# Une Capteur d'Image Reconfigurable dédié à l'Imagerie Rapide, au Traitement d'Image Linéaire et Réseaux Convolutifs

Jérôme Dubois, Dominique Ginhac et Michel Paindavoine Laboratoire LE2I, UMR-CNRS 5158 Université de Bourgogne, Aile des Sciences de l'Ingénieur, BP 47870 21078 DIJON Cedex

E-mail:. jerome.dubois@u-bourgogne.fr

#### Résumé

Un capteur CMOS permettant l'acquisition d'image à très haute cadence et intégrant des prétraitements, a été conçu, fabriqué et validé expérimentalement. Ce capteur d'image particulier peut être vu comme une architecture massivement parallèle permettant l'exécution d'opérations mathématiques au niveau même du pixel. Des traitements programmables de bas niveau comme l'extraction de gradients spatiaux, allant du simple masque de convolution  $2 \times 2$  aux masques plus complexes N×N en passant par des filtres spatiaux de type Laplacien ou Sobel. À cette fin, chaque pixel inclut une photodiode, un étage de mise en forme, deux cellules multiplexées à mémoire analogique et amplificateur (MA2M) et une unité de calcul analogique basée sur une architecture de multiplieurs quatre quadrants. Le circuit ainsi fabriqué, et mis en œuvre, est bien une preuve que le concept est réalisable. Une plateforme a été réalisée à l'aide d'un FPGA, d'une carte d'acquisition 50 Mech/s, interfacé à notre rétine de résolution 64×64. Des résultats expérimentaux ont été obtenu, en prise d'image à cadence élevé, jusqu'à 10 000 images par seconde, et intégrant des traitements convolutifs à des cadences allant jusqu'à 5 000 images par seconde.

# 1. Introduction

Aujourd'hui, l'essentiel des travaux sur les systèmes de vision CMOS se concentre sur l'intégration d'un processeur et d'une matrice photosensible sur le même substrat (approche "System On Chip"). Souvent, dans ces systèmes sur puce, des éléments de calculs sont intégrés au niveau des colonnes de la matrice [1], [2]. Dans les capteurs CMOS, le traitement au niveau du pixel est généralement écarté car la taille des pixels est souvent trop grande pour que le capteur en question puisse avoir un avenir industriel. Néanmoins, l'utilisation d'un élément de traitement par Pixel donne la possibilité de réaliser massivement des calculs parallèles et ainsi d'exploiter les possibilités d'un capteur d'image haute cadence [3]. Ceci au bénéfice de l'implémentation de nouvelles applications complexes a cadence vidéo ou à l'amélioration d'applications vidéo existantes, telles que l'évaluation de vecteurs de mouvement [4], l'adaptation au conditions lumineuses [5], la capture de mouvement [6], ou l'identification de modèles

[7]. Notre travail s'articule donc autour de la faisabilité de rétines artificielles présentant des tailles de pixels raisonnables tout en y intégrant des traitements bas nivaux. C'est pourquoi, nous nous imposons de rester dans le monde analogique pour que ces traitements comportent un nombre minimum de transistors. Nous ne pouvons raisonnablement extrapoler linéairement nos travaux vers des technologies plus fine (90 nm ou encore 65 nm) sans se heurter à des difficultés technologiques, surtout dans le cas de fonctions analogiques. C'est pourquoi, nous nous somme imposé une taille de pixel de  $35 \,\mu$ m.

Dans cet article, nous évoquerons l'implémentation matérielle d'un système de vision embarqué. À cette fin, nous avons conçu, fait fabriqué et testé une rétine analogique de résolution 64×64 intégrant des traitements programmables au niveau du pixel. Les principaux objectifs de notre capteur sont : (1) évaluer la vitesse du capteur, en particulier à la cadence très élevée de 10 000 images par seconde, (2) démontrer les possibilités de reconfigurations dynamiques de l'unité de traitement au sein du pixel, (3) proposer un système de vision embarqué. La suite du papier est organisé comme suit. La section 2 est consacrée à la description du principe opérationnel de la rétine analogique. Les caractéristiques principales de l'architecture du capteur sont décrites dans la section 3. Les détails sur les mémoires analogiques et l'unité de calcul y sont successivement décrits. En conclusion, quelques résultats expérimentaux d'acquisition à très haute cadence avec traitements bas niveau sont présentés dans la dernière section de cet article.

# 2. Algorithme embarqué au niveau du pixel

Traditionnellement, un capteur CMOS peut être vu comme une matrice de pixels indépendants, chacun comportant un photo-détecteur (PD) et un processeur élémentaire (PE). Les travaux existants sur les traitements analogiques au niveau du pixel peuvent être classés en deux catégories. Les premiers sont intra-pixel, les traitements sont effectués sur les différents pixels afin d'améliorer la qualité de l'image, tels les pixels actifs (APS) [1], comme le montre la figure 1(a). Les seconds sont interpixel, ici les traitements sont basés sur des voisinages de pixels. Notre travail se place dans cette deuxième catégorie. Cette approche nous oblige à réfléchir sur la distribution spatiale des ressources de traitements. En conséquence, chaque élément de traitement doit être placé au milieu de quatre photodiodes connexes, comme le montre la figure 1(b), afin de limiter d'une part les interconnexions entre processeurs élémentaires et d'autre part de réfléchir à une forme optimal de photodiode. Notre objectif principal est l'exécution de divers traitements d'image in situ basés sur des voisinages locaux, tels que gradients spatiaux, opérateurs Sobel et Laplacien.



Figure 1. Photosite avec processeur : (a) intra-pixel, (b) inter-pixel

## 2.1 Gradient spatial

Notre structure de site de traitement connexe à quatre photo-détecteurs se prête très bien à la mesure du gradient spatial comme l'atteste la figure 2.



Figure 2. Mesure du gradient spatial

En effet, les amplitudes, de la dérivée première discrétisée [8], des variations du niveau de gris suivant les axe x' et y' sont données par la relation (1).

$$\frac{\partial V}{\partial \vec{\xi}} = (V_2 - V_4)\cos(\beta) + (V_1 - V_3)\sin(\beta)$$
(1)

Où  $\beta$  est l'angle formé avec la direction du vecteur contour X et  $V_i$ ,  $i \in \{1;4\}$ , la luminance du pixel i, i.e., le signal issu de la photodiode, mis en forme et amplifié. De cette façon, le dérivé local dans la direction du vecteur contour est la combinaison linéaire, dans une base orthonormée, des dérivées premières discrétisées dans les directions x' et y'. En utilisant quatre multiplieurs quatre quadrants [9] (voir la section 3.2 pour les détails concernant cette structure), la fonction P, donnée par la relation (2) peut être facilement implémentée comme le montre la figure 3. Cette opération demanderait l'équivalent de quatre multiplications signées sur une architecture externe.

$$P = V_1 \cos(\beta) + V_2 \sin(\beta) - V_3 \sin(\beta) - V_4 \cos(\beta)$$
 (2)



Figure 3. Implémentation de 4 multiplieurs 4 quadrants

En conséquence, le traitement mis en œuvre au niveau du pixel réalise une combinaison linéaire de quatre pixels adjacents pondérée par quatre coefficients (coef<sub>i</sub>,  $i \in \{1 ; 4\}$ ). Ainsi pour évaluer P, les coefficients forment une matrice de rotation, donnée par la relation (3) :

$$\begin{pmatrix} coef1 & coef2\\ coef3 & coef4 \end{pmatrix} = \begin{pmatrix} sin(\beta) & cos(\beta)\\ -sin(\beta) & -cos(\beta) \end{pmatrix}$$
(3)

## 2.2 Opérateur Sobel

La structure de notre architecture est également bien adaptée à divers algorithmes de convolution basés sur des voisinages de pixels. L'évaluation des amplitudes horizontales et verticales du gradient de Sobel demande seulement deux reconfigurations de l'Unités Arithmétiques Analogiques (UA<sup>2</sup>, voir section 3.2).

Par construction, les coefficients appliqués sont les mêmes dans toute la matrice, ce qui permet de paralléliser les calculs. De plus, à l'issu de la phase d'acquisition, des mémoires analogiques conservent l'information luminance dans chaque pixel, il est donc possible de reconfigurer plusieurs fois les coefficients appliqués.

A l'issu de la phase d'acquisition, dans une première lecture trame, pour évaluer, par exemple, la moitié gauche le l'amplitude verticale du gradient de Sobel, les coefficients doivent être configurés comme suit :

$$\begin{pmatrix} coef1 & coef2\\ coef3 & coef4 \end{pmatrix} = \begin{pmatrix} -1 & 0\\ -1 & 0 \end{pmatrix}$$
(4)

Pour poursuivre l'évaluation du gradient, il faut reconfigurer les coefficients en conséquence et relire la trame comme le montre la figure 4. Il suffit ensuite d'accumuler, à l'aide d'un processeur externe, les valeurs lues pour évaluer les amplitudes.





Figure 4. Séquence de reconfiguration dynamique pour le filtre de Sobel verticale, (a) première configuration et lecture trame, (b) deuxième configuration et relecture.

Ainsi, les unités de calcul implantées au sein du pixel allègent considérablement le travail du processeur externe (FPGA). Son rôle se limite à accumuler et stocker les valeurs issues des unités de traitement. L'opérateur Laplacien se traite de la même façon avec quatre reconfigurations dynamiques. En suivant cette ligne stratégique, il est possible d'augmenter significativement la vitesse des traitements d'image, par exemple, dans la réalisation de réseaux convolutifs à pas variable, c'est le cas du premier étage de l'algorithme de détection de visage CFF [10]. Dans cet exemple, l'objectif est d'alléger le travail du processeur externe et surtout de diminuer le temps d'exécution de l'algorithme.

## 3. Architecture du circuit

Comme dans un capteur traditionnel, le cœur du circuit, présenté dans ce papier, est constitué d'une matrice bidimensionnelle de pixels à accès aléatoire via un décodeur d'adresse en ligne et un multiplexeur de colonne. Le lay-out complet est représenté en figure 5 et ses principales caractéristiques regroupées dans le tableau 1.



Figure 5. Lay-out du circuit complet

Chaque pixel contient une photodiode et 38 transistors assurant les fonctions suivantes : transduction luminance tension, deux mémoires analogiques indépendantes avec amplificateur et multiplexeur intégré, dédiées à l'imagerie rapide (voir section 3.1), et une unité de traitement analogique, comme évoquée plus haut, constituée de quatre multiplieurs quatre quadrants (voir section 3.2).

Technologie	CMOS 0,35 µm
Résolution	$64 \times 64$
Surface	11 mm²
Nombre de transistors	160 000
Transistors par pixel	38
Taille des pixels	$35 \mu m  imes 35 \mu m$
Taux de remplissage	25 %
Consommation	110 mW
Tension d'alimentation	3,3 V

Tableau 1. Caractéristiques du circuit

# 3.1 Mémoire Analogique, Amplificateur et Multiplexeur MA<sup>2</sup>M

Ce capteur a initialement été conçu pour réaliser des captures d'image à très haute cadence. Pour doubler la vitesse d'acquisition, une stratégie simple consiste à séparer la phase d'acquisition de la phase de lecture. Le temps minimal d'intégration du photo-courant est estimé à  $100\mu$ s, dans des conditions d'éclairement favorable. De plus, dans ce même intervalle temporel, une lecture séquentielle d'une trame précédemment acquise est tout à fait envisageable compte tenue de la résolution du capteur.

Ce principe, illustré sur la figure 6, est appliqué au sein même du pixel à l'aide de deux structures intégrant une mémoire analogique, un amplificateur et un multiplexeur.



Figure 6. Parallélisme de tâches entre séquence d'acquisition et lecture séquentielle

Ces deux structures sont précédées d'un étage de mise en forme dont la sortie est polarisée en tension autour de  $V_{DD}/2$ . L'amplificateur n'est autre qu'un inverseur CMOS, qui utilisé en petits signaux autour de  $V_{DD}/2$ , se comporte comme un amplificateur inverseur linéaire de fort gain. Le circuit d'acquisition, et les deux structures, comme le montre la figure 7, ne comportent que 15 transistors.



Figure 7. Schéma du circuit d'acquisition

#### 3.2 Unité Arithmétique Analogique UA<sup>2</sup>

L'unité arithmétique analogique est le processeur élémentaire du pixel. Interconnectée à ces trois proches voisins, cette structure, qui ne compte qu'une vingtaine de transistors, est capable de réaliser une combinaison linéaire de quatre pixels connexes. A cette fin, une nouvelle structure de multiplieur quatre quadrants a été validée. Le schéma structurel compte 22 transistors, comme le montre la figure 8.



Figure 8. Schéma de l'Unité Arithmétique Analogique

#### 4. Résultats expérimentaux

Un prototype de système de vision embarqué a été réalisé dans notre laboratoire, constitué de la rétine et d'une carte à FPGA (Spartan-3 et 32 Mo de SDRAM). L'essentiel des mesures réalisées sur le capteur sont listées sur le tableau 2.

Gain de conversion	54 µV/e <sup>-</sup> RMS
Sensibilité	0,5 V/lux.s
Bruit spatial fixe (noir)	225 µV RMS
Bruit de reset	68 µV RMS
Disparité de linéarité	4,3 %
Gain en tension	12
Dynamique	68 dB

Tableau 2. Mesures expérimentales

La figure 9 montre des résultats de capture d'image issu du capteur : En (a), l'image brute avec un temps d'intégration de 100  $\mu$ s; en (b), le gradient de Sobel horizontale obtenu à l'issu de deux configuration de l'UA<sup>2</sup>, de même en (c), le gradient de Sobel verticale, et en (d), le Laplacien.



Figure 9. (a) Image brute, (b) Sobel horizontale, (c) Sobel verticale, (d) Laplacien

Le capteur a ainsi été validé expérimentalement.

#### 5. Conclusions

Un capteur d'image dédié à l'imagerie rapide et au traitement d'image linéaire a été fabriqué, testé et validé expérimentalement. Deux structures à mémoire MA<sup>2</sup>M (voir section 3.1) permettent d'atteindre des cadences allant jusqu'à 10000 images par secondes. Une architecture originale de multiplieur quatre quadrants à 5 transistors a été implémentée au sein même du pixel, permettant de réaliser des post traitements de bas niveau.

De larges perspectives sont ouvertes, notamment sur l'implémentation du premier étage d'un réseau convolutif à pas variable, par exemple, pour accélérer l'exécution de l'algorithme de détection de visage CFF [10], actuellement implanté exclusivement sur processeur externe.

La prochaine étape de nos recherches s'oriente sur la conception d'un capteur CIF  $(352 \times 258)$  en technologie CMOS 130 nm, intégrant des traitements basés sur la morphologie mathématique à haute cadence.

#### Références

[1] O. Yadid-Pecht and A. Belenky, "In-Pixel Autoexposure CMOS APS", IEEE Journal of Solid-State Circuits, vol. 38, no. 8, pp. 1425–1428, Août 2003.

[2] M. Sakakibara, S. Kawahito, D. Handoko, N. Nakamura, M. Higashi, K. Mabuchi, and H. Sumi, "A High-Sensitivity CMOS Image Sensor With Gain-Adaptative Column Amplifiers", IEEE Journal of Solid-State Circuits, vol. 40, no. 5, pp. 1147–1156, Mai 2005.

[3] Y. Sugiyama, M. Takumi, H. Toyoda, N. Mukozaka, A. Ihori, T. kurashina, Y. Nakamura, T. Tonbe, and S. Mizuno, "A High-Speed CMOS Image With Profile Data Acquiring Function," IEEE Journal of Solid-State Circuits, vol. 40, pp. 2816–2823, 2005.

[4] D. Handoko, K. S, Y. Takokoro, M. Kumahara, and A. Matsuzawa, "A CMOS image sensor for local-plane motion vector estimation," in Symposium of VLSI Circuits, vol. 3650, pp. 28–29, Juin 2000.

[5] D. Stoppa, A. Somoni, L. Gonzo, M. Gottardi, and G.-F. Dalla Betta, "Novel CMOS Image Sensor With a 132-dB Dynamic Range," IEEE Journal of Solid-State Circuits, vol. 37, no. 12, pp. 1846–1852, Décembre 2002.

[6] X. Liu and A. El Gamal, "Simultaneous image formation and motion blur restoration via multiple capture," in IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 3, pp. 1841–1844, 2001.

[7] C.-Y. Wu and C.-T. Chiang, "A Low-Photocurrent CMOS Retinal Focal-Plane Sensor With a Pseudo-BJT Smoothing Network and an Adaptative Current Schmitt Trigger for Scanner Applications," IEEE Sensors Journal, vol. 4, no. 4, pp. 510–518, Août 2004.

[8] M. Barbaro, P. Burgi, A. Mortara, P. Nussbaum, and F. Heitge, "A 100x100 Pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding," IEEE Journal of Solid-State Circuits, vol. 37, no. 2, pp. 160–172, Février 2002.

[9] S. Liu and Y. Hwang, "CMOS Squarer and Four-Quadrant Multiplier", IEEE Transactions on Circuits and Systems-I:Fundamental Theory and Applications, vol. 42, no. 2, pp. 119–122, Février 1995.

[10] C. Garcia, M. Delakis, Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, Novembre 2004.