sensor technology makes it possible to capture and nondestructively read out, *i.e.*, without reset, multiple images within a normal exposure time [4]. Using this multiple capture capability one can in effect adapt the pixel exposure time to the lighting condition. For the examples in Figure 1, if we capture four images at τ, 2τ, 3τ, and $T = 4\tau$, the photocurrent for the high light pixel can be estimated using the images captured at τ and 2τ, while for the low light pixel can be estimated using the four images. Motion blur in the third case can be reduced by using the four captures to estimate photocurrent at the beginning of exposure time $i_{ph}(0)$. In the following section we derive an optimal recursive pixel photocurrent estimator from multiple captures and show how motion blur can be detected and reduced. Image (d) in Figure 2 of the moving square object is produced using four captures and the algorithm described in the next section.

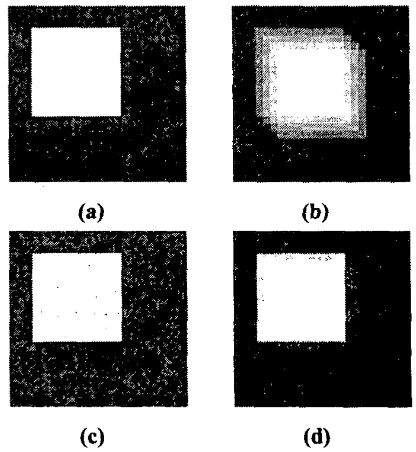

Figure 2: (a) Ideal image. (b) Long exposure time image. (c) Short exposure time image. (d) Image produced by applying our algorithm to 4 captures.

3. IMAGE FORMATION AND BLUR RESTORATION

Our image formation and motion blur restoration algorithm operates on n image captures at $\tau, 2\tau, \ldots, n\tau = T$ as follows:

1. Capture first image, set $k = 1$.
2. For each pixel: Use the *current estimation algorithm* to find the photocurrent estimate $\hat{I}_1$ from $Q(\tau)$.
3. Capture next image.
4. For each pixel: Use the *motion detection algorithm* to check if motion/saturation has occurred
   i. *Motion detected*: Set final photocurrent estimate $\hat{I}_n = \hat{I}_k$.
   ii. *No Motion detected or decision deferred*: Use the *current estimation algorithm* to find $\hat{I}_{k+1}$ from $Q((k+1)\tau)$ and $\hat{I}_k$ and set $k = k + 1$.
5. Repeat steps 3 and 4 until $k = n$.

In the following subsection we describe a recursive algorithm for estimating photocurrent, and in subsection 3.2 we describe a heuristic algorithm for performing motion detection.

3.1. Photocurrent Estimation

We simplify the derivations by neglecting dark current and reset noise. A detailed derivation of the estimation algorithm with reset noise included is presented in [6]. We also assume here that each pixel's photocurrent i is constant. In the following subsection we deal with the case when photocurrent changes during capture due to motion. Assuming n pixel values captured at $\tau, 2\tau, \ldots, n\tau = T$, the kth output charge is given by

$$Q_k = ik\tau + \sum_{j=1}^{k} U_j + V_k, \text{ for } 1 \leq k \leq n, \tag{3}$$

where V_k is the readout noise of the kth capture, U_j is the shot noise generated during time interval $((j-1)\tau, j\tau]$, and V_k and the U_js are all zero mean and independent, with

$$\begin{aligned} E(V_k^2) &= \sigma_V^2 > 0, \text{ for } 1 \le k \le n, \\ E(U_j^2) &= qi\tau, \text{ for } 1 \le j \le k. \end{aligned} \tag{4}$$

Define

$$\tilde{I}_k = \frac{Q_k}{k\tau} = i + \frac{\sum_{j=1}^{k} U_j}{k\tau} + \frac{V_k}{k\tau}, \text{ for } 1 \le k \le n. \tag{5}$$

At time $k\tau$, we wish to find the best linear unbiased MSE estimate of the parameter i given $\{\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_k\}$, *i.e.*, coefficients $a_1, a_2, \ldots, a_k$ such that

$$\hat{I}_k = \frac{1}{g_k} \sum_{j=1}^{k} a_j \tilde{I}_j, \tag{6}$$

where $g_k = \sum_{j=1}^{k} a_j$, minimizes

$$\Phi_k^2 = E(\hat{I}_k - i)^2, \tag{7}$$

subject to

$$E(\hat{I}_k) = i.$$

In recursive form, after the kth capture, the $(k+1)$st optimal estimate can be written as

$$\begin{aligned} \hat{I}_1 &= \tilde{I}_1, \\ \hat{I}_{k+1} &= \hat{I}_k + h_{k+1}(\tilde{I}_{k+1} - \hat{I}_k), \end{aligned} \tag{8}$$

where the gain h_{k+1} is given by

$$h_{k+1} = \frac{a_{k+1}}{g_k + a_{k+1}}, \tag{9}$$

and the a_k coefficients are given by

$$\begin{aligned} a_1 &= 1 \\ a_{k+1} &= (k+1)(1 + \tfrac{a_k}{k} + \tfrac{qi\tau}{\sigma_V^2} w_k), \end{aligned} \tag{10}$$

where $w_k = \sum_{j=1}^{k} \frac{a_j}{j}$.

The MSE error Φ_k^2 can also be expressed in a recursive form as

$$\begin{aligned} \Phi_1^2 &= \tfrac{qi}{\tau} + \tfrac{\sigma_V^2}{\tau^2} \\ \Phi_{k+1}^2 &= \tfrac{g_k^2}{g_{k+1}^2} \Phi_k^2 + \tfrac{1}{g_{k+1}^2}((2a_{k+1}g_k + a_{k+1}^2)\tfrac{qi}{(k+1)\tau} \\ &\quad + a_{k+1}^2 \tfrac{\sigma_V^2}{(k+1)^2\tau^2}) \end{aligned} \tag{11}$$

Note that computing the coefficients and the MSE require knowledge of i, which is not known. To solve this problem we replace i by its latest estimate. This, of course, makes the estimator suboptimal. Figure 3 shows that $\hat{I}_k$ improves SNR over $\tilde{I}_k$ by around 6dB at low light, when readout noise dominates. The SNR improvement is not significant at high light where shot noise dominates.

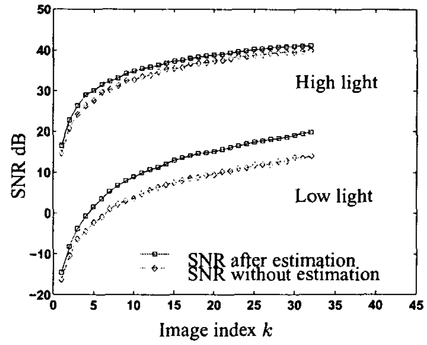

Figure 3: SNR vs. image index k with and without estimation.

3.2. Motion/Saturation Detection

The derivation of the recursive linear estimation algorithm in the previous section assumed that $i(t)$ is constant and that saturation does not occur before $k\tau$. In this section we describe an algorithm for detecting change in the value of $i(t)$ due to motion or saturation before the new image is used to update the photocurrent estimate. Since the statistics of the noise are not completely known and no motion model is specified, it is not possible to derive an optimal detection algorithm. Our algorithm is, therefore, based on heuristics. By performing the detection step prior to each estimation step we form a blur free high dynamic range image from the n captured images.

The algorithm operates on each pixel separately. After the kth capture, the best MSE linear estimate of i, $\hat{I}_k$, and its MSE, Φ_k^2, are computed as detailed in the previous subsection. If the current stays constant, the next observation $\tilde{I}_{k+1}^{pre}$ would be

$$\tilde{I}_{k+1}^{pre} = i + \frac{\sum_{j=1}^{k+1} U_j}{(k+1)\tau} + \frac{V_{k+1}}{(k+1)\tau}, \tag{12}$$

and the best predictor of $\tilde{I}_{k+1}^{pre}$ is $\hat{I}_k$ with the prediction MSE given by

$$\begin{aligned} \Delta_{k+1}^2 &= E\left((\tilde{I}_{k+1}^{pre} - \hat{I}_k)^2 | \hat{I}_k\right) \\ &= \Phi_k^2 - \tfrac{q\hat{I}_k}{(k+1)\tau} + \tfrac{\sigma_V^2}{(k+1)^2\tau^2}. \end{aligned} \tag{13}$$

Thus to decide whether the input signal i changed between time $k\tau$ and $(k+1)\tau$, we compare $\tilde{I}_{k+1} = \frac{Q_{k+1}}{(k+1)\tau}$ with $\hat{I}_k$. A simple decision rule would be to declare that motion has occurred if

$$|\tilde{I}_{k+1} - \hat{I}_k| \ge m\Delta_{k+1}, \tag{14}$$

and to use $\hat{I}_k$ as the final estimate of i, otherwise to use $\tilde{I}_{k+1}$ to update the estimate of i, *i.e.*, $\hat{I}_{k+1}$. The constant $m > 0$ is chosen to achieve the desired tradeoff between SNR and motion blur. The higher m the more motion blur if i changes with time, but also the higher the SNR if i is a constant, and vice versa.

One potential problem with this "hard" decision rule is that gradual drift in i can cause accumulation of estimation error resulting in undesired motion blur. To address this problem we propose the following "soft" decision rule.

Motion detection algorithm: For each pixel, after the $(k+1)$st capture:

1. If $|\tilde{I}_{k+1} - \hat{I}_k| \leq m_1\Delta_{k+1}$, then declare that *no motion detected.* Use $\tilde{I}_{k+1}$ to update $\hat{I}_{k+1}$ and set $L^+ = 0$, $L^- = 0$.
2. If $|\tilde{I}_{k+1} - \hat{I}_k| \geq m_2\Delta_{k+1}$, $L^+ = l_{max}$, or $L^- = l_{max}$, then declare that *motion detected.* Use $\hat{I}_k$ as the final estimate of i.
3. If $m_1\Delta_{k+1} < \tilde{I}_{k+1} - \hat{I}_k < m_2\Delta_{k+1}$, then *defer the decision* and set $L^+ = L^+ + 1$, $L^- = 0$.
4. If $-m_2\Delta_{k+1} < \tilde{I}_{k+1} - \hat{I}_k < -m_1\Delta_{k+1}$, then *defer the decision* and set $L^- = L^- + 1$, $L^+ = 0$.

The counters L^+, L^- record the number of times the decision is deferred, and $0 < m_1 < m_2$ and l_{max} are chosen to tradeoff SNR with motion blur.

4. SIMULATION RESULT

Figure 4 plots SNR versus i for conventional sensor operation, where the last sample $\tilde{I}_n$ is used, and using our estimation algorithm. Note that using our algorithm, SNR is consistently higher, due to the reduction in read noise. The improvement is most pronounced at low light. More significantly the sensor dynamic range defined as the ratio of the largest signal i_{max} to the smallest detectable signal i_{min} is increased. Assuming conventional sensor operation, $i_{max} = \frac{Q_{sat}}{T}$ and $i_{min} = \frac{\sigma_V}{T}$, which using the sensor parameters of the example in Figure 4 yields dynamic range of 47.4dB. Using our algorithm dynamic range is extended to 85.5dB – increasing 30.1dB at the high light end and 8dB at the low light end.

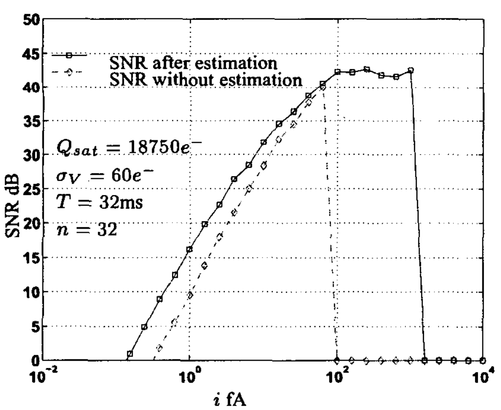

Figure 4: Estimation enhances the SNR and dynamic range

Figure 5 shows an example using a 20 frame image sequence captured by a high speed camera. Image k, for $1 \leq k \leq 20$, was constructed by summing the first k frames. Images (a) and (b) are the first and the last frames. Image (c) is the last image, which simulates a normal exposure time image. Image (d) is generated by applying our algorithm to the 20 images. Note that the image blur in (c) is almost completely eliminated in (d).

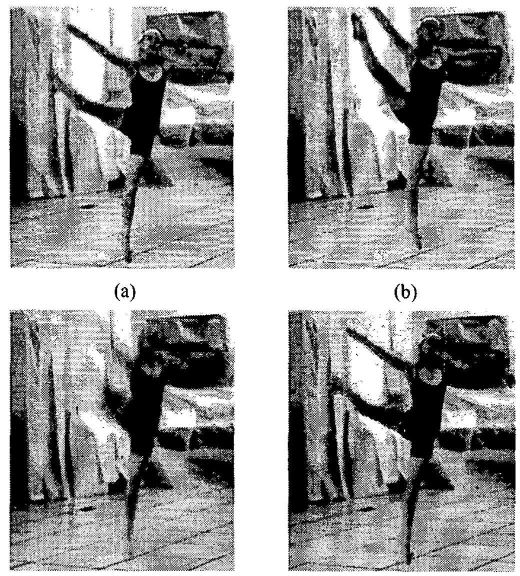

Figure 5: (a) First frame, (b) Last frame, (c) Simulation of the image generated by a conventional sensor. (d) Image generated by applying our algorithm.

5. ACKNOWLEDGMENT

The work reported in the paper is supported under the Programmable Digital Camera Program by Agilent, Canon, HP, Interval Research and Kodak. The authors wish to thank Chris Bregler for the high speed video sequence and Ting Chen, SukHwan Lim, and Khaled Salama for helpful discussions.

6. REFERENCES

[1] M.R. Banham and A.K. Katsaggelos, "Digital image restoration", *IEEE Signal Processing Magazine*, Vol. 14 no. 2, pp. 24 -41, March 1997.

[2] N. Stevanovic, M. Hillegrand, B. J. Hostica, and A. Teuner, "A CMOS image sensor for high speed imaging," *ISSCC Dig. Tech. Papers*, pp. 104-105, Feb 2000.

[3] S. Kleinfelder, S.H. Lim, X. Liu, and A. El Gamal, "A 10,000 frames/s 0.18 um CMOS digital pixel sensor with pixel-level memory", *ISSCC Dig. Tech. Papers*, Feb 2001.

[4] D. Yang, A. El Gamal, B. Fowler, and H. Tian, "A 640x512 CMOS image sensor with ultra-wide dynamic range floating-point pixel level ADC", *IEEE J. Solid-State Circuits*, Vol.34, no.12, pp. 1821-1834, Dec. 1999.

[5] S.H. Lim, and A. El Gamal, "Integration of Image Capture and Processing – Beyond Single Chip Digital Camera", *Proc. of SPIE*, Vol. 4306, March, 2001.

[6] X. Liu, and A. El Gamal, "Photocurrent estimation from multiple non-destructive samples in a CMOS image sensor", *Proc. of SPIE*, Vol. 4306, March, 2001.