Synthesis of High Dynamic Range Motion Blur Free Image From Multiple Captures

Xinqiao (Chiao) Liu, *Member, IEEE,* and Abbas El Gamal, *Fellow, IEEE*

***Abstract*—Advances in CMOS image sensors enable high-speed image readout, which makes it possible to capture multiple images within a normal exposure time. Earlier work has demonstrated the use of this capability to enhance sensor dynamic range. This paper presents an algorithm for synthesizing a high dynamic range, motion blur free, still image from multiple captures. The algorithm consists of two main procedures, photocurrent estimation and saturation and motion detection. Estimation is used to reduce read noise, and, thus, to enhance dynamic range at the low illumination end. Saturation detection is used to enhance dynamic range at the high illumination end as previously proposed, while motion blur detection ensures that the estimation is not corrupted by motion. Motion blur detection also makes it possible to extend exposure time and to capture more images, which can be used to further enhance dynamic range at the low illumination end. Our algorithm 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. Simulation and experimental results demonstrate the enhanced signal-to-noise ratio (SNR), dynamic range, and the motion blur prevention achieved using the algorithm.**

***Index Terms*—CMOS image sensor, dynamic range extension, motion blur restoration, motion detection, photocurrent estimation, saturation detection.**

I. Introduction

MOST of today's video and digital cameras use charge-coupled-device (CCD) image sensors [1], where the charge collected by the photodetectors during exposure time is serially read out resulting in slow readout speed and high power consumption. Also, CCDs are fabricated in a nonstandard technology, and as a result, other analog and digital camera functions such as A/D conversion, image processing, and compression, control, and storage cannot be integrated with the sensor on the same chip. Recently developed CMOS image sensors [2], [3], by comparison, are read out nondestructively and in a manner similar to a digital memory and can thus be operated continuously at very high frame rates [4]–[6]. A CMOS image sensor can also be integrated with other camera functions on the same chip ultimately leading to a single-chip digital camera with very small size, low power consumption, and additional functionality [7]–[10]. In [11], it is argued that the high frame-rate capability of CMOS image sensors coupled with the integration of processing with capture can enable the efficient implementations of many still and standard video imaging applications that can benefit from high frame rates, most notably, dynamic range extension.

CMOS image sensors generally suffer from lower dynamic range than CCDs due to their high readout noise and nonuniformity. To address this problem, several methods have been proposed for extending CMOS image sensor dynamic range. These include well-capacity adjusting [12], multiple capture [13], [14], [15], time to saturation [17], [18], spatially-varying exposure [16], logarithmic sensor [19], [20], and local adaptation [21]. With the exception of multiple capture, all other methods can only extend dynamic range at the high illumination end. Multiple capture also produces linear sensor response, which makes it possible to use correlated double sampling (CDS) for fixed pattern noise (FPN) and reset noise suppression, and to perform conventional color processing. Implementing multiple capture, however, requires very high frame-rate nondestructive readout, which has only recently become possible using digital pixel sensors (DPS) [6].

The idea behind the multiple-capture scheme is to acquire several images at different times within exposure time—shorter-exposure-time images capture the brighter areas of the scene, while longer-exposure-time images capture the darker areas of the scene. A high dynamic-range image can then be synthesized from the multiple captures by appropriately scaling each pixel's last sample before saturation (LSBS). In [22], it was shown that this scheme achieves higher signal-to-noise ratio (SNR) than other dynamic range-extension schemes. However, the LSBS algorithm does not take full advantage of the captured images. Since read noise is not reduced, dynamic range is only extended at the high illumination end. Dynamic range can be extended at the low illumination end by increasing exposure time. However, extending exposure time may result in unacceptable blur due to motion or change of illumination.

In this paper, we describe an algorithm for synthesizing a high dynamic range image from multiple captures while avoiding motion blur. The algorithm consists of two main procedures, photocurrent estimation and motion/saturation detection. Estimation is used to reduce read noise, and, thus, enhance dynamic range at the low-illumination end. Saturation detection is used

Manuscript received March 5, 2002; revised November 14, 2002. This work was supported in part by Agilent, in part by Canon, in part by Hewlett-Packard, in part by Interval Research, and in part by Kodak, all under the Programmable Digital Camera Program. This paper was presented in part at the 2001 SPIE Electronic Imaging conference, San Jose, CA, January 2001 [23] and at the IEEE International Conference on Acoustics, Speech, and Signal Processing, Salt Lake City, UT, May, 2001 [24]. This paper was recommended by Associate Editor B. E. Shi.

X. Liu was with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94304. He is now with Canesta Inc., San Jose, CA 95134 USA (e-mail: chiao@stanfordalumni.org).

A. El Gamal is with the Information Systems Laboratory, Department of Electrical Engineering, Stanford University, Stanford, CA 94305 (e-mail: abbas@isl.stanford.edu).

Digital Object Identifier 10.1109/TCSI.2003.809815

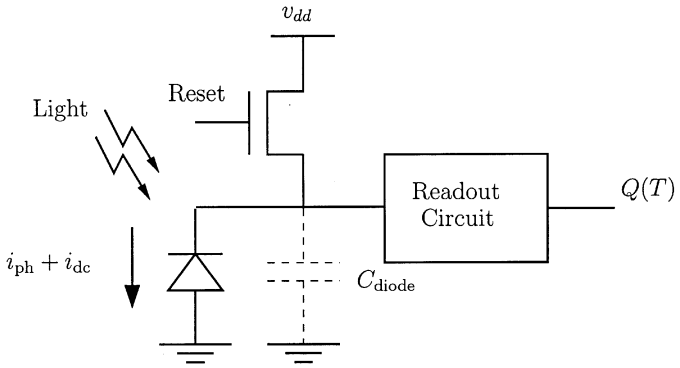


Fig. 1. CMOS image-sensor pixel diagram.

to enhance dynamic range at the high-illumination end as previously discussed, while motion blur detection ensures that the estimation is not corrupted by motion. Motion blur detection also makes it possible to extend exposure time and to capture more images, which can be used to further enhance dynamic range at the low illumination end. Our algorithm 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.

We present three estimation algorithms.

- An optimal recursive algorithm when reset noise and offset FPN are ignored. In this case, only the latest estimate and the new sample are needed to update the pixel photocurrent estimate.
- An optimal nonrecursive algorithm when reset noise and FPN are considered.
- A suboptimal recursive estimator for the second case, which is shown to yield mean-square error close to the nonrecursive algorithm without the need to store all the samples.

The later recursive algorithm is attractive since it requires the storage of only a constant number of values per pixel.

The motion-detection algorithm we describe in this paper detects change in each pixel's signal due to motion or change in illumination. The decision to stop estimating after motion is detected is made locally and is independent of other pixels signals.

The rest of the paper is organized as follows. In Section II, we describe the image-sensor signal and noise model we assume throughout the paper. In Section III, we describe our high-dynamic-range image-synthesis algorithm. In Section IV, we present the three estimation algorithms. In Section V, we present our motion-detection algorithm. Experimental results are presented in Section VI.

II. Image-Sensor Model

In this section, we describe the CMOS image-sensor operation and signal-and-noise model we use in the development and analysis of our synthesis algorithm. We use the model to define sensor SNR and dynamic range.

The image sensor used in an analog or digital camera consists of a 2-D array of pixels. In a typical CMOS image sensor [3], each pixel consists of a photodiode, a reset transistor, and several other readout transistors (see Fig. 1). The photodiode is reset before the beginning of capture. During exposure, the photodiode converts incident light into photocurrent $i_{\rm ph}(t)$, for $0 \le t \le T$, where T is the exposure time. This process is quite linear, and, thus, $i_{\rm ph}(t)$ is a good measure of incident light intensity. Since the photocurrent is too small to measure directly, it is integrated onto the photodiode parasitic capacitor $C_{\rm diode}$ and the charge $Q(T)$ (or voltage) is read out at the end of exposure time. Dark current $i_{\rm dc}$ and additive noise corrupt the output signal charge. The noise can be expressed as the sum of three independent components:

- Shot noise $U(T)$, which is normalized (zero mean) Poisson distributed. We assume here that the photocurrent is large enough and, thus, shot noise can be approximated by a Gaussian $U(T) \sim \mathcal{N}\left(0,\, q\int_0^T (i_{\rm ph}(t)+i_{\rm dc})\,dt\right)$, where q is the electron charge.
- Reset noise (including offset FPN) $C \sim \mathcal{N}(0,\, \sigma_C^2)$.
- Readout circuit noise $V(T)$ (including quantization noise) with zero mean and variance σ_V^2.

Thus, the output charge from a pixel can be expressed as

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

provided $Q(T) \le Q_{\rm sat}$, the saturation charge, also referred to as *well capacity*.

If the photocurrent is constant over exposure time, SNR is given by

$$\mathrm{SNR}(i_{\rm ph}) = 20\log_{10}\frac{i_{\rm ph}T}{\sqrt{q(i_{\rm ph}+i_{\rm dc})T+\sigma_V^2+\sigma_C^2}}. \quad (1)$$

Note that SNR increases with $i_{\rm ph}$, first at 20 dB per decade when reset and readout noise variance dominates, and then at 10 dB 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.

Dynamic range is a critical figure of merit for image sensors. It is defined as the ratio of the largest nonsaturating photocurrent to the smallest detectable photocurrent, typically defined as the standard deviation of the noise under dark conditions. Using the sensor model, dynamic range can be expressed as

$$\mathrm{DR} = 20\log_{10}\frac{i_{\max}}{i_{\min}} = 20\log_{10}\frac{Q_{\rm sat}-i_{\rm dc}T}{\sqrt{qi_{\rm dc}T+\sigma_V^2+\sigma_C^2}}. \quad (2)$$

Note that dynamic range decreases as exposure time increases due to the adverse effects of dark current. To increase dynamic range, one needs to either increase well capacity $Q_{\rm sat}$, and/or decrease read noise.

III. High-Dynamic-Range Image Synthesis

We first illustrate the effect of saturation and motion on image capture using the examples in Figs. 2 and 3. The first plot in Fig. 2 represents the case of a constant low light, where photocurrent can be well estimated from $Q(T)$. The second plot

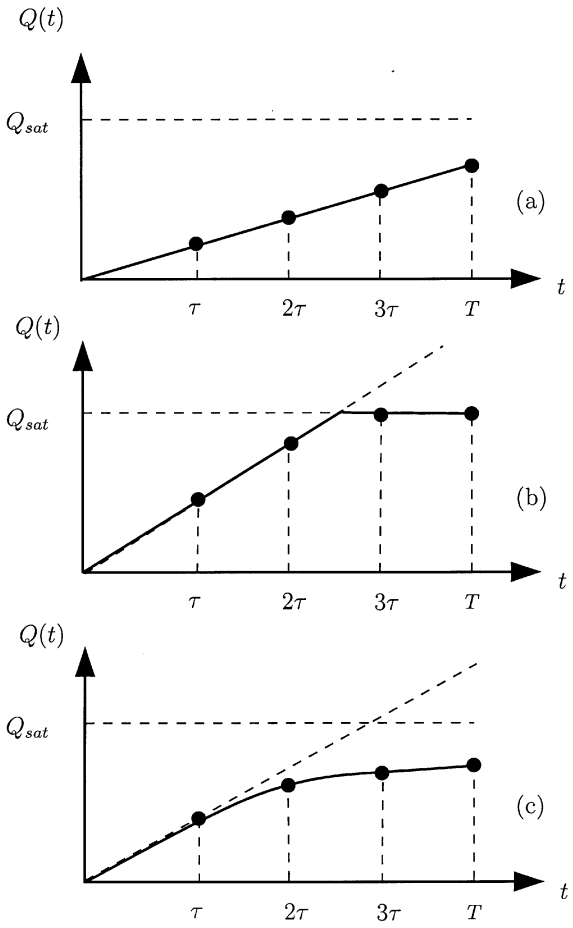


Fig. 2. $Q(t)$ versus t for three lighting conditions. (a) Constant low light. (b) Constant high light. (c) Light changing.

represents the case of a constant high light, where $Q(T) = Q_{\rm sat}$, 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 $i_{\rm ph}(0)$ again cannot be well estimated from $Q(T)$. To avoid saturation and the change of $i_{\rm ph}(t)$ due to motion, exposure time may be shortened, e.g., to τ in Fig. 2. Since in conventional sensor operations, 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 Fig. 3, 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 Fig. 3(b). On the other hand, if exposure time is set short, SNR deteriorates resulting in the noisy image of Fig. 3(c).

An important feature of several CMOS image-sensor architectures is nondestructive readout [23], [24]. Using this feature together with high-speed readout, several images can be captured without resetting during exposure. This is illustrated in the examples in Fig. 2, where each pixel signal is sampled at τ, 2τ, 3τ, and $T = 4\tau$. The estimation method described in [14] uses the LSBS to estimate photocurrent [$Q(T)$ in Fig. 2(a), $Q(2\tau)$ in Fig. 2(b)], and does not address motion blur. Applying this method to the example in Fig. 3, we get the same image as in

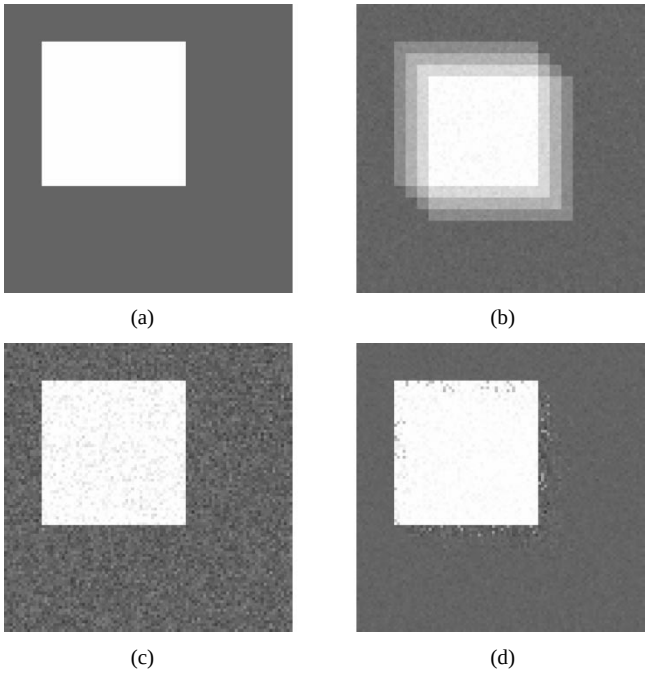


Fig. 3. (a) Ideal image. (b) Long-exposure-time image. (c) Short-exposure-time image. (d) Image produced using our algorithm.

Fig. 2(b). The algorithm we describe in this paper uses *all* the samples before saturation to estimate photocurrent at the beginning of exposure, so for the high light pixel example in Fig. 2, photocurrent is estimated using the images at τ and 2τ, while the photocurrent for the low light pixel is estimated using the four images. Motion blur in the third case can be reduced by using the first capture to estimate photocurrent at the beginning of exposure time $i_{\rm ph}(0)$. Applying our algorithm to the example in Fig. 3, we get the image (d), which is almost blur free and less noisy.

Our algorithm operates on $n \geq 1$ images,[1] captured at times τ, 2τ, ..., $n\tau = T$, as follows:

1) Capture first image, set $k = 1$.
2) For each pixel: Use the *photocurrent 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/saturation detected*: Set final photocurrent estimate
   $$\hat{I}_n = \hat{I}_k.$$
   ii) *No Motion/saturation detected or decision deferred*: Use the *photocurrent 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 $k = n$.

The following two sections provide details of the estimation and detection parts.

[1]Actually the algorithm operates on $n+1$ images, the first image, which is ignored here, is taken at $t = 0$ and is used to reduce reset noise and offset FPN as discussed in detail in Section IV.

IV. Photocurrent Estimation

Dynamic range at the low-illumination end can be enhanced using multiple captures by appropriately *averaging* each pixel's photocurrent samples to reduce readout noise. Since the sensor noise depends on the signal and the photocurrent samples are dependent, equal weight averaging may not reduce readout noise and can in fact be worse than simply using the LSBS to estimate photocurrent. In this section, we use linear mean-square error (MSE) estimation to derive the optimal weights to be used in the averaging. We first formulate the problem. We then present estimation algorithms for three cases: 1) when reset noise and offset FPN are ignored; 2) when reset noise and FPN are considered; and 3) a recursive estimator for case 2) without the need to store all the samples. Simulation results are presented and the performance of the three algorithms is compared in the last subsection.

A. Problem Formulation

We assume $n+1$ pixel charge samples Q_k captured at times $0, \tau, 2\tau, \ldots, n\tau = T$ and define the pixel current $i = i_{\rm ph} + i_{\rm dc}$. The kth charge sample is thus given by

$$Q_k = ik\tau + \sum_{j=1}^{k} U_j + V_k + C, \qquad \text{for } 0 \le k \le n \tag{3}$$

where V_k is the readout noise of the kth sample, U_j is the shot noise generated during the time interval $((j-1)\tau, j\tau]$, and C is the reset noise. U_j, V_k, and C are independent zero-mean random variables with

$$\begin{aligned} &E(V_k^2) = \sigma_V^2 > 0, \qquad \text{for } 0 \le k \le n \\ &E(U_j^2) = \sigma_U^2 = qi\tau, \qquad \text{for } 1 \le j \le k, \quad \text{and} \\ &E(C^2) = \sigma_C^2. \end{aligned}$$

We wish to estimate the photocurrent i from the $n+1$ samples. This is a parameter-estimation problem that can be formulated using several criteria, such as likelihood ratio and MSE [25]. Maximum-likelihood estimation achieves the smallest probability of error, but is generally difficult to derive and leads to nonlinear solutions that are not easy to implement in practice. In this paper we adapt the standard linear minimum mean-square parameter estimation (MMSE) methods (e.g., [25]) to our particular noise model.

Due to motion and/or saturation, the estimation may not use all the $n+1$ samples. The detection algorithm presented in the next section determines the LSBS/motion to be included in the estimation. Denoting the last sample to be included by k, $1 \le k \le n$, the linear MMSE problem is formulated as follows.

At time $k\tau$, we wish to find the best unbiased linear estimate, $\hat{I}_k$, of i given $\{Q_0, Q_1, \ldots, Q_k\}$, i.e., we wish to find $a_0^{(k)}, a_1^{(k)}, \ldots, a_k^{(k)}$ such that[2]

$$\hat{I}_k = \sum_{j=0}^{k} a_j^{(k)} Q_j \tag{4}$$

minimizes

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

subject to

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

B. Estimation Ignoring Reset Noise and FPN

Here, we ignore reset noise and offset FPN, i.e., set $C = 0$. Even though this assumption is not realistic for CMOS sensors, it is reasonable for high-end CCDs using very high-resolution A/D converters. As we shall see, the optimal estimate in this case can be cast in a recursive form, which is not the case when reset noise is considered.

To derive the best estimate, define the pixel *photocurrent samples* as

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

Thus, given the samples $\{\tilde{I}_1, \tilde{I}_2, \ldots, \tilde{I}_k\}$, we wish to find the best unbiased linear estimate of the parameter i, i.e., weights $a_1^{(k)}, a_2^{(k)}, \ldots, a_k^{(k)}$ such that[3]

$$\hat{I}_k = \sum_{j=1}^{k} a_j^{(k)} \tilde{I}_j$$

that minimizes

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

subject to

$$E(\hat{I}_k) = i. \tag{5}$$

The MSE Φ_k^2 is given by

$$\begin{aligned} \Phi_k^2 &= E(\hat{I}_k - i)^2 \\ &= \sum_{j=1}^{k} \left(\left(\sum_{l=j}^{k} \frac{a_l^{(k)}}{l} \right)^2 \frac{\sigma_U^2}{\tau^2} + \left(\frac{a_j^{(k)}}{j} \right)^2 \frac{\sigma_V^2}{\tau^2} \right). \end{aligned} \tag{6}$$

This is a convex optimization problem with a linear constraint as in (5). To solve it, we define the Lagrangian

$$F\left(a_1^{(k)}, a_2^{(k)}, \ldots, a_k^{(k)}\right) = \Phi_k^2 + \lambda \left(\sum_{j=1}^{k} a_j^{(k)} - 1 \right) \tag{7}$$

where λ is the Lagrange multiplier.

The optimal weights can be found using the conditions

$$\begin{aligned} &\nabla F = \left[\frac{\partial F}{\partial a_1^{(k)}} \;\; \frac{\partial F}{\partial a_2^{(k)}} \;\; \cdots \;\; \frac{\partial F}{\partial a_k^{(k)}} \right]^T = 0 \\ &\sum_{j=1}^{k} a_j^{(k)} = 1 \end{aligned} \tag{8}$$

[2]For the coefficient $a_j^{(k)}$, we use superscript (k) to represent the number of captures used and use subscript as the index of the coefficients for each capture.

[3]In this case, there is no need for sample Q_0 at $t = 0$, therefore, weights start with $a_1^{(k)}$.

and we get [26]

$$a_j^{(k)} = j a_1^{(k)} + \frac{j}{j-1} a_{j-1}^{(k)} + \frac{j\sigma_U^2}{\sigma_V^2}\left(\sum_{l=1}^{j-1} \frac{a_l^{(k)}}{l}\right)$$
$$\text{where } 2 \le j \le k \tag{9}$$

Now, to see that the optimal estimate can be cast in a recursive form, we define a new set of weights b_j, such that

$$b_1 = 1$$
$$b_j = j b_1 + \frac{j}{j-1} b_{j-1} + \frac{j\sigma_U^2}{\sigma_V^2}\left(\sum_{l=1}^{j-1} \frac{b_l}{l}\right), \qquad \text{for } j \ge 2 \tag{10}$$

and $a_j^{(k)}$ can be represented in terms of b_j as

$$a_j^{(k)} = \frac{b_j}{\sum_{l=1}^{k} b_l}, \qquad \text{for } 1 \le j \le k.$$

The optimal photocurrent estimate $\hat{I}_k$ can be written in a recursive form in terms of b_k, the latest photocurrent sample $\tilde{I}_k$, and the previous estimate $\hat{I}_{k-1}$ as

$$\hat{I}_k = \hat{I}_{k-1} + h_k(\tilde{I}_k - \hat{I}_{k-1}) \tag{11}$$

where

$$h_k = \frac{b_k}{g_k} \quad \text{and} \quad g_k = \sum_{l=1}^{k} b_l.$$

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

$$\Phi_k^2 = \frac{g_{k-1}^2}{g_k^2}\Phi_{k-1}^2 + \frac{1}{g_k^2}\left((2b_k g_{k-1} + b_k^2)\frac{\sigma_U^2}{k\tau^2} + b_k^2\frac{\sigma_V^2}{(k\tau)^2}\right). \tag{12}$$

This is important because Φ_k^2 will be used later in the motion-detection algorithm.

The first estimator $\hat{I}_1$ is approximated by $\tilde{I}_1$. In (10) and (12), $\sigma_U^2 = qi\tau$ is approximated using the latest estimate of i, $\hat{I}_k$, i.e., $\sigma_U^2 = q\hat{I}_k\tau$. We found that this approximation yields MSE very close to the optimal case, i.e., when i is known.

C. Estimation Considering Reset Noise and FPN

From (3) and (4), to minimize the MSE of the best estimator $\hat{I}_k$, we need

$$\frac{\partial \Phi_k^2}{\partial a_0^{(k)}} = 2a_0^{(k)}\sigma_V^2 + 2\left(\sum_{j=0}^{k} a_j^{(k)}\right)\sigma_C^2 = 0$$

which gives

$$a_0^{(k)} = -w\left(\sum_{j=1}^{k} a_j^{(k)}\right) \tag{13}$$

where

$$w = \frac{\sigma_C^2}{\sigma_C^2 + \sigma_V^2}. \tag{14}$$

Bring (13) into (4), we have

$$\hat{I}_k = \sum_{j=0}^{k} a_j^{(k)} Q_j = \sum_{j=1}^{k} a_j^{(k)}(Q_j - wQ_0). \tag{15}$$

Therefore, we redefine *photocurrent sample* $\tilde{I}_k$ as

$$\tilde{I}_k = \frac{Q_k - wQ_0}{k\tau}, \qquad \text{for } 1 \le k \le n \tag{16}$$

such that $\tilde{I}_k$ corresponds to an estimate with a *weighted* CDS operation. Compared with a conventional CDS operation [1], where

$$\tilde{I}_k = \frac{Q_k - Q_0}{k\tau} \tag{17}$$

the weighting w has the effect of reducing the additional readout noise due to CDS. In a conventional CDS as in (17), the readout noise power is always doubled, while in a weighted CDS as in (16), the readout noise power is $1 + w$, where $0 < w < 1$, as given in (14).

The pixel current estimate given the first k samples can be expressed as

$$\hat{I}_k = \mathbf{A}_k^T \tilde{\mathbf{I}}_k$$

where

$$\mathbf{A}_k = \begin{bmatrix} b_1^{(k)} & b_2^{(k)} & \cdots & b_k^{(k)} \end{bmatrix}^T \quad \text{and}$$
$$\tilde{\mathbf{I}}_k = \begin{bmatrix} \tilde{I}_1 & \tilde{I}_2 & \cdots & \tilde{I}_k \end{bmatrix}^T.$$

The optimal coefficient vector $\mathbf{A}_k$ is given by

$$\mathbf{A}_k = -\left(M_k\frac{\sigma_U^2}{\tau^2} + D_k\frac{\sigma_V^2}{\tau^2}\right)^{-1}\frac{\lambda}{2}L_k \tag{18}$$

where

$$M_k = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{k} \\ 1 & \frac{2}{2} & \frac{2}{3} & \cdots & \frac{2}{k} \\ 1 & \frac{2}{2} & \frac{3}{3} & \cdots & \frac{3}{k} \\ \cdots & & & & \\ 1 & \frac{2}{2} & \frac{3}{3} & \cdots & \frac{k}{k} \end{bmatrix}$$

$$L_k = \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ k \end{bmatrix}$$

$$D_k = \begin{bmatrix} w+1 & \frac{w}{2} & \frac{w}{3} & \cdots & \frac{w}{k} \\ w & \frac{w+1}{2} & \frac{w}{3} & \cdots & \frac{w}{k} \\ w & \frac{w}{2} & \frac{w+1}{3} & \cdots & \frac{w}{k} \\ \cdots & & & & \\ w & \frac{w}{2} & \frac{w}{3} & \cdots & \frac{w+1}{k} \end{bmatrix}$$

and λ is the Lagrange multiplier for the unbiased constraint.

In [26], we show that the above solution cannot be expressed in a recursive form, and, thus, finding $\hat{I}_k$ requires the storage of the vector $\tilde{\mathbf{I}}_k$ and inverting a $k \times k$ matrix.

D. Recursive Algorithm

Now, we restrict ourselves to recursive estimates, i.e., estimates of the form

$$\hat{I}_k = \hat{I}_{k-1} + h_k(\tilde{I}_k - \hat{I}_{k-1})$$

where again

$$\tilde{I}_k = \frac{Q_k - wQ_0}{k\tau}.$$

The coefficient h_k can be found by solving the equations

$$\frac{d\,\Phi_k^2}{d\,h_k} = \frac{d\,E(\hat{I}_k - i)^2}{d\,h_k} = 0 \quad \text{and} \quad E\hat{I}_k = i.$$

Define the MSE of $\tilde{I}_k$ as

$$\Delta_k^2 = E(\tilde{I}_k - i)^2 = \frac{1}{k^2\tau^2}\,(k\sigma_U^2 + (1+w)\sigma_V^2) \tag{19}$$

and the covariance between $\tilde{I}_k$ and $\hat{I}_k$ as

$$\Theta_k = E(\tilde{I}_k - i)(\hat{I}_k - i) = (1-h_k)\,\frac{k-1}{k}\,\Theta_{k-1} - \frac{(1-h_k)h_{k-1}}{k(k-1)\tau^2}\,\sigma_V^2 + h_k\Delta_k^2. \tag{20}$$

The MSE of $\hat{I}_k$ can be expressed in terms of Δ_k^2 and Θ_k as

$$\Phi_k^2 = (1-h_k)^2\Phi_{k-1}^2 + \frac{2(k-1)(1-h_k)h_k}{k}\,\Theta_{k-1} - \frac{2h_{k-1}(1-h_k)h_k}{k(k-1)\tau^2}\,\sigma_V^2 + h_k^2\Delta_k^2. \tag{21}$$

To minimize the MSE, we require that

$$\frac{d\Phi_k^2}{d\,h_k} = 0$$

which gives

$$h_k = \frac{\Phi_{k-1}^2 - \frac{(k-1)}{k}\,\Theta_{k-1} + \frac{h_{k-1}\sigma_V^2}{k(k-1)\tau^2}}{\Phi_{k-1}^2 - \frac{2(k-1)}{k}\Theta_{k-1} + \frac{2h_{k-1}\sigma_V^2}{k(k-1)\tau^2} + \Delta_k^2}. \tag{22}$$

Note that h_k, Θ_k, and Φ_k can all be recursively updated.

To summarize, the suboptimal recursive algorithm is as follows.

- Set initial parameter and estimate values as follows:

$$h_1 = 1$$
$$\tilde{I}_1 = \frac{(Q_1 - wQ_0)}{\tau}$$
$$\hat{I}_1 = \tilde{I}_1$$
$$\Delta_1^2 = \frac{\sigma_U^2 + (1+w)\sigma_V^2}{\tau^2}$$
$$\Phi_1^2 = \Delta_1^2$$
$$\Theta_1 = \Delta_1^2.$$

- At each iteration, the parameter and estimate values are updated as follows:

$$\tilde{I}_k = \frac{(Q_k - wQ_0)}{k\tau}$$
$$\Delta_k^2 = \frac{1}{k^2\tau^2}\,(k\sigma_U^2 + (1+w)\sigma_V^2)$$
$$h_k = \frac{\Phi_{k-1}^2 - \frac{(k-1)}{k}\,\Theta_{k-1} + \frac{h_{k-1}\sigma_V^2}{k(k-1)\tau^2}}{\Phi_{k-1}^2 - \frac{2(k-1)}{k}\,\Theta_{k-1} + \frac{2h_{k-1}\sigma_V^2}{k(k-1)\tau^2} + \Delta_k^2}$$
$$\Theta_k = (1-h_k)\,\frac{k-1}{k}\,\Theta_{k-1} - \frac{(1-h_k)h_{k-1}}{k(k-1)\tau^2}\,\sigma_V^2 + h_k\Delta_k^2$$
$$\Phi_k^2 = (1-h_k)^2\Phi_{k-1}^2 + 2h_k\Theta_k - {h_k}^2\Delta_k^2$$
$$\hat{I}_k = \hat{I}_{k-1} + h_k(\tilde{I}_k - \hat{I}_{k-1}).$$

Note that to find the new estimate $\hat{I}_k$ using this suboptimal recursive algorithm, only three parameters, h_k, Φ_k and Θ_k, the old estimate $\hat{I}_{k-1}$, and the new sample value $\tilde{I}_k$ are needed. Thus, only a small amount of memory per pixel is required independent of the number of images captured.

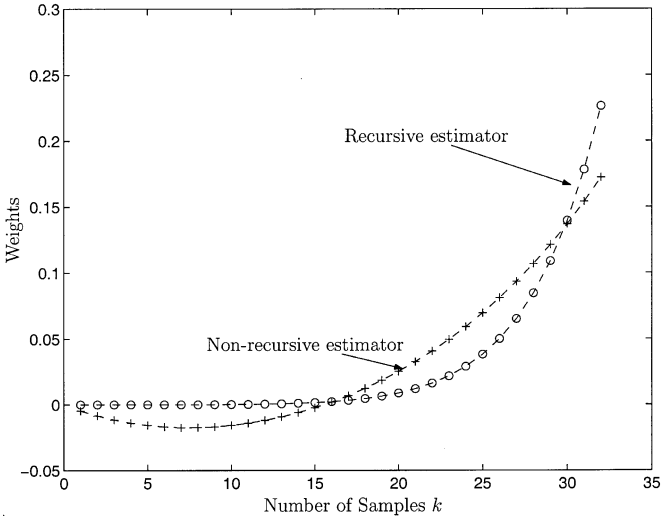


Fig. 4. Distribution of estimation weights among total 32 samples used in the nonrecursive and recursive algorithms.

E. Simulation Results

In this subsection, we present simulation results that demonstrate the SNR improvements using the nonrecursive algorithm described in Section IV-C, the recursive algorithm in Section IV-D, and the multiple capture scheme in [14].

The simulation results are summarized in Figs. 4–6. The sensor parameters assumed in the simulations are as follows

$$Q_{\text{sat}} = 18\,750 \text{ e-}$$
$$i_{\text{dc}} = 0.1 \text{ fA}$$
$$\sigma_V = 60 \text{ e-}$$
$$\sigma_C = 62 \text{ e-}$$
$$T = 32 \text{ ms}$$
$$\tau = 1 \text{ ms}.$$

Fig. 4 plots the weights for the nonrecursive and recursive algorithms in Sections IV-C and D, respectively. Note that with a typical readout noise rms (60 e^- in this example), later samples are weighted much higher than earlier ones since later samples have higher SNR. As read noise decreases, this becomes more pronounced—the best estimate is to use the last sample only (i.e., $\mathbf{A}_k = [0, \ldots, 0, 1]^T$) if sensor read noise is zero. On the other extreme, if shot noise can be ignored, then, the best estimate is averaging (i.e., $\mathbf{A}_k = [1/k, \ldots, 1/k]^T$). Also note that weights for the nonrecursive algorithm can be negative. It

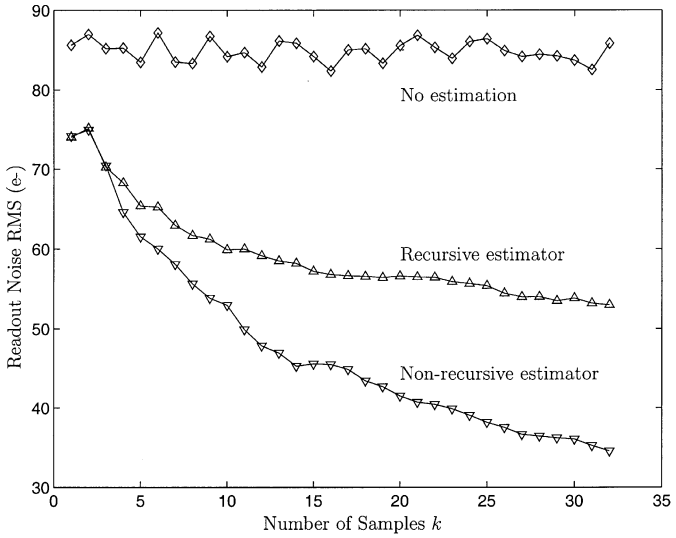


Fig. 5. Simulated equivalent readout noise rms value versus number of samples k.

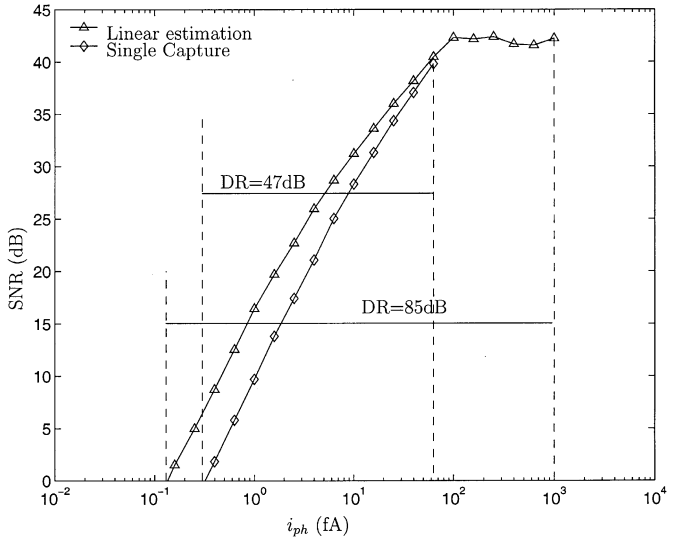


Fig. 6. Estimation enhances the SNR and dynamic range.

is preferred to weight the later samples higher since they have higher SNR, and this can be achieved by using negative weights for some of the earlier samples under the unbiased estimate constrain (sum of the weights equals one).

Fig. 5 compares the equivalent readout noise rms at low illumination level corresponding to $i_{\rm ph} = 2$ fA as a function of the number of samples k for conventional sensor operation and using the nonrecursive and the recursive estimation algorithms. As can be seen, the equivalent readout noise after the last sample is reduced from 86 e- when no estimation is used to 35.8 e- when the nonrecursive estimator is used and to 56.6 e- when the recursive estimator is used. Also note the drop in the equivalent readout noise rms due to the weighted CDS used in our algorithms.

Fig. 6 plots SNR versus i for conventional sensor operation, where the last sample $\tilde{I}_n$ is used, and using our estimation algorithm. Note that by using our algorithm, SNR is consistently higher, due to the reduction in read noise. The improvement is most pronounced at low light. In this example, sensor with single capture yields dynamic range of 47 dB. Using our algorithm, dynamic range is extended to 85 dB—an increase of 30 dB at the high illumination end and 8 dB at the low illumination. In general, the dynamic-range extension achievable at the low-illumination end depends on the read noise power and number of samples used in the estimation—the higher the read noise and the more samples are used, the greater the dynamic range is extended. On the other hand, the dynamic range extension achievable at the high-illumination end depends on the sensor readout time—the faster the readout, the greater the dynamic range is extended.

V. 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 $n+1$ 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 Section IV-D. If the current stays constant, the next observation $\tilde{I}_{k+1}^{\rm pre}$ would be

$$\tilde{I}_{k+1}^{\rm pre} = i + \frac{\sum_{j=1}^{k+1} U_j}{(k+1)\tau} + \frac{V_{k+1} - wV_0}{(k+1)\tau} + \frac{(1-w)C}{(k+1)\tau} \qquad (23)$$

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

$$\Delta_{\rm pre}^2 = E\left((\tilde{I}_{k+1}^{\rm pre} - \hat{I}_k)^2 \middle| \hat{I}_k\right) = \left(\frac{k}{k+1}\right)^2 \Delta_k^2 + \Phi_k^2 - \frac{2k}{k+1}\Theta_k + \frac{2h_k}{k(k+1)\tau^2}\sigma_V^2 + \frac{\sigma_U^2}{(k+1)^2\tau^2} \qquad (24)$$

where Δ_k^2, Θ_k, Φ_k^2, and h_k are given in (19)– (22), respectively.

Thus, to decide whether the input signal i changed between time $k\tau$ and $(k+1)\tau$, we compare $\tilde{I}_{k+1} = (Q_{k+1} - wQ_0)/((k+1)\tau)$ with $\hat{I}_k$. A simple decision rule would be to declare that motion has occurred if

$$|\tilde{I}_{k+1} - \hat{I}_k| \geq m\Delta_{\rm pre} \qquad (25)$$

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 the m, the more the motion blur if i changes with time, but also the higher the SNR if i is a constant, and vice versa.

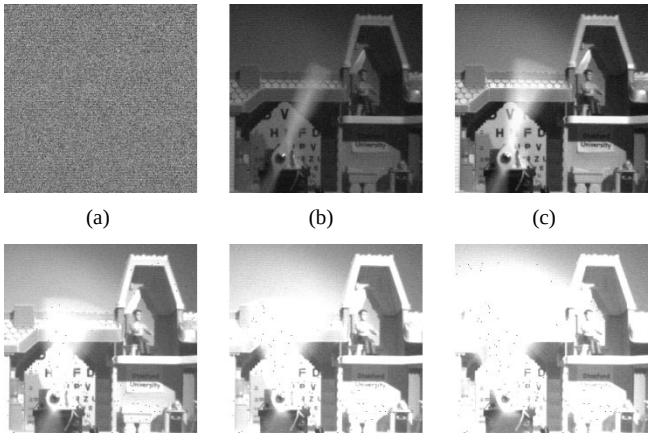


Fig. 7. Six of the 65 images of the high dynamic scene captured nondestructively at 1000 frames/s. (a) $t = 0$ ms. (a) $t = 0$ ms. (b) $t = 10$ ms. (c) $t = 20$ ms. (d) $t = 30$ ms. (e) $t = 40$ ms. (f) $t = 50$ ms.

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 $|\tilde{I}_{k+1} - \hat{I}_k| \leq m_1\Delta_{\rm pre}$, 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 $|\tilde{I}_{k+1} - \hat{I}_k| \geq m_2\Delta_{\rm pre}$, $L^+ = l_{\max}$, or $L^- = l_{\max}$, then declare that *motion detected*. Use $\hat{I}_k$ as the final estimate of i.
3) If $m_1\Delta_{\rm pre} < \tilde{I}_{k+1} - \hat{I}_k < m_2\Delta_{\rm pre}$, then *defer the decision* and set $L^+ = L^+ + 1, L^- = 0$.
4) If $-m_2\Delta_{\rm pre} < \tilde{I}_{k+1} - \hat{I}_k < -m_1\Delta_{\rm pre}$, 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.

VI. Experimental Results

In this section, we present experimental results performed using a PC-based high-speed CMOS imaging system [27] designed around the 10 000 frames/s CMOS DPS chip.

The high-dynamic-range scene used in the experiment comprised a doll house under direct illumination from above, and a rotating model airplane propeller. We captured 65 frames of the scene at 1000 frames/s nondestructively, and uniformly spaced over a 64-ms exposure time. Fig. 7 shows some of the images captured. Note that as exposure time increases, the details in the shadow area (such as the word "Stanford") begin to appear while the high-illumination area suffers from saturation and the area where the propeller is rotating suffers from significant motion blur.

We first applied the LSBS algorithm [14] to the 65 images to obtain the high dynamic range image in Fig. 8. While the image indeed contains many of the details in both low and high illumination areas, it suffers from motion blur and is quite noisy in the dark areas. Fig. 9 shows the high-dynamic-range motion blur free image synthesized from the 65 captures using the algorithm discussed in this paper. Note that the dark background is much smoother due to reduction in readout and FPN, and the motion blur caused by the rotating propeller in Fig. 8 is almost completely eliminated.

To illustrate the operation of our algorithm, in Fig. 10 we plot the sampled and estimated photocurrents for three pixels under different illumination levels. Note how motion blur is prevented in the third pixel using the motion-detection algorithm.

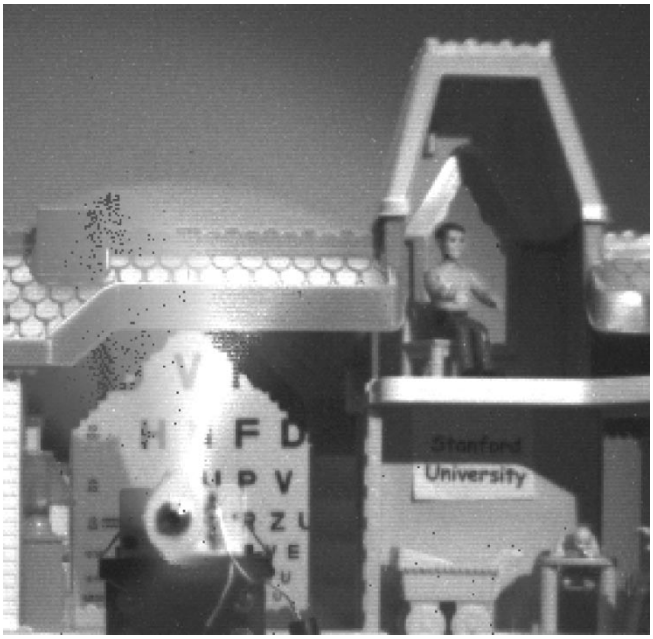


Fig. 8. High-dynamic-range image synthesized using the LSBS algorithm.

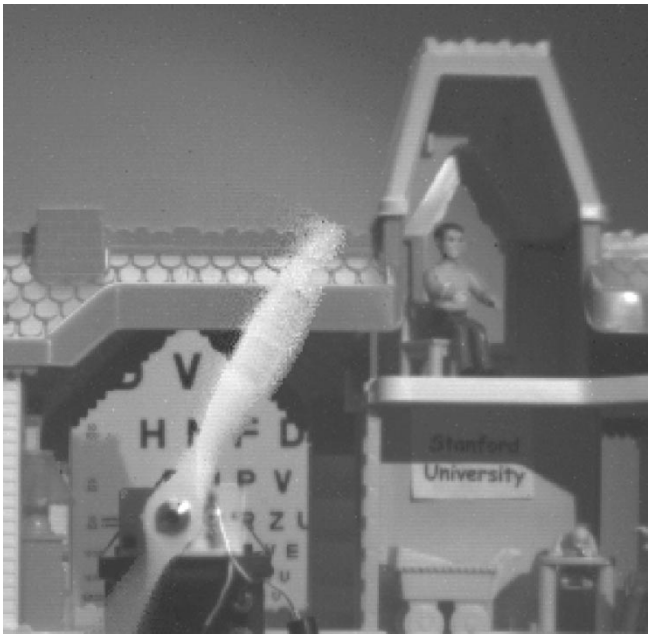


Fig. 9. High-dynamic-range motion blur free image synthesized from the 65 images.

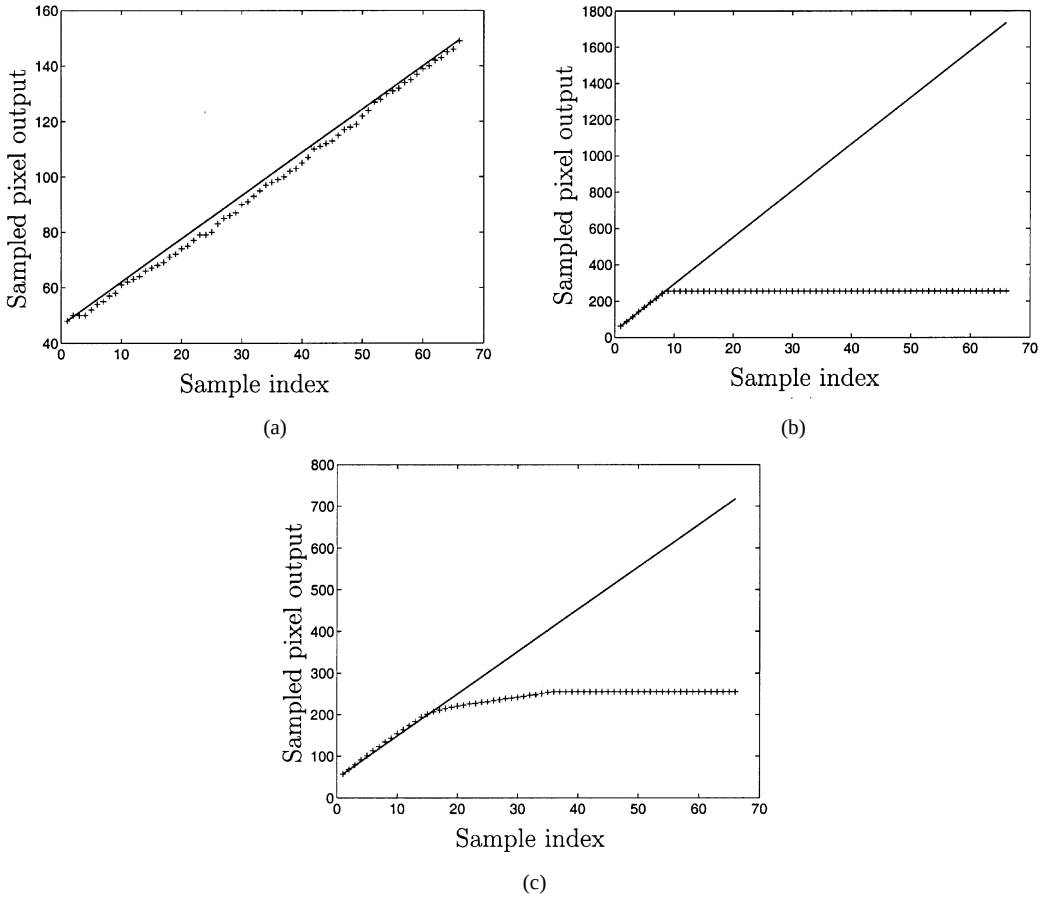


Fig. 10. Readout values (marked by "+") and estimated values (solid lines) for (a) pixel in the dark area, (b) pixel in bright area, and (c) pixel with varying illumination due to motion.

VII. Conclusion

The high frame-rate capability of CMOS image sensors makes it possible to nondestructively capture several images within a normal exposure time. The captured images provide additional information that can be used to enhance the performance of many still and standard video imaging applications [11]. The paper describes an algorithm for synthesizing a high dynamic range, motion blur free image from multiple captures. The algorithm consists of two main procedures, photocurrent estimation and motion/saturation detection. Estimation is used to reduce read noise, and, thus, enhance dynamic range at the low illumination end. Saturation detection is used to enhance dynamic range at the high illumination end, while motion blur detection ensures that the estimation is not corrupted by motion. Motion blur detection also makes it possible to extend exposure time and to capture more images, which can be used to further enhance dynamic range at the low illumination end. Experimental results demonstrate that this algorithm achieves increased SNR, enhanced dynamic range, and motion blur prevention.

Acknowledgment

The authors would like to thank T. Chen, H. Eltoukhy, A. Ercan, S. Lim, and K. Salama for their feedback.

References

[1] A. J. Theuwissen, *Solid-State Imaging With Charge-Coupled Devices*. Norwell, MA: Kluwer, May 1995.

[2] E. R. Fossum, "Active pixel sensors: Are CCDs dinosaurs," *Proc. SPIE*, vol. 1900, pp. 2–14, Feb. 1993.

[3] ——, "CMOS image sensors: Electronic camera-on-chip," *IEEE Trans. Electron Devices*, vol. 44, pp. 1689–1698, Oct. 1997.

[4] A. Krymski, D. Van Blerkom, A. Andersson, N. Block, B. Mansoorian, and E. R. Fossum, "A high speed, 500 frames/s, 1024 × 1024 CMOS active pixel sensor," in *Proc. 1999 Symp. VLSI Circuits*, June 1999, pp. 137–138.

[5] N. Stevanovic, M. Hillegrand, B. J. Hostica, and A. Teuner, "A CMOS image sensor for high speed imaging," in *Dig. Tech. Papers 2000 IEEE Int. Solid-State Circuits Conf.*, Feb. 2000, pp. 104–105.

[6] S. Kleinfelder, S. H. Lim, X. Liu, and A. El Gamal, "A 10 000 frames/s CMOS digital pixel sensor," *IEEE J. Solid-State Circuits*, vol. 36, pp. 2049–2059, Dec. 2001.

[7] M. Loinaz, K. Singh, A. Blanksby, D. Inglis, K. Azadet, and B. Ackland, "A 200mW 3.3V CMOS color camera IC producing 352 × 288 24b video at 30 frames/s," in *Dig. Tech. Papers 1998 IEEE Int. Solid-State Circuits Conf.*, Feb. 1998, pp. 168–169.

[8] S. Smith, J. Hurwitz, M. Torrie, D. Baxter, A. Holmes, M. Panaghiston, R. Henderson, A. Murray, S. Anderson, and P. Denyer, "A single-chip 306 × 244-pixel CMOS NTSC video camera," in *Dig. Tech. Papers 1998 IEEE Int. Solid-State Circuits Conf.*, Feb. 1998, pp. 170–171.

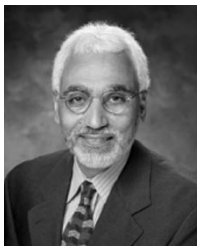
[9] S. Yoshimura, T. Sugiyama, K. Yonemoto, and K. Ueda, " A 48 kframes/s CMOS image sensor for real-time 3-D sensing and motion detection," in *Dig. Tech. Papers 2001 IEEE Int. Solid-State Circuits Conf.*, Feb. 2001, pp. 94–95.

[10] T. Sugiyama, S. Yoshimura, R. Suzuki, and H. Sumi, "A 1/4-inch QVGA color imaging and 3-D sensing CMOS sensor with analog frame memory," in *Dig. Tech. Papers 2002 IEEE Int. Solid-State Circuits Conf.*, Feb. 2002, pp. 434–435.

[11] S. H. Lim and A. El Gamal, "Integration of image capture and processing—Beyond single chip digital camera," *Proc. SPIE*, vol. 4306, pp. 219–226, Mar. 2001.
[12] S. J. Decker, R. D. McGrath, K. Brehmer, and C. G. Sodini, "A 256 × 256 CMOS imaging array with wide dynamic range pixels and column-parallel digital output," *IEEE J. Solid-State Circuits*, vol. 33, pp. 2081–2091, Dec. 1998.
[13] O. Yadid-Pecht and E. Fossum, "Wide intrascene dynamic range CMOS APS using dual sampling," *IEEE Trans. Electron Devices*, vol. 44, pp. 1721–1723, Oct. 1997.
[14] D. Yang, A. El Gamal, B. Fowler, and H. Tian, "A 640 × 512 CMOS image sensor with ultra-wide dynamic range floating-point pixel level ADC," *IEEE J. Solid-State Circuits*, vol. 34, pp. 1821–1834, Dec. 1999.
[15] O. Yadid-Pecht and A. Belenky, "Autoscaling CMOS APS with customized increase of dynamic range," in *Dig. Tech. Papers 2001 IEEE Int. Solid-State Circuits Conf.*, Feb. 2001, pp. 100–101.
[16] M. Aggarwal and N. Ahuja, "High dynamic range panoramic imaging," in *Proc. 8th IEEE Int. Conf. Computer Vision*, vol. 1, 2001, pp. 2–9.
[17] W. Yang, "A wide-dynamic range, low power photosensor array," in *Dig. Tech. Papers 1994 IEEE Int. Solid-State Circuits Conf.*, Feb. 1994, pp. 230–231.
[18] E. Culurciello, R. Etienne-Cummings, and K. Boahen, "Arbitrated address event representation digital image sensor," in *Dig. Tech. Papers 2001 IEEE Int. Solid-State Circuits Conf.*, Feb. 2001, pp. 92–93.
[19] M. Loose, K. Meier, and J. Schemmel, "A self-calibrating single-chip CMOS camera with logarithmic response," *IEEE J. Solid-State Circuits*, vol. 36, pp. 586–596, Apr. 2001.
[20] S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, and J. Bogaerts, "A logarithmic response CMOS image sensor with on-chip calibration," *IEEE J. Solid-State Circuits*, vol. 35, pp. 1146–1152, Aug. 2000.
[21] T. Delbruck and C. A. Mead, "Analog VLSI phototransduction," California Institute of Technology, Pasadena, CNS Memo no. 30, May 11, 1994.
[22] D. Yang and A. El Gamal, "Comparative analysis of SNR for image sensors with enhanced dynamic range," *Proc. SPIE*, vol. 3649, pp. 197–211, 1999.
[23] X. Liu and A. El Gamal, "Photocurrent estimation from multiple nondestructive samples in a CMOS image sensor," *Proc. SPIE*, vol. 4306, pp. 450–458, 2001.
[24] ——, "Simultaneous image formation and motion blur restoration via multiple capture," in *Proc. ICASSP2001*, vol. 3, Salt Lake City, UT, May 2001, pp. 1841–1844.
[25] H. Sorenson, *Parameter Estimation, Principles and Problems.* New York: Marcell Dekker, 1980.
[26] X. Liu, "CMOS image sensors dynamic range and SNR enhancement via statistical signal processing," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2002.
[27] A. Ercan, F. Xiao, X. Liu, S. H. Lim, A. El Gamal, and B. Wandell, "Experimental high speed CMOS image sensor system and applications," in *Proc. IEEE Sensors 2002*, Orlando, FL, June 2002, pp. 15–20.

Xinqiao (Chiao) Liu (S'97–M'02) received the B.S. degree in physics from the University of Science and Technology of China, Anhui, China, in 1993, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1997 and 2002, respectively.

In the summer of 1998, he worked as a Research Intern at Interval Research Inc., Palo Alto, CA, on image sensor characterization and novel imaging system design. He is currently with Canesta Inc., San Jose, CA, developing 3-D image sensors. At Stanford, his research was focused on CMOS image sensor dynamic range and SNR enhancement via innovative circuit design, and statistical signal processing algorithms.



Abbas El Gamal (S'71–M'73–SM'83–F'00) received the B.S. degree in electrical engineering from Cairo University, Cairo, Egypt, in 1972, the M.S. degree in statistics, and the Ph.D. degree in electrical engineering, both from Stanford University, Stanford, CA, in 1977 and 1978, respectively.

From 1978 to 1980, he was an Assistant Professor of Electrical Engineering at the University of Southern California, Los Angeles. He joined the Stanford faculty in 1981, where he is currently a Professor of Electrical Engineering. From 1984 to 1988, while on leave from Stanford, he was Director of LSI Logic Research Lab, Sunnyvale, CA, later, Cofounder and Chief Scientist of Actel Corporation, Sunnyvale, CA. From 1990 to 1995, he was a Cofounder and Chief Technical Officer of Silicon Architects, Mountainview, CA, which was acquired by Synopsis. He is currently a Principal Investigator on the Stanford Programmable Digital Camera project. His research interests include digital imaging and image processing, network information theory, and electrically configurable VLSI design and CAD. He has authored or coauthored over 125 papers and 25 patents in these areas.

Dr. El Gamal serves on the board of directors and advisory boards of several IC and CAD companies. He is a member of the ISSCC Technical Program Committee.