Hardware-friendly compressive imaging based on random modulations & permutations for image acquisition and classification

Wissam Benjilali, William Guicquero, Laurent Jacques, Gilles Sicard

HARDWARE-FRIENDLY COMPRESSIVE IMAGING BASED ON RANDOM MODULATIONS & PERMUTATIONS FOR IMAGE ACQUISITION AND CLASSIFICATION

Wissam Benjilali[⋆] *William Guicquero*[⋆] *Laurent Jacques*[†] *Gilles Sicard*[⋆]

[⋆] Univ. Grenoble Alpes, CEA, LETI, F-38000 Grenoble, France.
[†] ISPGroup, ICTEAM/ELEN, UCLouvain, Louvain-la-Neuve, Belgium.

ABSTRACT

This paper presents a new compressive sensing acquisition scheme well adapted for highly constrained hardware implementations. The proposed sensing model being basically designed to meet both theoretical (*i.e.*, Restricted Isometry Property) and hardware requirements (*i.e.*, power consumption, silicon footprint), is highly suitable for image sensors applications addressing both image rendering and embedded decision making tasks. In fact, for a pixels array, the proposed framework consists in applying for each row a random modulation ± 1 and a random permutation of the pixels, and then averaging the outputs by column to extract a compressed vector. This model is shown to be relevant as it has the same theoretical performance as a randomly generated sensing scheme as well as a low silicon footprint for physical implementation. Various numerical results and a discussion on possible implementations will be presented to show the robustness and the efficiency of the proposed model.

Index Terms— Compressive sensing, random modulations, random permutations, image sensor, machine learning

1. INTRODUCTION

With the rise of Internet-of-Things (IoT) [1] and data specific processing units [2] [3], the amount of data to sense and process has grown in leaps and bounds. To address the algorithm-hardware complexity involved by the data dimensionality, a compression technique is typically introduced in the signal processing pipeline [4]. To this end, a vector $\boldsymbol{x} \in \mathbb{R}^N$ (*e.g.*, an image with N pixels) is approximated by K non-zeros coefficients in a basis $\boldsymbol{\Psi} \in \mathbb{R}^{N\times N}$, *i.e.*, $\boldsymbol{x} = \boldsymbol{\Psi}\boldsymbol{\alpha}$, with $\|\boldsymbol{\alpha}\|_0 = |\operatorname{supp}(\boldsymbol{\alpha})|$ is the degree of sparsity of $\boldsymbol{x}$ in $\boldsymbol{\Psi}$.

Since the emergence of Moore's law, several works have focused on implementing near image sensor compression techniques [5]. However, these implementations exhibit high computational and memory costs mainly related to the transform coding (*e.g.*, DCT, DWT). On the other hand, Compressive Sensing (CS) [6] has emerged as a powerful framework for signal acquisition and sensor design based on random measurements. Indeed, CS theory demonstrated that a sparse signal can be recovered from a small set of those measurements. Mathematically, this corresponds to consider the sensinng matrix $\boldsymbol{\Phi} \in \mathbb{R}^{M\times N}$ to perform a signal independent dimensionality reduction mapping the signal $\boldsymbol{x} \in \mathbb{R}^N$ to $\boldsymbol{y} = \boldsymbol{\Phi}\boldsymbol{x} \in \mathbb{R}^M$ ($M \ll N$). CS allows one to alleviate hardware design constraints (*e.g.*, Analog-to-Digital conversion, power consumption) *e.g.*, using pseudo-random generators to generate on-the-fly the sensing matrix as a deterministic and reproducible process (*e.g.*, LFSR [7–9], Cellular Automaton [10, 11]). However, recovering the original signal from its CS measurements involves a costly reconstruction process. Taking advantage of the sparsity of the signal $\boldsymbol{x}$ in a basis $\boldsymbol{\Psi}$ (*e.g.*, DWT) to recover the sparsest signal $\hat{\boldsymbol{x}}$ such that $\boldsymbol{y}$ is very close to $\boldsymbol{\Phi}\hat{\boldsymbol{x}}$, we can then solve the following Basis Pursuit Denoising (BPDN) problem:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\|\boldsymbol{\Psi}^\top \boldsymbol{x}\right\|_1 \text{s.t.} \left\|\boldsymbol{y} - \boldsymbol{\Phi}\boldsymbol{x}\right\|_2^2 \leq \epsilon. \qquad (1)$$

Related works: The CMOS image sensor community has proposed several CS implementations to deal with either hardware or algorithm constraints for image rendering tasks. First, [7, 12] exploit random convolutions to extract CS measurements giving the priority to a fast and efficient image reconstruction at the expense of on-chip complexity. [8] reduces the measurements support and applies an identical pseudo-random measurement matrix for every column while performing charge summation to reduce power consumption. Moreover, [13] describes a scalable and low-complexity column based CS scheme based on a Cellular Automaton (CA) that shows a chaotic behavior to generate the sensing matrix during the acquisition process. Alternatively, the most promising way to perform CS during A/D conversion via incremental $\Sigma\Delta$ converters is proposed in [9, 14]. Summation/averaging operations are done during A/D conversion while well optimized 4T pixels and a canonical readout scheme are used optimizing that way the overall performance of the sensor in terms of noise, technological dispersion and power. These works yet all suffer from numerous drawbacks, limiting the use of CS to niche applications. This can mainly be summarized to the hardware complexity in terms of memory needs in [7, 12], and restricted measurements supports in [8, 9, 13].

LJ is funded by the Belgian F.R.S.-FNRS and by the project AlterSense (MIS-FNRS).

Contributions: In this paper we propose a hardware-friendly CS sensing scheme to provide more independent measurements by extending the measurements support to address highly constrained hardware design (*e.g.*, ultra-low power image sensor). The proposed framework is mathematically defined as a combination of a random modulation matrice and random permutation matrices. Thus, for a given pixels array, we first apply a random modulation to the sensed pixel values, and then perform a random column permutation which is different for each selected row before averaging the column output in a rolling shutter readout fashion [9]. The purpose of the modulation is to center CS matrix expectation and thus center measurements distribution, while the permutations increase the information content (diversity) of each measurement with uncorrelated measurements supports. Advantageously, the proposed sensing scheme enables the reuse of a standard rolling shutter acquisition scheme as well as an array of optimized pixels. In the rest of the paper, we first describe the proposed sensing scheme and then carry out some analytical and numerical studies to evaluate its robustness for both image rendering and decision making applications. Finally, we provide a discussion on the hardware implementation in the context of a highly constrained hardware for embedded near-sensor processing [15, 16].

2. PROPOSED SENSING FRAMEWORK

2.1. Mathematical model

Given $\boldsymbol{u} = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_{n_r})^\top \in \mathbb{R}^{n_r n_c}$ a per-row vectorized image sensed by a canonical $n_r \times n_c$ CIS with a rolling-shutter readout of its pixels array (n_r and n_c are the numbers of rows and columns respectively). The proposed sensing model $\boldsymbol{\Phi}$ corresponds to applying the following steps repeated over s snapshots. First, we apply a random modulation, *i.e.*, multiplication of each image by a ± 1 weights generated by a Bernoulli distribution. Then, an horizontal concatenation of n_r permutation matrices is built to accumulate randomly selected pixels from each row. For each snapshot, a compressed vector in $\mathbb{R}^{n_c}$ is extracted. Thus, for s snapshots, the sensing scheme is expressed with a normalization factor $\frac{1}{\sqrt{s}}$ as:

$$\boldsymbol{\Phi u} = \frac{1}{\sqrt{s}}((\sum_{j=1}^{n_r} \boldsymbol{P}_1^{(j)} \boldsymbol{\varphi}_1^{(j)} \boldsymbol{u}_j)^\top, \cdots, (\sum_{j=1}^{n_r} \boldsymbol{P}_s^{(j)} \boldsymbol{\varphi}_s^{(j)} \boldsymbol{u}_j)^\top)^\top, \tag{2}$$

where, for each $1 \leq i \leq s$ and $1 \leq j \leq n_r$, $\boldsymbol{\varphi}_i^{(j)}$ is a diagonal matrix whose main diagonal entries are the ± 1 weights applied to the j^{th} row at snapshot i, and $\boldsymbol{P_i} = (\boldsymbol{P}_i^{(1)}, \ldots, \boldsymbol{P}_i^{(n_r)}) \in \{0,1\}^{n_c \times n_c n_r}$ is the horizontal concatenation of n_r permutation matrices $\boldsymbol{P}_i^{(j)} \in \{0,1\}^{n_c \times n_c n_r}$ for the snapshot i. We stress that the modulation weights and permutation matrices are different for each snapshot. Theoritically, the design of CS sensing matrices has to satisfy Restricted Isometry Property (RIP) [17] to guarantee a stable embedding property. In particular, it was shown that a wide variety of randomly generated matrices (*e.g.*, Bernoulli and sub-Gaussian distributions) satisfies this property [18]. In the rest of this section, we first present the key concept to evaluate the robustness of a CS matrix, *i.e.*, Restricted Isometry Property (RIP), and then provide a quantitative and qualitative analysis of this property for the proposed CS scheme and compare its performances to a well known randomly generated sensing scheme [18].

2.2. Definitions

A key concept to evaluate the robustness of a CS sensing scheme is the RIP property. A sensing matrix $\boldsymbol{\Phi}$ is said to respect the RIP of order K if for all K-sparse vectors $\boldsymbol{x}$ there exists $\delta \in (0,1)$ such that:

$$(1-\delta)\|\boldsymbol{x}\|_2^2 \leq \|\boldsymbol{\Phi x}\|_2^2 \leq (1+\delta)\|\boldsymbol{x}\|_2^2. \tag{3}$$

When respecting the RIP, the mapping $\boldsymbol{\Phi}$ preserves the energy of the sensed signal and thus is said to be a stable embedding. Consequently, respecting the RIP over all $2K$-sparse vectors implies to preserve the pairwise distance between any two K-sparse vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, *i.e.*,

$$(1-\delta)\|\boldsymbol{u}-\boldsymbol{v}\|_2^2 \leq \|\boldsymbol{\Phi u}-\boldsymbol{\Phi v}\|_2^2 \leq (1+\delta)\|\boldsymbol{u}-\boldsymbol{v}\|_2^2. \tag{4}$$

2.3. RIP analysis

To show that the CS model presented in (2) respects the RIP property, let us present the following lemma.

Lemma 1. *Let $\boldsymbol{\Phi}$ in (2). For any K-sparse vector $\boldsymbol{u} \in \mathbb{R}^{n_r n_c}$,*

$$\mathbb{E}(\|\boldsymbol{\Phi u}\|_2^2) = \|\boldsymbol{u}\|_2^2.$$

Proof. The ℓ_2-norm of $\boldsymbol{\Phi u}$ can be expressed as:

$$\|\boldsymbol{\Phi u}\|_2^2 = \frac{1}{s}\sum_{i=1}^{s} \|\sum_{j=1}^{n_r} \boldsymbol{P}_i^{(j)} \boldsymbol{\varphi}_i^{(j)} \boldsymbol{u}_j\|_2^2,$$

equivalently,

$$\|\boldsymbol{\Phi u}\|_2^2 = \frac{1}{s}\sum_{i=1}^{s}\sum_{j=1}^{n_r}\sum_{k=1}^{n_r} (\boldsymbol{\varphi}_i^{(j)}\boldsymbol{u}_j)^\top (\boldsymbol{P}_i^{(j)})^\top \boldsymbol{P}_i^{(k)} (\boldsymbol{\varphi}_i^{(k)}\boldsymbol{u}_k).$$

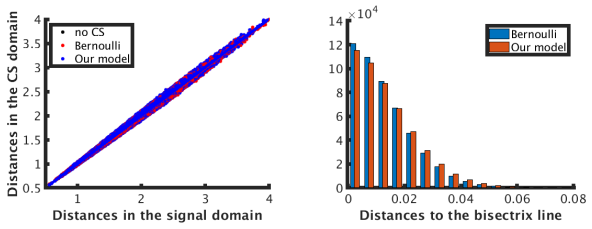
Because $\boldsymbol{P}_i^{(j)}$ are orthonormal matrices, $(\boldsymbol{P}_i^{(j)})^\top \boldsymbol{P}_i^{(j)} = \boldsymbol{I}_{n_c}$, and, the main diagonal entries of $\boldsymbol{\varphi}_i^{(j)}$ are Bernoulli variables with $\mathbb{E}(\operatorname{diag} \boldsymbol{\varphi}_i^{(j)}) = \boldsymbol{0}$, with $\boldsymbol{\varphi}_i^{(j)}$ and $\boldsymbol{\varphi}_k^{(l)}$ independent if $i \neq k$ or $j \neq l$, we can write:

$$\mathbb{E}(\|\boldsymbol{\Phi u}\|_2^2) = \frac{1}{s}\sum_{i=1}^{s}\sum_{j=1}^{n_r}\sum_{j'=1}^{n_c} \mathbb{E}((\boldsymbol{\varphi}_i^{(j)})_{j'}(\boldsymbol{u}_j)_{j'})^2.$$

Thus,

$$\mathbb{E}(\|\boldsymbol{\Phi u}\|_2^2) = \frac{1}{s}\sum_{i=1}^{s}\sum_{j=1}^{n_r}\sum_{j'=1}^{n_c} (u_j)_{j'}^2 = \|\boldsymbol{u}\|_2^2.$$

□



(a) (b)

Fig. 1: (a) Concentration of pairwise distances of our model and a Bernoulli distribution around the pairwise distances in the signal domain (bisectrix line) for $N = 1024$, $K = 10$ and $M = 128$. (b) Distances to the bisectrix line of our model and a Bernoulli distribution.

Note that this lemma can advantageously be generalized to $\mathbb{E}(\left\|\boldsymbol{\Phi\Psi u}\right\|_2^2) = \left\|\boldsymbol{\Psi u}\right\|_2^2 = \left\|\boldsymbol{u}\right\|_2^2$ by Parseval's identity, if $\boldsymbol{\Psi}$ is an orthonormal basis.

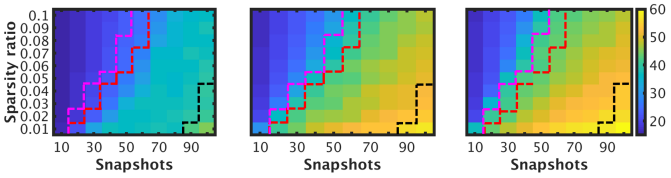
To estimate the RIP constant δ of our sensing matrix, an analytical estimation is carried out. A dataset of 1000 10-sparse signals are generated in the canonical basis. Each vector of length 1024 (*i.e.*, $n_c = n_v = 32$) has $K = 10$ non-zero coefficients normally generated on the support. This dataset is projected in the CS domain using either our model in (2) or a full Bernoulli random matrix to extract $M = 128$ measurements. The main idea behind this study is to show that the embedding performed by our model preserves the pairwise distance between any two K-sparse vectors of the generated dataset as well as the Bernoulli random matrix. Thus, to evaluate the distortion of the projected distances, Fig. (1(a)) exhibits a point cloud mapping the pairwise distances between the signal domain (X axis) and the CS domain (Y axis), for our model and a Bernoulli random matrix over 100 trials. As the point cloud of our model is well concentrated around the bisectrix line (blue line) and its regression line perfectly fits the Bernoulli's one, we can validate, qualitatively, the fact that our model preserve the pairwise distances with respect to a distortion constant, expressed as the RIP constant δ.

The RIP constant δ can be seen as the aperture of the cone defined by the point cloud related to our model. In Fig. (1(b)) we establish the histogram of distances to the bisectrix line for every point of Fig. (1(a)) to get an estimation of the RIP constant of our model and compare it to the Bernoulli random matrix. Thus at 3σ, the estimated $\delta = 0.033$ for our model and $\delta = 0.031$ for the Bernoulli matrix. This means that we can reasonably attest that the proposed model respects the RIP with a small distortion constant, and guarantee to recover signals from their compressed measurements.

3. NUMERICAL EXPERIMENTS

3.1. Reconstruction of sparse signals

In this section we evaluate the recovery performance of the proposed model for sparse signals. Here, we consider the *Barbara*, *Monkey*, *Boat*, *Cameraman* and *Lena* images resized to 128×128 with different level of sparsity (thresholded in the wavelet domain and back projected). The sparsifying basis $\boldsymbol{\Psi}$ is the Daubechies-6 wavelet basis and the UnLocBox [1] is used to solve the augmented Lagrangian [2] of (1). We then plot the average PSNR in dB for different sparsity level (*i.e.*, $\frac{K}{128^2}$) in function of the number of snapshots (*i.e.*, for each snapshot we extract 128 measurements). Fig. (2) reports the required amount of snapshots to perform (*i.e.*, minimum number of measurements to extract) to successfully recover the signal under the constraint of a certain PSNR using a per-column Bernoulli (denoted "Col-Bern") sensing scheme as used in [8], our model without permutations (denoted "W/o perm") and our model (denoted "modPerm"). One could say that a signal is properly reconstructed if the reconstruction error is lower than 10^{-4}, *i.e.*, a PSNR higher than 40 dB.



(a) Col-Bern. (b) W/o perm. (c) modPerm.

Fig. 2: Phase-transition diagrams. Black, red and magenta lines show the transitions to a success reconstruction above 40 dB for per-column Bernoulli (Col-Bern), our model without permutations (W/o perm) and our model sensing schemes (modPerm) respectivelly.

3.2. Reconstruction of compressible signals

In this section, the only considered image is the *Cameraman* image of size 512×512 (because of its intrisic image characteristics). Two regularization operators are used to recover this image instead of a simple ℓ_1-constraint in a wavelet basis: first the Total Variation (TV) and secondly the sum of the ℓ_1-norm of the gradient of the image in multiple wavelet domains (Db-2, Db-6 and Db-10 denoted mDWT-TV) as performed in [13]. The reconstruction from CS measurements is performed thanks to a FISTA algorithm [19] using our model and a per-column Bernoulli sensing [8] over 10 batches. As clearly seen in Fig. (3) and Fig. (4), our sensing scheme outperforms the alternative scheme.

3.3. Classification of compressible signals

Leveraging the cost of signal recovery, compressive signal processing [20] allows to perform signal processing (*e.g.*, filtering, detection and inference) in the CS domain thanks to the RIP property. In this section, we address an object recognition problem to evaluate the performance of our model on two object recognition databases: the MNIST handwritten database (10 classes, 28×28 pixels) [21] and the AT&T

[1] //epfl-lts2.github.io/unlocbox-html/

[2] with a very small regularization parameter λ

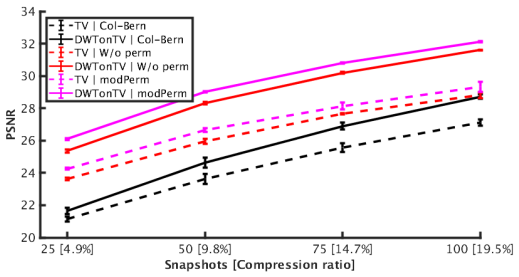


Fig. 3: Quality of reconstruction of our sensing model compared to a per-colon Bernoulli acquisition scheme over 10 batches.

(a) TV | Col-Bern. (b) mDWT-TV | Col-Bern.

(c) TV | modPerm. (d) mDWT-TV | modPerm.

Fig. 4: Quality of reconstruction of our sensing compared to a per-column Bernoulli acquisition scheme (75 snapshots, *i.e.*, 14.7% compression ratio).

face recognition database (40 classes, 92×112 pixels) [22]. The performance of a one-vs.-all Support Vector Machine (SVM) classifier [23] learned on CS measurements sensed by our model is reported in Fig. (5). It shows the accuracy in terms of the ratio of correct predictions to the total number of test samples and its standard deviation (10 randomly selected batches). By comparing the plots one can draw out the following conclusions regarding the data variability. First, one can see that in the case of the MNIST (small and low variable samples), we can achieve a no-loss classification compared to a no-CS setting from 10 snapshots, (*i.e.*, $\approx 35\%$ compression ratio). However, with more variability and informational content (*i.e.*, AT&T), we achieve the same performance from 10 snapshots, (*i.e.*, $\approx 10\%$ compression ratio).

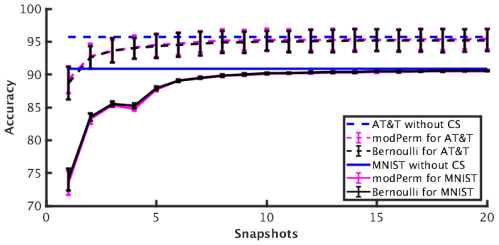


Fig. 5: Classification accuracy for the AT&T and MNIST databases.

4. DEDICATED HARDWARE IMPLEMENTATIONS AND CONCLUSION

A physical system implementation of the proposed sensing model is illustrated in Fig. (6) [15, 16]. Two possible embodiments are presented to perform on-chip pseudo-random modulations and permutations. To perform permutations, [15] proposes a pseudo-random generator to generate a pseudo-randomly permuted sequence to address the columns of the selected row. However, in [16] this task is achieved thanks to a butterfly network [24] controlled by a pseudo-random generator (PRG). In addition, the pseudo-random modulations and the sum are performed in both architectures by a dedicated A/D converter enabling to perform pseudo-random modulations, averaging and A/D conversion as in [9]. For multiple snapshots, our scheme requires non-destructive pixel readout both in the case of rolling shutter and global shutter acquisitions thanks to relaxed constraints on ADC speed. However, for specific inference tasks we can perform single scan.

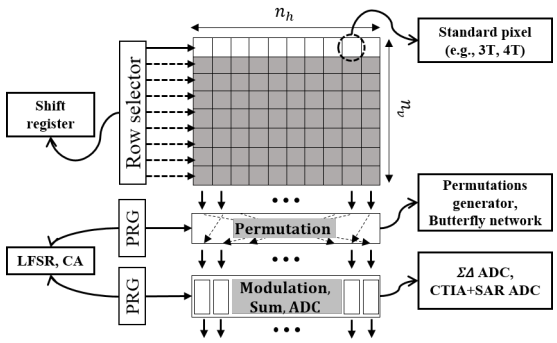


Fig. 6: Top-level architecture of a pseudo-random modulations & permutations compressive image sensor.

In conclusion, this paper proposes a new CS sensing scheme based on random modulations & permutations to meet highly constrained hardware tasks while respecting theoretical properties of a CS matrix. To this end, we demonstrate that our sensing model respects the RIP property with a distortion constant estimated, analytically. We also compared it to a randomly generated matrix and present some numerical results to evaluate its robustness regarding various applications. A key difference of the proposed model is its relevance in terms of silicon footprint and on-chip CMOS limitations.

5. REFERENCES

[1] T. Ernst *et al.*. Sensors and related devices for iot, medicine and s mart-living. In *2018 IEEE Symposium on VLSI Technology*, June 2018.

[2] K. Bong, S. Choi, C. Kim, D. Han, and H. J. Yoo. A low-power convolutional neural network face recognition processor and a CIS integrated with always-on face detector. *IEEE Journal of Solid-State Circuits*, 2018.

[3] L. Millet, S. Chevobbe, C. Andriamisaina, L. Benaissa, E. Deschaseaux, E. Beigne, K. B. Chehida, M. Lepecq, M. Darouich, F. Guellec, T. Dombek, and M. Duranton. A 5500-frames/s 85-gops/w 3-d stacked bsi vision chip based on parallel in-focal-plane acquisition and processing. *IEEE Journal of Solid-State Circuits*, 2019.

[4] S. Mallat. *A wavelet tour of signal processing: the sparse way*. Academic press, 2008.

[5] M. Zhang and A. Bermak. Cmos image sensor with on-chip image compression: A review and performance analysis. *Journal of Sensors*, 2010, 2010.

[6] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, March 2008.

[7] L. Jacques, P. Vandergheynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici. Cmos compressed imaging by random convolution. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2009.

[8] N. Katic, M. Hosseini Kamal, M. Kilic, A. Schmid, P. Vandergheynst, and Y. Leblebici. Column-separated compressive sampling scheme for low power cmos image sensors. In *2013 IEEE 11th International New Circuits and Systems Conference (NEWCAS)*, June 2013.

[9] Y. Oike and A. El Gamal. Cmos image sensor with per-column adc and programmable compressed sensing. *IEEE Journal of Solid-State Circuits*, 48(1), Jan 2013.

[10] S. Wolfram. *Cellular automata and complexity: collected papers*. CRC Press, 2018.

[11] M. Trevisi, A. Akbari, M. Trocan, . Rodrmguez-Vzquez, and R. Carmona-Galn. Compressive imaging using rip-compliant cmos imager architecture and landweber reconstruction. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019.

[12] V. Majidzadeh, L. Jacques, A. Schmid, P. Vandergheynst, and Y. Leblebici. A (256256) pixel 76.7mw cmos imager/ compressor based on real-time in-pixel compressive sensing. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, May 2010.

[13] W. Guicquero, A. Dupret, and P. Vandergheynst. An algorithm architecture co-design for cmos compressive high dynamic range imaging. *IEEE Transactions on Computational Imaging*, 2(3), Sept 2016.

[14] W. Guicquero, A. Verdant, and A. Dupret. High-order incremental sigmadelta for compressive sensing and its application to image sensors. *Electronics Letters*, 51(19), 2015.

[15] W. Benjilali, W. Guicquero, L. Jacques, and G. Sicard. A low-memory compressive image sensor architecture for embedded object recognition. In *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*, Aug 2018.

[16] W. Benjilali, W. Guicquero, L. Jacques, and G. Sicard. An analog-to-information vga image sensor architecture for support vector machine on compressive measurements. In *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2019. [Accepted for publication].

[17] E. J. Candès and T. Tao. Decoding by linear programming. *IEEE Transactions on Information Theory*, 2005.

[18] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constructive Approximation*, Dec 2008.

[19] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Img. Sci.*, March 2009.

[20] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk. Signal processing with compressive measurements. *IEEE Journal of Selected Topics in Signal Processing*, April 2010.

[21] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, Nov 1998.

[22] AT&T face database. Available [online]: `http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.zip`.

[23] C. M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. Springer, 2006.

[24] Y. Hilewitz and R. B. Lee. A new basis for shifters in general-purpose processors for existing and advanced bit manipulations. *IEEE Transactions on Computers*, 58(8), Aug 2009.