

OVERSAMPLED NOISY BINARY IMAGE SENSOR

Feng Yang and Martin Vetterli

LCAV - School of Computer and Communication Sciences
École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

ABSTRACT

We study the oversampled binary image sensor in [1] under noisy scenario. The binary image sensor is similar to traditional photographic film with pixel value equal to “0” or “1”. The potential application of the oversampled binary image sensor is high dynamic range imaging. Since the pixel value is binary, we model the noise as additive Bernoulli noise. We focus on the case that the threshold in the binary sensor is equal to a single photon. Because of noise, the dynamic range of the sensor is reduced. But the image sensor is quite robust to noise when the light intensity value is large. We use maximum-likelihood estimator (MLE) to reconstruct the light intensity field, and prove that when the threshold is a single photon, even if there is noise, the log-likelihood function is still concave, which guarantees to find the global optimal solution. Experimental results for 1-D signal and 2-D images verify our performance analysis and show the effectiveness of the reconstruction algorithm.

Index Terms— High dynamic range imaging, photon-limited imaging, diffraction-limited imaging, binary image sensor

1. INTRODUCTION

Moore’s law [2] claimed that the number of transistors could be placed on an integrated circuit doubled approximately every two years. There is a strong link between the pixel size of the CMOS image sensor and Moore’s law [3]. When the pixel size becomes small, the full-well capacity (*i.e.*, the maximum photon-electrons a pixel can hold) is reduced. This will result in low signal-to-noise ratios (SNRs) and poor dynamic ranges. To benefit from the shrinking of the pixel size, Fossum [4] proposed to build a binary sensor that was similar to photographic film in which the light intensity information was presented as the density of the opaque silver grains. Sbaiz et al. [5] proposed the concept of oversampled binary image sensor in the name of the gigavision camera.

We gave a theoretical analysis of the binary sensor in [1]. We showed that if we could build a binary sensor by modifying standard memory chip technology, the pixel size would be about 50 nm [6]. This value is far below the diffraction limit of the lens. Thus the imaging sensor is actually an oversampling device. We can use this spatial redundancy to compensate the information loss due to one-bit quantizer, as in oversampled analog-to-digital (A/D) conversions [7–10]. We also showed that the dynamic ranges of the binary sensor could be orders magnitude higher than those of conventional sensor.

One important thing missing in the previous work is the noise. In the previous work, we only considered shot noise. But dark current noise, thermal noise, and readout noise also effect the performance

of the image sensor. Since the pixel value in our sensor is binary, the influence of all these noise can be modeled as additive Bernoulli noise with a known parameter p_e , called noise rate. This can be estimated by covering the lens, taking a pictures, and computing the percentage of the “1”s in the binary image. In this paper, we focus on the case that threshold is a single photon. We present performance of the noisy binary sensor in terms of SNR. We show that the binary sensor is quite robust to noise for high light intensity values and the performance is only slightly worse due to noise. The binary sensor under noise still have much higher dynamic ranges comparing to the conventional sensor. We propose to use MLE to estimate the light intensity field, and show that when the threshold is equal to a single photon, the log-likelihood function is concave which ensures us to find the optimal solution.

In Section 2, we describe the noisy binary sensing model. In Section 3, we study the performance of the noisy binary sensor for estimating a constant light intensity. The MLE for estimating the light intensity field is presented in Section 4. Section 5 gives numerical results on both synthesized 1-D signal and images.

To simplify the notation in this paper, we focus our discussion on a one-dimensional (1-D) sensor array. All the results can be easily extended to the 2-D case.

2. IMAGING MODEL

We consider the problem of estimating the light intensity field using a binary image sensor as in [1]. Due to the low-pass effect of the lens, the light intensity field $\lambda(x)$ captured by the image sensor is a bandlimited signal. We use the same bandlimited light intensity field model as in [1], *i.e.*,

$$\lambda(x) = \frac{N}{\tau} \sum_{n=0}^{N-1} c_n \varphi(Nx - n), \quad (1)$$

where $\varphi(x)$ is a nonnegative interpolation kernel, N is a given integer, τ is the exposure time, $\{c_n : c_n \geq 0\}$ is a set of free variables, and the constant $\frac{N}{\tau}$ simplifies the expression in later analysis.

An image sensor with M binary pixels samples this light intensity field $\lambda(x)$. We define the oversampling factor K as the ratio between the number of pixels and the degree of freedom of the light intensity field $\lambda(x)$, *i.e.*, $K = \frac{M}{N}$. Let s_m be the light exposure of the m th pixel for given exposure time and surface area, $\beta(x)$ be the box function defined as,

$$\beta(x) \stackrel{\text{def}}{=} \begin{cases} 1, & \text{if } 0 \leq x \leq 1; \\ 0, & \text{otherwise,} \end{cases}$$

and define a discrete filter

$$g_m \stackrel{\text{def}}{=} \langle \varphi(x), \beta(Kx - m) \rangle, \quad m \in \mathbb{Z}.$$

This work was supported by the Swiss National Science Foundation under grant 200020-132730 and by the European Research Council under grant ERC-247006.

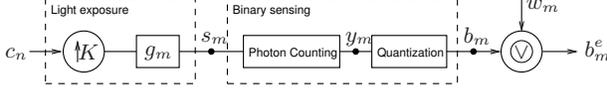


Fig. 1. The signal processing block diagram of the imaging model studied in this chapter. We upsample and filter the expansion coefficients c_n to get the light exposure value s_m at the m th pixel. Then, the binary image sensor converts $\{s_m\}$ into quantized measurements $\{b_m\}$. Due to noise $\{w_m\}$, we get the contaminated binary measurements $\{b_m^e\}$.

Then from [1], we know that the relation between the light exposure value for the m th pixel s_m and the free variables $\{c_n\}$ for the light intensity field $\lambda(x)$ is

$$s_m = \sum_n c_n g_{m-Kn}, \quad (2)$$

i.e., $\{s_m\}$ is obtained by first upsampling the free variables $\{c_n\}$ by a factor of K and then filtering with a discrete filter g_m as shown in Fig. 1.

Let $\mathbf{s} = [s_0, s_1, \dots, s_{M-1}]^T$ and $\mathbf{c} = [c_0, c_1, \dots, c_{N-1}]^T$, then the matrix-vector form for (2) is

$$\mathbf{s} = \mathbf{G} \mathbf{c}, \quad (3)$$

where \mathbf{G} is an $M \times N$ matrix denoting the combination of upsampling (by K) and filtering (by g_m) as depicted in Fig. 1. Each entry of \mathbf{s} is

$$s_m = \mathbf{e}_m^T \mathbf{G} \mathbf{c}, \quad (4)$$

where \mathbf{e}_m is the m th standard Euclidean basis vector.¹

Define y_m the number of photons received by the m th pixel. Then it is the realization of a Poisson random variable Y_m with intensity parameter s_m , i.e.,

$$\mathbb{P}(Y_m = y_m; s_m) = \frac{s_m^{y_m} e^{-s_m}}{y_m!}, \quad \text{for } y_m \in \mathbb{Z}^+ \cup \{0\}.$$

As shown in Fig. 1, we quantize y_m to get a binary pixel value. The quantizer is a binary quantizer with threshold q . If the number of photons received by the m th pixel y_m is larger or equal to q , the binary pixel value b_m will be “1”, otherwise “0”. For a single photon threshold, q is equal to “1”. The pixel value b_m is a realization of a random variable B_m . Introducing two functions,

$$p_0(s) \stackrel{\text{def}}{=} e^{-s} \quad \text{and} \quad p_1(s) \stackrel{\text{def}}{=} 1 - e^{-s}, \quad (5)$$

we can write

$$\mathbb{P}(B_m = b_m; s_m) = p_{b_m}(s_m), \quad b_m \in \{0, 1\}.$$

Since the pixel value b_m is binary, we model the noise w_m as the realization of a Bernoulli random variable W_m with parameter p_e , called noise rate, thus,

$$\mathbb{P}(W_m = w_m; p_e) = \begin{cases} p_e, & \text{if } w_m = 1, \\ 1 - p_e, & \text{otherwise,} \end{cases}$$

¹Here we use zero-based indexing. Thus, $\mathbf{e}_0 \stackrel{\text{def}}{=} [1, 0, \dots, 0]^T$, $\mathbf{e}_1 \stackrel{\text{def}}{=} [0, 1, \dots, 0]^T$, and so on.

The final output for the m th pixel is $b_m^e = b_m \vee w_m$, where $b_m^e \in \{0, 1\}$ and \vee is the disjunction binary operator. It is the realization of the random variable $B_m^e = B_m \vee W_m$. Introducing two functions,

$$p_0^e(s) \stackrel{\text{def}}{=} p_0(s)(1 - p_e) \quad \text{and} \quad p_1^e(s) \stackrel{\text{def}}{=} p_0(s)p_e + p_1(s), \quad (6)$$

then we have

$$\mathbb{P}(B_m^e = b_m^e; s_m, p_e) = p_{b_m^e}^e(s_m), \quad b_m^e \in \{0, 1\}. \quad (7)$$

3. PERFORMANCE ANALYSIS

In this section, we study the performance of the noisy binary image sensor for estimating the light intensity field and analyze the influence of the noise. We show that the noisy binary sensor is still better than the traditional sensor in terms of dynamic range with a reasonable noise rate. To simplify our analysis and derive closed-form solutions, we assume that the light intensity field is constant. Numerical results in Section 5 show that the conclusions hold for general linear models.

3.1. Closed-form MLE Solution for $q = 1$

In what follows, we derive the closed-form MLE solution when the threshold is $q = 1$. We assume that the light intensity is a constant value c .

Let $\mathcal{L}_b^e(c)$ be the likelihood function of observing K noisy binary sensor measurement $\mathbf{b}^e = [b_0^e, b_1^e, \dots, b_{M-1}^e]^T$. Then,

$$\begin{aligned} \mathcal{L}_b^e(c) &\stackrel{\text{def}}{=} \mathbb{P}(B_m^e = b_m^e, 0 \leq m < K; c, p_e), \\ &= \prod_{m=0}^{K-1} \mathbb{P}(B_m^e = b_m^e; c, p_e), \\ &= \prod_{m=0}^{K-1} p_{b_m^e}^e(c/K), \end{aligned} \quad (8)$$

where (8) is because each pixel counts the photons independently, and (9) is derived from (7). Denote by K_1 ($0 \leq K_1 \leq K$) the number of “1”s in the noisy binary sequence \mathbf{b}^e . Then (9) becomes

$$\mathcal{L}_b^e(c) = (p_1^e(c/K))^{K_1} (p_0^e(c/K))^{K-K_1}. \quad (10)$$

Given K noisy binary measurements \mathbf{b}^e , the MLE is to find the parameter c which can maximize the likelihood function $\mathcal{L}_b^e(c)$ in (10), i.e.,

$$\begin{aligned} \hat{c}_{\text{ML}}(\mathbf{b}^e) &\stackrel{\text{def}}{=} \arg \max_{0 \leq c \leq S} \mathcal{L}_b^e(c) \\ &= \arg \max_{0 \leq c \leq S} (p_1^e(c/K))^{K_1} (p_0^e(c/K))^{K-K_1}, \end{aligned} \quad (11)$$

where the upper and lower bound are used to make the solution physically meaningful, i.e., the light exposure value can not take negative value, and when the likelihood function $\mathcal{L}_b^e(c) = (p_1^e(c/K))^{K_1} (p_0^e(c/K))^{K-K_1}$ is monotonically increasing with respect to c , we can not set the light exposure value to be ∞ .

Lemma 1 When the threshold is $q = 1$, the solution to (11) is

$$\hat{c}_{\text{ML}}(\mathbf{b}^e) = \begin{cases} -K \ln \frac{K-K_1}{K(1-p_e)}, & \text{if } 0 < K_1 < K, \\ p_e < \min\{\frac{K_1}{K}, \frac{1}{2}\}, \\ 0, & \text{if } K_1 = 0 \text{ or } 0 < K_1 < K, \\ \frac{K_1}{K} \leq p_e < \frac{1}{2}, \\ S, & \text{if } K_1 = K \end{cases} \quad (12)$$

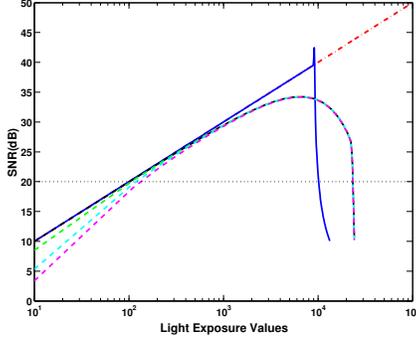


Fig. 2. Performance comparisons of three different sensing schemes (“BIN”, “IDEAL”, and “SAT”) over a wide range of light exposure values c (shown in logarithmic scale). The dash-dot line (in red) represents the “IDEAL” scheme with no quantization; The solid line corresponds to the “SAT” scheme with a saturation point set at $c = 9130$; The four dashed lines correspond to the “BIN” schemes with $q = 1$ and $K = 2^{12}$ and different noise rates (from far right to left, $p_e = 0, 0.001, 0.005$ and 0.01 , respectively).

where K_1 is the number of “1”s, and K is the total number of pixels.

3.2. The Influence of the Noise on the Dynamic Range

We denote our binary sensing scheme as “BIN”. We also compare our scheme with two other methods “IDEAL” and “SAT”. In the “IDEAL”, the pixel counts all the photons hitting on the pixel. The estimated light exposure value is just the number of the photons received by the pixel. The “SAT” scheme is similar to “IDEAL”, except that it has a saturation point C_{max} . We use signal-to-noise ratios (SNRs) to measure the performance. The SNR is defined as

$$\text{SNR} = 10 \log_{10} \frac{c^2}{\mathbb{E}[(\hat{c} - c)^2]},$$

where \hat{c} is the estimated light exposure value.

Let y be the number of photons impinging on a pixel. Then for the “IDEAL” scheme, as shown in [1], the MLE is $\hat{c}_{\text{IDEAL}}(y) = y$, and $\text{SNR}_{\text{IDEAL}} = 10 \log_{10}(c)$. For the “SAT” method, the sensor measurement is $y_{\text{SAT}} \stackrel{\text{def}}{=} \min\{y, C_{\max}\}$, and the estimator is $\hat{c}_{\text{SAT}}(y_{\text{SAT}}) = y_{\text{SAT}}$.

In Fig. 2, we show the SNR performance for “IDEAL”, “SAT”, and “BIN” with different noise rates. The dash-dot line in the figure corresponds to the “IDEAL” scheme. The solid line is for the “SAT” method. The four dashed lines represent the “BIN” scheme with fixed oversampling factor $K = 2^{12}$, and different noise rates (from far right to left $p_e = 0, 0.001, 0.005$, and 0.01 , respectively). We can see that the larger the noise rate p_e , the worse the SNR performance of the “BIN” scheme. We can also notice that the noise has more influence on lower light intensities. For the large light intensities, the SNR is almost the same for the noiseless and noisy cases. This indicates that our binary sensing method is quite robust to noise. Also note that the dynamic range of the noisy binary sensor still larger than the conventional sensor.

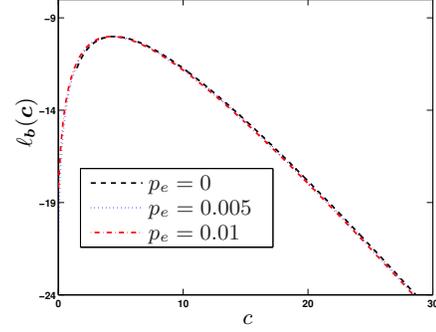


Fig. 3. The log-likelihood functions for constant light fields under the noisy case. The parameters are set as: threshold $q = 1$, oversampling factor $K = 20$, number of ones $K_1 = 4$.

4. IMAGE RECONSTRUCTION USING MAXIMUM-LIKELIHOOD ESTIMATOR

In the previous section, we derived the closed-form solution of the MLE for the constant light intensity field model when the threshold is $q = 1$. We extend the MLE to the general linear field model with arbitrary interpolation kernels. We show that for general light field model, the log-likelihood function is concave when $q = 1$. Thus we can find the optimal solution using iterative algorithms.

Similar to our derivations in (8) and (9), the likelihood function given M noisy binary measurements \mathbf{b}^e can be written as

$$\mathcal{L}_{\mathbf{b}}^e(\mathbf{c}) = \prod_{m=0}^{M-1} \mathbb{P}(B_m^e = b_m^e; s_m) = \prod_{m=0}^{M-1} p_{b_m^e}^e(\mathbf{e}_m^T \mathbf{G} \mathbf{c}), \quad (13)$$

where (13) follows from (7) and (4). We also define the log-likelihood function as

$$\ell_{\mathbf{b}}^e(\mathbf{c}) \stackrel{\text{def}}{=} \log \mathcal{L}_{\mathbf{b}}^e(\mathbf{c}) = \sum_{m=0}^{M-1} \log p_{b_m^e}^e(\mathbf{e}_m^T \mathbf{G} \mathbf{c}). \quad (14)$$

Given \mathbf{b}^e , the MLE is the parameter that maximizes $\mathcal{L}_{\mathbf{b}}^e(\mathbf{c})$, or $\ell_{\mathbf{b}}^e(\mathbf{c})$. Specifically,

$$\hat{\mathbf{c}}_{\text{ML}}(\mathbf{b}^e) \stackrel{\text{def}}{=} \arg \max_{\mathbf{c} \in [0, S]^N} \mathcal{L}_{\mathbf{b}}^e(\mathbf{c}) = \arg \max_{\mathbf{c} \in [0, S]^N} \ell_{\mathbf{b}}^e(\mathbf{c}) \quad (15)$$

The constraint $\mathbf{c} \in [0, S]^N$ means that every parameter c_n should satisfy $0 \leq c_n \leq S$, for some preset maximum value S .

Theorem 1 When the threshold is $q = 1$, for arbitrary noisy binary sensor measurements \mathbf{b}^e , the log-likelihood function $\ell_{\mathbf{b}}^e(\mathbf{c})$ defined in (14) is concave on the domain $\mathbf{c} \in [0, S]^N$.

Proof 1 When $q = 1$, $b_m^e = 0$, according to (6) and (5),

$$\log p_{b_m^e}^e(s) = \log((1 - p_e)e^{-s}) = \log(1 - p_e) - s,$$

which is a concave function.

When $q = 1$, $b_m^e = 1$,

$$\log p_{b_m^e}^e(s) = \log(1 - (1 - p_e)e^{-s}) = \log(e^s + p_e - 1) - s.$$

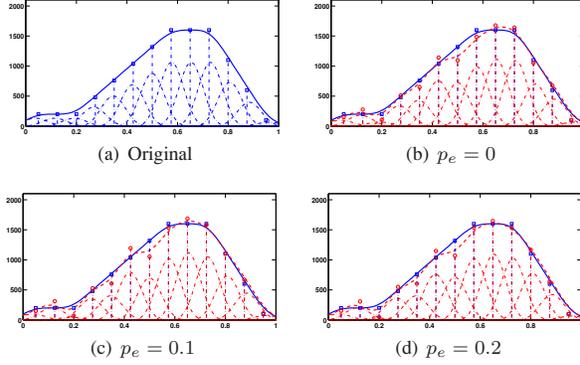


Fig. 4. Binary sensing and reconstructions of 1-D light fields with different noise rates. (a) The original light field $\lambda(x)$, modeled as a linear combination of shifted spline kernels. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1$, and 0.2 respectively. The oversampling factor is 1024.

The second derivative of this function is $\frac{e^s(p_e-1)}{(e^s+p_e-1)^2} \leq 0$, due to $0 \leq p_e \leq 1$. Thus when $q = 1$, $b_m^e = 1$, $\log p_{b_m^e}^e(s)$ is also a concave function.

Since the sum of concave functions is still concave and the composition of a concave function with a linear mapping ($s_m = e_m^T \mathbf{G} \mathbf{c}$) is still concave, we conclude that the log-likelihood function $\ell_b^e(\mathbf{c})$ defined in (14) is concave.

According to Theorem 1, we can find the optimal solution using iterative numerical methods even if there is noise.

For visualization purpose, we show the log-likelihood function $\ell_b^e(\mathbf{c})$ for a constant light intensity in Fig. 3. We can see that when the threshold is $q = 1$, the presence of noise affects the log-likelihood function only slightly. But the log-likelihood function is still concave as in the noiseless case.

5. NUMERICAL RESULTS

This section shows numerical results on synthesized 1-D signals and 2-D images. The results validate the performance analysis and the effectiveness of our proposed image reconstruction algorithm.

5.1. 1-D Synthetic Signals

Given expansion coefficients $\{c_n\}$ shown as blue dots in the Fig. 4(a), and the interpolation filter $\varphi(x)$ which is the cubic B-spline function, we generate a 1-D light field $\lambda(x)$, i.e., the blue line. As shown in Fig. 4(a), $\lambda(x)$ is a linear combination of the shifted kernel.

We first set the oversampling factor $K = 1024$. The reconstructed light intensity fields with the values of noise rate p_e are 0, 0.1, and 0.2 are shown in Fig. 4(a), Fig. 4(b), and Fig. 4(c) (in red), respectively. For comparison, the ground truth is given by the blue solid curve. We can see that when the noise rate increases, the performance becomes slightly worse. This obeys the performance analysis of Section 3.2, and shows the robustness of our proposed binary sensing scheme.

5.2. 2-D Synthetic Images

Consider a 2-D light intensity field as shown in Fig. 5(a). The values of the light intensity are in the range $[500, 2.5 \times 10^4]$. We simulate

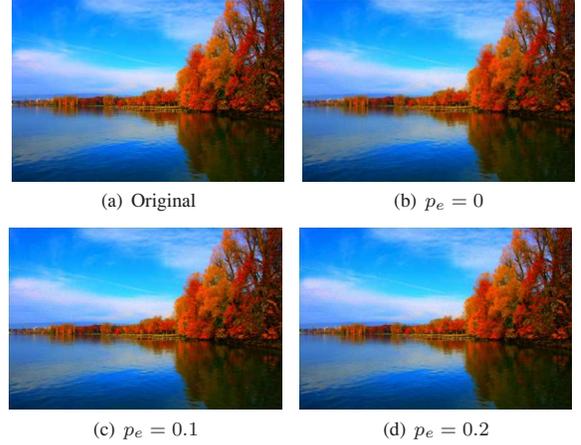


Fig. 5. Binary sensing and reconstructions of 2-D light fields with different noise levels. (a) The original light field. (b), (c), and (d) show the reconstruction results with different noise rates $p_e = 0, 0.1$, and 0.2 , respectively. The spatial oversampling factor is 8×8 , the temporal oversampling factor is 128.

the acquisition of this light intensity field using different noisy binary sensors. For the noise rate, we consider the cases $p_e = 0$, $p_e = 0.1$, and $p_e = 0.2$. The spatial oversampling factor of the binary sensor is set to 8×8 , and the temporal oversampling factor is 128 (i.e., 128 independent frames). Note that we have proved the equivalence between spatial and temporal oversampling in [1]. Similarly to our previous experiment on 1-D signals, we use a cubic B-spline kernel along each of the spatial dimensions. The reconstruction results for different noise cases are shown in Fig. 5. The MSE of the reconstruction results are shown in Table 1. We can see that the MSE of the noise case $p_e = 0.1$ is better than that of noiseless case. Since the noise can not change the binary measurements from “1” to “0”, the influence of noise when the light intensity is large is small. For a single experiment, there is a chance that the noise improves the estimation of large light intensity values. From the figures, we can see that our binary sensing scheme is quite robust to noise. We can hardly notice the presence of noise, although 10% or 20% of the binary measurements are contaminated by the noise. And this follows the analysis in the previous section.

Table 1. The MSE for different noise rates.

MSE	$p_e = 0$	$p_e = 0.1$	$p_e = 0.2$
$q = 1$	2.716×10^5	2.61×10^5	2.767×10^5

6. CONCLUSIONS

We worked on the noisy binary image sensor. The noise is modeled as additive Bernoulli noise with a known parameter, and it can only change the binary output from “0” to “1”. We showed that the noise had limited influence on the performance of the sensor and would slightly deteriorate the dynamic range. We used the MLE to estimate the light intensity function. When the threshold is a single photon, the log-likelihood function is still concave and the optimal solution can be computed using iterative algorithms. Future work may focus on the influence of arbitrary thresholds.

7. REFERENCES

- [1] F. Yang, Y.M. Lu, L. Sbaiz, and M. Vetterli, "Bits from photons: Oversampled image acquisition using binary Poisson statistics," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1421–1436, April 2012.
- [2] Gordon E. Moore, "Cramming more components onto integrated circuits," *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998.
- [3] A. J. P. Theuwissen, "CMOS image sensors: State-of-the-art," *Solid-State Electronics*, vol. 52, no. 9, pp. 1401–1406, Sept. 2008.
- [4] E. R. Fossum, "What to do with sub-diffraction-limit (SDL) pixels? – A proposal for a gigapixel digital film sensor (DFS)," in *IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, Nagano, June 2005, pp. 214–217.
- [5] L. Sbaiz, F. Yang, E. Charbon, S. Süssstrunk, and M. Vetterli, "The gigavision camera," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, April 2009, pp. 1093–1096.
- [6] Y. K. Park, S. H. Lee, J. W. Lee, et al., "Fully integrated 56 nm DRAM technology for 1 Gb DRAM," in *IEEE Symposium on VLSI Technology*, Kyoto, Japan, Jun. 2007.
- [7] J. C. Candy and G. C. Temes, *Oversampling Delta-Sigma Data Converters — Theory, Design and Simulation*, IEEE Press, New York, NY, 1992.
- [8] V. K. Goyal, M. Vetterli, and N. T. Thao, "Quantized overcomplete expansions in \mathbb{R}^n : Analysis, synthesis and algorithms," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 16–31, Jan. 1998.
- [9] P. T. Boufounos and A. V. Oppenheim, "Quantization noise shaping on arbitrary frame expansions," *EURASIP J. Applied Signal Processing*, pp. 1–12, 2006.
- [10] Z. Cvetković and I. Daubechies, "Single-Bit oversampled A/D conversion with exponential accuracy in the bit rate," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3979–3989, Nov. 2007.