INTERNATIONAL JOURNAL OF CIRCUIT THEORY AND APPLICATIONS
Int. J. Circ. Theor. Appl. (2013)
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cta.1969

Compressive image acquisition in modern CMOS IC design

Nikola Katic[1,*,†], Mahdad Hosseini Kamal[2], Alexandre Schmid[1], Pierre Vandergheynst[2] and Yusuf Leblebici[1]

[1]*Microelectronic Systems Laboratory (LSM), Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland*
[2]*Signal Processing Laboratory (LTS2), Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland*

SUMMARY

Compressive sampling (CS) offers bandwidth, power, and memory size reduction compared to conventional (Nyquist) sampling. These are very attractive features for the design of modern complementary metal-oxide semiconductor (CMOS) image sensors, cameras, and camera systems. However, very few integrated circuit (IC) designs based on CS exist because of the missing link between the well-established CS theory on one side, and the practical aspects/effects related to physical IC design on the other side. This paper focuses on the application of compressed image acquisition in CMOS image sensor integrated circuit design. A new CS scheme is proposed, which is suited for hardware implementation in CMOS IC design. All the main physical non-idealities are explained and carefully modeled. Their influences on the acquired image quality are analyzed in the general case and quantified for the case of the proposed CS scheme. The presented methodology can also be used for different CS schemes and as a general guideline in future CS based CMOS image sensor designs. Copyright © 2013 John Wiley & Sons, Ltd.

Received 15 May 2013; Revised 31 August 2013; Accepted 30 October 2013

KEY WORDS: CMOS image sensors; compressive sampling; image processing; image compression; IC design

1. INTRODUCTION

Modern applications of complementary metal-oxide semiconductor (CMOS) image sensors, cameras, and camera systems often require battery powered (remote) operation and a large amount of memory for the image/video storage. Introducing online image compression before image storage is very important for modern multi-camera systems where the required data bandwidth, power consumption, and the memory size can each form a severe bottleneck in the system design. This is especially the case in wireless video sensor networks and ultra-low-power biomedical applications (such as wireless camera-pills) where high sensor efficiency is required. In order to achieve higher sensor efficiency, several different methods of image compression on the focal plane have been presented in literature. Many sensors report traditional compression methods based on discrete cosine transform (DCT) and discrete wavelet transform (DWT). Although performing DCT and DWT is mainly reported on very small pixel arrays due to high hardware complexity, a few designs report successful compression on larger arrays (see [1] and [2]). Using alternate scanning patterns, such as the Z-curve or Hilbert curve [3], followed by quadtree decomposition results in the compressive image acquisition schemes presented in [4], [5], and [6]. The work presented in [7] uses a random scan pattern (chaotic scan) in order to achieve compressive acquisition. These techniques can result in very efficient compression levels of less than 1 bit per pixel. Finally, the sensors in [8] and [9]

*Correspondence to: Nikola Katic, Microelectronic Systems Laboratory (LSM), Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland.
†E-mail: nikola.katic@epfl.ch

use address event representation ([10]), in order to discard the redundant video information and only capture active events. This method is very convenient for machine vision and surveillance applications.

Compressive sampling (CS) intrinsically requires less bandwidth than the traditional (Nyquist) sampling and directly compresses images during the acquisition. Hence, CS reduces the power consumption, bandwidth, and the memory requirements. Because of the reduction in the overall image acquisition cost, CS theory has recently attracted significant attention [11], and the potential of CS in the integrated circuit (IC) design has been discussed in literature [12]. Being a relatively new concept and in spite of its expected benefits, CS is still far from the mainstream IC design practice. This work offers a practical approach to the theoretical features of CS from the perspective of IC designers. By exhaustive modeling, it explores the influence that physical implementation effects have on the compressed image acquisition. The presented results offer an insight on the potential benefits and limitations of compressive sampling.

Several image sensor designs based on CS have been presented in literature. The optical domain CS imager in [13] suffers from a large camera size and limited resolution. The CS imager in [14] is very sensitive to mismatch and process variation effects and has limited frame rate because the pixel output currents are responsible for charging and discharging the column-line parasitic capacitance. The separable-transform image sensor in [15] allows high readout speed and camera resolution but suffers from device mismatch effects in the analog-matrix generation as well as high readout noise. Another separable-transform imager [16] can perform CS; however, it is limited to very small pixel blocks. The most promising design proposed in [17] implements the first reported single-shot imaging CS acquisition that is suitable for video applications. The design benefits from CS in terms of high frame rate and overall power reduction but still consumes relatively high power for some applications (93–96 mW, depending on the frame rate). All the reported designs gain in terms of power consumption, compression factor (bandwidth reduction) and/or memory usage, but trade these benefits for relatively poor raw image quality. This is expectable considering that existing CS algorithms cannot yet support the utilization of CS in high quality imaging applications. However, techniques such as multiple capture can use high oversampling and different exposure times to achieve high video quality even with low quality raw images (frames). Moreover, high frame rates are desirable for the implementation of many novel post-processing video applications such as motion vector estimation, multiple capture, motion capture, pattern recognition, and so on.

2. COMPRESSIVE SAMPLING

2.1. General methods

Sparse signals can be reconstructed from significantly fewer measurements (coefficients or samples) than required using Nyquist sampling, which only constrains signals to be band-limited. Although images are not sparse in their natural (spatial) domain, they are known to be sparse in other domains. For example, as they mainly contain low-frequency components, the DCT can be used to represent the images in the DCT domain where they are sparse (the main principle of JPEG compression). DWT can also be used to represent images in the wavelet domain where the image sparsity is even higher (JPEG2000). If a signal is represented in a domain where it is sparse (or mainly 'consists of' zero values), the formed concentrated signal can be sampled with incoherent (spread-out and random) measurements. This process can be understood as a dual case to Nyquist sampling, where the signal, which is spread-out in a certain domain, is sampled by the measurement window, which is concentrated in the same domain (e.g., time). Hence, instead of band-limitation, CS exploits signal sparsity to perform an accurate signal acquisition from a small number of linear measurements.

Based on the previous discussion, the complete CS model is illustrated in Figure 1 (see also [18]). An image $\mathbf{x} \in \mathbb{R}^n$ is represented as a vectorized $n \times 1$ version of a $N_{row} \times N_{col}$ image frame with a total number of pixels equal to n ($n = N_{row} \times N_{col}$, where $N_{row/col}$ is the number of image rows/columns). An orthonormal basis (e.g., DCT or DWT) $\Psi \in \mathbb{R}^{n \times n}$ is used to transform the image into a domain where it forms a sparse signal (vector) α, $\mathbf{x} = \Psi\alpha$. Incoherent (noisy) linear measurements $\mathbf{y}$ are obtained by

DOI: 10.1002/cta

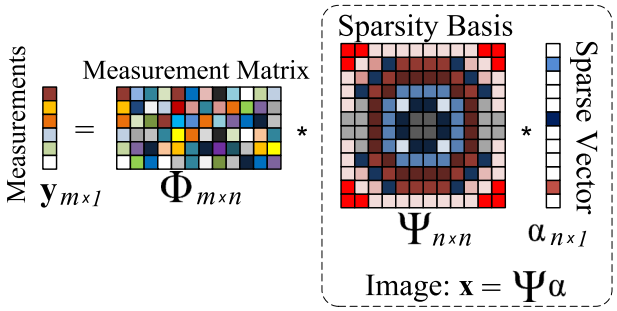

Figure 1. Compressive sampling.

applying a random measurement matrix Φ to the original signal $\mathbf{x}$. If the number of required measurements is equal to m, then $\mathbf{y} \in \mathbb{R}^m$ and $\Phi \in \mathbb{R}^{m\times n}$. Because α is a sparse vector and has a small number of unknowns, it is possible to reconstruct the signal with a relatively small number of measurements $m \ll n$. The ratio between m and n represents the compression rate of the CS, because the number of samples required by Nyquist image sampling is equal to n (total pixel count).

The CS theory states that the measurement matrix $\mathbf{\Phi}$ should satisfy the restricted isometry property (see [18]) in order to perfectly recover a signal from m linear samples. This constraint is fulfilled by matrices that are randomly generated from a Gaussian distribution or several other random distributions, such as Bernoulli and Rademacher distributions. In this paper, the Rademacher distribution is used to generate the measurement matrix. However, the conventional method is modified, and an improved scheme is used and analyzed instead. The modification improves the speed of the reconstruction algorithm and allows an efficient hardware implementation, which is compliant with common CMOS image sensor design techniques. The reconstruction algorithm benefit is explained in the next subsection. In Section 6, the benefits of the new scheme related to the hardware implementation are discussed.

The image reconstruction from the acquired measurements requires solving an inverse equation: $\mathbf{x} = \Phi^{-1} \cdot \mathbf{y}$. Thanks to the sparsity of α, the problem reduces to finding the sparsest α such that $\Phi\Psi\alpha = \mathbf{y}$. The measure of signal (or vector) sparseness can be mathematically defined by using the l_0 or l_1 norms in the Ψ domain. Minimizing the norm and performing an iterative algorithm finally results in finding the correct α. It should be mentioned that using the l_0 norm results in a combinatorial NP-hard optimization problem. Hence, defining the sparseness using the l_1 norm is significantly more effective as it results in a typical linear programming basis pursuit problem. Still, the CS reconstruction algorithms are complex and require a significant amount of time in order to converge to the correct solution. The conversion time depends on the specific algorithm, convergence criterion, image size and content, and the measurement noise intensity. In practice, the frame reconstruction can last from minutes to hours depending on the mentioned parameters. As a consequence, image frame reconstruction is performed off-line using any of the currently existing algorithms.

For images, the sparseness can be alternatively defined by using the total variation (TV), which is the measure of image gradient. Using TV eliminates the need for the sparsity basis Ψ and solves the optimization problem by finding the image $\mathbf{x}$ directly. Minimizing the image TV also results in the sparsest solution with the same algorithm complexity (basis pursuit) as minimizing the l_1 norm. A formal link can also be made between the TV and l_1 norm, because the TV can be interpreted as the l_1 of the appropriately discretized gradient. In this work, the image $\mathbf{x}$ is directly recovered from the measurements, using TV minimization:

$$\arg\min_{x\in\mathbb{R}^N} \left\{ \|x\|_{TV} + \lambda \|y - \Phi x\|_2^2 \right\}, \tag{1}$$

where λ is an empirical regularization parameter and $\|\mathbf{x}\|_{TV}$ is the TV semi norm [19]. Parameter λ controls image smoothing and must be adapted to the overall image noise. Typically, the values

DOI: 10.1002/cta

used in this work are in the range [0.05,0.1]. The iterative reconstruction algorithm is directly defined by using a set of updating rules from the fast iterative shrinkage-thresholding algorithm (FISTA). The details of the FISTA implementation and the exact set of updating rules can be found in [20].

2.2. *Modified method*

In the conventional Rademacher method, the matrix Φ is completely populated with either 1 or -1 (Rademacher distribution). This results in a dense and large matrix Φ with size $m \times n$, where $n = N_{row} \times N_{col}$ is the total pixel count. In each iteration, the reconstruction algorithm requires matrix multiplications with Φ and Φ^T resulting in the multiplication complexity of the order $O(N^4)$ (without loss of generality, we can define $N = N_{row} = N_{col}$). To reduce the order of the required multiplication, the modified method applies the same random vector to each image column. This is illustrated in Figure 2, where the square-board corresponds to a single image frame, which has each of its pixels multiplied by either $+1$ (black square) or -1 (white square). Expressed in mathematical form, this technique results in the modified matrix Φ:

$$\Phi_{m\times n} = \begin{bmatrix} \Xi & 0 & 0 & \cdots & 0 \\ 0 & \Xi & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \Xi \end{bmatrix}, \Xi = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,N} \\ \phi_{2,1} & \phi_{2,2} & \cdots & \phi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{M,1} & \phi_{M,2} & \cdots & \phi_{M,N} \end{bmatrix} \tag{2}$$

where $\phi_{ij} = \pm 1/N$ and $m = M \times N$. The block-diagonal structure of the new measurement matrix reduces the multiplications with Φ and Φ^T into N multiplications with Ξ and Ξ^T, which results in the total order of $O(N^3)$. Although the reduction from the order $O(N^4)$ to the order $O(N^3)$ does not appear significant when defined mathematically, the actual required program memory and the reconstruction time of the algorithm are substantially reduced. In practice, the reconstruction time can be reduced from several hours to only a few minutes using this technique while maintaining the same performance as the conventional Rademacher method. The benefit of the new scheme improves with the size of the pixel array, which allows hardware design scalability. Moreover, the column-parallel measurement technique allows a straightforward hardware implementation and reduces in-pixel hardware and matrix generation hardware complexity, which is further explained in Section 6.

The simulation results are shown in Figure 3 for the two different conventional methods and the newly proposed method using the sample image 'Cameraman', showing the peak signal-to-noise ratio (PSNR) values for each. Random selection forms the measurements by randomly multiplying the pixel values by either 1 or 0 and creating a sum over the entire image. The matrix Φ is therefore randomly populated with ones and zeros from the Bernoulli distribution, which is very convenient for hardware implementation. Random convolution forms the random measurements by using a diagonal random phase matrix and the discrete Fourier matrix (see [21]). As expected, random convolution (Figure 3(c)) provides the best image quality as it is a more optimal compression model

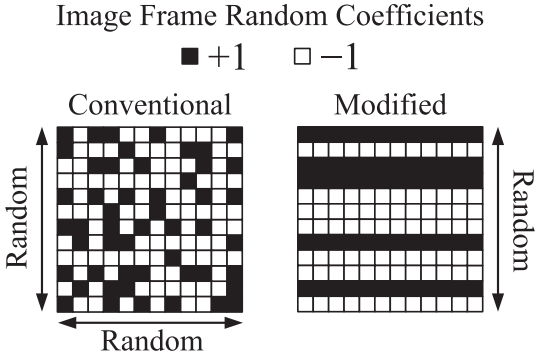

Figure 2. Modified Rademacher method.

DOI: 10.1002/cta

(a) (b)

(c) (d)

Figure 3. (a) Original image; (b) random selection: $PSNR = 23.6$ dB; (c) random convolution: $PSNR = 26.3$ dB; and (d) proposed scheme: $PSNR = 25.2$ dB.

than the random selection (Figure 3(b)) and the presented Rademacher method. However, random convolution is significantly more complex to implement in hardware. It requires randomization of the phase, and thus the measurement matrix in the form $F^{-1} \cdot D \cdot F$, where D is a diagonal phase randomization (random) matrix and F is the discrete Fourier matrix. The matrix generation and the multiplication with the full-frame are hard to obtain in hardware and do not result in a well-scalable design. The proposed modification of the Rademacher method (Figure 3(d)) provides the same result in terms of PSNR (25.2 dB) as the conventional Rademacher scheme, which is noticeably higher than the random selection PSNR (23.6 dB). The reconstructed image of the proposed scheme (Figure 3(d)) also preserves more details of the scene than the random selection while some details also appear more clear than in the case of random convolution. Even though the overall noise is low compared to random selection, for the human eye, a single frame (Figure 3(d)) may appear to be less smooth. In video applications, this appears as high-frequency low-correlated temporal noise and is mainly filtered by the low-pass characteristics of the human eye. Moreover, the level of image smoothing can be controlled by the parameter λ (see expression (1)), in which case it will be traded versus the detail preservation.

The proposed scheme represents a very good compromise between the possibility of compact hardware implementation and the final image quality. In the following sections, the details about the modeling of the physical design effects and the hardware implementation are presented. Physical modeling is developed for the proposed scheme (Sections 4 and 5), but the methodology and the majority of the results can be easily reused for any other compressive sampling scheme.

3. THEORETICAL AND MODELING FUNDAMENTALS

The physical design of CMOS image sensors introduces a set of non-idealities that affect the final performance of the image acquisition. These non-idealities such as gain errors or temporal noise are well characterized and understood both intuitively and theoretically for the case of Nyquist sampling. However, compressive sampling represents a completely new acquisition method to

DOI: 10.1002/cta

which the effects of non-idealities are not yet well characterized. For some of these effects, such as pixel offset and pixel temporal noise, a detailed mathematical analysis can be derived independently of the image content. For other effects, such as pixel gain error or dark current, a-priori knowledge of the image content is required, and hence, only a statistical estimate can be obtained. In order to mathematically describe the influence of the pixel-level non-idealities that depend on the image content, the image signal amplitude is assumed to be normalized and uniformly distributed over the interval [0,1] (i.e., $U(0,1)$) in the following analysis. This assumption basically defines an image as a set of normalized and limited amplitude values, where every amplitude may appear with an equal probability. The actual nature of the image content is modeled very closely under this assumption. Furthermore, it allows performing a general analysis and the statistical estimation of the non-ideality effects, which depend on the image content.

The mathematical solution or statistical estimates are always bound to the specific CS scheme. More precisely, they are bound to the sparsity basis, the applied reconstruction method, and the probability distribution from which the measurement matrix is generated. Hence, in this work, the obtained results are only applied to the example of the proposed scheme (Subsection 2.2). However, the methodology for obtaining the theoretical results and the described model are applicable for any CS scheme with the adjustment of the measurements matrix Φ probability distribution. The modeling derivations presented in this work assume the Rademacher distribution of the measurement matrix and the uniform distribution of image amplitudes as previously explained. In the developed circuit model, the pixel array has a size of 256×256, and the readout follows a column-parallel scheme, which is a common trend in modern CMOS image sensors. In addition, the proposed modified Rademacher scheme is used, which performs the CS on each column separately with the same random sequence. The reconstruction is performed using the FISTA algorithm [20]. All the major non-idealities are addressed in the following sections. They are separated in two subgroups: pixel-level non-idealities and readout-level non-idealities.

4. PIXEL-LEVEL NON-IDEALITIES

Pixel-level non-idealities comprise all physical photo-detector and pixel circuit imperfections that cause deviations from the ideal signal. Intuitively, because they modify the image before the image compression by CS, they should cause similar effects as in the Nyquist sampling schemes. Theoretically, as long as they preserve the image sparsity, the image should be reconstructed correctly. However, careful modeling is required to quantitatively characterize the effect on image quality.

Every pixel-level non-ideality will cause a certain error in the pixel output. In each pixel, the error is multiplied by a corresponding random value of the measurement matrix drawn from a certain distribution. The output error e_{out} can therefore be formed as a product of two random variables where variable e_X is the normalized pixel error and variable W is the random coefficient of the measurement matrix. Because random variables e_X and W are independent, the variance of the product can be calculated from the following expression:

$$Var(e_{out}) = E(e_X)^2 Var(W) + E(W)^2 Var(e_X) + Var(e_X)Var(W). \tag{3}$$

Expression (3) provides two important hints for the circuit design phase. In order to minimize the effect of the pixel-level non-idealities, the mean of the random measurement matrix should preferably be equal to 0 ($E(W)=0$). Hence, the Rademacher distribution provides more attenuation of the pixel-level errors than, for example, the Bernoulli distribution matrices with $E(W)\neq 0$. Furthermore, it is preferable that the mean of the non-ideality error e_X is close to 0 as well. This is further discussed within the context of the pixel offset fixed pattern noise (FPN) (Subsection 4.3.), where it relates to the minimization of the systematic offset in the design phase. If $E(W)=0$ and the pixel-level non-ideality error has approximately zero mean, expression (3) reduces to $Var(e_{out}) \approx Var(e_X)Var(W)$.

DOI: 10.1002/cta

In the CS scheme, the measurements are formed as the sum of multiple pixel randomized outputs. Because the random measurement matrix values are mutually independent, the randomized pixel outputs can also be considered independent. According to the central limit theorem (CLT), the measurement error caused by the sum of multiple pixel errors is approximately normally distributed. Moreover, the variance of the total measurement error is the sum of the individual variances:

$$Var(e_y) = \sigma_e^2 = N \cdot Var(e_{out}) = N \cdot \sigma_{out}^2. \tag{4}$$

Substituting expression (3) into (4), the final result can be interpreted as follows: *If the random measurement matrix values are uncorrelated, any pixel-level error in any CS scheme can be approximately modeled as an output-referred White Gaussian Noise (WGN) of the image measurements y (e_y), with the variance given as*

$$Var(e_y) = N \cdot \left[E(e_X)^2 Var(W) + E(W)^2 Var(e_X) + Var(e_X) Var(W) \right], \tag{5}$$

and the mean given as

$$E(e_y) = N \cdot E(e_X) \cdot E(W), \tag{6}$$

where variable e_X is the pixel error due to physical non-ideality, variable W is the random coefficient of the measurement matrix and N is the number of the summed pixel outputs for the single measurement.

In CS schemes, pixel amplitudes are multiplied by random coefficients and summed together to form signal measurements. This enables amplitudes to be scaled in a similar way as previously described for errors (noise). If the image is assumed to be uniformly distributed ($U(0,1)$), and X is considered to be the image signal, according to the CLT, the image measurements y follow a normal distribution with the variance:

$$Var(y) = N \cdot \left[E(X)^2 Var(W) + E(W)^2 Var(X) + Var(X) Var(W) \right], \tag{7}$$

and the mean

$$E(y) = N \cdot E(X) \cdot E(W). \tag{8}$$

Hence, assuming $E(W) = 0$, the measurement amplitudes also scale with a factor $\sqrt{N}$ as errors do, if N pixels are combined together. Moreover, most of the amplitudes are in the range of $[-3\sigma_y, 3\sigma_y]$ where $\sigma_y = \sqrt{Var(y)}$. Therefore, $3\sigma_y$ is used as a maximum measurement amplitude (y_{max}) in the following analysis for normalizing the output-referred noise (error) relatively to the measurement amplitude. This is also required considering that the readout circuit has a limited output voltage range. Choosing $[-3\sigma_y, 3\sigma_y]$ to cover the full swing at the output is a reasonable design choice. More detail explanation related to the readout output saturation is covered in Subsection 5.5.

4.1. Dark current

Dark Current is a physical effect typical to all semiconductor photo-detectors. It represents the current (signal) level observed when the received light is equal to 0, that is, when there is no charge carrier generation induced by photons. Dark current is caused by multiple physical phenomena such as diffusion of carriers, carrier generation in the depleting region and carrier tunneling. Even though it has a complex physical origin, dark current can be simply modeled by setting a fixed boundary to the minimum detectable signal. More specifically, if the image signal amplitude u satisfies $u < \Delta i_{dark}$, then the resulting signal amplitude is $u = \Delta i_{dark}$, where Δi_{dark} is the normalized dark current with respect to the normalized maximum signal amplitude (equal to 1). If the image is

DOI: 10.1002/cta

assumed to be uniformly distributed ($U[0,1]$), the error caused by dark current can therefore directly be represented by the following probability density function (PDF):

$$p_{dark}(x) = \begin{cases} (1-\Delta i_{dark})\delta(x) & \text{if } x = 0 \\ 1 & \text{if } 0 < x \leq \Delta i_{dark} \\ 0 & elsewhere, \end{cases} \tag{9}$$

where $\delta(x)$ is the Dirac delta function. The given PDF is derived from the fact that the dark current error is uniformly distributed in the range $(0, \Delta i_{dark})$ and is equal to 0 for all the image signal amplitudes higher than Δi_{dark} (Dirac function).

In the case of the proposed scheme, random measurement coefficients are generated from the Rademacher distribution ($E(W)=0, Var(W)=1$). The resulting variance of the randomized pixel output error can be derived using the resulting dark current error PDF (Figure 4 shows the PDF after multiplication with random coefficients) and expression (3):

$$Var(e_{out}) = \sigma_{dark}^2 = \frac{1}{2}\Delta i_{dark}^2 + \frac{1}{3}\Delta i_{dark}^3 \tag{10}$$

Because the Rademacher distribution has a mean equal to 0, the mean value of the randomized error is also equal to 0 (see expression (6) and Figure 4). Moreover, for relatively small values of Δi_{dark}, the output error variance can be considered to be $Var(e_{out}) \approx \frac{1}{2}\Delta i_{dark}^2$. In the case of the proposed scheme, where N pixels are summed within the column, the resulting measurement error variance is

$$Var(e_y) = \sigma_e^2 = \frac{N}{2}\Delta i_{dark}^2 \tag{11}$$

with the zero mean and normal distribution. For the Rademacher scheme, using expression (7), and assuming the uniform distribution of the input image as discussed, the maximum measurement amplitude is

$$y = 3\sigma_y = \sqrt{3}\sqrt{N}\cdot\max(x) \tag{12}$$

where $\max(x)=1$ is the maximum of the normalized pixel amplitude. The error caused by the dark current therefore results in the output-referred noise with zero mean and the normalized standard deviation:

$$\frac{\sigma_e}{y} = \frac{\Delta i_{dark}}{\sqrt{6}}. \tag{13}$$

For the proposed CS scheme, the derived theoretical result (13) of the output-referred noise is shown in Figure 5 versus the conventional (directly implemented) model of the dark current simulated on a

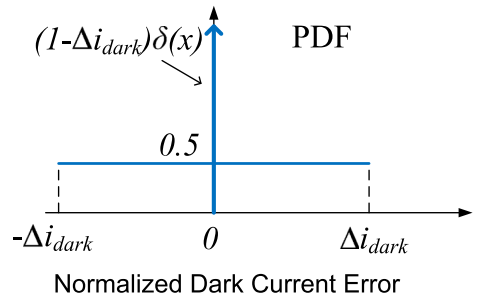

Figure 4. Dark current error probability density function (PDF).

DOI: 10.1002/cta

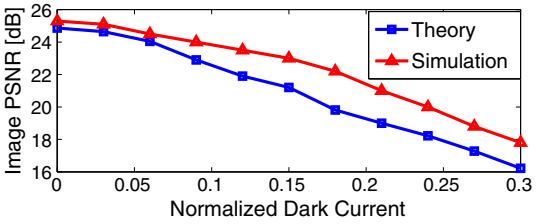

Figure 5. Dark current influence on image peak signal-to-noise ratio (PSNR).

sample image 'Cameraman'. The results are still stochastic thanks to the random nature of the measurement matrix and are averaged over five independent image frame reconstructions. The horizontal axis represents Δi_{dark} (normalized dark current value), while vertical axis depicts the resulting image PSNR. The derived theoretical result shows a relatively pessimistic PSNR compared to the direct dark current model simulation due to relatively low number of 'dark' pixels in the benchmark image (Cameraman). Nevertheless, the theoretical curve is a more reliable option in a conservative design approach as it depends less on the chosen benchmark image and is closer to the general case. Nevertheless, both curves show a similar trend. The influence of the dark current causes a noticeable decrease in image quality only at very high dark current levels (10% of the maximum signal), which is almost never observed in practice except for some specific applications such as astronomical imaging, or if high-temperature operation is required. This low sensitivity to the dark current is an expected result from the perspective that the image quality is initially relatively low due to compression.

4.2. *Dark-signal-non-uniformity*

Dark-signal-non-uniformity (DSNU) is the statistical variation of the photo-detector dark current due to fabrication processing imperfections such as doping variation or impurity atoms in the silicon lattice. Alike the dark current, DSNU also has to be taken into account when calculating the overall signal-to-noise ratio (SNR) and the dynamic range of the image sensor. Physically, DSNU depends on multiple uncorrelated random variables and can therefore be approximately modeled as a random variation of dark current generated from the Gaussian distribution. The direct modeling of the DSNU is included in the model developed in this work. However, the isolated influence of the DSNU on the image quality for the existing CS schemes is negligible, because it is a small variation of the dark current, which initially has a very low effect on image quality. Hence, there is no practical benefit in deriving the isolated expressions for output-referred noise caused by DSNU and presenting the resulting PSNR, which is practically unaffected.

4.3. *Offset fixed pattern noise*

Offset FPN is the pixel response spatial variation that results in pixel-to-pixel offset in the image frame. It is caused by device parameter variations due to fabrication imperfections in the pixel circuit (readout mismatch is discussed separately in the next section). A minor portion of the pixel offset FPN is a consequence of DSNU, which was already discussed. In the case of conventional active pixels, the main contributor to the overall offset FPN is the pixel source-follower output buffer. Physically, the main causes of offset FPN are standard metal-oxide semiconductor device variations such as variations in doping, oxide thickness, and device dimension variation, which can be considered as uncorrelated random variables. Because it originates from multiple uncorrelated random processes in multiple devices, the overall output error can be approximately modeled as a random variation extracted from the Gaussian distribution (according to CLT). In the pixel design phase, if the pixel circuit and the fabrication process are known, the offset FPN standard deviation can be estimated from circuit-level Monte-Carlo simulations. For the proposed scheme, the error caused by offset FPN is multiplied with the random coefficients from the Rademacher distribution. The variance of e_{out} can be derived using expression (3):

DOI: 10.1002/cta

$$Var(e_{out}) = \sigma_{offset}^2 = \mu_{ex}^2 + \sigma_{ex}^2 \tag{14}$$

where μ_{ex} and σ_{ex} are the pixel offset mean and standard deviation, respectively (both normalized relatively to the maximum image signal amplitude $\max(x)=1$ because $x \sim U(0,1)$). Based on equation (14), it should be noted that the systematic offset μ_{ex} has the same effect on the overall pixel error as the random offset σ_{ex}. Because systematic offset can be controlled by appropriate design practices, in CS schemes it should be minimized during the design phase to reduce the variation of the pixel output error. This represents a major difference compared to Nyquist sampling, where the systematic offset can be easily canceled by post-fabrication calibration (if not already canceled by circuit design techniques). In CS, unless it is minimized or canceled by circuit design techniques, the systematic offset appears as noise in the image measurements and therefore in the reconstructed image as well. Hence, no post-fabrication calibration of the systematic offset is possible in CS.

Summing N pixel outputs to form the CS measurement results in

$$Var(e_y) = \sigma_e^2 = N \cdot \left(\mu_{ex}^2 + \sigma_{ex}^2\right). \tag{15}$$

If the systematic offset is considered to be minimized ($\mu_{ex}^2 \approx 0$), and the CS measurement amplitude is given by (12), the output-referred normalized measurement error standard deviation due to pixel offset FPN is

$$\frac{\sigma_e}{y} = \frac{\sigma_{ex}}{\sqrt{3}}. \tag{16}$$

The influence of offset FPN on the proposed scheme image quality is presented in Figure 6, for a sample image 'Cameraman'. The results are averaged over five independent image frame reconstructions. The x-axis represents the normalized standard deviation of the pixel offset σ_{ex}. The theoretical curve based on (16) and the simulated curve based on direct modeling of the pixel offset match significantly better than the dark current curves do (Figure 5). This is an expected result because the overall dark current error depends on the benchmark image content, while the pixel offset error is scene-independent.

For the sake of comparing various effects, the 2 *dB* of image PSNR loss is chosen as an acceptable limit, because in the used examples it does not cause an observable difference to the human eye. A 2 dB PSNR loss is observed when the offset FPN standard deviation is approximately 5% of the maximum signal amplitude. In CMOS image sensors, offset FPN is mainly canceled using offset and noise cancelation techniques such as correlated double sampling (CDS). If CDS is used, the offset FPN is not a major concern in CS image acquisition. Without offset cancelation techniques, a careful pixel design is required to minimize the systematic offset and keep the random offset due to device mismatch below the preferable limit for the targeted image quality.

4.4. Photo-response non-uniformity

Photo-response non-uniformity is another form of FPN (or spatial variation), which manifests itself as a random pixel gain error due to device mismatch in the fabrication process. PRNU cannot be canceled

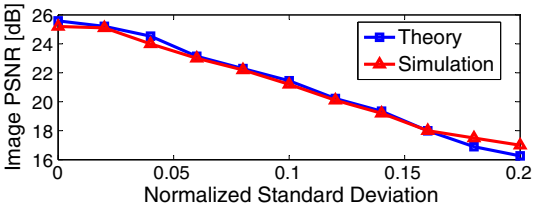

Figure 6. Pixel offset fixed pattern noise (FPN) influence on image peak signal-to-noise ratio (PSNR).

DOI: 10.1002/cta

by using conventional FPN cancelation techniques as for offset FPN. However, it can be reduced by more complex circuit design techniques. Nevertheless, similarly as offset FPN, it also depends on multiple independent random variables. Hence, it can approximately be modeled as a random gain variation with the Gaussian distribution. The normalized gain error of each pixel can be represented as $\frac{\Delta G}{G}$, if the gain error is represented as ΔG, and the nominal gain value is G. The corresponding signal error is therefore the product of the normalized signal amplitude and the normalized gain error $\frac{\Delta G}{G}$. Hence, if the image is considered to be uniformly distributed ($U(0,1)$), the signal error is also uniformly distributed over the interval $\left[0, \frac{\Delta G}{G}\right]$. From the signal error PDF, the randomized pixel output signal error variance can be calculated using (3):

$$Var(e_{out}) = \mu_{e_x}^2 + \sigma_{e_x}^2 = \frac{\left(\frac{\Delta G}{G}\right)^2}{4} + \frac{\left(\frac{\Delta G}{G}\right)^2}{3} = \frac{7}{12}\left(\frac{\Delta G}{G}\right)^2. \tag{17}$$

Summing N pixel outputs to form the CS measurement results in

$$Var(e_y) = \sum_{i=1}^{N} Var(e_{out,i}) = \frac{7}{12}\sum_{i=1}^{N}\left(\frac{\Delta G_i}{G}\right)^2, \tag{18}$$

where $e_{out,i}$ and $\frac{\Delta G_i}{G}$ are the randomized output error and the normalized gain error of pixel i, respectively. Assuming that the gain error has zero mean, the expected variance is

$$E\left[Var(e_y)\right] = \frac{7}{12}\sum_{i=1}^{N} E\left[\left(\frac{\Delta G_i}{G}\right)^2\right] = \frac{7}{12}\cdot N\cdot\sigma_g^2, \tag{19}$$

where σ_g is the normalized gain error standard deviation. Using (12), the normalized measurement error standard deviation is expressed as

$$\frac{\sigma_e}{y} = \frac{\sqrt{7}}{6}\sigma_g. \tag{20}$$

As mentioned for the offset FPN standard deviation σ_{e_x}, the gain error standard deviation σ_g can also be extracted from Monte-Carlo circuit-level simulations during the design phase.

The simulated effect of PRNU on the proposed scheme's image quality is presented in Figure 7. The x-axis represents the normalized gain error standard deviation σ_g. The theoretical curve is based on (20), while the simulations are performed by modeling the gain error directly from the Gaussian distribution and including it into the system model. Similar as for dark current, because of the simulation result dependence on the benchmark image, the results of direct PRNU simulation show slightly higher PSNR. Nevertheless, for both curves, the PSNR loss of 2 dB is observed when the PRNU standard deviation is approximately 10%. This value can be considered as a relatively large gain error margin, which is the result of reduced constraints on image quality.

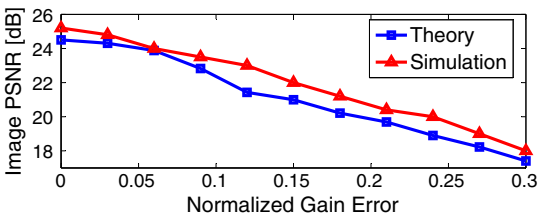

Figure 7. Photo-response non-uniformity (PRNU) influence on image peak signal-to-noise ratio (PSNR).

DOI: 10.1002/cta

However, because pixel gain error calibration (or cancelation) would be very difficult in systems using CS, PRNU must be estimated and precisely limited during the sensor design phase.

4.5. Pixel-response nonlinearity

Pixel-response nonlinearity is a fixed characteristic of the pixel, as pixel response is never perfectly linear. A nonlinear response is desirable in many applications (e.g., high dynamic range). If the pixel topology is known, the pixel-response nonlinearity can be modeled by fitting the circuit-level response curve to the specific polynomial function. Because the nonlinearity of the pixel response preserves image sparsity, even highly distorted images can be correctly reconstructed. Therefore, after the CS reconstruction, the image is mapped to the initial nonlinear image without any error directly caused by pixel nonlinearity. If compared to the ideal (completely linear) pixel response, the effect of the pixel error due to pixel response nonlinearity could be modeled as output-referred WGN by using a similar analysis as for PRNU. However, because the initial image would be nonlinear itself, calculating this error would not be meaningful from the perspective of the circuit design.

The effect of pixel response nonlinearity is illustrated in Figure 8 where two highly distorted images are reconstructed while the nonlinearity is modeled as $y = (e^{Kx} - 1)/(e^K - 1)$ with x being the normalized pixel intensity value within the range (0,1] and parameter K defining the exponential slope (for $K = \{+5, -5\}$). The effect that the pixel response nonlinearity has on the CS reconstructed images is the same as the effect it has on the original image, or the image signal sampled by Nyquist sampling.

4.6. Pixel temporal noise

Pixel temporal noise is a random time-dependent variation of the pixel response. It is caused by the combination of multiple physical phenomena observed in all integrated circuits such as thermal noise, electron-shot noise, and flicker noise. The overall pixel temporal noise power depends on the type and the geometry of the photo-detector, the overall pixel topology, and the fabrication process parameters. The temporal noise can be approximately modeled as white Gaussian noise with zero mean, assuming that the sources of noise are independent from pixel to pixel on a single image frame. For the proposed scheme, the pixel output error caused by the temporal noise is multiplied with the coefficients from the Rademacher distribution and therefore using (3)

$$Var(e_{out}) = \sigma^2_{temp} = \sigma^2_{e_X}, \tag{21}$$

where σ_{e_X} is the normalized pixel temporal noise standard deviation. Using (5) and (12), the normalized measurement error standard deviation due to pixel temporal noise can be expressed as

(a) (b)

Figure 8. Reconstruction of distorted (nonlinear) images: (a) $K = 5$ and (b) $K = -5$.

DOI: 10.1002/cta

$$\frac{\sigma_e}{y} = \frac{\sigma_{ex}}{\sqrt{3}}. \tag{22}$$

The effect of temporal noise on the image quality for the proposed scheme is shown in Figure 9, where the x-axis represents the normalized temporal noise standard deviation. As for offset FPN, because the temporal noise does not depend on the benchmark image, the simulated and the theoretical curve match very well. The PSNR loss of 2 dB is observed when the temporal noise standard deviation is approximately 6% of the maximum signal amplitude. CMOS active pixel sensors (APS) typically achieve a dynamic range of more than 60 dB and can easily achieve an SNR of 40 dB, and thus comfortably fulfill the mentioned requirement. However, as for Nyquist sampling, the pixel temporal noise is a pixel non-ideality that requires the highest attention. Performing CDS can reduce flicker noise and eliminate reset noise which is common for standard APS. However, as for PRNU, the residual of the temporal noise has to be estimated and precisely limited during the sensor design phase.

5. READOUT-LEVEL NON-IDEALITIES

In CS schemes, the readout circuit is responsible for forming the image measurements **y** (Figure 1). If the multiplication with the measurement matrix Φ is performed within the pixel circuit, the readout circuit should only form the sum of the pixel outputs. It is also possible to form the measurement matrix multiplication within the readout circuit [17]. Irrespective of the selected case, any error caused by any physical non-ideality in the readout circuit, directly causes an error in the image measurements. Consequently, these non-idealities can be easily understood, controlled, and modeled by directly referring them to the output. Hence, in this section, only the results based on direct modeling will be presented, because all the error effects are already output-referred. Nevertheless, because they cause errors in the image measurements and not in the actual image signal, they cannot be added to the pixel-level effects simply using superposition. This is especially important for effects such as readout nonlinearity and readout output saturation. In Nyquist sampling schemes, the readout nonlinearity would distort the image signal directly, which can be tolerated by the human eye and in the general case is not a major concern. Moreover, saturation of the readout circuit output only sets the upper limit of the detectable signal and therefore defines the brightest possible signal that can be detected. In CS schemes, both of these effects cause errors in the image measurements that reduce the accuracy of the actual image reconstruction. Hence, they reduce the image quality independently of the actual image content. All the readout circuit non-ideality effects and their consequences on the circuit design considerations are discussed in more detail in the following subsections.

5.1. *Offset column-fixed pattern noise*

Offset column-FPN appears in the column-parallel readout due to fabrication mismatch of the devices from the different columns. For Nyquist sampling, column-FPN introduces random variation in the form of vertical stripes, which are very uncomfortable to the human eye. In CS schemes, column-FPN causes an error in the measurement values and depending on the scheme does not necessarily result in the vertical (nor horizontal) stripes. However, it causes high image noise and has to be handled properly. It is modeled as a random variation of the readout output with a Gaussian

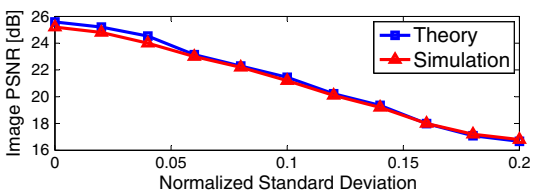

Figure 9. Pixel temporal noise influence on image peak signal-to-noise ratio (PSNR).

DOI: 10.1002/cta

distribution. The relation between the reconstructed image PSNR and the offset column-FPN for the proposed scheme is presented in Figure 10. The standard deviation is normalized to the maximum measurement amplitude ($3\sigma_y$). The PSNR drop introduced by this non-ideality is higher than any other physical effect discussed to this point. More than 2 dB drop is observed for column-FPN standard deviation at 3% of the maximum signal amplitude. As already mentioned, post-fabrication calibration would not be feasible in most CS schemes, so offset column-FPN has to be carefully addressed in the design phase. If satisfying column-to-column device matching is not achievable, offset cancelation techniques such as delta double sampling or CDS must be implemented.

5.2. *Gain error column-fixed pattern noise*

Gain error column-FPN occurs if the column-parallel readout circuits perform amplification of the signal. It is also caused by fabrication mismatch between the devices of different columns, causing discrepancies in the applied gain. Similarly to PRNU, it can be modeled as a random gain variation with a Gaussian distribution and cannot be canceled using conventional FPN cancelation techniques. If the circuit topology and the fabrication process are known, as for any other kind of FPN including PRNU, the standard deviation can be extracted from circuit-level Monte-Carlo simulations. The effect of gain error column-FPN on the reconstructed image quality for the proposed scheme is shown in Figure 11. The PSNR drop of 2 dB is observed for a normalized gain error standard deviation of approximately 5%, which is generally not difficult to achieve in CMOS. However, alike PRNU, gain error column-FPN must be carefully estimated and limited during the sensor design phase.

5.3. *Readout nonlinearity*

Readout nonlinearity has a very different effect on image acquisition using CS than the nonlinearity of the pixel response. It is also different compared to the Nyquist sampling where the nonlinearity of the readout directly introduces distortion of the image signal. The direct image signal nonlinearity is not a concern for the majority of applications. In Nyquist sampling schemes, the readout is often nonlinear. Moreover, a common design practice in image sensor analog-to-digital converter (ADC) design is to consider and characterize SNR as a relevant parameter without considering distortion. For general-purpose ADCs, signal-to-noise and distortion ratio (SNDR) is crucial parameter as it includes the ADC nonlinearity and defines the effective number of conversion bits (ADC resolution). In CS schemes, however, the nonlinearity of the readout introduces distortion of the measurements that essentially causes measurement errors. Errors in measurements reduce the accuracy of the image reconstruction. They finally appear as noise in the reconstructed image, hence reducing the image quality. Stated otherwise, CS transforms readout nonlinearity into random noise after the

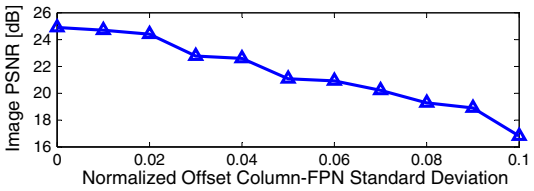

Figure 10. Offset column-fixed pattern noise (FPN) influence on image peak signal-to-noise ratio (PSNR).

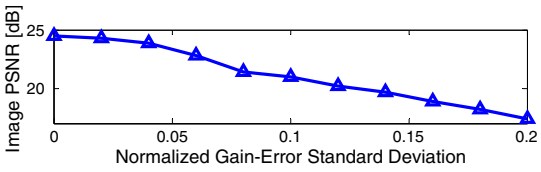

Figure 11. Gain error fixed pattern noise (FPN) influence on image peak signal-to-noise ratio (PSNR).

DOI: 10.1002/cta

image reconstruction. Figure 12 presents the effect of the nonlinearity modeled as the exponential equation: $y=(e^{Kx}-1)/(e^{K}-1)$ for the values of the parameter K between -1 and 1, which corresponds to the relative measurement error of $K/10$ (e.g. $K=1$ provides a relative error of 10%). The nonlinear function is applied to the measurements normalized (scaled) to the range of [0,1]. The drop of 2 dB in image PSNR is observed for $K\approx 0.1$, which approximately corresponds to 1% of relative error. In practice, this directly relates to the total harmonic distortion of the column-parallel readout and SNDR of the ADCs. In the case of the proposed CS scheme, if less than 2 dB of PSNR loss is required, the SNDR of the column-parallel ADC should be higher than 40 dB. Hence, there is a design constraint on the minimum SNDR of the ADC. It should be noted, however, that the required level of linearity to preserve acceptable image quality is easily achievable in CMOS design. Nevertheless, considering the nonlinearity of the readout, a fundamental difference exists compared to Nyquist sampling, which requires careful design consideration.

5.4. *Readout temporal noise*

Readout temporal noise is mainly caused by thermal and flicker noise in the readout circuit. Nevertheless the signal cross-talk, substrate coupling, power supply noise, electromagnetic interference, and several other sources also contribute to the overall temporal noise. Because all these sources can be considered as independent and uncorrelated, for estimating the image quality, the readout temporal noise can be modeled as WGN added to the image measurements.

The effect of the readout temporal noise on the reconstructed image quality for the proposed scheme is presented in Figure 13. The temporal noise standard deviation is normalized versus the maximum image measurement amplitude ($3\sigma_y$). Compared to the other physical non-idealities, readout temporal noise has the highest impact on image quality reduction. In a single frame, the effect is quantitatively very similar to the effect of the readout offset FPN. Nevertheless, the temporal noise is significantly harder to reduce by using circuit design cancelation techniques than the offset FPN. Some components such as reset noise (in APS) and flicker noise can be reduced by CDS. However, all the other components of the temporal noise remain unchanged. Therefore, readout temporal noise is the most important concern for designing image acquisition with CS schemes. This, however, does not represent a major difference from the Nyquist sampling, where the readout temporal noise is also considered as a fundamental factor limiting the achievable image quality.

5.5. *Readout output voltage range*

Readout output voltage range is fundamentally limited by the readout circuit supply voltage. Moreover, because the linearity of the readout is important in CS schemes, the output swing limit is

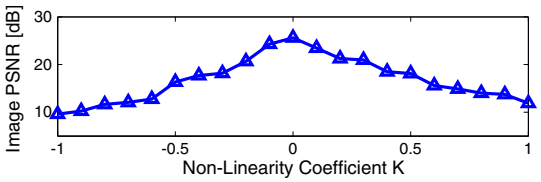

Figure 12. Readout nonlinearity versus image peak signal-to-noise ratio (PSNR).

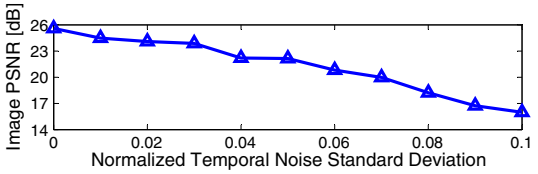

Figure 13. Readout temporal noise versus image peak signal-to-noise ratio (PSNR).

DOI: 10.1002/cta

usually reduced below the power supply limit. The image measurements, which represent the output of the readout, are typically formed as sums of the outputs originating from a large number of pixels. Conveniently, if these outputs are multiplied with a measurement matrix with zero mean (such as the Rademacher distribution), the final output amplitude has zero mean as well, scales with $\sqrt{N}$ and has a stochastic nature. The differential output range of $[-3\sigma_y, 3\sigma_y]$, where σ_y is the amplitude standard deviation (equation (12)), is sufficient for accurate image reconstruction. If the readout circuit temporal noise is relatively high, a further increase of the measurement amplitude can be considered. The results obtained from the readout circuit saturation modeling and simulation are presented in Figure 14. If the measurement amplitude is increased such that the available output range only covers the range of $[-2\sigma_y, 2\sigma_y]$, the drop in the image quality is still acceptable and below the empirical 2 dB limit. However, a very steep roll-off in image PSNR is observed after the amplitude is limited by output saturation to $2\sigma_y$. Hence, the maximum measurement amplitude can be increased to some extent, in the case of high readout noise. Nevertheless, the trade-off should be carefully weighted to avoid a high image quality reduction due to saturation in the region below $2\sigma_y$. The distribution of image measurement amplitudes for an actual benchmark image (Cameraman) is shown in Figure 15. The theoretical measurement amplitude PDF is shown as well, if the image amplitude is considered to be uniformly distributed. The majority of image measurement amplitudes are in the range of $[-3\sigma_y, 3\sigma_y]$, and thus the output saturation at this level does not cause a significant measurement error. If the range approaches $[-2\sigma_y, 2\sigma_y]$, many amplitudes will be saturated and be subjected to an error. The image quality will consequently decrease.

5.6. *Quantization*

Quantization is last in the chain of the design non-idealities. For the CS schemes, it can be modeled by restricting the measurement values to the predefined quantization levels. The number of quantization levels depends on the number of bits in the ADCs. The effect of quantization on the proposed scheme image PSNR is presented in Figure 16. For an image quality above 23 dB, approximately six conversion bits are sufficient. This result is closely related to the error caused by the readout nonlinearity. For a similar image quality, 40 dB of SNDR was needed, which in turn requires an ADC of $ENOB \approx 6.37$.

5.7. *Compression ratio*

Compression ratio is a controlled parameter that determines how many CS measurements are taken per image frame compared to the total number of pixels. Even though formally it is not a non-ideality, the compression ratio defines the loss of information in the CS scheme and therefore influences the

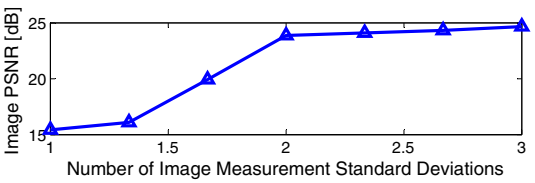

Figure 14. Saturation versus image peak signal-to-noise ratio (PSNR).

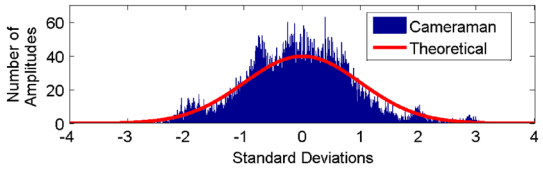

Figure 15. Histogram of measurement amplitudes.

DOI: 10.1002/cta

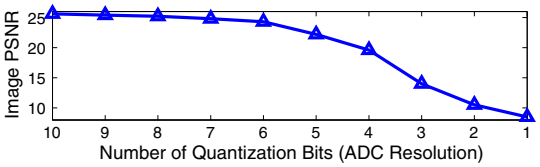

Figure 16. Quantization versus image peak signal-to-noise ratio (PSNR).

reconstructed image quality. The image PSNR dependence to the compression ratio (factor) for the proposed scheme is shown in Figure 17. Combined with the number of quantization bits, the compression ratio defines the compression level and the memory and/or bandwidth requirements for the specific compression scheme. For example, a compression ratio of 0.25 with 6-bit measurement quantization (Figure 16) results in compression level of 1.5 bit per pixel. This directly defines the memory size required to store the image frame. Depending on the required frame rate, the required image transmission bandwidth can also be defined accordingly.

6. DISCUSSION

The general architecture of the CS acquisition system can be represented by the diagram in Figure 18 based on the CS scheme proposed in Section 2.2 and the analysis presented in Sections 4 and 5. The presented CS architecture has several benefits compared to conventional CS CMOS image sensor design approaches. As explained in Section 2, it transforms the measurement matrix into the block-diagonal structure, which increases the speed of the image reconstruction. Furthermore, the measurement matrix generation hardware can be placed outside the physical area of the pixel array. This effectively reduces the pixel size compared to the methods where the measurement matrix is generated directly on the focal plane using additional in-pixel hardware. Moreover, only horizontal metal lines are required to route the random matrix coefficients toward the pixels, which further reduces the effective pixel size compared to other methods. Using the same random coefficient (bit) for the full image row (Figure 18) results in the same spatial random vector in every image column. As shown in Section 2, this technique provides a similar image quality as in the case where the full image array is random (Figures 2 and 3). Consequently, a completely new measurement matrix can be generated in only a single clock cycle of the random sequence generators (Figure 18). Hence, single-shot (or single clock) image measurement is achievable, which allows very high frame rates. The random sequences can easily be generated by linear feedback shift registers (LFSRs), in a way that a completely new measurement matrix is formed in every LFSR clock cycle. An example of the complete circuit that implements the proposed scheme to obtain high frame-rate operation is given in [22].

A general comparison of the proposed scheme with other compression schemes that can be implemented on the focal plane is shown in Table I. The proposed scheme shows comparable image quality to conventional CS schemes and schemes based on custom scan patterns (for a comparable compression level), while providing a more scalable and compact hardware implementation. The schemes based on DCT/DWT and address event representation can provide better image quality and higher compression ratios but typically suffer from higher hardware complexity and lower scalability.

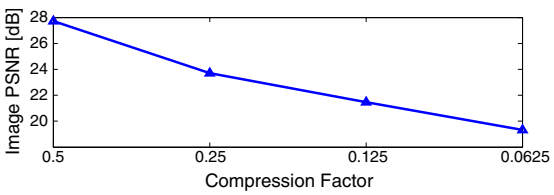

Figure 17. Image peak signal-to-noise ratio (PSNR) versus compression factor.

DOI: 10.1002/cta

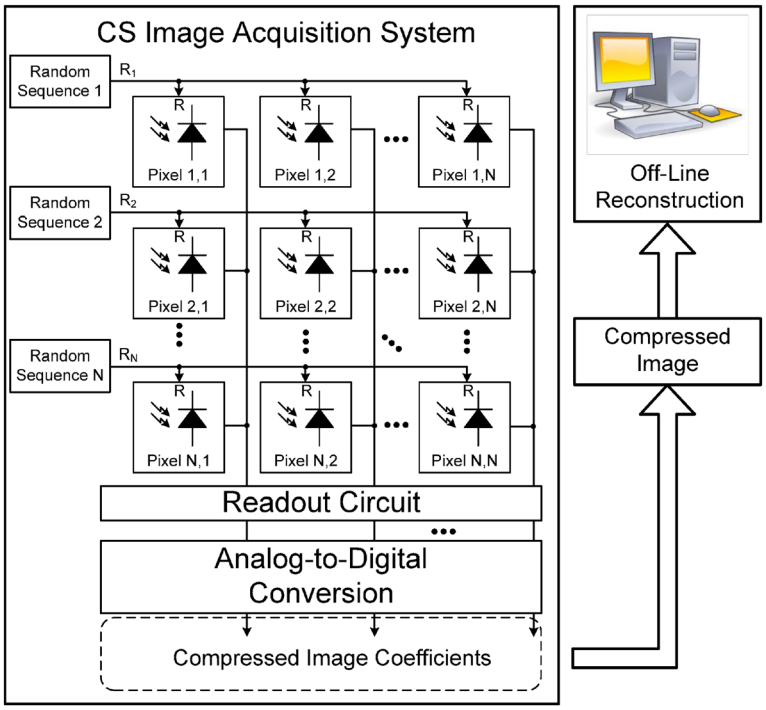

Figure 18. Top level view.

Table I. Comparison with related work.

	Typical PSNR [dB]	Compression Level [BPP]	Hardware Complexity	Hardware Scalability	Image Reconstruction
DCT/DWT [1] [2]	20–50	0.5–6	High	Poor	Real-time
Custom scans [4] [5] [6]	20–35	0.5–1	Medium	Medium	Off-line
AER [8] [9]	40–50	N/A[a]	High	Medium	Real-time
Traditional CS [17]	20–35	0.5–2	High	Poor	Off-line
Proposed scheme	20–25	0.375–1.5	Low	High	Off-line

[a]AER compression depends on active events in the scene.
BPP = bit-per-pixel; DCT = discrete cosine transform; DWT = discrete wavelet transform; PSNR = peak signal-to-noise ratio.

The physical non-ideality modeling methodology and guidelines presented in this work can easily be adapted for any CS scheme. Depending on the measurement matrix probability distribution, the results may slightly differ in the final error mean and variance values, because the random matrix mean and variance may be different. Nevertheless, the general approach does not differ. Even if the image measurements are not performed over columns, but rather on a full pixel array, the only difference would be the number of pixels that are combined to form the image measurement. However, a special care is required in that case, because there is a fundamental limitation to the actual number of pixels that can be summed together to form the image measurement. This limitation is caused by the limited output swing of the readout circuit. As shown in Section 5, the amplitude of the image measurements increases proportionally to the $\sqrt{N}$, where N is the number of combined pixels (equation (12)). Hence, increasing N increases σ_y. The same $\sqrt{N}$ proportionality is valid for the circuit noise and for all the other pixel non-ideality errors (as shown in Section 4). However, the maximum measurement amplitude is limited by the available readout circuit output swing and the fact that it cannot be lower than $2\sigma_y$ (Figure 14). If σ_y raises to a level such that the

DOI: 10.1002/cta

output swing limit of $2\sigma_y$ is reached, the effective signal amplitude originating from every pixel must be deceased to allow any further increase of N. On the other hand, pixel noise floor is a fixed pixel performance metric, thus the decrease in pixel signal amplitude results in decreasing the pixel SNR, hence decreasing the image quality. Therefore, for larger pixel arrays, combining pixel columns instead of full array in order to form image measurements represents a very reasonable design choice. This can be considered as a side-benefit of the proposed scheme. It should be noted that instead of decreasing the pixel amplitude, a large number of pixels can be combined by time-multiplexing the pixel charge transfers, using switch capacitor circuits. However, time-multiplexing would simply trade the image quality versus the achievable frame rate. Moreover, single-shot (or single cycle) measurements are not feasible if the time-multiplexing of charge transfer is performed.

7. CONCLUSION

A novel compressive sampling acquisition scheme convenient for hardware implementation in CMOS image sensor is presented in this work. The scheme uses an identical random vector for every pixel column, which reduces the required metal lines within the pixel array, thereby reducing the effective pixel size. Moreover, it enables simple and efficient measurement matrix generation outside of the pixel array physical area. The full matrix can be generated in a single clock cycle, which enables high frame rate operation.

The CS theoretical concepts and modeling fundamentals are presented, which can be used for any CS image acquisition scheme design. On the example of the proposed scheme, the most important physical non-idealities related to IC design are modeled and analyzed. All the modeling steps are presented in detail; hence, the proposed methodology can be used during the system design phase of any other compressive sensing scheme. All the physical IC design effects are characterized and discussed in the context of circuit design constraints. The final results are useful for the future design practice in CS image acquisition.

The final results show that the proposed or similar CS schemes can be used for applications where relatively low image quality is acceptable. Many of the design constraints are relaxed because of the image quality reduction, but the further performance benefit can be obtained in terms of circuit power consumption, frame rate, and memory usage. For higher video quality, additional signal quantization levels can be extracted from the temporal domain as a benefit of high oversampling. Moreover, in video applications, image quality can be improved using smart reconstruction algorithms that exploit the temporal sparsity. Potentially, these algorithms, which are under development [23] [24], offer an additional degree of compression and therefore higher performance of CS in video acquisition.

ACKNOWLEDGEMENT

This work was supported by the Swiss National Foundation (SNSF), under the project number 200020-146369.

REFERENCES

1. Bandyopadhyay A, Hasler P et al. MATIA: a programmable 80 μW/frame CMOS block matrix transform imager architecture. *IEEE Journal of Solid-State Circuits (JSSC)* 2006; **41**:263–272.
2. Nilchi A, Genov R et al. Focal-plane algorithmically-multiplying CMOS computational image sensor. *IEEE Journal of Solid-State Circuits (JSSC)* 2009; **44**:1829–1839.
3. Pajarola R, Widm P. An image compression method for spatial search. *IEEE Transactions on Image Processing* 2000; **9**:357–365.
4. Artyomov E, Yadid-Pecht O et al. Morton (Z) scan based real-time variable resolution CMOS image sensor. *IEEE Transactions on Circuits and Systems for Video Technology (TCASVT)* 2005; **15**:947–952.
5. Zhang M, Bermak A. Quadrant-based online spatial and temporal compressive acquisition for CMOS image sensor. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems* 2011; **19**:1525–1534.
6. Shoushun C, Bermak A et al. Adaptive-quantization digital image sensor for low-power image compression. *IEEE Transactions on Circuits and Systems-I: Regular Papers* 2007; **54**:13–25.

DOI: 10.1002/cta

7. Dogaru R, Kim H et al. Chaotic scan: a low complexity video transmission system for efficiently sending relevant image features. *IEEE Transactions on Circuits and Systems for Video Technology (TCASVT)* 2010; **20**:317–321.
8. Ruedi PF, Nussbaum P et al. A 128× 128 pixel 120 dB dynamic range vision sensor chip for image contrast and orientation extraction. *IEEE International Solid-State Circuits Conference (ISSCC)* 2003; **1**:226–227.
9. Posch C, Wohlgenannt R et al. A QVGA 143 dB dynamic range asynchronous address-event PWM dynamic image sensor with lossless pixel-level video compression. *IEEE International Solid-State Circuits Conference (ISSCC)* 2010; **1**:400–401.
10. Boahen KA. A burst-mode word-serial address-event linkI: transmitter design. *IEEE Transactions on Circuits and Systems-I: Regular Papers* 2004; **51**:1269–1280.
11. Candes EJ, Wakin MB. An introduction to compressive sampling. *IEEE Signal Processing Magazine* 2008; **25**:21–30.
12. Zhang M, Bermak A. Compressive acquisition CMOS image sensor: from the algorithm to hardware implementation. *IEEE Transactions on Very Large Scale Integration Systems* 2010; **18**:490–500.
13. Duarte MF, Davenport MA et al. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine* 2008; **25**:83–91.
14. Majidzadeh V, Vandergheynst P, Leblebici Y et al. A (256x256) pixel 76.7 mW CMOS imager/compressor based on real-time in-pixel compressive sensing. *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS)* 2010; **1**:2956–2959.
15. Robucci R, Romberg J, Hasler P et al. Compressive sensing on a CMOS separable-transform image sensor. *Proceedings of the IEEE* 2010; **98**:1089–1101.
16. Chi YM, Abbas A et al. An active pixel CMOS separable transform image sensor. *IEEE International Symposium on Circuits and Systems (ISCAS)* 2009; **1**:1281–1284.
17. Oike Y, Gamal AE. A 256x256 CMOS image sensor with ΣΔ based single-shot compressed sensing. *IEEE International Solid-State Circuits Conference (ISSCC)* 2012; **1**:386–388.
18. Candes EJ, Romberg J, Tao T. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics* 2006; **59**:1207–1223.
19. Chambolle A. An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision* 2004; **20**:89–97.
20. Beck A, Teboulle M. Fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* 2009; **2**:183–202.
21. Romberg J. Sensing with random convolution. *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMPSAP)* 2007; **1**:137–140.
22. Katic N, Leblebici Y, et al. Column-separated compressive sampling scheme for low-power CMOS image sensors. *IEEE International New Circuits and Systems Conference* (*NEWCAS*), June 2013; **1**.
23. Wakin MB, Baraniuk RG et al. Compressive imaging for video representation and coding. *Picture Coding Symposium*, April 2006; **1**.
24. Fowler JE, Tramel EW et al. Block-based compressed-sensing of images and video. *Foundations and Trends in Signal Processing* 2012; **4**:297–416.

DOI: 10.1002/cta