

EXPLORING INFORMATION RETRIEVAL USING IMAGE SPARSE REPRESENTATIONS: FROM CIRCUIT DESIGNS AND ACQUISITION PROCESSES TO SPECIFIC RECONSTRUCTION ALGORITHMS

THÈSE N° 6587 (2015)

PRÉSENTÉE LE 10 SEPTEMBRE 2015

À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE DE TRAITEMENT DES SIGNAUX 2
PROGRAMME DOCTORAL EN GÉNIE ÉLECTRIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

William GUICQUÉRO

acceptée sur proposition du jury:

Prof. J.-Ph. Thiran, président du jury
Prof. P. Vandergheynst, Dr A. Dupret, directeurs de thèse
Prof. L. Jacques, rapporteur
Prof. Y. Leblebici, rapporteur
Prof. A. B. Rodríguez Vázquez, rapporteur



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

The computer can't tell you the emotional story.
It can give you the exact mathematical design,
but what's missing is the eyebrows.
— Frank Zappa

To my family and friends. . .

Acknowledgements

First, I would like to express my gratitude to both my supervisor Pierre Vanderghenst and my co-supervisor Antoine Dupret. They have always provided me sound advice, despite their respective job responsibilities. In particular, Pierre for helping me to define the right directions to take and Antoine for theoretical and technical approaches to follow. More generally, their perspicacity and knowledge about the research world considerably helped me. Thank you both for your unfailing support.

A sincere gratitude goes to my former colleagues and co-PhD students at the EPFL with whom I would have liked to keep more contacts. Unfortunately, this may be due to my own particular status related to co-supervision between the CEA and the EPFL. However, I want to warmly non-exhaustively salute Fab, Johann, Kirell, Mahdad, Nath' and Vassilis. I sincerely thank David, my former "co-desktooper", for his enthusiasm for new technologies and his thoughts about mixing art&science.

I would like to wink my CEA colleagues and "roommates", for their patience and their understanding. The advice of Laurent the first and Numerobis has been of a precious help, especially regarding microelectronic design issues without neglecting the Haar wavelets (that will lead the world one day, who knows ?). Arnaud V.&P. ideas and fervour have been a source of inspiration and creativity particularly when referring to $\Sigma\Delta$ s, of order 17. I am glad to have had the opportunity to work with a really great nematic guy (twisted nematic ? I never know...). Thank you Timothe. Coffee breaks and discussions with Clem', Paco and their champion have allowed me to think about something else than my current work, at least once a day and even if it is about something millimetre sized. This remark also works for Guillaume for the "spare times" spent together. I am greatly thankful to Michael (nowadays called Fabrice ?) who accepted me in his lab (two times...) and for his trust in me. It has been a pleasure for me to work with the DACLE team thanks to the great atmosphere and working conditions.

I also would like to thank my lovely girlfriend who have been of unconditional support all along this long period. This support have been without failure. Thank you for having followed me even if it has implied two relocations, meaning job changes.

I will never forget my friends who have always been kind and willing to help me, allowing me to have so many prolific and stimulating discussions about science and technology.

Last but not least, my parents and my grandparents have been of a great help, giving me the motivation and the perseverance to lead this work to its end, what probably explains why I dedicate this thesis to them.

I also thank all the people I did not mention in this too short acknowledgement page.

Grenoble, Mai 2014

W. G.

Abstract

New advances in the field of image sensors (especially in CMOS technology) tend to question the conventional methods used to acquire the image. Compressive Sensing (CS) plays a major role in this, especially to unclog the Analog to Digital Converters which are generally representing the bottleneck of this type of sensors. In addition, CS eliminates traditional compression processing stages that are performed by embedded digital signal processors dedicated to this purpose. The interest is twofold because it allows both to consistently reduce the amount of data to be converted but also to suppress digital processing performed out of the sensor chip. For the moment, regarding the use of CS in image sensors, the main route of exploration as well as the intended applications aims at reducing power consumption related to these components (i.e. ADC & DSP represent 99 % of the total power consumption).

More broadly, the paradigm of CS allows to question or at least to extend the Nyquist-Shannon sampling theory. This thesis shows developments in the field of image sensors demonstrating that is possible to consider alternative applications linked to CS. Indeed, advances are presented in the fields of hyperspectral imaging, super-resolution, high dynamic range, high speed and non-uniform sampling.

In particular, three research axes have been deepened, aiming to design proper architectures and acquisition processes with their associated reconstruction techniques taking advantage of image sparse representations. How the on-chip implementation of Compressed Sensing can relax sensor constraints, improving the acquisition characteristics (speed, dynamic range, power consumption) ? How CS can be combined with simple analysis to provide useful image features for high level applications (adding semantic information) and improve the reconstructed image quality at a certain compression ratio ? Finally, how CS can improve physical limitations (i.e. spectral sensitivity and pixel pitch) of imaging systems without a major impact neither on the sensing strategy nor on the optical elements involved ?

A CMOS image sensor has been developed and manufactured during this Ph.D. to validate concepts such as the High Dynamic Range - CS. A new design approach was employed resulting in innovative solutions for pixels addressing and conversion to perform specific acquisition in a compressed mode. On the other hand, the principle of adaptive CS combined with the non-uniform sampling has been developed. Possible implementations

Acknowledgements

of this type of acquisition are proposed. Finally, preliminary works are exhibited on the use of Liquid Crystal Devices to allow hyperspectral imaging combined with spatial super-resolution.

The conclusion of this study can be summarized as follows: CS must now be considered as a toolbox for defining more easily compromises between the different characteristics of the sensors: integration time, converters speed, dynamic range, resolution and digital processing resources. However, if CS relaxes some material constraints at the sensor level, it is possible that the collected data are difficult to interpret and process at the decoder side, involving massive computational resources compared to so-called conventional techniques. The application field is wide, implying that for a targeted application, an accurate characterization of the constraints concerning both the sensor (encoder), but also the decoder need to be defined.

Key words: Signal Processing, Microelectronics, Compressive Sensing, CMOS image sensor, High Dynamic Range, Regularization algorithm, Non-uniform Sampling, Hyperspectral Image, Super-resolution.

Résumé

Les nouvelles avancées dans le domaine des capteurs d'images (particulièrement en technologie CMOS) tendent à remettre en cause les méthodes classiques employées pour acquérir l'image. C'est en cela que l'acquisition compressive joue un rôle majeur notamment afin de désengorger les convertisseurs Analogique-Numérique, représentant généralement le goulot d'étranglement de ce type de capteurs. De plus, l'acquisition compressive permet de s'affranchir des traitements de compression effectués traditionnellement par un traitement numérique du signal et réalisés par des composants dédiés. L'intérêt est double car cela permet à la fois de réduire de manière conséquente la quantité de données à convertir mais aussi de se passer des traitements numériques externes au capteur. Dans le domaine des capteurs d'image, la principale voie d'exploration ainsi que les applications visées sont basées sur l'emploi de l'acquisition compressive pour diminuer la consommation de ces composants (i.e. 99% de la consommation totale).

Contrairement, ou plutôt, plus largement ; le paradigme de l'acquisition compressive permet de remettre en question ou tout du moins d'étendre la théorie relative aux limites de Nyquist-Shannon. Cette thèse propose des développements dans le champ des capteurs d'images démontrant qu'il est possible d'envisager des applications alternatives liées au paradigme de l'acquisition compressive. En effet, y sont présentées des avancées dans les champs de l'imagerie hyperspectrale, de la super-résolution, de la grande dynamique, de la grande vitesse et de l'échantillonnage non uniforme.

En particulier, trois axes de recherche ont été approfondis, visant à concevoir des architectures ou processus d'acquisition avec leurs propres techniques de reconstruction associés exploitant les représentations parcimonieuses de l'image. Comment la mise en œuvre, en pratique, de l'acquisition compressive peut-elle relâcher les contraintes des capteurs et améliorer leurs caractéristiques inhérentes (vitesse, plage de dynamique, consommation d'énergie) ? Comment l'acquisition compressive peut-elle être combinée avec de simples analyses d'image fournissant des descripteurs permettant la réalisation d'applications de plus haut niveau (en ajoutant de l'information sémantique) ainsi que l'amélioration de la qualité de l'image reconstruite ? Enfin, comment l'acquisition compressive peut-elle réduire les contraintes physiques (sensibilité spectrale et pas pixel) des systèmes d'imagerie classiques tout en limitant l'impact sur la stratégie d'acquisition ainsi que sur les éléments optiques impliqués ?

Acknowledgements

Un capteur d'image en technologie CMOS a été développé et fabriqué lors de cette thèse afin de valider certains concepts tels que l'High Dynamic Range - Compressive Sensing. Pour ce faire, une nouvelle approche architecturale a été employée débouchant sur des solutions innovantes concernant l'adressage ainsi que la conversion des pixels lors d'une acquisition en mode compressif. D'autre part, le principe de l'acquisition compressive adaptative couplé à l'échantillonnage non-uniforme a été étudié et approfondi. Des implémentations permettant de mettre en oeuvre à moindre coût ce type d'acquisition est proposé dans ce manuscrit. Enfin, des travaux préliminaires sont exposés concernant l'utilisation d'un certain type de cristaux liquides afin de permettre la réalisation d'une acquisition hyperspectrale et super-résolue spatialement à partir d'un capteur d'image CMOS dit standard.

La principale conclusion de cette étude peut être résumée de la façon suivante : l'acquisition compressive doit désormais être considérée comme une boîte à outils permettant de définir plus aisément des compromis entre les différentes caractéristiques des capteurs : temps d'intégration, vitesse des convertisseurs, dynamique, résolution et ressources de traitement numérique disponibles. Cependant, si l'acquisition compressive permet de relâcher certaines contraintes matérielles au niveau du capteur, il se peut que les données collectées soit difficilement interprétables par la suite et requérant des charges en termes de calculs et de mémoire plus importantes par rapport aux techniques dites classiques. Lié au fait que le champ applicatif soit si vaste, il est alors primordial de définir de façon précise les contraintes pour une application donnée concernant à la fois le capteur (codeur) mais aussi le décodeur.

Mots clefs : Traitement du signal, Micro-électronique, Acquisition compressive, Capteur d'image CMOS, Extension de dynamique, Algorithme de régularisation, Échantillonnage non-uniforme, Image hyperspectrale, Super-Résolution.

Contents

Acknowledgements	i
List of Figures	ix
List of Tables	xvii
List of Symbols	xix
1 Introduction	1
Organization of the thesis	3
Image sensor world, the domination of CMOS	5
Compressive Sensing, when whitening is beautiful	10
Randomization and sparsity: “New fields of applications and actual limitations”	13
2 State Of The Art	15
2.1 Compressive Sensing background	15
2.1.1 A brief note on notations	16
2.1.2 The CS paradigm	18
2.1.3 Reconstruction algorithms	21
2.2 Hardware CS implementations and field of applications	22
2.2.1 Acquisition systems and CS-CIS	23
2.2.2 Reconstruction, signal recovery and information interpretation . .	38
2.3 Limitations and current development	41
3 Smart Automaton Compressive Sensing Imager	45
3.1 Preliminary study	46
3.1.1 Type of measurements, supports, and expectancy reduction	46
3.1.2 Possible implementations and addressed applications	50
3.2 CS image sensor architecture and design (SACSI)	51
3.2.1 Pseudo-Random Generator (PRG)	53
3.2.2 Specific but common pixel design	60
3.2.3 End-of-column circuitry	64
3.3 The information considered as a group of measurements	70
3.3.1 Still image reconstruction	70

Contents

3.3.2	Video reconstruction, the time into slices	80
3.3.3	High Dynamic Range (HDR) reconstruction	86
3.3.4	Artifacts and non-linearity compensation	90
4	Adaptive Compressive Sensing imager	95
4.1	Adaptive acquisition and Compressive Sensing, the trade-off	96
4.1.1	Sparsity estimator, block variance	96
4.1.2	Block sparsity variability	99
4.1.3	Adaptative CS, the use of block variances	99
4.2	Top-level implementation of an adaptive CS imager	104
4.2.1	Variance calculation	105
4.2.2	Adaptation and variable density sampling	106
4.3	Specific adaptive CS applications	108
4.3.1	Compressive sensing versus pixel random selection	109
4.3.2	Tracking using variance and particle filtering	114
4.3.3	Face recognition using adaptive CS measurements	116
5	Beyond spectral limitations	121
5.1	Hyperspectral CS	122
5.1.1	Hyperspectral CS imaging systems	122
5.1.2	Study on 3-Dimensionnal constraints	124
5.2	Multiple blurring kernels for joint SR and hyperspectral	126
5.2.1	Sensing Scheme based on multiple blurring kernels	126
5.2.2	Joint Super-Resolution and Hyperspectral reconstruction	129
5.3	Liquid Crystal Devices for SRHS imaging	131
6	Main outlines and perspectives	135
	Bibliography	155
	Curriculum Vitae	157

List of Figures

1.1	Applications involving digital image sensors [source: Yole développement].	2
1.2	Compound Annual Growth Rate (CAGR) for main image sensor applications [source: Yole développement].	3
1.3	Top-level functional view of a canonical CMOS Image Sensor.	5
1.4	Photodiode schematic model and voltage variation across the photodiode during photo integration (autointegration reverse biased working mode). .	6
1.5	Passive CMOS pixel (1T pixel).	7
1.6	Active CMOS pixel using an in-pixel voltage follower (3T pixel).	7
1.7	Active CMOS pixel using an in-pixel voltage follower combined with a pinned-photodiode (4T pixel).	8
1.8	The classical acquisition chain (a) revisited by the Compressive Sensing (b). 11	
1.9	The lossy compression by structured transform (a) revisited by the Compressive Sensing using non structured projections (b).	12
1.10	The main limitations and constraints related to CMOS Image Sensors (CIS). Areas of study and fields of improvements regarding CS-CIS (and CS imaging systems more broadly).	14
2.1	Single pixel camera using DMDs [61]. The scene is focused on the DMDs plane then reflected (Spatial Light Modulation) and finally focused on a single photodiode	24
2.2	Reconstruction results which are presented in [61]. A comparison between ideal and exact sparse images (top row) and the SPC acquisitions for two different compression ratios.	25
2.3	Single pixel camera using LCDs [70]. The scene is focused on the LCDs plane then transmitted (Spatial Light Modulation) and finally focused on a single photodiode.	26
2.4	Reconstruction results presented in [70] for different compression ratios, from left to right, 0.25, 0.5, 0.75 and 1). The results are notably better than the ones presented for the single pixel camera using DMDs [61], probably due to the improvements in terms of optical calibration	27

List of Figures

2.5	Coded strobing scheme presented in [73]. This technique employs an aperiodic time modulation of the pixel integration. In time domain those successive modulations for each frame are respectively depicted in light blue, green, pink and orange.	28
2.6	Mathematical model of the Coded strobing system [73].	28
2.7	Camera setup used for experiments presented in [76]. The use of a combination of random mirrors with a flat mirror allows spatial modulation of the image before being acquired by the common camera's sensor. . . .	29
2.8	CS Imaging system based on the optical scattering of some media [77]. . .	29
2.9	Reconstruction results presented in [77]. Each point represents the probability of successful recovery for a k sparse image of N pixels using M measurements (around 50 trials per point).	30
2.10	Lensless structure using controllable LCD [79]. a) presents a light rays cone for a single aperture element. b) shows a example of aperture assembly configuration corresponding to a specific command.	30
2.11	Mathematical description of separable transform operation [82].	32
2.12	Top view and pixel schematic of the separable transform image sensor [82].	33
2.13	Top-level functional view of the random convolution based CS imager [88].	34
2.14	Pixel schematic (a) and local processing (b) used in the proposed design of [88].	34
2.15	An array of pixels in which only randomly selected pixels are read and converted [92].	35
2.16	Top-level functional view of the Sigma Delta based CS imager [93]. . . .	36
2.17	CS-MUX proposed in [93] to perform CS measurements on each supports (blocks of pixels).	36
2.18	The proposed low power column-based CS image sensor. a) Top-level functional view of row binary measurement based CS image sensor [98]. b) Schematic view of the pixel and the readout circuitry of [98].	38
2.19	Summary of the main CS imaging techniques at the encoder and the decoder parts.	42
3.1	Equivalence between matrix to vector multiplication and matrix to matrix multiplication in the case of column-based CS.	47
3.2	Example of masks used in the case of Bernoulli block-based CS or Bernoulli column-based CS.	48
3.3	Impact of using a sparser sensing matrix in the case of column-based CS. The expectancy of non zero values in measurement vectors linearly depends on the p value.	49
3.4	Impact of using a sparser sensing matrix in the case of block-based CS. The expectancy of non zero values in measurement vectors linearly depends on the p value.	49
3.5	Impact of using a specific sensing scheme for three different p value. . . .	50

3.6	Top-level view of the proposed architecture SACSI (Smart Automaton Compressive Sensing Imager).	52
3.7	Generic top view of a one-dimension cellular automaton whose the transition function only takes left and right neighbors outputs as inputs.	54
3.8	Transition function used in the cellular automaton performing the Wolfram's Rule 30.	55
3.9	Generic top view of a one-dimension cellular automaton followed by an expectancy reduction layer.	56
3.10	Expectancy reduction function used in the proposed PRG.	56
3.11	Simulation of the proposed PRG structure (cf. figure 3.9) for a size of the cellular automaton of 64. Form left to right, the correlation matrix between generated column CS measurements, number of activated rows for each time index and the number of activations for each row.	57
3.12	Activation matrix representing the compressive sensing matrix in its reduced form (i.e. column-based CS). Points represent the indexes of activated rows in the function of time and their relative gray level represents the number of activated pixel a its time index.	58
3.13	Transient simulation of the CMD_i^t generation in column-based CS mode.	59
3.14	Transient simulation of the CMD_i^t generation in uncompressed mode.	59
3.15	Schematic of the digital adder used in CS mode for counting the number of activated rows.	60
3.16	The two different working modes of the 3T pixel.	61
3.17	Pixel current response. The pixel output current in function of the photodiode voltage during the current reading operation.	62
3.18	Proposed pixel layout. Fill factor: 32%	62
3.19	Distribution of the pixel output current, after an arbitrary time of integration resulting of a Monte-Carlo simulation of the pixel.	64
3.20	Schematic top view of the reconfigurable end-of-column circuitry.	65
3.21	Single-stage mirrored cascode OTA followed by a source follower output stage.	65
3.22	Bode plot of the proposed OTA with its output stage for a typical load and input voltage.	66
3.23	Resistance network composed of 4 switchable resistances.	66
3.24	Layout of the resistance network.	67
3.25	Histogram of Monte-Carlo simulations evaluating the effect of technological dispersions at the RTIA level.	67
3.26	Average output from Monte-Carlo simulations for a normalized input current in the case of 1 activated pixel and 12 activated pixels.	68
3.27	Transient extracted simulation of a CS acquisition during the pixel reading operation, for a single activated pixel and several input photocurrents. From top to bottom: output of the RTIA V_{out} (V), RST command, RS command (SEL), Photodiode voltage (V).	69

3.28	Output voltage of the I to V converter in function of the input photo current in the case of a unique activated pixel. The different colored curves represents different integration time: Yellow ($t_{int} = 4ms$), Purple ($t_{int} = 3.5ms$), Light blue ($t_{int} = 2.5ms$), Red($t_{int} = 1.5ms$), Blue($t_{int} = 0.5ms$).	69
3.29	Average Peak Signal to Noise Ratios computed between original greyscale images and the reconstructed ones using multiple different prior operators.	72
3.30	Average Structural Similarity Image Measurements computed between original greyscale images and the reconstructed ones using multiple different prior operators.	73
3.31	Reconstruction of the Darkhair test image using CS measurement scheme of the SACSI for a compression ratio of 11%. From left to right and top to bottom, the prior operators are TV , TV_{4axes} , $mDWT$, $TVmDWT$, $mDWTonTV$ and $mDWTonTV_{4axes}$	74
3.32	Average PSNR (a) and SSIM (b) of the reconstructions with noise and dispersion on the CS measurements.	75
3.33	Normalized transfer function of the pixel between the final readout (output of the RTIA) and the incoming light on the photodiode.	76
3.34	Average PSNR (a) and SSIM (b) of the reconstructions with noise and dispersion on the CS measurements taking into account non-linearity of the pixel.	78
3.35	Bayer Pattern of Color filter Array (CFA).	78
3.36	Reconstruction of the colored Lena test image with an equivalent greyscale compression ratio of 12.5%. The PSNR of the reconstructed image is of 26.9dB.	80
3.37	Average SSIM over successive frames (88) for the different prior operators ($TVmDWT_{2D}$ (red), $TV_{SO}mDWT_{2D}$ (green), $TV_{SO}TV_{2D}$ (blue), $mDWT_{SO}TV_{2D}$ (black)). Four 256×256 center cropped video test sequences have been used: Mother and Daughter (a), Coastguard (b), Hall Monitor (c) and Foreman (d).	83
3.38	Hall monitor CS reconstructions for a compression ratio of 6.25% using different prior operators. From left to right, from top to bottom, $TVmDWT_{2D}$, $TV_{SO}mDWT_{2D}$, $TV_{SO}TV_{2D}$, $mDWT_{SO}TV_{2D}$	83
3.39	Coarse to Fine CS reconstruction scheme	84
3.40	Comparison between a temporal averaging of the clock video sequence and a CS reconstruction performed for the same sequence.	85
3.41	Comparison between several uncompressed cropped frames and their related CS reconstructions ($CR = 1/256$, integration time=1/fps).	86
3.42	Example of time diagram which can be used for a HDR-CS acquisition.	87
3.43	Original HDR 32-bit image.	88
3.44	Histogram of the original HDR image.	88
3.45	The positive 8 images corresponding to each time of integration.	89

3.46	Tone mapped HDR reconstructions. a) is a traditional multicapture reconstruction and b) is the proposed HDR-CS reconstruction.	89
3.47	Histogram of the tone mapped reconstruction.	90
3.48	Reconstruction without compensation of the model parameters.	92
3.49	Reconstruction with compensation of the model parameters.	92
3.50	Image restoration simulation results in terms of PSNR in function of the relative error on the acquisition model parameters. Standard deviations of the reconstructed PSNR for multiple trials is represented by the vertical bars.	93
4.1	Analysis of block standard deviation and block operator $l1$ norms (compressibility estimator) from a set of test images and for four different block sizes, a) 4×4 , b) 8×8 , c) 16×16 , d) 32×32	97
4.2	Analysis of block variance and block coefficients entropy (compressibility estimator) from a set of test images and for four different block sizes, a) 4×4 , b) 8×8 , c) 16×16 , d) 32×32	98
4.3	Correlation between image block variance and its sparsity in function of the block size (side of a square block).	98
4.4	Histogram of distribution for the different block compressibility estimators and block sizes.	99
4.5	Histogram of block variances for the cameraman image (blue). Ratio of the number of measurements to perform (green)	100
4.6	Example of extracted features for the cameraman test image.	101
4.7	Reconstructed Cameraman image using $mDWTonTV$. From left to right, top to bottom: Regular CS (CR_1), Regular CS ($CR_1 + CR_2$), Adaptive CS (CR_1) and Adaptive Var-reg ($CR_1 + CR_2$).	103
4.8	Top view of the proposed CIS architecture. Three functioning modes (a) b) c)) are enabled. All of those mode use reconfigurable processing blocks (first order $\Sigma\Delta s$, column decoders).	105
4.9	Error in number of codes for a resolution of 10bits between the conversion of an analog squared input its the digital squaring after a linear $\Sigma\Delta$ conversion.	106
4.10	A column decoder performing adaptive pixel random selection associated to a single column of blocks.	107
4.11	Expectancy of pixel selection in function of the output of the column decoder with pixel random selection. $1/2$, $1/4$, $1/8$ and $1/16$ respectively represent the outputs out2, out4, out8 and out16 presented in figure 4.10. The different graphs are for different block sizes.	108
4.12	Acquisition steps. From top to bottom, from left to right: the original lena image, the block variances, the 2bits quantized block variances, Resulting sensed pixels I_s (black areas corresponds to non acquired pixels).	110

List of Figures

4.13	Reconstructed images for the different sensing-reconstruction couples ($CR = 7\%$). From left to right, top to bottom: original image, PS using $Inpaint$, PS with TV , PS with $mDWTonTV$, CS with TV and CS with $mDWTonTV$	111
4.14	Reconstructions of the 128th frame of the hall monitor video sequence (compression ratio = 0.14). From left to right: Original frame, Interpolated frame from a 96×96 averaged image, FBF reconstruction with filtering and GOF with 3D filtering.	113
4.15	Example of frame composed of camera extracted data only.	114
4.16	Example of person tracking using only simulated measurements performed by the proposed sensor architecture using 4×4 pixel block variance feature extraction and adaptive pixel selection ($CR = 22\%$).	115
4.17	Example of person tracking using only simulated measurements performed by the proposed sensor architecture using 4×4 pixel block variance feature extraction and adaptive pixel selection ($CR = 22\%$). The results presented in this figure also takes advantage of frame-to-frame differences as additional feature to reweight the histograms.	115
4.18	General overview of the program structure.	117
4.19	Classification results obtained for the Development set and the Test set.	119
5.1	CS reconstructions of Urban a) and Mexican gulf b) hyperspectral test images for a compression ratio of 1.6% and using different constraint operators. From left to right, from top to bottom, $TVmDWT_{2D}$, $TV_{SO}mDWT_{2D}$, $TV_{SO}TV_{2D}$, $mDWT_{SO}TV_{2D}$. Notice that these images has been color-interpolated to render a colorized image from the hyperspectral cube.	124
5.2	Average PSNR over all spectral bands for the different constraint operators ($TVmDWT_{2D}$ (red), $TV_{SO}mDWT_{2D}$ (green), $TV_{SO}TV_{2D}$ (blue), $mDWT_{SO}TV_{2D}$ (black)). a) and b) corresponds to scenes 6 and 8 of dataset [198]. c) and d) respectively correspond to Urban and Gulf of Mexico datasets.	125
5.3	Proposed multiple blurred acquisitions system.	127
5.4	Multiple spectral dependent blurring kernels acquisition model (block operators notations).	128
5.5	Multiple spectral dependent blurring kernels acquisition model (algebraic notations).	128
5.6	a) Example of blurring kernels used for the simulation presented in figure 5.7 and b) the resulting 16 blurred images binned by a factor of 4×4 ($s = 4$) where each image corresponds to a blurring kernel of a).	129

5.7	For left to right: the original Lena test image, 4×4 ($s = 4$) super-resolution based on a single low resolved image using the proposed regularization, SR reconstruction without compression (the 16 low resolution images have been used at the reconstruction, cf. figure 5.6), SR reconstruction with a compression of factor 2 (only 4 low resolution images are used).	130
5.8	a) is the original hyperspectral image from the CAVE database, b) is the reconstructed image with a Super-Resolution of 2×2 ($s = 2$) and a compression factor of 2. e) shows PSNR results for different settings. . . .	131
5.9	Colorized blurring kernels.	132
5.10	For left to right: the original Lighthouse test image, demosaiced image using bilinear filtering, demosaiced image using a $mDWTonTV_{4axes}^{RGB}$ regularization, SRHS reconstruction without compression (as much data as a mosaiced image).	133
6.1	Packaged SACSI chip.	136
6.2	Multi-scale block variance image of Lena.	137

List of Tables

1.1	Comparison in terms of features and performances between CCD and CMOS image sensors.	8
1.2	Comparison of typical ADCs features for CIS.	9
2.1	Comparison between the main imaging systems using optically performed CS.	31
2.2	Comparison between the main CS-CIS developed in the literature.	39
3.1	In-pixel MOS transistor sizes (in μm) for the proposed 3T pixel.	63
3.2	Phase and Gain margins of the proposed OTA for extreme configurations.	67
3.3	Comparison of resulting PSNRs for different reconstruction algorithms and sensing schemes.	74
4.1	PSNR & SSIM reconstruction results for several test images.	104
4.2	Reconstruction results for several images, CR varies from 7% to 12%.	112
4.3	Reconstruction results for several test video sequences (256 frames).	113
4.4	Results on the AT&T face recognition dataset for different input of the classifier (450 true examples and 50 false examples in the test set).	119
5.1	Average RGB-PSNR of the reconstructed color images.	133

List of Symbols

- ADC: Analog to Digital Conversion / Converter
- APS: Active Pixel Sensor
- BCS-SPL: Block-Based Compressed Sensing - Smoothed Projected-Landweber iterative reconstruction
- BM: Background Model
- BN: Bayesian Network
- CASSI: Coded Aperture Snapshot Spectral Imager
- CBSC: Comparator-Based Switched Capacitor
- CCD: Charge-Coupled Device
- CCTV: Closed-Circuit Television
- CDS: Correlated Double Sampling
- CFA: Color Filter Array
- CIS: Cmos Image Sensor
- *CMD*: Binary output of the Pseudo Random Generator
- CMOS: Complementary Metal Oxide Semiconductor
- CPPCA: Compressive-Projection Principal Component Analysis
- CR: Compression Ratio
- CS: Compressive Sensing, Compressed Sensing
- CSC: Current Squaring Circuits
- CTIA: Capacitive TransImpedance Amplifier
- DCT: Discrete Cosine Transform

List of Symbols

- DD-CASSI: Double Disperser CASSI
- DMD: Digital Micromirror Device
- DSP: Digital Signal Processor
- DWT: Discrete Wavelet Transform
- EER: Equal Error Rate
- ERL: Expectancy Reduction Layer
- FAR: False Acceptance Ratio
- FBF: Frame By Frame
- FISTA: Fast ISTA
- FPN: Fixed Pattern Noise
- FRR: False Rejected Ratio
- FT : Fidelity Term
- GND: Ground (i.e. reference voltage point)
- GOF: Group Of Frames
- HDR: High Dynamic Range
- HTER: Half Total Error Rate
- IOT: Internet Of Things
- IR: InfraRed
- ISTA: Iterative Shrinkage Thresholding Algorithm
- JPEG: Joint Photographic Experts Group (commonly used method of lossy compression for digital images)
- kTC noise: Thermal noise on capacitors
- LCD: Liquid Crystal Device
- LFSR: Linear Feedback Shift Register
- $M1/M2/M3/M_{top}$: Metal layers
- $mDWT$: Multiple Discrete Wavelet Transform $l1$ (reconstruction prior)
- $mDWT_{onTV}$: Multiple Discrete Wavelet Transform on horizontal and vertical gradients $l1$ (reconstruction prior)

- $mDWTonTV_{4axes}$: Multiple Discrete Wavelet Transform on horizontal, vertical and diagonal gradients $l1$ (reconstruction prior)
- Mpix: Mega pixels (more than 10^6 pixels)
- MRI: Magnetic Resonance Imaging
- MSE: Mean Squared Error
- NP-Hard: Non-deterministic Polynomial-time hard
- Op : Prior Operator
- OTA: Operational Transconductance Amplifier
- PaFiMoCS: Particle Filtered Modified Compressive Sensing
- PCA: Principal Component Analysis
- PD: PhotoDiode
- PN: Interface between p-type and n-type semiconductor materials
- PPS: Passive Pixels Sensor
- PRG: Pseudo-Random Generator
- PRNU: Photo-Response NonUniformity
- PS: Pixel Selection
- PSF: Point Spread Function
- QTD: Quadrant Tree Decomposition
- RGB: Red-Green-Blue color space
- RIP: Restricted Isometry Property
- RS : Row Select pixel signal
- RST : Row Reset pixel signal
- RTIA: Resistive TransImpedance amplifier
- RTS: Random Telegraph Noise
- SACSI: Smart Automaton Compressive Sensing Imager
- SAR: Successive Approximation Register, by extension type of ADC using SAR
- SD-CASSI: Single Disperser CASSI

List of Symbols

- SLM: Spatial Light Modulator
- SNR: Signal to Noise Ratio
- SPC: Single Pixel Camera
- SPIHT: Set Partitioning In Hierarchical Trees
- SR: Super Resolution
- SRHS: Super Resolution HyperSpectral
- SSIM: Structural SIMilarity
- THz: Terahertz spectral band
- TV : Total Variation $l1$ (reconstruction prior)
- TV_{4axes} : Total Variation following horizontal, vertical and digonals gradients $l1$ (reconstruction prior)
- TV_{aniso} : Anisotropic Total Variation $l1$ (reconstruction prior)
- $TV_{SOMDWT_{2D}}$: Second Order TV (in 3^{rd} axis) combined with mDWT (over 2D separte signals) $l1$ (reconstruction prior)
- $TV_{SO}TV_{2D}$: Second Order TV (in 3^{rd} axis) combined with anisotropic TV (over 2D separte signals) $l1$ (reconstruction prior)
- $TVmDWT$: anisotropic TV combined with DWT (reconstruction prior)
- TX: Transfert Gate or Pass Gate command in 4T pixel
- UV: UltraViolet spectral band
- V_{bias} : Polarization Voltage of the pixel
- VGA: Video Graphics Array
- VPIC: Visual Pattern Image Coding
- VRST: Reset Voltage
- X-ray: X spectral band
- YCbCr: Luminance/chrominances color space
- $\Sigma\Delta$: Type of ADC based on summation and thresholding operations

1 Introduction

For the last two decades the importance of solid state image sensors has dramatically increased with impacts on many different fields. The number of tomorrow's applications involving imagers will be incredibly high. As example, nowadays digital image sensors can be found in surgical units, in artificial satellites but also in everyone's pocket because of the widespreading of cellphones and smartphones. For the moment, most of the applications only uses a single image sensor. With the rise of the Internet Of Things (IOT), sensor networks will be of increasing importance. The global market share of image sensors is estimated to reach more than \$10 billion in few years. Mobile phone cameras currently represent between 60% to 80% of it. Yet, emerging applications such as machine vision, wearable sensors and smart surveillance systems lead the market to grow even faster in the future. Those systems, involving new imaging devices, tend to affect ever more the market in the near future, according to their own high growth rate. Such systems will probably require a new kind of smart integrated sensors to properly address their own dedicated functions. Apart from the visible spectrum domain, image sensors dedicated to different specific wavelengths will also take place in a large variety of systems.

As predicted by E. Fossum in [1] and in [2] later on, Complementary Metal Oxide Semiconductor (CMOS)-based imaging systems on a single chip did reflect the evolution of microelectronics. When he wrote in 1998: "We can expect the emergence of a plethora of imaging applications and the insertion of image capture systems in many aspects of our personal lives"; at this time, he was totally right. Active Pixels Sensors (APS) are now incorporated in most of mass consumer electronic products (such as laptops, televisions and cellphones). A few years later (in 2002), [3] showed for instance that "multi-mega" pixel sensors will be soon enabled by new modifications of the technological processes. More recently, [4] predicts that novel architectures tend to overcome the common limitations of standard design approaches and to promote new functionalities. This meets old concepts and works presented in [5] (1997) that refer to the design of smart and/or bio-inspired image sensors with specific features. In parallel, [6] expresses



Figure 1.1 – Applications involving digital image sensors [source: Yole développement].

that the trend of advanced technology generation pixel manufacturing will continue, probably even below the sub-diffraction limit of the pixel.

However, several problems arise because of the increasing need for both high resolution and high frame rate, leading to a considerable amount of data (e.g. the Nokia smartphone Lumia 1020 with its 41Mpix image sensor). Fast and accurate image and video compression algorithms have been developed to handle with those data in order to represent them in a compact form. In particular, the structural nature of the signals so as their time and space correlation are generally used to perform the compression. In addition to the increase of resolution and frame rate, the physical limitations due to the sensor technology tend to be reduced. For instance, the dynamic range, the noise floor, the noise level and the spectrum range constitute the most important points which are required to be improved nowadays. High Dynamic Range (HDR) acquisition modes are also emerging in mass-market products (smartphones, automotive, etc.). On the other hand, to reduce production costs of Infrared / UV / THz / X-ray imaging, technology gaps will likely be bridged modifying the way the current sensors work. In parallel, the evolution of image and video processing has provided a large set of top level applications. For instance, it refers to detection, classification, tracking and interpretation techniques which tend to be close to the artificial intelligence field of study. A current and future trend is to build smarter sensor systems in order to produce either images of higher quality, decision-making tools or autonomous systems. Regarding those considerations, the range of applications goes from video surveillance to industrial imaging passing through personal home/biomedical monitoring (see figure 1.2).

In the mid-2000s, a new theory called Compressive Sensing (CS) emerged thanks to the work of Terence Tao, Emanuel Candès and David Donoho. This exhibits a new paradigm that extends the classical Nyquist-Shanon sampling theorem, providing a way to acquire signals based on alternative fashions. Since the first papers reporting this recent signal processing technique [7] [8], the field of CS has been intensively studied.

Both theoretical limits and potential applications continue to be explored. This technique takes advantage of sparsity assumptions of the signal for yielding an effective sensing using only few “measurements”, replacing traditional regular “samples”. Compared to a traditional sampling, the main interest of CS is that the signal can be directly acquired in a compressed form relaxing constraints on Analog to Digital Converters (ADCs) and digital processing. At the opposite, regarding the classical lossy compression techniques, intensive digital processing is generally required to transform the digital signal into a traditional compressed format. Indeed, CS is now considered as a good candidate for designing new generations of CMOS Image Sensor (CIS) architectures and for cutting-edge applications.

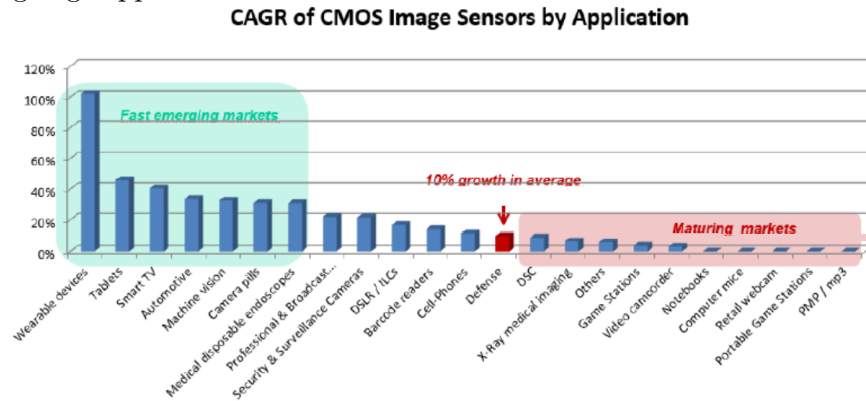


Figure 1.2 – Compound Annual Growth Rate (CAGR) for main image sensor applications [source: Yole développement].

Organization of the thesis

The rest of this introduction chapter is organized as follows: first, a brief introduction to CMOS image sensor is given with a presentation of the basic CIS architectures, then the principle of Compressive Sensing is introduced with its potential applications. In particular, it will deal with the necessity on signal structure assumptions and sparsity requirements. This chapter mainly focuses on the motivations and the reasons why it is particularly relevant to investigate the interest of CS for image sensors.

Chapter 2 summarizes the state of the art of first the most important results regarding the CS theory and then the main architectures of image sensors that are already implementing it. Imaging systems presented in this study are of two different natures. Some of them consist in imaging systems composed of specific optical elements combined with a common CIS whereas the others consist of solid state CS-dedicated CIS. We can thus differentiate two different ways to perform CS measurements in the case of image acquisition, either by applying optical modulation or by performing analog on-chip processing before the ADC. This analysis of the state of the art aims at providing key information for understanding CS-related limitations in terms of both acquisition and reconstruction. The three following chapters present three different ways of considering CS as a good opportunity for helping to design new classes of sensors relaxing current sensor constraints.

Chapter 3 reviews an implementation of a CS imager called "Smart Automaton Compressive Sensing Imager". Preliminary studies are exposed and detailed to explain and motivate taken decisions in terms of implementation. First, a study deals with measurement supports and type of measurements that are appropriate for a real-world CS implementation. Then, we discuss about possible applications related to the specific measurement scheme that has been adopted. Secondly, the proposed architecture is presented from top level (functional view) to layout level (extracted simulation) for the main functional blocks. Each specific and critical blocks are detailed. The proposed image sensor exhibits well adapted properties to work as a high frame rate CS imager and a HDR CS imager. At last, simulation and experimental results are provided to demonstrate the relevance of using the CS for relaxing image sensor constraints.

In Chapter 4, the emerging concept of adaptive CS for images is presented and reviewed. It generally consists in coupling a simple and easy pre-evaluation of the image content in order to accordingly tune the compression ratio. This technique allows to perform a smart acquisition providing output image features as well as compressively acquired images. Some non canonical architectures of imagers have recently been developed to address specific purposes. These architectures generally perform analog embedded processing (filtering, time difference, motion detection...). In the case of CS imagers, analog processing is typically required to perform CS measurements. Combining these two aspects of smart acquisition leads to the design of an adaptive CS CIS. This way, a compromise is taken in order to improve the quality of CS reconstructed images in relation to the compression ratio at the expense of supplementary but light embedded processing. Moreover, it allows to design image sensors with different possible working modes: features only, non adaptive CS, adaptive CS, adaptive CS plus features and uncompressed mode. This kind of architecture is particularly well suited for being integrated into a system which needs to adapt its working mode to the content of the observed scene.

Chapter 5 deals with multi-spectral CS acquisition and how it can be performed without multiple bulky optical devices. In this chapter we first presents the state of the art of main CS imaging systems dedicated to multi-spectral. Regarding limitations of the classical hyperspectral imaging systems, a preliminary study on using specific Liquid Crystals Devices (LCDs) for performing CS multi-spectral imaging is presented.

Finally, Chapter 6 concludes this thesis by summarizing the main outlines. A discussion about potentialities of CS when implemented in a CIS is opened. In particular, what tools CS can bring into the CIS designer world. The conclusion is that CS should not be considered as a miracle solution to fix limitations met in the CIS world but as a great opportunity to find alternatives in terms of sensing schemes, letting the designer decide the best compromises for specific applications.

Image sensor world, the domination of CMOS

Invented in the 60', the almost first solid-state camera design, uses a Charge-Coupled Device (CCD) sensor. This class of visible image sensor was widely used in digital cameras until the rise of CMOS image sensors in the last 90'. The CMOS technology have partly replaced CCDs except in some niche high-end applications. Indeed, in practice, CCD imagers are less power efficient and have a less modular structure. Nowadays, CIS approximatively represents 90% of the yearly number of shipped image sensor. This thesis will mainly deals with this technology. Indeed, this is the main technology that allows a certain flexibility in the design for integrating processing directly in or next to the focal plane.

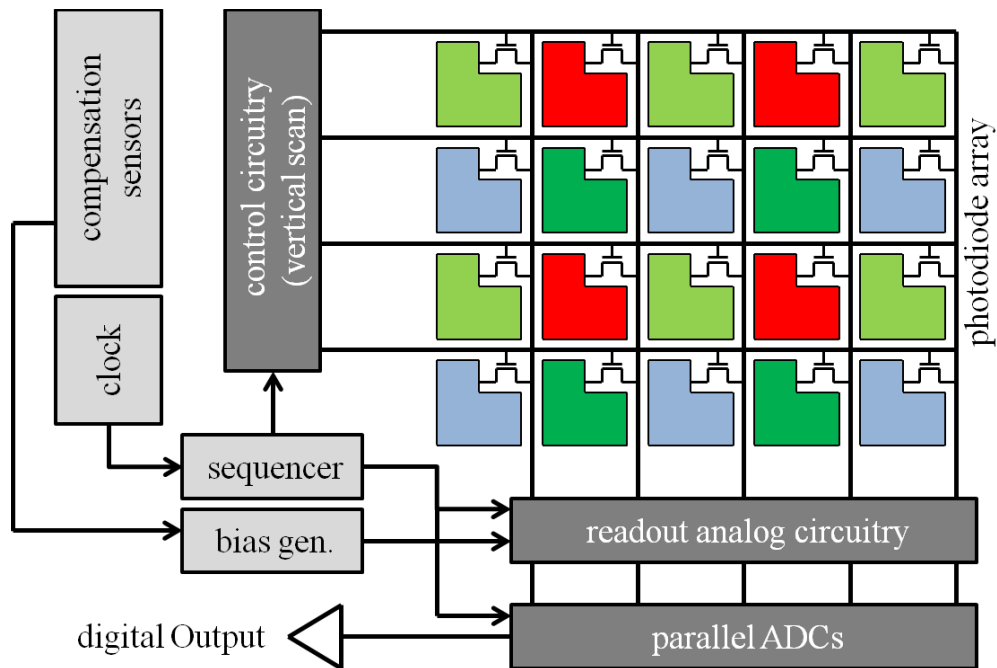


Figure 1.3 – Top-level functional view of a canonical CMOS Image Sensor.

CMOS image sensor architectures generally consist of arrays whose units are photo-sensitive sensors (pixels). As shown in figure 1.3, a modern imager is thus usually composed of a matrix of photodiodes, control and readout circuits, Analog to Digital Converters (ADCs) and digital outputs. It may embed digital processing components for specific applications such as compression. Bias controls, sequencers and possible compensation sensors (such as a temperature sensor) can be added to handle noise and technology dispersion issues. Logic control circuits and multiplexers are also involved to generate and route the different control signals. Traditionally, CMOS readout circuitry is column parallelized to speed the reading operations up, thus increasing the pixel rows time consistency. This most common architecture uses a parallelized Rolling Shutter readout, meaning that each row of pixels are read and converted sequentially.

The common model of a photodiode (the main component of a typical pixel) is shown in figure 1.4. A photodiode is a semi-conductor component formed of a PN junction enabling the conversion of incoming photons energy into electrical charges. This way, when a photon of a sufficient energy is absorbed in the reversed biased PN junction of the photodiode, an electron-hole pair is generated. The recombination of the electron and the hole is prevented by the inner electric field, thus leading to a charge displacement. Moreover, the reverse biased PN junction forms a capacitor of capacitance C . As a result, the photonic flow reaching the photodiode generates a current called the photo current i_p . Because of the relatively small value of this current (typically $\sim pA$ per μm^2 under a $100lux$ illumination), an integration of the charges Q_i is required during a certain time of integration t_{int} to collect a charge large enough to be read with a good signal to noise ratio ($Q_i = i_p \times t_{int}$). The current i_p discharges the intrinsic capacitance C of the photodiode PD and, at the end of the integration time, the total charge Q_i induces a Voltage drop $V_i = C \times Q_i$, which is the physical value that is ultimately read (see figure 1.4). At the end, it produces a signal, which is a voltage, proportional to the integration of the average level of light over an integration time slot. For more details, the articles [9] and [10] depict the evolution of pixel structures involved in CIS.

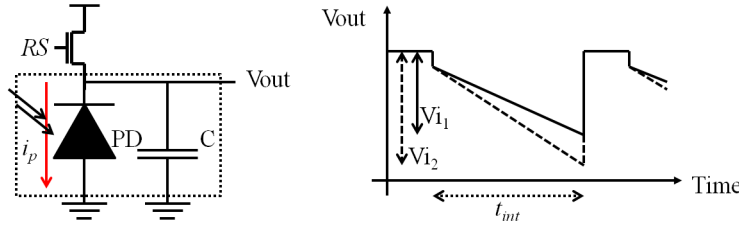


Figure 1.4 – Photodiode schematic model and voltage variation across the photodiode during photo integration (autointegration reverse biased working mode).

Before the definition of Active Pixel Sensors (APS) by E. Fossum in 1993, 1T Passive Pixel Sensors (PPS) were widely used but with limited sensor resolution and image quality. Indeed, this pixel structure (see figure 1.5) induces a very high Fixed Pattern Noise (FPN). It is also limited by the impact of the charge capacitance of the column bus connected to the readout circuitry. The main reading steps of a pixel acquisition for PPS are listed below.

- Reset: The photodiode is reverse biased via activation of RS .
- Integration: The incoming photons decrease the voltage across the photodiode.
- Read: The canonical output, in the form of a charge is measured (can be turned into a current or a voltage output).

The first compact structure of an APS, which is based on 3 in-pixel transistors, is commonly called the 3T-pixel (see figure 1.6). This 3T structure reduces the main drawbacks of the past 1Ts by adding the amplification stage at the pixel level and

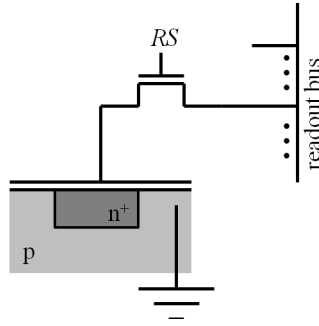


Figure 1.5 – Passive CMOS pixel (1T pixel).

separating the reset circuitry from the readout circuitry. 3T reading steps are listed below.

- Reset: The photodiode is reverse biased via the activation of RST .
- Integration: The incoming photons decrease the voltage.
- Read: The voltage output is measured via amplification of the voltage follower and activating RS .

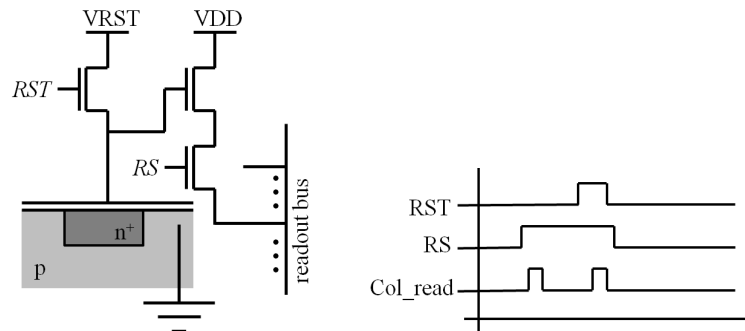


Figure 1.6 – Active CMOS pixel using an in-pixel voltage follower (3T pixel).

Pixels designed with a 4T structure partly overcome problems due to the 3T structure. Notice that, in the rest of this thesis 4T pixels will refer to 4T pinned photodiode pixels (alternative 4T pixels structures exist). The main interest lies in the possibility to implement a proper Correlated Double Sampling (CDS) which is used to remove the undesired offset introduced by the Voltage follower transistor (which is a source of FPN) and the time varying reset. The CDS is performed making the difference between the “black” Voltage of the pixel (i.e. the reset Voltage) and the Voltage at the end of the integration time. This 4T pixel has the same structure as the 3T pixel with an additional pinned photodiode which is connected to the sensing node via a transfer gate TX . This way, the photodiode is isolated from the sensing node. 4T reading steps are listed below.

- Reset: The pinned photodiode is reset via the activation of TX .
- Integration: The incoming photons are converted via the integration of charges in the pinned photodiode.
- Reference read: the readout node is reset by RST and then read by activating RS .
- Transfer: The activation of TX fully empties the photodiode by transferring all the charges to the sensing node.
- Signal read: A second measurement is then performed on the sensing node via a second activation of RS .
- CDS: The two measurements are subtracted from each other.

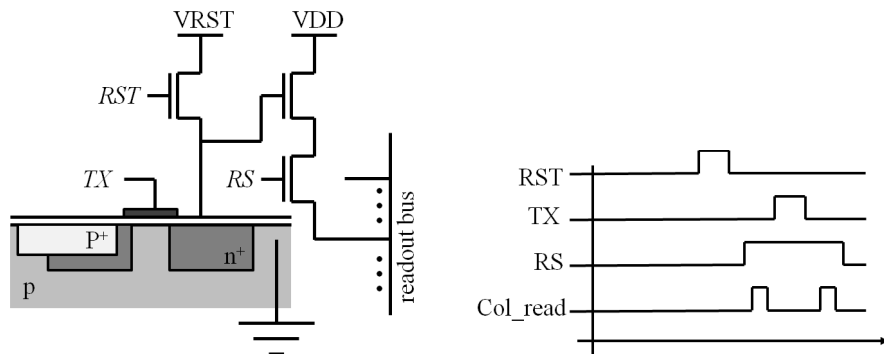


Figure 1.7 – Active CMOS pixel using an in-pixel voltage follower combined with a pinned-photodiode (4T pixel).

Those typical pixel structures are the most used in CIS because of their compact design. The following table 1.1 summarizes advantages and drawbacks of those main image sensor architectures.

	CCD	CMOS (1T)	CMOS (3T)	CMOS (4T)
Noise	Low	High	Moderate	Moderate to low
System complexity	High	Low	Low	Low
Sensor complexity	Low	Moderate	High	Higher
Sensor Modularity	Low	High	Highest	High
Fill factor	High	High to Moderate	Low	Lower to Higher (1.75T)
Chip output	Analog	Digital	Digital	Digital
Uniform shuttering	Possible	Impossible	Impossible	Possible w/o CDS
Cost	High	Low	Low	Low
Pwr consumption	High	Low	Low	Low

Table 1.1 – Comparison in terms of features and performances between CCD and CMOS image sensors.

In the case of CIS, column ADCs are used to parallelize the pixel readout operation. It puts special strains on the ADC design, especially regarding layout constraints. Indeed, ADC should first fit within the pixel pitch and the conversion rate should be at least equal to the row reading frequency. There exists multiple architectures of ADCs but generally one of the following 4 (Single slope, multi-slope, SAR, $\Sigma\Delta$) is employed in the end-of-column circuitry to convert the pixel values. Each of those 4 types of ADC has its own pros and cons which are reported below in table 1.2.

ADC	Status
Single slope	Low complexity Conversion rate limited by ramp generator No linearity issues
Multi-Slope	Enables finer bi resolution Linearity issues between slopes of coarse and fine steps
Successive Approximation	Possible lowest consumption Requires complex circuitry Limited by the DAC precision and noise
$\Sigma\Delta$	Very fast, high oversampling enables high bit depth Complex analog/digital circuitry Generally high power consumption Enables processing and filtering operations

Table 1.2 – Comparison of typical ADCs features for CIS.

However, many different limitations arise when one considers an image sensor based on one of those typical architectures. Even if not exhaustive, one can list the main limitations as follow (commonly employed as solution in italics):

- **Skew effect** due to rolling shutter (time inconsistency between each row). *Global shutter, $4T/5T$ pixels.*
- **Blurring effect** due to a too long time of integration relatively to the motion in the scene. *Large photodiodes, high sensitivity, low noise, short time of exposure.*
- **Resolution and frame rate** due to data output limited bandwidth and maximum clock frequency. *Binning, data compression.*
- **Power consumption** which is generally correlated to the total number of pixels (i.e. the resolution times the frame rate). *Binning, event based, charge injection.*
- **Dynamic range** due to intrinsic nature of measuring an integrated value and the sensibility of the pixels. *Bracketing, log pixels, thresholded pixel sensor(SPAD), integration time appropriate tuning.*
- **Fixed Pattern Noise** due to technological dispersion. *Proper design, large pixels, post-processing.*
- **Temporal Noise** which corrupts analog information driven during the acquisition. *Proper analog design (e.g. CDS).*

As explained in [10], the noise can be of different natures: temporal noise or spatial noise. Temporal noise refers to thermal reset -kTC- noise (which can be suppressed by CDS) [11], Random Telegraph Noise (RTS) noise [12], dark current shot noise [13], 1/f noise [14], photon shot noise [10], power supply noise... Spatial noise refers to dark fixed pattern, light fixed pattern, column fixed pattern, row fixed pattern and defect/malfunctioning pixels. The effect of temporal and spatial noises is generally reduced using specific on-chip analog implementations (e.g. a 4T pixel for kTC noise). Spatial noise can also be partly removed using digital post-processing if a calibration has been properly performed.

To summarize, canonical image sensor architectures have been somehow improved during the last two decades to limit noise effects and allowing to design imagers with higher resolutions and frame rates. As a result, the amount of data to be converted increases over years, implying an important need of data compression. Most of CIS systems uses a classical digital compression which is performed by a dedicated processing unit. Lossy compressions (using mathematical transforms) are generally employed for both still images and video sequences. In particular, for video compression, frame to frame differences and motion compensation are also used because of considerably helping to reduce the amount of data to store. Unfortunately, those operations require large processing units and consume a lot of resources (power, silicon, time,...).

Compressive Sensing, when whitening is beautiful

Regarding highly strained embedded systems, the power consumption is a sticking point. This power consumption is mainly due to analog to digital conversions and digital data compression. Digital processing units and ADCs respectively represent 90% and 9% of the global power consumption of a traditional digital CMOS camera. CS seems to be a good candidate to solve these two issues at the same time. Indeed, CS revisits the classical acquisition chain as shown in figure 1.8. The aim of this thesis is to show how the CS paradigm can help to relax on-chip constraints reducing the multiple acquisition limitations without addressing power consumption issues only. Moreover, it will state that CS could provide image features as well, for “features only” imaging.

Introduced in the mid-2000’, Compressive (or Compressed) Sensing (or Sampling) [7], [8], [15] and [16] states that a sparse (or compressible) signal can be efficiently sensed by relatively few linear projections if appropriately designed. Natural and physical signals are generally sparse, meaning that they can be represented in a certain basis with only few coefficients. This property has been intensively used in signal processing and more particularly regarding data compression purposes. For instance, a sinusoidal signal is sparse in the Fourier basis. A sinusoidal signal can be described by only 3 values: its frequency, its phasis and its amplitude. Typically, the Fourier Transform is widely used for applications in engineering and science to provide understandable interpretation of signal natures. The sparsity property is used in many different well-known compression

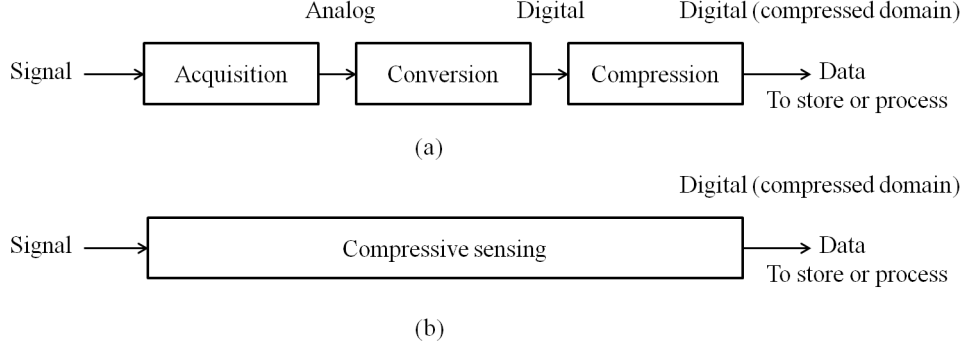


Figure 1.8 – The classical acquisition chain (a) revisited by the Compressive Sensing (b).

formats like MP3 (windowed- DCT), JPEG (2D- DCT) and JPEG 2000 (2D- DWT). The underlying goal of CS is to extract and acquire only the relevant information directly at the sensing stage to avoid waste of power and time due to complex signal analysis if involved. To efficiently perform such an acquisition, the measurements need to spread the information content of the signal. This way, each measurement contains a part of the information, maximizing the entropy of the measurements. This can be optimized when linear projection vectors are as incoherent from the sparsity basis as possible, and uncorrelated between each others. This consideration is directly linked to the Restricted Isometry Property (RIP) described in [17]. For instance, the Spike basis (Identity) and the Fourier basis are two incoherent matrices which means that a sinusoidal (or harmonic) signal can be efficiently sensed by taking samples at random positions (which corresponds to use as measurement vectors randomly picked up rows of the spike matrix). Alternatively, to efficiently sense a signal which only contains a relatively low number of peaks, one can take few randomly chosen coefficients in its Fourier representation. Regarding any physical signal that exhibits a certain underlying structure, an universal sensing can be performed using universal measurement matrices. Indeed, a certain class of randomly generated matrices can be used to acquire any sparse signal via CS. Figure 1.9 summarizes the difference between the principle of lossy compression by transform and compressive sensing. The notation we used will be defined more formally in the next chapter. In this figure, x corresponds to the original signal, Ψ the sparsity basis and ϕ the CS measurement matrix. The signal in its compressed form is denoted y_c for both classical lossy compression and CS. x_c thus represents the final reconstructed image after the decompression operations. Some works still tend to extend the basic concepts of CS first stated in [17]. The article [18] deals with quantization noise and saturation issues when CS is performed. Recently published works in the article [19] show that under a certain extension of the RIP, non-linearity of the measurements is a problem that now becomes tractable. This last article shows that if sufficient assumptions are satisfied on the measurements, CS can be performed via non-linear projections.

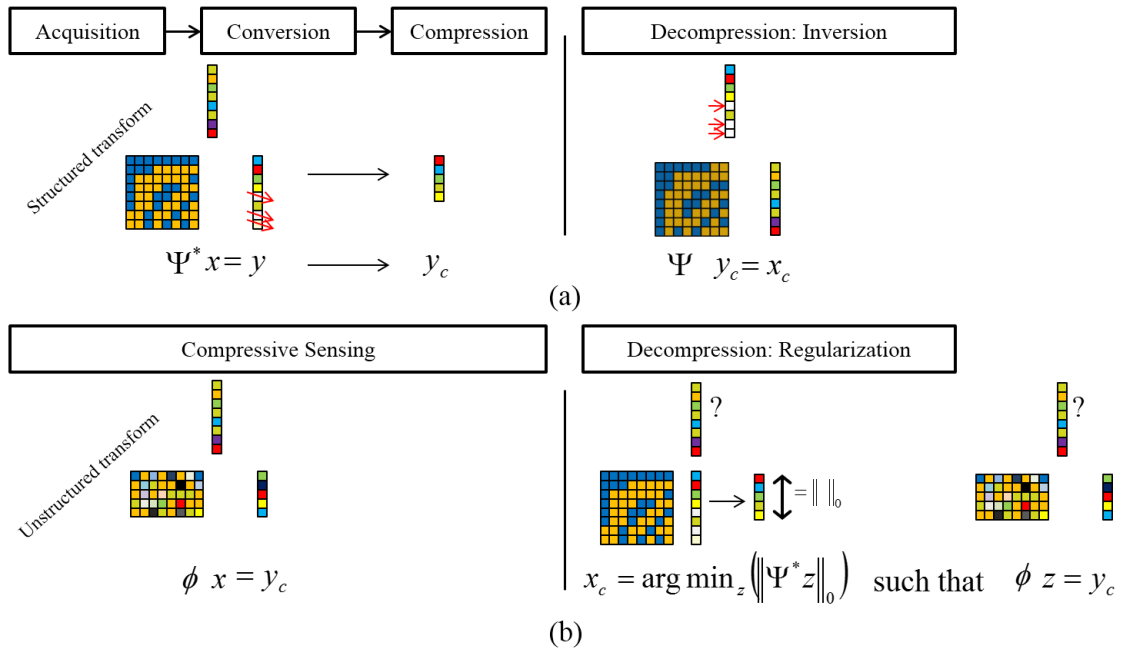


Figure 1.9 – The lossy compression by structured transform (a) revisited by the Compressive Sensing using non structured projections (b).

Both techniques (transform based compression and Compressive Sensing) take advantage of sparsity but when a classical lossy transform analyzes the signal to keep the essential information at the compression, CS spreads the information in less samples and reconstruct it by synthesis at the decompression stage using sparsity priors. The CS paradigm thus reverses the classical point of view of compression consisting in analysing the signal at the coder facilitating the decoding stage. CS is based on the acquisition of a smaller sized whitened signal containing all the information of the original signal but using a more complex synthesis procedure for decoding. This way, in the case of CS, the main computational cost is deported to the decoder. It leads to the fact that the main goal of the CS implementation is to reduce the global power consumption of sensing systems. Other limitations due to the sensor architectures could potentially be addressed by CS because of its low level, close to the sensor implementation. In the next chapter, theoretical and experimental results on CS will be discussed more precisely. In addition, the main already developed CS imaging systems will be presented in this state-of-the-art.

Randomization and sparsity: “New fields of applications and actual limitations”

Even if the CS principle works on any type of signal, it has been intensively studied in the case of 2D signals such as images. In terms of theoretical performance, the larger the signal the more attractive the CS. Indeed, in different fields of application, a noticeable evolution of compression techniques is required because of the constant increase of image sensor resolution. On the other hand, a major problem arises when large signals are involved because of issues related to the generation and the employment of appropriate CS measurement vectors, which is therefore the case when dealing with CS on still images, videos and multidimensional images. In addition, CS applications reveal technical and functional limitations for displaying, reconstructing or interpreting CS data. These issues are even worst when the signal size increases. Due to the intrinsic nature of the projections, it is somehow difficult to deal with acquired data. It is sometime impossible to reconstruct a compressively sensed signal with common optimization procedures because of the non-realistic requirements on memory and processing units. In general, the CS measurement matrix is of a pseudo-random nature to fulfil CS requirements and needs to be known for decoding. Therefore, it has to be stored in local memory to perform the reconstruction. It is thus more powerful if an implicit definition of the sensing matrix is available for reconstruction purposes. Yet, it generally implies a more complicated implementation at the sensing stage. For the moment, this trade-off dramatically limits potential CS applications. Targeted applications are thus generally related to niche markets. It often means to define a specific sensing scheme implementation (generate, store and apply the sensing matrix). This also hints to design dedicated algorithms to handle with those specific CS measurements. Furthermore, CS measurements generally do not provide as much understandable information as traditional compressed domains. However, in the literature, several applications have been explored using directly CS measurements as inputs, detection [20], classification [21], prediction [22] and tracking [23]. Indeed, the idea is that since CS measurements contain all the information, it is possible to design applications and specific algorithms that are taking advantage of this representation thanks to its intrinsic dimensionality reduction. Beyond these considerations, and thanks to RIP [24], close signals exhibit close representations in the compressed domain.

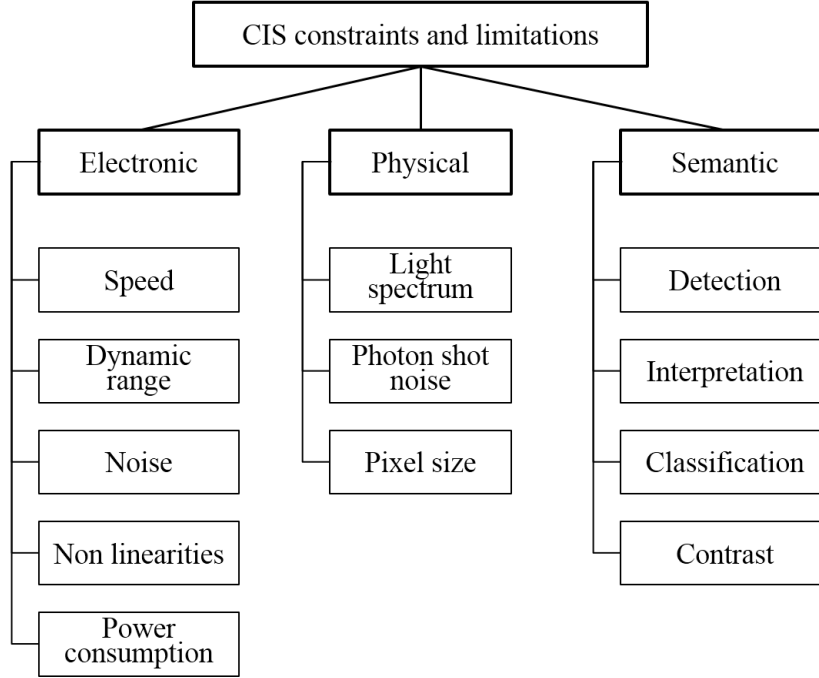


Figure 1.10 – The main limitations and constraints related to CMOS Image Sensors (CIS). Areas of study and fields of improvements regarding CS-CIS (and CS imaging systems more broadly).

This thesis focuses on three different focal areas (cf. figure 1.10) aiming to design proper architectures and acquisition processes with their associated reconstruction techniques taking advantage of image sparse representations. How on-chip implemented CS can relax sensor constraints, improving the acquisition characteristics (speed, dynamic range, power consumption) ? How CS can be combined with simple analysis to provide useful image features for high level applications (adding semantic information) and improve the reconstructed image quality at a certain compression ratio ? How CS can improve physical limitations (i.e. spectral sensitivity) of imaging systems without a major impact neither on the sensing strategy nor on the optical elements involved ?

2 State Of The Art

In the last century, with the development of electronic systems and more recently of computer science, signal processing have tended to take a crucial place in engineering. A large variety of sensing devices have been designed to acquire, measure and interpret the physical world, making the signal processing inescapable. Even if not exhaustive we can list: radars, medical imagery devices, cameras, seismic detectors, microphones, etc. Nowadays, the total amount of sensed data considerably increases on a daily basis [25] (as example: 60 million images per day on Instagram website –source: <http://instagram.com/press> 14/10/2014 –). Since traditional sensors aims at acquiring meaningful information, theoretical limitations often force a very fine sampling. For the sake of simplicity and implementation purposes, traditional sensors use a periodical discretization of the signal (whatever spaces involved). It thus leads to the famous Shannon-Nyquist theorem [26] implying that recovering a band-limited signal is possible only if sampled with a sampling rate at least two times higher than its bandwidth. Imaging does not deviate from this law, from the moment when digital cameras became available in the market, orthogonal periodic sampling was employed (squared pixels). Compressive sensing paradigm claims that more efficient sensing schemes can be defined for specific applications if certain assumptions on the signal can be made. In a certain sense, this paradigm tends to extend the Nyquist-Shanon theorem replacing the concept of band-limitation by sparsity. This state of the art first deals with the theoretical background on CS, then with the main community contributions regarding the designs of imaging systems and associated applications. The interests and limitations related to the implementation of such systems for real-world end-users and specific applications (i.e. astronomy, video-surveillance, medical imagery) are discussed.

2.1 Compressive Sensing background

Physical signals generally consist of a combination of structured waveforms even if the underlying information is hidden. Indeed, even if we restrict this to linear transforms,

most of natural signals are sparse. This concept of signal sparsity has been intensively used for compression purposes as noticed in the introduction section. Formally, what we call a sparse signal is a signal that can be represented by only few coefficients in a certain basis. It means that in this sparse representation the signal information is concentrated on only few peaks and the rest of coefficients are considered small or close to zero if the signal is said to be compressible. In particular, this interesting property lets showing-off intrinsic information of the signal when considered in the proper basis. It results that the CS paradigm claims that if a signal is sufficiently sparse it can be acquired with only few appropriate linear measurements compared to the signal size. More generally, CS can be interpreted as an extension of the Shanon-Nyquist sampling theorem [27]. The main theoretical results are presented after a brief note on the notations that will be used in the rest of this thesis.

2.1.1 A brief note on notations

For the sake of simplicity, we define here the notations used in all the following. Normal letters will be used for scalars (e.g. x), bold letters for column vectors (e.g. \mathbf{x}). For instance, $\mathbf{x} \in \mathbb{R}^N$ is a real N-dimensional vector composed of components x_i for $i \in \{1, \dots, N\}$. Capital letters represent matrices (e.g. X) and bold capitals are orthogonal matrices (e.g. \mathbf{X}). The i^{th} column vector of a matrix is noted X_i and $X_{i,j}$ denotes the entry on the i^{th} row and j^{th} column of X . For a given vector or matrix we denote its transposition (or adjoint) as $(\cdot)^*$. The scalar product between two vectors \mathbf{u} and \mathbf{v} in \mathbb{R}^N is thus denoted as $\mathbf{u}^* \mathbf{v} = \langle \mathbf{u}, \mathbf{v} \rangle$. For a given matrix X we respectively denote X^{-1} its inverse (if it exists) and $X^\dagger = (XX^*)^{-1}X$ its pseudo-inverse. $\#S$ denotes the cardinality of a set of elements S . The l_p norm (or pseudo-norm) of a vector is denoted by $\|\cdot\|_p$ for p between zero and the infinity and defined as:

$$\|\mathbf{x}\|_0 = \#supp(\mathbf{x}) = \#\{i : \mathbf{x}_i \neq 0, 1 < i < N\}, p = 0 \quad (2.1)$$

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^N |\mathbf{x}_i|^p \right)^{1/p}, 0 < p < \infty \quad (2.2)$$

$$\|\mathbf{x}\|_\infty = \max_{1 < i < N} |\mathbf{x}_i|, p = \infty \quad (2.3)$$

We also define coherence μ of the matrix $X \in \mathbb{R}^{M \times N}$ as follows:

$$\mu(X) = \sup_{i,j} |\langle X_i, X_j \rangle| \quad (2.4)$$

If we denote two orthonormal bases $\mathbf{X} \in \mathbb{R}^{N \times N}$ and $\mathbf{Y} \in \mathbb{R}^{N \times N}$; their mutual coherence μ is defined as follows and can be reached with particular couple of basis (e.g. Identity/DCT, Identity/Hadamard, Wavelet/Noiselet):

$$\mu(\mathbf{X}, \mathbf{Y}) = \sup_{i,j} |\langle X_i, Y_j \rangle| \geq 1/\sqrt{N} \quad (2.5)$$

CS intensively uses minimization/regularization problems, we therefore define "argmin $f(\mathbf{x})$ ", the procedure that returns the vector \mathbf{x} for which the function $f(\cdot)$ is minimized. In the case of constrained optimization, the shorthand "wrt." is used to mean "with respect to". Even if not a perfect metric, the Peak Signal to Noise Ratio (PSNR) will be intensively used for image quality assessment. The PSNR between two signals \mathbf{x} and \mathbf{x}_r (reference signal) whose the considered dynamic is between 0 and d , is defined as below.

$$PSNR(\mathbf{x}, \mathbf{x}_r) = 10 \log_{10} \left(\frac{d^2}{MSE(\mathbf{x}, \mathbf{x}_r)} \right) \quad (2.6)$$

$$MSE(\mathbf{x}, \mathbf{x}_r) = \frac{1}{\#\mathbf{x}_r} \|\mathbf{x}_r - \mathbf{x}\|_2^2 \quad (2.7)$$

The Structural Similarity Image Measurement (*SSIM*) [28] is also used in this thesis to assess image reconstructions. This quality index returns a value between 0 and 1 (can be expressed in %). A quality index of 1 means that $\mathbf{x} = \mathbf{x}_r$. *SSIM* is defined by the equation 2.8 with $\bar{\mathbf{x}}$ mean of \mathbf{x} , $\sigma_{\mathbf{x}}^2$ variance of \mathbf{x} and $\sigma_{\mathbf{x}, \mathbf{x}_r}$ the covariance of \mathbf{x} and \mathbf{x}_r . Compared to the *PSNR*, *SSIM* provides an image quality assessment which is more sensitive to structural distortions and less sensitive to the energy of errors. It is also more correlated to visual perception of an image quality which explain why it is often considered more accurate than the *PSNR*.

$$SSIM(\mathbf{x}, \mathbf{x}_r) = \frac{4\bar{\mathbf{x}}\bar{\mathbf{x}}_r\sigma_{\mathbf{x}, \mathbf{x}_r}}{(\bar{\mathbf{x}}^2 + \bar{\mathbf{x}}_r^2)(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{x}_r}^2)} \quad (2.8)$$

2.1.2 The CS paradigm

The sparsity is an inherent characteristic of many physical signals and particularly in the case of multidimensional signal such as images. It makes the use of lossy compression by transforms an efficient way to compress any sparse signal. The image is called k -sparse in a certain domain when only k expansion coefficients are significant and the other $(N - k)$ coefficients are negligible. Typically, images are sparse in a large variety of transform domains (e.g. Laplacian, Wavelet, Contourlet, Discrete Cosine,...). Let \mathbf{x} be a signal that we want to acquire using CS. Let's consider \mathbf{x} as a vector composed of N real values ($\mathbf{x} \in \mathbb{R}^N$). If \mathbf{x} is considered exactly k -sparse, it can be represented by a linear combination of k row-elements taken in ψ_n , $n \in \{1, \dots, N\}$ of a certain orthonormal basis Ψ^* ($\Psi \in \mathbb{R}^{N \times N}$), i.e., if $\|\alpha\|_0 = k$ is the compressed representation.

$$\mathbf{x} = \sum_{n=1}^N \alpha_n \psi_n \quad (2.9)$$

Thanks to sparsity, α_n equals zero $(N - k)$ times for $n \in \{1, \dots, N\}$. The \mathbf{x} signal representation in this basis Ψ^* is thus α , a coefficient vector which contains only k non-zeros entries. In matrix-to-vector multiplication form it gives the following:

$$\mathbf{x} = \Psi \alpha, \|\alpha\|_0 = k \quad (2.10)$$

Now, let Φ be the sensing matrix composed of M row measurement vectors Φ_i . y thus corresponds to the resulting measurements of \mathbf{x} by Φ . Notice that $\Phi \in \mathbb{R}^{M \times N}$ and $\mathbf{y} \in \mathbb{R}^M$.

$$\mathbf{y} = \Phi \mathbf{x} = \Phi \Psi \alpha \quad (2.11)$$

Considering the case when $M \ll N$, recovering \mathbf{x} from \mathbf{y} and Φ is an ill-posed problem because Φ is non invertible. In this case, an infinity of possible solutions satisfies the previous equation. To simplify the following we denote: $\Theta = \Phi \Psi \in \mathbb{R}^{M \times N}$. The CS handles this issue by taking advantage of prior knowledge of the level of sparsity of the signal \mathbf{x} in the basis Ψ . Therefore, the optimal/theoretical CS reconstruction is based on looking for the sparsest signal α' among all possible solutions of \mathbf{u} such that $\mathbf{y} = \Theta \mathbf{u}$.

If the basis is well suited to the signal to acquire, there exists a unique solution.

$$\alpha' = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbf{u}\|_0 \text{ wrt. } \mathbf{y} = \Theta \mathbf{u} \quad (2.12)$$

Because of the l_0 norm presence, solving this problem is nearly impossible (known as NP-Hard). In the CS literature [29], it is proposed to replace the l_0 norm by the l_1 norm in order to relax this optimization procedure. The l_1 regularization problem is the closest convex relaxation of l_0 . Indeed, using the Basis Pursuit algorithm, the problem becomes solvable. In particular, this l_1 minimization problem is now a convex problem which is much easier to solve. Under a certain constraint on the level of sparsity (i.e. $\|\alpha\|_0 < \frac{1}{2}(1 + \mu(\Theta)^{-1})$) [30], the solution of the l_1 minimization equals the solution of the l_0 . Many other theoretical results have been stated for the exact recovery for a CS acquisition implying the number of measurements, the signal size and the coherence between the sensing and the sparsifying matrices Φ and Ψ [8].

$$\alpha' = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbf{u}\|_1 \text{ wrt. } \mathbf{y} = \Theta \mathbf{u} \quad (2.13)$$

In practice, most physical signals are not exactly sparse but considered compressible meaning that there exists a sparse approximation of the signal. For instance, we denote the k -sparse approximation of \mathbf{x} , $\mathbf{x}_k = \Psi \alpha_k$ such that $\|\alpha_k\|_0 = k$. Then we can define the l_2 k -sparse approximation error as $\epsilon(k) = \|\mathbf{x} - \mathbf{x}_k\|_2 \leq \|\mathbf{x} - \mathbf{x}'\|_2$ for any \mathbf{x}' such that $\|\alpha'\|_0 \leq k$. Considering this, in the compressible case it is more suited to solve the following equation than the previous one, with an $\epsilon \in \mathbb{R}$ properly set.

$$\alpha' = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbf{u}\|_1 \text{ wrt. } \|\mathbf{y} - \Theta \mathbf{u}\|_2^2 \leq \epsilon \quad (2.14)$$

Due to computational issues, this previous problem is generally reformulated in its augmented Lagrangian form (*known as synthesis-based sparsity*):

$$\alpha' = \underset{\mathbf{u}}{\operatorname{argmin}} \|\mathbf{u}\|_1 + \lambda \|\mathbf{y} - \Theta \mathbf{u}\|_2^2 \quad (2.15)$$

This problem can be also reformulated to work directly in the physical domain and not

in its sparse domain (*known as analysis-based sparsity*:

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} \|\Psi^T \mathbf{u}\|_1 + \lambda \|\mathbf{y} - \Phi \mathbf{u}\|_2^2 \quad (2.16)$$

This previous problem can be extended to any optimization that uses a constraint operator Op (see equation 2.17) since it promotes either sparsity or at least exhibits the intrinsic structure of the signal. In such cases, a formal proof of convergence, convexity or exact reconstruction can not always be demonstrated but experimental results often show improvements compared to a traditional, basic CS reconstruction using a single sparsity representation of the signal.

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} Op(\mathbf{u}) + \lambda \|\mathbf{y} - \Phi \mathbf{u}\|_2^2 \quad (2.17)$$

$FT(\mathbf{u}, \mathbf{y}) = \|\mathbf{y} - \Phi \mathbf{u}\|_2^2$ is called either the data-fitting function or the fidelity term whereas the term $Op(\mathbf{u})$ is called the prior term. λ is the regularization parameter which allows a certain degree of tuning the weight between prior term and fidelity term. Theoretically, this term has not much impact on the reconstruction if the signal is exactly sparse. In practice, it has to be tuned carefully because of non exact sparsity and noise. Regarding the basic CS problem (i.e. equation 2.14), the Restricted Isometric Property (RIP) criteria is typically used to guarantee the exact and stable reconstruction of the original signal. The matrix Θ satisfies the RIP of order S with restricted isometry constant δ_S if for every submatrix $\Theta_S \in \mathbb{R}^{M \times S}$ of $\Theta \in \mathbb{R}^{M \times N}$ and for every vector \mathbf{z} there exists a constant $\delta_S \in (0, 1)$ such that:

$$(1 - \delta_S) \|\mathbf{z}\|_2^2 \leq \|\Theta_S \mathbf{z}\|_2^2 \leq (1 + \delta_S) \|\mathbf{z}\|_2^2 \quad (2.18)$$

The RIP states that each set of S columns of Θ represents an orthonormal system. The RIP of the CS matrix with $\delta_S \leq 0.2$ guaranties the exact recovery of a $S/3$ -sparse signal when using convex optimization problem 2.13 [31]. Other theoretical results exists henceforth, with better constraints. As example a part from many others, [24] shows that if $\delta_{2S} \leq \sqrt{2} - 1$ the recovery of a S -sparse signal is exact.

2.1.3 Reconstruction algorithms

A very large variety of algorithms now exists to solve the above optimization problems. In particular, convex programming is a very efficient way to solve CS-related reconstruction problems. As presented in [32], we propose to classify the reconstruction algorithms into six different types of methods that are listed below.

- **Greedy Iterative Algorithms** find a solution in an iterative fashion using a step by step procedure [33]. The common principle is to select iteratively each column of Θ that correlates most with \mathbf{y} , minimizing the least square error (the fidelity term) [34]. Despite its intrinsic precision (when the signal is exactly sparse), this kind of algorithm is hardly suitable for a real application purpose because of being highly computational demanding.
- **Convex Relaxation algorithms** correspond to the main, first proposed algorithms to perform the CS reconstruction in the non-exactly sparse noise situation [35]. It is based on solving a convex optimization problem using linear programming. Like greedy algorithm, it is quite efficient in terms of requirements on the number of measurements but it is often considered too computationally complex.
- **Iterative Thresholding Algorithms** perform CS recovery faster than the two previous classes of algorithms [36]. This class of algorithms is attractive because of its simplicity of implementation, however basic implementations are known to converge relatively slowly. Slight modifications can be added preserving the simplicity of the algorithm but dramatically improving the global rate of convergence [37]. In [38] and [39] new iterative-thresholding algorithms are inspired by belief propagation approaches. It derives an other algorithm class called message-passing algorithms.
- **Bayesian Network Algorithms** are related to algorithms that consider CS from a Bayesian perspective [40] and when measurements can be seen as observations.
- **Bregman Iterative Algorithms** are good candidates in the future because of their computational speed and its simplicity but could be unsteady in certain condition [41].
- **Combinatorial Algorithms** are based on group testing. It has been shown that this class of algorithms allows very fast reconstructions but needs specific requirements on the sensing matrix to be proficient.
- **Non Convex Minimization Algorithms** are probably the most interesting field of upcoming research but because of its non convexity, global minimization is not always guaranteed. However, in practice, it seems that a local minimum of a non convex problem is sometimes a better solution than the global solution the equivalent convex relaxation problem [42].

In this thesis, we will mainly deal with Iterative Thresholding Algorithms because of their modularity, speed, efficiency and reliability. The article [43] reviews the proximal

splitting methods, gathering several well-known iterative minimization algorithms in a unifying framework. More particularly, the Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) [37] of the Unlocbox [44] will be mainly used to solve minimization problems related to the equation 2.17 (see algorithm 1). This algorithm differs from the naive Iterative Shrinkage-Thresholding Algorithm (ISTA) due to the fact that the shrinkage operator $p_L(\cdot)$ is not directly applied to the previous computed \mathbf{x}_{k-1} but to \mathbf{y}_k which is a local additional side variable resulting of the linear combination of \mathbf{x}_{k-1} and \mathbf{x}_{k-2} .

Algorithm 1 FISTA: Fast Iterative Shrinkage Thresholding Algorithm

Goal: find α such that $\alpha = \underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x}) + g(\mathbf{x})$
Input L , $\mathbf{y}_1 = \mathbf{x}_0$, $t_1 = 1$
While convergence is not reached
 $\mathbf{x}_k = p_L(\mathbf{y}_k) = \underset{\mathbf{u}}{\operatorname{argmin}} g(\mathbf{u}) + \frac{L}{2} \|\mathbf{u} - \left(\mathbf{y}_k - \frac{1}{L} \nabla f(\mathbf{y}_k)\right)\|^2$
 $t_k k + 1 = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$
 $\mathbf{y}_{k+1} = \mathbf{x}_k + \left(\frac{t_k - 1}{t_{k+1}}\right) (\mathbf{x}_k - \mathbf{x}_{k-1})$
Endwhile
 $\alpha = \mathbf{x}_k$

A wide range of applications have thus been presented/developed by the signal processing community since the early days of CS. Theoretical results so as practical progress are born from the original CS statements. Even if computation time dedicated to CS reconstructions still remains a bottleneck, the main limitation resides in the hardware implementations of CS and its application in a real case scenario. Indeed the signal processing community is waiting for working with really accurate data coming from existing sensor systems. Listed below, the most important CS based imaging systems which have been developed in the last decade are presented.

2.2 Hardware CS implementations and field of applications

Since the emergence of CIS, numerous works did focus on the implementation of intelligence close to the sensor. Indeed, the CMOS technology allows to add specific processing directly in the focal plane. For instance, a review of CMOS image sensors with an integrated on-chip image compression is provided in [45]. A large variety of techniques and architectures have been proposed in the last decades, we can chronologically list: JPEG compression [46], raw Discrete Cosine Transform (DCT) [47], Predictive Coding [48][49], Set Partitioning In Hierarchical Trees (SPIHT) [50], Discrete Wavelet Transform (DWT , Haar wavelet) [51], Quadrant Tree Decomposition (QTD) [52] and Visual Pattern Image Coding (VPIC) [53]. Unfortunately those implementations generally exhibit severe drawbacks compared to an external dedicated digital implementation, even if a noticeable gain in terms of power consumption can be observed. However, the implementation of CS would change this traditional point of view providing a very efficient way to perform

the compression at low level of the sensing device and dramatically improving the power consumption of the entire imaging system. In addition, the implementation of CS in an image sensor would also increase its capabilities compared to traditional compression techniques because of relying on simple linear projections.

2.2.1 Acquisition systems and CS-CIS

Two main technological evolutions are currently leading the image sensor world. The image sensors resolution still increases and with it the availability for parallelizing tasks (e.g. 3D stacked technology [54] [55], smart sensors [4]). Thanks to the image sparsity, Compressive Sensing is thus expected to have an important impact on Visible light sensors. More largely, Compressive Sensing can be applied to many different types of sensors (i.e. Magnetic Resonance Imaging (MRI) [56], Terahertz (THz) [57], Time of Flight (TOF) [58], Radars [59],...). It has already shown that it can dramatically reduce the time of acquisition for MRI while not impacting on the image quality [60]. Regarding imaging systems, we can discern two different kinds of Compressive Sensing implementations in the literature. Those two types of CS implementations can be differentiated by the domain in which the measurements are performed, either in the optical domain or in the electronic domain. First, the Rice Single Pixel Camera (SPC) [61] has shown that with appropriate optical modulations it is possible to perform CS measurements directly before the sensing operation. Different applications have been developed in that scope reusing the concept of the SPC and for a large range of spectral bands (see [62] [63] [64]). Then, several works arise in relation with CMOS image sensor dedicated to the acquisition of compressively sensed images. Those two different point of views aim at addressing distinct problems. The use of a single photodetector is interesting in the case of very expensive transducers or for those that are not compliant with a dense array implementation. CS has found extensive applications because of its impact in reducing the whole number of required photosensitive cells for a given image resolution. Compared to a full sequential scan, the main interest is to reduce the acquisition time. Regarding CS CMOS Image Sensor (CS-CIS), the goal is to implement a form of compression before or during the A/D conversion in order to relax constraints on power consumption, data bandwidth and data throughput at the sensor chip level.

Optically performed linear measurements

Optically performed CS applies the CS paradigm following a theoretically compliant framework thanks to the scalar products that are realized in the continuous domain before the transduction and the analog to digital conversion. Indeed, the use of appropriate optical elements to perform the linear combinations is one of the most elegant way to apply the CS concept which corresponds to the acquisition of the signal directly in its compressed form.

Rice Single Pixel Camera (SPC): DMD

Nowadays, digital cameras generally acquire images in spatial domain by using a large number of photodetectors. The array of photodetectors forms the pixel matrix of the image. Therefore, the resolution of the acquired image is directly linked to number of photodetectors. A binning operation between photodetectors can be performed to average by block of neighboring pixels reducing noise impact on the image but decreasing at the same time the output sensor resolution. The Single Pixel Camera (SPC) [61][65] was the first proposed implementation of a CS imaging system. Compared to binning, it operates reversely. It is based on the use of a single photodetector for achieving a multiple pixel image acquisition. Indeed this single detector is used to capture different compressive measurements performed thanks to a Digital Micromirrors Device (DMD). The required linear projections are thus performed using optical devices put in front of the sensing device. The final image resolution is thus defined by those optical devices. Figure 2.1 shows the proposed generic implementation of a single-pixel camera that uses a single photodiode as photodetector.

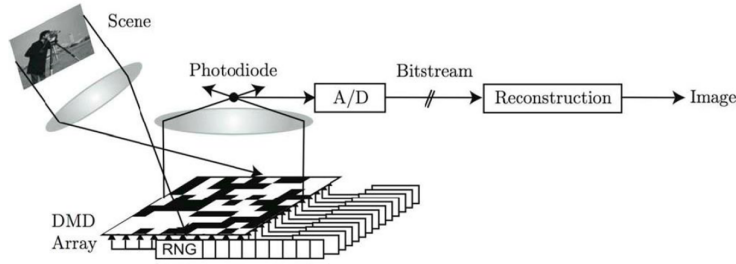


Figure 2.1 – Single pixel camera using DMDs [61]. The scene is focused on the DMDs plane then reflected (Spatial Light Modulation) and finally focused on a single photodiode

The main interest of using a single detector is to get rid of problems and constraints that arise when an array of pixels is involved. For instance, practical and technological problems are generally due to the following issues:

- technological dispersion
- spatial and temporal noise
- power consumption and current peaks
- production cost

For some specific types of sensors (THz, Infrared and Lidar) the production cost related to the chip area is so high that the use of an intermediate spatial modulator (DMD in [66] and later in [67] and [68]) is an easy way to increase the resolution of such sensors without increasing the size of silicon area. The resolution is therefore only limited by the resolution of the modulator. Figure 2.2 shows reconstruction results obtained when using

the SPC. In this figure, CS reconstructions are compared to the sparse approximations using both Haar wavelet as sparsity basis. At this time, reconstruction algorithms were not as optimized as today's and provided images of relatively poor quality at the end. In addition, as can be seen in figure 2.2 (d), the original image reflected by the DMD is also suffering of a poor quality, probably caused by misalignment of the optical system and focus troubles. One can notice that misfocus can make the reconstruction completely fail because of being unknown [69].

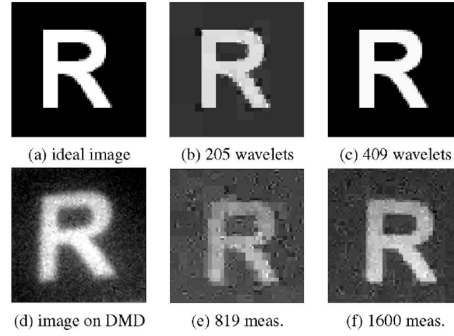


Figure 2.2 – Reconstruction results which are presented in [61]. A comparison between ideal and exact sparse images (top row) and the SPC acquisitions for two different compression ratios.

However, this sensing system is not well suited for all types of acquisitions. Indeed, one CS measurement corresponds to the averaged level of light after being integrated by the photodiode. Those measurements are thus achieved by controlling many different pseudo-random arrangements of the micro-mirrors. All the single-pixel cameras are based on serial measurements. Typically, all the measurements are performed during consecutive time intervals. These considerations causes several sticking points. First, no parallelization of the measurements is made possible. The sequential nature of the acquisition thus limits possible video acquisition (moving object). Then, the power consumption of such a system is not negligible. The generation of the pseudo-random sequences has to be moved aside. In addition, controlling an array of DMD is power consuming because of the number of DMD activations for a single image acquisition ($\mathcal{O}(M \times N)$). Lets compare the power consumption of a traditional CIS with a SPC. The power consumptions of the pixel readout and its conversion are respectively denoted P_{pixel} and P_{ADC} . The activation of a mirror belonging to the DMD is noted P_{DMD} . If the SPC consumes less than a traditional sensor for given M and N , thus we have the equation 2.19 which leads to the equation .

$$MN \times (P_{DMD}) + M \times (P_{pixel} + P_{ADC}) \leq N \times (P_{pixel} + P_{ADC}) \quad (2.19)$$

$$P_{DMD} \leq (P_{pixel} + P_{ADC})(1/M - 1/N) \quad (2.20)$$

It thus means that to make the single pixel camera attractive in terms of power consumption, the power consumption linked to the activation of DMDs needs to be very insignificant. For instance, considering a compression a ratio of 0.10 and a VGA (640×480) format, the activation of approximatively ten thousands DMDs has to consume less than an unique pixel read-out and its A/D conversion.

Transmissive SPC: LCD

The Transmissive SPC is based on the same principle as the SPC using DMD. The aim of this technique is to operate the acquisition in a transmissive mode rather than reflective. A reason of this choice is to reduce the optical instrument congestion. For that purpose, Liquid Crystal Device (LCD) is used as the spatial modulator, replacing commonly used DMD. Figure 2.3 shows the proposed scheme of the developed camera [70].

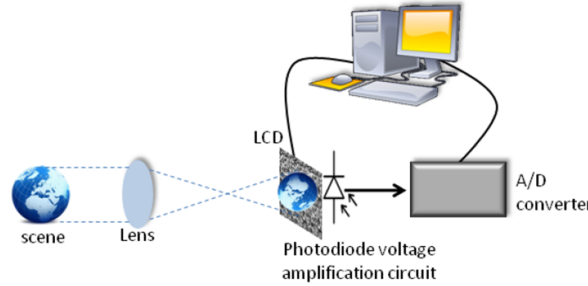


Figure 2.3 – Single pixel camera using LCDs [70]. The scene is focused on the LCDs plane then transmitted (Spatial Light Modulation) and finally focused on a single photodiode.

The sensing scheme used in this paper [70] is based on the Hadamard-based random binary codes. It reveals that this kind of sensing matrix is particularly efficient because those generated pseudo-random binary codes can be build on-the-run during both the sensing stage and reconstruction stage. However it limits the range of possible applications related to time modulations (because of reaching cycles in the Hadamard based modulation vectors) for video acquisition and it still requires a connection between the LCD and a computer to properly control the modulations during the acquisition. Figure 2.4 shows several reconstructions using the proposed system and for a set different compression ratio. It can be noticed that a limitation of this system is its poor light transmission level which represents around 20% of the incoming light at best. More generally LCDs have intreresting properties as well to spread the light spectrum (see chapter 5).

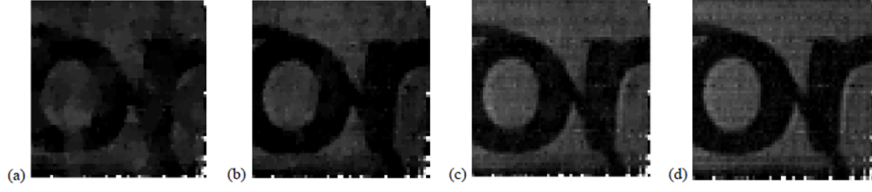


Figure 2.4 – Reconstruction results presented in [70] for different compression ratios, from left to right, 0.25, 0.5, 0.75 and 1). The results are notably better than the ones presented for the single pixel camera using DMDs [61], probably due to the improvements in terms of optical calibration

Coded aperture

A entire class of CS imaging systems are based on the coded aperture technique. Cameras generally use a combination of apertures and lenses to focus the incoming light from a targeted object onto the sensor focal plane. This way each photodetector in the imager array corresponds to a single pixel of the reconstructed image. However, in the case of coded aperture the optical field from the scene is modulated. In the CS literature we can found three main applications to this technique, super-resolution [71], acquisition of scenes with fast motion [72] and acquisition of high speed periodic videos [73]. More recently, it has also been proposed for surveillance systems [74], in this work background extraction model and object tracking is directly incorporated at the reconstruction stage.

In [71] coded aperture is used for high resolution imaging without increasing the resolution of the image sensor (i.e. the number of physical pixels). The aperture has been properly designed so that different CS measurements are performed by the different pixels. Compared to the single-pixel imager, CS measurements are parallelized. It particularly improves the time consistency between all the measurements. Later, in [75], the sequential nature of common CS is questioned. Indeed, it proposes a one-shot procedure allowing to employ a fixed sensing pattern instead of a complex controllable device (i.e. a DMD or a LCD).

The works presented in [73] focus on the temporal periodicity of the pixels in periodic videos. The acquisition technique is called coded strobing. The principle is to see the values of each pixel over time as a single dimensional signal. Since this signal is considered sparse in the frequency domain (because of signal periodicity), CS can be designed to take advantage of the acquisition of such signals meaning that the strobing operation needs to be incoherent with the Fourier basis to work the best way. Figures 2.5 and 2.6 shows the acquisition model for the proposed coded strobing scheme.

This application particularly takes advantage of the periodical intrinsic nature of each pixels independently from each others. The sparsity prior here is only through the time axis (a single dimension). In particular, if local random apertures were used, image

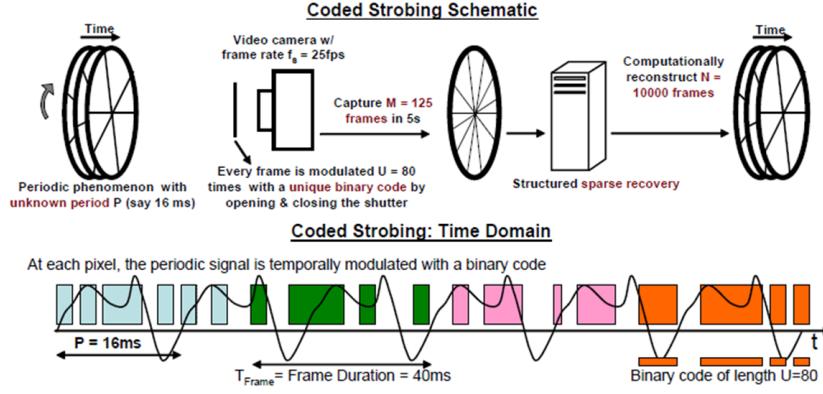


Figure 2.5 – Coded strobing scheme presented in [73]. This technique employs an aperiodic time modulation of the pixel integration. In time domain those successive modulations for each frame are respectively depicted in light blue, green, pink and orange.

sparsity priors could be added to the time periodicity prior. However, this would result into a far more complex imaging acquisition system and would imply an increasing complexity at the reconstruction too.

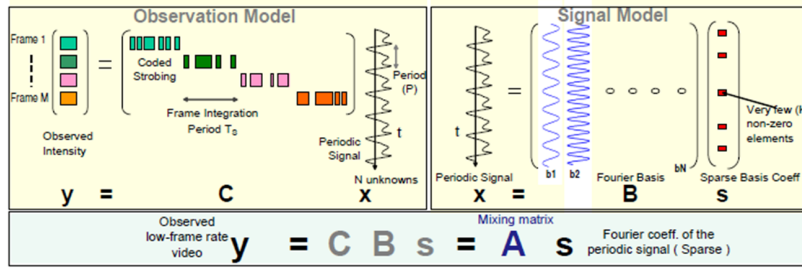


Figure 2.6 – Mathematical model of the Coded strobing system [73].

This coded aperture approach has also been used for spectral imaging what will be discussed in details in chapter 5. Using this type of approach, any other applications such as depth imaging, light field and super-resolution have been addressed too.

Random lens

The CS imaging system developed in [76] uses random lens based on multi-faceted mirrors. In this system, the light rays are randomized thanks to a random direction reflective surface. Compared to a traditional camera, the only additional part is constituted of the randomizer mirror. The camera setup which is presented in figure 2.7 shows that such a system can be easily integrated into a mass-market product. Super-resolution and depth imaging are the main potential applications which are depicted in [76]. It takes advantage of the randomization of the light rays as CS operator. The problem is that a complex calibration is required which represents the major drawback of this system. In

addition, this technique hardly allows to modify the patterned mirrors for instance for video acquisition.



Figure 2.7 – Camera setup used for experiments presented in [76]. The use of a combination of random mirrors with a flat mirror allows spatial modulation of the image before being acquired by the common camera’s sensor.

Imaging with nature

A direct extension of the use of random lens is the use of natural materials as scattering media. This technique called “imaging with nature” is presented for the first time in [77], 7 years after [76]. Naturally occurring randomness of wave propagation is exploited to design an optimal CS scheme. This compressive imaging system relies on scattering theory in order to perform appropriate random projections for CS in the case of active imaging using a coherent monochromatic light. Figure 2.8 shows the proposed framework for compressive imaging using multiply scattering media to record the deformation of a light propagating front.

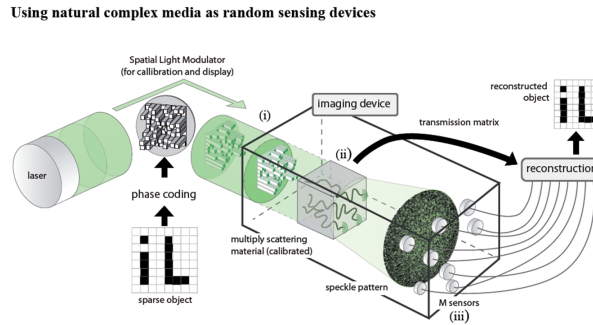


Figure 2.8 – CS Imaging system based on the optical scattering of some media [77].

A traditional image sensor can be used for the acquisition of resulting light intensities. A relatively low number of randomly selected pixel values thus corresponds to CS measurements. The number of selected pixels directly defines the compression ratio. The first advantage of this type of CS implementation is that it dramatically enables the parallelization of the measurements without reducing the theoretical efficiency of the

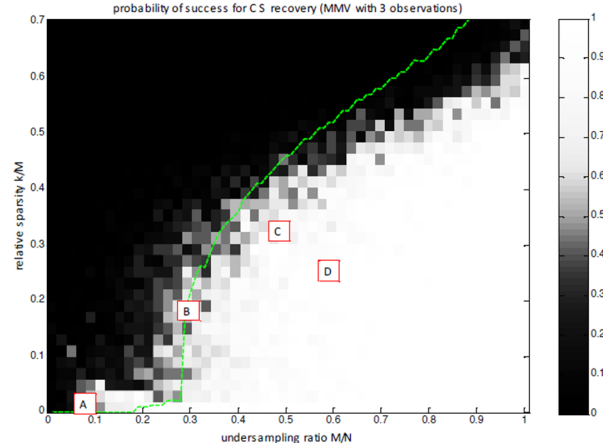


Figure 2.9 – Reconstruction results presented in [77]. Each point represents the probability of successful recovery for a k sparse image of N pixels using M measurements (around 50 trials per point).

acquisition. In addition, this concept can be directly adapted in other domains of optics, the only constraint is to find a media that dispatches information in a random way.

Figure 2.9 presents the experimental results of image successful recovery using the system introduced in [77]. Recently, [78] investigates possibilities of such an acquisition taking advantage of the intrinsic randomness nature of speckle patterns.

Lensless imaging

In the article [79], the authors propose a lensless compressive imaging architecture. As in [70], this architecture is composed of a single photodetector and an array of aperture elements, each independently controllable. The prototype of the sensor uses a LCD panel for the aperture assembly performing the linear CS projection.

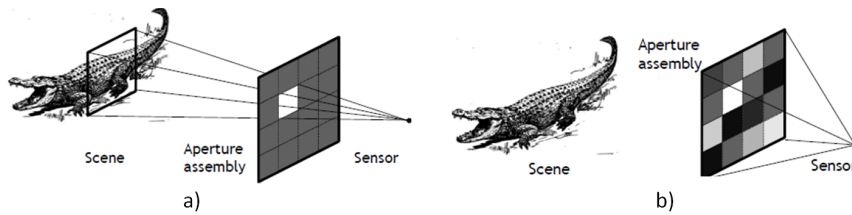


Figure 2.10 – Lensless structure using controllable LCD [79]. a) presents a light rays cone for a single aperture element. b) shows a example of aperture assembly configuration corresponding to a specific command.

Each CS measurement performed by the photodetector corresponds to the integration of rays in the cones modulated by the transmittance of the aperture elements. Indeed each aperture element has its own cone of observation. At the expense of post processing and

2.2. Hardware CS implementations and field of applications

taking advantage of the absence of lens, the proposed architecture shows particularly interesting features. First, thanks to the absence of lens, the imaging device can be really compact if the distance between the aperture elements and the photodetector were not a limitation. Secondly, the focus can be performed at the reconstruction and there is no blurring artifact introduced by a lens. At last, multi-view imaging is easily enabled by using multiple photodetectors on the back of the aperture elements. This last feature has been particularly exploited in [80] showing that multi-view CS can enhance the quality of reconstructed images.

Comparison of the different optically performed CS

Reference	[61]	[70]	[71]	[77]	[79]
Year	2006	2011	2008	2013	2013
Optical element	SLM (DMD)	SLM (LCD)	Random phase shift	Nature (paint)	SLM (LCD)
Programmable	yes	yes	yes	no	yes
Parallelization	no	no	yes	yes	available
Special feature	none	none	super-resolution	multispectral	3D
Compactness	poor	moderate	moderate	poor	moderate
Optical calibration	complex	complex	required	complex	moderate
Reconstruction calibration	none	none	moderate	complex	complex

Table 2.1 – Comparison between the main imaging systems using optically performed CS.

In the above systems, randomness of the measurement vectors is generated outside the sensor. CS measurements are performed in the optical domain, sequentially. Yet, most suitable architectures propose to parallelize the measurements in order to speed up the acquisition. In addition, the use of relatively complex optical elements induces difficult calibrations and can dramatically increase the whole camera size. Applying optical modulations still remains the most elegant way to perform the CS because of being performed before the actual acquisition/conversion. It yet implies many previously listed limitations. An additional limitation arises since the size of the optical unit elements are very small because of diffraction issues (e.g. in the case of LCDs and DMDs). Therefore, performing the CS at the sensor level can be considered as the most attractive alternative for real application purposes.

CS-dedicated CMOS image sensor

In the following, on-chip, on focal-plane CS implementations will be presented. It will refers to CS-CIS (i.e. Compressive Sensing CMOS Image Sensors). Obviously, the CS measurements operator could be implemented after the A/D conversion, in the digital domain, but if this is the case, CS is not fully exploited in the sense that more than needed and useful information is being processed. Therefore, analog implementations are generally preferred for both decreasing the amount of the on-chip storage area and reducing power consumption.

Separable transform

Before developing CS-dedicated image sensor, several works have been presented on analog computing imagers [48] [50] [81]. The authors of the article [82] proposed in 2010 a CIS architecture performing any separable transform directly in the focal plane by analog operations (i.e. before the ADCs). A two dimensional separable transform can be represented by a 2D basis function defined as the product of two 1D basis functions. For images it corresponds to apply successively a transform on the rows first and then on the columns, independently. For an input image P_σ and A , B respectively column and row transform matrices, the output of the separable transform is denoted $Y_\sigma = A^T P_\sigma B$ (see figure 2.11). In [82], this architecture is used as a CS encoder whose transform coefficients are chosen appropriately. This architecture is particularly interesting because of the plurality of possible working modes. Indeed, it can be tuned to perform either classical transform coding using quantization and then entropy coding or, alternatively, directly CS. In the case of CS, it thus simultaneously combines the sensing and compression which is an advantage because of the absence of the need for an entropy coder. Notice that, the separability property of the transform makes the reconstruction procedure easier in terms of algorithmic complexity.

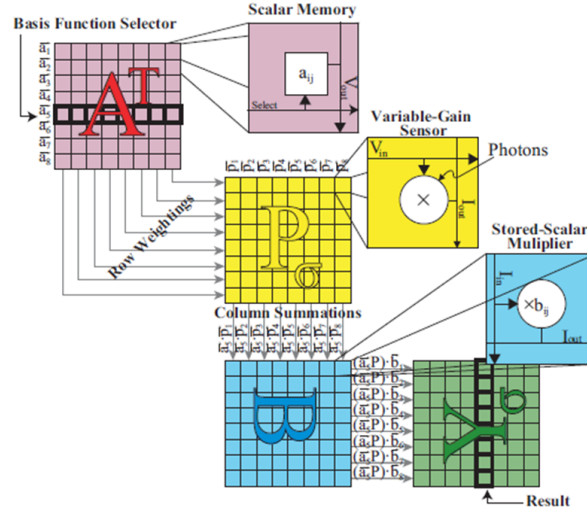


Figure 2.11 – Mathematical description of separable transform operation [82].

The inner products are computed in the analog domain in the focal plane level. As shown in figure 2.12, the two 1D-transforms are sequentially computed using operation at the array of pixels level for the first transform coefficients and at the end-of-column circuitry for the second. Block diagram of the CS imaging system which is presented in figure 2.12 presents the computational sensor element and how each pixel is connected with each other. In this implementation, only the compressed version of the image is readout, but requiring a lot of analog computations. Unfortunately, the structure of the pixel which is close to a 1T pixel is relatively poor in terms of noise. Even if this drawback can

be compensated in some sense with intensive post-processing performing noise removal, an other negative point is that the silicon space dedicated to processing operation is considerable compared to the entire CIS (cf. 2.12 c)).

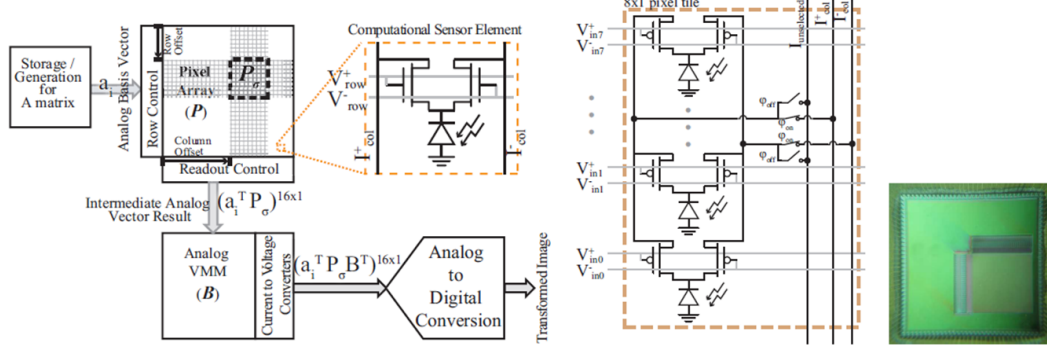


Figure 2.12 – Top view and pixel schematic of the separable transform image sensor [82].

The image transformation performed by the described imager architecture is operated on separate sub-regions (8×8 pixel blocks). More recently, [83] proposes an architecture based on the same principle but using 16×16 pixel blocks. Several other implementations have been proposed in the literature to perform specific transform before A/D conversion [51] [84] [85] [86]. Most of these implementations suffers from considerable requirements either on silicon area and/or acquisition time.

Random convolution

Previously presented architectures do not exhibit a CS dedicated structure but rather a more generic structure. It means that they can be used for several other potential applications (transforms that are not incoherent with the sparsity basis for example). This generic point of view has its own pros and cons because it complexifies the architecture even if all possible configurations are not used in practice. On the other hand, the method of applying random coefficients to the pixels represents the major difficulty. In [87], J. Romberg proposed a new measurement strategy called random convolution. As theoretically explained in [87], the M measurements correspond to M picked values at random positions in the convolution of an image by a random filter. In particular, this filter can be defined as a Rademacher sequence of ± 1 (which is an easy way to apply random convolution).

CIS architectures based on this sensing scheme are presented in [89] and [88]. Figure 2.13 presents a top-level view of the proposed architecture of [88]. As can be seen in this figure, this architecture requires an array of pixels composed of 3T pixels and local memories (i.e. figure 2.14). Those in-pixel memories and digital adders are used to perform the CS measurements. In practice, for this architecture, in-pixel memories are necessary to keep the ± 1 coefficients inside the pixels in order to control it.

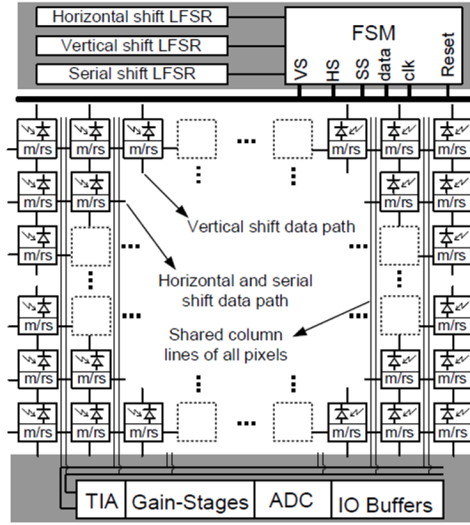


Figure 2.13 – Top-level functional view of the random convolution based CS imager [88].

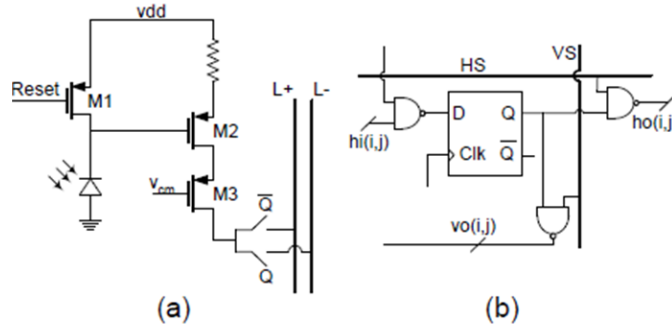


Figure 2.14 – Pixel schematic (a) and local processing (b) used in the proposed design of [88].

The main interesting point regarding this kind of architecture is that the sensing method facilitates the reconstruction. Indeed, the random convolution in the Fourier domain corresponds to a pointwise multiplication which is dramatically faster than a full matrix-to-vector multiplication. Thanks to the topology choices, this architecture also allows to change the measurement pattern in only few clock ticks. At the expense of complexifying the on-chip circuitry this architecture provides well suited measurements for the efficiency of the reconstruction in terms of algorithmic complexity. The major drawback resides on the low fill factor induced by adding a local memory inside the pixel. In addition, the FPN which is not discussed in the paper is probably important due to the impact of using a complex read-out data path inside the array of pixels.

However, from the algorithmic point of view, random convolution acquisition has been used in several works such as light field CS using camera arrays [90] and spectral imaging [91] because of its particularly interesting properties in the Fourier domain.

Random pixel selection

The common way to produce CS measurements is to project the signal using random or pseudo random vectors. As we have seen, this operation is maybe the most difficult part of the work. The authors of [92] propose to rather deal with randomly selected pixels as CS measurements which reduces the implementation complexity. Since the number of sensed pixels is reduced compared to a full scan, it performs a compressive acquisition. The reconstruction thus consists of an inpainting problem which is less computational demanding than a classical CS reconstruction. The proposed architecture is dedicated to this task, random pixel selection and sampling. The overall sensor array is divided into 4×4 pixel blocks whereby only one randomly selected pixel is sampled. Figure 2.15 shows the principle of this sensing scheme where grey pixels correspond to selected and converted pixels $X_{i,j}$.

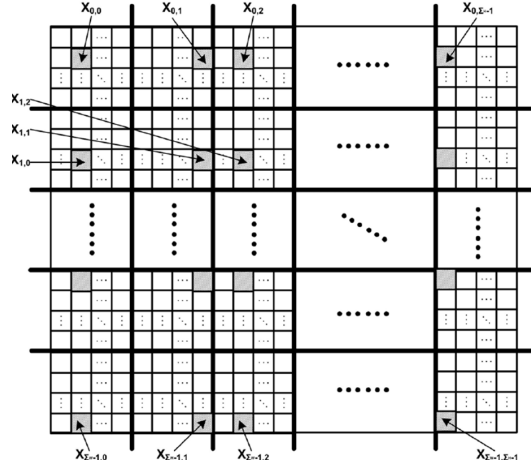


Figure 2.15 – An array of pixels in which only randomly selected pixels are read and converted [92].

A major issue of this architecture –which is not clearly expressed in [92]– is that since the pixels are acquired, the classical CS reconstruction can be replaced by alternative and faster reconstruction methods based on inpainting.

Compressive $\Sigma\Delta$

In the paper [93], the authors describe a CIS in which CS is performed at the A/D conversion level reusing the concept of CS- $\Sigma\Delta$ [94]. The compression and the conversion are both performed at the same time using CS-dedicated ADCs based on a first order $\Sigma\Delta$ architecture. The first order incremental $\Sigma\Delta$ architecture is used because it has the advantage of combining the both operations of averaging and converting an input composed of successive scrambled pixel values. This sensor architecture addresses in

particular the limitation on the sequential nature of the measurements. The proposed architecture has also the advantage of using low noise pixels known as 4T pixels. Figure 2.16 presents the main functional blocks of the proposed architecture which are the array of pixels, the CS mux (performing the pseudo-random successive selections) and column-parallel $\Sigma\Delta$ s. This CIS takes advantage of a CS-dedicated end-of-column circuitry without changing the overall topology of a traditional CIS which does thus neither impact on temporal noise issues nor on the fixed pattern noise.

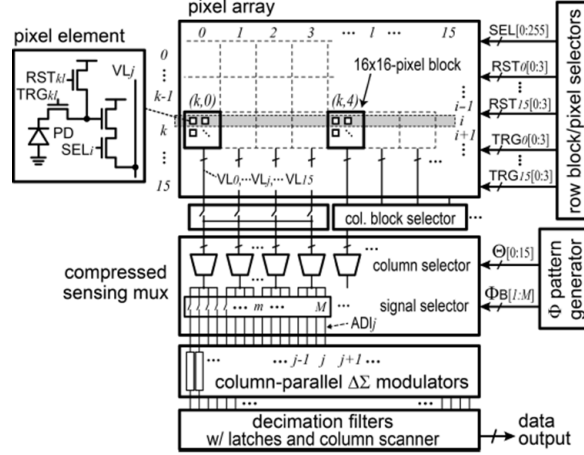


Figure 2.16 – Top-level functional view of the Sigma Delta based CS imager [93].

The proposed CS measurements are performed over pixel blocks of size 16×16 by the so-called CS-MUX which is a multiplexer with a higher frequency than the conversion frequency. Using pseudo random codes, this CS-MUX scrambles the successive inputs of a first order $\Sigma\Delta$ ADC. This way the conversion directly and intrinsically performs the summation of the inputs during a conversion time. In addition, the compression ratio and the relative frame rate can be tuned accordingly if required.

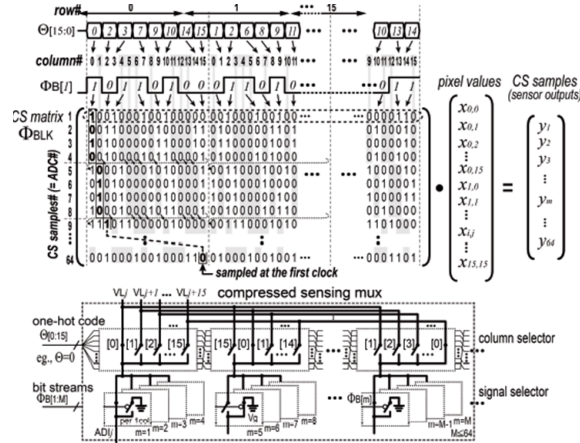


Figure 2.17 – CS-MUX proposed in [93] to perform CS measurements on each supports (blocks of pixels).

The main advantage of this structure is based on the fact that the CS-MUX structure allows to perform in parallel, as many CS measurement/conversions as the number of $\Sigma\Delta$ s integrated in a column of pixels. Figure 2.17 depicts the way the CS-MUX works to ordering the successive inputs of the $\Sigma\Delta$ s.

Capacitive measurements

Two recent implementations of CS-CIS propose to use pixel local capacitances to perform the measurements. This promising way to perform CS can lead to facilitate the design of future CS-CIS. [95] and [96] propose to use Capacitive TransImpedance Amplifier (CTIA) to perform the summation in the CS measurements. Different topologies are proposed, with the advantage of not modifying the 3T pixel structure. They use a combination (summation) of randomly selected pixels at the column based circuitry by a CTIA. The 3T pixel is slightly differently connected from the common way for this purpose to use a local block capacitance. Unfortunately, this architecture is not scalable at all for multiple reasons. First, it requires AA each pixel paths of pixel blocks. They propose to use 4×4 pixel block which means that 16 read out paths is necessary for a column of blocks. Increase the block size would mean to increase the number of paths accordingly which is not affordable because of the impact on the fill factor and the number of available metal layers. In addition, the impact of parasitic capacitance of the wires referring to coupled capacitance and grounded capacitance have not been properly studied. Notice that those problems can arise when larger resolution are involved. This will lead to increase dramatically the required capacitance (and thus size) at the end-of-column circuitry and inside pixel blocks. At last, they do not evaluate the impact of non linearities introduced by the pixel behavior during the integration time.

[97] and [98] propose to use a local capacitor inside the pixel and two separate column lines for column parallel pixel outputs. This architecture use external Linear Feedback Shift Registers (LFSRs) to generate measurement vectors following a ± 1 Rademacher distribution. Instead of applying pseudo-random sequences either on the whole array of pixels or on block of pixels, those sequences are applied in a parallel fashion to each column of pixels at the same time (see figure 2.18 a)). It theoretically limits the interest of CS but dramatically facilitates and improves possible implementations of it. In this case each measurement is performed at the end-of-column circuitry thanks to charges transfer using a Comparator-Based Switched-Capacitor circuitry. This parallelization of the operations allow simultaneous activations of the pixels. However for each set of column measurements a control signal needs to be propagated along the imager to reload the additional capacitor inside each pixel. This implementation choice is particularly appropriate for low power purposes.

However, this proposed architecture have several drawbacks. In particular the complexity inside the pixel is noticeable, thus reducing the fill factor accordingly. Indeed, 8 supple-

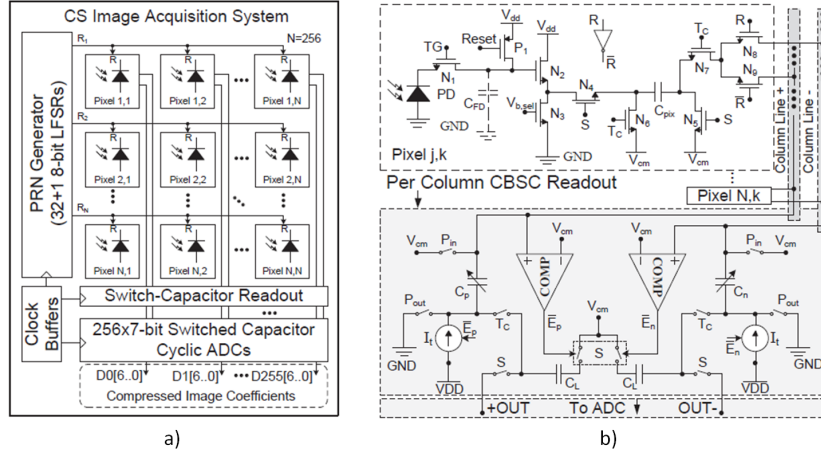


Figure 2.18 – The proposed low power column-based CS image sensor. a) Top-level functional view of row binary measurement based CS image sensor [98]. b) Schematic view of the pixel and the readout circuitry of [98].

mentary transistors are included in the pixel with an additional capacitor (see figure 2.18 b)). In addition, this pixel design requires 8 connections: 6 command wires, a power supply, a ground and the two column lines. Non overlapping clock have to be carefully provided to each pixel to avoid possible short circuit. At last, all the end-of-column circuitry should fit within the column pitch which can be difficult depending on the capacitors they use.

CS-CIS comparison summary

As it has been presented, several architectures of CIS performing CS have emerged during the last five years. Each of those presented ones have their own pros and cons. Major limitations still reside on the complexity inside the pixel (reducing the fill factor and impacting on the final resolution at a given silicon area) and the parallelization of the measurements (requiring a parallel and scalable topology). The table 2.2 reports the main features of the early presented CS-CIS. The architecture presented in [93] exhibits the relatively best features. The main advantage of this architecture is that it slightly differ from traditional 4T-pixel CIS with a by-column first order $\Sigma\Delta$ ADC. In terms of power consumption, the architecture presented in [98] clearly exhibits the best results because it only uses small charge transfers to operate CS measurements.

2.2.2 Reconstruction, signal recovery and information interpretation

As mentioned in [27], Compressive Sensing can be considered a “science 2.0” byproduct. Indeed, one can notice the exponential evolution of the number of articles which takes advantage of CS results since 2006. In particular, thanks to reproducible research and the

2.2. Hardware CS implementations and field of applications

Reference	[82]	[88]	[93]	[98]	[96]
Year	2010	2010	2013	2013	2014
Pixel type	PPS (2T)	APS (3T)	APS (4T)	APS (4T)	APS (3T)
Pixel size	$8\mu m \times 8\mu m$	large	$5.5\mu m \times 5.5\mu m$	large	$20\mu m \times 20\mu m$
Tech. node (μm)	0.35	0.18	0.15	0.18	0.13
Array size	256×256	256×256	256×256	256×256	16×16
Frame rate (fps)	-	60	1920	8000	-
Additional components	none	≈ 30 transistors	none	8 transistors and 1 capacitor	Connections (wires)
Fill factor	high	very low	moderate	very low	low
Compression ratio	as wanted	3.3	16	4	moderate
ADC	-	11b	12b	7b	-
Pwr $pJ/pixel$	-	19505	765	50	-
Special feature	Transform calculation	Other convolution binary kernels	Programmable compression ratio	Acquisition speed fast reconstruction	Time consistency
Technique	S.Transform by block	Random convol	CS-MUX	Column weighted sum	Block summation
Off-array complexity	very high	high	moderate	moderate	moderate

Table 2.2 – Comparison between the main CS-CIS developed in the literature.

exponential growth of scientific exchanges the number of proposed applications using CS concepts has dramatically exploded. Indeed, CS breaks the conventional way to tackle the problem of sensing and compression. Even if CS remains difficult to be implemented in real-world sensing devices, the related reconstruction problem - which is a smart mixture of signal processing and algorithmic techniques - has been deeply studied. As it has already been presented, a large number of algorithms can be used to handle a classical CS minimization problem (equations 2.12 to 2.16). Only regarding imaging applications the number of alternatives to reconstruct a signal from CS measurements is considerable. One of the most used and basic alternative to the l_1 minimization in the sparsity basis as prior operator $Op(\cdot)$ (i.e. equation 2.16) is the Total Variation (TV) minimization (i.e. equation 2.21). There exists two definitions of the TV for images, an isotropic version (equation 3.7) and an anisotropic version (equation 2.23). Particularly, TV minimization reconstruction is based on the assumption that neighboring pixels are close. If we denote U an image in its matrix form, u is its vectorized form. In particular, the gradient operator ∇ which can be defined on a matrix (here an image U) can be of different orientation. Basic TV operates using horizontal and vertical gradients respectively denoted ∇_h and ∇_v . We thus have $(\nabla_h U)_{i,j} = U_{i,j+1} - U_{i,j}$ and $(\nabla_v U)_{i,j} = U_{i,j} - U_{i+1,j}$. Notice that border effects are not detailed here. For the sake of simplicity in terms of notations, the vectorization inside the l_1 norm is implicit.

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} TV(\mathbf{u}) + \lambda \|\mathbf{y} - \Phi \mathbf{u}\|_2 \quad (2.21)$$

$$\text{TV}_{\text{aniso}}(U) = \|\nabla_h U\|_1 + \|\nabla_v U\|_1 \quad (2.22)$$

$$\text{TV}_{\text{iso}}(U) = \sum_{i,j} \sqrt{(\nabla_h U)_{i,j}^2 + (\nabla_v U)_{i,j}^2} \quad (2.23)$$

A direct extension of the l_1 minimization in a certain sparsity basis was the use of multiple sparsity basis at the same time for sparsity averaging. Instead of using a unique wavelet transform, [99] employs a multi-wavelet transform approach. The article [100] even presents an extension where the sparsity basis is replaced by a redundant dictionary. Indeed, this point view considers that the signal structure can be better described using a redundant dictionary than using an unique orthonormal transform. The article [101] shows that sparsity basis can be chosen according to the type of region where it is applied, this method is called spatially adaptive sparsity. This paper experimentally demonstrates that it achieves improved image quality when the couple of sparsity basis is *DCT/DWT*. On the other hand, [102] presents a collaborative sparsity approach, which enforces multiple sparsity simultaneously in an adaptive hybrid space-transform domain. An additional notion of sparsity has also been introduced called the group/structured-sparsity which assumes that sparse coefficients exhibit a certain structure. For instance, [103] proposes to impose a group-sparsity constraint on the CS analysis prior formulation with a redundant transform. Different works also try to improve the *TV* regularization using additional assumptions. For instance [104] shows a modified *TV* algorithm that removes typical *TV* artifacts exploiting *DCT* and contourlet transforms properties as additional priors. More straightforwardly [56] uses a joint *TV-DWT* l_1 minimization which show interesting results and represents an easy combination of *TV* and l_1 typical reconstruction methods. Recently, [105] presents a reweighted *TV* joint with non local sparsity; this technique enforces a sharper *TV* regularization thanks to the reweighting operation.

Most conventional CS recovery approaches take advantage of a l_1 minimization algorithm but using different priors and projections under l_1 norms. Few alternative techniques have also been developed in the literature. As example, [106] proposes a genetic algorithm for solving the reconstruction problem, [107] uses local geometric features as CS priors and finally [108] exploits a low rank regularization which is applied non locally. In particular, non-local patch correlation and local piecewise smoothness are used as priors in [109]. Even if providing better results, most of those techniques still suffer from a sever computational cost. It thus leads to the idea that the reconstruction method chosen for a specific application is necessarily a trade-off between the computational and memory loads and the quality of reconstructed signals rendering.

As expressed before, several high level applications, based on the use of CS measurements as inputs instead of a standard sampled signal, have also been developed in the literature. Rather than performing the original signal reconstruction, those applications directly interpret the information related to the compressed measurements. Thanks to the intrinsic nature of CS, the information contained in the CS measurements is already in a compressed form (lower dimension) meaning that for large scale problems the size of the signals that the algorithm have to deal with is smaller. This dimensionality reduction property is especially exploited in [110] for Compressive-Projection Principal Component Analysis whose motivations is to design a system involving a very light encoder/sensor. Article [111] first studied signal classification using CS measurements as input, later [112] proposed a CS-Mahalanobis semi-supervised classifier. Recently, in [113] and [114], texture classification algorithm using block-based CS where learning and classification operations are carried out directly in a compressed domain. [20] also takes advantage of block based CS measurements for feature extraction using in particular the information contained in spatial histograms. Regarding computer vision analysis, [115] presents a new way to perform motion estimation from highly compressively sensed images. [116] studies time adaptive CS whose final presented applications are background subtraction and tracking. Specifically, a technique for background subtraction using CS is detailed in [117]. Regarding tracking, the article [23] proposes a new algorithm called Particle Filtered Modified Compressive Sensing (PaFiMoCS) which is designed for tracking signal sequences. One can notice that video surveillance system is a major field of application which is targeted by CS [118] [119] [74]. Some applications such as face recognition have also been investigated in [120] and later in [121]. Most of or even every of those papers only propose algorithms without considering the implementation of the sensing scheme on a sensor. Yet, at the end, a system that would employ CS should be designed taking into account the constraints at both the coder (sensors and data management) and the decoder (reconstruction or advanced application).

2.3 Limitations and current development

A wide variety of CS-based image acquisition systems have been proposed in the literature (see figure 2.19). The only point in common between them is that they all take advantage of the core principle of CS which is spreading the information as most as possible and then take sparse samples. In particular, regarding image sensors, CS relaxes the constraints on the main bottleneck which is the A/D conversion and that limits both the frame rate and the resolution. Despite the large number of publications on this fashionable topic, no killer application neither regarding the acquisition nor the reconstruction has been clearly defined yet. But potentialities are almost infinite if some sticking points are broken down. Indeed, the acquisition complexity and or its calibration remains the main obstacle for the sensing stage whatever if performed using optical devices or using analog computations performed in the focal plane of a CIS. Moreover, the community of image

sensor designers is hardly ready to accept the CS paradigm because it can question the way to tackle the noise and the technological dispersion issues. In the sense that, for the moment, the development of such an architecture implies that common optimizations of the architecture design need to be put in question. Unfortunately, it can cause real damages on the real abilities and characteristics of the image sensor at the end in terms of maximum quality of the image it can reach for a fixed number of pixels. However, CS have to be considered as an appropriate solution for defining compromises. Indeed, used as an alternative working mode without modifying the entire sensor architecture, CS could bring an opportunity for improving sensor specific performances. For instance, multiple CS alternative working modes could be used in a single sensor to address each a targeted application: compressed still imaging, high frame rate imaging, transient detection, high dynamic range imaging.

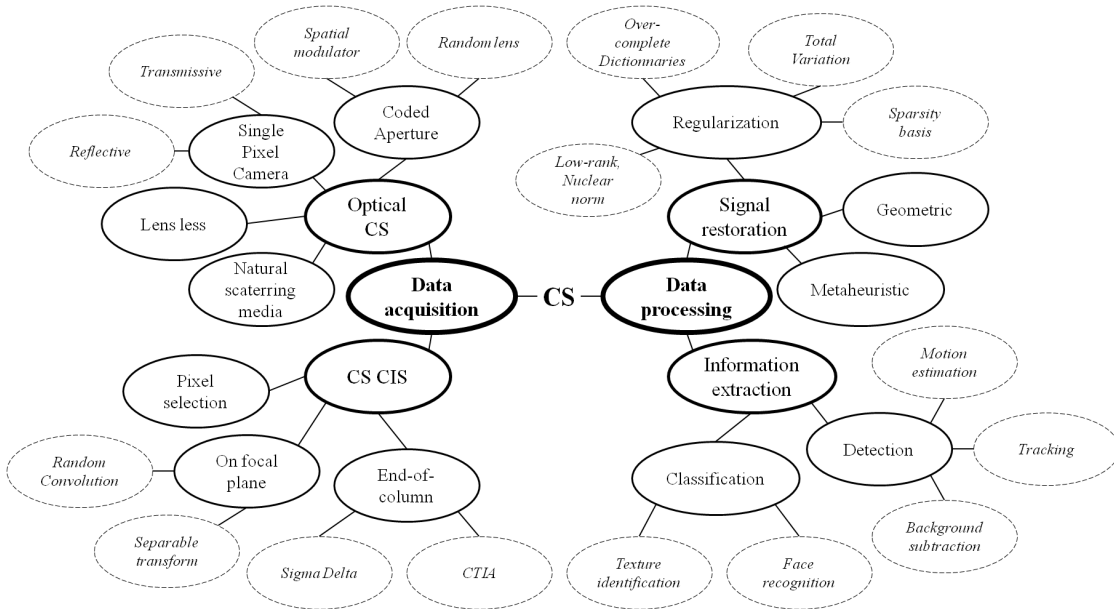


Figure 2.19 – Summary of the main CS imaging techniques at the encoder and the decoder parts.

From this study of the state of the art, we conclude that the need of a joint design of a CS system and its dedicated algorithms is a crucial point. Indeed, nowadays and with the rise of CS systems the acquisition and the reconstruction should no more be considered as separate problems. Different ways of defining proper trades off have been studied leading to three separate focal areas. First, the design of solid state CS image sensor have to be considered as modular as possible (in CS sense) when relaxing at most the on-chip complexity. Secondly, we state that adding simple embedded processing inside a CS imager can solve issues, severely improving the overall efficiency of a CS based imaging system composed of a specific imager and a fast reconstruction. Finally, we tried to find out an alternative solution for relaxing classical constraints in terms of congestion of an optical CS based system by using a single optical element (for the moment without real implementation) and an traditional imager used for parallelizing the measurements.

3 Smart Automaton Compressive Sensing Imager

One of the main issue of Compressive Sensing –which is generally evoked in the literature– is the sensor power consumption savings thanks to the reduced number of A/D conversions. In addition to this, since the data are output from the sensor in their CS format, common compression stages are no more required, which implies considerable power and silicon area savings too. It thus means that the data can be directly stored, processed or transmitted right after the acquisition. In [93], the implementation of CS helps to dramatically increase the frame rate and the compression ratio accordingly. Indeed, the only limitation is the data throughput of the image sensor chip outputs. If the compression ratio is doubled therefore the frame rate can be increased accordingly by a factor of two as well. In particular, the Compressive Sensing paradigm facilitates the design of high speed imaging devices by reducing the requirements on bandwidth of A/D Converters. Moreover, video sequences are sparser than still images thanks to temporal correlations of the pixels and time redundancy of the information. Traditional video encoding schemes already take advantage of intensive processing in the temporal axis (even in some cases multi-pass processing). In addition, some applications in industrial imaging or machine vision require only few information on the state of image frame components but at a very high frame rate. For instance, the goal of such an application can be of finding a transient event in a video sequence but with a very precise time index. On the other hand, one can notice that the traditional readout mode (Rolling Shutter) introduces the skew effect and a high sensitivity to the vibrations because of the lack of time consistency between pixels of the same column. Rolling Shutter can be replaced by a Global Shutter at the expense of supplementary components inside the pixel [122] [123]. Those issues could be handled by specific sensing scheme if based on CS, which therefore improves time consistency of the measurements. From this point of view, regarding pixel integration drawbacks, linear image sensors with a traditional pixel structure often suffer from dynamic range limitations. High Dynamic Range imaging represents a set of techniques used either to design specific sensor well suited for HDR acquisition at the expense of complex design and power consumption [124] or to extend the dynamic using post processing operations [125].

Therefore, the number of intrinsic limitations due to the image sensor structure is substantial. This chapter mainly deals with the CS implementation in CIS for relaxing those typical on-chip constraints (i.e. power consumption, frame rate and dynamic range). In the case of CIS, the implementation of CS has been investigated to address applications that are commonly limited by classical architectures and design constraints. In previous works, CS architectures exhibit features and performances referring as a breakthrough technology. The hope is that exploring many directions in terms of applications would be prolific to define which of those are well suited to incorporate the CS framework. Indeed, it has been determined that the CS paradigm should not be confined to simple compression purposes only.

This chapter is organized as follow: first, the efficiency of different measurement schemes is evaluated in terms of both reconstruction and on-chip implementation complexity, secondly a CS-CIS architecture named SACSI (Smart Automaton Compressive Sensing Imager) is presented in detail with the motivations related to each component blocks. Then, applications which are addressed by this architecture are depicted with their dedicated reconstruction algorithms. Lastly, the acquisition non-linearity issues and unknown parametrization of the pixel behavior are discussed with related calibration requirements.

3.1 Preliminary study

Since a CS system combines the acquisition, the conversion and the compression in a single stage implying specific priors on the reconstruction, it appears that a joint design between the coding and the decoding operations are necessary to ensure that the system is properly working. Trade-offs between the theoretical efficiency of an entire system and its complexity are generally required. Indeed, a compromise is always taken when defining where to place the complexity cursor, rather on the coder side or on the decoder side. In particular, what one calls block-based CS is used to implement CS sensors allowing multiple relaxation on the overall system.

3.1.1 Type of measurements, supports, and expectancy reduction

As previously presented, different random sequences (measurement vectors) concatenated into a sensing matrix Φ can be used to perform compressed measurements potentially constituting a universal sensing matrix. We present here several possible orientations for developing an appropriate sensing scheme in the perspective of designing a CS-CIS. We can play on different characteristics to define a suitable and functional sensing method. First, four main types of sensing matrix constructions constitute well known acquisition methods for compressive sensing [27]. Those ones are known as Random Gaussian sequences, Binary $\{-1, 1\}$ Rademacher sequences, any discrete Bernoulli distribution

based sequences and random convolution. Secondly, due to the reconstruction limitations and for achieving parallelization of the measurement during the sensing stage, this study is limited to what we call “parallelized CS”. In particular, regarding images, two different sensing schemes are feasible to design a parallelized CS which are block-based CS and column-based CS. This “parallelized” term means that the same set of measurement vectors is applied at the same time –or not– to each image block or image column respectively. This parallelization also operates simplifications on needed calculations performed by the reconstruction algorithm. For instance, if we denote Φ_R a measurement matrix applied to each column of an image X thus the sensing operation corresponds to $\mathbf{y} = \Phi \mathbf{x}$ where \mathbf{x} is the unfolded version of X and Φ is a diagonal concatenation of multiple Φ_R . This notation is equivalent to $Y = \Phi_R X$ (see figure 3.1).

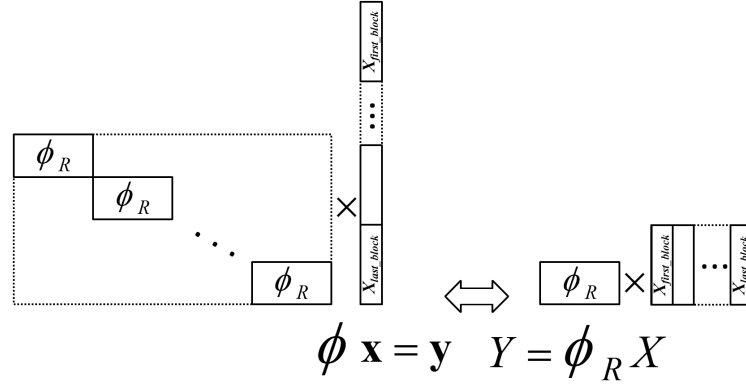


Figure 3.1 – Equivalence between matrix to vector multiplication and matrix to matrix multiplication in the case of column-based CS.

Two examples of parallelization of the measurements in the case of CS imaging is given by figure 3.2. In both cases (block-based CS or column-based CS), this figure depicts Bernoulli distributed measurement vectors.

This consideration is crucial when large image sizes or video sequences are involved. Indeed, if a full $\Phi \in \mathbb{R}^{M \times N}$ matrix were used, the memory requirements would be considerably larger so as the time of reconstruction. For instance the CS matrix Φ for an image of size $N = 512 \times 512$ and a compression ratio of $1/8$ would be composed of more than 8 Giga elements ($M \times N$). In the case of random convolution, this problem does not exist because of the intrinsic nature of the sensing scheme. The fact is that random convolution has a formal definition for being performed in the Fourier domain which implies almost no requirement in terms of memory and computations. But some difficulties arise concerning the implementation on a chip. In [88], random convolution has been implemented at the expense of the need of sequentiality and of a major growing complexity of the pixel. As proposed in the literature [27] the CS matrix can also be build using a Bernoulli Binary distribution of $\{0, 1\}$ with any success probability p ($p = \mathbb{P}(\Phi_{i,j} = 1)$). This property is particularly interesting when the expectancy p is close to zero. First,

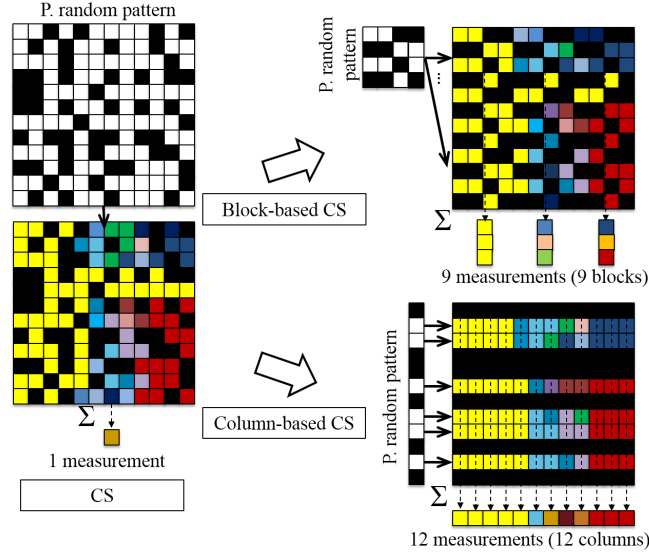


Figure 3.2 – Example of masks used in the case of Bernoulli block-based CS or Bernoulli column-based CS.

it can sparsify Φ which could speed up the reconstruction, secondly it could facilitate the on-chip implementation. However, if the random sequence is generated following a Bernoulli law with a very small p , the maximum correlation between measurement vectors dramatically increases. In addition, it increases the probability of generating a null vector. Considering it, the effect in practice of reducing p has been evaluated for both block-based CS and column-based CS. The following figures present the reconstruction results in terms of PSNR for different compression ratios. The reconstruction consists of an isotropic TV for which a stopping criteria is common for all the simulations, iteration steps and regularization parameter have been tuned depending on the compression ratio and the level of sparsity of the sensing matrix. A set of 5 test images of size 512×512 (Lena, Pirate, Cameraman, Darkhair, Blonde) have been used to obtain averaged results. Figure 3.3 presents the results when column-based CS is employed. Figures 3.3 a) b) and c) respectively present the results when different random sequences are used, Bernoulli (with parameter $p/2$), Gaussian where coefficients are forced to zero with a ratio of $(1 - p)$ and Bernoulli (with parameter $1/2$) where only regularly spaced coefficients are preserved and the other are forced to zero (the distance between consecutive non zeros value is at least $1/p$). This last technique that is employed to reduce the number of considered pixels for each measurement will be called Downsampling.

In the case of column-based CS, we can observe that the values p has an impact only when it is under $1/32$ for any compression ratio and distribution used. It means that for natural images the number of pixels taken into account for each measurement can be efficiently reduced. In addition, when p is relatively high, the reconstruction is more complex because of the density of the sensing matrix, involving worse reconstruction results because of using

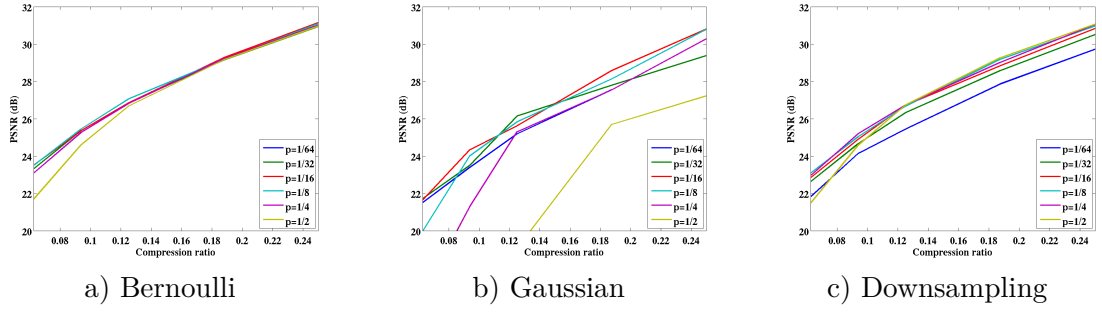


Figure 3.3 – Impact of using a sparser sensing matrix in the case of column-based CS. The expectancy of non zero values in measurement vectors linearly depends on the p value.

the same parameters related to stopping criteria (e.g. number of iterations, threshold on relative objective function...). Concerning the different distributions, we can observe that either using Bernoulli or Downsampling, the results are almost identical. The PSNR of reconstructions using Gaussian distribution are a bit lower than the two other, particularly when the compression ratio is under 0.1. Figure 3.4 presents the same results as figure 3.3 but when a block-based CS scheme is employed. Image blocks we employed are of size 32×32 pixels.

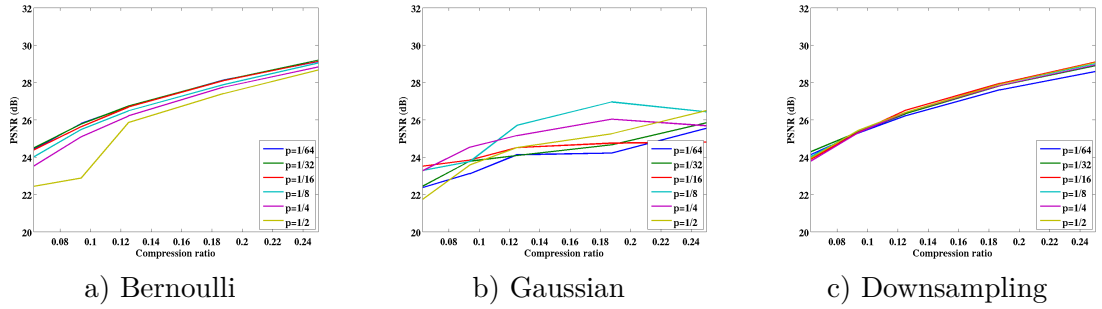


Figure 3.4 – Impact of using a sparser sensing matrix in the case of block-based CS. The expectancy of non zero values in measurement vectors linearly depends on the p value.

Compared to column-based CS, block-based CS works worse but for low compression ratio under 0.1. This effect is due to two separate antagonist reasons. First, performing CS measurements over blocks is less efficient than over columns because of the fact that the disparity of the pixels belonging to a single column is higher than the disparity in a single block. However the block size we used is two times higher than the size of a column so block-based CS should works better, in particular when compression ratio goes down. Figure 3.5 shows the results in terms of reconstruction PSNR of the different proposed sensing schemes for $p = 1/2$ a), $p = 1/16$ b) and $p = 1/64$ c). From these results we can conclude that the better configuration is when column-based CS is used with a Bernoulli distribution.

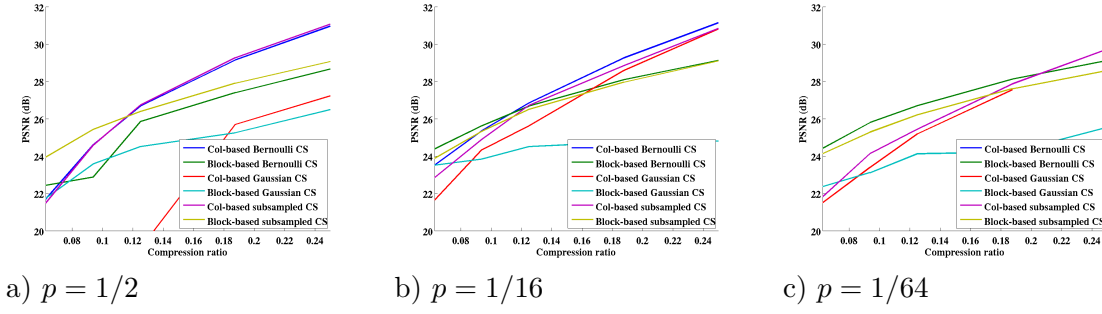


Figure 3.5 – Impact of using a specific sensing scheme for three different p value.

Notice that every consideration depicted here are valid only for 512×512 natural images. Indeed those results could depend on the size of the image but also on the reconstruction method. For this preliminary study, the main motivation to use TV regularization is that it is the most commonly used CS reconstruction assuming that the image gradient is sparse (which is not a strong prior for natural images).

3.1.2 Possible implementations and addressed applications

Considering the previously presented results, the measurement scheme that have been retained is the column-based one, using a Bernoulli distribution. Indeed, it seems to be the most appropriate for being implemented in an image sensor chip for multiple reasons. First, it is relatively easy to generate pseudo-random binary values with a Bernoulli distribution ($p = 1/2$). When a Bernoulli distribution is used instead of a Rademacher, no subtraction operation is required and a simple switch can be used to either activate or not a pixel for a specific measurement. Using this sensing scheme we also can hope to integrate the operators needed for CS without changing the traditional imager topology, with a column read out circuitry. In addition, keeping this topology means that the image sensor architecture could be easily scalable. The main idea is to find a way to design an architecture that can work either in CS mode or in traditional rolling shutter uncompressed mode. Without changing the traditional structure, the architecture would address two main limitations which are the time consistency of the measurements and the sequentiality of the pixel activations. This way, problems related to rolling shutter are solved and frame rate limitations are reduced thanks to the frequency of column measurements in CS mode that could be equal to the frequency of a row readout in full rolling shutter mode. It thus implies a direct link between the compression ratio and the frame rate. Indeed, in this specific CS mode, an image frame can be acquired $1/CR$ times faster with a compression ratio of CR . Finally, if the measure of a pixel intensity is not destructive we can imagine multiple applications implying specific sequencing in the sensing scheme. In particular, a dynamic range extension application will be presented.

3.2 CS image sensor architecture and design (SACSI)

Under the constraint of keeping the traditional 3T pixel structure and taking into account the considerations stated before, different architectures can be imagined. The design of such an imager first depends on the pixel working mode. One can distinguish 3 different working modes of a 3T pixel. It can either work in voltage mode (as commonly used) [10], in charge mode [126] [127] or in current mode [128].

In charge mode, several problems arise. The size of the column bus can not be too large which makes the architecture hardly scalable. Indeed if a Capacitance TransImpedance Amplifier (CTIA) is used at the end-of-column which is connected to each pixel of the column, the effect of the bus ground capacitance will be significant, otherwise the CTIA circuitry requires a capacitor with very large capacitance. To compensate this effect, an intra block capacitor is required at the interblock level as proposed in the architecture proposed in [96]. It thus affects the array of pixels topology, introducing a considerable FPN. Yet, the structure used in [98] based on Comparator-Based Switched-Capacitors (CBSC) could be exploited to ensure that all the charges are efficiently collected into a capacitor at the end-of-column circuitry. However if a transistor diffusion capacitance is directly used to carry the information corresponding to the illumination of the pixel, this capacitance is still too small compared to the column bus capacitance. In addition, since the capacitance information is read, a new integration is needed otherwise the corresponding time of integration will have changed. The solution proposed in [98] is to use a 4T pixel allowing to store the information at the floating diffusion point combined with an additional in-pixel capacitor which is reloaded and then read for each CS measurement.

In voltage mode, two variants are possible to perform the summations. The first variant is to make it sequentially by integrating successively activated pixels. The overly long time due to the sequential nature of this variant makes it impracticable. Moreover, it requires that each CS measurement corresponds to the activation of the same number of pixels to manage sequencing at the end-of-column circuitry for the ADCs. The second alternative is to use multiple paths combined with a summing amplifier circuit at the end-of-column circuitry. But it is not conceivable because it would induce a large number of column buses making this structure absolutely not scalable.

In current mode (as in [89]), the advantage is that the summation is already performed if pixels outputs are connected together thanks to Kirchhoff laws. Indeed, it does not require additional components compared to the traditional architecture. However, several practical issues have to be taken into account: the non-linearity, bus resistivity and power consumption and added noise.

The proposed architecture mainly addresses the reconfigurability and low power functioning issues. Indeed, the predominant goal here is to provide an architecture which is as

close as possible from a classical 3T APS imager in order to profit from the traditional optimizations when working in uncompressed mode. Moreover as will be presented, this architecture can be used for both high frame rate acquisition and dynamic range extension depending on the sequencing of the measurements. An overview of the proposed architecture is given in figure 3.6.

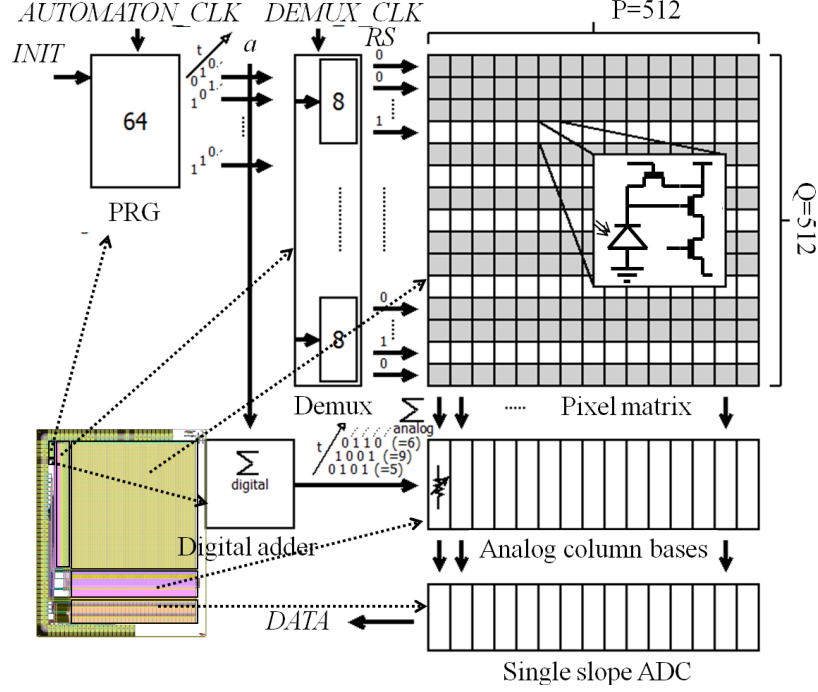


Figure 3.6 – Top-level view of the proposed architecture SACSI (Smart Automaton Compressive Sensing Imager).

The proposed architecture is composed of 6 main operating blocks: a Pseudo Random Generator (PRG), an Expectancy Reduction Layer (ERL), an array of 3T pixels (4 wired connexions), a digital adder, configurable end-of-column circuits and single ramp 10b-ADCs with output buses. When CS mode is activated, this architecture performs column-based CS using a Bernoulli-like distribution. The ratio of 1 binary values of each measurement vector is relatively low using downsampling combined with local rolling shutters and a small p value of the generated Bernoulli distribution. The structure of this image sensor has been chosen because of being very closed to a classical one. Therefore, it allows the imager to work either in CS mode or uncompressed mode allowing the reuse of common optimizations. To keep temporal consistency, in CS mode, a snapshot current driven readout has been retained. Two modes of acquisition can be selected, a classical uncompressed mode or a CS mode. Multiple time diagrams of the control signals can be employed to address specific CS acquisitions. Notice that the CS mode is also compliant with the 4T pixel architecture in the same fashion but for the sake of simplicity the proposed design uses classical 3T pixels. The use of 4T pixels would especially improve the uncompressed mode thanks to lower levels of noise and the possibility of a

CDS. Several approaches have been envisaged and studied. We have chosen a current functioning mode since it ensures a more robust functioning than a charge approach while alleviating design constraints on the end-of-column readout circuits (as expressed before). Reconfiguration is then controlled by applying different voltages on the source follower transistor of the classical 3T pixel design and on the readout circuitry. The CS mode involves measurements acquisition in a by column fashion which does not have a considerable impact on the final reconstruction quality. A vector of binary pseudo-random values is generated by a boolean cellular automaton with a chaotic behavior followed by an Expectancy Reduction Layer. We can notice that keeping the same 3T structure and using a current driven readout circuitry, only one measurement per column can be performed for each integrated image. With such an architecture we can not expect to have good results for still images (in particular because of photon shot noise and motion blurs). Indeed, this architecture is particularly well suited for video CS with a very high compression ratio. This circuit has been designed in *TOWERJAZZ* 0.18 μ CMOS Image Sensor process. The analog voltage supply is of 3.3V. The final layout of the chip is presented in figure 3.6. The chip is back from foundry and will be tested as soon as possible.

3.2.1 Pseudo-Random Generator (PRG)

In the case of CS, the PRG is one of the most important operating block in the image sensor. This operating block must generate predictable set of pseudo-random vectors (in our case binary vectors). Indeed, the decoder needs a computationally efficient routine to regenerate the sensing matrix for the reconstruction stage. Particularly, the interest of the generation of measurement vectors in the case of column-based CS is that it can easily be deported outside the array of pixels, replacing the common row selector. In figure 3.6 the PRG is positioned on the left side of the pixel matrix. To provide an architecture with a row based structure, the use of a cellular automaton with a chaotic behavior is used. When pixels are used in current mode the activation of half the pixels is not possible because it would induce a too large current ($> 100mA$ for a 512×512 pixel array and with $1\mu A$ per pixel). Therefore, a stage of AND logical gates (ERL) is thus used to activate far less than one over two rows of the entire matrix, to reduce the output current resulting of the current summation. The goal is to determine appropriate boolean functions to reduce the number of activated pixels by column and its variance of this number between each measurements.

Cellular Automaton

The structure of Linear Feedback Shift Registers (LFSRs) is generally used to generate a sequence of binary values following a pseudo random distribution [129]. LFSRs are well known for their properties such as length of cycles and level of correlation of successive

samples. Instead of using a LFSR structure, we propose to use a cellular automaton to generate sequences of pseudo-random binary values (as in [130]). A cellular automaton consists of multiple cells at a certain state at a certain time instant. At each discrete time instant t , each cell state a_i^t is updated using a transition function defining the general behavior of the cellular automaton. This transition function which is identical for each cell defines the transition from a certain state to an other between two successive time instants t and $t + 1$. Thanks to the topology of a single dimensional cellular automaton (see figure 3.7), it is easily scalable without modification. Moreover, in the case of column-based CS the parallelization of the generated values is a clear advantage for such a structure compared to a LFSR structure.

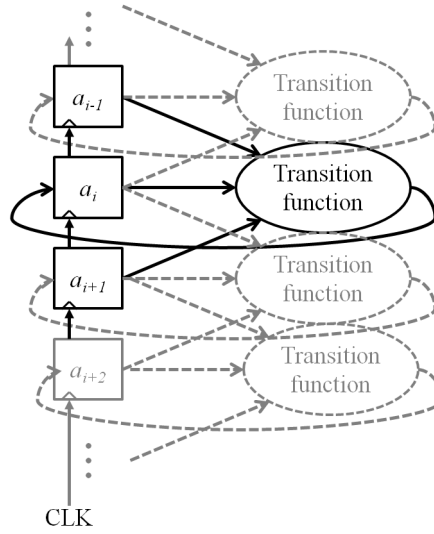


Figure 3.7 – Generic top view of a one-dimension cellular automaton whose the transition function only takes left and right neighbors outputs as inputs.

Following the Wolfram classification of transition rules [131], multiple rules can be used to implement a cellular automaton with a chaotic behavior. Due to implementation constraints, we focused our work on the Rule 30 (cf. 3.8) which exhibits proper chaotic properties. For instance, this is highly sensitive to the initial conditions meaning that two slightly different initial conditions would result in a considerable divergence of state patterns. This basic transition function only takes the 2 neighbor states of each cell to compute the next state which is also a considerable advantage regarding its implementation and layout wired connections compared to the use of independent LFRSs each connected to a single row.

To switch between CS mode and uncompressed rolling shutter mode, the transition function of the equation 3.1 can be simply replaced by the equation 3.2. This way, the operators and the latches used for the cellular automaton can be reused to form a Linear Shift Register for performing a row scan instead of multiple random rows activation. Indeed, this reconfigurability relaxes the requirements in terms of silicon area.

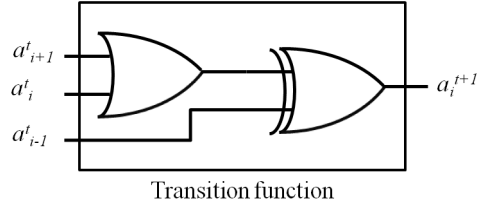


Figure 3.8 – Transition function used in the cellular automaton performing the Wolfram’s Rule 30.

$$a_i^{t+1} = (a_i^t + a_{i+1}^t) \oplus a_{i-1}^t \quad (3.1) \quad a_i^{t+1} = a_{i-1}^t \quad (3.2)$$

Expectancy Reduction Layer and local rolling shutter

This is relatively easy to design a cellular automaton providing pseudo-random binary values in a parallel fashion. Those values follow a distribution which is generally close to a Bernoulli one with $p = 1/2$ (minimizing the maximum correlation between generated binary vectors). As example, Rule 30 defined by the transition function stated by the equation 3.1 exhibits this property. Regarding the CS implementation we target, this property is not appropriate because it would imply that on average, half of the pixels would be activated at the same time. Adapting the expectancy of a binary sequence implies to generate a binary sequence following a Bernoulli distribution with $p \neq 1/2$. Actually, this can be generated from a sequence following a Bernoulli distribution with $p = 1/2$. In our case this stage is necessary to do not activate too many pixels at the same time, thus requiring an expectancy reduction. To reduce the expectancy of a random sequence, it only needs to combine successive samples using AND logical gates assuming that successive bits are uncorrelated. This way, and if successive samples are sufficiently independent, the expectancy is divided by a factor of two when using a single AND gate. The expectancy could be increased too, by the use of OR logical gates for instance which is not the point in our case. Figure 3.9 provides a generic overview of the proposed structure used to generate a pseudo-random binary sequence with a lowered expectancy. It is composed of a cellular automaton keeping in memory the two previous states and one dimensional array of buffer/inverter defining the initial state and an expectancy reduction layer.

However, a problem arises because in the practical case of a cellular automaton, the correlation between successive and neighbor states can not be neglected. It implies that the expectancy reduction is not operated as efficiently as theoretically compared to a proper Bernoulli distribution. Therefore, several combinations of AND logical gates connected to the output of the cellular automaton have been evaluated to fulfil each requirement (listed below).

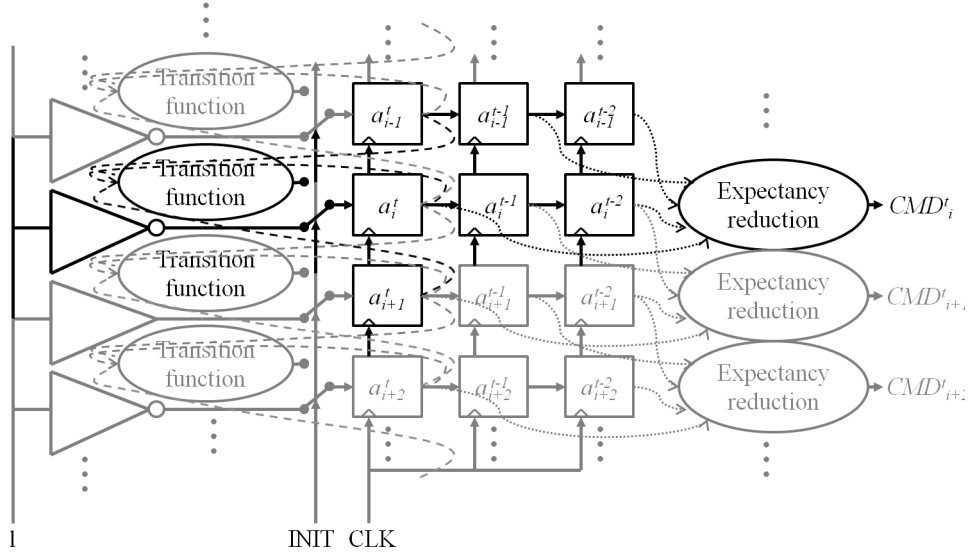


Figure 3.9 – Generic top view of a one-dimension cellular automaton followed by an expectancy reduction layer.

- The expectancy reduction around 1/8.
- The implementation complexity (connections between the cells).
- The number of AND logical gates.
- The correlation of the output sequences (in time and space).
- The low variability on the number of activated rows between each time instant.
- The low variability on the number of activations between each row.

The optimal function we empirically found is depicted in figure 3.10 and is only composed of 3 AND logical gates. This function has also the advantage of taking inputs close to each other which is particularly interesting regarding layout issues.

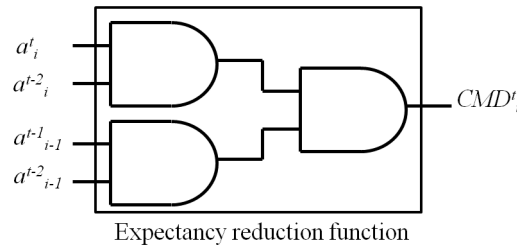


Figure 3.10 – Expectancy reduction function used in the proposed PRG.

The resulting equation 3.3 of this expectancy reduction function is simply replaced by the equation 3.4 if uncompressed mode is activated. This way, a same row selector block is used for both modes, only local reconfigurations are needed.

$$CMD_i^t = a_{i-1}^{t-2} \times a_{i-1}^{t-1} \times a_i^t \times a_i^{t-2} \quad (3.3) \quad CMD_i^t = a_i^t \quad (3.4)$$

Figure 3.11 presents the simulation results for an architecture composed of 64 distinct cells ($i \in \{1, \dots, 64\}$). In practice, the 64th cell is obviously connected to the first one, making the automaton a cyclic structure. For this simulation and the final implementation, a specific optimized initialization is used to fulfill the requirements expressed before and leading to a longer cycle (i.e. more than 10 million time steps). As shown in figure 3.11, the normalized correlation matrix is close to an identity matrix. Maximum correlation and average correlation between generated vectors (normalized scalar product) are respectively 0.75 and 0.01. In particular, values are almost identical for a computer generated sequence. The average number of activated rows equals the number of activations of each row. This number is approximately 6 (over 64 and thus 512) which is fitting with the fixed requirements. Moreover, the first generated vector (after expectancy reduction) containing no activation arrives at time 6274 which corresponds to more than 12 successive frame acquisition with a compression ratio of 1.

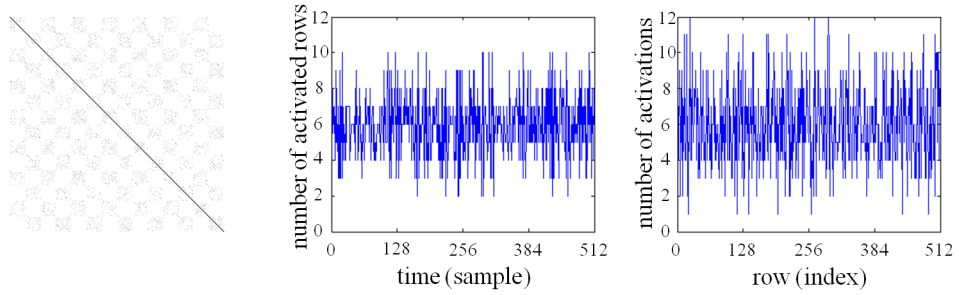


Figure 3.11 – Simulation of the proposed PRG structure (cf. figure 3.9) for a size of the cellular automaton of 64. From left to right, the correlation matrix between generated column CS measurements, number of activated rows for each time index and the number of activations for each row.

In addition to the Expectancy Reduction Layer, a demultiplexer is used to sequentially address one row out of eight. Each group of eight RS_i^t commands (i.e. row selection signals) is controlled by one of the 64 binary outputs of the pseudo-random generator CMD_i^t . This way, two active rows have at least seven non active rows between them. It reduces randomness of the sensing matrix but ensures to spread more potentially activated rows. This operating block which is called Demux in figure 3.6 can be controlled either using the same clock control as the PRG or a different one depending on the application to address.

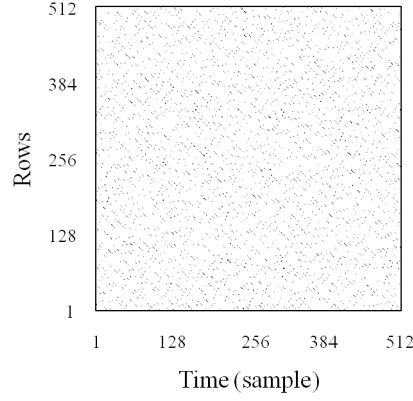


Figure 3.12 – Activation matrix representing the compressive sensing matrix in its reduced form (i.e. column-based CS). Points represent the indexes of activated rows in the function of time and their relative gray level represents the number of activated pixel a its time index.

The demultiplexer is also used to sequentially reset one row out of eight. This way, a local rolling shutter is performed when CS mode is activated. Indeed, reset rows are the ones that have been potentially be read just before. It leads to a reduction of the time consistency of all the measurements but allows to use an exposure time 8 times longer than if there were no rolling shutter at all and for the same measurement rate. Removing the local demultiplexers would imply either that each group of measurements is related to a specific time of exposure or that each group of column measurements is performed after a time of integration. In the case of a shared clock between the Demux and the PRG, the resulting RS_i^t are reported in figure 3.12.

Top-level simulations

The Cadence/Eldo simulation of the PRG provides the results presented in figure 3.14 for uncompressed mode and figure 3.13 for the proposed column-based CS mode. Figure 3.13 shows the activations of a_i in function of time when column-based CS is activated. As can be seen in this figure, multiple a_i can be activated a the same time and the chaotic nature of the PRG visually appears.

By contrast, figure 3.14 exhibits a simple rolling shutter reading scheme that can be employed when CS is off.

As will be presented in the following, the architecture requires a digital adder to normalize the output measurements according to the number of activated rows. Indeed, to efficiently perform the conversion of the measurements it is necessary to strain the output measurement level between two known levels. An operating block calculating how many

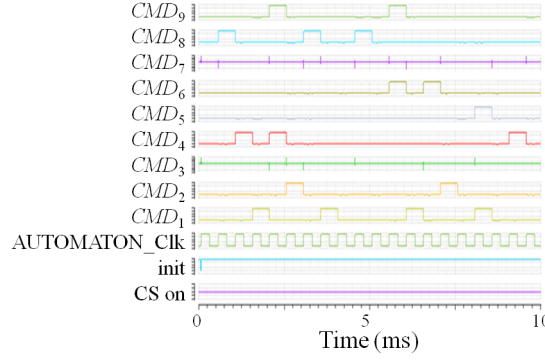


Figure 3.13 – Transient simulation of the CMD_i^t generation in column-based CS mode.

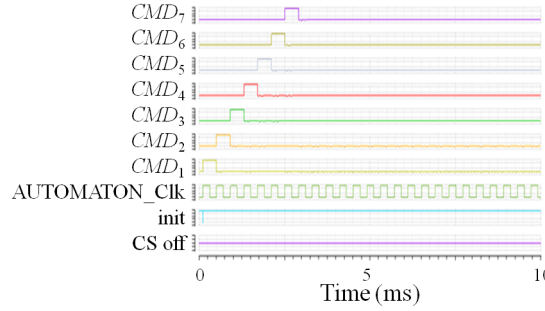


Figure 3.14 – Transient simulation of the CMD_i^t generation in uncompressed mode.

rows of pixels will be activated for each group of measurements is thus needed. This operating block has been designed to work sequentially because of layout constraints. The implementation of an adder on the 512 rows RS would not be an efficient way. Since we know that the maximum number of activated rows is 12, a 4-bit counter is enough to figure out the number of activated rows. This counter operates sequentially on the outputs CMD_i using a delayed token. Then, 4 cascaded full adder/D-latches blocks are used for the sequential counting. Figure 3.15 presents the structure of the digital adder employed for counting the number of activated rows. Moreover, even if being sequential, the frequency of the successive reading of PRG output is not limiting because of being in any case lower than the frequency used to evacuate the converted data at the output bus.

In the proposed architecture, the pixel commands for achieving pixels summation (i.e. CS measurements) are performed using the RS signals propagated through the lines buses inside the pixels array. When CS mode is activated the pixels must thus work in an alternative mode which address multiple requirements (i.e. a current readout, a small current output level and an acceptable technology dispersion).

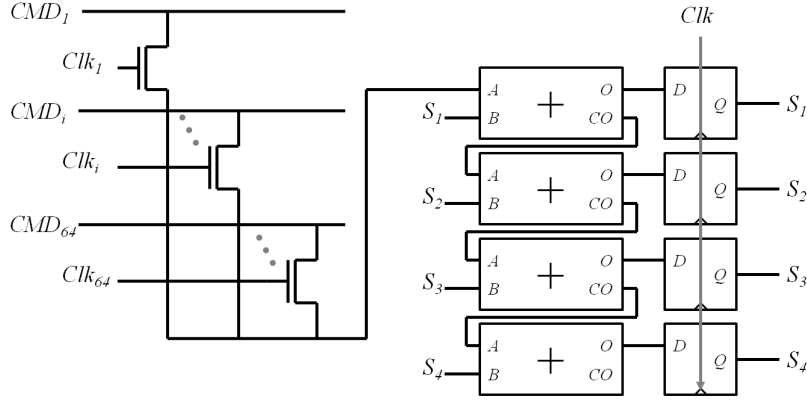


Figure 3.15 – Schematic of the digital adder used in CS mode for counting the number of activated rows.

3.2.2 Specific but common pixel design

To perform the summation operation needed to operate the measurements we propose to turn the pixel working mode into a current readout mode instead of a voltage readout mode. Indeed, if we take the example of the simplest APS (i.e. a 3T pixel structure), several works have presented CIS architecture involving pixels providing current output level as information representing the level of incoming light [132].

3T pixels but multiple readout modes

Figure 3.16 shows the two different operating modes of the pixel. In both cases the structure is the same, the modification only resides on the bias during the reading operation and the conversion at the end-of-column circuitry. A traditional 3T pixel structure takes advantage of the M2 transistor working as source follower to amplify the photodiode signal. For the proposed CS working mode, the M2 transistor now works as a current source.

This current source can either work in linear region or in saturation region. In particular, we can recall the Shockley first order transistor model (cf. equation 3.5). This model defines three different working regions known as Cutoff, Linear and Saturated. The driven drain-to-source current I_{ds} (which correspond in figure 3.16 to a part of I_{out}) relies on the threshold voltage V_t and β which is the transistor gain parameter depending

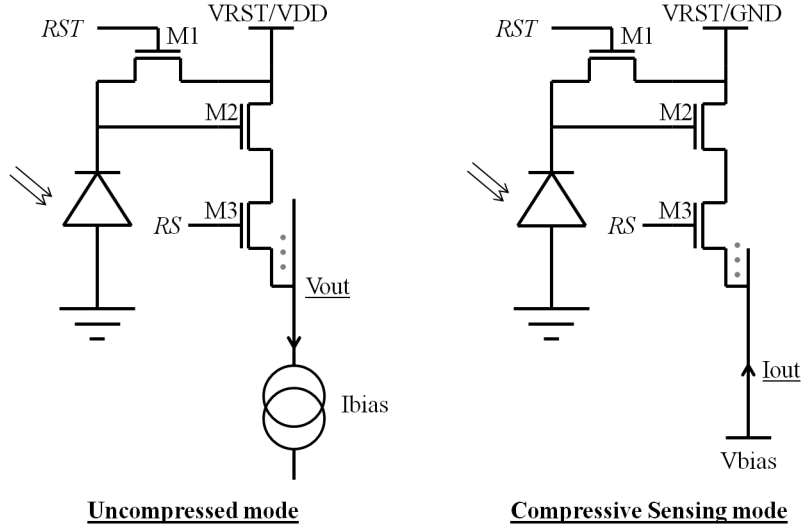


Figure 3.16 – The two different working modes of the 3T pixel.

on the mobility, the gate oxide capacitance per unit area C_{ox} and its size W and L :

$$I_{ds} = \begin{cases} 0 & V_{gs} < V_t \\ \beta(V_{gs} - V_t - V_{ds}/2)V_{ds} & V_{ds} < (V_{gs} - V_t) \\ \frac{\beta}{2}(V_{gs} - V_t)^2 & V_{ds} > (V_{gs} - V_t) \end{cases} \quad (3.5)$$

In this work, we mainly focus on the most linear working mode to linearly convert the voltage of a photodiode into a current at least for low light pixels. As an alternative, we can also use the $M2$ transistor to work in the saturation region over the entire dynamic for non-linear imaging. Indeed, it allows to perform a global tone mapping at the acquisition stage for dynamic range compression purposes. Figure 3.17 presents the simulation results for different V_{bias} voltage levels forced by the end-of-column circuitry. It results a linear curve describing the behavior of the pixel output current in function of the photodiode voltage. As can be observed in this figure, a large difference exists between schematic and layout extracted simulations. It results a wider linear region. This figure also depicts the impact of a $4mV$ fall of the V_{bias} that can be caused by the bus resistivity in the worst case (this will be discussed later).

Noise and technological dispersions

Several constraints have been taken into account to design such a pixel. We designed the pixel in order to reduce the impact of noise and dispersion issues under the constraint

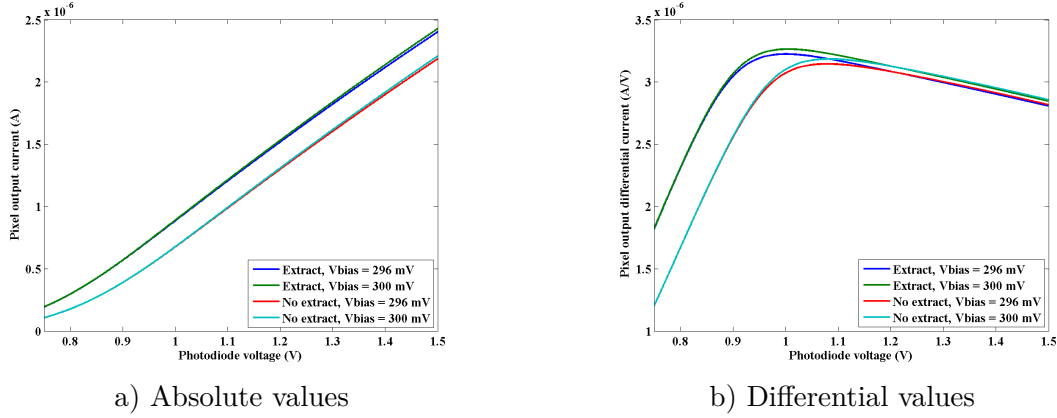


Figure 3.17 – Pixel current response. The pixel output current in function of the photodiode voltage during the current reading operation.

of a pixel pitch of $5.6 \times 5.6 \mu\text{m}$. The resulting layout of the pixel is presented in figure 3.18. A parasitic resistance, grounded capacitance and coupling capacitance extraction of this layout has been performed for retro-anotated simulation purposes. The simulations based on extracted layouts will be called extracted simulations. In particular, in figure 3.17, we can observe the effect of parasitic components on the simulation results.

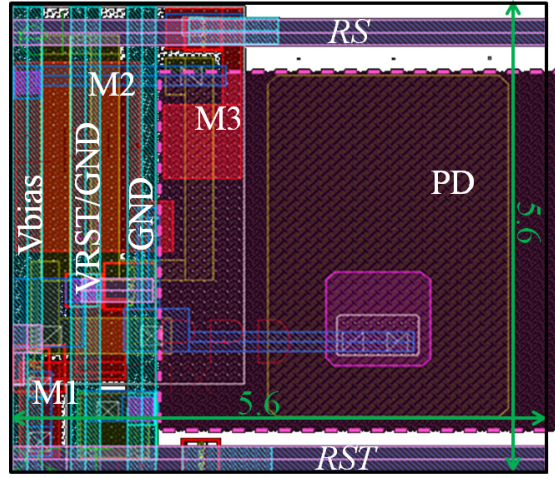


Figure 3.18 – Proposed pixel layout. Fill factor: 32%

The different size of the transistor are reported below. Indeed to reduce possible technology dispersions, $W \times L$ of M2 transistor is chosen relatively large. Indeed, the Pelgrom's law [133] explicits a direct correlation between technological dispersion of a transistor and the product of its W and L . Under the layout constraints we also minimize the W/L ratio in order to make the M2 have a small β thus reducing the generated pixel current output to limit the current driven on each column of pixels. The resulting fill factor is of 32%. However, a microlens (dashed pink layer) is placed on the top of each photodiode to concentrate the light flux.

3.2. CS image sensor architecture and design (SACSI)

	W	L
M1	0.42	0.52
M2	0.30	4.44
M3	0.30	0.86

a) Our design

	W	L
M1	0.42	0.52
M2	1.0	0.5
M3	1.0	0.35

b) A Classical design

Table 3.1 – In-pixel MOS transistor sizes (in μm) for the proposed 3T pixel.

First, the reset kT/C noise is theoretically evaluated for a photodiode capacitance of $C = 5fF$ which gives a thermal noise standard deviation of $910\mu V$. Under a dynamic of $1.6V$ it thus gives a SNR of $65dB$. Using a hard reset (i.e. applying a reset voltage of $2.4V$) the standard deviation of the reset noise of the extracted simulation is of $205\mu V$ leading to a SNR of $78dB$. The readout transient noise impacts more largely the current output with a SNR of $44dB$ under a dynamic of $3\mu A$. Actually, most of this noise is induced by M2 working as a resistance. Removing temporal noise by the low pass filtering operation naturally performed by the self capacitance of the column thus leads to a SNR of $60dB$. This result takes into account all possible noise sources (RST and RS control and V_{bias} are also noisy) during an entire cycle of pixel reading including reset, integration and readout. A Fixed Pattern Noise is induced by the variation of the voltage bias V_{bias} between the top and the bottom of the array of pixels. In particular, the TowerJazz datasheet for metal resistivity gives in the worst case: $\rho_{M3} = 80m\Omega/sq$ and $\rho_{MTop} = 40m\Omega/sq$. Two metal layers (M3 and MTop) are connected together forming the column read out bus. Those buses are sized such that their form factor are $sq(M3) = 512 \times 5.6/0.3$ and $sq(MTop) = 512 \times 5.6/0.47$. It results that the maximum resistances of the buses connections between a pixel and the end-of-column circuitry are respectively $R_{M3} = \rho_{M3} \times sq(M3) = 765\Omega$ and $R_{MTop} = \rho_{MTop} \times sq(MTop) = 244\Omega$. The equivalent resistance is thus $R_{eq}^{col} = 185\Omega$. For instance, under a current of $20\mu A$, the maximum voltage drop is about $4mV$. In the case of maximum current outputs, $12 \times 2.5 = 30\mu A$ impact $5.5mV$ inducing divergence of the driven current of 1.3% . On average if the summed current is of $6 \times 1.25 = 7.5\mu A$ the divergence will be of 0.1% . However, taking into account the distribution of activated rows, there is less impact of the divergence of the pixel behavior on the summed output current. Figure 3.17 presents the results in the case of two different V_{bias} which corresponds to the maximum divergence of the behavior. Yet, the FPN is also and mainly due to technology dispersions. For instance, figure 3.19 presents the results of a Monte-Carlo simulation. This simulation shows the dispersion of output pixel currents for an input considered as a static photocurrent. This consideration has been considered for the pixel design, impacting on the size of the M2 transistor. This dispersion mainly impacts on the final PRNU which has been evaluated to be of $44dB$ according to Monte-Carlo simulations. This noise will represent the major contributor of the final readout noise.

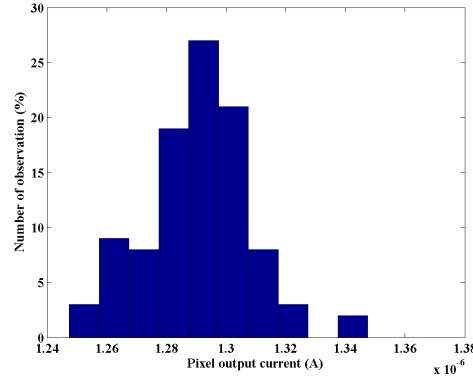


Figure 3.19 – Distribution of the pixel output current, after an arbitrary time of integration resulting of a Monte-Carlo simulation of the pixel.

3.2.3 End-of-column circuitry

The end-of-column circuitry needs to convert a level of current into a voltage with a variable gain. Indeed, one of the main goal of this architecture is to be able to work either in CS mode or in uncompressed mode. To keep the traditional end-of-column in the case of the uncompressed mode, the ADC employed converts a voltage level into a digital value. An auto-zero single ramp ADC is used as in most of CIS. For the current to voltage conversion, the architecture uses a Resistive TransImpedance Amplifier (RTIA). A RTIA structure has been retained instead of a Capacitive TIA because of its ability for dealing with noise issues regarding the current levels which are involved and required silicon area. Moreover, capacitance of the column bus ($3pF$) would impact on the conversion if not fully loaded. Figure 3.20 presents a schematic view of the proposed end-of-column circuitry. This structure is repeated over each column and placed between the column of pixels and the ADC providing bias to the column readout bus and forming the voltage signal level for the ADC.

Tunable gain RTIA

A RTIA is employed to convert the summed pixel output current into a voltage level depending on the number of activated pixels. In order to optimize the systematic offset of the OTA and minimize mismatches issues, we thus choose to use a mirrored cascode [134] (see figure 3.21) with relatively large differential pair transistors. This specific OTA structure has been retained for its high gain and large output swing which however affects the speed. Transient simulation yet shows that even the slew rate is not limiting the application purposes. In addition, this structure allows a large adaptability in the input common mode range, offering a flexible input common mode range for testing different operating parameters. Notice that this RTIA could be further optimized in terms of power consumption once the true characterization of the pixel array.

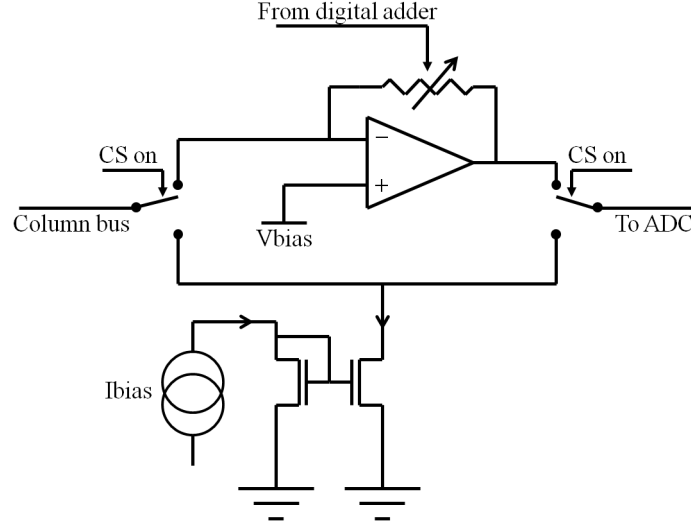


Figure 3.20 – Schematic top view of the reconfigurable end-of-column circuitry.

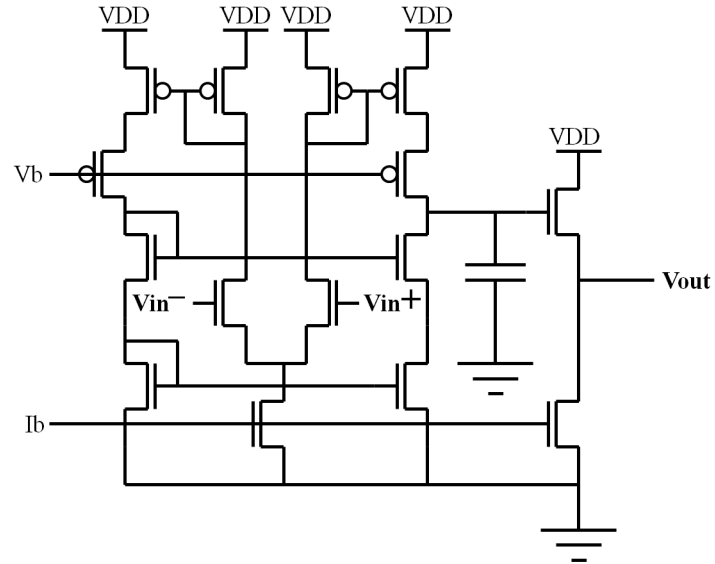


Figure 3.21 – Single-stage mirrored cascode OTA followed by a source follower output stage.

Figure 3.22 presents the Bode magnitude and phase plot of the proposed OTA for a specific function point.

This structure is finally combined with an adaptive resistance in the feedback loop. The goal of the adaptive resistance is to adapt properly the level of output voltage according to the number of pixels. The output of the digital adder is thus used to control this resistance. A resistance matching network has been designed to minimize the level of possible technological dispersions. The top-level schematic of the resistance network is

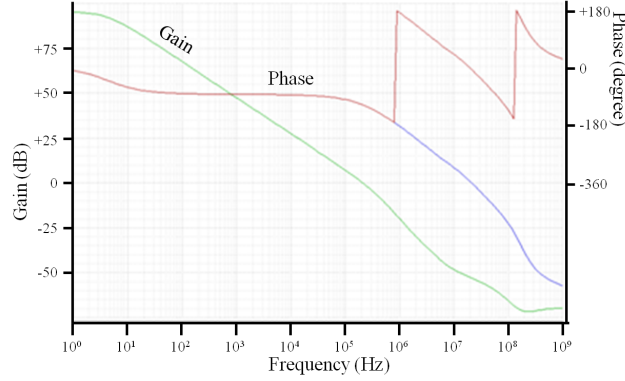


Figure 3.22 – Bode plot of the proposed OTA with its output stage for a typical load and input voltage.

presented in figure 3.23. In particular the transistor switch sizes have also been tuned to preserve proper ratios between the resistivity of the different branches.

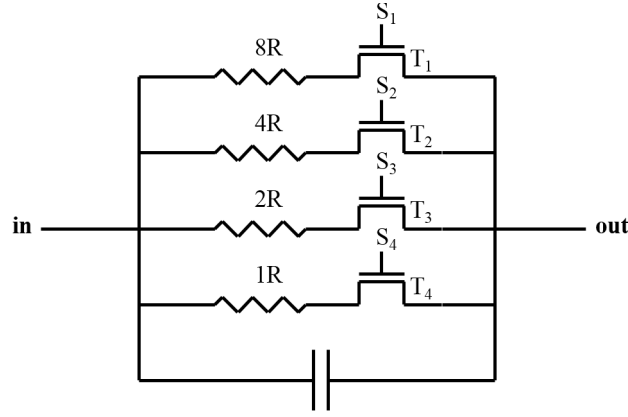


Figure 3.23 – Resistance network composed of 4 switchable resistances.

The unitary resistance is fixed to $R = 66k\Omega$ in order to convert inputs from $0\mu A$ to $3\mu A$ per activated pixel. For this range of inputs, it induces an output between V_{bias} and $V_{bias} + 3 \cdot 10^{-6} \times 528 \cdot 10^3$, for $V_{bias} = 300mV$ it thus leads to $1.884V$. Therefore, the maximum resistance used in the matching network is $R_M = 8R = 528k\Omega$ (used lonely if a single pixel is activated). The following equation 3.6 describes the I to V conversion law depending on the number of activated pixels $N_a = \sum_{i=1}^4 2^{i-1} S_i$.

$$V_{out} = R_{eq} I_{out} = \frac{R_M}{N_a} I_{out} = \left(\frac{S_1}{8R} + \frac{S_2}{4R} + \frac{S_3}{2R} + \frac{S_4}{R} \right)^{-1} I_{out} \quad (3.6)$$

Phase and gain margins for the loaded circuit are reported in table 3.2 for two specific and extreme configurations (12 activated pixels and 1 activated pixel).

3.2. CS image sensor architecture and design (SACSI)

	$R_{eq} = 528k\Omega, V_{in} = 300mV$	$R_{eq} = 44k\Omega, V_{in} = 300mV$
Phase margin	66.8°	74.7°
Gain margin	$-22.5dB$	$-26.5dB$

Table 3.2 – Phase and Gain margins of the proposed OTA for extreme configurations.

To minimize the mismatch, a specific layout (cf. figure 3.24) has been designed with an unitary resistance size and with proper switch transistor sizes. This was made possible because of the binary ratio between the resistances.

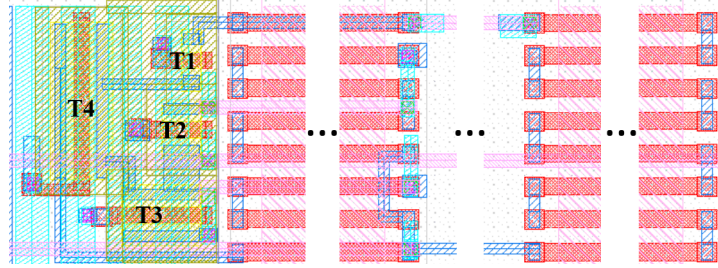


Figure 3.24 – Layout of the resistance network.

Figure 3.25 shows two histograms from Monte-Carlo simulations of the technological dispersion after extraction of the entire I to V converter. Those two histograms represents the output distribution for a fixed input. The first histogram a) is for a single activated pixel (using the largest resistance) and the second histogram b) is in the case of 12 activated pixels (using the smallest resistance).

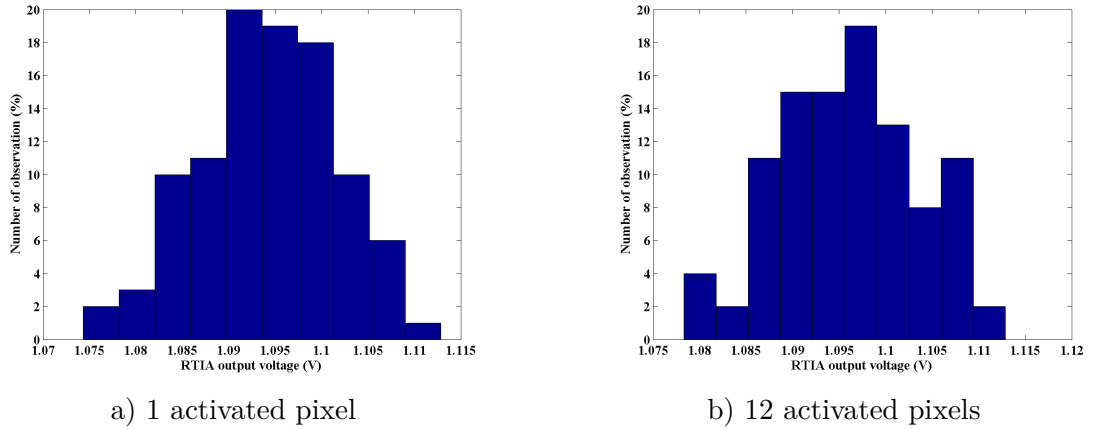


Figure 3.25 – Histogram of Monte-Carlo simulations evaluating the effect of technological dispersions at the RTIA level.

The resulting SNRs for both 1 pixel activation and 12 pixels activation respectively are $46.5dB$ and $46.2dB$. As the technological dispersion observed at the pixel level, the dispersion of the I to V conversion which is mainly due to the dispersion of the resistances

impacts on the PRNU (in a by column fashion). Indeed it affects the gain of conversion leading to a slightly different gain between each column. However, this dispersion is largely less important than the one induced at the pixel level. Figure 3.26 presents the I to V responses of the system. Those curves shows that this response does not depend on the number of pixels. Indeed, the small difference (especially when the input current is high) is negligible compared to other noise and dispersion sources.

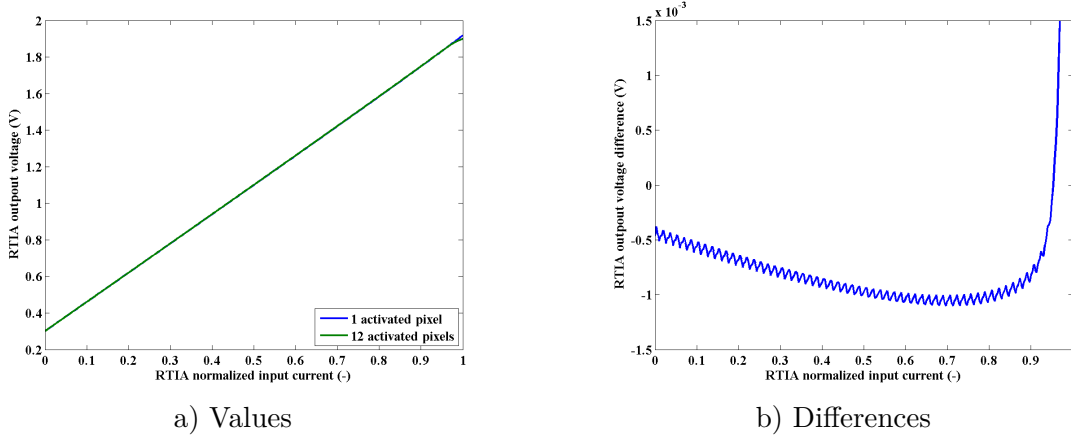


Figure 3.26 – Average output from Monte-Carlo simulations for a normalized input current in the case of 1 activated pixel and 12 activated pixels.

The temporal noise induced by the I to V converter which is given by extracted simulations leads to a SNR of 80.6604dB. In addition, temporal noise introduced by the current source inside the pixel is also filtered by the I to V converter. Finally, a top-level Eldo simulation (see figure 3.27) shows the output of the I to V converter during the reading operation for multiple input photocurrents.

From multiple extracted simulations, the response of the entire reading chain is reported for multiple times of integration. Those results are reported in figure 3.28. The different curves of this figure which describe the entire chain, exhibit non-linearities. Those non-linearities are mainly introduced by both the non-linearity due to the in-pixel photodiode capacitance and the M2 transistor working as current source with a low photodiode voltage.

Regarding power consumption, the top level (with the Inputs/Outputs ring) simulations show an average equivalent power consumption of around $50nJ/measurement$ which is an upper boundary. Indeed, we took into account the worst case (black image) where the generated pixel currents are the highest. Notice that this result is of the same order as the power consumption claimed in [93] (i.e. around $12nJ/measurement$). This preliminary evaluation is also far away from possible optimizations reducing the power consumption, in particular the dynamic one. It would be improved by properly adapting voltage biases and pixel reading time without impacting the column establishing time and readout linearity. This will be tuned by playing on those parameters with the fabricated test

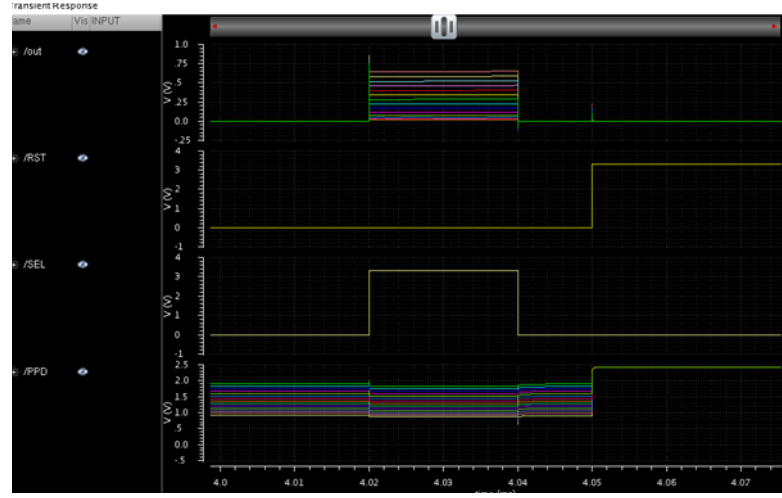


Figure 3.27 – Transient extracted simulation of a CS acquisition during the pixel reading operation, for a single activated pixel and several input photocurrents. From top to bottom: output of the RTIA V_{out} (V), RST command, RS command (SEL), Photodiode voltage (V).

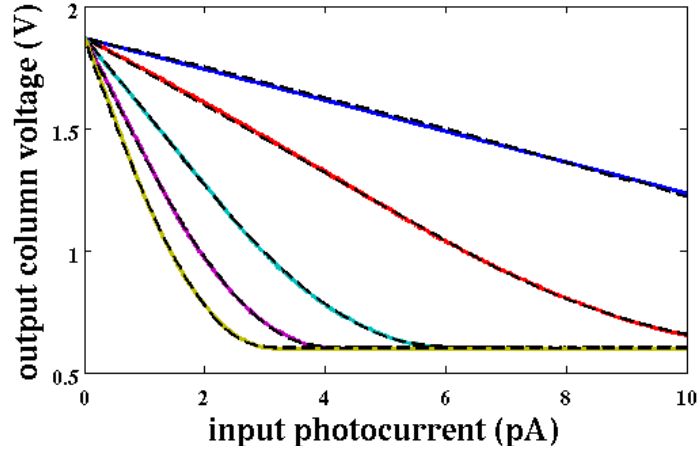


Figure 3.28 – Output voltage of the I to V converter in function of the input photo current in the case of a unique activated pixel. The different colored curves represents different integration time: Yellow ($t_{int} = 4ms$), Purple ($t_{int} = 3.5ms$), Light blue ($t_{int} = 2.5ms$), Red ($t_{int} = 1.5ms$), Blue ($t_{int} = 0.5ms$).

chip. What should be mainly retained from this overall design is that it proposes a compact solution for performing row parallel CS measurements without destructive pixel readouts and multiple specific row selection commands possible schedules. It also allows to keep an in-pixel design close to a traditional one. With this specific design image CS is no more considered a static measurements but dynamic measurements because of performing successive measurements without storing any information inside the pixel.

3.3 The information considered as a group of measurements

The main goal of this work is to show that CS can be an alternative acquisition mode which can be easily implemented in a CIS without large modifications of the overall traditional architecture. As presented in the introduction, possible applications of such an acquisition are numerous. In this chapter, three different applications will be presented before a brief discussion on a solution to get rid of unknown non-linearity issues due to the acquisition model improving the robustness of the reconstruction algorithm. The first application corresponds to the acquisition of still image. In this case, the proposed architecture is not perfectly suited for this application but remains interesting compared to other proposed CS CIS because of its small impact on the entire CIS design. An evaluation of noise impacts on the reconstruction will be provided according to the noise and dispersions given by extracted simulations.

3.3.1 Still image reconstruction

There have been a major interest since now in the signal processing community to define appropriate algorithms to reconstruct still image from CS measurements. In this section we demonstrate that the proposed CIS architecture when working in CS mode can be used as a CS still image sensor. In particular, we provide suitable reconstruction algorithms which have been developed jointly with the sensor architecture to work properly. No intensive comparison with state-of-the-art results will be detailed because of no relevance with the sensing scheme we use.

Proposed prior operators

In the literature, plenty of CS reconstruction methods have been developed for greyscale images. This work mainly deals with reconstruction based on the equation 2.17. In that scope, multiple basic prior operators have been evaluated. The prior operators defined by the equations 3.7 to 3.12 have been tested and compared. In particular, most of those operators uses multiple gradient operators which are respectively denoted ∇_h , ∇_v , ∇_{d1} and ∇_{d2} representing the horizontal gradient, the vertical gradient and the two main diagonal gradients. For instance, for an image U (in its matrix form), $(\nabla_{d1}U)_{i,j} = U_{i+1,j+1} - U_{i,j}$ and $(\nabla_{d2}U)_{i,j} = U_{i+1,j-1} - U_{i,j}$. In the following every time where it will be question of gradient operator a zeros padding operator has been preferred to a cyclic point of view for boundaries. Discrete Wavelet Transform (DWT) are also used as mathematical operator in the described p operators. In particular, we take advantage of multiple DWT (mDWT) consisting of the decomposition of a signal into multiple wavelet basis. For example, here Ψ_i denotes matrices performing Daubechies Wavelet Transforms of different orders. In this subsection Ψ_1 represents D2 transform (i.e. Haar transform), Ψ_2 D6 transform and Ψ_3 D10 transform. The use of this three

different orders of Daubechies wavelet have been chosen according to empirical results [135]. Indeed, when combining multiple DWT it is better to employ orders that are the less the possible multiple of each others to reduce the impact of over fitting which generally induces ringing artifacts.

The first prior operator we used is the classic anisotropic TV. This prior has been widely employed in the literature and with multiple variations. In addition, fast implementations to perform TV regularizations already exist providing relatively good results compared to its processing time and live memory requirements.

$$TV_{aniso}(U) = \|\nabla_h U\|_1 + \|\nabla_v U\|_1 \quad (3.7)$$

The second operator is the combination of multiple wavelet transform l1 norm. This operator promotes an averaged sparsity over multiple wavelet basis which are based on different scaling functions of different lengths. This way a compromise is offered during the reconstruction between high frequency and low frequency components of the image. This operator shows little improvements compared to a single wavelet transform.

$$mDWT(U) = \sum_{i=1}^3 \|\Psi_i^T U\|_1 \quad (3.8)$$

The next prior operator is a simple extension of the anisotropic TV. Indeed the TV_{4axes} characterizes the total variation not only using vertical and horizontal gradients but adding the information on the variations of the image along the diagonals. This point of view still can be extended by the use of high order gradients but without highly significant improvements. For instance article [136] presents an iterative directional TV refinement for CS recovery which seems to improve the results compared to a basic TV by on average 1dB in terms of PSNR.

$$TV_{4axes}(U) = \|\nabla_h U\|_1 + \|\nabla_v U\|_1 + \|\nabla_{d1} U\|_1 + \|\nabla_{d2} U\|_1 \quad (3.9)$$

The prior operator based on the joint TV and $mDWT$ is designed because those two priors works in an antagonist way. The ringing artifacts related to the $mDWT$ tend to be compensated by the cartoon artifacts introduced by the TV regularization.

$$TVmDWT(U) = TV_{aniso}(U) + mDWT(U) \quad (3.10)$$

Finally, the last two priors operator that have been studied are based on the combination of TV (or TV_{4axes}) and $mDWT$. Those priors are based on the idea that the gradient of the image is sparse in the spatial domain (TV) but largely sparser in the wavelet domain ($mDWTonTV$). In fact, it also relax priors of the DC and low frequency components which often penalize the reconstructions based on DCT or DWT because they are not sparse in general for natural images.

$$mDWTonTV(U) = \sum_{i=1}^3 \|\Psi_i^T \nabla_h U\|_1 + \|\Psi_i^T \nabla_v U\|_1 \quad (3.11)$$

$$mDWTonTV_{4axes}(U) = \sum_{i=1}^3 \|\Psi_i^T \nabla_h U\|_1 + \|\Psi_i^T \nabla_v U\|_1 + \|\Psi_i^T \nabla_{d1} U\|_1 + \|\Psi_i^T \nabla_{d2} U\|_1 \quad (3.12)$$

Based on the acquisition scheme described in the previous section which is performed by the SACSI, we reported several results of reconstruction using the different prior operators. Figure 3.29 presents the average reconstruction PSNR for a set of 5 test greyscale images (Lena, Darkhair, Cameraman, Blonde and Pirate). Figure 3.30 presents the average reconstruction SSIM for the same set of images.

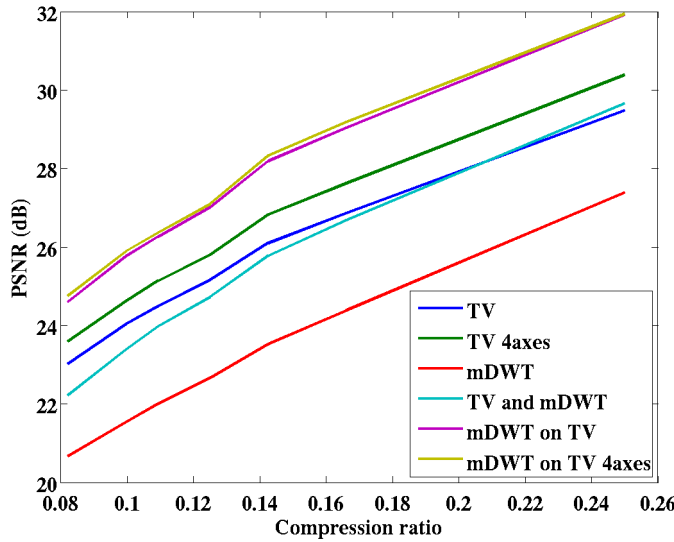


Figure 3.29 – Average Peak Signal to Noise Ratios computed between original greyscale images and the reconstructed ones using multiple different prior operators.

As we can observe on those figures, basic TV works relatively well compared to $mDWT$ but when both are used an improvement is noticeable when the compression ratio is higher

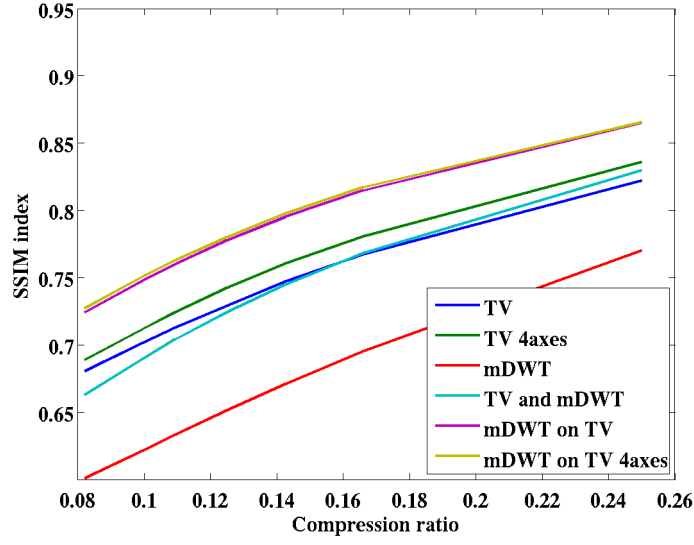


Figure 3.30 – Average Structural Similarity Image Measurements computed between original greyscale images and the reconstructed ones using multiple different prior operators.

than 0.2. The use of TV_{4axes} increases the output PSNR of $1.2dB$ and the SSIM of 0.03 on average compared to the basic TV which is considerable for such a slight modification on the overall algorithm complexity. Finally the best results are obtained when using the combination of TV_{4axes} with $mDWT$ (i.e. $mDWTonTV_{4axes}$). In fact, the improvement is considerable, it represents -at constant compression ratio- more than $2dB$ in terms of PSNR and 0.1 in terms of SSIM compared to TV_{aniso} . However applying this last prior at the reconstruction increase the time of reconstruction by a large factor (10 more or less). The conclusion of this study is that at the expense of a more complex, more time consuming prior operator, the classical TV regularization can be improved without large modifications. For example, figure 3.31 presents the reconstructions obtained using the different prior operators of the Darkhair test image for a compression ratio of 11%. As we can clearly see, the visual quality is largely improved when using the $mDWTonTV_{4axes}$ prior operator compared to the others. Even if providing worse results in terms of SSIM and PSNR, the combination of multiple wavelet with TV (i.e. $TVmDWT$) still improves the overall visual quality compared to either TV or $mDWT$. The use of 4 directional gradients relaxes constraints on straight angle edges when TV regularized when wavelet transforms are employed.

A brief comparison of those results with state-of-the-art techniques is provided in the table 3.3. the reconstruction PSNRs presented in [137] for a block-Based acquisition using Total Variation (BCS-TV) and Smoothed Projected Landweber Dual-Tree Discrete Wavelet Transform (BCS-SPL-DDWT).



Figure 3.31 – Reconstruction of the Darkhair test image using CS measurement scheme of the SACSI for a compression ratio of 11%. From left to right and top to bottom, the prior operators are TV , TV_{4axes} , $mDWT$, $TVmDWT$, $mDWTonTV$ and $mDWTonTV_{4axes}$.

method	CR	<i>BCS</i>		<i>SACSI</i>	
		<i>TV</i>	<i>SPL-DDWT</i>	<i>TV</i>	<i>mDWTonTV_{4axes}</i>
Barbara	10%	22.5	22.9	20.7	21.9
Barbara	20%	23.6	24.3	23.2	24.7
Lena	10%	27.9	28.3	24.5	26.6
Lena	20%	30.6	31.4	28.3	30.9

Table 3.3 – Comparison of resulting PSNRs for different reconstruction algorithms and sensing schemes.

Noise impacts

The previous reconstructions have been performed without taking into account noise and dispersion issues that have already been estimated using extracted simulations of the SACSI architecture (previous section). In order to be as close as possible to the real situation of having CS measurements performed by the proposed architecture, we added quantification noise (10bit), an overall PRNU of 40dB (the worst case) and a reset noise (mainly due to kTC noise) of 60dB. In addition, the model of divergence of the pixel behavior between the top rows and the bottom ones has been added (also taking the worst case which correspond to a maximum resistivity of the column buses). The pixels are supposed to work linearly in an acceptable dynamic range which corresponds, in the

reconstructions, to the range of codes between 0 and 255 for the 8bit test images.

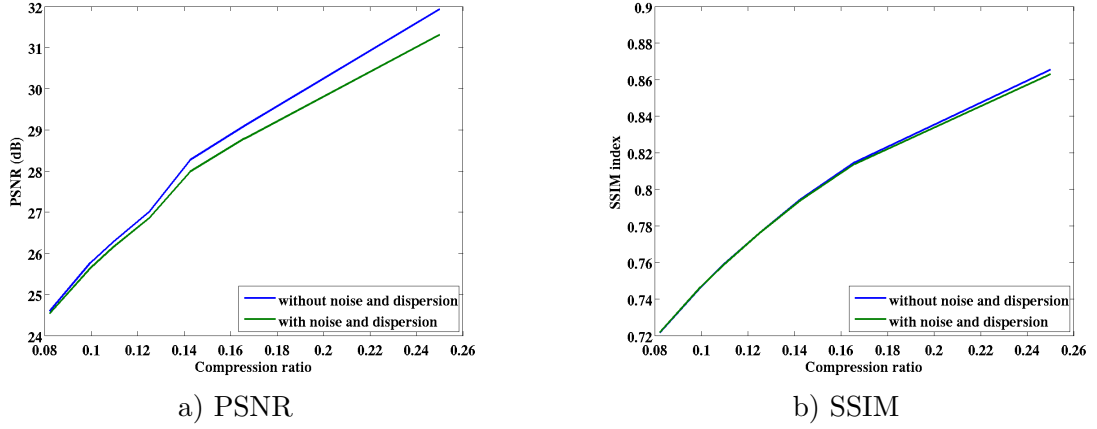


Figure 3.32 – Average PSNR (a) and SSIM (b) of the reconstructions with noise and dispersion on the CS measurements.

Figure 3.32 presents averaged results for reconstructions performed using measurement inputs with noise and dispersion. For these results we used the $mDWT_{onTV_{4axes}}$ prior operator. As one can observe, both curves are close to the ones obtained without noise. It thus means that technological dispersions and noises that have been estimated have only a small impact regarding the reconstruction. In the case of a traditional acquisition this would have an impact because of high fidelity of the acquisition, in the case of CS this has almost no impact on the quality. This observation is directly linked to the intrinsic nature of CS. Indeed, assuming the image not exactly sparse and noisy, the reconstruction intrinsically performs denoising. Yet, noise does not correspond to the only artifact generated at the acquisition. Indeed, in order to properly model the acquisition chain we should also add readout circuit non-linearities and photon shot noise. For the moment, the photon shot noise will be neglected because of its negligible impact since supposing that the measurements are from the same time of integration and for normalized image (maximum value of the image corresponds to the saturation level). Indeed, as it will be presented, it only impacts on the reconstruction for multicapture HDR-CS.

Non linearity and pixel model

Now, we thus want to evaluate the impact of non-linearity of the pixel. Still, thanks to the extracted simulations of the entire reading chain, the pixel integration behavior can be modeled. We use spline functions to model the pixel output after a certain time of integration in function of the input photocurrent. The curves presented in 3.33 (spline functions) correspond to the normalization of the dashed black curves that are plotted in figure 3.28. The spline model we employ corresponds to second order spline functions which are described by the equations 3.13 and 3.14. The construction of this model is based on the realistic assumption that the integration behavior is composed of 3 distinct

regions: linear, quadratic and fully saturated. Moreover, one can reasonably assume that the function is continuous and differentiable. We thus obtain the equation 3.15 that is used in the reconstruction algorithms based on FISTA in our case. In particular those spline functions are parameterized by only two parameters the integration slope τ and the linearity threshold α .

$$\begin{cases} a &= \frac{\tau^2}{\alpha} \\ b &= -(\tau + \frac{2a(1-\alpha)}{\tau}) \\ c &= \frac{b^2}{4a} \end{cases} \quad (3.13)$$

$$f(x) = \begin{cases} 1 - \tau x & \text{if } x < \frac{1-\alpha}{\tau} \\ ax^2 + bx + c & \text{else if } x < -\frac{b}{2a} \\ 0 & \text{otherwise} \end{cases} \quad (3.14)$$

$$f'(x) = \begin{cases} -\tau & \text{if } x < \frac{1-\alpha}{\tau} \\ 2ax + b & \text{else if } x < -\frac{b}{2a} \\ 0 & \text{otherwise} \end{cases} \quad (3.15)$$

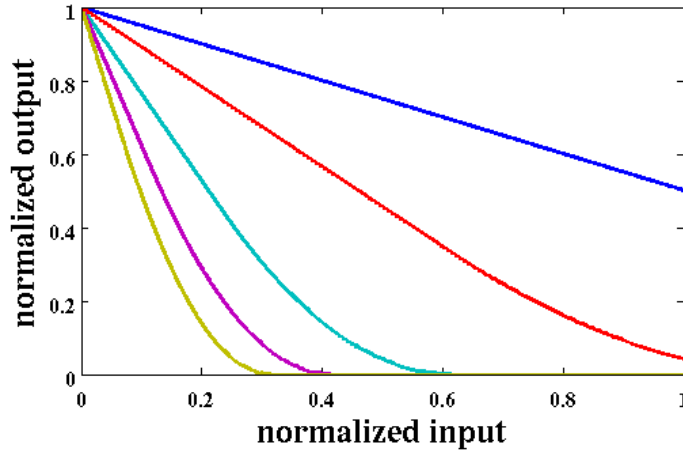


Figure 3.33 – Normalized transfer function of the pixel between the final readout (output of the RTIA) and the incoming light on the photodiode.

At the reconstruction, this non-linearity can be handled with 2 different techniques. Either the reconstruction is performed in the non-linear domain and since the optimization

procedure has converged, an inverse non-linear function is applied to the resulting image or the reconstruction directly works in the non-linear domain. We have experimentally observed that working directly in the non-linear domain provides better results if the non-linear function is convex (which is the case of spline functions). Moreover, in the case of simulation the quality assessment is performed between the original image and the reconstructed image, therefore the prior operator better work when applied directly on this image. In particular, [19] presents theoretical results if a point-wise non-linear function is incorporated for performing compressed measurements. The only supplementary constraint is that the derivative of this function have to be bounded (i.e. by its Lipschitz constant). The fidelity term $FT(\mathbf{u}, \mathbf{y})$ is now replaced by the following equation 3.16 when considering the non-linearity.

$$FT(\mathbf{u}, \mathbf{y}) = \|\mathbf{y} - \Phi f(\mathbf{u})\|_2^2 \quad (3.16)$$

f' exists (cf. equation 3.15) and is bounded, therefore we can apply it to directly derive the expression we use inside the reconstruction algorithm:

$$FT'(\mathbf{u}, \mathbf{y}) = -2f'(\mathbf{u}) \cdot \Phi^T(\mathbf{y} - \Phi f(\mathbf{u})) \quad (3.17)$$

Figure 3.33 provides the conversion models for a given time of integration. For simulation purposes, we propose to use the following parameters of the spline function: $\alpha = 0.5$ and $\tau = 1.5$. This way, and because of the normalization, there is no saturation region of the pixel taken into account. Yet, an important part of the curve follows a quadratic law. Figure 3.34 presents the results in terms of PSNR and SSIM in the case of a non-linear, noisy and technology dispersed sensing. One can observe a slight degradation (0.5dB) on those results compared to the ones that are not taking into account pixel non-linearities. It is mainly due to the fact the dynamic is strained in some region thus increasing the impact of the quantization and because of the low number of activated pixels (SACSI sensing scheme).

Color limitations

The human eye has the ability of colorization of images because of acquiring both chrominance and luminance information into a scene. To reproduce this color information at a screen display or to print a digitally sensed image, the acquisition needs to be somehow spectral dependent. However, traditional photosensitive elements (photodiodes) do not discriminate spectral information, these are only converting photons energy into

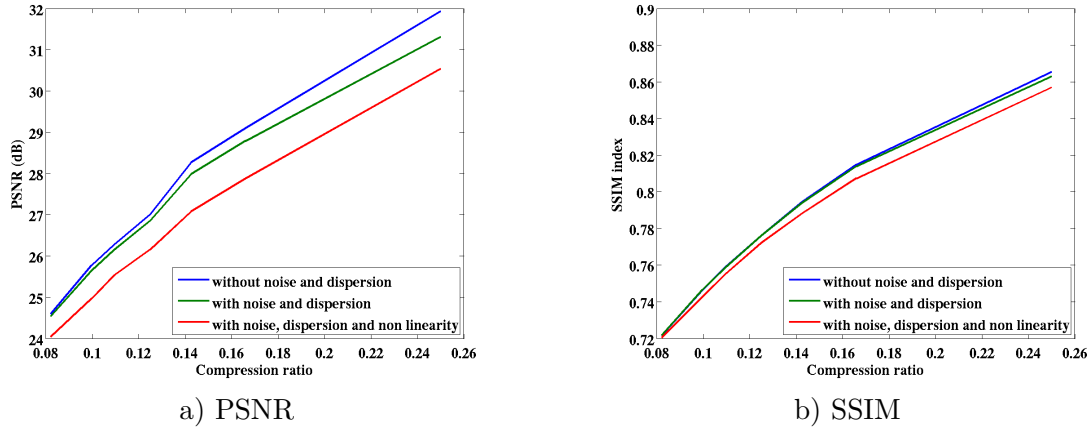


Figure 3.34 – Average PSNR (a) and SSIM (b) of the reconstructions with noise and dispersion on the CS measurements taking into account non-linearity of the pixel.

a photo current. Color CISs thus commonly use an additional Color Filter Array (CFA) which is a mosaic of color filters placed on the top of the pixel sensors.



Figure 3.35 – Bayer Pattern of Color filter Array (CFA).

Therefore, these CFA maps color indexes to spatial positions in the image. One can consider that CFA allows a compromise between spectral information and spatial information because of sensing only one color component per image pixel. To mimic the physiology of the human eye, a CFA is generally based on a small pattern composed of 50% of green filters, 25% of red filters and 25% of blue filters. This pattern is then repeated on the entire matrix of pixels to form the CFA. A particular arrangement of color filter called Bayer filter is widely used because of its simplicity both in terms of hardware implementation and color reconstruction algorithm. This pattern is depicted in figure 3.35. In terms of signal, classical RGB color images are defined by a concatenation of three grayscale images representing respectively the Red, Green and Blue bands. This kind of images can thus be considered as 3D signals. When an image is acquired using a CFA, the RGB image is full of undetermined values ($2/3$) and is called a mosaiced image. Plenty of methods have been proposed in the literature to restore the undetermined color information. This operation named demosaicing can be basically performed using a bilinear interpolation of the unknown values but finally inducing many artifacts that are mainly related to aliasing issues (due to a lack of spatial information compared to spatial highest frequency). Because of being an undetermined linear problem, it can also be formulated as a CS problem. In particular, [138] proposes demosaicing techniques using a CS framework. [139] and [103] present models for performing color CS combining a CS acquisition scheme and a CFA. [103] proposes a color CS reconstruction based on redundant transforms and group sparsity which exhibits good empirical results compared to separate band reconstruction. In particular, group sparsity is well suited because of

correlation of the image structure between color band images. The effect of the structure of CS-appropriate panchromatic CFAs compared to traditional Bayer CFAs have also been investigated in [140] and [141].

In the case of the SACSI combined with a Bayer CFA on top of the focal plane of the sensor, it means that the problem to solve is to reconstruct directly the colored image from CS mosaiced measurements. To optimally achieve this operation, a specific prior operator which is using priors based on the 3D representation of color images is defined. To mimic classical compression methods, the proposed reconstruction exploits an alternate representation of color RGB images which is the Luminance/Chrominances representation. This “color space”, called YCbCr make the 3D color image signal sparser than in its RGB representation. We thus use for the color reconstruction the operator $mDWTonTV_{4axes}$ applied simultaneously on the Luminance and Chrominances components (cf. equation 3.18). Better results can even be obtained when adding different weights on each components, enforcing more sparsity on Chrominances parts (just as the quantization proceeds in the case of JPEG). We denotes $U_{R,G,B}$ the color image (i.e. 3D signal in its RGB format) and $U_{Y,Cb,Cr}$ the same image in its YCbCr format. Each color component is denoted $U_{component}$ (like the luminance of U is U_Y).

$$mDWTonTV_{4axes}^{RGB}(U_{R,G,B}) = \sum_{S=\{Y,Cb,Cr\}} mDWTonTV_{4axes}(U_S). \quad (3.18)$$

The following equation 3.19 presents the convention we use for the color space transform (considering normalized images between 0 and 1).

$$\begin{cases} U_Y &= 0.299U_R + 0.587U_G + 0.114U_B \\ U_{Cb} &= -0.1687U_R - 0.3313U_G + 0.5U_B + 0.5 \\ U_{Cr} &= 0.5U_R - 0.4187U_G - 0.0813U_B + 0.5 \end{cases} \quad (3.19)$$

Figure 3.36 shows an experimental result of the reconstruction of the colored Lena test image with a compression ratio of 12.5%. We use a model of the traditional Bayer color filter as a mask of the original colored image before performing SACSI CS measurements. One can consider a relatively high fidelity of the reconstruction even if the ratio of acquired information represents only 4.2% of the reconstructed information.



Figure 3.36 – Reconstruction of the colored Lena test image with an equivalent greyscale compression ratio of 12.5%. The PSNR of the reconstructed image is of $26.9dB$.

As example, compared to existing reconstruction techniques, and for a compression ratio of 25% the proposed algorithm reaches a PSNR of $29.86dB$ when the method of [139] provides a PSNR of $28.16dB$.

3.3.2 Video reconstruction, the time into slices

Another application addressed by CS –which have been investigated by the signal processing community– is the compressed video acquisition. Indeed, thanks to the specific nature of video sequences and its highly compressible property –that is already exploited in traditional video compression techniques– one can feel that CS can be easily employed in that scope. Frame to frame correlation have to be considered when video compression is involved, in comparison with still image compression. For instance qJPEG which is based on a frame by frame JPEG compression shows largely worse results than MPEG compression which takes into account motion estimation for coding a video sequence. In the case of CS, the measurement corresponding to each frame is non-adaptive but providing sufficient information for the decoder. Before of being implemented in an image sensor or even being studied as a way to acquire video, CS have been investigated as a compression framework lighting the encoder in terms of computational requirements. Those compression schemes generally use reference frames which are less compressed and between frames which are more compressed [142] [137] [143]. Adaptive CS frameworks have also been proposed to adapt the compression ratio locally according to the scene motion [144]. In particular, the article [145] reviews the different video coding techniques using CS. Regarding the reconstruction, the proposed techniques rely on Group Of Frame

(GOF) decomposition. [146] proposes to iteratively update and learn a dictionary which is used for the next frame as a sparsity representation. [147] uses an adaptive low rank and sparse decomposition at reconstruction. [148] and [137] focus more on block-based acquisition which is more relevant in practical perspectives to implement CS on a imaging device. [137], [143] propose to reuse the concept of motion estimation (ME) / motion compensation (MC) during the CS reconstruction. Indeed, since we can assume that motion vectors in a scene are “sparse”, it appears relevant to add a frame to frame prior taking into account the motion.

Video, time gradient priors

We empirically observe that when a too simple block motion estimator is employed and without reference frames, adding a block motion compensation prior tends to introduce block artifacts. This is also due to the fact that if the reconstructed frame is not of sufficient quality the block motion estimation can provide wrong information for the motion compensation. Yet, for instance, in [143] a complex motion estimator is employed combined with reference frames. Even if providing particularly interesting results, this approach is not well suited for an in-line CS camera for two main reasons. First, regarding the SACSI sensing scheme the measurement vectors are regular and non adaptive (which is totally not the case in [143]). Secondly, the reconstruction which is used does not require to be camera motion resilient supposing the camera is fixed for instance for CCTV. Without this motion compensation, two orientations have been investigated. A reconstruction based on the simultaneous reconstruction of a Group Of Frames (GOF) and an in-line reconstruction in which the reconstruction of each frame successively exploits the content of the previous reconstructed frame. This second reconstruction method yet requires an initial frame of a good quality (uncompressed), then one can assume that CS measurements only provide innovation information.

In this section, we will only deal with the GOF (p frames) reconstruction. We thus looked at the option of working directly on the 3D video signal (V), using a prior operator which is based on frame-to-frame priors. Indeed, for fixed camera (but not only), the gradient on the time axis is a sparse image. We first adapted the TV which traditionally works in 2D to deal with a 3D signal. The results we obtained were noticeable compared to a TV performed frame by frame. However, better results have been obtained with the following three prior operators. The first one, $TV_{SO}TV_{2D}$ consists of applying a TV regularization for both the time gradient and each frame (cf. equation 3.20).

$$TV_{SO}TV_{2D}(V) = \sum_{k=2}^p TV_{aniso}(U^k - U^{k-1}) + \sum_{k=1}^p TV_{aniso}(U^k) \quad (3.20)$$

The second one, $TV_{SO}mDWT_{2D}$ which is depicted by the equation 3.21 is based on combining multiple wavelet transforms on each frame with a TV on the time gradient.

$$TV_{SO}mDWT_{2D}(V) = \sum_{k=2}^p TV_{aniso}(U^k - U^{k-1}) + \sum_{k=1}^p mDWT(U^k) \quad (3.21)$$

Finally, $mDWT_{SO}TV_{2D}$, the combination of a TV on each frame with a multiple wavelet transform on the time gradient (cf. equation 3.22) exhibits the best results. Figure 3.38 shows the reconstruction for a compression ratio of 6% using this prior operator.

$$mDWT_{SO}TV_{2D}(V) = \sum_{k=2}^p mDWT(U^k - U^{k-1}) + \sum_{k=1}^p TV_{aniso}(U^k) \quad (3.22)$$

In order to speed up the reconstruction of 3D signals, a 2-dimensional TV regularization is performed first constituting an initialization point. Then, during the reconstruction using a FISTA-based algorithm, the regularization parameter λ is iteratively refined. This way, a coarse reconstruction is first performed using a small value of λ and then improved by iteratively updating (increasing) its value to reach a final targeted λ value. This algorithm that we employed is based on the idea that for a high fidelity of the reconstruction (low noise and thus high proper λ) the algorithm faster converges when the prior is already enforced to the signal. It can be empirically observe that for large signals, it is faster to partly solve multiple cascaded minimization problems than only one but directly fully.

Figure 3.37 presents reconstruction results in terms of SSIM. The conclusion of this test bench is that for the 4 video sequences we tried, the best prior is $mDWT_{SO}TV_{2D}$ if no large camera motion (i.e. Foreman). We observed that the collaborative reconstruction using wavelet decompositions of the frame to frame differences highly improves the overall quality of the image. In terms of visual quality assessment, $mDWT_{SO}TV_{2D}$ clearly beats the other prior operators.

One of the major interest of using SACSI is that the CS measurements can be performed at the same frequency than a pixel row readout in a traditional rolling shutter based CIS. The pixels can thus be read multiple times during its integration time or be reset every 8 CS column-based measurements using the local rolling shutter.

3.3. The information considered as a group of measurements

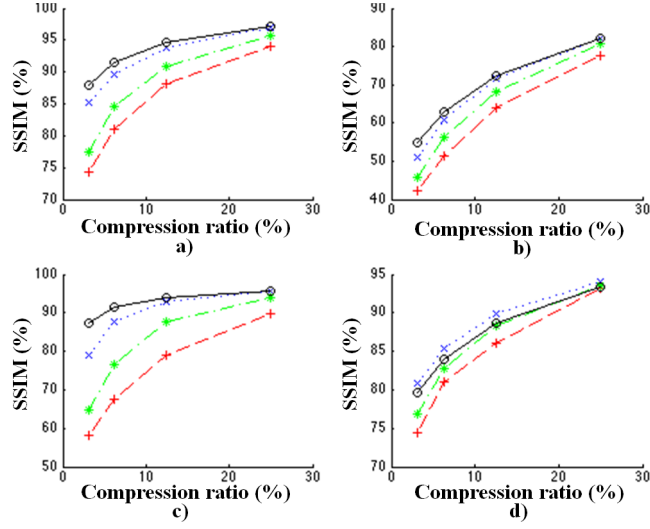


Figure 3.37 – Average SSIM over successive frames (88) for the different prior operators ($TVmDWT_{2D}$ (red), $TV_{SO}mDWT_{2D}$ (green), $TV_{SO}TV_{2D}$ (blue), $mDWT_{50}TV_{2D}$ (black)). Four 256×256 center cropped video test sequences have been used: Mother and Daughter (a), Coastguard (b), Hall Monitor (c) and Foreman (d).



Figure 3.38 – Hall monitor CS reconstructions for a compression ratio of 6.25% using different prior operators. From left to right, from top to bottom, $TVmDWT_{2D}$, $TV_{SO}mDWT_{2D}$, $TV_{SO}TV_{2D}$, $mDWT_{50}TV_{2D}$.

High frequency reconstruction, coarse to fine

Generally, CISs are drastically limited in terms of frame rate due to the time of integration, and the internal clock but especially because of data bandwidth limitations. Because of the intrinsic nature of the measurements performed by the SACSI, the by-frame compression ratio can reach $1/512$ which is less than 0.2%. In the case of still image, it is not relevant to employ such a compression ratio because of producing really poor images if the goal of such acquisition is photography and the image size is 512×512 . Yet, it makes sense in the case of an almost static video sequence. In this context and when an image is acquired in CS mode, it can either be reconstructed as a still image or as a video. It thus allows compromises between resolution in time and space. Therefore, it is possible from SACSI measurements –at the extend of additional and intensive processing– to reconstruct a video sequence corresponding to a high frame rate according to its compression ratio. We propose a reconstruction for such an acquisition based on a coarse to fine time step restoration in order to relax the computation load at the decoder. The construction first consider that the entire set of measurement corresponds to a single image, then to two separate images, then to four images and so on, until reaching the compression ratio of $1/512$ which corresponds to the reconstruction of 512 images. At each step the reconstructed image(s) is(are) duplicated to form the initialization of the next step. Figure 3.39 illustrates this CS reconstruction scheme for videos acquired with SACSI.

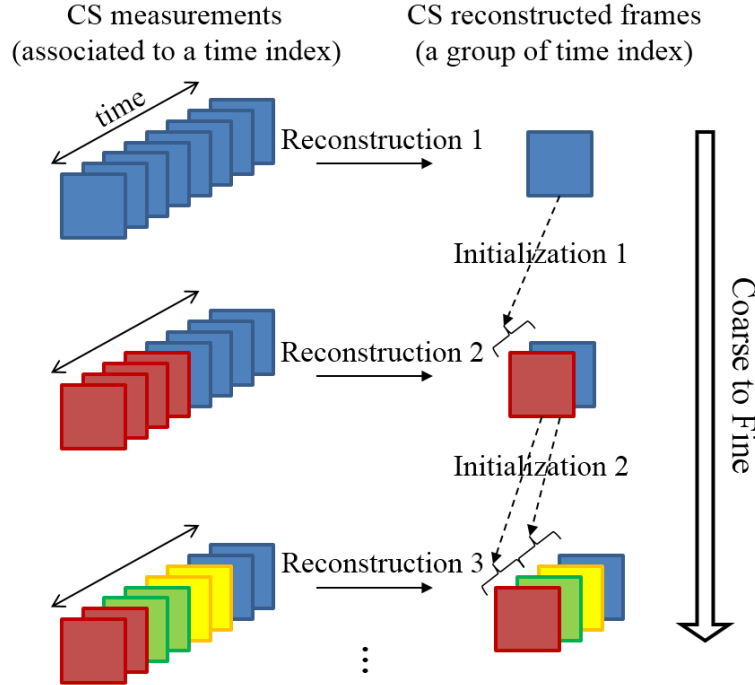


Figure 3.39 – Coarse to Fine CS reconstruction scheme

To show the efficiency of this method we took an original high frame rate video (X fps)

of a very poor quality (with a high level of temporal noise). The video sequence captures the motion of a hand of an analog clock. At the reconstruction, with a compression ratio of $1/2$, the algorithm first reconstructs a single image (see figure 3.40). This single image reconstruction is of good quality compared to the average frame over time ($PSNR = 33dB$) but still exhibits considerable artifacts on the region where the hand is moving during the sequence. Line artifacts caused by temporal noise (non consistency of successive pixels that should be with a constant value) are also introduced due to the particular SACSİ sensing scheme.



a) Original time-averaged test video (mean of 256 consecutive frames).



b) Reconstructed image with a compression ratio of $1/2$.

Figure 3.40 – Comparison between a temporal averaging of the clock video sequence and a CS reconstruction performed for the same sequence.

Yet, the CS measurements still contains the information of two different states in our specific case of video sequence. Indeed, the time index of the measurements can be exploited to refine the time precision of the reconstructed images. A coarse to fine reconstruction shows a real advantage for refining this video, this way the motion of the second hand of the clock can be restored (see figure 3.41). However, some troubles can arise in the case of using the SACSİ sensing scheme because of the low number of activated pixel by measurements, not detecting variation over the entire set of rows (only a small part of it).

In particular, future applications will address the problem of detecting the precise time step when a specific transient happens and without the need of reconstructing the entire video sequence. This is a clear advantage compared to a classical acquisition because SACSİ lets the opportunity at the decoder to extract exact time information. Supposing that the available master clock is $100MHz$ then due to the 10-bit single ramp ADCs, the frequency of conversion is around $97kHz$. In uncompressed rolling shutter mode the frame rate would be of $190fps$ with a maximum integration time of $5.3ms$. In the case of CS mode, the frame rate which is reachable is thus of $97kfps$ when the compression

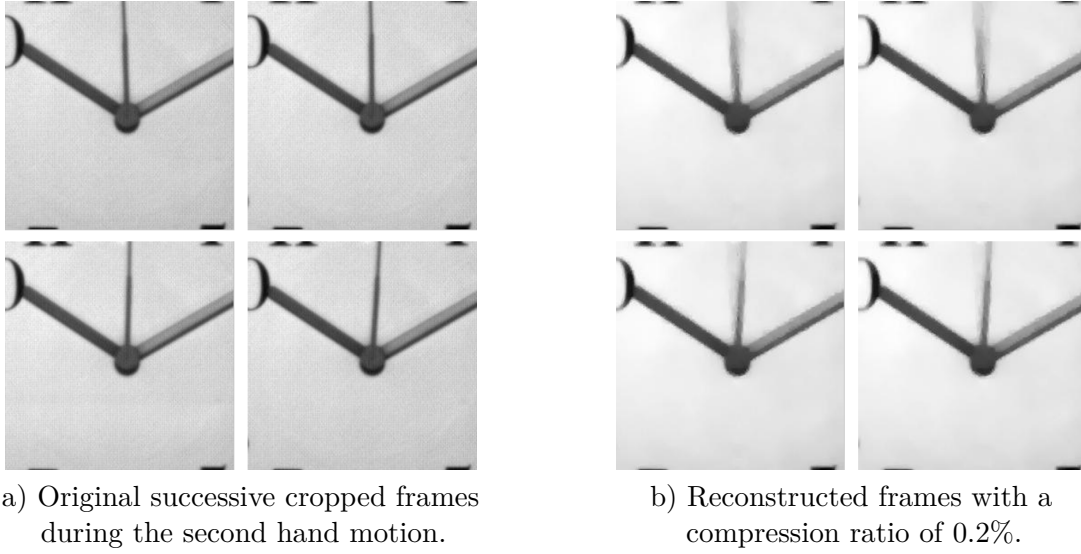


Figure 3.41 – Comparison between several uncompressed cropped frames and their related CS reconstructions (CR = 1/256, integration time=1/fps).

ratio is considered to be of 1/512. Yet, thanks to the local rolling shutter, the maximum integration time is of $82.5\mu s$. This consideration is particularly important because it improves the low light response compared to if there would be no local rolling shutter and the maximum integration time would be of $10.3\mu s$.

3.3.3 High Dynamic Range (HDR) reconstruction

Standard image sensors feature dynamic range about 60dB to 70dB while the light flux of natural scenes may be over 120dB. Most imagers that address such dynamic ranges, need specific, large pixels. With respect to standard pixels, dedicated HDR pixels suffer from many limitations such as poor low-light sensitivity [149] or larger silicon area [150] [151] induced by their extra-transistors [152] [153] [154]. Conventional imagers can also be used for High Dynamic Range (HDR) multi-capture to compensate saturation non-linearity at the expense of the need for large memory requirements or heavy digital embedded processing. The overall acquisition time is thus increased because of data throughputs accordingly. Indeed, multi-capture HDR, which is based on standard imager, implies to multiply power consuming analog to digital conversions and also to store full frame images or interleaved images. We thus hope that CS could reduce these drawbacks proving that it is possible to reconstruct a HDR image from relatively few measurements being performed on images corresponding to different integration times. The proposed technique (multicapture HDR-CS) that exploits the specific SACSI architecture, takes advantage of the knowledge of non-linearities introduced at the sensing stage by the pixel integration response depending on different integration times m (see figure 3.28). This technique is fully non adaptive which differs from the works presented in [155] where

3.3. The information considered as a group of measurements

a feedback loop is added to control a conversion gain and where a SPC is used. In particular, 8 (M_f) different integration times have been considered for the simulations corresponding to binary multiples of the shortest integration time $T1$. Integration times are thus considered relatively among the others and with normalized input light intensity for simulation purposes. An example of time diagram is proposed in figure 3.42. Notice that other CS architectures [61] [93] could be employed for HDR-CS even if not well suited for this purpose.

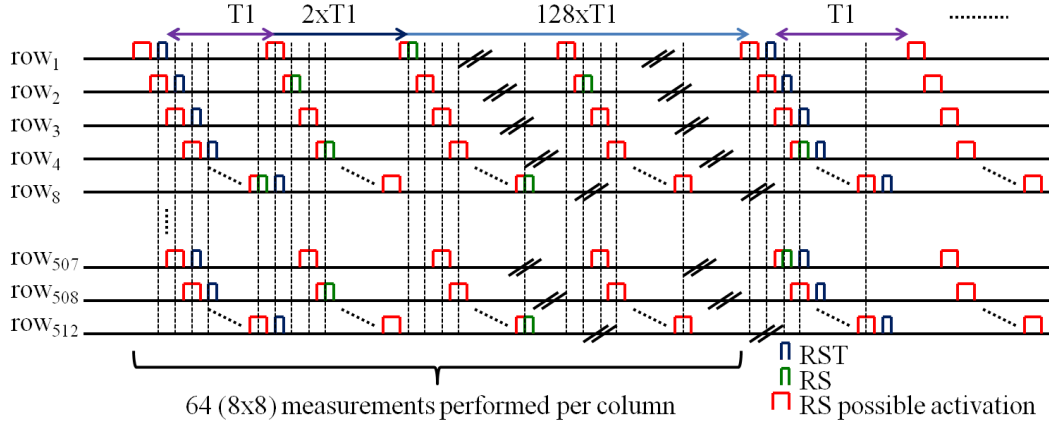


Figure 3.42 – Example of time diagram which can be used for a HDR-CS acquisition.

The reconstruction algorithm takes advantage of knowing pixel integration functions (cf. figure 3.33) related to each integration time. It compensates non-linearities due to saturation in the low and high illumination regions. It combines all the measurements corresponding to each integration time. The proposed reconstruction is performed by an iterative algorithm (algorithm 2) that directly performs a global tone-mapping related to the content of the scene. This algorithm iteratively evaluates the image contrast in order to accordingly tune a γ correction. This γ correction considered as the tone-mapping function $tm(\cdot)$ is then taken into account for the next iteration of the algorithm. This operation flatten the histogram to emphasize low light details also enforcing a better denoising. At each iteration the algorithm takes into account an integration time more (from m to M) for solving the TV regularization in the tone-mapped domain. Notice that the first iteration works as an initialization using only one time of integration and without tone-mapping. For each m^{th} integration time, corresponding measurement matrix and pixel integration function are respectively noted Φ_m and f_m . Robustness of the algorithm is obviously improved by reusing at each iteration the previously solved x as an initialization for solving the minimization problem. In particular, this algorithm provides even better results when TV regularization is replaced by a better prior operator such that $mDWTonTV_{4axes}$ but still at the expense of a larger computational load.

Algorithm 2 Tone-mapped HDRCS reconstruction

Goal: find \mathbf{x} such that $\underset{\mathbf{x}}{\operatorname{argmin}} TV(\mathbf{x}) + \lambda \sum_{m=1}^{M_f} \|\Phi_m f_m(tm^{-1}(\mathbf{x})) - \mathbf{y}_m\|_2^2$

with $tm(\cdot)$ an adaptive tone-mapping function based on a gamma correction

init $tm(\cdot)$, \mathbf{x}

for M in $1..M_f$ **do**

$$FT(\cdot) \leftarrow \sum_{m=1}^M \|\Phi_m f_m(tm^{-1}(\cdot)) - \mathbf{y}_m\|_2^2 \quad (5)$$

$$\mathbf{x} \leftarrow \underset{\mathbf{x}}{\operatorname{argmin}} TV(\mathbf{x}) + \lambda FT(\mathbf{x}) \quad (6)$$

$$\gamma \leftarrow \left\lceil \frac{\ln(tm^{-1}(\mathbf{x}))}{\ln(3)} \right\rceil \quad (7)$$

$$tm(\cdot) \leftarrow \left(\frac{\cdot - \min(tm^{-1}(\mathbf{x}))}{\max(tm^{-1}(\mathbf{x})) - \min(tm^{-1}(\mathbf{x}))} \right)^{\frac{1}{\gamma}} \quad (8)$$

end for

Figure 3.43 shows the original 32-bit image used for the simulations. Its histogram (figure 3.44) presents a saturation in the dark with some very high values. Notice that the contrast of this image is particularly poor.



Figure 3.43 – Original HDR 32-bit image.

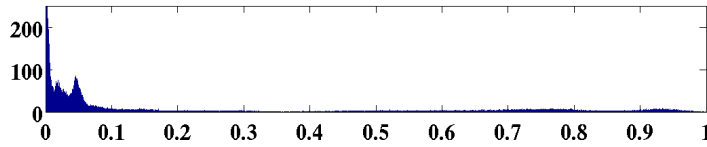


Figure 3.44 – Histogram of the original HDR image.

Figure 3.45 shows the 8 equivalent images corresponding to the different integration times if they were acquired by a common image sensor. The ratio between the shortest integration time T_1 and the longest one T_2 is 1 : 128. Considered as bracketing, it possibly results into a dynamic range enhancement of $42dB$. For the simulations, those 8 images have been compressively acquired with a compression ratio of 1 : 8 to keep the total number of compressive measurements equal to the number of reconstructed pixels.

3.3. The information considered as a group of measurements

Notice that previously estimated electronic noises, technological dispersion and photon shot noise have been considered to simulate the SACSI measurements. This technique thus improves image dynamic range by increasing saturation levels while keeping small quantization steps for low illumination.



Figure 3.45 – The positive 8 images corresponding to each time of integration.

In this simulation, we added noise and pixel non-linearity that have already been discussed. In addition, the photon shot noise that is added to the 8 original images largely impacts on the final quality of the reconstructed image because of being of different level/nature for each time of integration on each pixel. Indeed, photon shot noise “highly” impacts the consistency of the measurements from a time of integration to another. The final reconstructed image is shown in figure 3.46.

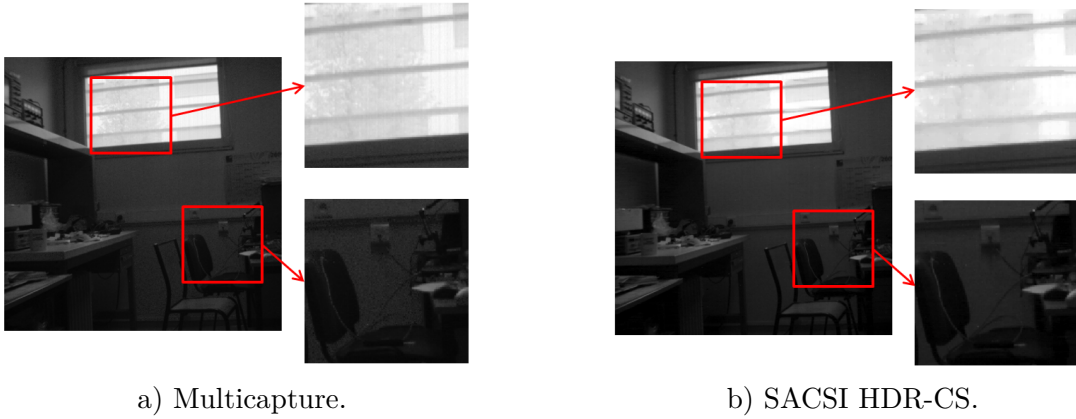


Figure 3.46 – Tone mapped HDR reconstructions. a) is a traditional multicapture reconstruction and b) is the proposed HDR-CS reconstruction.

This HDR-CS technique allows to extend the dynamic range of SACSI CS acquisition without increasing the number of data compared to an uncompressed acquisition. However, this dynamic range extension is being performed but at the extend of some loss on the quality of the image in the light regions (because of being more compressed). In dark regions a high fidelity of the reconstruction is reached thanks to denoising and

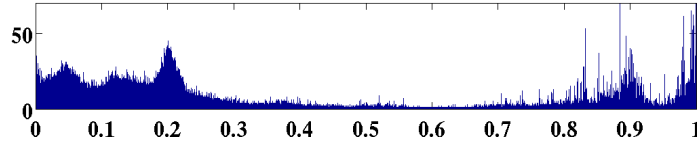


Figure 3.47 – Histogram of the tone mapped reconstruction.

regularization. This technique thus provides a new possible trade-off between dynamic range, image quality and compression ratio. Typically, we have shown that using a specific SACSI sensing scheme, a $40dB$ enhancement of the dynamic range can be performed without increasing the amount of acquired information compared to a single integration time acquisition. Moreover, one can further investigate other (better) time diagrams enabled by SACSI architecture in HDR-CS mode. Indeed, we have stated that better sensing schemes could be designed for instance by combining in each measurement, different pixels corresponding to different times of integration. The final reconstructed images would be improved if the ratio of short time of integration is higher than longer ones. However, in practice, problems (non convergence of the algorithm) can arise when too large mismatches exist between assumptions made on the pixel behavior and the real one. It motivates the following, presenting a way to make the CS reconstruction or more generally a synthesis robust against differences existing between the acquisition model and the way the signal is really acquired.

3.3.4 Artifacts and non-linearity compensation

Generally, there are unfortunately many differences between the behavior of a simulated analog circuit and the really built one. Significant mismatches are commonly observed between a test-chip and its simulations. For large analog circuits especially such as imagers, the simulations can only take into account a part of the design because of simulation time and memory issues. In the case of CS image reconstruction or image synthesis, wrong assumptions on the pixel integration will dramatically impact on the final result. Algorithmic convergences troubles arise when the behavior of the pixel is different from the one which assumed. This problem is even amplified when multiple sensing parameters (integration times in our case) are implied. Moreover, the model parameters of the non-linear behavior are not accurately known because of components variability due to technological dispersions and approximative transistor model simulations. It generally severely degrades the reconstruction accuracy and even can prevent the reconstruction algorithm convergence. A compensation regularization technique is proposed to reconstruct an image of good visual rendering, suppressing possible artifacts induced by wrong assumptions. The proposed algorithm 3 is composed of a three steps loop: regularize the image, fit the model parameters describing unknown pixel behaviors and update the regularization coefficient for the next iteration. At the first iteration when using a basic TV regularization, the first reconstructed image is

3.3. The information considered as a group of measurements

largely blurred because of a small weight on the fidelity term, then a first approximation of the model parameters is performed and finally regularization parameters are updated enlarging the weight on the fidelity term for the next iteration. Therefore, the algorithm iteratively refines the model parameters evaluated on images which are less and less regularized.

Algorithm 3 Image restoration with partially known acquisition model parameters

```

while Stopping criterion is not reached do
    Reconstruct or denoise image by regularization
    Find the best fitting model parameters
    Update regularization parameters
end while

```

In particular, in the case of imagers the pixel behavior and more largely the non-linearity introduced in HDR mode is partially unknown. This is the reason why we propose to introduce an algorithm compensating those unknown parameters for HDR SACSI reconstruction (algorithm 4). This algorithm is based on the assumption that even if unknown, the pixel integration behavior still follows a predefined spline model. In that scope, we recall the fidelity term depending on the measurements, and the acquisition model parameters (α and τ):

$$FT(\cdot, \alpha, \tau, \mathbf{y}) = \sum_{m=1}^M \|\Phi_m f_m(tm^{-1}(\cdot)) - \mathbf{y}_m\|_2^2, f_m(\cdot) = f(\cdot, \alpha_m, \tau_m). \quad (3.23)$$

To find the best fitting model parameters α and τ , a greedy algorithm is employed. The approximative outputs are sufficient to let the algorithm converge to an appropriate solution for restoring the image from an iteration to the next one.

Algorithm 4 HDRCS reconstruction with spline model parameters compensation

```

while convergence of  $\mathbf{x}$  is not reached do
     $\mathbf{x} \leftarrow \text{HDRCS}(\mathbf{x}, \lambda, \alpha, \tau)$ 
     $\{\alpha, \tau\} \leftarrow \underset{\alpha_m \in 1..M_f, \tau_m \in 1..M_f}{\text{argmin}} FT(\mathbf{x}, \alpha, \tau, \mathbf{y})$ 
     $\lambda \leftarrow 2\lambda$ 
end while

```

Figures 3.48 and 3.49 show an example of reconstructions respectively performed without and with a compensation on the model parameters. For this simulation we use an input error of 5% on those parameters for the initialization. As one can observe, the proposed compensation algorithm clearly outperforms a basic reconstruction. Even if the fidelity of the reconstruction is not perfect because of the modifications on both the image and the model parameters, the final visual rendering is largely improved. Notice that these

results are based on a simulation where CS measurements are performed without noise.



Figure 3.48 – Reconstruction without compensation of the model parameters.



Figure 3.49 – Reconstruction with compensation of the model parameters.

The algorithm presented above takes advantage of the fact that mismatches between model parameters of the acquisition and those that are used at the reconstruction creates artifacts reducing the level of sparsity or the smoothness of the reconstructed image. By performing multiple iterative regularizations, the image is made sparse, until reaching a convergence point. This is made possible by enforcing priors on the image more through iterations. For the moment, there is no proof neither on the convergence nor on the convexity of this kind of algorithms. However experimental results validate the

improvements related to this kind of algorithm. For instance, figure 3.50 shows the comparison in terms of reconstruction PSNRs between both algorithms, with and without compensation of the model parameters.

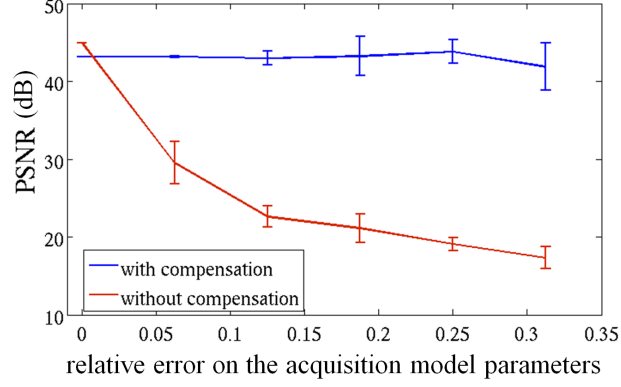


Figure 3.50 – Image restoration simulation results in terms of PSNR in function of the relative error on the acquisition model parameters. Standard deviations of the reconstructed PSNR for multiple trials is represented by the vertical bars.

Conclusion

As it has been shown in this chapter, the SACSI can address multiple applications related to CS. In particular the flexibility on the control of the activated pixels and their resets to form particular time diagrams and thus specific sensing schemes is an advantage of this image sensor. The SACSI benefits from its high level of reconfigurability (CS HDR mode, CS high frame rate, uncompressed, etc.). Regarding the SACSI in the field of sensor networks, this reconfigurability is an asset in order to enable commanding the working mode by external routines. More than a technique to relax sensing limitations, CS should be considered as a pragmatic and powerful tool to define compromises regarding the different acquisition characteristics of images. On the other hand, one trend in sensor networks is to build and integrate smarter sensors. On the contrary, CS do the opposite deporting huge processing loads to the decoder parts leading to sensors without any intelligence corresponding to “simple” information transducers. One possible application of this imager is for wireless CCTV. First, it would considerably reduce the number of data to transmit at each sensor node (in CS mode). Secondly, specific time diagrams can be designed to appropriately sense the time information as the dynamic information, leading to a video HDR-CS system. The next chapter deals with a new way to implement CS in an alternative way to render CS sensors smarter keeping in mind the perspective of sensors networks.

4 Adaptive Compressive Sensing imager

CS measurements generally do not provide as much understandable information as traditional compressed domains. On the other hand, with the rise of CS dedicated CIS, it becomes necessary to define relevant sensing schemes in terms of on-chip implementation but also efficient regarding the reconstruction. For instance, the SACSI sensing scheme is particularly well suited because it can be used just as an alternative sensing mode to the traditional rolling shutter without a considerable amount of on-chip supplementary materials. Applying the same linear projection to reduced signal supports (generally known as block-based CS) also optimizes the reconstruction time and makes the sensing strategy more suitable at the sensor level. This chapter aims at defining basic procedures to efficiently adapt the sensing strategy directly in the image sensor chip. Those procedures are based on local sparsity estimation to adapt the sensing scheme accordingly. It mainly refers to specific features which are computed before the CS, optimizing the sensing operation and helping at the reconstruction. Since those additional measurements become practical on a real sensor at the focal plane level, we will show that it also improves the use of CS by providing some useful inputs for high level applications (i.e. tracking and face recognition). In addition, those features significantly upgrade the reconstruction quality when considered as CS measurements. The raised question can be formulated as follow: which tools CS can provide to design a smart imager and is there a way to define simple procedures for making a CS imager self-reconfigurable. The underlying goal being to prospect how to add a little bit of adaptivity and intelligence at the sensor level in the case of a CS imager without keeping in mind power consumption issues regarding the overall sensing stage.

This chapter is organized as follow: first we motivate the use of adaptive CS due to the intrinsic nature of natural images and provide examples of application of such a sensing scheme, secondly we propose a top-level implementation for the local block variance adaptive CS, finally we present different applications taking advantage of the already on-chip computed local features.

4.1 Adaptive acquisition and Compressive Sensing, the trade-off

One trend in CS consists in providing alternative ways to acquire the information. Moreover, the complexity of the sensing scheme plays an important role for designing a sensor but also for the reconstruction algorithm. Many of those limitations are reduced adopting a block-based approach [137]. This approach works those limitations out by parallelizing the problem and multiplexing available resources. Indeed, a block-based implementation consists of applying the same sensing matrix on each sub-support of the signal (i.e. as SACSI works). It obviously simplifies the acquisition method. At the reconstruction stage, a fast implementation can be found taking advantage of redundancy in the whole corresponding sensing matrix. Some works are even performing the reconstruction separately on those supports before combining them to form the final signal. However, those techniques often suffer from blocking artifacts. This chapter focuses more on how to adapt the number of measurements performed on each blocks. Since a block-based approach is adopted, it is unfair to acquire the same number of measurements on each block because of the sparsity. For instance, the classical JPEG compression standard typically takes advantage of block DCT entropy variation by compressing each block with an adaptive ratio. Indeed, the entropy coding performed on the resulting transform coefficients shows that the required data bandwidth varies from an image block to another.

4.1.1 Sparsity estimator, block variance

In general, the image sparsity can largely vary from an image to another. Regarding block-based processing, the sparsity also varies in the same image from a block to another. This sparsity level generally depends on the block size too. The size of blocks used for CS implies to limit the theoretical efficiency of CS when both acquisition and reconstruction techniques are relaxed because of possible parallelizations. When block-based CS is preferred, a lot of information is lost because of taking the same number of CS measurements for each block. Starting from the idea that the implementation of a sparsity estimator using a structured transform at the sensor level is not relevant because of its complexity, an alternative, correlated estimator of the sparsity based on a statistical character is thus employed for this purpose. In particular, this statistical character corresponds to the block variance. For a large set of greyscale images and for different size blocks we evaluate both the block variance and its compressibility (see figure 4.1 and figure 4.2). Figure 4.1 presents a point cloud for TV , TV_{4axes} , $Haar$ coefficients $l1$ norm and $l1$ norm in the D6 space of horizontal and vertical gradients in function of the block standard deviation (std). In particular, for these $l1$ norm operators, the DC component has been removed to provide more relevant results. The density of those points reflects the sparsity property of the image blocks. Still, alternatively, the degree of sparsity (compressibility) has been defined as the entropy of coefficients using the same specific operators as previously used for $l1$ norms. Figure 4.2 presents the same

results as figure 4.1 but in the case of this alternative compressibility definition. In these two figures, the drawn points are more concentrated close to the origin. We can also observe that there exists a noticeable correlation between the block standard deviations and the estimations of the compressibility (thus the sparsity) by block that have been considered. The larger the block size the sparser the points are when far from the origin.

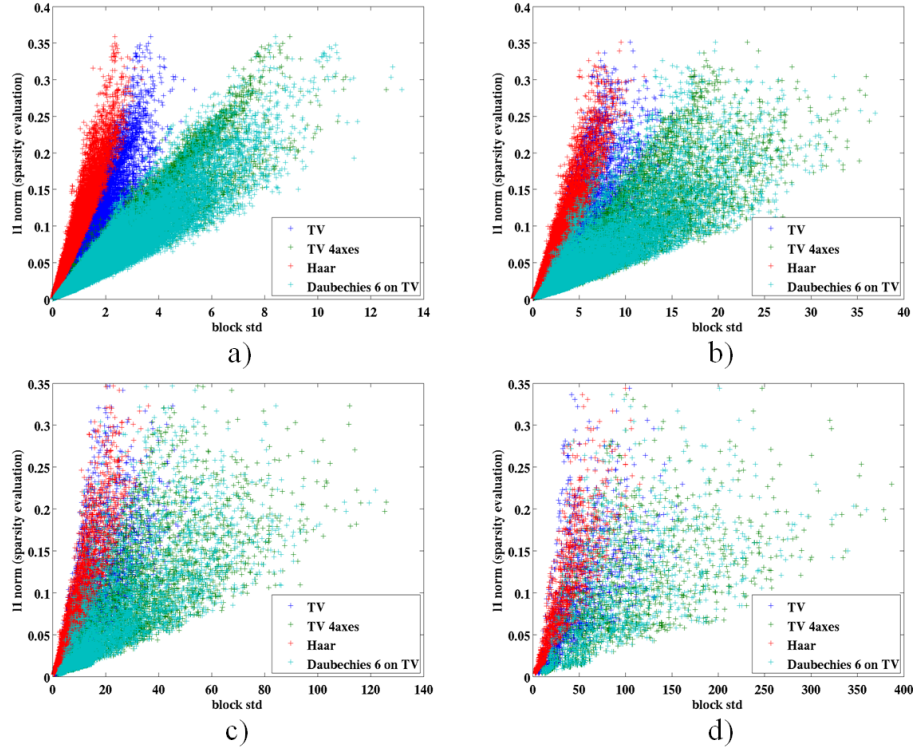


Figure 4.1 – Analysis of block standard deviation and block operator $l1$ norms (compressibility estimator) from a set of test images and for four different block sizes, a) 4×4 , b) 8×8 , c) 16×16 , d) 32×32 .

Figure 4.3 represents the Pearson correlation coefficients between the image block standard deviations and the previously presented compressibility estimators for different block sizes. This coefficient is relatively high meaning that the block standard deviation is largely correlated with the degree of sparsity. Yet, this consideration remains relevant only regarding natural images. We can observe that this correlation is generally decreasing with the block size increase. In particular, for image blocks larger than 16×16 , this coefficient is smaller than 0.7 whatever the estimator considered. However, this study is based on 512×512 pixel images and these presented results might for higher resolution images. Indeed the size of block for which the standard deviation is correlated to compressibility would reasonably be larger.

The problem can be stated as follow: if we were able to evaluate and characterize the sparsity level in each processed block we would probably be able to perform a traditional

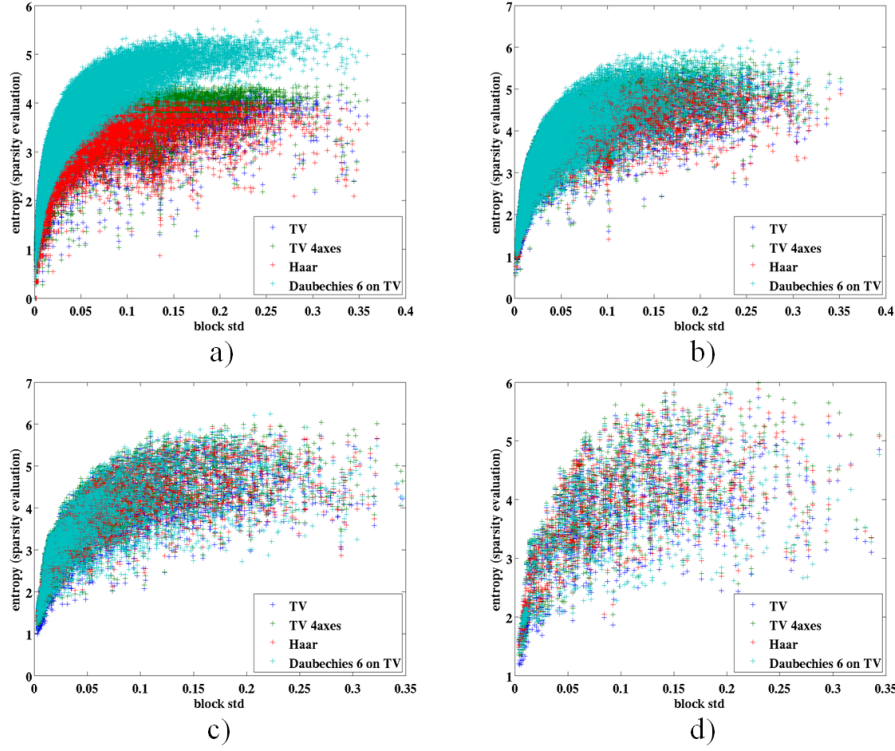


Figure 4.2 – Analysis of block variance and block coefficients entropy (compressibility estimator) from a set of test images and for four different block sizes, a) 4×4 , b) 8×8 , c) 16×16 , d) 32×32 .

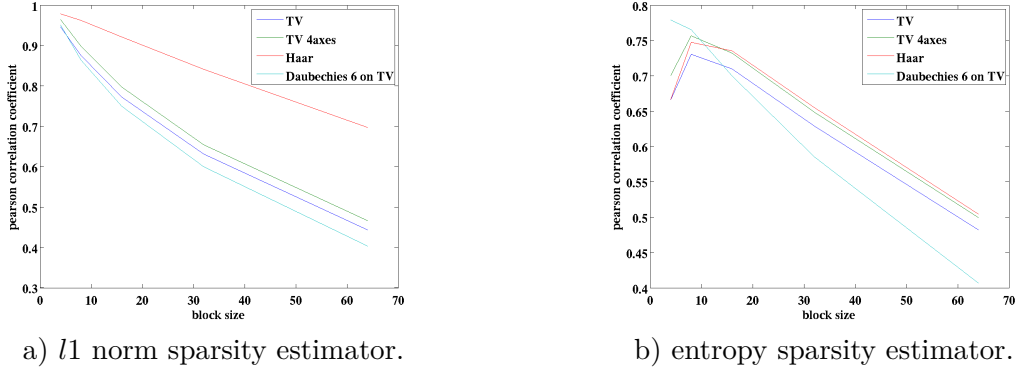


Figure 4.3 – Correlation between image block variance and its sparsity in function of the block size (side of a square block).

lossy compression based on this specific sparsity basis. However, it would require to embedded demanding computations. In the case of adaptive CS, we thus need to define features which evaluates the degree of sparsity without implying much computations and complex processing. In that scope and according to this preliminary study, we can reasonably expect that block variance is a good estimator of the block sparsity.

4.1.2 Block sparsity variability

Adaptive CS has not been deeply studied since aims of CS is theoretically to be absolutely generic and non adaptive. Yet, since block-based CS is employed it is absolutely not relevant to suppose that each image block contains the same level of information and thus the same degree of sparsity. But still, the main interest of block-based CS is the efficiency of the acquisition implementation and the reconstruction speed up. A first compromise must be taken because the smaller the blocks the bigger the sparsity ratio and on the other hand the bigger the blocks the more processing requirements. The same trade-off applies at the reconstruction favoring either the processing time using small blocks or the reconstruction quality versus the compression ratio using larger blocks. Figure 4.4 shows the variability of the compressibility estimators in function of the block size. One can observe in this figure that the smaller the blocks the more variable the compressibility. Indeed, the interest of using adaptive CS is increased when small blocks are involved meaning that the compressibility considerably varies between blocks.

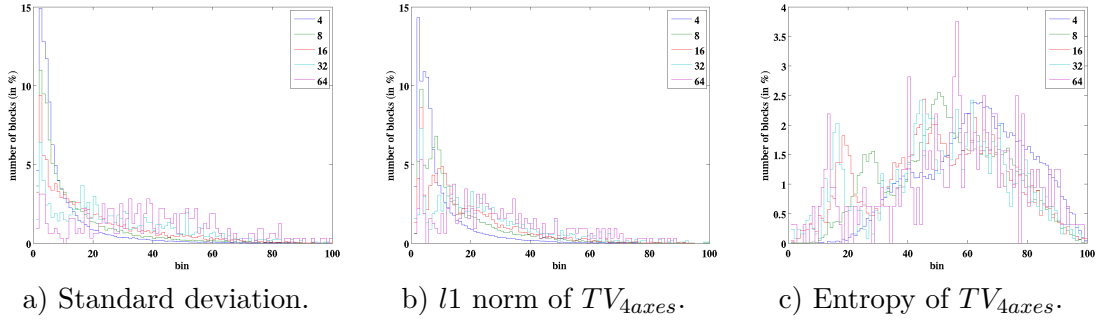


Figure 4.4 – Histogram of distribution for the different block compressibility estimators and block sizes.

In the following, this work attempts to demonstrate that block spatial variance can properly estimate the number of required measurements for each block in order to improve the quality of image at a constant compression ratio. In particular, this statistical feature identifies and highlights edges and high contrast regions which are often corresponding to less sparse blocks in the case of natural images.

4.1.3 Adaptive CS, the use of block variances

Block adaptive CS is just an extension of block-based CS where the number of measurements performed by block is tuned according to some external information. In our case this external information is a feature which has been preprocessed on the block. More particularly and thanks to the preliminary study depicted above, the block variance is used as block feature. For more formalism, Φ_{cs} denotes the corresponding full CS measurement matrix employed according to the adaptability of the number of measurements performed by blocks. This matrix is thus composed of a concatenation of block

sensing matrices as presented in the previous section for block-based CS. Yet, those block sensing matrices now correspond to the concatenation of a selected number of rows of a square matrix Φ_B which is common for all the blocks. The adaptation is thus performed by setting these numbers of rows independently for each block. Adaptive CS strategy appears appropriate since the necessary number of measurements theoretically depends on the sparsity and because it has been shown that this sparsity (or compressibility) highly depends on the blocks. The interests of using those specific statistical characters (block variances) are multiple. Hence, we propose to use the variance in order to evaluate at low cost the sparsity of image blocks [94]. Those variance measurements can also help at the reconstruction stage as additional measurements to the traditional CS ones. In any case, it is necessary to send the block variance estimations to the decoder in order to properly rebuild the CS matrix Φ_{cs} before the reconstruction because they have been used to control the number of measurements performed by block support. In that sense, those additional measurements increase the number of data thus reducing the compression ratio. However, as will be presented, it still ameliorates the acquisition considering a same total amount of stored/transmitted data.

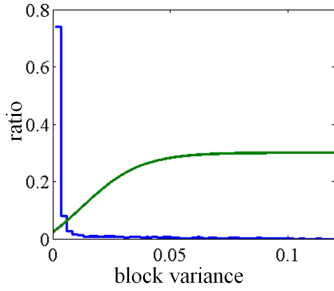


Figure 4.5 – Histogram of block variances for the cameraman image (blue). Ratio of the number of measurements to perform (green)

For natural images, spatial block variances are strongly correlated with total variation since block sizes are small enough. In Figure 4.5, the blue curve represents the distribution of the block variance values for cameraman. For compression purpose, we therefore propose to attribute a certain number of measurements to a level of variance. This is represented by the green curve in Figure 4.5. This green curve represents a sigmoid function. A sigmoid function is being considered because of allowing to define low and high thresholds to limit the number of measurements in high variance regions and set a lower bound for very low variances. The transition region of this curve can be tuned as well by controlling the slope. Three degrees of freedom are available to define the ‘best’ shape for a specific acquisition. In addition, based on the previously stated considerations, the compromise we found tends to favour large blocks for applying the CS

measurements and smaller blocks for evaluating the block variances. Therefore, the function used to derive the number of measurements is applied to averaged values of block variances over non overlapped groups of 4×4 blocks of 8×8 pixels. This trade-off is very important because it increases the size of the blocks on which CS measurements are performed keeping a finer spatial resolution for variance estimations. It finally gives the number of measurements by 32×32 pixel block. As example, figure 4.6 shows the simulation of the resulting number of measurements that have to be performed by block.

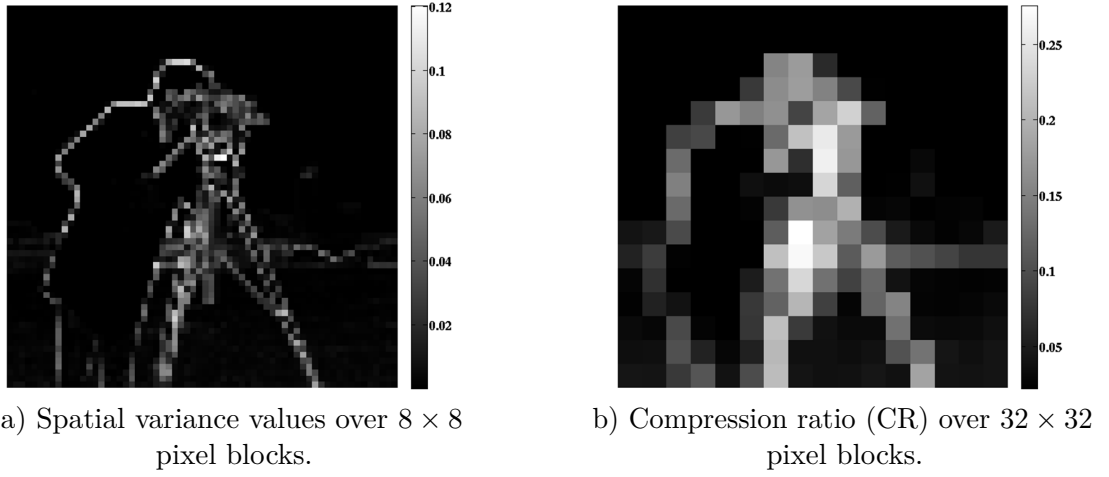


Figure 4.6 – Example of extracted features for the cameraman test image.

The use of additional non-linear compressive measurements

Since an adaptive sensing strategy is adopted, the number of measurements performed by block is required at the reconstruction to rebuild the CS matrix Φ_{cs} used for the acquisition. This can be evaluated using the value of 8×8 pixel block variances and simulating the acquisition procedure. The proposed reconstruction method also considers these values as non-linear measurements [19]. The reconstruction problem that is stated as equation 4.1 can thus be summarized as a generalized CS problem.

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} Op(\mathbf{u}) + \lambda FT(\mathbf{u}, \mathbf{y}_{cs}, \mathbf{y}_v) \quad (4.1)$$

The measurements on the image \mathbf{x} are thus composed of block-based CS measurements that we consider for the moment as Gaussian White Noise measurements (\mathbf{y}_{cs}) and block variances (\mathbf{y}_v). At the reconstruction, those two groups of measurements are used to form the fidelity term $FT(\mathbf{u}, \mathbf{y}_{cs}, \mathbf{y}_v)$ defined by the equations (4.2) and (4.3).

$$FT(\mathbf{u}, \mathbf{y}_{cs}, \mathbf{y}_v) = FT_v(\mathbf{u}, \mathbf{y}_v) + FT_{cs}(\mathbf{u}, \mathbf{y}_{cs}) \quad (4.2)$$

$$FT(\mathbf{u}, \mathbf{y}_{cs}, \mathbf{y}_v) = \|\Phi_v(\mathbf{u}) - \mathbf{y}_v\|_2^2 + \|\Phi_{cs}\mathbf{u} - \mathbf{y}_{cs}\|_2^2 \quad (4.3)$$

Φ_v can be decomposed as the second order moment (equation (4.4)) if Φ_m denotes the

linear operator performing block means. Φ_m thus returns a vector of the block means when multiplied by \mathbf{x} . The Hadamard product (or point wise product) noted (\cdot) is thus used to redefine Φ_v with linear operators only.

$$\Phi_v(\mathbf{u}) = \Phi_m \mathbf{u}^{\cdot 2} - (\Phi_m \mathbf{u})^{\cdot 2} \quad (4.4)$$

$$FT_v(\mathbf{u}) = \|\Phi_m \mathbf{u}^{\cdot 2} - (\Phi_m \mathbf{u})^{\cdot 2} - \mathbf{y}_v\|_2^2 \quad (4.5)$$

The gradient of FT_v can be derived from this equation in order to be used in a gradient based minimization algorithm. In addition we empirically observe that when enforcing signal positivity, it ensures convexity and thus convergence.

$$\frac{dFT_v(\mathbf{u})}{d\mathbf{u}} = 4\mathbf{u} \cdot \Phi_m^T \left(\Phi_m \mathbf{u}^{\cdot 2} - (\Phi_m \mathbf{u})^{\cdot 2} - \mathbf{y}_v \right) - 4\Phi_m^T \left((\Phi_m \mathbf{u}) \cdot \left(\Phi_m \mathbf{u}^{\cdot 2} - (\Phi_m \mathbf{u})^{\cdot 2} - \mathbf{y}_v \right) \right) \quad (4.6)$$

Results and discussions

In the following simulations, Φ_B is composed of Gaussian random measurement vectors. In particular, figure 4.7 presents the cameraman reconstructions using the operator $mDWTonTV$. The global visual quality of this reconstruction is remarkably good knowing that the compression ratio is below 8%. In the rest of this chapter and for the sake of consistency, all the compression ratios denoted by CR are evaluated between 8-bit original images and measurements performed using a 8-bit quantization (except when it is explicitly expressed). Notice that the edges are well preserved without ringing artifacts. The sharpness of the face and camera are particularly well restored. The most important point is that salient areas are not damaged at all by compression effects even if for instance the visual rendering of the grass ground of this image is somehow degraded.

Results in terms of PSNR for several test images are reported in table 4.1. This figure shows that the PSNR is increased when adaptive strategy is adopted against non adaptive CS at the same compression ratio CR_1 . CR_2 denotes the compression ratio related to the variance measurements, the resulting compression ratio when adaptive CS is employed is thus of $CR_1 + CR_2$. Those results also show a clear improvement when using $mDWTonTV$ against the two other operators. The use of block variances as measurements improves the final rendering as well. However, the improvement related to

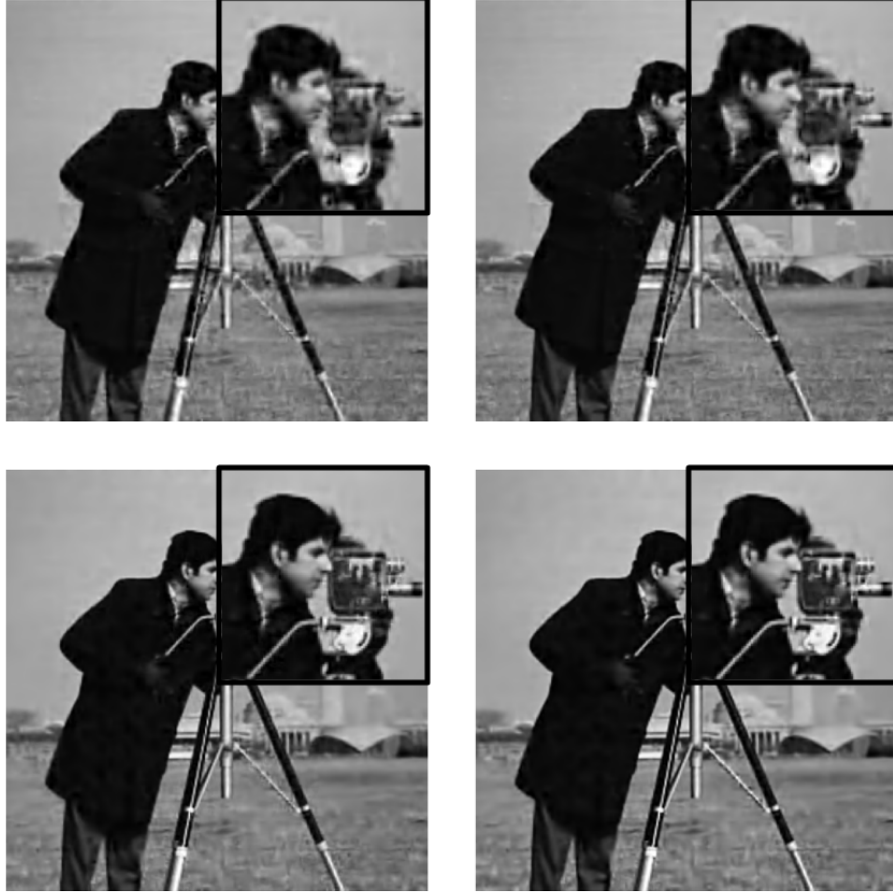


Figure 4.7 – Reconstructed Cameraman image using $mDWTonTV$. From left to right, top to bottom: Regular CS (CR_1), Regular CS ($CR_1 + CR_2$), Adaptive CS (CR_1) and Adaptive Var-reg ($CR_1 + CR_2$).

this remarkably depends on the image. For instance, regarding Cameraman and Darkhair images, this improvement is large because of the main edges and the localization of the sharpness. On the contrary, for Blonde and Pirate this improvement is not that high because of the complexity of those test images (related to the smaller variability of the compressibility of image blocks). Results in terms of SSIM are also reported in table 4.1. The SSIM assessment globally leads to the same conclusions as the PSNR. However, the function used to derive the number of measurements has been adapted to the histogram of cameraman which is probably not optimized. For better results this function should be adaptive as well to properly fit with the image histogram. Even if not adapted at the acquisition, for a specific application this function should be predefined and designed with respect to potential content of the image (block sparsity histogram).

In this section, we thus experimentally demonstrate that block variance is a good candidate for tuning the number of measurements in the context of adaptive block-based CS. It is also shown that block variances can be directly used as non-linear CS measurements

method		TV		TVmDWT		mDWTonTV	
	CR	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
	Cameraman						
Regular CS (CR_1)	6.2%	25.4	79.8	24.5	78.1	27.1	83.1
Adaptive CS (CR_1)	6.2%	28.5	82.9	27.9	81.9	30.5	86.1
Regular CS ($CR_1 + CR_2$)	7.8%	26.2	81.8	25.5	80.4	28.2	85.5
Adaptive Var-reg ($CR_1 + CR_2$)	7.8%	29.3	84.4	29.6	85.1	31.1	87.1
	Darkhair						
Regular CS (CR_1)	3.3%	28.9	82.9	28.7	85.4	31.5	87.8
Adaptive CS (CR_1)	3.3%	28.7	81.2	30.1	86.5	33.0	88.7
Regular CS ($CR_1 + CR_2$)	4.8%	30.4	85.2	30.5	87.1	32.9	89.3
Adaptive Var-reg ($CR_1 + CR_2$)	4.8%	30.6	84.7	31.3	87.8	33.4	89.3
	Blonde						
Regular CS (CR_1)	5.3%	24.7	64.7	24.4	66.0	25.4	67.9
Adaptive CS (CR_1)	5.3%	25.6	66.2	25.4	67.5	26.4	69.5
Regular CS ($CR_1 + CR_2$)	6.9%	25.3	66.9	25.1	68.0	26.0	70.2
Adaptive Var-reg ($CR_1 + CR_2$)	6.9%	25.8	67.7	26.1	71.0	26.6	71.1
	Pirate						
Regular CS (CR_1)	5.7%	24.8	63.6	24.1	62.1	25.3	66.7
Adaptive CS (CR_1)	5.7%	25.5	64.7	24.6	62.9	26.0	67.6
Regular CS ($CR_1 + CR_2$)	7.2%	25.6	66.6	24.8	64.8	25.9	69.5
Adaptive Var-reg ($CR_1 + CR_2$)	7.2%	25.7	66.9	25.4	67.8	26.1	69.4
	Lena						
Regular CS (CR_1)	5.0%	25.9	73.4	25.4	73.8	27.0	77.1
Adaptive CS (CR_1)	5.0%	27.0	75.2	26.7	75.9	28.5	79.4
Regular CS ($CR_1 + CR_2$)	6.6%	26.7	75.6	26.3	76.0	27.9	79.4
Adaptive Var-reg ($CR_1 + CR_2$)	6.6%	27.2	76.4	27.7	79.1	28.7	80.8

Table 4.1 – PSNR & SSIM reconstruction results for several test images.

for the reconstruction. In addition, such a statistical feature could be implemented in the analog domain before the analog to digital conversion into an image sensor. Those features can also be used as features for tracking or classification purposes as it will be presented at the end of this chapter.

4.2 Top-level implementation of an adaptive CS imager

The design of an adaptive CS architecture is also made of compromises. Indeed, adaptive CS implies the computation of the features allowing the number of measurements by block adaptation, before performing the effective CS measurements. Regarding the image sensor constraints, this added operation should not be too much complex and silicon area consuming. Moreover, this feature extraction should neither imply to reduce the frame rate of the imager nor dramatically impact on the overall power consumption. Multiple architecture implementations for adaptive CS have been investigated, particularly regarding the variance calculation. First, an implementation in the focal plane inside the array of pixels have been considered using analog computations based on the quadratic current behavior of the MOS transistors in the saturation region. This structure could use a specific pixel structure using two distinct current outputs combined with biases and a CSC structure [156] to directly compute the variance value at the block of pixel level. However several problems arise regarding this technique. First, technological

dispersions dramatically impact on the resulting calculation or should be handled by design at the expense of increasing the size of the transistors. Even if being compact, this structure implies to connect together each pixel belonging to the same block thus leading to the increase of the number of wires. Then, adding a block structure in the array of pixels would also increase the FPN. For all of those reasons we opt for an end-of-column implementation leading to the architecture presented in figure 4.8. This architecture employs first order incremental $\Sigma\Delta$ s as ADCs such as in [93]. The interest of using this ADC is that it intrinsically performs the average of a variable input by converting -during a time of conversion- the DC value of its input. This property is used for both variance calculation and CS.

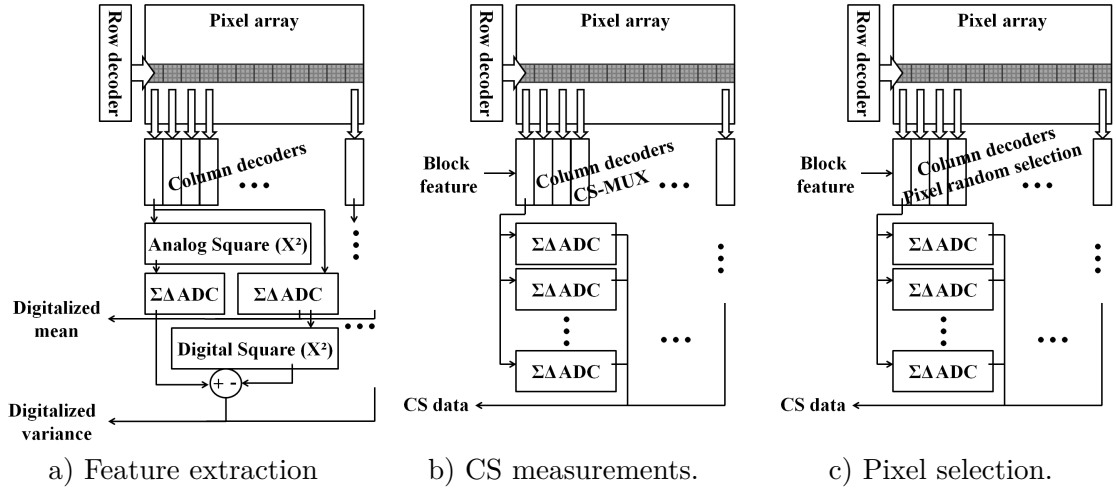


Figure 4.8 – Top view of the proposed CIS architecture. Three functioning modes (a) b) c)) are enabled. All of those mode use reconfigurable processing blocks (first order $\Sigma\Delta$ s, column decoders).

4.2.1 Variance calculation

This top-level architecture has the advantage of keeping the traditional structure of a CIS. The pixel used by this architecture can be of a canonical nature such as 3T or 4T pixels. To perform the so called adaptive CS acquisition, a compact implementation for block variance calculation is needed. The processing blocks used for the two acquisition steps, a) and then b) or a) and then c). It is composed of a row decoder and a column decoder used to properly address and read requested pixels of currently processed blocks for the two acquisition steps. For each row of blocks that are read simultaneously, two successive acquisitions are performed. During the first step, all the pixels are sequentially read to perform variance calculation. However, this operation is interestingly performed during a single conversion time of the first order $\Sigma\Delta$ s (i.e. in 2^{bits} where $bits$ corresponds to the number of bits of resolution). In this way, the variance of each block of the activated block row is acquired and can be exploited for the next step. The proposed circuit evaluates the variance by computing for each block B : $var(B) = \overline{pix^2} - \overline{pix}^2$ where pix

corresponds to all the pixels belonging to B . This structure has been chosen because it allows to evaluate block variances of a row of blocks during a single conversion time and without intensive processing. The analog squarer unit is based on the schematic presented in [157], and has been optimized to maximize the voltage input range (from $592mV$ up to $1.385mV$) and the output range when preserving at most the quadratic law over the entire dynamic. The digital squarer has been inspired from the work presented in [158] employing a rearrangement of digital circuitry to optimize the squaring operation compared to a traditional multiplication. Indeed many simplifications can be performed on the digital operations when modifying a traditional booth multiplier to a squaring operator. Alternatively the digital squarer proposed in [159] could be employed regarding low power issues. This joint analog/digital processing circuit is relatively compact and can fit within the classical sensor end-of-column. In particular, we evaluate for a 10bit squarer the number of gates to 136 AND, 89 XOR and 44 OR. Notice that these numbers of gates are considerably reduced when the number of bit of resolution is only 8. Figure 4.9 shows resulting errors at the outputs when comparing for a fixed input the two distinct paths: analog squaring then conversion and conversion then digital squaring. It results that the maximum acceptable resolution of the variance would be of 8 bits which is sufficient for the targeted application.

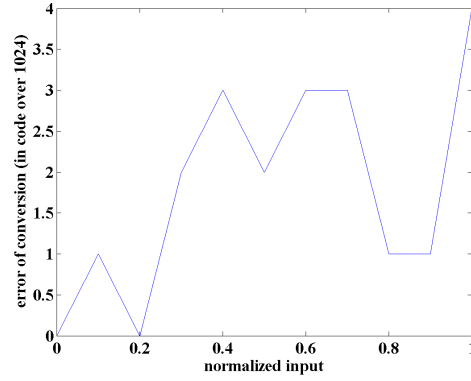


Figure 4.9 – Error in number of codes for a resolution of 10bits between the conversion of an analog squared input its the digital squaring after a linear $\Sigma\Delta$ conversion.

4.2.2 Adaptation and variable density sampling

Since the variance calculation is done, the second step of block acquisition can start either using a traditional block-based CS with adaptive number of measurements or using a random pixel selection with adaptive expectancy. Due to the fact that the imager architecture is close to the one depicted in [93], a customized CS-MUX (figure 4.8 b)) can easily be employed to perform CS measurements but stopping the number of conversion according to the precalculated variance. Figure 4.8 c) yet proposes an alternative sensing strategy based on capturing random pixels from each image block with an adaptive expectancy. This sensing strategy consists in adapting the number of acquired pixels

(CR in this case) by block according to the block variance. It dramatically facilitates the acquisition by removing a complex multiplexer replaced by a cellular automaton allowing the selection or not of each pixel, sequentially. As proposed in the previous section, this acquisition is an adaptive CS which is not a Block-based CS. Indeed, the selected pixels can be of different indexes in each block even if the compression ratio is the same. Assuming that natural images are sparse in the Fourier domain this sensing scheme is particularly efficient (because of the incoherence between Fourier basis and Spike basis). Such a compressed acquisition using an inpainting like reconstruction has been proposed in [160] but requires a complex edge detection and image analysis which is not the case in the proposed acquisition method. However, the proposed architecture still uses block variances as features for identifying blocks containing significant information such as edges and texture details. As for SACSI, a pseudo-random binary number generator based on a cellular automaton with chaotic behavior (Wolfram's Rule 30) is employed. The expectancy of this bit generator (the ratio between 1 and 0 outputs) can be set to $\{1/2, 1/4, 1/8, 1/16\}$ according to a 2-bit value. It is performed using an AND-gate layer working as an expectancy reducer (see figure 4.10). This 2-bit value is obtained by a 3-level thresholding of the 10-bit (or 8-bit) variance output that has been computed just before for this specific block. Indeed, the 2-bit quantized variance is then used to select the output of the pseudo random generator (out2, out4, out8 or out16). During this second step, each pixel row of the block is read sequentially just as a standard row-parallel rolling shutter readout. The selected binary outputs of the pseudo-random generator are thus used to activate or not the conversion of the currently read pixel. Between each row readout, the cellular automaton needs to be updated to activate different rows. A specific initialization can also be addressed to each block independent cellular automaton in order to decrease the correlation between of activated pixels.

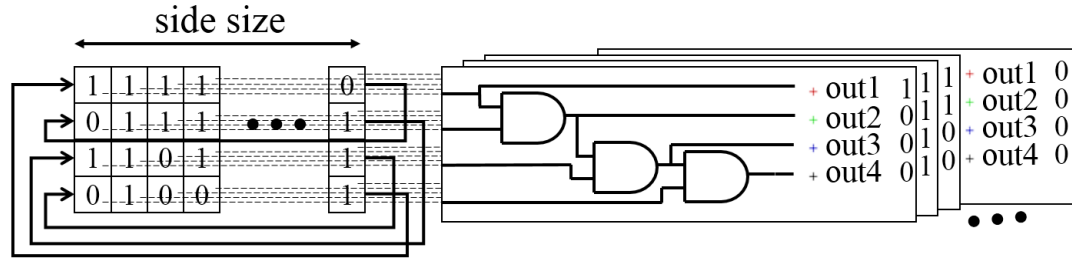


Figure 4.10 – A column decoder performing adaptive pixel random selection associated to a single column of blocks.

Figure 4.11 presents the resulting expectancy of activation of the proposed pseudo random generator for several block sizes. This simulation shows that the number of selected pixels can efficiently be tuned using the different outputs of the structure presented in figure 4.10. In particular, the proposed architecture can be composed of 4×4 pixel blocks. Finally, each column of blocks is thus associated to a pseudo random generator, an analog squaring circuit, 4 $\Sigma\Delta$ ADCs, a digital squaring circuit and a digital subtracter.

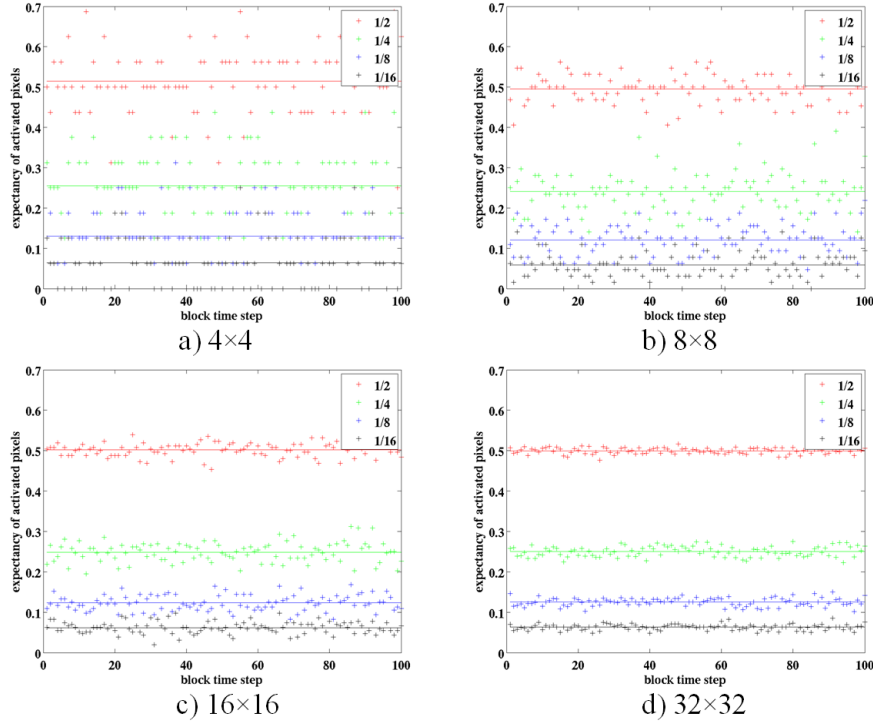


Figure 4.11 – Expectancy of pixel selection in function of the output of the column decoder with pixel random selection. 1/2, 1/4, 1/8 and 1/16 respectively represent the outputs out2, out4, out8 and out16 presented in figure 4.10. The different graphs are for different block sizes.

This adaptive random selection of the pixels has two main advantages. The first one is related to the simplicity of its implementation at the end-of-column circuitry. Indeed the proposed architecture is easily modular to be designed for any block size to perform adaptive CS based on fixed quantizations of the block variances. Secondly, it provides CS measurements which are some how easier to deal with. In particular, the following section presents different applications taking advantage of this architecture and its different outputs (both features and CS measurements).

4.3 Specific adaptive CS applications

In this section, three different applications using outputs of the previously presented image sensor architecture are depicted. First, a fast reconstruction algorithm based on the use of randomly selected pixels of images and videos (voxels) is presented. Secondly, an object/person tracking algorithm which is working with the features provided by the imager is presented. Finally, the face recognition problem is addressed when pictures are taken using adaptive CS.

4.3.1 Compressive sensing versus pixel random selection

Recently developed Compressive Sensing image sensor architectures tend to provide compact on-chip implementations of alternative acquisitions. On the other hand, the time of reconstruction generally limits possible applications taking advantage of those sensing schemes. An adaptive CS CIS provides a compromise between the sensing scheme efficiency for relaxing on-chip constraints and the reconstruction complexity/quality. For instance [161] proposes a transmission system that scan in a chaotic way the pixels over an image to extract relevant image features. In our case, one possible sensing scheme allowed by the proposed architecture performs an adaptive block-based sensing which is particularly well suited for video acquisition and high level applications.

Reconstruction for still images

The proposed acquisition scheme provides a sort of CS measurements (randomly selected pixels). To demonstrate the relevance of the proposed sensing scheme in terms of CS, it is compared with binary block measurements. We denote this type of acquisition *CS* and the proposed pixel random selection *PS*. In both cases same notations are used. The sensing matrix is noted Φ and the resulting measurements \mathbf{y} . In particular, if the original image is noted I (\mathbf{i} when vectorized), $\mathbf{y} = \Phi \mathbf{i}$. In the case of the proposed acquisition this notation can be replaced by $I_s = \Omega \cdot I$ where \cdot is the Hadamard product operator, Ω is the representation of Φ as a pixel mask and I_s the resulting sensed pixels. An example of such an acquisition is shown in figure 4.12. The CS reconstruction can be performed by solving the regular CS optimization problem of the equation 2.17.

One of the most important issue concerning this kind of algorithm is that it can take a very long time to converge to an accurate solution. The time of reconstruction is often considered as a sticking point regarding CS. Yet, adopting a random pixel sampling, the reconstruction also corresponds to an inpainting problem. A fast inpainting algorithm based on successive convolutions is proposed (i.e. Algorithm 5) for recovering missing pixels. Due to the intrinsic nature of the acquisition where missing pixels are well distributed all over the image, convolution algorithms properly work. Previous works have shown the efficiency of such algorithms [162]. Another interesting point is that since the measurements are stored one can select a wide variety of algorithms to favor either the reconstruction time [163] or the image quality [164]. Much like in [162], the proposed algorithm reported below uses the small convolution kernel h (equation 4.7). At the end, a multi-wavelet filtering is performed on the output image to reduce the induced particle noise.

$$h = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (\text{Von Neumann neighborhood}) \quad (4.7)$$

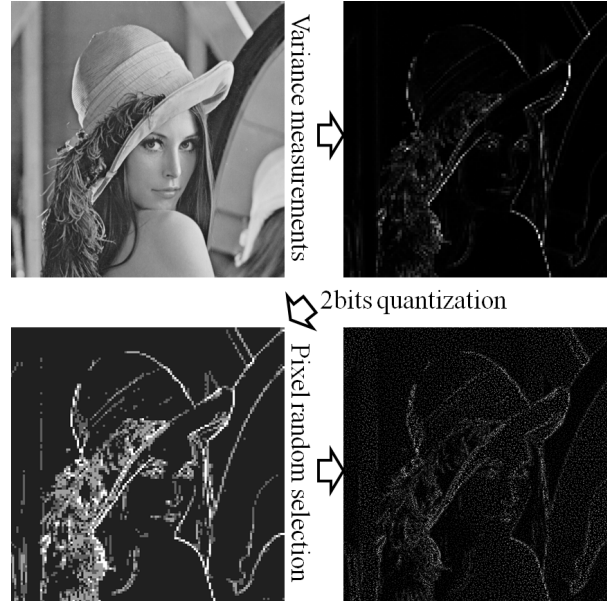


Figure 4.12 – Acquisition steps. From top to bottom, from left to right: the original lena image, the block variances, the 2bits quantized block variances, Resulting sensed pixels I_s (black areas corresponds to non acquired pixels).

Algorithm 5 Fast Inpainting by successive convolutions (*Inpaint*)

```

init  $M \leftarrow \Omega(\text{pixel mask})$  ,  $I_r \leftarrow I_s(\text{sensed pixels})$ 
while  $rel_{err} \geq th$  do
     $I_p \leftarrow I_r$ 
     $I_r \leftarrow I_r * h$ 
     $M \leftarrow M * h$ 
     $I_r(x,y) \leftarrow I_r(x,y) / M(x,y), \forall x, y | M(x,y) \neq 0$ 
     $I_r(x,y) \leftarrow I_s(x,y), \forall x, y | \Omega(x,y) \neq 0$ 
     $M(x,y) \leftarrow 1, \forall x, y | M(x,y) \neq 0$ 
     $rel_{err} \leftarrow \sum_{x,y} |I_p(x,y) - I_r(x,y)|$ 
end while
return  $I_r$ 

```

The inpainting algorithm is working as follows, it first takes into account known pixels and diffuses them following their Von Neumann neighbours. Averaging operations are then performed where it is necessary. If the known pixels have been modified during the previous operations they are replaced by their original values. Since all these operations are done during an iteration, a new iteration starts with a new convolution of the resulting image by h , and so on.

Table 4.2 reports the reconstruction results for 5 different 512×512 test images. Power Signal to Noise Ratios (PSNRs) and reconstruction times are reported in this table. In terms of computation time, the *CS* reconstruction using *TV* requires one minute whereas

the use of $mDWTonTV$ is about seven minutes (without initialization). This difference is mainly due to the complexity of the prior operator. In addition, artifacts arise when $mDWTonTV$ is used with CS because in our case the blocks are very small (4×4 pixels). Notice that this size of blocks is not well suited for CS but for variance evaluation. In the case of PS, $mDWTonTV$ yields far better results than TV in terms of PSNR and especially image quality (as can be observed in figure 4.13).



Figure 4.13 – Reconstructed images for the different sensing-reconstruction couples ($CR = 7\%$). From left to right, top to bottom: original image, PS using *Inpaint*, PS with TV, PS with $mDWTonTV$, CS with TV and CS with $mDWTonTV$.

Finally, the proposed inpainting algorithm *Inpaint* combined with the filtering stage takes only 1.4 seconds at most. In that case, the resulting output quality is still acceptable in terms of PSNR. Considering the visual rendering for an identical compression ratio, only PS using $mDWTonTV$ provides a better reconstruction than PS using *Inpaint*. For instance, compared to results of MS-BCS-SPL (Multiscale - Block-Based Compressive

method	<i>CS</i>		<i>PS</i>		
	<i>TV</i>	<i>mDWTonTV</i>	<i>TV</i>	<i>mDWTonTV</i>	<i>Inpaint</i>
Pirate (CR = 0.12)					
reconstruction time (s)	60	602	65	437	1.4
PSNR (dB)	29.7	29.0	27.6	29.1	28.1
Lena (CR = 0.10)					
reconstruction time (s)	73	720	62	395	1.4
PSNR (dB)	30.9	30.2	29.6	31.9	30.2
Cameraman (CR = 0.10)					
reconstruction time (s)	73	698	61	422	1.2
PSNR (dB)	32.7	32.4	30.3	33.7	30.6
Blonde (CR = 0.11)					
reconstruction time (s)	60	707	67	471	1.3
PSNR (dB)	29.2	28.6	27.7	29.3	28.3
Darkhair (CR = 0.07)					
reconstruction time (s)	68	690	55	990	1.4
PSNR (dB)	35.3	34.7	33.1	36.4	34.5

Table 4.2 – Reconstruction results for several images, CR varies from 7% to 12%.

Sensing with Smoothed Projected-Landweber) presented in [137], the reconstruction PSNR of Lena using *mDWTonTV* is 0.3dB higher but it takes 8.5 times more time. By contrast, *PS* using *Inpaint* takes 30 times less time while not dramatically reducing the image quality. Those results are obtained using an Intel Xeon CPU quad core @ 2.67GHz. The interest of using such an acquisition yet resides in video sensing.

Reconstruction for videos

The authors of [165] and [142] have proposed CS schemes for video acquisition which require frame to frame analysis. To keep consistency and preserve a simple sensing architecture, no time domain processing is performed in the proposed acquisition. In the field of CS, the larger the signal the more relevant. Yet in general, the time of reconstruction considerably increases when video sequences are involved. In addition of dealing with larger signals, efficient prior operators have to be far more complex in order to take into account time dimension (cf. previous chapter). For instance even if [143] proposes a low complexity CS video reconstruction, it still results in a high complexity algorithm. Block-based processing often helps reducing the computation time but limits the quality because of introducing artifacts. Our work proposes to extend the proposed algorithm (algorithm 5) for image inpainting to three dimensional signals. This algorithm can be used either to reconstruct frame by frame (FBF) or by group of frames (GOF) to take profit from the time dimension. The results for 256×256 cropped version of test video sequences are reported in table 4.3. As expected, PSNR results are far better when PS_{GOF} is used and there is no large camera motion. In comparison with [165] and [142] our CS system has almost the same performances in terms of PSNR for a given compression ratio but requires much less reconstruction time.

method	PS_{FBF}	PS_{GOF}
Hall monitor (CR = 0.14)		
reconstruction time (s/f)	0.23	0.39
PSNR (dB)	27.3	29.4
Coastguard (CR = 0.17)		
reconstruction time (s/f)	0.22	0.37
PSNR (dB)	25.6	26.5
Foreman (CR = 0.10)		
reconstruction time (s/f)	0.25	0.48
PSNR (dB)	29.8	27.0
Mother Daughter (CR = 0.08)		
reconstruction time (s/f)	0.24	0.42
PSNR (dB)	31.8	33.4

Table 4.3 – Reconstruction results for several test video sequences (256 frames).



Figure 4.14 – Reconstructions of the 128th frame of the hall monitor video sequence (compression ratio = 0.14). From left to right: Original frame, Interpolated frame from a 96×96 averaged image, FBF reconstruction with filtering and GOF with 3D filtering.

The sensing scheme and its image sensor architecture are proposed. For this application, the sensing is based on capturing random pixels over blocks of 4×4 pixels. The number of acquired pixels by block depends on the local spatial variance. The implementation does not require complex on-chip processing while improving a lot the acquisition performances. A dedicated reconstruction is proposed and refers to inpainting. The result of using a fast reconstruction algorithm based on successive convolutions followed by a filtering operation shows that such a system is relevant in the case of an application requiring very fast still image reconstruction. A particular interest is on the application of video CS because of its reconstruction speed. Typically, for a 1:10 video compression, the reconstruction lasts less than half a second per frame and the average PSNR is about 30dB. We can reasonably expect that such an algorithm can be implemented in an embedded system to reach far improved performance in terms of time of reconstruction leading to a real time but delayed reconstruction at a typical frame rate such as 60fps.

4.3.2 Tracking using variance and particle filtering

To further demonstrate the capabilities of such an acquisition, a person/object tracking algorithm has been developed. This application is only using adaptively random selected pixels, block variances and block means as inputs. The proposed algorithm is based on a Particle Filter (Sequential Importance Resampling) as in [166]. We aim at showing that compressed input data can directly be employed in a classical tracking framework. The proposed algorithm uses a dynamical model consisting to a constant speed model, thus considering the acceleration as noise [167] [168]. To limit the algorithm complexity we restrict our study to a single rectangular window with a static shape (i.e. length/width ratio) whose scale factor is variable. This window is defined by a user at the beginning of the sequence.

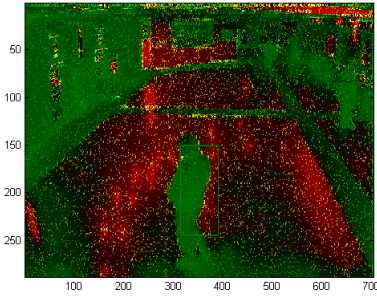


Figure 4.15 – Example of frame composed of camera extracted data only.

Regarding the proposed algorithm, the histogram of selected pixels, block variances and block means are used as object (person) features. Those features are thus replacing the common color histogram as it has been first used in [169]. For instance, figure 4.15 shows an image frame extracted from a sequence (CR=0.22). In this colored picture, for each pixel its luminance, blue and red chrominances respectively corresponds to the pixel extracted value, the block variance and the block mean. As when using color histograms [170] those features are invariant against rotation, deformation and scale. A Bhattacharya distance is then calculated between reference histograms and the current ones to evalu-

ate the likelihood. The proposed technique also adds a histograms updating stage. Indeed, to overcome illumination changes the different reference histograms are updated from frame to frame. This updating procedure is simply performed using a low pass filter (as in [171]) and dramatically improves the accuracy of the tracking when the object change its pose. The results of this tracking algorithm are presented in figure 4.16 using a metro video [172]. This colored scene, for which the acquisition has been simulated according to the adaptive pixel selection based on block variance values, has been compressed by a factor of 5 (4×4 pixel blocks involved). Since the tracking algorithm inputs are only those adaptive CS measurements, it thus speeds up the histogram calculations. In addition, if we qualitatively compare the accuracy of tracking in our case compared to a basic implementation of particle filter in the color space, it is almost as precise for the selected acquisition parameters, leading to a compression ratio of 5.

This basic technique can largely be improved using background subtraction or just frame to frame differences as additional information. Taking into account potential feature-fusion [173], the proposed algorithm can still be improved by using edge orientations



Figure 4.16 – Example of person tracking using only simulated measurements performed by the proposed sensor architecture using 4×4 pixel block variance feature extraction and adaptive pixel selection (CR = 22%).



Figure 4.17 – Example of person tracking using only simulated measurements performed by the proposed sensor architecture using 4×4 pixel block variance feature extraction and adaptive pixel selection (CR = 22%). The results presented in this figure also takes advantage of frame-to-frame differences as additional feature to reweight the histograms.

as additional features that can easily be extracted from block variances. For the sake of simplicity, the frame to frame block means absolute differences have been employed to reweight the Bahattacharya distances at each time step assuming that the tracked object most likely concentrates this frame to frame differences by defining a part of the foreground. The resulting improvement can be noticed in figure 4.17.

The presented results are obtained for specific parameters of the acquisition leading to a compression ratio of 22% but these can be easily be tuned to reach for instance $CR = 5\%$ when increasing block size but without a large drop in terms of tracking accuracy. The tracking robustness can also be improved if the imager provides mosaiced images thus giving some information on the color for all the data. Indeed it would provide block variances/means and selected pixels for each RGB band, leading to 9 histograms.

4.3.3 Face recognition using adaptive CS measurements

We finally present a last application that can be addressed by this adaptive CS sensing scheme, face recognition. Indeed, both dimensionality reduction which is intrinsically performed by CS and features extraction are especially appropriate for this purpose. To demonstrate the interest of such measurements we propose an example of algorithm that perform face recognition on the AT&T database using bob python toolbox [174] (examples of face images are presented in figure 4.18). This multi-class classification is thus considered as a supervised classification problem. As in [120], the proposed method shows that an improvement in terms of recognition accuracy can be reached when using a CS approach compared to most popular subspace analysis approaches such as Eigenfaces, Fisherfaces, and Laplacianfaces. More recently, [175] proposes a comparison of l_1 -min alternative algorithms based on sparse representation for face recognition that are well suited to take advantage of CS measurements as inputs.

General recognition framework

The face recognition framework we employed can be described as follows: first a Background Model (BM) for all the learning database is computed, then an enrolment of each client is performed, finally a probing stage evaluates a distance between a set of clients and a certain client model to return a similarity score (see figure 4.18). The resulting classification obviously returns the model index that yields the best similarity score. The proposed algorithm takes advantage of a Principal Component Analysis (PCA) to reduce dimensionality of the data to deal with, whatever the input data. This operation is performed at the background stage. The similarity measurement performed at the probing stage uses a distance called a Linear Discriminant Distance. In particular, this distance takes profits of knowing intraclass/interclass covariance matrices which are computed at the enrollment stage. The proposed framework is summarized like:

- The inputs data can be of different nature depending on which system of acquisition (and/or feature extraction) we want to evaluate.
- The resulting data can still be dimensionally reduced using a PCA.
- The projections on this reduced PCA space are used to perform similarity measurements between a client and a class model.

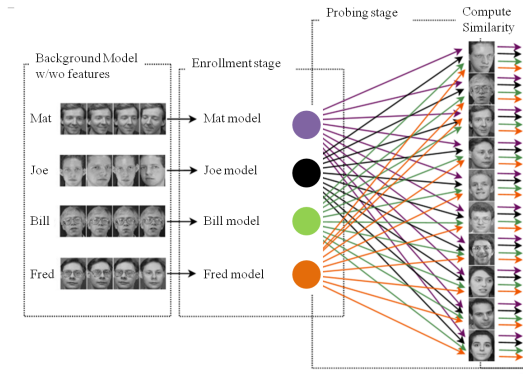


Figure 4.18 – General overview of the program structure.

However, we mainly focus on the impact of using block means/variances features and CS measurements instead of direct pixel values as input. Indeed, we want to show that this face recognition system could work directly with specific imager's outputs reducing a lot the processing complexity and even improving recognition accuracy. In the case of large scale problems, it could really help to reduce the algorithmic complexity by reducing the input data size, for instance making the PCA easier to compute.

Principal Component Analysis

PCA corresponds to a linear orthogonal transform. PCA dimensionally reduction is performed by keeping only the most significant eigen vectors which preserve most of the information and removing irrelevant and insignificant vectors. In the proposed implementation, the PCA is directly performed using a Singular Value Decomposition ($X = USV^*$), where X represents all the data arranged by column, U and V matrices where columns represent eigenvectors of respectively X^*X and XX^* and S a diagonal matrix containing the eigenvalues of XX^* . The proposed algorithm keeps only the 30 first components of the PCA almost corresponding to 95% of the total energy. It means that U_r –the reduced version of U – is composed of only 30 columns. The following presented results are all based on this number of kept components. The matrix U_r^* is thus used to project a vector in the PCA subspace.

Linear Discriminant Distance

Based on the Mahalanobis distance and the Fisher's discriminant analysis, a new specific distance is defined to maximize the interclass variance and minimize the intraclass variance. This metric is thus defined as follows (equation 4.8) between a class k and a column vector x corresponding to a PCA-projected version of an input vector (in our particular case, of size 30 due to the PCA).

$$D_k(x) = (x - \mu_k)^T \Sigma^{-1} \Sigma_b (x - \mu_k) \quad (4.8)$$

where μ_k represents the k -class mean, Σ_b the “between” class (or interclass) covariance matrix and Σ the average intraclass covariance matrix. If we note μ the mean of μ_k for k in the range of classes K and N_k the number of example of the k -class, it thus gives the equations 4.9 and 4.10.

$$\Sigma_b = \frac{1}{K} \sum_{k=1}^K (\mu_k - \mu)(\mu_k - \mu)^T \quad (4.9)$$

$$\Sigma = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k} \sum_{i=1}^{N_k} (x - \mu_k)(x - \mu_k)^T \quad (4.10)$$

Classification results

The proposed algorithm has been tested for inputs of different natures. As a “reference”, the classification accuracy has been evaluated without any feature extraction with direct pixel values (i.e. leading to eigenfaces analysis). Three other alternative inputs have been evaluated in terms of accuracy regarding a development set and a test set.

- **Block means&2bit variances** consist of means of 4×4 pixel blocks and 2bit quantized variances of 4×4 pixel blocks.
- **CS measurements** consist of 40 measurements performed per column of pixels. The random matrix has been chosen for its efficiency in terms of classification (using an a posteriori selection).
- **CS measurements + Block means& 2bit variances** consists of the concatenation of the two previous inputs

Figure 4.19 presents the results in terms of False Acceptance Ratio (FAR) and False Rejected Ratio (FRR) at a given threshold for those different inputs.

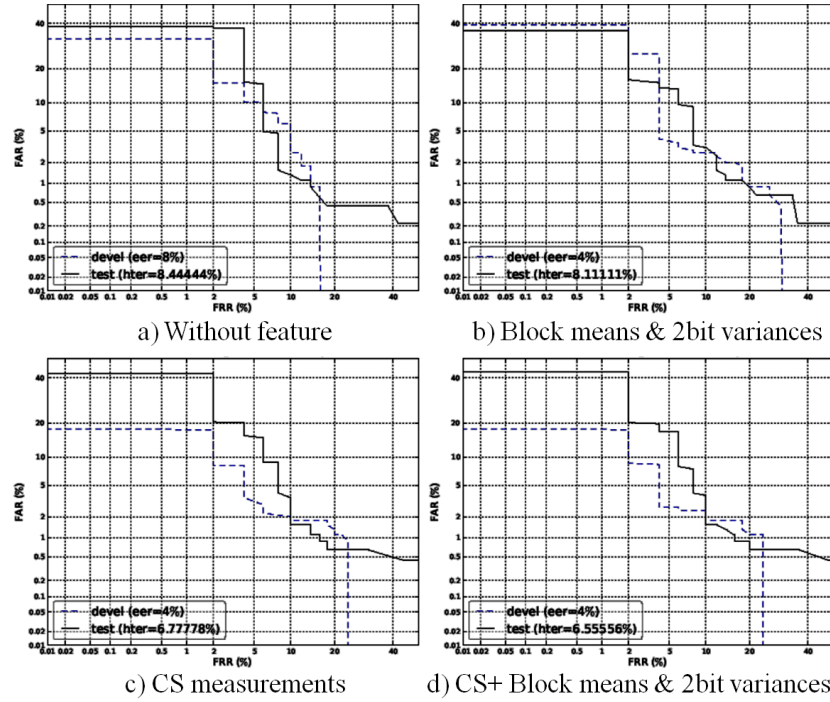


Figure 4.19 – Classification results obtained for the Development set and the Test set.

Those results that are summarized in table 4.4 show an improvement by a factor of 2 regarding the Equal Error Rate (EER) on the development set when CS measurements + Block means& 2bit variances are used as inputs compared to the “reference”. This Error Rate is often used because it defines the operating threshold for the acceptance vs. rejection for which the probability of false acceptance and false rejection are equal. But this improvement is not that important regarding the Half Total Error Rate (HTER, the average between false acceptance and false rejection rates) on the test set for which only 3 falsely-accepted clients are now really-rejected. An interesting point is that if using compressed data, it does not produce any degradation of the efficiency of the classification, it even slightly improves the overall classification accuracy. It thus make sense to implement CS for such applications in the case of power constrained systems.

	CR	Development set			Test set		
		FAR	FRR	EER	FAR	FRR	HTER
Without feature	100%	8.0 %	8.0 %	8.0 %	10.9 %	6.0 %	8.4 %
Block means&2bit variances	8%	4.0 %	4.0 %	4.0 %	10.2 %	6.0 %	8.1 %
CS measurements	35%	4.0 %	4.0 %	4.0 %	5.6 %	8.0 %	6.8 %
CS measurements + Block means&2bit variances	44%	4.0 %	4.0 %	4.0 %	5.1 %	6.0 %	6.6 %

Table 4.4 – Results on the AT&T face recognition dataset for different input of the classifier (450 true examples and 50 false examples in the test set).

Conclusion

In this chapter, we thus show that a large variety of applications can be addressed by adaptive compressive sensors. Indeed, the use of simple features, easy to compute and light in terms of memory load can be employed to locally adapt the way the CS is performed. For instance, for a block-based approach this adaptation consists in tuning the compression ratio for each block. We can also imagine that alternative features could be employed to adapt the type of acquisition depending on the local structure. This way of thinking comes from the fact that Block-based/local CS is generally needed for large scale signal acquisition allowing parallelization at both the encoder and the decoder. In the field of smart image sensors, it becomes a general trend to design architectures that have multiple working modes. For instance, the architecture of adaptive CS imager which is proposed in this chapter is well suited for providing as data outputs either features only, features and CS measurements or direct pixels... It results that the overall number of data can be optimized for a specific application, adapting the working mode according to the scene for example (awakable/smart sensors). The block variances, which are relatively easy to compute on the focal plane using dedicated mixed signal operating units, can additionally be considered as image features as well for high level applications (e.g. tracking and face recognition). We also demonstrate that alternative sensing schemes based on random pixel selection can dramatically boost the overall performances of the imager combined with its dedicated reconstruction algorithm. Finally we present that the semantic dimension of CS data can be used directly in raw format for high-level applications (tracking, face recognition).

5 Beyond spectral limitations

In the two previous chapters, two CS CIS architectures are depicted. Those architectures tend to either relax on-chip electrical limitations or extend capabilities of CS-based acquisitions promoting a certain degree of adaptability at the sensing stage. They do not alleviate intrinsic physical limitations due to the nature of CISs which are the size of the pixels and their spectral sensitivity. However, recent works show possibilities and extensions of the traditional acquisitions of non trivial image information such as depth (e.g. using multiple image sensors) or hypersepctral (e.g. using dispersive optical elements). Nowadays, ever more applications require hyperspectral images. Indeed, spectroscopy analysis and hyperspectral image segmentation play a growing role in many different fields of applications (astronomy, biology, geology, etc.). A recent trend consists in designing new acquisition methods for multi-spectral signals in order to relax various implementation constraints. Multi-spectral images generally require complex processing and expensive optical sensing systems. Fortunately, CS paradigm claims that a compressible signal can be reconstructed from very few linear measurements, if properly designed. By performing such an acquisition, the information of the signal is spread over the resulting measurements. This way, the information is directly compressed at the sensing stage. CS is especially appropriate in the case of hyperspectral images in which spectral information is highly redundant. Indeed, the sparser the signal the more relevant the CS. As example, [176] shows that hyperspectral images can significantly be compressively sensed without much lost of information. On the other hand, there is a continuously increasing need of high resolution images as for astronomical imaging. In the case of some applications, captured images are not sufficiently clean and resolved enough. Super-Resolution (SR) techniques are thus often used to enhance the resolution of the original captured images. There exists both single-frame [177] [178] and multi-frame [179] [71] SR techniques. It either takes profit of intrinsic image properties (sparsity, non local means) and/or takes advantage of some assumptions on the acquisition of multiple frames (spatial shifts, multiple blurs). This chapter tends to demonstrate that hyper-spectral sensing can be combined with a SR acquisition by using multiple spectral dependent blurring kernels that are generated by a compact implementation

regarding optical elements. This chapter is organized as follow. First, a state-of-the-art on hyperspectral imaging system taking advantage of CS is presented. Then, a hyperspectral acquisition model that naturally performs spectral multiplexing by intrinsically summing the results of blurred single band images at the sensor level is depicted with simulation results. Finally, a brief study on the use of Liquid Crystal Device to generate proper blurring kernels for this purpose is provided with outlined discussions on the efficiency and reliability of this approach regarding practical issues.

5.1 Hyperspectral CS

Hyperspectral imaging involves the sensing of a large amount of information. This is mainly due to the three-dimensional nature of such signals which relies on spatial resolution times spectral resolution. Multiple orientations are possible to acquire an hyperspectral image, but two main techniques are commonly used. The first technique consists in scanning a scene using a spectrometer and the second one which corresponds of doing the opposite, based on successive acquisitions of the scene using different bandpass filters. In both cases, the technique requires to acquire the entire hyperspectral cube by using a sequential sensing, involving an important memory load. It explains why the CS approach has shown to be remarkably appropriate in the case of hyperspectral images.

5.1.1 Hyperspectral CS imaging systems

Limitations arise when using the randomness of some media because of a high hardware integration complexity or complex optical calibrations. It results that only few CS applications are now suitable in practice. In addition, the sequential nature of most of the proposed sensing schemes also limits their performances. Yet, [75] shows that parallelization of the measurements can possibly be implemented without much modification with classical sensing systems. In this section we present a short state of the art of Hyperspectral CS techniques, for acquiring the information and for restoring it under some specific assumptions.

Acquisition methods

Two main architectures have been developed in the scope of hyperspectral CS which are the Dual-Disperser Coded Aperture Snapshot Spectral Imager (DD-CASSI) [180] and the Single Disperser Coded Aperture Snapshot Spectral imaging (SD-CASSI) [181]. A DD-CASSI consists of two separate dispersive elements working in an antagonist way so that the dispersion of the first element is properly compensated by the second one. The image cube that the imager receive is thus finally not twisted. In this case the CS modulation is performed between these two dispersers by using either a reflective or a transmissive SLM. The interest of using a SD-CASSI is to reduce the number of required optical elements,

providing a simpler way (in terms of required optical devices) to perform hyperspectral CS, reducing the impact of misalignment of the optical elements. [180] provides an evaluation of the DD-CASSI in terms of efficiency and robustness. [182] and later [183] propose an optimization on the sensing scheme to employ in CASSI for improving CS acquisition performances. [183] also deals with SR issues when reconstructing the hyperspectral image assuming a proper calibration of the CASSI. Alternatively, Compressive spectral imaging based on colored coded apertures, using multiple colored masks is depicted in [184]. [185] extends the CASSI capabilities by replacing the traditional blocking-unblocking coded apertures by a set of colored coded apertures. Instead of processing spectral mixing in the spatial domain, the article [186] proposes a CS spectrometry system based on liquid crystal devices. This technique can obviously be extended to CS hyperspectral imaging by using multiple arrayed spectrometers. [187] claims that CASSI is theoretically suboptimal in terms CS because of employing random sequences multiplexing the signal in the spatial domain. In this article, the authors propose to perform hyperspectral imaging by random separable projections in both the spatial and the spectral domains which improve the theoretical efficiency of CS acquisition. Simulations show that the best performances are achieved when the compression between the spatial domain and the spectral domain is properly balanced. [188] details the effect of the spatial compression versus the spectral compression in that case of application.

Reconstructions

From CS measurements operated over the hyperspectral cube, a large variety of algorithms are competing to promote a proper reconstruction. We can easily differentiate the reconstructions whose role is to reconstruct each pixel of each spectral from those which directly unmix or segment the hyperspectral image [189]. As example [190] [191] both present joint segmentation/unmixing and reconstruction methods of coded aperture hyperspectral data. The reconstruction method presented in [176] is based on low-rank and joint-sparse priors on the hyperspectral image you want to reconstruct, assuming a low number of different signal signatures. Using some a priori information on the spectral signatures, the joint segmentation and reconstruction can be efficiently performed [192] [193]. This technique is particularly proficient for specific applications such as target detection [194] or classification [195] based on an already learned dictionary. Alternatively, [196] presents a CS reconstruction algorithm based on compressive-projection principal component analysis (CPPCA). CPPCA is extended to incorporate both spectral and spatial partitioning to address classification and unmixing tasks. Finally, [197] presents a blind CS reconstruction for data acquired by CASSI. Indeed, the use of a Bayesian approach and extending dictionary learning to blind CS reconstruction is employed to improve the robustness of the algorithm against miscalibration of the acquisition system.

5.1.2 Study on 3-Dimensionnal constraints

Similarly to video sequences, hyperspectral images have high structural similarities between consecutive bands. [176] shows that for specific hyperspectral images, almost perfect reconstructions can be achieved with only 3% of CS measurements. Indeed, a low rank regularization is employed to minimize the number of different spectral signatures in the image. This kind of prior adds a global constraint over the whole hyperspectral cube which seems particularly relevant in the case of wide range hyperspectral images (a high number of bands compared to the resolution). Regarding multi spectral acquisition in the visible range (for instance) and without a large number of bands, the low rank assumption is generally not properly verified. On the contrary, we thus propose to take advantage of local constraints supposing that the structuralism between consecutive bands is relatively high. In that scope, we propose to reuse the constraint operators (equation 3.20 to equation 3.22) that have already been employed in the context of video CS reconstruction in chapter 3. The reconstruction of URBAN and GULF OF MEXICO hyperspectral test images are presented in figure 5.1. For this assessment of the different constraint operators, we use 2D random convolution measurements performed on each band separately which easy to deals with and generate for simulation purposes. Notice that URBAN dataset and GULF OF MEXICO respectively consist of 155 bands and 289 bands covering a large spectral range. Figure 5.1 explicitly shows that the best results in term of visual rendering in terms of contrast are achieved using constraint operator $mDWT_{SO}TV_{2D}$. For instance, the average PSNR of compressed URBAN reconstructed bands is of 25dB for a compression ratio of 1/64 when using this operator.

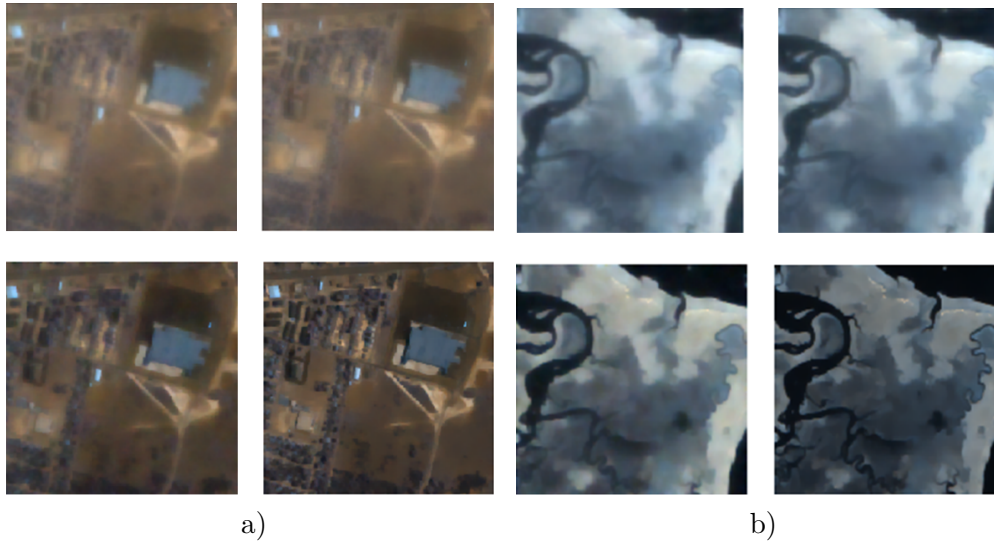


Figure 5.1 – CS reconstructions of Urban a) and Mexican gulf b) hyperspectral test images for a compression ratio of 1.6% and using different constraint operators. From left to right, from top to bottom, $TVmDWT_{2D}$, $TV_{SO}mDWT_{2D}$, $TV_{SO}TV_{2D}$, $mDWT_{SO}TV_{2D}$. Notice that these images has been color-interpolated to render a colorized image from the hyperspectral cube.

The figure 5.2 reports the average PSNR reconstruction results for several hyperspectral images and compression ratios. The first two images are composed of 33 spectral bands in the visible range [198] (scenes 6,8), whereas the others are Urban and Gulf of Mexico datasets respectively consisting of 155 bands and 289 bands that cover a large spectral range. As expected, the operator $mDWT_{SO}TV_{2D}$ provides the best results both in terms of visual rendering and PSNR thus leading to the same conclusions as for video signals. Yet, the interest of using such a constraint operator in the case of hyperspectral is even higher because of higher correlation in the spectral axis than in the time axis for videos. Indeed, if there still exists more variations of the intensity and on the contrast between spectral bands of an hyperspectral image than between frames in a video, there is less modification of the scene (motion).

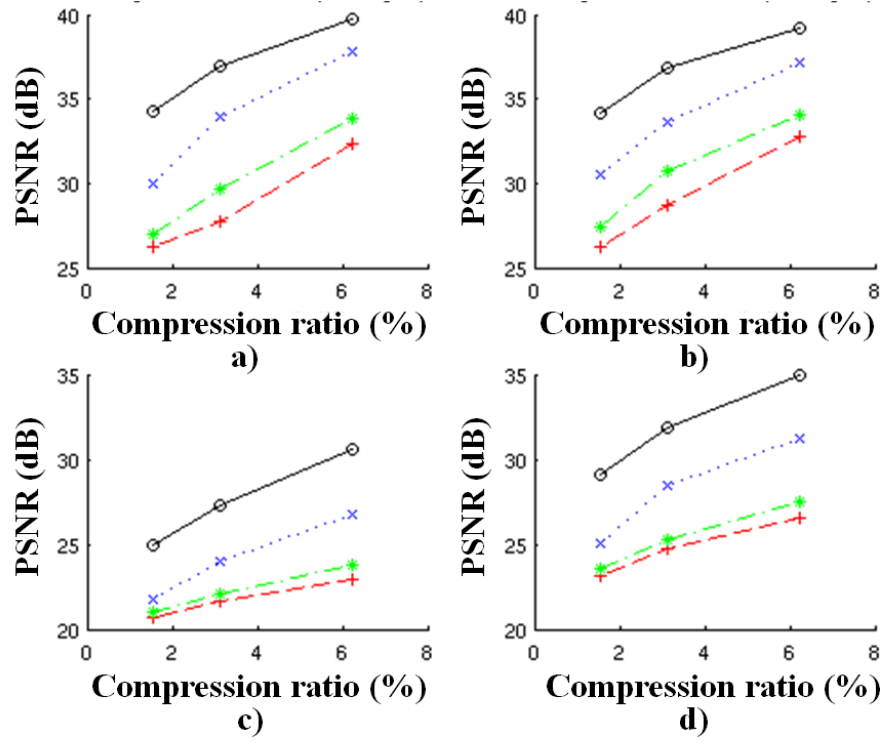


Figure 5.2 – Average PSNR over all spectral bands for the different constraint operators ($TVmDWT_{2D}$ (red), $TV_{SO}mDWT_{2D}$ (green), $TV_{SO}mDWT_{2D}$ (blue), $mDWT_{SO}TV_{2D}$ (black)). a) and b) corresponds to scenes 6 and 8 of dataset [198]. c) and d) respectively correspond to Urban and Gulf of Mexico datasets.

5.2 Multiple blurring kernels for joint SR and hyperspectral

A new approach to perform the acquisition and the reconstruction of joint Super-Resolved and hyperspectral images is presented. The proposed hyperspectral sensing strategy is based on acquiring several low-resolved grayscale images following a specific acquisition scheme which takes profit from different spectral dependent blurring kernels. The proposed model describes how output grayscale pixels are related to input multi-spectral pixels in a CS fashion. By its nature, an optical scattering media with particular spectral properties can be used to generate such appropriate Point Spread Functions (PSFs). This proposed technique claims that for well chosen scattering media, a common grayscale image sensor can be employed to acquire super-resolved hyperspectral images. This way the capabilities of a traditional CIS is extended at the expense of the use of an optical element operating properly suited blurs which exhibit a dependency on the frequency band. The proposed reconstruction is performed by a regularization algorithm adding appropriate constraints on the hyperspectral cube (see previous section). More generally, this can be considered as a hyperspectral CS problem since the number of measurements can be less than the total amount of reconstructed information. This study is restricted to the acquisition of discretized hyperspectral bands and to further demonstrate our approach, considered blurring kernels are with normalized power. This way, there is no modulation of the amplitude power in the spectral axis.

5.2.1 Sensing Scheme based on multiple blurring kernels

CS commonly requires a certain media to produce randomness if not performed by the image sensor itself. Even though not ideally suited for CS, multiple blurring kernels acquisition have already been used for SR [179]. In parallel, [71] shows that a judicious selection of the type of distortion induced by measurement systems dramatically improves the abilities of performing SR from coded aperture image. For that purpose, generated kernels have to be incoherent between each others as much as possible and need to be defined with a finer resolution than the resolution of the image sensor. Compared to CS, the resolution of the reconstructed image does not depend on the resolution of a Spatial Light Modulator (DMDs in the case of the single pixel camera [65]) or of the image sensor itself [93] but of the level of discretization of the simulation of calibrated PSFs. It means that in the case of multiple blurring kernels acquisition, the only limitation lies in the maximum PSF resolution. An other interesting point is that the entire sensor is thus dedicated to the acquisition, e.g. as proposed in [199] for CS. In comparison with [65], the proposed acquisition in [199] is parallelized and thus performed more efficiently. Under those considerations, successive captures of a single hyperspectral image which is optically processed by different controlled or fixed spectral dependent optical media appears suitable for performing a joint hyperspectral and SR acquisition because of producing multiple spectral dependent blurs. As example, this can be operated by different scattering media [77]. Since the proposed acquisition

5.2. Multiple blurring kernels for joint SR and hyperspectral

being generic, the reconstruction can be suited to specific applications (segmentation, classification) independently of the acquisition scheme. In this section, we only deal with pixel reconstruction of the entire cube to render a hyperspectral output image. This study considers the assumption of slow variation over the pixel intensities between consecutive bands.

More formerly, our approach consists in acquiring multiple grayscale images with known spectral blurs in order to reconstruct a single hyperspectral image with a higher spatial resolution than the CIS original resolution. A set of blurring kernels which corresponds to the natural concatenation of PSFs over the spectral axis is employed. Each kernel is being specific to a certain acquisition and a spectral band. The general overview of the proposed acquisition system is presented in figure 5.3. Multiple acquisitions of the same scene are thus sequentially performed with different scattering media but with the same image sensor and without any coded aperture system.

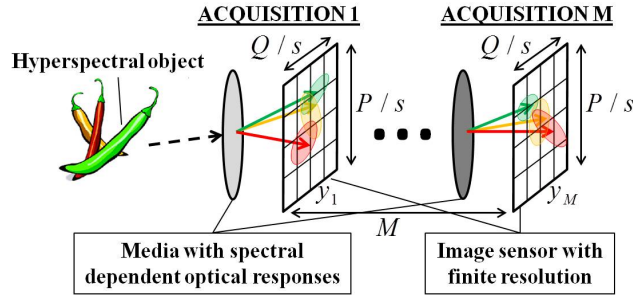


Figure 5.3 – Proposed multiple blurred acquisitions system.

In figure 5.4, the acquisition model is presented as an acquisition chain. Let X be a multi-spectral image of size P, Q, K and $Y_{m \in [1, M]}$ M low resolution grayscale images. Considering a discrete spectrum of K wavelengths constituted of $\lambda_{k \in [1, K]}$. Each X_{λ_k} represents the original image related to the spectral band λ_k of X . h_{m, λ_k} is the 2D blurring kernel generated by the medium, associated to the k^{th} spectral band and the m^{th} acquisition. D_s represents the decimation operator by a factor of s . Regarding SR, the resolution of produced PSFs is thus considered $s \times s$ times larger than the spatial resolution of the image sensor. This way, coupling the information of the different blurred images is used for SR. The final reconstructed resolution reaches the resolution of the discretization (i.e. sampling) of the PSFs. The acquisition procedure is thus described by equation 5.1 for each acquired low resolution image Y_m .

$$Y_m = D_s \left(\sum_{k=1}^K h_{m, \lambda_k} * X_{\lambda_k} \right), m \in \{1, \dots, M\} \quad (5.1)$$

The matrix to vector multiplication defined by the equation (5.2) derives from the

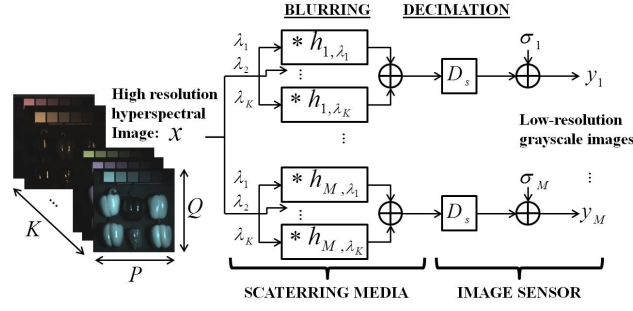


Figure 5.4 – Multiple spectral dependent blurring kernels acquisition model (block operators notations).

equation (5.1). Concatenating all the operators gives the global acquisition expression of equation (5.3). \mathbf{x} , \mathbf{y} and \mathbf{y}_m are now respectively representing X , Y and Y_m in their vectorized forms. D_s denotes also the shrinking operator in its matrix form. H_m denotes the matrix notation of a set of blurring kernel for a single acquisition. D_s^M denotes a block diagonal matrix composed of M D_s matrices. H denotes the vertical concatenation of the M H_m matrices. For the sake of simplicity, figure 5.5 shows a schematic representation of this algebraic notation.

$$\mathbf{y}_m = D_s H_m X \quad (5.2)$$

$$\mathbf{y} = D_s^M H \mathbf{u} \quad (5.3)$$

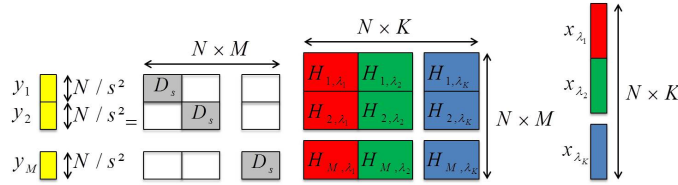


Figure 5.5 – Multiple spectral dependent blurring kernels acquisition model (algebraic notations).

The reconstruction problem is either a determined or an undetermined inverse problem depending on the number of measurements. If the total number of acquired pixels equals the number of reconstructed pixels, the problem is invertible since $D_s^M H$ is nonsingular. More generally, it is not computationally affordable to operate this matrix pseudo-inversion because of its size ($NK \times NK$ in the case of the invertible problem where $N = P \times Q$). Thanks to convexity, it is more appropriate to solve the equation (5.4) using (5.5) opting for a gradient descent algorithm (i.e. find \mathbf{u} for which $FT(\mathbf{u}) = 0$). When the number of acquired pixels is less than the number of reconstructed ones, the number of solutions of the optimization (5.5) is infinite. A regularization term $Op(\mathbf{u})$ thus can be added to the objective function in the minimization to make the original ill-posed problem solvable, mimicking CS framework. In particular, a property of the convolution is employed for solving the CS equivalent problem (equation 5.6): the adjoint

operator of a convolution by the kernel h corresponds to the convolution by π -rotated h .

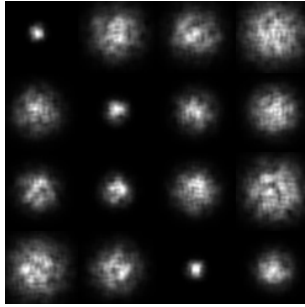
$$\underbrace{\sum_{m=1}^M \|D_s \left(\sum_{k=1}^K h_{m,\lambda_k} * \mathbf{u}_{\lambda_k} \right) - y_m\|_2^2}_{FT(\mathbf{u})} = 0 \quad (5.4)$$

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} \quad FT(\mathbf{u}) \quad (5.5)$$

$$\mathbf{x}' = \underset{\mathbf{u}}{\operatorname{argmin}} \quad (Op(\mathbf{u}) + \alpha FT(\mathbf{u})) \quad (5.6)$$

5.2.2 Joint Super-Resolution and Hyperspectral reconstruction

CS-like reconstruction is thus used to manage the restoration of an hyperspectral and super resolved image. To further validate the proposed sensing scheme, we first show its capabilities in terms of SR and compression applied to grayscale image acquisition only. We then demonstrate the efficiency of our approach applied to joint hyperspectral SR acquisition. For the presented simulations, PSFs are generated as Gaussian blurs of random radius on which a random binary patterned mask is applied. An example of those blurring kernels is shown in figure 5.6 a) and resulting blurred images in figure 5.6 b).



a) Simulated PSFs



b) Low resolved blurred images

Figure 5.6 – a) Example of blurring kernels used for the simulation presented in figure 5.7 and b) the resulting 16 blurred images binned by a factor of 4×4 ($s = 4$) where each image corresponds to a blurring kernel of a).

For the reconstruction of grayscale images, the $mDWTonTV_{4axes}$ constraint operator is used with only Haar wavelet as DWT to re-enforce the sharpness of the image. The

results of the first simulation are reported in figure 5.7 for a 4×4 SR acquisition of Lena image. Even through not a perfect metric, PSNR in the hyperspectral cube is used to assess the reconstruction efficiency. This shows that the SR reconstruction is almost perfect if there is no compression (PSNR = 39.2dB). In the case of a compression by a factor of 4, the reconstructed PSNR is of 33.4dB. In those both cases the only visible artifact is the alteration of very high frequency details on Lena's hat. Edges are perfectly reconstructed as can be observed on the shoulder. It can be compared to the reconstruction using the same algorithm of a single low resolved image. Despite of requiring much more data and calibration, the proposed acquisition technique provides far better results.



Figure 5.7 – For left to right: the original Lena test image, 4×4 ($s = 4$) super-resolution based on a single low resolved image using the proposed regularization, SR reconstruction without compression (the 16 low resolution images have been used at the reconstruction, cf. figure 5.6), SR reconstruction with a compression of factor 2 (only 4 low resolution images are used).

The second simulation presented in figure 5.8 shows that the proposed approach succeeds in joint hyperspectral super-resolved acquisition. For those simulation we use a modified version of $mDWT_{SO}TV_{2D}$ using oriented gradients and the Haar wavelet only as mDWT. The 512×512 original hyperspectral image we use is from the CAVE database [<http://www.cs.columbia.edu/CAVE/databases/multispectral/>] and its range of wavelength bands is from $400nm$ to $700nm$ with a $50nm$ step. For a SR of 2×2 and a compression by a factor of 2 the reconstruction PSNR is of 47.3dB. It is important to notice that spatial details are perfectly preserved and there is almost no residual spectral artifacts. Yet, Spectral artifacts can occurs when the spectral discretization step is not fine enough. Figure 5.8 e) shows PSNR results for different compression ratios and s values. Notice that the greater s the better CS. Compressive Sensing and Super-Resolution can thus be performed when acquiring hyperspectral images by taking

advantage of the optical scattering processes of some media combined with local averaging of phototransduced signals performed at pixel level. A media producing appropriate Point Spread Functions can be employed to properly spatially spread the spectral information in the sense of Compressive Sensing. Based on presented experiments, promising results are achieved for Super-Resolution and Super-Resolution on hyperspectral images. However, the proposed acquisition system requires to find out a scattering media with proper spectral characteristics and without too much dispersion due to temperature (for example). Tunable Liquid Crystals Devices (LCDs) could maybe used to control spectral dependent PSFs instead of using multiple different natural transparent media. Under certain conditions, specific LCDs exhibit well suited properties for that purpose as it going to be presented in the next section.

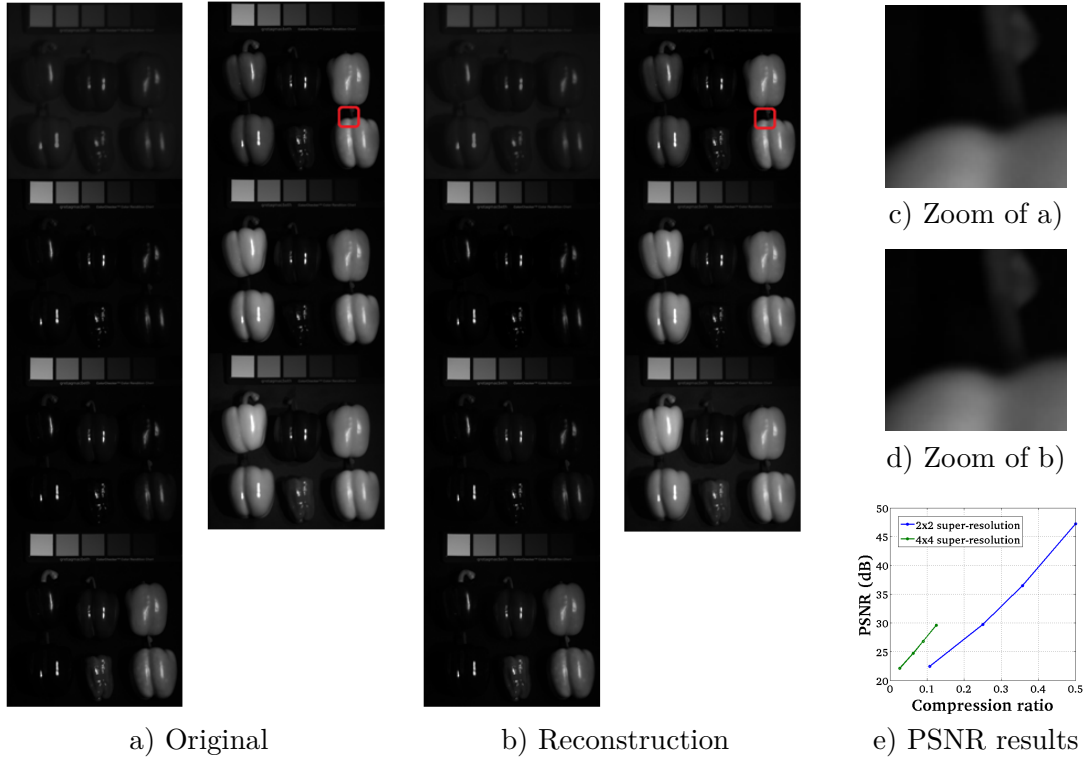


Figure 5.8 – a) is the original hyperspectral image from the CAVE database, b) is the reconstructed image with a Super-Resolution of 2×2 ($s = 2$) and a compression factor of 2. e) shows PSNR results for different settings.

5.3 Liquid Crystal Devices for SRHS imaging

Liquid crystals materials are well known for generating convolution kernels (Point Spread Functions) [200] [201] [202]. They also have been used in several CS acquisition systems. In that scope, they have been employed to spatially modulate images. Our preliminary study on a real implementation of the SRHS described above tends to show that LCDs

could be used for that purpose. In fact, nematic liquid crystals can be used to modulate also spectrally the images before being traditionally captured by a digital grayscale camera. For instance a liquid crystal such as the E44 exhibits appropriate properties according to its different refraction indexes depending on both the wavelength and a local control voltage. We thus propose to use this specific media to produce proper spectral dependent blurring kernels for hyperspectral imaging with super-resolution in the spatial domain (increase the original resolution of the sensor). In such configuration, greater the wavelength dependency of the response, better the optical media. Since the Liquid Crystals that is considered features a sufficient wavelength dependency and refraction index excursion, two different approaches are possible for the generation of proper blurring kernels. First, a random command could be employed which makes the proposed acquisition close to a random convolution based system. In this case, calibration for SR would be almost impossible. Indeed, it would mean that the resolution of the resulting kernels should be higher than the pixel pitch of the sensor, in the random case is hardly possible.

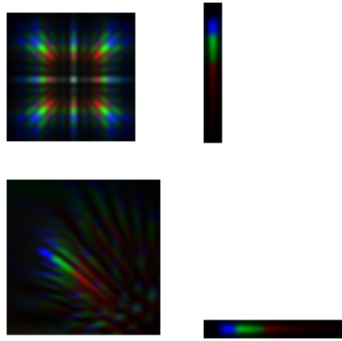


Figure 5.9 – Colorized blurring kernels.

Instead of using this random point of view, the second option is to generate more “predictable” blurring kernels; for instance using Zernike polynomials to form specific wavefronts. This way and by regularization of the calibration, we can synthesize a super-resolved version of the generated blurring kernels for each spectral band we are interested in. As example, a simulation of such kernels is shown in figure 5.9. In this figure, 4 different kernels are presented with their own spectral components (colorized in red, green and blue). Using those generated kernels we simulate the acquisition and the reconstruction of a color images considering a discrete spectrum (3 separate bands). An example of reconstructed color image is presented in figure 5.10 for $s = 2$. It results

the same amount of data than a color acquisition using a Bayer CFA. The proposed acquisition reduce the aliasing effect compared to a Bayer mosaicing combined with a bilinear filtering reconstruction. However it introduces blur and artifacts related to the specific blurring kernels which are employed.

Table 5.1 reports the PSNR for several images, comparing a CFA-based color acquisition and the proposed acquisition based on the blurring kernels of figure 5.9. We can notice that the proposed acquisition outperforms a traditional Bayer filter considering a simple bilinear filter for demosaicing. However, it still is far less efficient compared to the CFA when an advanced demosaicing algorithm is employed.

The major interest of the proposed acquisition is to allow the selection a posteriori of the different bands and its number that we want to reconstruct since the acquisition is performed. The number of bands is only limited by the number of low resolution images



Figure 5.10 – For left to right: the original Lighthouse test image, demosaiced image using bilinear filtering, demosaiced image using a $mDWTonTV_{4axes}^{RGB}$ regularization, SRHS reconstruction without compression (as much data as a mosaiced image).

acquisition	CFA-Bayer		LCDs
reconstruction	Bilinear	$mDWTonTV_{4axes}^{RGB}$	
Zoneplate	15.8	36.5	17.7
Barbara	27.6	34.0	30.2
Peppers	31.7	35.3	33.7
Lighthouse	28.1	40.2	32.5

Table 5.1 – Average RGB-PSNR of the reconstructed color images.

that have been acquired and strained by the variations of intensities in the spectral axis. Future works will be to define appropriate controls of the LCDs taking into account the following sticking points:

- the inter-correlation of the kernels between bands for the same commands,
- the inter-correlation of the kernels generated by different commands,
- the effect of the temperature,
- the effect of the misalignment,
- the effect of temporal noise both due to the sensor and the LCDs,
- the calibration of the system and possible compensation (mechanical, electronic and algorithmic).

To improve the presented results, a proper study and tests of liquid crystals of a certain type should be performed. In that perspective, a partnership should be developed with a laboratory whose field of study is optics and optoelectronics.

6 Main outlines and perspectives

In this thesis, we investigate new ways to acquire image signals in a Compressive Sensing fashion. The main objective is to explore the different applications that can be properly addressed by CS in the image sensing field. We focus on alternatives to traditional ways of thinking CS in order to define specific sensing schemes and dedicated reconstruction that consist of compromises between Nyquist sampling and CS. This final chapter summarizes the main contributions and potential impacts in the field of image sensor design field. Finally, we depict what could be the future research lines in the specific domain of smart image sensors possibly taking advantage of integrating CS.

CS reverses the traditional processing flow both with regard to the sensing and displaying parts. Indeed, it lightens the coder when loading the decoder. It generally refers to the opposition of the analysis versus the synthesis. When traditional signal processing techniques tend to analyze the signal in order to compress it or to interpret it by understanding the underlying information, CS only captures a certain amount information in a “compressed format”. As a counterpart, it supposes that some prior or assumption has been done on the nature of the signal. However, since the analysis is performed, the traditional techniques do not involve major processing to restore the signal in its original form. It leads to the fact that great opportunities for managing resources in the case of constrained systems are related to the adaptation of sensor work modes in function of some external requests when the use of CS or not is enabled. A wide range of applications require such a working mode adaptability to ensure smart power management or event detection and reconfiguration for instance.

In this thesis, we study the impact of CS to limit electrical circuit-related constraints, extract semantic information and reduce physical limitation of CMOS Image Sensors.

SACSI: In chapter 3 a novel imager architecture (see figure 6.1) is proposed which implement an an array of 3T pixels that are able to work either in CS mode or in traditional rolling shutter. We demonstrate that the impact of adding the CS feature to a traditional 3T-pixel imager is relatively negligible in terms of silicon area due to smart

circuit rearrangements and reconfigurability. Several trade-off are considered such as the measurement support, the number of activated pixels per CS measurement, the time of integration against the maximum frame (column measurements) rate. All of these trade-offs lead to an imager that is able to work in a plurality of working modes to address specific applications such as CS-HDR and CS-HighFrameRate imaging for transient detection.

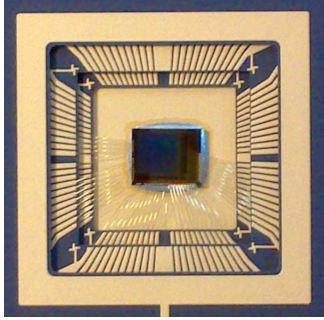


Figure 6.1 – Packaged SACSI chip.

In parallel, the dedicated algorithms employed to reconstruct the CS data have been implemented to demonstrate that the image sensor can successfully address those applications. In that scope, we develop the novel concept of CS-HDR based on multiple CS acquisitions with different times of integration. We also evaluate the impact on the reconstructions of circuit implementation issues such as technology dispersions and matching, temporal noise and nonlinearities. Finally, a reconstruction algorithm framework has been proposed to achieve a correct reconstruction even if the sensing model is corrupted. We demonstrate the efficiency of this algorithm for small but non negligible deviation of the sensing model against the real one in the case of CS-HDR. Unfortunately,

no precise test-chip characteristics evaluation in terms of power consumptions and efficiency has not been performed yet. However we estimate –according to a very pessimistic top-level (including I/O ring) simulation of the imager– an average equivalent power consumption of around $50nJ/measurement$. It thus implies to optimize multiple biases and time diagrams for reducing static/dynamic power consumption. A special attention should be paid to CS-HDR specific time diagrams in order to estimate the interest of performing measurements in which multiple times of integration are used.

Adaptive CS: In chapter 4 the concept of adaptive CS is developed. We demonstrate that in the case of Block-based-CS which is commonly used for designing a CS-dedicated CIS, adaptive CS could play an important role to define smart compromises between the sensor complexity and the reconstruction complexity. We could say that a bit of adaptability in the non adaptive world of CS can help a lot because of compressibility variability in case of Block-based CS. We first show that the local block variance in the spatial domain is an appropriate feature to estimate the compressibility of a natural image. Indeed, it is simple but efficient which leads to the fact that compact implementations can be found to compute on-chip those block features. In addition, we also show that the use of nonlinear measurements (block variance measurements in our case) can improve the CS reconstruction by being added to the fidelity term. A specific architecture implementation using $\Sigma\Delta$ s converters is proposed with an application on adaptive random pixel selection. The proposed implementation is lucky of being without large requirements regarding silicon area, power consumption and frame rate. This novel sensing scheme can be combined to fast inpainting algorithms based on iterative convolutions. We demonstrate that the proposed couple of encoder/decoder is both fast and efficient. In addition, we

further show that the outputs of adaptive CS, based on the estimation of the block variance, can be employed for high level application such that tracking and face recognition.



Figure 6.2 – Multi-scale block variance image of Lena.

Future works would define and implement, proper circuits to perform adaptive CS for imaging. As example, the use of adaptive CS combined with CS-HDR would very appropriate in the case of applications such as radars, Lidars or multiview imaging. Proposed design and implementation of variance calculation has a wide area of possibility for improvement in terms of hardware implementation. For instance, multiple cascaded digital integrations could be used at the output of the single order $\Sigma\Delta$ for calculating statistical moments of different orders. The use of higher order statistical moments both as features and as measurements should also be investigated. Additionally, we started an investigation on multi-scale Block-based CS using multi-scale block variances –which provided very promising preliminary results (see figure 6.2)– should be continued for defining smart hardware implementations and proper reconstruction algorithms to correctly evaluate the interest of it.

Physical limitations: In chapter 5 an alternative sensing scheme for Hyperspectral imaging in a CS fashion is proposed. Aims are at optimizing the number of required optical element to perform a CS based Hyperspectral acquisition of images. In particular, we demonstrate that if a controllable optical media can produce spectral dependent blurring kernel would be possible from multiple grayscale acquisition to reconstruct a super resolved, hyperspectral image. A brief study is then provided on the potential use of specific LCDs to perform the needed optical operations. Regarding super-resolution issues, optimized parameters are provided to reach an optimal acquisition, assuming that the LCDs can properly follow their theoretical behavior. Still, an important work remains on how optimizing the control signal that could be employed for such an acquisition system when the spectral dimension is involved. The two main issues are first the impact of noise and perturbation related to temperature variations and secondly the calibration and misalignment of the proposed system. Several works tend to provide tools to compensate unknown –or approximatively known– acquisition model parameters (such as the algorithm 3) but in the case of using LCDs for super-resolution, a proper calibration is required to determinate blurring kernels with a sub-pixel resolution. Thanks to the provided preliminary study on the LCD-generated blurring kernels, different assumptions on the blurring shapes related to a certain command could be employed for that purpose.

Bibliography

- [1] E. R. Fossum, “Active pixel sensors: Are ccds dinosaurs?,” in *IS&T/SPIE’s Symposium on Electronic Imaging: Science and Technology*, pp. 2–14, International Society for Optics and Photonics, 1993.
- [2] E. Fossum, “Digital camera system on a chip,” *IEEE Micro*, vol. 18, no. 3, pp. 8–15, 1998.
- [3] A. El Gamal, “Trends in CMOS image sensor technology and design,” in *Electron Devices Meeting, 2002. IEDM ’02. International*, pp. 805–808, 2002.
- [4] A. Dupret, M. Tchagaspian, A. Verdant, L. Alacoque, and A. Peizerat, “Smart imagers of the future,” in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2011*, pp. 1–6, 2011.
- [5] A. Moini, “Vision chips or seeing silicon, 1997,” URL <http://www.eleceng.adelaide.edu.au/Groups/GAAS/Bugeye/visionchips>, vol. 240.
- [6] R. Fontaine, “Recent innovations in CMOS image sensors,” in *Advanced Semiconductor Manufacturing Conference (ASMC), 2011 22nd Annual IEEE/SEMI*, pp. 1–5, 2011.
- [7] D. L. Donoho, “Compressed sensing.”
- [8] E. Candes and J. Romberg, “Sparsity and incoherence in compressive sampling,” *Inverse problems*, vol. 23, no. 3, p. 969, 2007.
- [9] M. Bigas, E. Cabruja, J. Forest, and J. Salvi, “Review of cmos image sensors,” *Microelectronics journal*, vol. 37, no. 5, pp. 433–451, 2006.
- [10] A. J. Theuwissen, “Cmos image sensors: State-of-the-art,” *Solid-State Electronics*, vol. 52, no. 9, pp. 1401–1406, 2008.
- [11] H. Tian, B. Fowler, and A. E. Gamal, “Analysis of temporal noise in cmos photodiode active pixel sensor,” *Solid-State Circuits, IEEE Journal of*, vol. 36, no. 1, pp. 92–101, 2001.

- [12] X. Wang, P. R. Rao, A. Mierop, and A. J. Theuwissen, "Random telegraph signal in cmos image sensor pixels," in *Electron Devices Meeting, 2006. IEDM'06. International*, pp. 1–4, IEEE, 2006.
- [13] C. Jung, M. Izadi, M. La Haye, G. Chapman, and K. Karim, "Noise analysis of fault tolerant active pixel sensors," in *Defect and Fault Tolerance in VLSI Systems, 2005. DFT 2005. 20th IEEE International Symposium on*, pp. 140–148, 2005.
- [14] H. Tian and A. El Gamal, "Analysis of 1/f noise in switched mosfet circuits," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 48, no. 2, pp. 151–157, 2001.
- [15] R. Baraniuk, "Compressive sensing," 2007.
- [16] P. Boufounos, M. Duarte, and R. Baraniuk, "Sparse signal reconstruction from noisy compressive measurements using cross validation," in *Statistical Signal Processing, 2007. SSP '07. IEEE/SP 14th Workshop on*, pp. 299–303, 2007.
- [17] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [18] J. N. Laska, P. T. Boufounos, M. A. Davenport, and R. G. Baraniuk, "Democracy in action: Quantization, saturation, and compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 429–443, 2011.
- [19] T. Blumensath, "Compressed sensing with nonlinear observations and related nonlinear optimization problems," *Information Theory, IEEE Transactions on*, vol. 59, no. 6, pp. 3466–3474, 2013.
- [20] Q. Hou, H. Pan, J. Li, and T. Wu, "Image feature extraction based on compressive sensing with application of image denoising," in *Electrical and Control Engineering (ICECE), 2010 International Conference on*, pp. 1154–1157, 2010.
- [21] M. Davenport, P. Boufounos, M. Wakin, and R. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, 2010.
- [22] S.-J. Kim and G. Giannakis, "Cognitive radio spectrum prediction using dictionary learning," in *Global Communications Conference (GLOBECOM), 2013 IEEE*, pp. 3206–3211, 2013.
- [23] S. Das and N. Vaswani, "Particle filtered modified compressive sensing (pafimocs) for tracking signal sequences," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*, pp. 354–358, 2010.

-
- [24] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, pp. 589–592, May 2008.
- [25]
- [26] A. J. Jerri, "The shannon sampling theorem—its various extensions and applications: A tutorial review," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1565–1596, 1977.
- [27] L. Jacques and P. Vandergheynst, "Compressed sensing: When sparsity meets sampling," tech. rep., Wiley-Blackwell, 2010.
- [28] Z. Wang and A. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9, no. 3, pp. 81–84, 2002.
- [29] E. Candes, N. Braun, and M. Wakin, "Sparse signal AND image recovery from compressive samples," in *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on*, pp. 976–979, 2007.
- [30] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse over-complete representations in the presence of noise," *Information Theory, IEEE Transactions on*, vol. 52, no. 1, pp. 6–18, 2006.
- [31] J. A. T. D. Needell, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," 2008.
- [32] S. Qaisar, R. M. Bilal, W. Iqbal, M. Naureen, and S. Lee, "Compressive sensing: from theory to applications, a survey," *Communications and Networks, Journal of*, vol. 15, no. 5, pp. 443–456, 2013.
- [33] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [34] Y. C. Pati, R. Rezaifar, and P. Krishnaprasad, "Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition," in *Signals, Systems and Computers, 1993. 1993 Conference Record of The Twenty-Seventh Asilomar Conference on*, pp. 40–44, IEEE, 1993.
- [35] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM journal on scientific computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [36] T. Blumensath and M. E. Davies, "Iterative hard thresholding for compressed sensing," *Applied and Computational Harmonic Analysis*, vol. 27, no. 3, pp. 265–274, 2009.
- [37] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm with application to wavelet-based image deblurring," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 693–696, 2009.

- [38] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18914–18919, 2009.
- [39] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *Information Theory Workshop (ITW), 2010 IEEE*, pp. 1–5, IEEE, 2010.
- [40] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [41] J. Yan and W.-S. Lu, "Fast dual-based linearized bregman algorithm for compressive sensing of digital images," 2012.
- [42] F. Xu, S. Wang, and Z. Xu, "A hybrid hard thresholding algorithm for compressed sensing," in *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*, pp. 192–196, 2011.
- [43] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*, pp. 185–212, Springer, 2011.
- [44] N. Perraudin, D. Shuman, G. Puy, and P. Vandergheynst, "Unloibox a matlab convex optimization toolbox using proximal splitting methods," *arXiv preprint arXiv:1402.0779*, 2014.
- [45] M. Zhang and A. Bermak, "Cmos image sensor with on-chip image compression: A review and performance analysis," *Journal of Sensors*, vol. 2010, 2010.
- [46] K.-Y. Min and J.-W. Chong, "A real-time JPEG encoder for 1.3 mega pixel CMOS image sensor SoC," in *Industrial Electronics Society, 2004. IECON 2004. 30th Annual Conference of IEEE*, vol. 3, pp. 2633–2636, 2004.
- [47] A. Bandyopadhyay, J. Lee, R. Robucci, and P. Hasler, "Matia: a programmable 80 μ W/frame CMOS block matrix transform imager architecture," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 3, pp. 663–672, 2006.
- [48] W. Leon-Salas, S. Balkir, K. Sayood, N. Schemm, and M. Hoffman, "A CMOS imager with focal plane compression using predictive coding," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 11, pp. 2555–2572, 2007.
- [49] W. Leon-Salas, S. Balkir, N. Schemm, M. Hoffman, and K. Sayood, "Predictive coding on-sensor compression," in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on*, pp. 1636–1639, 2008.
- [50] Z. Lin, M. Hoffman, N. Schemm, W. Leon-Salas, and S. Balkir, "A CMOS image sensor for multi-level focal plane image decomposition," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, no. 9, pp. 2561–2572, 2008.

-
- [51] A. Nilchi, J. Aziz, and R. Genov, "Focal-plane algorithmically-multiplying CMOS computational image sensor," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 6, pp. 1829–1839, 2009.
- [52] S. Chen, A. Bermak, and Y. Wang, "A CMOS image sensor with on-chip image compression based on predictive boundary adaptation and memoryless qtd algorithm," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 19, no. 4, pp. 538–547, 2011.
- [53] D. Chen, F. Tang, M.-K. Law, and A. Bermak, "A 12 pJ/pixel analog-to-information converter based 816 \times 640 pixel CMOS image sensor," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1210–1222, 2014.
- [54] F. Raymundo, P. Martin-Gonthier, R. Molina, S. Rolando, and P. Magnan, "Exploring the 3D integration technology for CMOS image sensors," in *Electronics, Control, Measurement, Signals and their application to Mechatronics (ECMSM), 2013 IEEE 11th International Workshop of*, pp. 1–5, 2013.
- [55] Y. Oike, Y. Oike, M. Ikeda, A. Theuwsen, J. Solhusvik, J. Chang, and T. Kuroda, "F2: 3D stacking technologies for image sensors and memories," in *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2014 IEEE International*, pp. 512–513, 2014.
- [56] S. Ma, W. Yin, Y. Zhang, and A. Chakraborty, "An efficient algorithm for compressed mr imaging using total variation and wavelets," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pp. 1–8, 2008.
- [57] H. Shen, N. Newman, L. Gan, S. Zhong, Y. Huang, and Y. Shen, "Compressed terahertz imaging system using a spin disk," in *Infrared Millimeter and Terahertz Waves (IRMMW-THz), 2010 35th International Conference on*, pp. 1–2, 2010.
- [58] L.-W. Liu, Y. Li, L.-H. Wang, D.-X. Li, and M. Zhang, "Tof depth map super-resolution using compressive sensing," in *Image and Graphics (ICIG), 2013 Seventh International Conference on*, pp. 135–138, 2013.
- [59] L. Anitori, A. Maleki, M. Otten, R. Baraniuk, and P. Hoogeboom, "Design and analysis of compressed sensing radar detectors," *IEEE Transactions on Signal Processing*, vol. 61, no. 4, pp. 813–827, 2013.
- [60] F. Razzaq, S. Mohamed, A. Bhatti, and S. Nahavandi, "Non-uniform sparsity in rapid compressive sensing MRI," in *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on*, pp. 2253–2258, 2012.
- [61] M. Wakin, J. Laska, M. Duarte, D. Baron, S. Sarvotham, D. Takhar, K. Kelly, and R. Baraniuk, "An architecture for compressive imaging," in *Image Processing, 2006 IEEE International Conference on*, pp. 1273–1276, 2006.

- [62] W. L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 12, p. 121105, 2008.
- [63] H. Chen, N. Xi, B. Song, H. Chen, and K. Lai, "Single pixel infrared camera using a carbon nanotube photodetector," in *Sensors, 2011 IEEE*, pp. 1362–1366, 2011.
- [64] Z. Sadeghipoor, Y. Lu, and S. Susstrunk, "A novel compressive sensing approach to simultaneously acquire color and near-infrared images on a single sensor," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pp. 1646–1650, 2013.
- [65] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 83–91, 2008.
- [66] L. Li, M. Mirza, V. Stankovic, L. Li, L. Stankovic, D. Uttamchandani, and S. Cheng, "Optical imaging with scanning MEMS mirror - a single photodetector approach," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*, pp. 2685–2688, 2009.
- [67] M. Shams, Z. Jiang, S. Rahman, J. Qayyum, L.-J. Cheng, H. Xing, P. Fay, and L. Liu, "Approaching real-time terahertz imaging with photo-induced coded apertures and compressed sensing," *Electronics Letters*, vol. 50, no. 11, pp. 801–803, 2014.
- [68] G. A. Howland, P. B. Dixon, and J. C. Howell, "Photon-counting compressive sensing laser radar for 3d imaging," *Applied optics*, vol. 50, no. 31, pp. 5917–5920, 2011.
- [69] W. Dang, A. Pezeshki, and R. Bartels, "Analysis of misfocus effects in compressive optical imaging," in *Signals, Systems and Computers (ASILOMAR), 2014 Conference Record of the Forty Eighth Asilomar Conference on*, 2014.
- [70] F. Magalhães, M. Abolbashari, F. Farahi, F. M. Araújo, and M. V. Correia, "A compressive sensing based transmissive single-pixel camera," in *International Conference on Applications of Optics and Photonics*, pp. 80010B–80010B, International Society for Optics and Photonics, 2011.
- [71] R. Marcia and R. Willett, "Compressive coded aperture superresolution image reconstruction," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 833–836, 2008.
- [72] D. Reddy, A. Veeraraghavan, and R. Chellappa, "P2C2: Programmable pixel compressive camera for high speed imaging," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pp. 329–336, 2011.

- [73] A. Veeraraghavan, D. Reddy, and R. Raskar, “Coded strobing photography: Compressive sensing of high speed periodic videos,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 671–686, 2011.
- [74] J. Chen, Y. Wang, and H. Wu, “Coded aperture compressive imaging array applied for surveillance systems,” *Journal of Systems Engineering and Electronics*, vol. 24, no. 6, pp. 1019–1028, 2013.
- [75] T. Bjorklund and E. Magli, “A parallel compressive imaging architecture for one-shot acquisition,” in *Picture Coding Symposium (PCS), 2013*, pp. 65–68, 2013.
- [76] R. Fergus, A. Torralba, W. T. Freeman, R. Fergus, A. Torralba, and W. T. Freeman, “Random lens imaging,” 2006.
- [77] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron, “Imaging with nature: A universal analog compressive imager using a multiply scattering medium,” *arXiv preprint arXiv:1309.0425*, 2013.
- [78] X. Shao, W. Dai, T. Wu, F. Liu, and H. Li, “Imaging through disordered media by reconstruction with compressed sensing theory,” in *Imaging Systems and Applications*, pp. ITh3C–3, Optical Society of America, 2014.
- [79] G. Huang, H. Jiang, K. Matthews, and P. Wilford, “Lensless imaging by compressive sensing,” in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 2101–2105, 2013.
- [80] H. Jiang, G. Huang, and P. Wilford, “Multi-view in lensless compressive imaging,” in *Picture Coding Symposium (PCS), 2013*, pp. 41–44, 2013.
- [81] M. Zhang and A. Bermak, “Digital pixel sensor with on-line spatial and temporal compression scheme,” in *Integrated Circuits, ISIC '09. Proceedings of the 2009 12th International Symposium on*, pp. 155–158, 2009.
- [82] R. Robucci, J. Gray, L. K. Chiu, J. Romberg, and P. Hasler, “Compressive sensing on a CMOS separable-transform image sensor,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1089–1101, 2010.
- [83] L.-l. Xiao, K. Liu, and D.-p. Han, “Cmos low data rate imaging method based on compressed sensing,” *Optics & Laser Technology*, vol. 44, no. 5, pp. 1338–1345, 2012.
- [84] R. Njuguna and V. Gruev, “Low power programmable current mode computational imaging sensor,” *IEEE Sensors Journal*, vol. 12, no. 4, pp. 727–736, 2012.
- [85] L. Alacoque, L. Chotard, M. Tchagaspian, and J. Chossat, “A small footprint, streaming compliant, versatile wavelet compression scheme for cameraphone imagers,” in *Proceedings of the 2009 International Image Sensor Workshop, IISW*, vol. 9, 2009.

- [86] S. S. Mahant-Shetti, "Cmos sensor camera with on-chip image compression," July 9 2002. US Patent 6,417,882.
- [87] J. Romberg, "Sensing by random convolution," in *Computational Advances in Multi-Sensor Adaptive Processing, 2007. CAMPSAP 2007. 2nd IEEE International Workshop on*, pp. 137–140, 2007.
- [88] V. Majidzadeh, L. Jacques, A. Schmid, P. Vanderghelynst, and Y. Leblebici, "A (256×256) pixel 76.7mW CMOS imager/ compressor based on real-time in-pixel compressive sensing," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, pp. 2956–2959, 2010.
- [89] L. Jacques, P. Vanderghelynst, A. Bibet, V. Majidzadeh, A. Schmid, and Y. Leblebici, "CMOS compressed imaging by random convolution," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 1113–1116, 2009.
- [90] M. Hosseini Kamal, M. Golbabaee, and P. Vanderghelynst, "Light field compressive sensing in camera arrays," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 5413–5416, 2012.
- [91] Y. Wu and G. Arce, "Snapshot spectral imaging via compressive random convolution," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pp. 1465–1468, 2011.
- [92] M. Zhang, Y. Wang, and A. Bermak, "Block-based compressive sampling for digital pixel sensor array," in *Quality Electronic Design (ASQED), 2010 2nd Asia Symposium on*, pp. 9–12, 2010.
- [93] Y. Oike and A. El Gamal, "CMOS image sensor with per-column $\Sigma\Delta$ ADC and programmable compressed sensing," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, pp. 318–328, 2013.
- [94] H.-T. Wang and W. Leon-Salas, "An incremental sigma delta converter for compressive sensing applications," in *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, pp. 522–525, 2011.
- [95] M. Dadkhah, M. Deen, and S. Shirani, "Block-based compressive sensing in a CMOS image sensor," *IEEE Sensors Journal*, to be published. Early Access.
- [96] M. Dadkhah, M. Deen Jamal, and S. Shirani, "Block-based cs in a CMOS image sensor," *IEEE Sensors Journal*, vol. 14, no. 8, pp. 2897–2909, 2014.
- [97] N. Katic, M. Hosseini Kamal, M. Kilic, A. Schmid, P. Vanderghelynst, and Y. Leblebici, "Column-separated compressive sampling scheme for low power CMOS image sensors," in *New Circuits and Systems Conference (NEWCAS), 2013 IEEE 11th International*, pp. 1–4, 2013.

-
- [98] N. Katic, M. Hosseini Kamal, M. Kilic, A. Schmid, P. Vandergheynst, and Y. Leblebici, "Power-efficient CMOS image acquisition system based on compressive sampling," in *Circuits and Systems (MWSCAS), 2013 IEEE 56th International Midwest Symposium on*, pp. 1367–1370, 2013.
 - [99] F. Yang, S. Wang, and C. Deng, "Compressive sensing of image reconstruction using multi-wavelet transforms," in *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference on*, vol. 1, pp. 702–705, 2010.
 - [100] H. Rauhut, K. Schnass, and P. Vandergheynst, "Compressed sensing and redundant dictionaries."
 - [101] Q. She, Z. Luo, Y. Zhu, H. Zou, and Y. Chen, "Spatially adaptive image reconstruction via compressive sensing," in *Asian Control Conference, 2009. ASCC 2009. 7th*, pp. 1570–1575, 2009.
 - [102] J. Zhang, D. Zhao, C. Zhao, R. Xiong, S. Ma, and W. Gao, "Image compressive sensing recovery via collaborative sparsity," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 380–391, 2012.
 - [103] A. Majumdar and R. Ward, "Compressive color imaging with group-sparsity on analysis prior," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 1337–1340, 2010.
 - [104] M. Dadkhah, S. Shirani, and M. Deen, "Compressive sensing with modified total variation minimization algorithm," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 1310–1313, 2010.
 - [105] W. Dong, X. Yang, and G. Shi, "Compressive sensing via reweighted TV and nonlocal sparsity regularisation," *Electronics Letters*, vol. 49, no. 3, pp. 184–186, 2013.
 - [106] Z. Lin, "Image adaptive recovery based on compressive sensing and genetic algorithm," in *Computer Science and Automation Engineering (CSAE), 2012 IEEE International Conference on*, vol. 2, pp. 346–349, 2012.
 - [107] R. Gupta, P. Indyk, E. Price, and Y. Rachlin, "Compressive sensing with local geometric features."
 - [108] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3618–3632, 2014.
 - [109] X. Shu, J. Yang, and N. Ahuja, "Non-local compressive sampling recovery," in *Computational Photography (ICCP), 2014 IEEE International Conference on*, pp. 1–8, 2014.

- [110] J. Fowler, "Compressive-projection principal component analysis," *IEEE Transactions on Image Processing*, vol. 18, no. 10, pp. 2230–2242, 2009.
- [111] J. Haupt, R. Castro, R. Nowak, G. Fudge, and A. Yeh, "Compressive sampling for signal classification," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pp. 1430–1434, 2006.
- [112] P. Barbano and R. Coifman, "Compressive mahalanobis classifiers," in *Machine Learning for Signal Processing, 2008. MLSP 2008. IEEE Workshop on*, pp. 345–349, 2008.
- [113] L. Liu and P. Fieguth, "Texture classification using compressed sensing," in *Computer and Robot Vision (CRV), 2010 Canadian Conference on*, pp. 71–78, 2010.
- [114] L. Liu and P. Fieguth, "Texture classification from random features," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 3, pp. 574–586, 2012.
- [115] V. Thirumalai and P. Frossard, "Motion estimation from compressed linear measurements," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 714–717, 2010.
- [116] G. Warnell, D. Reddy, and R. Chellappa, "Adaptive rate compressive sensing for background subtraction," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 1477–1480, 2012.
- [117] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, "Compressive sensing for background subtraction," in *Computer Vision–ECCV 2008*, pp. 155–168, Springer, 2008.
- [118] D. Venkatraman and A. Makur, "A compressive sensing approach to object-based surveillance video coding," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 3513–3516, 2009.
- [119] H. Xu and F. Yu, "Improved compressive tracking in surveillance scenes," in *Image and Graphics (ICIG), 2013 Seventh International Conference on*, pp. 869–873, 2013.
- [120] N. Vo, N. Vo, S. Challa, and B. Moran, "Compressed sensing for face recognition," in *Computational Intelligence for Image Processing, 2009. CIIP '09. IEEE Symposium on*, pp. 104–109, 2009.
- [121] S. Dongcheng, X. Yidan, and D. Guangyi, "Research on technology of compressed sensing for face recognition," in *Image and Graphics (ICIG), 2013 Seventh International Conference on*, pp. 505–508, 2013.

-
- [122] A. Belenky, A. Fish, A. Spivak, and O. Yadid-Pecht, "Global shutter CMOS image sensor with wide dynamic range," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 54, no. 12, pp. 1032–1036, 2007.
- [123] A. Xhakoni and G. Gielen, "A 132-dB dynamic-range global-shutter stacked architecture for high-performance imagers," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 61, no. 6, pp. 398–402, 2014.
- [124] F. Guezzi-Messaoud, A. Dupret, A. Peizerat, and Y. Blanchard, "High dynamic range image sensor with self adapting integration time in 3D technology," in *Electronics, Circuits and Systems (ICECS), 2012 19th IEEE International Conference on*, pp. 340–343, 2012.
- [125] P.-J. Lapray, B. Heyrman, M. Rosse, and D. Gin hac, "A 1.3 megapixel FPGA-based smart camera for high dynamic range real time video," in *Distributed Smart Cameras (ICDSC), 2013 Seventh International Conference on*, pp. 1–6, 2013.
- [126] D. Baiko and J. Swab, "Linearity of active pixel charge injection devices," *IEEE Transactions on Electron Devices*, vol. 52, no. 8, pp. 1923–1926, 2005.
- [127] K. Yasutomi, S. Itoh, and S. Kawahito, "A two-stage charge transfer active pixel CMOS image sensor with low-noise global shuttering and a dual-shuttering mode," *IEEE Transactions on Electron Devices*, vol. 58, no. 3, pp. 740–747, 2011.
- [128] R. Philipp, D. Orr, V. Gruev, J. Van der Spiegel, and R. Etienne-Cummings, "Linear current-mode active pixel sensor," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 11, pp. 2482–2491, 2007.
- [129] P. P. Chu and R. E. Jones, "Design techniques of fpga based random number generator," in *Military and Aerospace Applications of Programmable Devices and Technologies Conference*, pp. 1–6, 1999.
- [130] R. Dogaru, "Hybrid cellular automata as pseudo-random number generators with binary synchronization property," in *Signals, Circuits and Systems, 2009. ISSCS 2009. International Symposium on*, pp. 1–4, 2009.
- [131] P. S. Bruno Martin, "Pseudo-random sequences generated by cellular automata," 807.
- [132] F. Tang and A. Bermak, "An 84 pW/frame per pixel current-mode CMOS image sensor with energy harvesting capability," *IEEE Sensors Journal*, vol. 12, no. 4, pp. 720–726, 2012.
- [133] M. Pelgrom, A. C. Duinmaijer, and A. Welbers, "Matching properties of MOS transistors," *IEEE Journal of Solid-State Circuits*, vol. 24, no. 5, pp. 1433–1439, 1989.

- [134] F. Maloberti, F. Maloberti, and F. Maloberti, *Analog design for CMOS VLSI systems*. Springer, 2001.
- [135] R. E. Carrillo, J. D. McEwen, D. Van De Ville, J.-P. Thiran, and Y. Wiaux, “Sparsity averaging for compressive imaging,” *Signal Processing Letters, IEEE*, vol. 20, no. 6, pp. 591–594, 2013.
- [136] X. Fei, Z. Wei, and L. Xiao, “Iterative directional total variation refinement for compressive sensing image reconstruction,” *IEEE Signal Processing Letters*, vol. 20, no. 11, pp. 1070–1073, 2013.
- [137] J. E. Fowler, S. Mun, and E. W. Tramel, “Block-based compressed sensing of images and video,” *Foundations and Trends in Signal Processing*, vol. 4, no. 4, pp. 297–416, 2012.
- [138] A. Moghadam, M. Aghagolzadeh, M. Kumar, and H. Radha, “Compressive framework for demosaicing of natural images,” *IEEE Transactions on Image Processing*, vol. 22, no. 6, pp. 2356–2371, 2013.
- [139] P. Nagesh and B. Li, “Compressive imaging of color images,” in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 1261–1264, 2009.
- [140] M. Aghagolzadeh, A. Moghadam, M. Kumar, and H. Radha, “Compressive demosaicing for periodic color filter arrays,” in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pp. 1693–1696, 2011.
- [141] H. Ilbegi and A. Gurbuz, “Demosaicking with compressive sensing,” in *Signal Processing and Communications Applications Conference (SIU), 2012 20th*, pp. 1–4, 2012.
- [142] Y. Hou and F. Liu, “A low-complexity video coding scheme based on compressive sensing,” in *Computational Intelligence and Design (ISCID), 2011 Fourth International Symposium on*, vol. 2, pp. 326–329, 2011.
- [143] M. Asif, F. Fernandes, and J. Romberg, “Low-complexity video compression and compressive sensing,” in *Signals, Systems and Computers, 2013 Asilomar Conference on*, pp. 579–583, 2013.
- [144] X. Yuan, J. Yang, P. Llull, X. Liao, G. Sapiro, D. Brady, and L. Carin, “Adaptive temporal compressive sensing for video,” in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 14–18, 2013.
- [145] I. Wahidah, A. Suksmono, and T. Mengko, “A comparative study on video coding techniques with compressive sensing,” in *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*, pp. 1–5, 2011.

- [146] H.-W. Chen, L.-W. Kang, and C.-S. Lu, "Dictionary learning-based distributed compressive video sensing," in *Picture Coding Symposium (PCS), 2010*, pp. 210–213, 2010.
- [147] F. Yang, H. Jiang, Z. Shen, W. Deng, and D. Metaxas, "Adaptive low rank and sparse decomposition of video using compressive sensing," in *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pp. 1016–1020, 2013.
- [148] T. Do, X. Lu, and J. Sole, "Compressive sensing with adaptive pixel domain reconstruction for block-based video coding," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 3377–3380, 2010.
- [149] B. Choubey and S. Collins, "Low dark current logarithmic pixels," in *Circuits and Systems, 2005. 48th Midwest Symposium on*, pp. 376–379, 2005.
- [150] S. Kavusi and A. El Gamal, "Quantitative study of high-dynamic-range image sensor architectures," in *Electronic Imaging 2004*, pp. 264–275, International Society for Optics and Photonics, 2004.
- [151] S. Kavusi, H. Kakavand, and A. El Gamal, "Quantitative study of high dynamic range sigma-delta-based focal plane array architectures," in *Defense and Security*, pp. 341–350, International Society for Optics and Photonics, 2004.
- [152] S. Vargas-Sierra, G. Linan-Cembrano, and A. Rodriguez-Vazquez, "A 148dB focal-plane tone-mapping qcif imager," in *Circuits and Systems (ISCAS), 2012 IEEE International Symposium on*, pp. 1616–1619, 2012.
- [153] S. Vargas-Sierra, G. Liñán-Cembrano, and A. Rodriguez-Vazquez, "A 151db high dynamic range cmos image sensor chip architecture with tone mapping compression embedded in-pixel," 2013.
- [154] S. Vargas-Sierra, G. Linan-Cembrano, and A. Rodriguez-Vazquez, "A 151 dB high dynamic range CMOS image sensor chip architecture with tone mapping compression embedded in-pixel," *IEEE Sensors Journal*, vol. 15, no. 1, pp. 180–195, 2015.
- [155] M. Abolbashari, F. Magalhães, F. M. M. Araújo, M. V. Correia, and F. Farahi, "High dynamic range compressive imaging: a programmable imaging system," *Optical Engineering*, vol. 51, no. 7, pp. 071407–1, 2012.
- [156] K. Bult and H. Wallinga, "A class of analog CMOS circuits based on the square-law characteristic of an MOS transistor in saturation," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 3, pp. 357–365, 1987.
- [157] E. Seevinck and R. Wassenaar, "A versatile CMOS linear transconductor/square-law function," *IEEE Journal of Solid-State Circuits*, vol. 22, no. 3, pp. 366–377, 1987.

- [158] A. Deshpande and J. Draper, "Squaring units and a comparison with multipliers," in *Circuits and Systems (MWSCAS), 2010 53rd IEEE International Midwest Symposium on*, pp. 1266–1269, 2010.
- [159] J. Moore, M. Thornton, and D. Matula, "Low power floating-point multiplication and squaring units with shared circuitry," in *Circuits and Systems (MWSCAS), 2013 IEEE 56th International Midwest Symposium on*, pp. 1395–1398, 2013.
- [160] D. Liu, X. Sun, F. Wu, S. Li, and Y.-Q. Zhang, "Image compression with edge-based inpainting," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 10, pp. 1273–1287, 2007.
- [161] R. Dogaru, I. Dogaru, and H. Kim, "Chaotic scan: A low complexity video transmission system for efficiently sending relevant image features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 2, pp. 317–321, 2010.
- [162] M. M. O. B. B. Richard and M. Y.-S. Chang, "Fast digital image inpainting," in *Appeared in the Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001), Marbella, Spain*, pp. 106–107, 2001.
- [163] M. M. Hadhoud, K. A. Moustafa, and S. Z. Shenoda, "Digital images inpainting using modified convolution based method," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 7340, p. 26, 2009.
- [164] H. Noori, S. Saryazdi, and H. Nezamabadi-pour, "A convolution based image inpainting," in *1st International Conference on Communication and Engineering*, 2010.
- [165] Z. Liu, H. Zhao, and A. Elezzabi, "Block-based adaptive compressed sensing for video," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, pp. 1649–1652, 2010.
- [166] A. Abdel-Hadi, "Real-time object tracking using color-based Kalman particle filter," in *Computer Engineering and Systems (ICCES), 2010 International Conference on*, pp. 337–341, 2010.
- [167] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, 2002.
- [168] H. Van Trees and K. Bell, "A tutorial on particle filters for online nonlinear/non-gaussian Bayesian tracking," in *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, pp. 723–737, Wiley-IEEE Press, 2007.
- [169] J. Czyz, B. Ristic, and B. Macq, "A color-based particle filter for joint detection and tracking of multiple objects," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, vol. 2, pp. 217–220, 2005.

-
- [170] Z. Zivkovic and B. Krose, "An em-like algorithm for color-histogram-based object tracking," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 1, 2004.
- [171] M. Fotouhi, A. Gholami, and S. Kasaei, "Particle filter-based object tracking using adaptive histogram," in *Machine Vision and Image Processing (MVIP), 2011 7th Iranian*, pp. 1–5, 2011.
- [172] J. Yao and J.-M. Odobez, "Multi-camera multi-person 3d space tracking with mcmc in surveillance scenarios," in *Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications-M2SFA2 2008*, 2008.
- [173] Z. Wang, H. Zhao, H. Shang, and S. Qiu, "An improved particle filter for multi-feature tracking application," in *Imaging Systems and Techniques (IST), 2012 IEEE International Conference on*, pp. 522–527, 2012.
- [174] A. Anjos, L. E. Shafey, R. Wallace, M. Günther, C. McCool, and S. Marcel, "Bob: a free signal processing and machine learning toolbox for researchers," in *20th ACM Conference on Multimedia Systems (ACMMM), Nara, Japan*, ACM Press, Oct. 2012.
- [175] A. Yang, Z. Zhou, A. Balasubramanian, S. Sastry, and Y. Ma, "Fast ℓ_1 -minimization algorithms for robust face recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 8, pp. 3234–3246, 2013.
- [176] M. Golbabaee and P. Vandergheynst, "Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pp. 2741–2744, 2012.
- [177] P. Sen and S. Darabi, "Compressive image super-resolution," in *Signals, Systems and Computers, 2009 Conference Record of the Forty-Third Asilomar Conference on*, pp. 1235–1242, 2009.
- [178] W. Dong, D. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2011.
- [179] S. Farsiu, M. Robinson, M. Elad, and P. Milanfar, "Fast and robust multiframe super resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, 2004.
- [180] M. Busuiocanu, D. W. Messinger, J. B. Greer, and J. C. Flake, "Evaluation of the cassi-dd hyperspectral compressive sensing imaging system," in *SPIE Defense, Security, and Sensing*, pp. 87431V–87431V, International Society for Optics and Photonics, 2013.

- [181] A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Applied optics*, vol. 47, no. 10, pp. B44–B51, 2008.
- [182] H. Arguello and G. Arce, "Code aperture agile spectral imaging (caasi)," in *Imaging Systems and Applications*, p. ITuA4, Optical Society of America, 2011.
- [183] G. Arce, D. Brady, L. Carin, H. Arguello, and D. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2014.
- [184] H. Rueda, H. Arguello, and G. R. Arce, "Compressive spectral imaging based on colored coded apertures," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pp. 7799–7803, 2014.
- [185] H. Arguello and G. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions on Image Processing*, vol. 23, no. 4, pp. 1896–1908, 2014.
- [186] Y. August and A. Stern, "Compressive sensing spectrometry based on liquid crystal devices," *Optics letters*, vol. 38, no. 23, pp. 4996–4999, 2013.
- [187] Y. August, C. Vachman, Y. Rivenson, and A. Stern, "Compressive hyperspectral imaging by random separable projections in both the spatial and the spectral domains," *Applied optics*, vol. 52, no. 10, pp. D46–D54, 2013.
- [188] Y. August, C. Vachman, and A. Stern, "Spatial versus spectral compression ratio in compressive sensing of hyperspectral imaging," in *SPIE Defense, Security, and Sensing*, pp. 87170G–87170G, International Society for Optics and Photonics, 2013.
- [189] G. Martin, J. M. B. Dias, and A. J. Plaza, "A new technique for hyperspectral compressive sensing using spectral unmixing," in *SPIE Optical Engineering+ Applications*, pp. 85140N–85140N, International Society for Optics and Photonics, 2012.
- [190] C. Li, T. Sun, K. Kelly, and Y. Zhang, "A compressive sensing and unmixing scheme for hyperspectral data processing," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1200–1210, 2012.
- [191] D. S. Kittle, D. J. Brady, S. Prasad, Q. Zhang, and R. Plemmons, "Joint segmentation and reconstruction of coded aperture hyperspectral data," in *Computational Optical Sensing and Imaging*, p. CMD1, Optical Society of America, 2011.
- [192] Q. Zhang, R. Plemmons, D. Kittle, D. Brady, and S. Prasad, "Joint segmentation and reconstruction of hyperspectral data with compressed measurements," *Applied Optics*, vol. 50, no. 22, pp. 4417–4435, 2011.

-
- [193] Q. Zhang, R. Plemmons, D. Kittle, D. Brady, and S. Prasad, "Reconstructing and segmenting hyperspectral images from compressed measurements," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on*, pp. 1–4, 2011.
- [194] K. Krishnamurthy, M. Raginsky, and R. Willett, "Hyperspectral target detection from incoherent projections," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 3550–3553, 2010.
- [195] S. Lim, C. Cho, A. Das, and S. Chai, "Coded aperture snapshot imaging based spectral classification," in *Computational Optical Sensing and Imaging*, pp. CTu2C–5, Optical Society of America, 2014.
- [196] N. Ly, Q. Du, and J. Fowler, "Reconstruction from random projections of hyperspectral imagery with spectral and spatial partitioning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 466–472, 2013.
- [197] A. Rajwade, D. Kittle, T. han Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing."
- [198] D. H. Foster, S. NASCIMENTO, and K. Amano, "Information limits on neural identification of colored surfaces in natural scenes," *Visual neuroscience*, vol. 21, no. 03, pp. 331–336, 2004.
- [199] T. Björklund and E. Magli, "A parallel compressive imaging architecture for one-shot acquisition," *arXiv preprint arXiv:1311.0646*, 2013.
- [200] Y. Takaki, "Electro-optical implementation of learning architecture to control point spread function of liquid crystal active lens," in *SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation*, pp. 205–214, International Society for Optics and Photonics, 1995.
- [201] V. Laude, J.-P. Huignard, M. Defour, and P. Refregier, "Optical image processing with the liquid crystal active lens," in *Lasers and Optics in Manufacturing III*, pp. 139–145, International Society for Optics and Photonics, 1997.
- [202] M. R. Beversluis, L. Novotny, and S. J. Stranick, "Programmable vector point-spread function engineering," *Optics express*, vol. 14, no. 7, pp. 2650–2656, 2006.



William Guicquero

Education

- 2011–2015 **Ph.D. in Electrical Engineering**, *Swiss Federal Institute of Technology in Lausanne (EPFL)*.
- 2009–2011 **M.Sc. in Computer Science and Applied Mathematics**, *Grenoble Institute of Technology - Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble (Grenoble INP - ENSIMAG)*.
- 2008–2009 **B.Sc. in Engineering Science**, *Grenoble Institute of Technology (Grenoble INP)*.
- 2006–2008 **Preparatory courses for French Elite Schools**, *Physics and Engineering Sciences (MPSI-PSI)*, Blaise Pascal, Orsay.

Research Interests

- Signal Processing Sparse representations, Non-linearity, Compression techniques, Image processing
- Algorithmic Optimization, Machine learning, Tracking
- μ -electronics Design of integrated circuits for imaging in CMOS technology

Ph.D. thesis

- Title *Exploring Information Retrieval using Image Sparse Representations: From Circuit Designs and Acquisition Processes to Specific Reconstruction algorithms*
- Supervisors Prof. Pierre Vanderghenst, Dr. Antoine Dupret
- Achievements Design of CMOS image sensors, Demonstrator, PCB, & FPGA programming
Theoretical study and practical implementation of image processing algorithms

M.Sc. thesis

- Title *Digital processing for a physical reconfigurable layer between three Radio-Frequency communication standards at 2.4GHz.*
- Supervisors Prof. Jean-Marc Brossier, Laurent Ouvry
- Achievements Reconfigurable multi-standards demodulator
RF Communication standard classifier

Experience

Vocational

2014–now **Research engineer**, *CEA-Leti*, Grenoble.
Design of CMOS image sensor & programming of image processing tools

2011–2014 **Ph.D. student**, *CEA-Leti*, Grenoble.
CMOS Compressive Sensing imaging & alternative sensing schemes

Internships & scholar projects

January–August 2011 **Internship**, *CEA-Leti*, Grenoble.
Digital processing for demodulation and supervised machine learning in the field of Radio-Frequency communications

Fall semester 2010 **Semester research project**, *EPFL*, Lausanne.
Evaluation and algorithmic optimisation of music similarities in mp3 compressed domain

Summer 2010 **Internship**, *CSTB*, Grenoble.
Optimisation of dynamical audio rendering for reverberant environments (GPU programming, OpenCL)

May 2010 **Scholar project**, *ENSIMAG*, Grenoble.
Simulation in C of error detection and correction performed at the demodulation stage during Compact Discs reading (CD & CD-ROM)

January 2010 **Scholar project**, *ENSIMAG*, Grenoble.
Programming in ADA of a compiler for an object-oriented “Java-like” language

Summer 2009 **Internship**, *Gipsa-Lab*, Grenoble.
Programming of an audio processing platform, embedded in a DSP

Teaching

2014 **Internship supervisor**, *CEA-Leti*, Grenoble.
Design and validation of an innovative CMOS image sensor architecture performing adaptive Compressive Sensing

Spring semester 2013 **Teacher Assistant**, *EPFL*, Lausanne.
Probabilities and statistics (B.Sc.), prof. S. Morgenthaler

Spring semester 2012 **Teacher Assistant**, *EPFL*, Lausanne.
Signal and Systems (B.Sc.), prof. P. Vandergheynst

Languages

French **Native speaker**
English **Proficient in spoken and written**
German **Beginner**

Computer skills

Programming Languages	Matlab, C, C++, QT, Python, Ada, OpenCL, Bash	Description languages	LaTeX, VHDL
Operating Systems	TangoStudio (CAM on Linux), Ubuntu, Windows 7	Softwares	Cadence-Virtuoso, Code::Blocks, Powerpoint

9 bis Rue Bizanet – 38000 Grenoble – France

☎ +33 (682) 890 229 • ☎ (pro)+33 (438) 780 957

✉ william.guicquero@cea.fr • Date of birth: 12th of November 1988

Personal Interests

- Sport Tai-Juitsu (more than 10 years of practice), Hiking, Skiing
Musique Western concert flute, Saxophone, Electric Guitar
Others Vinyl collection (Jazz-World-Electro), Cooking (bread draw), Card games (Coinche)

References

P. Vandergheynst, *Professor*, Institute of Electrical Engineering, EPFL, Switzerland.

E-mail: pierre.vandergheynst@epfl.ch

A. Dupret, *Doctor (HDR)*, Commissariat à l'Energie Atomique et aux Energies Alternatives, CEA-Leti, France.

E-mail: antoine.dupret@cea.fr

L. Alacoque, *Doctor*, Commissariat à l'Energie Atomique et aux Energies Alternatives, CEA-Leti, France.

E-mail: laurent.alacoque@cea.fr

Publications & Patents

W. Guicquero, A. Verdant, A. Dupret, and P. Vandergheynst, "Nonuniform sampling with adaptive expectancy based on local variance," in *accepted to 11th international conference on Sampling Theory and Applications (SampTA 2015)*, 2015.

A. Peizerat, F. Guezzi, M. Benetti, A. Dupret, R. Jalby, L. Bruno de Sa, W. Guicquero, and Y. Blanchard, "A 3t or 4t pixel compatible dr extension technique suitable for 3d-ic imagers : a 800x512 and 5 μ m pixel pitch 2d demonstrator," in *accepted to Circuits and Systems, 2015. ISCAS 2015. IEEE International Symposium on*, 2015.

W. Guicquero, P. Vandergheynst, T. Laforest, and A. Dupret, "On adaptive pixel random selection for compressive sensing," in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*, 2014, pp. 712–716. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7032211>

W. Guicquero, P. Vandergheynst, T. Laforest, A. Verdant, and A. Dupret, "On multiple spectral dependent blurring kernels for super-resolution and hyperspectral imaging," in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*, 2014, pp. 717–721. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7032212>

W. Guicquero, M. Benetti, A. Peizerat, A. Dupret, and P. Vandergheynst, "On compensating unknown pixel behaviors for image sensors with embedded processing," in *Signals, Systems and Computers, 2014 Asilomar Conference on*, 2014.

9 bis Rue Bizanet – 38000 Grenoble – France

☎ +33 (682) 890 229 • ☎ (pro)+33 (438) 780 957

✉ william.guicquero@cea.fr • Date of birth: 12th of November 1988

A. Dupret, A. Verdant, L. Millet, L. Alacoque, A. Peizerat, M. Benetti, T. Laforest, W. Guicquero, and M. Tchagaspian, "Processing next to the pixel: Lessons from some developments at cea-leti," in *Cellular Nanoscale Networks and their Applications (CNNA), 2014 14th International Workshop on*, 2014, pp. 1–2. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6888641>

W. Guicquero, A. Dupret, and P. Vandergheynst, "Multi-capture high dynamic range compressive imaging," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, 2013, pp. 143–147. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6810247>

—, "An adaptive compressive sensing with side information," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, 2013, pp. 138–142. [Online]. Available: <http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6810246>

W. Guicquero, T. Laforest, A. Verdant, and A. Dupret, "Système et procédé d'acquisition d'images hyperspectrales," Patent FR 1 454 453, 2014.

W. Guicquero and A. Dupret, "Dispositif d'acquisition compressive d'une image," Patent FR 1 358 672, 2013.

—, "Dispositif et procédé d'acquisition compressive d'images," Patent FR 1 358 674, 2013.