HAL
open science

Conception en technologie CMOS d'un Système de Vision dédié à l'Imagerie Rapide et aux Traitements d'Images

Jérôme Dubois

▶ **To cite this version:**

Jérôme Dubois. Conception en technologie CMOS d'un Système de Vision dédié à l'Imagerie Rapide et aux Traitements d'Images. Micro et nanotechnologies/Microélectronique. Université de Bourgogne, 2008. Français. NNT : . tel-00365950

HAL Id: tel-00365950

https://theses.hal.science/tel-00365950v1

Submitted on 5 Mar 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Numero d'Ordre : xxxx
EDSPIC : xxx

UNIVERSITÉ DE BOURGOGNE

ÉCOLE DOCTORALE

E2S - ENVIRONNEMENT - SANTÉ/STIC

Formation Doctorale :

Instrumentation et Informatique de l'Image

Thèse

présentée par

Jérôme Dubois
(Normalien, Agrégé)

pour obtenir le grade de

DOCTEUR D'UNIVERSITÉ

(SPÉCIALITÉ : Micro-Électronique)

Conception en technologie CMOS d'un Système de Vision dédié à l'Imagerie Rapide et aux Traitements d'Images

Soutenue publiquement le 27 août 2008 devant le jury :

Monsieur	P. Magnan	Président
Monsieur	B. Zavidovique	Rapporteur
Monsieur	G. Cathébras	Rapporteur
Monsieur	M. Paindavoine	Directeur de thèse
Monsieur	D. Ginhac	Co-directeur de thèse

A ma famille.

Remerciements

Les Travaux présentés dans cette thèse ont été effectués au sein du Laboratoire LE2I (Laboratoire d'Électronique, d'Informatique et d'Image), UMR CNRS 5158 de l'Unniversité de Bourgogne. Cette thèse, qui a officiellement débuté le 1er septembre 2005, a été financée par une allocation couplée du contingent de l'ENS de CACHAN. Le thème de cette thèse est en parfaite adéquation avec mes travaux de stage de DEA III (Instrumentation et Informatique de l'Image) réalisés à l'issue de ma quatrième année de scolarité ENS.

Tout d'abord, je tiens à remercier Monsieur J.M. BILBAUT, Professeur à l'Université de Bourgogne et directeur du Laboratoire LE2I, pour son accueil au sein du laboratoire.

Je remercie Monsieur B. ZAVIDOVIQUE, Professeur à l'Université Paris-Sud, ainsi que Monsieur G. CATHEBRAS, Professeur à l'Université de Montpellier II, pour l'intérêt qu'ils ont porté à ce travail en acceptant d'en être rapporteurs.

Enfin, merci du fond du coeur à Monsieur M. PAINDAVOINE et Monsieur D. GINHAC respectivement Professeur et Maître de conférences à l'Université de Bourgogne pour l'attention constante avec laquelle ils ont suivi ces travaux, pour l'autonomie qu'ils ont su me donner et pour les qualités humaines dont ils ont su faire preuve durant cette période de ma vie.

Enfin, je ne terminerai pas sans remercier toutes les personnes, qui ont contribué, à divers titres , au bon déroulement de cette thèse et en particulier l'ensemble des membres du laboratoire LE2I.

Remerciements Personnels

Après Les remerciements officiels, viennent les remerciements personnels. Tous d'abord je tiens à remercier de fond du cœur mon épouse Mme DUBOIS Christelle qui a fait tant de sacrifices pour que nous quittions la région parisienne afin de s'installer quatre années à Dijon dans le but j'obtienne plus facilement une allocation couplée " les normaliens de Cachan sont gentiment priés de faire leur thèse en province ".

En effet cette grande dame a démissionné de son poste d'assistante de direction (groupe Carrefour) pour me suivre dans cette aventure de thésard qui a durée quatre longues années (DEA inclus). Déjà quatre ans de mariage et pour fêter nos noces de cire, cette grande dame vient de déposer à nouveau sa démission pour me suivre non loin de là où nous étions il y a quatre ans à Villebon sur Yvette, dans l'Essonne Nord. Je tiens à te dire ma chérie merci et je t'aime de tous mon cœur.

Je tiens également à remercier ma mère et ma sœur pour leur amour et leur soutient tout au long de ce périple.

J'adresse ma sympathie à tous les doctorants que j'ai eu l'occasion de croiser, Ludovic Journaux, Sami Boukhechem, Cedric Lemaitre, Nicolas Farrugia, Francois Tavin, Ambroise Marin, Romain Marinelle, Romuald Mosqueron et tous ceux que j'oublie dans cette liste non exhaustive.

Merci à tous et bonne lecture !

Table des matières

Liste des figures **5**

Liste des tableaux **11**

Introduction **15**

1 Imagerie rapide et capteurs "intelligents" **19**

1.1 Introduction 19

1.2 De Muybridge aux caméras rapides 19

1.2.1 Introduction 19

1.2.2 Traitements d'images dans les imageurs CMOS rapides 22

1.3 Résolution, tailles des pixels, les contraintes industrielles . . . 23

1.3.1 Introduction 23

1.3.2 Taille des pixels 24

1.3.3 La technologie CMOS s'essouffle 26

1.4 Pourquoi des capteurs "intelligents" ? 29

1.4.1 Les rétines artificielles 32

1.4.2 Rétine avec traitements dédiés 33

1.4.3 Rétines programmables 35

1.5 Conclusion 39

2 Technologie des capteurs CMOS **41**

2.1 Introduction 41

2.2 Chaîne de l'information 41

2.3 Photodiode verticale de type N 42

2.3.1 Premier modèle 43

2.3.2 Second modèle 46
2.4 Pixels analogiques 48
2.4.1 Pixel passif PPS 48
2.4.2 Pixel actif APS 49
2.5 Pixel numérique DPS 52
2.6 Les convertisseurs analogiques numériques dans les capteurs CMOS 56
2.6.1 Introduction 56
2.6.2 Technologies de CAN 58
2.7 Caractéristiques physiques du pixel à intégration 60
2.7.1 Réponse spectrale et efficacité quantique 60
2.7.2 Facteur de conversion 61
2.7.3 Courant d'obscurité 62
2.7.4 Dynamique 62
2.8 Le bruit dans les imageurs CMOS de type APS 63
2.8.1 Le bruit de grenaille (shot noise) 64
2.8.2 Le bruit en 1/f (flicker noise) 64
2.8.3 Le bruit de génération - recombinaison 65
2.8.4 Le bruit d'avalanche 65
2.8.5 Bruit thermique et de lecture (thermal/reset noise) . . 66
2.8.6 Bruit spatial fixe (BSF) 67
2.9 Compensation du bruit dans les imageurs CMOS à intégration 68
2.10 Conclusion 69

3 Architecture globale et algorithmes embarqués sur HISIC 71
3.1 Introduction 71
3.2 Architecture globale du système de vision artificielle 72
3.3 Traitements et algorithmes embarqués 76
3.3.1 Gradient spatial 77
3.3.2 Opérateur Sobel 80
3.3.3 Détecteur du second ordre : Laplacien 83
3.3.4 Généralisation 86
3.3.4.1 Filtres spatiaux 86
3.3.4.2 Application aux réseaux convolutifs 87

3.3.5 Détection de mouvement 90
3.4 Conclusion 92

4 Conception microélectronique d'HISIC 93
4.1 Introduction 93
4.2 Conception du pixel 94
4.2.1 Photo-détecteur 95
4.2.1.1 Choix technologique 95
4.2.1.2 Réflexion sur la géométrie de la photodiode . 96
4.2.1.3 Modèle théorique déduit des différences finies 101
4.2.2 Électronique de lecture et de traitement au sein du pixel 104
4.2.2.1 Transduction éclairement/signal 105
4.2.2.2 Mémoires Analogiques, Amplificateurs et Multiplexeurs (MA^2M) 107
4.2.2.3 Unité Arithmétique Analogique (UA^2) 109
4.3 Architecture du convertisseur analogique-numérique 117
4.3.1 Architecture du comparateur 118
4.3.1.1 Étage OTA : Amplificateur Opérationnel à Transconductance 120
4.3.1.2 Étage PD : Paire Différentielle 122
4.3.1.3 Étage AI : Amplificateur Inverseur 123
4.3.2 Stratégie de commande 123
4.4 Décodeur de lignes et multiplexeur de colonnes 125
4.5 Conclusion 127

5 Réalisations et résultats expérimentaux 129
5.1 Introduction 129
5.2 Projet HISIC 131
5.2.1 Plate-forme expérimentale d'instrumentation et de caractérisation du circuit HISIC.I 131
5.2.1.1 Erreurs d'antenne 131
5.2.1.2 Effets d'antenne 133
5.2.1.3 Améliorations 134
5.2.1.4 Exploitation 134
5.2.2 Système de vision artificielle HISIC.II 135

5.3 Caractérisation du système 138
5.3.1 A l'échelle du pixel 138
5.3.1.1 Validation expérimentale de la forme octogonale 138
5.3.1.2 Circuit d'acquisition et cellules MA^2M 140
5.3.1.3 Unité arithmétique analogique, multiplieurs quatre quadrants 144
5.3.2 A l'échelle du capteur 146
5.3.2.1 Bruit spatial fixe et bruit temporel 146
5.3.2.2 Compensation du bruit par CDS 147
5.3.2.3 Convertisseur analogique numérique 149
5.4 Traitements d'images embarqués et exemples d'applications . . 152
5.4.1 Capture d'images à haute cadence 152
5.4.1.1 Texture d'une goutte de lait 152
5.4.1.2 Ventilateur 154
5.4.1.3 Chenillard à DEL 155
5.4.1.4 Tapette à souris 156
5.4.2 Seuillage 158
5.4.3 Gradients spatiaux 159
5.4.4 Gradient morphologique temporel 160
5.4.5 Comparaison avec une architecture classique 161
5.5 Conclusion 162

Conclusion et perspectives **163**

Bibliographie **167**

Publications **177**

Glossaire **179**

Annexe **179**

A Paramètres et grandeurs physiques **183**

B Carte XESS 3S1000 et pilote SVGA **185**
B.1 Description du système 185

B.2 Le pilote SVGA 185

Table des figures

1 Marqueurs placés sur un rat 16

1.1 Séquence d'images d'un cheval au galop 20
1.2 Évolution des parts de marché entre les technologies CCD et CMOS (source : iSupply) 21
1.3 Sur-échantillonnage et traitements d'images à l'échelle temporelle de la cadence vidéo 23
1.4 Angle de champ d'une lentille convergente 24
1.5 Taille du pixel dans les deux technologies CCD et CMOS (source : Micron) 25
1.6 Coupe transversale d'un transistor NMOS en technologie standard CMOS 26
1.7 Synoptique d'un système de vision classique 29
1.8 Coupe transversale de l'œil humain 32
1.9 (a) Dessin de masque d'une rétine à résolution fovéale , (b) agrandissement de la partie centrale 34
1.10 Synoptique d'un processeur analogique 35
1.11 Synoptique d'une architecture distribuée dans un capteur CMOS 36
1.12 Pixel numérique programmable 36

2.1 Architecture d'un imageur CMOS 42
2.2 Effet photoélectrique dans un semi-conducteur 43
2.3 Photodiode verticale de type N, (a) phénomène contribuant au courant photonique , (b) caractéristique I=f(V_d), (c) schéma équivalent dans le 3eme quadrant. 44
2.4 Densités de courant dans une photodiode N verticale 46
2.5 Architecture du pixel passif (PPS) 49
2.6 Architecture du pixel actif (APS) 50

2.7 Phase de préchargement puis intégration du photocourant . . 51
2.8 Architecture du pixel numérique (DPS) 52
2.9 Synoptique d'un convertisseur sigma-delta du premier ordre . 53
2.10 Graphe du bit de poids faible (bit 0) et schéma structurel . . . 54
2.11 Schéma à transistors du pixel numérique DPS de Kleinfelder . 55
2.12 Dessin de masque de 4 pixels connexes (technologie 0,18 μm) . 55
2.13 Hiérarchie architecturale dans un capteur CMOS 57
2.14 Synoptique du capteur CMOS industriel MTM9413 de Micron 57
2.15 Synoptique d'un convertisseur flash 58
2.16 Synoptique d'un convertisseur pipeline 59
2.17 Réponses caractéristiques de différents capteurs 62
2.18 (a) Bruit spatial fixe de pixel , (b) bruit spatial fixe de colonne , (c) bruit spatial fixe total 68
2.19 Technique de réduction du bruit dans les pixels APS 69

3.1 Architectures de rétines CMOS programmables 72
3.2 Schéma bloc du circuit HISIC 73
3.3 Schéma bloc du pixel d'HISIC 74
3.4 Synoptique du pixel dédié à l'imagerie rapide 74
3.5 Parallélisme de tâches entre séquences d'acquisitions et lectures séquentielles (stratégie Ping-Pong) 75
3.6 Photosite avec processeur : (a) intra-pixel, (b) inter-pixel . . . 77
3.7 Mesure du gradient spatial 78
3.8 Implémentation d'une Unité Arithmétique Analogique ou UA^2, composée de quatre multiplieurs quatre quadrants, au niveau du pixel 79
3.9 (a) Voisinage de pixels, (b) masque 3×3 appliqué localement autour de quatre éléments de calcul 80
3.10 Schéma bloc du système d'acquisition 81
3.11 Séquence de reconfiguration dynamique pour la composante verticale du filtre de Sobel 83
3.12 Position du point courant dans le masque de base 2×2 dans le cas d'une convolution de dimension 3×3 85
3.13 Calcul d'un masque quelconque de dimension 3×3 87
3.14 Stratégies de configuration appliquées au premier étage d'un réseau convolutif 89

3.15 Synoptique d'un capteur HDV rapide et intelligent 90

4.1 Courbe d'évolution du nombre de transistors par pixel dans les architectures distribuées dans le cas de processeurs analogiques et de processeurs numériques de 1995 à 2012, pour des pixels de 5 μm de côté et un facteur de remplissage de 30% . . 95
4.2 Vue en coupe de la technologie de fabrication CMOS AMS 0,35 μm (double-poly, quadruple-métal) 96
4.3 Règles de dessin appliquées à une photodiode de type N 98
4.4 (a) forme carrée, (b) forme en croix, (c) forme octogonale . . . 99
4.5 Graphe de la surface passive *Spa* pour les trois formes géométriques en fonction du paramètre a, la longueur du côté intérieur de la forme carrée 100
4.6 Topologie 4-connexe avec une photodiode octogonale 101
4.7 (a) Matrice de pixels basée sur une géométrie octogonale, (b) repère orthonormé pour l'évaluation du gradient spatial 101
4.8 Vue en trois dimensions d'une photodiode verticale de type N 102
4.9 Distribution du potentiel électrostatique dans une photodiode verticale de type N 103
4.10 Schéma électronique du premier étage et son dessin des masques (à droite) 106
4.11 Schéma du premier étage du pixel et des deux structures MA^2M 108
4.12 Layout du pixel avec le circuit d'acquisition et les deux cellules MA^2M 109
4.13 L'unité Arithmétique Analogique (UA^2) 110
4.14 ((a)Transistors NMOS et PMOS connectés en charge active, (b) circuit équivalent petits signaux 111
4.15 Schéma complet de l'Unité Arithmétique Analogique (UA^2) : 4 multipieurs 4 quadrants réalisant la combinaison linéaire de 4 pixels adjacents 112
4.16 MOS en régime triode, I_D faible 113
4.17 MOS contrôlé comme une résistance linéaire 114
4.18 Layout d'un pixel et de ces quatre photodiodes connexes . . . 116
4.19 Synoptique de l'architecture de convertisseur 118
4.20 Dessin des masques en technologie CMOS 0,35μm du convertisseur analogique numérique 119
4.21 Schéma électronique du comparateur 119

4.22 Modèle petits signaux avec prise en compte de l'effet de substrat 121
4.23 Capacité de charge de l'OTA en fonction de la tension de saturation 121
4.24 Signal appliqué en entrées des comparateurs 124
4.25 Architecture du décodeur de lignes 126
4.26 Architecture du multiplexeur de colonne 126

5.1 Dessin des masques de 6 pixels de HISIC.I 129
5.2 Dessin des masques du circuit HISIC.I 130
5.3 HISIC.I et son interface pour l'instrumentation 132
5.4 Correction des erreurs d'antenne 133
5.5 Dessin des masques de la rétine HISIC.II 135
5.6 Schéma bloc de la platforme expérimentale 136
5.7 Illustration photographique de l'ensemble des circuits du système 136
5.8 (a) Illustration photographique du prototype de caméra rapide et intelligente, (b) Circuit HISIC.II, sous boîtier LCC48. . . . 137
5.9 Dessin des masques des pixels tests de HISIC.I 139
5.10 Réponse spectrale en A/W des photodiodes aux formes carrées et octogonales 140
5.11 Graphe des signaux caractéristiques du pixel test octogonal (HISIC.I) avec un temps d'intégration de 100μs, pour une longueur d'onde de 680 nm et un éclairement de 5000 lux. 141
5.12 Couplage capacitif entre deux segments de métal de niveau différent et qui se croisent. 142
5.13 Résultats expérimentaux du multiplieur quatre quadrants . . . 145
5.14 Correction du bruit à l'aide des cellules MA^{2}M et de l'UA^2 . . 148
5.15 Acquisition d'image en l'absence de lumière avec un temps d'intégration de 100 μs, (a) image de bruit fixe et temporel sans correction, (b) image corrigée par CDS. 148
5.16 Résultats expérimentaux sur le temps de réponse du comparateur en fonction de la tension différentielle d'entrée 150
5.17 Erreurs différentielles et intégrales du convertisseur en fraction de LSB, en fonction de la sortie codée sur 8 bits. 150
5.18 Plate-forme expérimentale d'acquisition d'une goutte de lait . 152
5.19 Résultats expérimentaux d'acquisitions haute cadence d'une goutte de lait 153

5.20 Image du ventilateur à l'arrêt 154
5.21 16 images consécutives d'un ventilateur en rotation à 3750 tr/min, avec un temps d'intégration de 500 μs. 155
5.22 Chenillard à 12 DEL 155
5.23 16 images consécutives du chenillard, avec un temps d'intégration de 100 μs 156
5.24 Image de la tapette à souris à l'arrêt 156
5.25 Déploiement de la tapette filmée à 2000 im/s 157
5.26 12 images consécutives du déploiement de la tapette avec saturation et un temps d'intégration de 500 μs. 157
5.27 16 images consécutives du ventilateur en rotation à 3750 tr/min, avec un temps d'intégration de 500 μs et une opération de seuillage sur les images. 158
5.28 12 images consécutives de la tapette à souris, avec un temps d'intégration de 100 μs et une opération de type Laplacien sur les images, soit au total 500 μs par trame. 159
5.29 12 images consécutives de la tapette à souris, avec un temps d'intégration de 100 μs et une opération de morphologie temporelle sur les images, soit au total 500 μs par trame. 160

B.1 Illustration photographique de l'intérieur du système de vision 186
B.2 Illustration photographique de l'utilisation du système de vision 186
B.3 Schéma de connexion du port VGA 187

Liste des tableaux

1.1 Évolution des paramètres physiques de la technologie CMOS (analogique) de 2003 à 2016 (source : ITRS) 28
1.2 Comparaisons de capteurs d'images programmables 38

2.1 Caractéristiques du capteur à DPS de Kleinfelder 55
2.2 Récapitulatif des différentes implémentations de CAN 59

3.1 Algorithme morphologique spatio-temporel 91

5.1 Caractéristiques du circuit HISIC.II 137
5.2 Comparaisons avec quelques capteurs d'images 143
5.3 Récapitulatif des mesures expérimentales sur le pixel test octogonal (HISIC.I) et sur la rétine HISIC.II, avec un temps d'intégration de 100 μs, pour une longueur d'onde de 680 nm et un éclairement de 5000 lux. 146
5.4 Comparaisons de capteurs d'images 151
5.5 Coefficient de corrélation entre les images traitées par la rétine et les images traitées par une unité de calcul numérique externe 161

A.1 Paramètres et grandeurs physiques en technologie CMOS AMS 0,35 μm 184

Introduction

La science et la technologie permettent d'apprécier l'infini, l'invisible, l'incommensurable. Après la cinématographie rapide au XIXe siècle, l'imagerie rapide électronique dans la deuxième moitié du XXe siècle est devenue l'une des technologies permettant de voir l'invisible. En effet, la persistance rétinienne de l'homme est au mieux de 18 ms. Ce qui signifie que sans assistance technologique, on est incapable de distinguer nettement un objet mobile qui oscille à une certaine fréquence (à partir de 28 Hz). Cette propriété est utilisée dans tous les systèmes d'affichages électroniques allant de la simple montre à cristaux liquides jusqu'aux écrans à plasma.

Cette partie de la science, l'imagerie rapide, est exploitée dans nombre de secteurs économiques comme l'industrie du sport, la médecine ou encore la recherche en matériaux. L'imagerie rapide est très utile pour l'étude de nombreuses technologies comme l'aviation, la robotique ou l'automobile. Dans le cas de l'aviation, on étudie le vol des insectes pour s'en inspirer et concevoir des drones[1] de plus en plus performants. Dans cet autre exemple, une séquence d'images rapides représente un rat qui se déplace à vitesse élevée sur un tapis roulant (cf. figure 1). Le but est d'étudier sa locomotion qui, dans le cadre de la recherche fondamentale en neurologie, constitue un paramètre essentiel, notamment lorsqu'il s'agit d'évaluer les potentialités thérapeutiques de futurs médicaments. On peut ainsi détecter les premiers signes d'une pathologie, quantifier l'ampleur d'un déficit et finalement mesurer un niveau de récupération. Dans ce contexte, l'utilisation de systèmes d'analyse d'images rapides donne accès aux différents paramètres de la locomotion chez le rongeur, tels que la cadence de la marche, la longueur et la durée du pas, la durée relative des deux phases du cycle (appui et oscillation), les variations angulaires des différents segments anatomiques de l'appareil locomoteur pendant la marche, etc.

Dans les années 1970, l'âge d'or du transistor bipolaire, des capteurs

[1]Un drone ou un UAV (Unmanned Air Vehicle) est un aéronef inhabité, piloté à distance, semi-autonome ou autonome

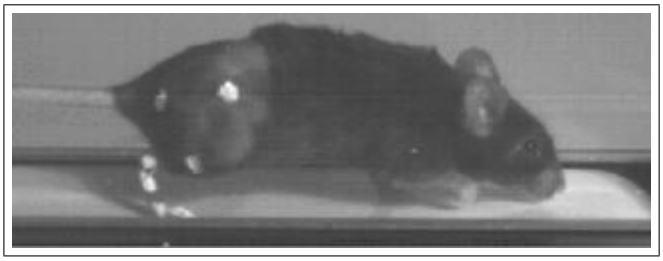


FIG. 1 – Marqueurs placés sur un rat

d'images dits CCD (Charge Coupled Devices) à transfert de charges font leur apparition. Plus récemment, ils se voient détrônés par la technologie CMOS (Complementary Metal Oxyde Semiconductor) qui présente l'avantage d'intégrer sur la même puce la fonction de prise d'images et de pré-traitement de l'information. Il est possible aujourd'hui d'industrialiser des circuits qui intègrent un capteur d'images et une électronique mixte de contrôle et de traitement sur la même puce. Dans un futur proche, pour quelques euros et cela grâce à la production de masse, nous disposerons de capteurs d'images rapides et programmables.

Néanmoins, il est raisonnable d'émettre certaines réserves quant à la quantité de traitements implantables dans le site de traitement, au sein du pixel. En effet, la première originalité de ce travail est de ne pas figer dans le silicium des traitements dédiés à une application, afin de garder une souplesse et une polyvalence maximale du capteur créé. Cependant, la petite capacité des éléments de calcul constitue un frein réel pour réussir efficacement l'intégration d'algorithmes multiples au sein du même capteur. Ces limitations impliquent inévitablement de restreindre notre champ d'investigation à des classes particulières d'applications modélisables selon un principe commun : le domaine des systèmes linéaires appliqués au traitement des images.

HISIC (High Speed Image Capture with processing at pixel level), du français Capture d'Images à Haute Cadence avec traitement au niveau du pixel, est l'acronyme utilisé pour désigner notre rétine. HISIC est le nom des deux circuits HISIC.I et HISIC.II que nous avons réalisés au cours de cette thèse. Le projet HISIC se propose de rechercher, développer et démontrer de nouvelles solutions qui permettront de résoudre les problèmes rencontrés dans la conception de capteurs d'images à haute résolution et à haute cadence. Le but de ce projet est la réalisation d'ici à 2013 d'un capteur d'images de résolution HDV 1440×1080 à la cadence de 10 000 images par seconde et intégrant des traitements d'images (reconnaissance de formes et suivi d'objets en mouvements rapides), en vue de son éventuelle industrialisation future. Pour l'heure l'originalité de nos travaux est de proposer un capteur d'images

qui soit à la fois rapide et qui intègre des traitements programmables ou reconfigurables.

Nous avons conçu, fait fabriquer et testé, une rétine analogique, en technologie CMOS 0,35 μm (double-poly, quadruple-métal), intégrant des traitements programmables au niveau du pixel. Ce capteur d'images 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 niveaux comme l'extraction de gradients spatiaux, allant du simple masque de convolution 2×2 aux masques plus complexes N×N en passant par des filtres spatiaux de type Laplacien ou Sobel. À cette fin, chaque pixel de taille 35μm×35μm inclut une photodiode, un étage de mise en forme, deux cellules mémoires multiplexées (MA^{2}M) et une unité de calcul analogique basée sur une architecture de multiplieurs quatre quadrants (UA^2). Cette rétine analogique permet l'accès aléatoire aux pixels via trois bus analogiques asynchrones : la première sortie est dédiée aux traitements, les deux autres à l'imagerie rapide ou à l'échantillonnage corrélé. C'est une machine massivement parallèle de 4096 processeurs arithmétiques analogiques inter-connectés selon une grille 64×64 en topologie 4-connexe. Le circuit ainsi fabriqué et mis en oeuvre est un premier résultat encouragent. Concrètement, nous avons réalisé un système embarqué à l'aide d'une carte à FPGA interfacée à notre "rétine" CMOS. Nous avons obtenu des résultats expérimentaux en prise d'images à la cadence de 10 000 images par seconde et intégrant des traitements convolutifs à des cadences allant jusqu'à 2 000 images par seconde.
Les principaux objectifs de nos travaux sont :

1. Evaluer la vitesse du capteur, en particulier à la cadence de 10 000 images par seconde avec un éclairement inférieur à 5000 lux.
2. Démontrer les possibilités de reconfigurations dynamiques de l'unité de traitement au sein du pixel.
3. Proposer un démonstrateur pré-industriel de système de vision artificielle.

Ce mémoire est sub-divisé en cinq chapitres de taille comparable. En premier lieu, un bref historique de l'imagerie rapide donne des ordres de grandeur et nous positionne dans l'étude des capteurs d'images rapides. Les concepts d'imagerie rapide et de rétines artificielles sont abordés dans ce premier chapitre. Le second décrit analytiquement la photodiode verticale de type N, le photodétecteur utilisé dans la conception de notre capteur. Ce chapitre porte également sur les technologies de pixels, les convertisseurs analogiques numériques et s'attarde sur les caractéristiques du pixel à intégration. Le

troisième chapitre présente l'architecture globale de notre capteur et les traitements spatiaux/spatio-temporels pouvant y être exécutés. Le quatrième insiste sur la conception microélectronique du circuit HISIC et enfin, un dernier chapitre est consacré aux résultats de mesures expérimentales et à des exemples d'applications de notre rétine.

Chapitre 1

Imagerie rapide et capteurs "intelligents"

1.1 Introduction

Notre travail s'articule autour de la conception d'un capteur d'images en technologie CMOS qui se veut à la fois rapide et "intelligent" au sens où des pré-traitements de l'image sont implémentés au sein même du pixel. Dans la littérature, il existe des réalisations de capteurs rapides [KVBA+99, Sei00, Lit01] et d'autres solutions où l'intégration de traitements au plus près du pixel est privilégiée [LSB+98, SHT+98, EGYF99]. Cependant, rares sont les capteurs intégrant les deux spécificités notamment dans le domaine de l'imagerie rapide où de tels capteurs sont quasi inexistants.

Dans ce chapitre, nous présentons succinctement un historique de l'imagerie rapide et son intérêt pour le traitement d'images à la cadence vidéo. Une troisième partie relate les contraintes industrielles et l'évolution de la technologie CMOS. Enfin une dernière partie évoque les raisons pour lesquelles l'étude des capteurs "intelligents" reste toujours séduisante.

1.2 De Muybridge aux caméras rapides

1.2.1 Introduction

La cinématographie rapide débute en 1872 avec les travaux célèbres du photographe américain Eadweard Muybridge, pionnier de la photographie en instantané, qui obtint dès 1878 une séquence d'images d'un cheval au galop (cf. figure 1.1). Pour réaliser son expérience, il aligna 24 appareils

photo et les déclencha l'un après l'autre en une fraction de seconde. Il put ainsi montrer que pendant un court instant, un cheval au galop ne touche pas le sol. Muybridge considérait ses clichés comme une série d'images distinctes mais il venait de tourner le premier film de l'histoire. Même s'il ne concevait pas ses travaux comme des films, il avait accompli quelque chose d'important. Il fut l'un des premiers à comprendre que la manipulation du temps nous donne accès à de nouveaux aspects de la réalité et nous révèle des beautés qui nous sont imperceptibles.

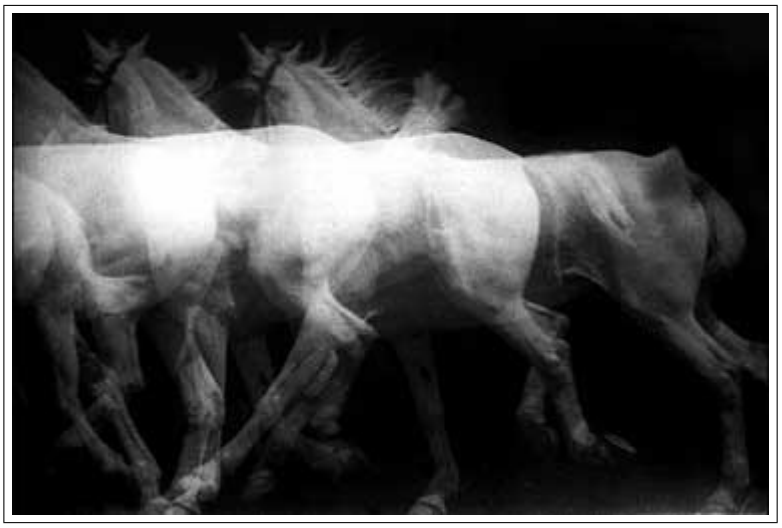


FIG. 1.1 – Séquence d'images d'un cheval au galop

Avant la deuxième moitié du XXe siècle, des caméras mécaniques à miroirs tournants répondaient à la plupart des besoins scientifiques. Les cadences d'acquisition pouvaient atteindre 10 000 images par seconde mais nécessitaient des moyens techniques importants. Le nombre d'images acquises se réduisait à une centaine et le développement du film photographique imposait une manipulation en chambre noire.

Au cours des 30 dernières années, ces deux inconvénients ont été éliminés par l'apparition des caméras utilisant des matrices à transfert de charges CCD. La lecture séquentielle des pixels (les uns après les autres) de cette matrice limitait la cadence des caméras rapides, mais de nouveaux capteurs d'images CCD à sorties multiples ont été proposés par les constructeurs. Par exemple, Eastman KODAK Company dépose en janvier 2008 un brevet d'invention [Par08] concernant un capteur d'images couleur à couplage de charge à sorties multiples. En reliant des circuits adjacents horizontalement, l'inventeur propose un capteur CCD à accès parallèle en lignes.

Durant ces vingt dernières années, les capteurs CMOS [AW96] à accès

aléatoires en lignes et en colonnes ont supplanté la technologie CCD. Du fait des progrès technologiques importants de l'industrie des semi-conducteurs, de tels capteurs atteignent aujourd'hui des performances équivalentes à celles des capteurs CCD, jusqu'à les dépasser dans des domaines applicatifs (acquisitions vidéos) où le CCD régnait en maître jusqu'à présent. Pierre MAGNAN propose en 2003, une étude comparative des deux technologies [Mag03]. Il explique notamment que dans le cas des pixels à photodiodes, le schéma CMOS se simplifie puisqu'il n'y a pas de transfert de charge (CCD) et que la photodiode détermine à la fois le rendement de détection et le gain de conversion par la valeur de sa capacité.

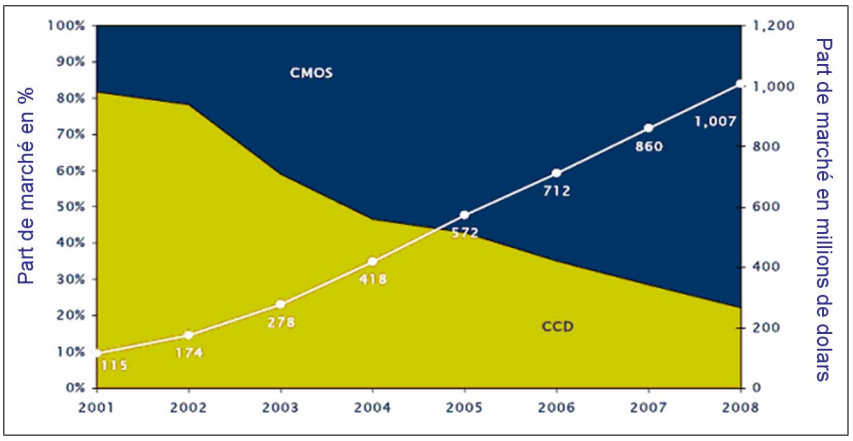


FIG. 1.2 – Évolution des parts de marché entre les technologies CCD et CMOS (source : iSupply)

La tendance du marché des capteurs d'images s'est inversée dans le courant de l'année 2004. La figure 1.2 montre l'évolution des parts de marché entre les deux technologies. Aujourd'hui le marché des capteurs CMOS pèse près d'un milliard de dollars et représente plus de 85% des ventes de capteurs d'images dans le monde.

Le développement des technologies sub-microniques à large intégration est un des facteurs convergents vers la recherche et développement sur les capteurs d'images CMOS. Aujourd'hui, les capteurs intégrés monolithiques à pixels actifs (MAPS) présente l'avantage intrinsèque d'être aussi granulaires et minces que les CCD. L'apport crucial de la technologie CMOS sur la technologie CCD est, d'une part l'accès aléatoire aux pixels et d'autre part, la possibilité d'intégrer des fonctions électroniques au sein même du pixel [Fos93, KL95, Fos97, Sei00, Lit01]. Les capteurs CMOS ont des avantages déterminants : coût très favorable, résolution potentiellement élevée, faible consommation électrique. La possibilité d'intégrer des traitements

au sein du pixel est particulièrement important pour l'implémentation de systèmes complets de caméras digitales requérant un processeur de traitement [LSB+98, SHT+98, EGYF99]. Des caméras CMOS rapides ont ainsi été réalisées. A titre d'exemple Krymski [KVBA+99] conçoit une caméra rapide CMOS en 1999. La résolution est alors de 1024×1024 et la fréquence trame de 500 images par seconde. En France, une des premières rétines opérationnelles de grande taille est proposée par Thierry Bernard et Bertrand Zavidovique en 1993 [BZD93].

Depuis de nombreuses années, le laboratoire LE2I (Laboratoire d'Électronique et d'Informatique de l'Image, UMR CNRS 5158) de l'Université de Bourgogne s'est intéressé à la mise en oeuvre de systèmes électroniques et informatiques dédiés au traitement du signal et de l'image et plus particulièrement à la conception de capteurs d'images CMOS dédiés à la vision artificielle. On peut citer par exemple la rétine d'intercorrélation [ABLYV+04] et la rétine de vidéo rapide dédiée à l'acquisition de phénomènes lumineux rapides [BLLYVC03] (équipe creusotine) avec la collaboration de Guy Cathébras. On peut citer aussi la rétine dédiée à la localisation et à la reconnaissance de visages [GPPH05] et plus récemment, le développement d'un système de vision dédié à l'analyse de mouvement rapide [Mos06] (équipe dijonaise).

Le défi de ces dix prochaines années sera d'associer vitesse d'acquisition et traitement de l'image au sein même des capteurs industriels "intelligents".

1.2.2 Traitements d'images dans les imageurs CMOS rapides

L'intérêt de faire des acquisitions à haute cadence est la possibilité d'échantillonner l'information optique à une fréquence très supérieure à la cadence vidéo de 25 images par seconde. La cadence vidéo est directement liée à l'appréciation d'une séquence d'images traitée par le cerveau humain. Il existe une vision simple et objective par l'œil et une vision plus complexe et subjective réalisée par l'analyseur, dit cognitif, situé dans le cortex occipital à l'arrière du cerveau. La première est définie par la persistance rétinienne qui est au mieux de 18 ms, c'est une vision simple et objective par l'œil. La seconde est une vision cérébrale qui est beaucoup plus lente (50 à 100 ms par image). C'est pourquoi un flot d'images à la cadence vidéo (40 ms/image) nous semble continu. L'ambition des systèmes d'acquisition à base de rétines CMOS est de profiter de la vitesse du capteur pour réaliser, à la cadence vidéo, des traitements sur les images acquises [Lim03] (cf. figure 1.3). Afin de raréfier les données avant qu'elles ne soient transmises à une unité de cal-

cul externe à la rétine, des traitements spatiaux et/ou spatio-temporels sont effectués sur le flot d'images haute cadence.

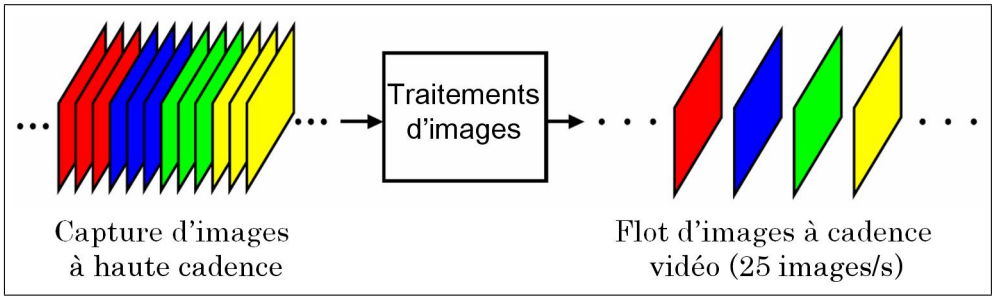


FIG. 1.3 – Sur-échantillonnage et traitements d'images à l'échelle temporelle de la cadence vidéo

Le sur-échantillonnage d'images permet l'implémentation de nouvelles applications complexes à cadence vidéo ou l'amélioration d'applications vidéo existantes, telles que l'évaluation du flot optique [HST+00, LEG01a, LEG01c], l'adaptation aux conditions lumineuses [HST+00, LEG01a, LEG01c], la capture de mouvements [LEG01b], ou l'identification de modèles [WC04]. La recherche actuelle s'oriente aussi vers le pixel "numérique" [YFEG99, KLLEG01a, HAB+05, CMC+06]. Dans les applications liées à l'automobile, l'ambition des capteurs d'images rapides et intégrant des traitements est la capacité à réagir plus rapidement que le cerveau humain, par exemple, dans le contrôle de trajectoires pour éviter un obstacle sur la route.

Une rétine CMOS est avant tous un capteur d'images réalisé en technologie CMOS. Dans ce travail de thèse, nous nous intéressons à la conception de systèmes de vision artificielle basé sur la technologie des capteurs CMOS.

1.3 Résolution, tailles des pixels, les contraintes industrielles

1.3.1 Introduction

L'avènement du format haute résolution HDV (1280x720 et 1440x1080) a accéléré le pas ces dernières années. Les industriels investissent dans la recherche et le développement de capteurs à haute résolution, de grande dynamique ($\geq$120dB), rapides (plusieurs centaines d'images pas seconde) et intelligents au sens qu'ils intègrent des pré-traitements sur l'image acquise.

1.3.2 Taille des pixels

La première contrainte à laquelle nous sommes confrontés lors de la conception d'un capteur "intelligent" est la taille des pixels. Elle est fixée par l'optique de la caméra. Les objectifs des caméras sont des assemblages complexes de lentilles, qui se comportent comme une lentille convergente à focale variable. Une focale dite normale est une focale dont l'angle de champ se rapproche de celui de l'œil humain, pour une vision à distance normale. Il est d'approximativement de 50 degrés. L'angle de champ, noté β (cf. figure 1.4), est donné par la relation 1.1.

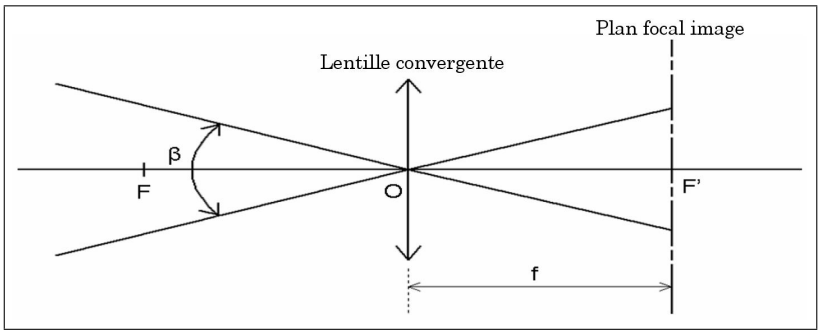


FIG. 1.4 – Angle de champ d'une lentille convergente

$$\tan(\frac{\beta}{2}) = \frac{D}{2\ f} \tag{1.1}$$

Où f est la distance focale équivalente de l'objectif et D est la diagonale du capteur d'images. Les industriels proposent des objectifs pour capteurs numériques, les formats standards sont répertoriés en fonction de la diagonale du capteur. La diagonale du capteur est exprimée en pouces. Les formats numériques standards commencent au 1/4" soit une diagonale de capteur de 6,35 mm.

Cependant, certains constructeurs n'hésitent pas à fabriquer des capteurs au format 24×36 mm, le standard des pellicules argentiques. Il y a deux avantages à procéder ainsi :

1. Les optiques existantes n'ont pas besoin d'être modifiées.
2. La taille des pixels permet d'envisager l'intégration de traitements intelligents.

Ces capteurs d'images sont généralement utilisés dans les appareils photos numériques. Le constructeur Nikon propose le D3 au format FX (23,9×36 mm). C'est un appareil photo numérique, de résolution 4256×2832,

soit 12 mégapixels effectifs, avec des photosites de 8,5 μm de côté. 80 à 90% des besoins d'un photographe sont satisfaits avec une telle résolution, par exemple, sortir une double page magazine, autoriser un recadrage sans perte de qualité ou réaliser un tirage 60×80 cm. Avec des photosites de 8,5 μm de côté, le constructeur intègre des traitements qui améliorent considérablement la qualité de l'image et la dynamique du capteur. Cette nouvelle technologie rattrappe les performances des meilleurs appareils argentiques.

Du côté des caméras numériques à haute résolution, il existe deux formats, le DV (720x576) et le HDV (1280x720 et 1440x1080). Les optiques des caméras numériques ont leur propre standard (≥1/4"). Avec ce format numérique et les contraintes précitées, la taille du pixel ne doit donc pas dépasser 4,4 μm (6,35 mm/1440). Les progrès des procédés de fabrication CMOS permettent la réalisation de capteurs PPS avec une taille de pixel en deçà de 2 microns, comme le montre le graphe d'évolution de la taille du pixel (cf. figure 1.5), ce qui est par ailleurs aussi vérifié en technologie CCD.

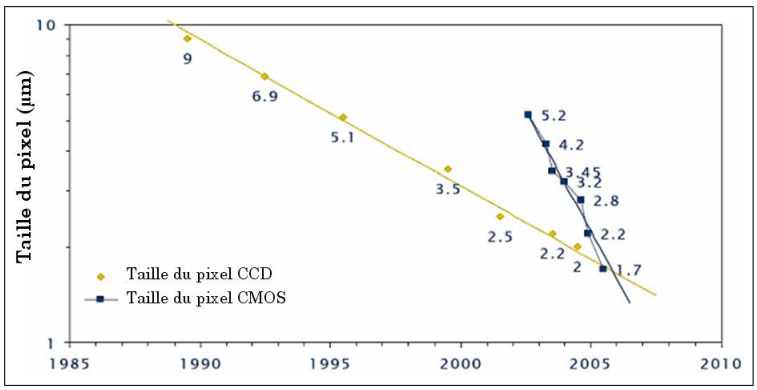


FIG. 1.5 – Taille du pixel dans les deux technologies CCD et CMOS (source : Micron)

Cependant ces pixels sont passifs car seule l'acquisition du signal optique est réalisée au sein du pixel. Avec une telle définition de l'image il est difficile de faire des traitements d'images temps réel compte tenu du flot importants d'informations en résolution HDV (1440×1080×8-bits×25 images par seconde). La question reste donc entière sur ce qu'il est possible d'intégrer comme traitements dans le pixel, tout en restant dans le gabarit industriel et si possible en utilisant la technologie standard CMOS pour des questions de coût.

1.3.3 La technologie CMOS s'essouffle

Les travaux de recherches actuels sur les imageurs utilisent la technologie CMOS 0,18μm [KLLEG01a, BGE03, Lim03, EGE05] en vue de futures applications industrielles. Il faut donc se projeter dans l'avenir sur une dizaine d'années et réfléchir sur l'interpolation possible des travaux de recherches vers des technologies plus fines (65, 45 ou 25 nm).

Nous savons que les industriels imposent des tailles de pixels n'excédant pas 4,4 microns. Or nous travaillons avec des technologies faiblement submicroniques pour limiter le coût de fabrication des prototypes, et bien que la loi de Moore[1] était une règle ces 30 dernières années, la technologie CMOS finira inéluctablement par atteindre ses limites physiques. En effet, plus la largeur de grille (L) d'un transistor[2] diminue (cf. figure 1.6), dans le procédé standard de fabrication, plus la surface de la capacité d'oxyde de grille, à même largeur de canal (W), diminue.

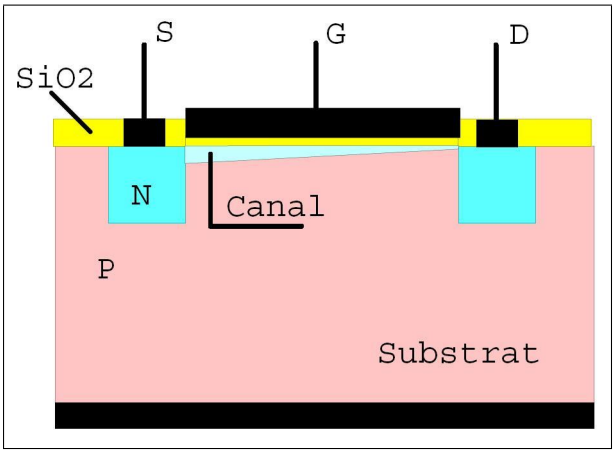


FIG. 1.6 – Coupe transversale d'un transistor NMOS en technologie standard CMOS

Cette capacité d'oxyde de grille, dont l'expression (celle d'un condensateur plan) est donnée par la relation 1.2, varie avec la technologie, sauf si l'épaisseur d'oxyde de grille (T_{ox}=7,6 nm, voir tableau A.1) est modifiée dans les même proportions. Par exemple, si la technologie passe d'une largeur de grille de 0,35 μm à une largeur de grille de 0,18 μm alors la surface de la capacité d'oxyde de grille est réduite dans les mêmes proportions. Dans cet exemple, si on ne fait rien au niveau de l'épaisseur d'oxyde de grille (T_{ox}),

[1]Selon cette loi, la densité d'intégration double tous les deux ans

[2]Par définition, la largeur de grille d'un transistor est la valeur de référence de la technologie

la capacité d'oxyde de grille est divisée par environ deux. Or il faut rendre invariante cette capacité d'oxyde de grille afin de maintenir les performances du transistor.

$$C_{gate} = \frac{\epsilon_{SiO2}\ \epsilon_0}{T_{ox}}\ W\ L \tag{1.2}$$

En effet, le couplage capacitif entre grille et canal contribue à la densité de charge d'inversion (un canal de porteurs existe entre drain et source, en régime d'inversion), et donc au courant de saturation. La silice, ou molécule de dioxyde de silicium SiO_2, est l'oxyde de grille naturel et de référence qui a rendu possible le succès fulgurant du silicium comme matériau de base de la microélectronique. Pour des épaisseurs supérieures à 20 Å, le courant qui traverse l'oxyde, sous l'action du champ vertical de grille (c'est un champ électrique créé par la différence de potentiel V_{GS} entre grille et source du transistor) et par effet tunnel, reste trop faible pour dégrader le courant de fuite global du transistor. L'effet tunnel est un phénomène aléatoire qui traduit le passage de charges majoritaires du canal vers la grille de polysilicium. Le courant de fuite global caractérise les fuites du transistor et donc sa consommation statique. Les premiers signes de faiblesse du SiO_2 se manifestent au dessous de 20 Å car le courant tunnel devient la principale composante du courant de fuite.

En technologie 0,35 μm, il n'y a pas encore d'effet tunnel, l'épaisseur d'oxyde est de 7,6 nm soit 76 Å. Les spécifications technologiques projetées par l'ITRS (International Technological Roadmap for Semiconductors) pour les transistor LSTP (Low Stand-by Power) sont illustrées dans le tableau 1.1 [Cha05]. Ce tableau illustre l'évolution de la largeur et de l'épaisseur de grille des transistors LSTP. Le LSTP est un transistor qui privilégie une faible consommation tout en maintenant un courant de saturation assez élevé et est utilisé pour les fonctions analogiques. On constate que loi de Moore n'est plus tout à fait vérifiée à partir de 2003 et que l'année 2016 marquera un tournant. En effet, en deçà de 10 Å, outre les problèmes de réalisation technologique, la fuite par la grille est de l'ordre de grandeur du courant de saturation. Après cette date la question reste entière quand au passage vers les nano-technologies.

Jusqu'en 2016, nous pouvons interpoler les travaux réalisés, dans des technologies sub-microniques, vers des technologies fortement sub-microniques. Mais il faut prendre quelques précautions car on ne peut pas directement interpoler linéairement les travaux sans modifier les comportements des fonctions électroniques analogiques intégrées. Pour ne prendre qu'un exemple, la plage de linéarité d'une paire différentielle n'est pas la même en technolo-

TAB. 1.1 – Évolution des paramètres physiques de la technologie CMOS (analogique) de 2003 à 2016 (source : ITRS)

Année de mise en production	2003	2004	2007	2010	2013	2016
Largeur de grille L (nm)	75	65	45	25	18	13
Épaisseur d'oxyde de grille (nm)	2,2	2,1	1,6	1,3	1,1	1
V_{DD} (V)	1,2	1,2	1,1	1	0,9	0,8

gie 0,35 μm qu'en technologie 45 nm (différences de courants de saturation, capacité de transconductance de grille, courant tunnel, etc.).

En conséquence, l'interpolation ne doit donc pas être strictement linéaire, c'est pourquoi nous proposons d'appliquer arbitrairement un facteur de sécurité de deux, comme on parle de marge de sécurité en automatique. Dans cet exemple, le rapport des largeurs de grille est d'environ 5, mais la taille du nouveau circuit équivalent, dans la technologie 65 nm, ne doit pas être divisée par 5 mais par 2,5. A l'heure actuelle, la technologie 45 nm (LSTP, analogique) est bien maîtrisée, au niveau de la production de masse. Dans le cas des capteurs d'images, la surface des pixels équivalents obtenue par le passage de la technologie 0,35 μm vers la technologie 45 nm, serait alors divisée par 16 en appliquant cette règle. Concrètement, un pixel analogique, dessiné en technologie 0,35 μm, de 35 μm de côté, aurait une taille de 8,75 μm de côté, en technologie 45 nm, avec des caractéristiques suffisamment proches du circuit initial. Ceci n'est qu'une estimation et non une affirmation, et ainsi nous proposons seulement d'appliquer une marge de sécurité afin de proposer un dimensionnement car à mesure que la largeur de grille des transistors baisse, il faut ajouter des transistors pour compenser les défauts de la nouvelle technologie.

On peut aussi mettre en avant ces défauts pour réaliser d'autres fonctions comme par exemple des fonctions non linéaires. Dans tous les cas, la faisabilité industrielle de rétines artificielles présentant des tailles de pixels raisonnables tout en y intégrant des traitements de bas nivaux visant à améliorer les caractéristiques techniques du capteur reste un défi scientifique et technique.

1.4 Pourquoi des capteurs "intelligents" ?

Le terme "intelligent" est ici exagéré car on est encore loin de prétendre concevoir des capteurs "intelligents". Devrait-on dire plutôt interactif ? En effet, certaines réalisations méritent ce titre au sens où le capteur est capable d'adaptation, aux conditions lumineuses par exemple [HA01]. C'est cette interaction avec les éléments et les caractéristiques physiques de la scène observée qu'il est intéressant de développer dans les capteurs d'images en technologie CMOS. On parle surtout de capteurs programmables et adaptatifs.

Les systèmes de vision classique (cf. figure 1.7), associent un capteur d'images, un convertisseur analogique numérique (placé à l'extérieur de la matrice de pixels), et une unité de traitement numérique (processeur arithmétique).

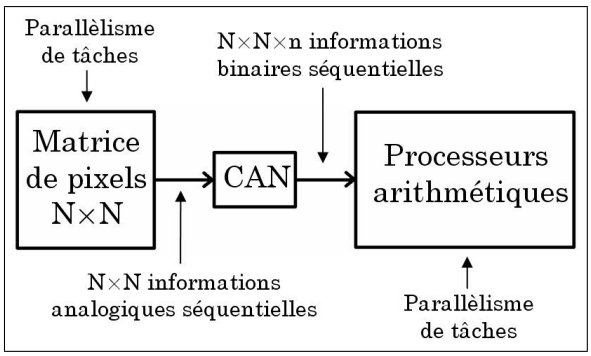


FIG. 1.7 – Synoptique d'un système de vision classique

Dans cette configuration, un goulot d'étranglement existe entre le capteur d'images et le processeur arithmétique, qui limite les performances des traitements temps réel. Par nature, un capteur d'images est une matrice d'éléments photo-sensibles, donc une architecture parallèle. Au lieu de profiter de ce parallélisme naturel, le système de vision classique sérialise les données analogiques de la matrice de pixels pour les injecter dans un CAN. Souvent les informations sont ensuite de nouveau parallélisées pour être traitées par plusieurs processeurs arithmétiques, il est alors dommage de ne pas profiter du parallélisme présent au niveau du capteur d'images.

Des capteurs intégrant des traitements sur plan focal, irréalisables avec la technologie CCD qui ne permet pas l'accès aléatoire aux pixels et l'intégration de fonctions électroniques au niveau du pixel, sont nommés "rétines artificielles". Ainsi, une rétine artificielle CMOS peut être vue comme une grille bidimensionnelle de "cellules" [TH01], chaque cellule étant formée de

l'association d'un élément photosensible (photorécepteur), et d'un élément de calcul (processeur élémentaire). Du fait de son caractère massivement parallèle, une rétine peut avoir des performances en vitesse, en intégration et en consommation meilleures que les solutions classiques (capteurs puis traitements logiciels et/ou matériels) [HH03]. Les approches conventionnelles pour la conception de machines de vision sont en général basées sur des architectures connectées à une caméra. L'approche proposée dans ce travail de thèse consiste à associer, dans un même circuit - une rétine CMOS -, les photo-capteurs et des fonctions de pré-traitements de l'image. C'est ce que l'on nomme communément l'approche ViSOC (Vision System On Chip ou système de vision sur puce) [PMB99]. Nous nous sommes donc intéressés à la faisabilité et à la conception en technologie CMOS d'un système de vision artificielle dédié à l'imagerie rapide et aux traitements d'images linéaires.

La plupart des travaux existant sur les capteurs d'images CMOS portent sur l'intégration monolithique d'une matrice photosensible et d'une unité de traitement, totalement ou partiellement dissociée de la matrice en question. Il existe des réalisations où les éléments de calculs sont intégrés au niveau du pixel ou d'un groupe de pixels ou encore d'une colonne de pixels du capteur [YPB03, ASIS04, KRB+05, SKH+05]. Dans les capteurs CMOS, le traitement au niveau du pixel est généralement écarté, car il conduit à une taille de pixels 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 des calculs massivement parallèles et ainsi d'exploiter les possibilités d'un capteur d'images haute cadence [KN03, CRVG+04, LMJM05, STT+05].

Par définition, une rétine artificielle est un capteur d'images capable de calculer et de fournir une information de plus haut niveau que l'éclairement ou la couleur (longueur d'onde) de chaque pixel. Les principaux avantages d'une rétine CMOS par rapport à un système de vision classique sont :

1. La vitesse des traitements : la vitesse des traitements des capteurs "intelligents" est plus grande que celle de la combinaison d'un capteur d'images avec le(s) processeur(s) externe(s) à la matrice de pixels. Dans un système de vision classique, le transfert d'information se produit par voie série entre le capteur d'images et les processeurs, alors que dans un capteur intelligent, les données entre les différents niveaux de traitements peuvent être transférées et traitées en parallèle.
2. Un simple circuit intégré ViSOC : un capteur "intelligent" monolithique comporte sur la même puce, la matrice de pixels photo-sensibles et les couches de bas et haut niveaux de traitement sur le même plan focal image. Par exemple, un circuit d'un centimètre carré de surface peut

remplacer l'équivalent d'un système complexe caméra-processeur.

3. Adaptation : dans beaucoup de capteurs "intelligents", la dynamique est grandement améliorée par une adaptation locale aux conditions lumineuses. La technologie APS (voir section 2.4.2) permet, en agissant sur le temps d'intégration, de réaliser un contrôle automatique et rapide du gain (sensibilité) et de l'offset (différence qui existe entre la réponse du pixel et une valeur de calibration), en fonction de l'éclairement du pixel. Dans les caméras conventionnelles le gain est compensé en bout de chaîne de l'information, son contrôle est donc plus lent.

4. Puissance consommée : dans les rétines, souvent l'information est raréfiée avant d'être déplacée ce qui contribue à une baisse de la consommation électrique. De plus, les rétines utilisent souvent des fonctions analogiques qui opèrent en régime de faible inversion ou d'inversion modérée. En outre, une grande partie de la puissance est dissipée dans les amplificateurs de sortie, les micro-fils de connexions entre le circuit intégré et le boîtier, et les inter-connexions hors boîtier. En plaçant le capteur d'images et les processeurs dans un même boîtier, les longueurs d'inter-connexions sont considérablement réduites, ce qui diminue d'autant la puissance dissipée dans ces amplificateurs et lignes de connexions.

5. Taille et coût : l'implémentation monolithique de processeurs et d'un capteur d'images réduit considérablement la taille du circuit intégré. Par conséquent, l'intégration d'un circuit de traitement sur le même plan focal que le capteur d'images réduit globalement le coût de fabrication comparativement à un traitement externe.

Bien que la conception d'un capteur "intelligent" monolithique soit une idée séduisante, elle fait face à plusieurs limites et inconvénients :

1. Fiabilité des traitements : les circuits de traitements des capteurs "intelligents" utilisent souvent des circuits analogiques non conventionnels qui ne sont pas toujours bien caractérisés. Par conséquent, les circuits de traitements ont parfois peu de précision qui est affectée par de nombreux facteurs incontrôlables. En conséquence, si pendant la conception du capteur, ces inexactitudes ne sont pas prises en compte, la fiabilité des traitements peut être dégradée.

2. Conception dédiée : des circuits analogiques non conventionnels sont souvent imaginés dans le domaine de la conception de capteurs "intelligents". Les circuits issus de la bibliothèque de composants sont rarement utilisés et donc de nouvelles fonctions sont élaborées spéci-

fiquement. Ceci augmente inévitablement les risques de mauvais fonctionnements du circuit ainsi conçu.

3. Programmabilité : certains capteurs "intelligents" ne sont pas reconfigurables pour effectuer différentes tâches de vision, il sont plutôt dédiés à une application spécifique. Plus généralement, la gamme des traitements est limitée mais le principal objectif de ces architectures est de déplacer l'information une fois devenue rare. Ce manque de programmabilité est parfois indésirable, toutefois la recherche avance dans ce domaine et propose de plus en plus de capteurs à grande reconfigurabilité.

Même avec de tels inconvénients, les rétines CMOS sont toujours attractives, principalement en raison de leur coût, de leur taille et de la vitesse de fonctionnement. C'est pourquoi le concept de rétines artificielles suscite de nombreux travaux de recherches à l'heure actuelle. Dans notre réalisation, nous privilégions la vitesse d'acquisition et la programmabilité au sein même du pixel du capteur.

1.4.1 Les rétines artificielles

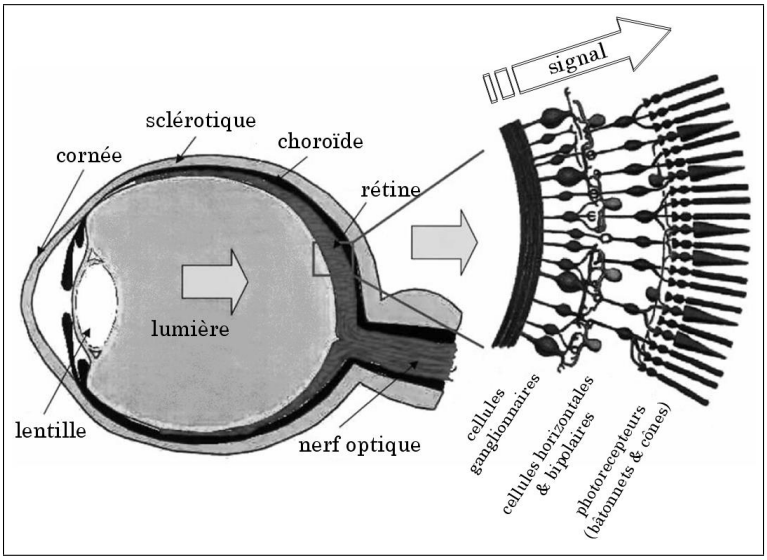


FIG. 1.8 – Coupe transversale de l'œil humain

Le concept de rétines artificielles est d'abord inspiré du modèle biologique, à la fin des années 1980, avec les travaux de Mead [MM88, Mea89]. Les traitements sont proches de ceux réalisés par les cônes, les bâtonnets, les

cellules horizontales et bipolaires. Les bâtonnets informent sur la luminance et les cônes sur la chrominance, ce qui constitue une première opération (couche d'entrée) au niveau du pixel. Il modèlise la connexion cône - cellule horizontale par une résistance et la connexion entre cellules horizontales par une autre résistance. Des amplificateurs sont utilisés pour réaliser des cellules bipolaires. Des condensateurs modélisent l'effet intégrateur de la membrane des cellules horizontales. La coupe transversale de l'œil humain (cf. figure 1.8) montre bien que la rétine contient des fonctions de prétraitement du signal photonique. Le modèle de Mead a permis la réalisation de différentes versions analogiques de rétines artificielles et des travaux récents continuent à être publiés dans ce sens [Mey02, DL04].

Cet axe de recherche est souvent privilégié car l'implémentation de prétraitements au niveau du pixel limite les transferts d'informations entre la rétine et le processeur externe, ce qui contribue à une augmentation de la vitesse globale du système d'acquisition et à une moindre consommation d'énergie.

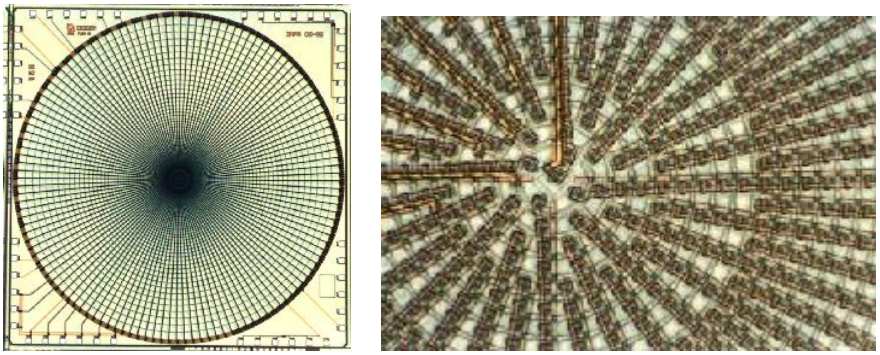
Insérer, au plus près de l'élément photosensible, une unité de calcul simple, mais capable de transformer et traiter le signal afin de raréfier les données avant d'être déplacées est une approche souvent privilégiée dans la conception des rétines artificielles.

1.4.2 Rétine avec traitements dédiés

Parmi les traitements dédiés, on distingue les traitements spatiaux et les traitements spatio-temporels. La première catégorie est bien représentée, citons par exemple la rétine de Standley [Sta91] en 1991 qui permet de déterminer la position du centre de gravité et l'orientation d'un objet très contrasté. La rétine extrait le moment géométrique d'une image à l'aide d'une grille résistive. Boahen et Andreou [AB95] en 1995, proposent une architecture de rétine analogique qui inclut un contrôle local du gain et un réhaussement de contraste sur l'image acquise. Le circuit de 48 000 pixels et 590 000 transistors a un surface de 9,5 mm $\times$ 9,3 mm et est réalisé dans une technologie standard CMOS 1,2 μm (double métal, double poly).

L'ambition de ces rétines est de réaliser des fonctions de traitements spatiaux comme les opérations de filtrage, parmi lesquelles l'extraction de contours [NZDA93, FNM+97, TH01], la détection d'objets et leur orientation [BBM+02] ou encore les opérations morphologiques spatiales. La position d'un objet peut être extraite d'une scène [NHIA00], mais la scène observée doit être relativement simple et très contrastée.

Les rétines à résolution fovéale s'inspirent de la biologie. La fovéa est la zone centrale de la macula, c'est la zone de la rétine où la vision des détails est la plus précise. La fovéa est peuplée quasi uniquement de cônes, les bâtonnets étant répartis sur la rétine périphérique. Sandini [SQS+00] propose en 2000 une rétine à résolution fovéale. Les 33 000 pixels du capteur sont distribués radialement (cf. figure 1.9) avec une décroissance logarithmique (résolution log-polaire) dans une technologie CMOS 0,35 μm. Le principal avantage de cette architecture est la faible consommation électrique et la vitesse de lecture réduites d'un facteur 35.



(a) (b)

FIG. 1.9 – (a) Dessin de masque d'une rétine à résolution fovéale , (b) agrandissement de la partie centrale

Park [PKK+03] propose en 2003 une application de ce concept à une rétine de 1024 pixels incluant chacun quatre transistors MOSFET et deux photo-transistors. Une adaptation à l'éclairement incident est réalisée au sein même du pixel. Le circuit est réalisé en technologie CMOS 0,6 μm (double-poly, triple métal) et mesure 8 mm×8 mm. Il peut détecter les contours d'une scène à grande dynamique de luminosité. Ce circuit pourrait trouver des application dans le domaine du suivi et de la reconnaissance d'objets.

D'autres rétines répondent aux besoins liés aux problématiques classiques de la vision artificielle, comme l'amélioration du contraste (égalisation d'histogramme), l'augmentation de la dynamique [Ni01], la détection de mouvement ou de déformation [Bel03] ou encore l'adaptation aux conditions lumineuses [Gou95, HA01, LMS01].

Toutes ces rétines ont leurs traitements figés, dédiés à une application particulière. Or souvent un système de vision artificielle se veut flexible, programmable ou reconfigurable.

1.4.3 Rétines programmables

Parmi les rétines à traitements programmables, Chen propose en 1990 la rétine PASIC [CADS90], un capteur d'images CMOS programmable qui inclut une matrice de 128×128 pixels photo-sensibles. Un CAN par colonne assure la conversion analogique numérique et sur le même substrat une matrice de 128×128 processeurs arithmétiques booléens permet de réaliser des traitements massivement parallèles sur les images acquises.

Dans certaines réalisations, les données déplacées se réduisent à une décision comme la rétine d'intercorrélation [ABLYV+04] d'Olivier Aubreton du LE2I. Une rétine dite "à masques" est utilisée pour réaliser une opération de reconnaissance de formes basée sur les moments de Zernike.

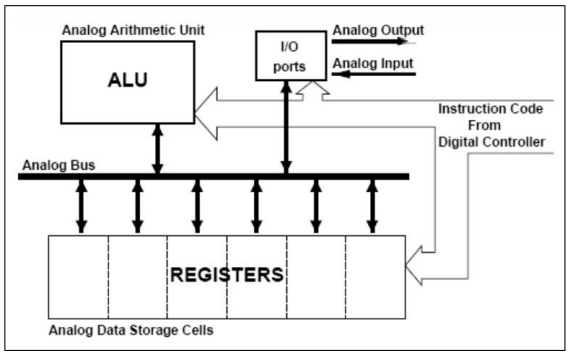


FIG. 1.10 – Synoptique d'un processeur analogique

Une architecture de processeur analogique est proposée par Dudek [DH00] avec son AμP, capable d'exécuter séquentiellement des échanges de données analogiques avec les plus proches voisins et des calculs sur ces données analogiques. La figure 1.10 illustre le synoptique de l'AμP. Cette architecture de processeur n'a pas été implémentée au niveau même du pixel, mais au niveau d'un groupe de pixels, car le nombre de transistors de la structure est relativement important (>150) et le contrôle numérique de la rétine s'avère complexe.

A l'Institut d'Électronique Fondamentale (IEF) le projet PARIS [KDMS99] se propose d'étudier en 1999 une nouvelle génération de rétines programmables en mêlant unités de calcul analogique et unités de calcul booléen. Cependant, la complexité du processeur ne permet pas d'envisager son intégration au sein du pixel, il est donc implémenté en bordure de la matrice sous la forme d'une ligne de processeurs SIMD. Les applications proposées sont l'exécution d'algorithmes de type CNN ou l'implantation de filtres anisotropes.

Intel Corporation propose le CPA [MA99] en 1999, avec un processeur 8 bits implémenté par blocs de 8 colonnes dans une matrice de pixels de type APS. Un convertisseur est implémenté au niveau de chaque colonne de la matrice de pixels. Cette architecture offre une grande programmabilité et permet de réaliser des traitements de bas niveau temps réel comme la détection de contours ou le suivi d'objets.

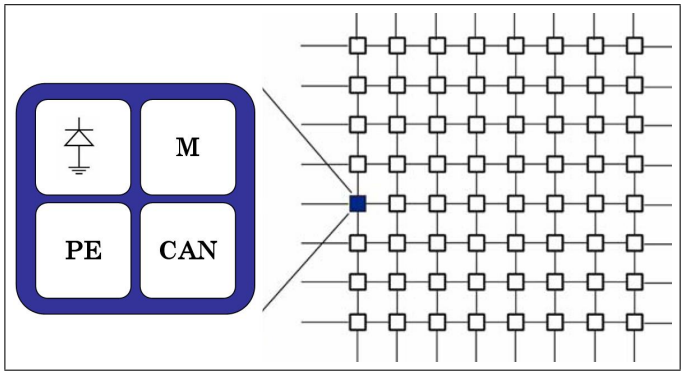


FIG. 1.11 – Synoptique d'une architecture distribuée dans un capteur CMOS

L'architecture de la rétine qui privilégie les traitements au voisinage du pixel est dite distribuée. Dans cette configuration (cf. figure 1.11), sont implémentés au sein même du pixel, le photo-détecteur, un processeur élémentaire (traitement analogique et/ou numérique), un convertisseur analogique numérique et de la mémoire CMOS.

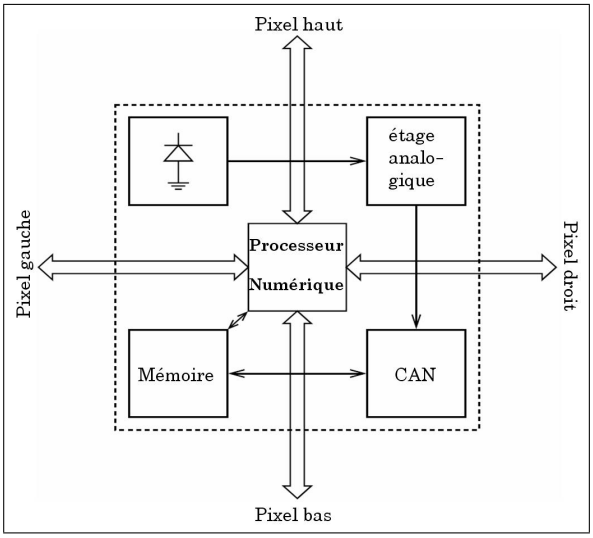


FIG. 1.12 – Pixel numérique programmable

Les capteurs d'images CMOS intelligents intègrent souvent des traitements programmables au plus près du pixel. Parmi ces rétines CMOS, on trouve la rétine d'El Gamal [EGYF99] en 1999, où le pixel contient un photodétecteur, un étage de mise en forme analogique, un convertisseur analogique numérique, de la mémoire et un processeur booléen programmable (cf. figure 1.12). L'intégration de ces fonctionnalités au sein de chaque pixel implique un nombre de transistors très conséquent (>200) ; ce qui limite fortement la résolution spatiale de ces architectures avec une technologie standard CMOS.

La rétine SCAMP-3 [Dud05, DC06] est une matrice de processeurs qui implémente une variété de traitements d'images bas niveaux à haute cadence (1000 im/s). Il existe une réalisation où des processeurs de traitements d'images "MIMD IP" (Multiple Instruction Multiple Data Image Processing) [ECKC01] permettent d'exécuter à haute cadence (9600 im/s) des convolutions locales utilisant des noyaux de dimensions variables (de 3×3 à 11×11).). La configuration, la dimension et les coefficients du noyau sont programmables. Le circuit ACE16k [RVEDC+96, CRVG+04, RVLCC+04] intègre un imageur et une matrice de processeurs SIMD (Single Instruction Multiple Data) configurés en réseaux de neurones cellulaires CNN (Cellular Neural Network) opérant sur les images acquises en niveaux de gris. La rétine VCS-IV [KIIY03, KKI04] est une matrice de processeurs SIMD programmables pouvant réaliser divers traitements d'images tels que la détection de contours, le lissage ou le filtrage par reconfigurations dynamiques. La rétine PVLSAR34, dernière née de la famille des rétines PVLSAR (Programmable and Versatile Large Size Artificial Retina) [Pai01], développée par Thierry Bernard à l'ENSTA en 2005, est un exemple de rétine numérique programmable. Le circuit se comporte comme une machine massivement parallèle SIMD à entrée optique. Chaque pixel contient un processeur booléen programmable, capable de communiquer avec ses plus proches voisins. Cette topologie est idéale pour réaliser des traitements locaux à haute cadence comme par exemple les gradients spatiaux. Le tableau 1.2 compare diverses réalisations de rétines en technologie CMOS et proposées par plusieurs équipes. Dans ce tableau, le taux de remplissage est défini par le rapport entre la surface photosensible du photo-détecteur et la surface du pixel.

Certaines solutions technologiques ont été industrialisées comme la rétine NSIP (Near Sensor Image Processing) [ESA95], les circuits LAPP (Linear Array Picture Processor) [RVB98] et la rétine Mapp (Matrix Array Picture Processor) [JISF05]. Ces rétines sont le résultat de recherches menées par l'équipe de l'Université de Linköping, en Suède.

La topologie 4-connexe, parmi les architectures distribuées, semble être

TAB. 1.2 – Comparaisons de capteurs d'images programmables

Capteur	SCAMP-3	MIMD IP
Auteur/année	Dudek (2005)	Cummings (2001)
Technologie	0,35 μm	1,2 μm
Résolution	128×128	80×78
Taille des pixels	49,35×49,35 μm^2	45,6×45 μm^2
Taux de remplissage	5,6%	33%
Transistors/PE	128 T	9 T
Type PE	analogique	analogique
fréquence trame	1000 im/s	9600 im/s
traitements d'image	bas-niveau	spatiaux
ACE16k	VCS-IV	PVLSAR34
Vazquez (2004)	Komuro (2004)	Bernard (2005)
0,35 μm	0,35 μm	0,8 μm
128×128	64×64	128×128
75,5×75,3 μm^2	67,4×67,4 μm^2	60×60 μm^2
6%	10%	30%
198 T	84 T	50 T
analog/num	numérique	numérique
1000 im/s	1000 im/s	1000 im/s
spatiaux	bas-niveau	bas-niveau

un bon compromis entre la complexité d'interconnexion (4 pixels connexes) et les possibilités de traitements au voisinage du pixel. Parmi les applications de ces rétines, on peut citer la détection de mouvements [Ric06], la morphologie mathématique [Man00] et aussi la détection de segments significatifs [BB06]. L'objectif est toujours est toujours le même : implémenter des pré-traitements au sein du pixel afin de raréfier les données avant de les transmettre à une unité de calcul de plus haut niveau.

1.5 Conclusion

Dans ce premier chapitre, nous avons mis en avant l'intérêt de faire des acquisitions à haute cadence. En effet, l'imagerie rapide trouve des applications dans l'analyse de mouvements rapides (vol d'insectes ou motricité d'un rongeur). De plus, être capable de faire des acquisitions pendant un temps bref ($\leq$1 ms) laisse du temps au système global de vision pour réaliser des traitements sur les images acquises.

Néanmoins, nous avons vu que les traitements implémentés sous forme numérique au sein du pixel demandent beaucoup de ressources matérielles (un grand nombre de transistors, souvent supérieur à 200), et que ce type d'architecture est difficile à contrôler (accès aux registres, exécution d'instructions). Ceci constitue les limites de ces architectures numériques. Ne pourrait-on pas se demander s'il ne serait pas préférable d'axer la recherche vers les rétines analogiques ? Certes, il est difficile de maîtriser les comportements des fonctions analogiques implémentées au sein des pixels (dérives en fonction de la température, dynamique, linéarité, etc.), mais outre ces difficultés, le gain en terme de densité d'intégration ainsi que les facilités liées au contrôle de la structure peuvent faire pencher la balance.

Un grand nombre de transistors implique une grande surface de pixel, or nous avons vu que la technologie CMOS s'essouffle et qu'elle cessera d'évoluer, telle qu'on l'a connaît, d'ici huit ans. La question du chercheur en systèmes de vision est : est-ce que mon pixel a un avenir industriel ? Car bien souvent les traitements réalisés au sein du pixel sont prometteurs mais la taille du pixel est souvent trop grande pour réaliser un capteur de résolution HDV dont la diagonale serait de l'ordre du centimètre (pixel de 4μm). Nous sommes donc confrontés à la fois à une contrainte technologique et à une contrainte industrielle sans parler de la contrainte économique.

Les travaux réalisés aujourd'hui en technologie standard CMOS 0,35 μm ou 0,18 μm peuvent être interpolés vers la technologie fortement submicronique de 2016, mais dans le cas particulier de processeur "analogique", l'interpolation ne peut être linéaire. D'où la marge de sécurité arbitraire que nous proposons dans cette étude. Nous n'affirmons pas que cette marge à elle seule suffira à garantir un dimensionnement projeté vers l'avenir, néanmoins cette approche permet de faire un première estimation de la surface de la rétine analogique dans ces technologies de demain.

Pour toutes ces raisons, nous choisissons de nous placer dans le cadre

des architectures distribuées en topologie 4-connexe avec un processeur élémentaire de type "analogique" incluant un nombre restreint de transistors. Compte tenu de la technologie utilisée (standard CMOS 0,35 μm) et des contraintes industrielles (pixel de taille inférieure à 4 μm), projetons nous par exemple en 2013. La technologie des transistors LSTP sera alors parfaitement maîtrisée au niveau de la production de masse avec une largeur de grille des transistors LSTP de 18 nm. Supposons que les contraintes d'encombrement reste les mêmes qu'en 2008 (capteur HDV de l'ordre du centimètre). La question est : dans quelle mesure les travaux réalisées dans des technologies standard CMOS supérieures à 0,1 μm peuvent être réutilisés dans une technologie plus moderne (LSTP 18 nm) ? Il n'y a pas de réponse précise à cette question et nous le savons. Nous devons étudier la nouvelle technologie, ses limites, ses nouvelles règles de dessin et essayer d'appliquer directement les travaux de conception pour en mesurer les défauts (linéarité des structures, influence du courant tunnel, etc.). Ceci n'est qu'un préalable car les résultats expérimentaux sont souvent très éloignés des résultats obtenus dans l'ancienne technologie. C'est pourquoi il est souvent nécessaire d'ajouter des étages ou de créer des structures télescopiques comme se fut le cas des amplificateurs opérationnels dans les années 1990 lors du passage de la technologie micronique à la technologie submicronique. Pour simplifier nous proposons d'appliquer une marge de sécurité sous la forme d'un facteur d'interpolation de deux. L'idée est de surdimensionner comme cela se fait en conversion d'énergie (alimentation électrique). Outre les problèmes de diaphonie optique[3], l'estimation de la taille du pixel est calculée sur la base d'une interpolation linéaire, de la technologie sub-micronique vers les technologies futures, affectée par exemple d'un coefficient 2. C'est donc pourquoi nous avons fixé la taille de notre pixel à 35 μm. De cette façon, la taille estimée du pixel en 2013 en production de masse sera de 3,6 μm soit une diagonale de capteur HDV (1440×1080) inférieure à 10 mm.

Dans la littérature, il existe différents types de technologies de capteurs d'images, c'est pourquoi il est important de faire état de quelques solutions qui ont suscité un certain intérêt dans ce domaine. Le chapitre qui suit se propose donc d'étudier la technologie des capteurs d'images CMOS et notamment les éléments de la chaîne de l'information (photodétecteur, technologie du pixel, etc.).

[3]Les rayons éloignés de l'axe optique, arrivant obliquement sur la surface photosensible, peuvent atteindre la zone sensible du pixel voisin provoquant sur ce pixel voisin une modulation parasite ou diaphonie par un rayon de lumière qui ne lui est pas destiné.

Chapitre 2

Technologie des capteurs CMOS

2.1 Introduction

L'Emploi de la technologie CMOS, dans le domaine des capteurs d'images, a débuté dans les années 1980 pour satisfaire aux besoins de la science, des technologies ou même de l'art. Dans ce chapitre, nous présentons les éléments théoriques, caractéristiques et matériels des capteurs d'images CMOS, souvent traités lorsque l'on se plonge dans la conception de tels dispositifs microélectroniques.

En particulier, sont passées en revue les caractéristiques de la photodiode, élément photosensible que nous utilisons dans notre capteur. Les différents types de capteurs (passif, actif et numérique) et certains convertisseurs analogiques numériques sont présentés également dans ce chapitre. Enfin les caractéristiques physiques du pixel à intégration (pixel actif) et les sources dominantes de bruits dans les capteurs d'images du même nom permettent de clore ce chapitre.

A titre d'information, le tableau A.1, en annexe A, regroupe toutes les grandeurs physiques utiles dans ce mémoire et leurs valeurs dans l'unité du système international. Les paramètres, liés au procédé de fabrication (standard CMOS 0,35 μm), sont extraits des documents constructeurs de la firme autrichienne AMS (AutriaMicroSystem).

2.2 Chaîne de l'information

La constitution des imageurs CMOS permet l'accès aléatoire aux pixels du capteur, a contrario des capteurs CCD qui par nature transfèrent sé-

quentiellement l'information vers la sortie du capteur. L'élément central de l'architecture matérielle d'un capteur CMOS est une matrice de pixels à entrées optiques, adressée suivant les lignes via un décodeur de lignes et suivant les colonnes via un multiplexeur de colonnes. Comme l'atteste la figure 2.1, on rencontre aussi dans beaucoup d'architectures, des amplificateurs de colonnes en amont du multiplexeur. Le dernier élément clé de la chaîne de l'information est le CAN ou Convertisseur Analogique Numérique.

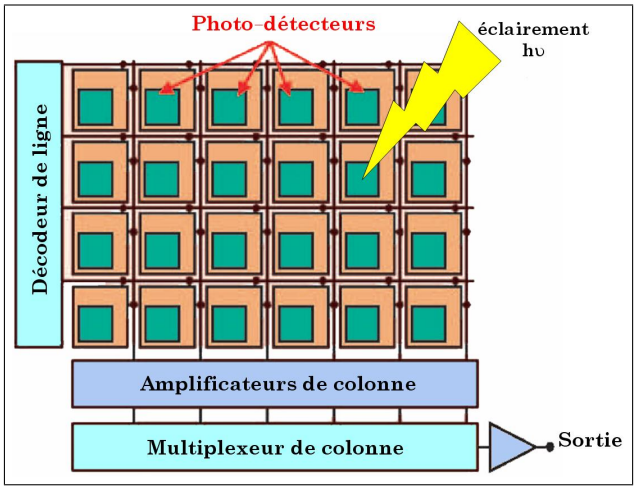


FIG. 2.1 – Architecture d'un imageur CMOS

Il existe, à ce propos, trois configurations possibles : (1) le CAN est placé en sortie du capteur, (2) le CAN est placé au niveau de la colonne, ou (3) le CAN est placé au niveau même du pixel. Dans ce dernier cas, la taille du convertisseur a une importance capitale, son architecture doit être la plus simple possible. La deuxième configuration semble être un bon compromis entre performances du convertisseur (vitesse de conversion, non linéarité intégrale, etc.) et densité d'intégration (surface occupée par le convertisseur).

La section qui suit se propose d'étudier un type particulier de photodétecteur, la photodiode, un composant semi-conducteur capable de transformer un rayonnement lumineux en un courant électrique et qui constitue l'élément photosensible de notre pixel.

2.3 Photodiode verticale de type N

La transduction éclairement courant ou éclairement tension utilise l'effet photoélectrique. Le silicium est le matériau de base de la microélectronique. Au niveau de la partie photosensible, le silicium a la propriété physique de

convertir les photons, incidents dans le plan focal, en électrons (ou photo-électrons). L'effet photoélectrique (cf. figure 2.2) se produit lorsqu'un photon est absorbé par un atome de silicium qui libère ainsi un de ses électrons de valence par une transition de bande (passage entre la bande de valence et la bande de conduction).

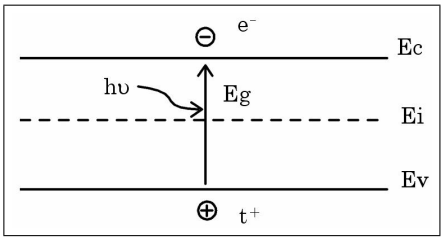


FIG. 2.2 – Effet photoélectrique dans un semi-conducteur

La transition de bande ne s'opère que si l'énergie du photon est supérieure à l'énergie de gap du silicium hν>Eg=Ec-Ev=1,12 eV [WS85]. Cette stricte inégalité entraîne l'existence d'une longueur d'onde maximale $\lambda_c = \frac{h \times c}{Eg} = 1,1\ \mu m$, qui correspond au proche infrarouge. Pour information, le germanium (Ge), qui a une énergie de gap de 0,7 eV [SHO04], est donc sensible jusqu'à $\lambda = 1,8\ \mu m$. Le silicium est parfaitement maîtrisé au niveau de la technologie de fabrication et ce matériau est en abondance dans la nature.

Il existe plusieurs modèles qui permettent de dimensionner une photodiode verticale de type N. Un premier modèle approximatif se réduit à une simple fraction rationnelle et un second plus complet se propose de déterminer la densité de courant en prenant en compte des phénomènes de diffusions.

2.3.1 Premier modèle

Dans une photodiode, afin de collecter les photo-électrons générés par effet photoélectrique, un champ électrique est appliqué à travers une jonction PN. La jonction PN polarisée en inverse collecte des photoélectrons lorsqu'elle est éclairée par un flux photonique. La technologie CMOS permet de concevoir des photodiodes verticales de type N, basées sur une structure N^+/P, dont la région P correspond au substrat et la région N^+, est fabriquée de la même manière qu'un drain ou une source de transistor NMOS. Le procédé standard permet aussi la réalisation de photodiodes verticales de type P, constituées d'un caisson N plus dopé que le substrat P, mais la présence de ce caisson N impose des distances minimales importantes. Dans cette technologie, la distance minimale entre le caisson N et une diffusion N^+ est de 2,6 μm.

La photodiode étant polarisée en inverse, le champ électrique résultant $\vec{E} = -\overrightarrow{grad}\ V$ entraîne l'apparition d'une zone neutre où il n'y a plus de porteurs pour assurer la conduction électrique. C'est une zone dite dépeuplée ou zone de déplétion. Dans cette Zone de Charge d'Espace (ZCE), on peut considérer que tous les porteurs libres créés sont séparés sauf ceux qui se recombinent et contribuent alors au bruit de génération/recombinaison. Les porteurs de charge libres ainsi créés contribuent au courant principal, le courant de transit (I_{tr}).

Dans les zones N^+ et P, la création d'électrons et de trous déséquilibre la relation entre porteurs minoritaires et majoritaires, et de ce fait les porteurs minoritaires de chaque région diffusent vers le bas. On a donc une composante du courant photonique qui est due aux trous dans la zone N (I_{pdiff}), et aux électrons dans la zone P (I_{ndiff}). Cette diffusion est fortement dépendante de la durée de vie τ_n et τ_p des porteurs minoritaires de chaque région [Sol96]. Le courant total est la somme des ces trois courants (I_{ph}=I_{tr}+I_{pdiff}+I_{ndiff}). En première approximation, les courants de diffusion peuvent être négligés et seul le courant de transit peut être considéré.

Comme le montre la figure 2.3(a), la zone de déplétion se situe presque exclusivement dans la région P (sur environ 900 nm), en raison du fort dopage de la région N^+ ($212.10^{15}\ cm^{-3}$ en technologie CMOS 0,35 μm d'après les données du constructeurs AMS). Donc, pour arriver dans la zone photosensible, les photons doivent d'abord traverser la région quasi neutre (N^+) qui influe sur la réponse spectrale de la photodiode. Dans les photodiodes verticales de type P, le caisson N est plus dopé que le substrat P, donc la ZCE s'étend dans de moindres proportions, c'est pourquoi cette solution technologique est souvent écartée.

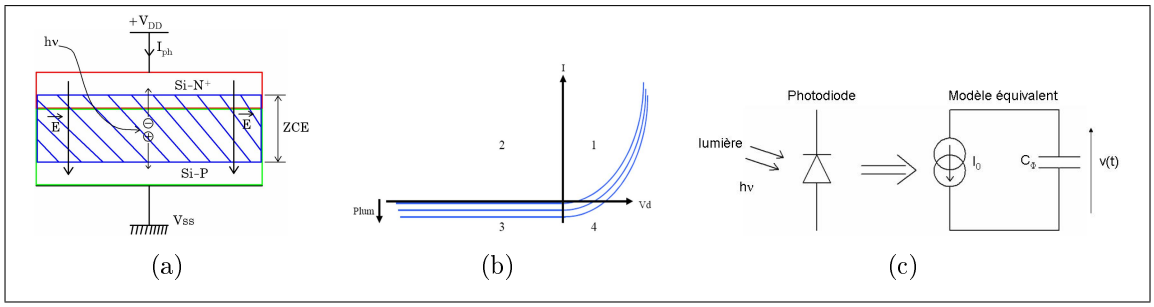


FIG. 2.3 – Photodiode verticale de type N, (a) phénomène contribuant au courant photonique , (b) caractéristique I=f(V_d), (c) schéma équivalent dans le 3eme quadrant.

Le courant photonique I_{ph} est proportionnel au produit du flux (i.e. le

nombre de photons incidents par unité de temps) par le rendement quantique $\eta \approx 70\%$, et il circule dans le sens inverse de la diode passante (cf. figure 2.3(a)), *i.e.*, de la région N^+ vers la région P.

La polarisation de la photodiode dans le troisième quadrant (cf. figure 2.3(b)) est intéressante pour l'allure constante du courant. Dans ces conditions, la photodiode est assimilable à un générateur de courant I_{ph} en parallèle avec la capacité de jonction C_ϕ, entre la diffusion et le substrat (cf. figure 2.3(c)). Cette capacité de dépeuplement s'apparente à un condensateur plan, mais son expression (cf. équation 2.1) formelle est très différente, car l'épaisseur de la ZCE dépend de constantes intrinsèques et de la tension de polarisation inverse V_d. L'épaisseur de la ZCE dans la diffusion N^+ est négligée, l'expression dépend donc de la concentration N_A d'atomes accepteurs dans le substrat P, de la tension de seuil V_{seuil} de la diode, de la charge élémentaire q, de la constante de Boltzmann k et de la surface active de la photodiode A_{ph}.

$$C_\phi = \sqrt{\frac{q\ \epsilon_{SiO2}\ \epsilon_0\ N_A}{2\ \left(V_{seuil} + V_d - \frac{2\ k\ T}{q}\right)}}\ A_{ph} \tag{2.1}$$

Où la tension de seuil a pour expression :

$$V_{seuil} = \frac{k\ T}{q}\ \ln(\frac{N_A\ N_D}{N_i^2}) \tag{2.2}$$

Ce modèle ne tient pas compte du périmètre de la photodiode. Pour compléter ce modèle, le fondeur donne les capacités surfaciques C_s et périmétriques C_p pour une tension de polarisation V_d nulle et propose l'expression de la capacité $C_{\phi 0}$. En considérant A_{ph}, l'aire de la diffusion et P_{ph}, son périmètre, la capacité de jonction est donnée par la relation (2.3).

$$C_{\phi 0} = A_{ph}\ C_s + P_{ph}\ C_p \tag{2.3}$$

Dans une photodiode de type P, la ZCE est moins épaisse que dans une photodiode verticale de type N, donc sa capacité surfacique est d'autant plus grande. Or cette capacité est une capacité parasite car elle contribue en grande partie au bruit de lecture dans les pixels de type APS (technologie particulière de pixel).

Le photocourant I_{ph} dépend exclusivement de la longueur d'onde λ et de la puissance optique P_{opt} reçue au niveau du plan focal image. En considérant h la constante de Planck, η le rendement quantique, c la célérité de la lumière

et q la charge élémentaire, l'expression du courant photonique est donnée, en première aproximation, par la relation (2.4).

$$I_{ph} = \eta \frac{\lambda\ q}{h\ c}\ P_{opt} \tag{2.4}$$

Enfin, la puissance optique reçue est le produit de l'éclairement E (en $W.m^{-2}.(100\,lux)^{-1}$) incident au plan focal image, par la surface active S_{ph} (en m^{-2}) de la photodiode, d'où l'expression du courant photonique :

$$I_{ph} = \eta \frac{\lambda\ q}{h\ c}\ E\ S_{ph} \tag{2.5}$$

Ce modèle approximatif permet de fixer les ordres de grandeurs, préalables à l'application d'un modèle plus complet qui considère les composantes de diffusions du photo-courant.

2.3.2 Second modèle

Un modèle plus complet se propose d'exprimer la densité de courant au travers de la jonction d'une photodiode verticale de type N. La densité de courant possède alors deux composantes (transit et diffusion). La figure 2.4 illustre les densités de courant qui existent dans une photodiode verticale de type N.

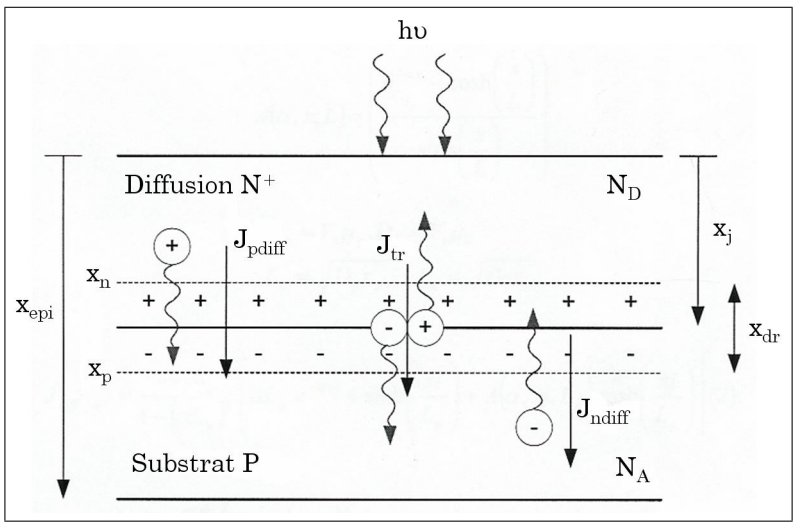


FIG. 2.4 – Densités de courant dans une photodiode N verticale

Le courant de transit, dans la zone de déplétion, a pour expression :

$$J_{tr} = -\alpha\ P_{opt}\ \frac{\lambda\ q}{h\ c} \int_{x_j - x_n}^{x_j + x_p} e^{-\alpha\ x} dx \tag{2.6}$$

Où $\alpha = \left(\frac{67{,}2916}{\lambda(en\ nm)}\right)^2 .m^{-1}$. Après intégration de l'expression 2.6, il vient :

$$J_{tr} = P_{opt}\ \frac{\lambda\ q}{h\ c}\ e^{-\lambda\ x_j}(1 - e^{-\lambda\ x_{dr}}) \tag{2.7}$$

Où $x_j \approx 200\,nm$, en technologie CMOS AMS 0,35 μm, est la profondeur de diffusion N_{diff} et x_{dr} est l'épaisseur de la zone de charge d'espace qui est donnée par la relation 2.8.

$$x_{dr} = \sqrt{\frac{2\ \epsilon_{SiO2}\ \epsilon_0 (V_{seuil} + V_d)}{q} \left(\frac{1}{N_A} + \frac{1}{N_D}\right)} \tag{2.8}$$

L'épaisseur x_{dr} de la zone de charge d'espace est fonction de la concentration de dopage et de la tension de polarisation inverse appliquée a la photodiode. Cette propriété est utilisée dans notre étude de dimensionnement de la photodiode. Les densités de courants de diffusion en dehors de la zone de charge d'espace ont pour expressions :

$$J_{ndiff} = \frac{q\ \phi_0\ L_p}{1-(\alpha\ L_p)^2} \left[\alpha\ L_p\ e^{-\alpha\ x_j} + \sinh\left(\frac{x_j}{L_p}\right) + A(\alpha, x_j, L_p)\cosh\left(\frac{x_j}{L_p}\right)\right]$$

$$J_{pdiff} = \frac{q\ \phi_0\ L_n}{(\alpha\ L_n)^2 - 1}\ e^{-\alpha(x_j + x_{dr})}\ [A(\alpha, x_{epi} - (x_{dr} + x_j), L_n) + \alpha\ L_n]$$

Avec, $\phi_0 = \alpha\ P_{opt}\ \frac{\lambda}{h\ c}$, $D_p = \frac{k\ T\ \mu_p}{q}$, $D_n = \frac{k\ T\ \mu_n}{q}$

les longueurs de diffusion : $L_p = \sqrt{D_p\ \tau_p}$, $L_n = \sqrt{D_n\ \tau_n}$

et, $A(\alpha, x, L) = \left(\frac{e^{-\alpha\ x} - \cosh\left(\frac{x}{L}\right)}{\sinh\left(\frac{x}{L}\right)}\right)$

(2.9)

La densité de courant de diffusion des trous en dehors de la ZCE dépend de l'épaisseur de la couche épitaxiale du substrat faiblement dopée en trous ($x_{epi} \approx 20\,\mu m$ en technologie CMOS AMS 0,35 μm). La densité de courant totale, au travers de la jonction PN, est la somme des trois contributions. Le courant total, qui traverse la photodiode, est obtenu en intégrant les densités de courant sur la surface active de la photodiode.

$$I_{ph} = \int\int_{S_{active}} (J_{tr} + J_{ndiff} + J_{pdiff})\ dS \tag{2.10}$$

Ce modèle, plus précis que le premier, permet de prendre en compte les courants de diffusion qui sont générés en dehors de la zone de charge d'espace et qui diffusent à travers la jonction [YPEC04]. Néanmoins, le courant de transit est la composante principale du courant photonique, ce qui justifie le premier modèle. Dans le quatrième chapitre, nous verrons que le potentiel électrostatique dans la couche épitaxiale a une distribution poissonienne et comment une analyse numérique peut compléter encore davantage ce modèle.

2.4 Pixels analogiques

Il existe principalement deux façons de réaliser la commande électronique de la transduction éclairement - tension au niveau du pixel.

1. La lecture de l'information est réalisée de manière passive (pixel PPS), en lisant simplement une valeur de courant (photocourant), via un transistor de sélection.
2. La lecture est active (pixel APS), alors le photocourant est d'abord intégré pour traduire une différence de potentiel proportionnelle à l'éclairement.

Dans ce même ordre, cette partie décrit les deux technologies en insistant particulièrement sur la technologie APS dans la mesure où nous nous en inspirons dans notre réalisation de rétine.

2.4.1 Pixel passif PPS

Le pixel analogique passif ou PPS (Passive Pixel Sensor) [Den95, Fuj00], présente l'architecture la plus simple, dans l'ensemble des imageurs CMOS. Comme le montre la figure 2.5, ce pixel inclut une photodiode et un transistor de sélection [EGE05]. Un convertisseur courant-tension est placé au bas de chaque colonne de l'imageur. Pendant l'intégration, le photocourant est régi par l'équation 2.11.

$$i(t) = C\frac{d(v_{ref} - v_{sortie})(t)}{dt} = -C\frac{dv_{sortie}(t)}{dt} \tag{2.11}$$

A la fin de l'intégration, le potentiel V_{sortie} est une image de l'éclairement du pixel. Le condensateur de l'intégrateur est ensuite court-circuité via le

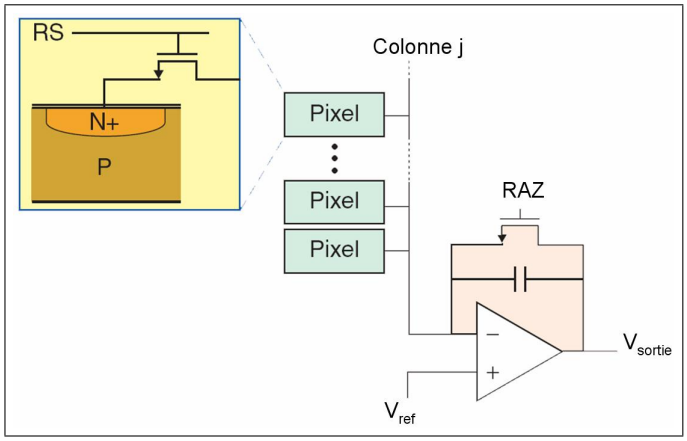


FIG. 2.5 – Architecture du pixel passif (PPS)

transistor RAZ, pour en dissiper la charge, avant la lecture de la ligne suivante. Cette architecture de pixel présente l'avantage de ne comporter qu'un seul transistor, ce qui permet de réaliser des pixels de très petite taille avec un taux de remplissage[1] proche de 100%.

Cependant, cette structure de pixel est assez susceptible aux perturbations électromagnétiques conduites internes à l'imageur. Le bruit de colonne, induit par les non uniformités des convertisseurs courant-tension, s'avère non négligeables, ce qui limite fortement leur utilisation. Les pixels APS résolvent ce problème, mais au prix de transistors supplémentaires.

2.4.2 Pixel actif APS

Les architectures APS (Active Pixel Sensor) ont un intérêt particulier dans la recherche et le développement de capteurs d'images et nous nous en inspirons dans la conception du premier étage de notre rétine. Le pixel analogique actif ou 3-T APS (à 3 Transistors) [Won96, Cho00, EGE05] inclut un transistor de préchargement, un étage suiveur (drain commun) et un transistor de sélection piloté en ligne (cf. figure 2.6).

Le transistor de préchargement (Reset) polarise en inverse la photodiode. Il se forme alors une zone de charge d'espace où les photo-porteurs donnent naissance au photo-courant. Le principe de fonctionnement d'une photodiode, en mode tension, est basé sur le phénomène de charge et de décharge de la capacité équivalente de la photodiode. En fonctionnement normal, la diode

[1]Le taux (ou facteur) de remplissage est défini comme étant le rapport entre la surface active du photo-détecteur et celle du pixel.

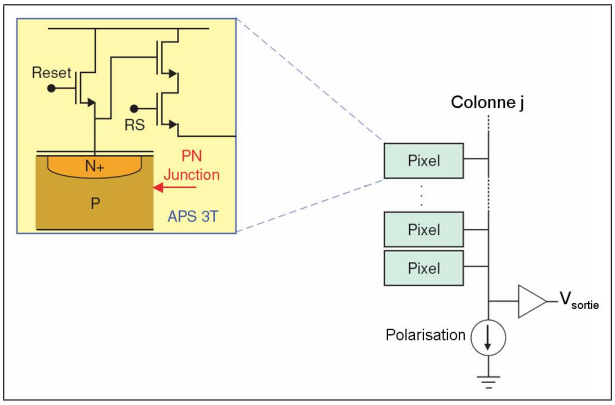


FIG. 2.6 – Architecture du pixel actif (APS)

est préchargée à la tension connue V_{DD} (cf. figure 2.3(a)), le potentiel d'alimentation le plus élevé.

Lorsque la phase d'initialisation est terminée (quelques centaines de ns), la phase d'intégration est amorcée, la décharge de la photodiode se fait proportionnellement à l'intégrale des variations du photocourant. La capacité se décharge à courant constant si l'éclairement au niveau du pixel est constant. En faisant quelques approximations sur la non-linéarité de la capacité, la décharge est assimilable à une droite dont la pente est proportionnelle à la puissance lumineuse (cf. figure 2.7).

En première approximation, la photodiode se décharge à courant constant. Son évolution est régie par la relation (2.12). On note C_{int}, la capacité effective d'intégration.

$$I_{ph} = C_{int} \frac{\partial v}{\partial t} \tag{2.12}$$

On peut réaliser une lecture de la tension au bornes de la photodiode à intervalles réguliers ΔT. ΔT est alors le temps d'intégration, c'est un paramètre de réglage, il dépend principalement des conditions lumineuses. Si l'éclairement est insuffisant, la période d'intégration peut être augmentée, à contrario, diminuée. Si la cadence de l'imageur doit être augmentée, le temps d'intégration peut être au contraire diminué mais il faut augmenter, en contre partie, la luminance de la scène. La différence de tension ΔV est donnée par la relation (2.13) :

$$\Delta V \approx \frac{I_{ph}}{C_{int}} \Delta T \tag{2.13}$$

Pour isoler le nœud flottant cathodique du circuit de lecture, un étage

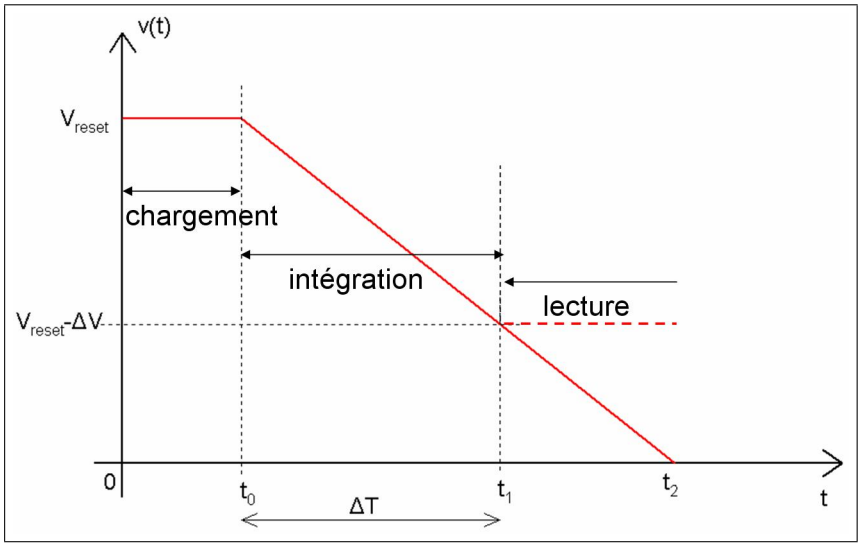


FIG. 2.7 – Phase de préchargement puis intégration du photocourant

suiveur est interposé. C'est un simple transistor NMOS, monté en drain commun, qui a la particularité de présenter une forte impédance d'entrée, une impédance de sortie très faible et un gain voisin de l'unité. Le circuit de lecture est un simple interrupteur, un transistor NMOS de sélection.

Les avantages de cette structure par rapport aux pixels PPS sont :

1. Une sucsceptibilité moindre aux perturbations électromagnétiques conduites internes à l'imageur.
2. Une variance de bruit de colonne moindre également.
3. La possibilité d'agir sur le temps d'intégration permet de réaliser des capteurs "intelligents" capables de s'adapter aux conditions lumineuses [HST+00, LEG01a, HA01].

L'inconvénient de cette structure est l'importance du bruit de lecture, dû au pré-chargement de la capacité parasite de la photodiode. Pour palier cette difficulté, il faut diminuer la capacité parasite ou compenser le bruit par correction. Pour diminuer la capacité parasite, on peut soit changer de technologie, soit changer la géométrie dans le procédé standard de fabrication. La première solution est onéreuse, la seconde est un axe de recherche qui à été privilégié dans ce travail de thèse. Dans notre capteur, nous empruntons à l'architecture APS, l'idée du mode de lecture actif, à savoir l'intégration du photocourant d'une photodiode préchargée à la tension d'alimentation du circuit.

Une autre technologie de pixel permet de palier l'inconvénient du bruit de lecture, c'est le pixel numérique, souvent privilégié dans la conception de

capteurs d'images rapides.

2.5 Pixel numérique DPS

Le DPS (Digital Pixel Sensor) a fait coulé beaucoup d'encre ces dernières année avec une architecture de pixel numérique permettant la réalisation de capteurs d'images rapide et de grande taille.

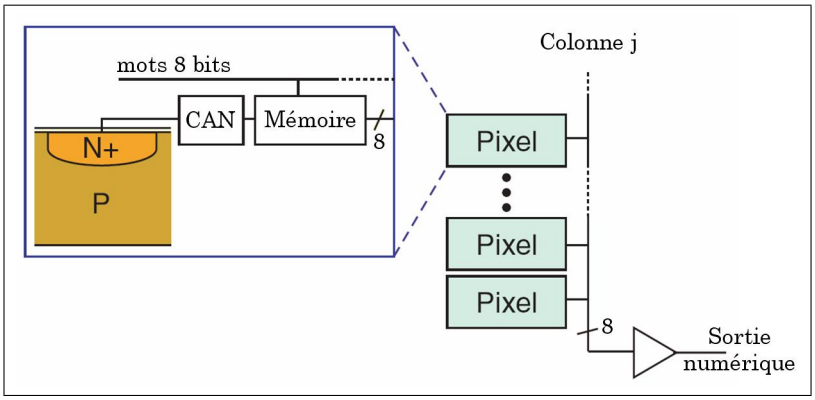


FIG. 2.8 – Architecture du pixel numérique (DPS)

Dans cette architecture, l'information lue au niveau du pixel est numérique, codée généralement sur 8 bits. Le principal avantage de cette architecture par rapport aux pixels analogiques est qu'elle intègre une électronique analogique simple [Bel03]. Le DPS intègre une conversion analogique numérique (cf. figure 2.8). Il offre plusieurs avantages par rapport aux pixels analogiques. Il élimine le bruit spatial fixe de colonne et le bruit de lecture. Intégrer un CAN et huit points mémoires, au niveau même du pixel, permet d'obtenir à l'échelle du capteur, une conversion analogique-numérique massivement parallèle. Cette architecture permet une lecture numérique de l'image et cela à haute cadence. La principale difficulté est de concevoir un CAN à haute densité d'intégration. Cette méthode requiert un nombre de transistors forcément conséquent. Une équipe de l'Université de Stanford travaille depuis 1992 sur la mise au point de capteurs CMOS numériques dédiés à l'imagerie rapide. Trois prototypes de DPS ont été développés avec des architectures de convertisseurs analogiques-numériques différentes.

Le première génération de circuits [FEGY94, Fow95, FYG95, YFEG96] est une matrice de 128×128 pixels avec un CAN sigma-delta du premier ordre, implémenté par groupe de quatre pixels connexes, en technologie standard CMOS 0,8 μm. Ce type de convertisseur [KHE+86] est basé sur le principe

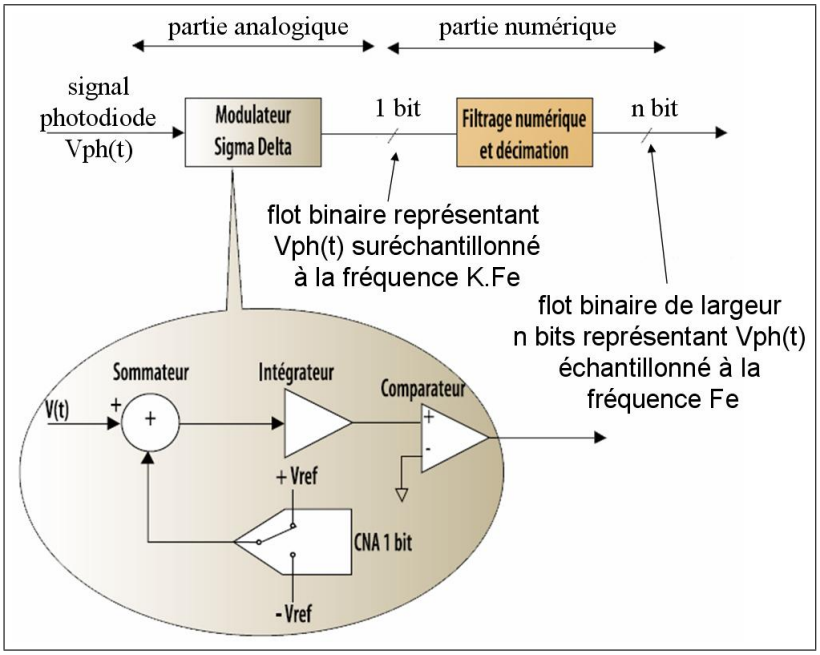


FIG. 2.9 – Synoptique d'un convertisseur sigma-delta du premier ordre

du sur-échantillonnage d'un signal d'entrée. Comme le montre la figure 2.9, la partie analogique d'un convertisseur $\Sigma\Delta$ constitue le modulateur. Un modulateur de premier ordre comprend un sommateur, un intégrateur, un comparateur (CAN 1 bit), et un CNA 1 bit. Le principe d'un modulateur Sigma-Delta du premier ordre est simple. Un comparateur est en général utilisé pour convertir sur un bit la différence (delta) entre le signal d'entrée et le résultat de la conversion (0=plus petit, 1=plus grand). Si par exemple le résultat le comparaison est 0, le CNA a un niveau "-Vref", le signal en sortie de celui-ci tendra à faire baisser la tension en sortie de l'intégrateur, jusqu'au prochain basculement du comparateur. Donc plus la tension Vph est élevée (proche de "+Vref"), plus le nombre de 1 sera important en sortie du modulateur, et réciproquement. La valeur moyenne contenue dans le flux binaire ainsi généré est donc représentative de la valeur de la tension Vph d'entrée. Ce système asservi (où la sortie est rebouclée sur l'entrée) fait osciller la valeur de l'intégrale du signal à convertir autour d'une valeur de référence (le résultat de la conversion). La sortie numérique du comparateur est sur un bit à haut débit (la fréquence de suréchantillonnage K.Fe). Elle est filtrée par le décimateur ce qui augmente la résolution tout en réduisant la fréquence de suréchantillonnage à la fréquence Fe. Le filtre numérique supprime les composantes fréquentielles les plus élevées, il n'y a aucune raison de conserver une fréquence d'échantillonnage aussi élevée en sortie qu'en entrée. De ce fait, cette

fréquence d'échantillonnage est réduite d'un certain facteur, appelé "facteur de décimation" K. Le principal problème est que le décimateur doit être implémenté en dehors de la matrice de pixels ce qui implique souvent un débit important dans la boucle de retour de l'asservissement.

La seconde génération [FEGY98, YFEG98, YEGFT99a, YFEG99] de DPS résout ce problème en implémentant, par groupe de quatre pixels connexes, un convertisseur 8 bits qui utilise une tout autre technique. Un capteur de résolution 742×554 a été industrialisé sur ce principe, en technologie 0,18 μm [BGE03].

Un convertisseur 8 bits associe la technique de conversion simple rampe [YEGFT99b, WCD+99] avec un codage grey binaire pour simplifier le contrôle numérique du capteur. La figure 2.10 illustre le principe du convertisseur à partir du bit de poids faible. Avec cette méthode, il faut donc 256 cycles d'horloge pour convertir le signal analogique sur 8 bits.

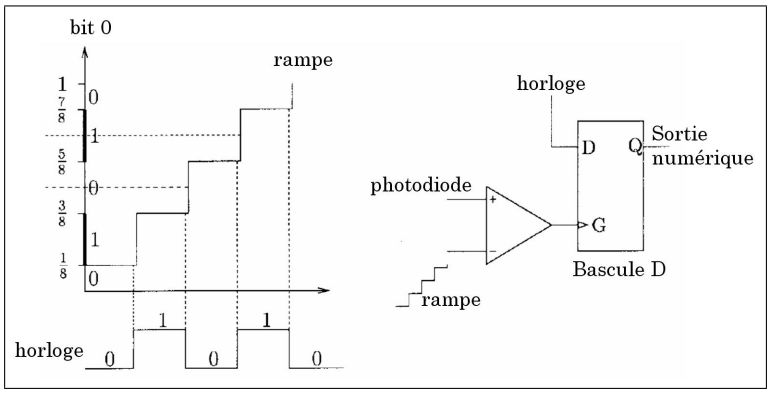


FIG. 2.10 – Graphe du bit de poids faible (bit 0) et schéma structurel

Récemment, dans la même équipe de Stanford, le capteur CMOS à DPS de Stuart Kleinfelder [KLLEG01a, KLLEG01b, Lim03] comporte 37 transistors par pixel (cf. figure 2.11). Cette architecture de capteur est l'une des plus performantes [EGE05]. La mémoire intégrée dans le pixel est composée de huit points mémoires dynamiques (DRAM pour Dynamic Random Access Memory) à trois transistors chacun, soit 24 transistors(cf. figure 2.11). Le circuit est une matrice de 288×352 pixels (format QCIF). Pour la première fois, l'acquisition est continue à la cadence de 10 000 images par seconde soit un débit de un gigapixels par seconde.

La technologie, utilisée ici, est le procédé standard CMOS 0,18 μm. La taille du pixel est de 9,4 μm de côté avec un taux de remplissage de 15%. Le tableau 2.1 regroupe les caractéristiques du capteur. La photodiode ne fait

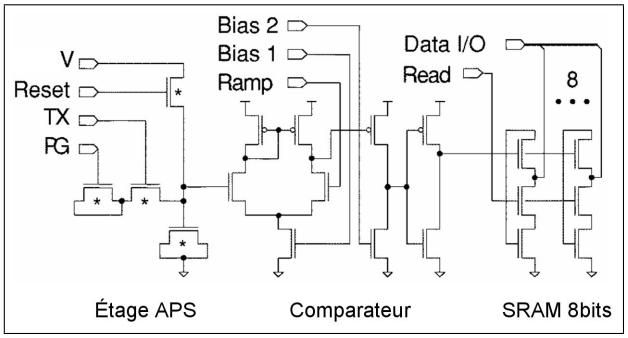


FIG. 2.11 – Schéma à transistors du pixel numérique DPS de Kleinfelder

TAB. 2.1 – Caractéristiques du capteur à DPS de Kleinfelder

Technologie	CMOS 0.18μm
Résolution	352 × 288
Surface	$\approx 25\ mm^2$
Nombre de transistors	3 751 000
Nombre de transistors / pixel	37
Tailles des pixels	9,4 μm × 9,4 μm
Taux de remplissage	15 %
Erreur différentielle du CAN	0,56 LSB
Consommation	50 mW à 10 000 im/s
Tension d'alimentation	1,8 V

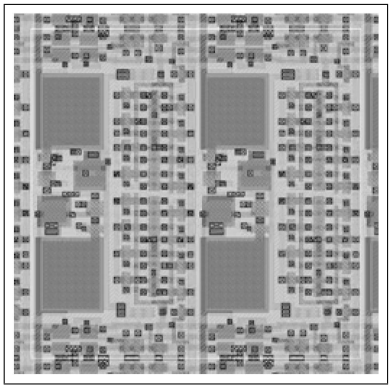


FIG. 2.12 – Dessin de masque de 4 pixels connexes (technologie 0,18 μm)

que 3,6×3,6 μm^2 de surface. Le dessin de masque[2] est illustré figure 2.12.

La technologie DPS constitue une avancée significative dans le domaine des capteurs d'image CMOS rapides. C'est une solution certes séduisante et efficace mais il est possible de trouver d'autres solutions tout aussi efficaces en évitant l'implantation d'un convertisseur au sein même du pixel. En effet, il est envisageable d'implanter des fonctions analogiques, moins gourmandes en transistors, en vue de réaliser des fonctions de pré-traitements de l'image au plus près du pixel et d'effectuer la conversion analogique numérique à posteriori. Avant de détailler cette solution, une réflexion doit être menée sur les convertisseurs analogiques numériques dans les capteurs CMOS.

2.6 Les convertisseurs analogiques numériques dans les capteurs CMOS

2.6.1 Introduction

Dans notre capteur, nous avons implémenté un convertisseur analogique numérique au niveau de la colonne du capteur mais il existe d'autres façons de procéder. Comme évoqué dans la section 2.2, il existe trois façons d'implémenter le convertisseur analogique numérique dans un capteur CMOS. En fonction de la configuration choisie, un compromis doit être fait entre la taille du convertisseur et ses performances. Il y a globalement trois niveaux dans la hiérarchie architecturale d'un capteur d'images CMOS (cf. figure 2.13).

L'élément de base du premier niveau est la plus petite partie de l'imageur, son quantum, le pixel. Les colonnes constituent le second niveau et la matrice de pixels, le troisième. Le niveau hiérarchique est proportionnel à la complexité et à la rapidité du convertisseur analogique numérique :

1. Le CAN est placé en sortie du capteur, dans ce cas peu importe la taille du convertisseur, mais il doit être très rapide relativement à la cadence de l'imageur (de l'ordre du milliard d'échantillons par seconde sur 8, 10 ou 14 bits).
2. Le CAN est placé au niveau de la colonne, dans ce cas la taille du convertisseur ne doit pas dépasser le pas spatial du capteur, mais sa vitesse de conversion peut être N fois moins grande que celle du CAN de la première configuration, avec N le nombre de colonnes.

[2]De l'anglais Layout ou mask design, ne pas confondre avec le stick diagram, le dessin de masque représente les géométries planaires des couches successives du procédé de fabrication et le stick diagram n'en est qu'une grossière représentation.

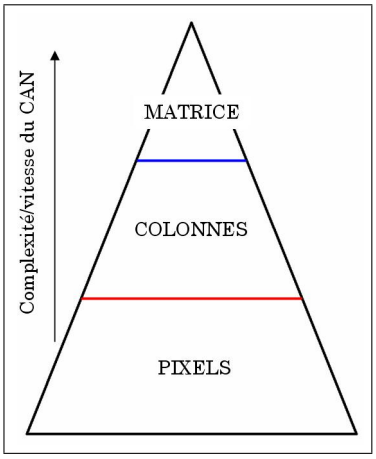


FIG. 2.13 – Hiérarchie architecturale dans un capteur CMOS

3. Le CAN est placé au niveau même du pixel (voir section 2.5). Dans ce cas précis, la taille du convertisseur a une importance capitale, son architecture doit être la plus simple possible tout en ayant une période de conversion suffisamment faible (≈100 μs). En effet la fréquence trame de l'imageur est alors directement liée à la période de conversion.

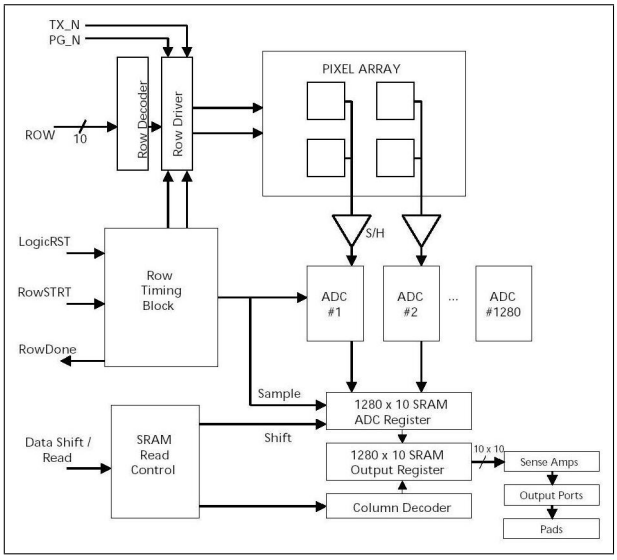


FIG. 2.14 – Synoptique du capteur CMOS industriel MTM9413 de Micron

Le milieu de la pyramide semble être un bon compromis entre performances du convertisseur (vitesse de conversion, non linéarité intégrale, etc) et densité d'intégration (surface occupée par le convertisseur). C'est cette

configuration qui est utilisée dans le circuit MTM9413 de Micron (voir le synoptique figure 2.14). C'est un capteur d'images CMOS rapide (500 images pas seconde) haute résolution (1280×1024). Ce capteur intègre la conversion analogique numérique (sur 10 bits) au niveau des colonnes, soit 1280 CAN pipeline 10 bits cadencés à la fréquence de 66 MHz.

2.6.2 Technologies de CAN

Il existe trois types de convertisseurs CMOS rapides : Le convertisseur flash, le convertisseur pipeline et le convertisseur Sigma Delta.

Le convertisseur Sigma Delta est décrit dans la précédente section 2.5. Le convertisseur flash décode directement les résultats d'un réseau de comparateurs (au nombre de 2^n-1, avec n la résolution du convertisseur) connectés à des références de tension (Vref et réseau de résistances). Les comparateurs comparent une valeur de référence avec le signal à convertir (cf. figure 2.15). L'avantage de cette architecture est sa vitesse de conversion, quelques centaines de picosecondes. Les inconvénients sont l'encombrement, quelques mm^2 en technologie CMOS 0,35 μm, et la consommation électrique, 2,7 W pour un convertisseur 8-bits à 400MHz.

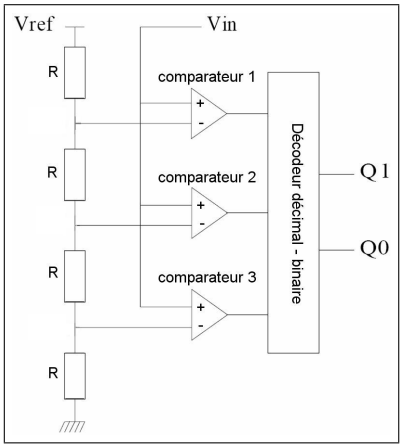


FIG. 2.15 – Synoptique d'un convertisseur flash

Le convertisseur pipeline est aussi un convertisseur rapide mais à moindre encombrement. Le cas particulier du convertisseur pipeline N étages 1-bit est intéressant pour sa simplicité et son efficacité. La figure 2.16 illustre le premier étage du convertisseur. Le principe est simple, le premier étage donne le bit de poids fort. Il consiste à comparer la valeur à convertir à la tension de référence divisée par deux, le résultat de la comparaison pilote un aiguilleur

(commutateur CMOS) qui connecte l'entrée négative du soustracteur à 0V ou à Vref/2. Si par exemple la valeur à convertir est supérieure à Vref/2 alors l'entrée est retranchée de Vref/2 avant d'être multipliée par 2 et injectée dans l'étage suivant. Chaque étage réalise un changement d'échelle de tension. La multiplication par 2 est une opération analogique qui introduit une certaine erreur qui se propage et s'amplifie à travers chaque nouvel étage. Il faut un certain temps de latence pour que le signal soit complètement converti et ce temps de latence est directement proportionnel à la résolution (environ 10 ns pour une résolution de 10-bits).

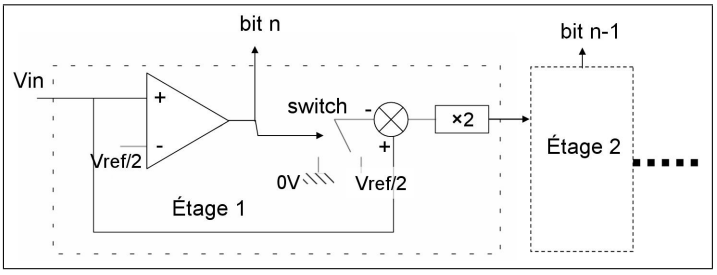


FIG. 2.16 – Synoptique d'un convertisseur pipeline

Ces convertisseurs (flash, pipeline et sigma-delta) sont généralement gourmands en transistors, donc généralement utilisés au niveau des colonnes ou de la matrice. La structure la plus simple est le convertisseur simple rampe, déjà implémenté dans le DPS [KLLEG01a] (voir section 2.5). Le tableau 2.2 liste les différentes possibilités d'implémentations de CAN dans un capteur CMOS.

TAB. 2.2 – Récapitulatif des différentes implémentations de CAN

Type de conversion	Précision	Position dans la pyramide
Flash	10-16 bits	Matrice
Pipeline	10-14 bits	Colonnes ou Matrice
Sigma-delta	8-12 bits	Pixels ou Colonnes
Simple rampe	8 bits	Pixels ou Colonnes

La recherche actuelle est orientée vers le pixel numérique qui permet la réalisation de capteurs ultra-rapides à haute résolution. Cependant, il est possible de faire d'autres compromis pour atteindre des performances certainement comparables, comme choisir par exemple d'implémenter un convertisseur par groupe de pixels formant ainsi une mosaïque dans une architecture

de capteur à haute résolution. Cette solution est mise en lumière dans le chapitre perspectives de ce manuscrit. Néanmoins dans notre réalisation, nous nous sommes placés au deuxième niveau de la hiérarchie en implémentant un convertisseur par colonne du capteur.

2.7 Caractéristiques physiques du pixel à intégration

Dans ce travail de thèse, notre capteur présente des similitudes avec les caractéristiques des capteurs d'image CMOS à intégration. En effet, le premier étage de notre pixel réalise l'intégration du photocourant d'une photodiode de type N. Dans cette section, nous présentons les éléments théoriques des principales caractéristiques des capteurs d'images CMOS de type APS. Un pixel ou photosite est constitué d'un élément photosensible, le photodétecteur. Le photo-détecteur privilégié dans la conception de capteurs CMOS à intégration est la photodiode verticale de type N pour deux raisons.

1. Les photodiodes verticales de type N ont un meilleur rendement que les photodiodes de type P.
2. Les règles de dessin sont beaucoup moins contraignantes car elles ne nécessitent pas de puits N qui imposent des distances minimales importantes [Gou95].

L'efficacité quantique, le facteur de conversion, le courant d'obscurité et la dynamique du photosite sont autant de paramètres fondamentaux qui sont mis en lumière dans cette partie de ce chapitre.

2.7.1 Réponse spectrale et efficacité quantique

Dans un pixel APS, les photons d'énergie $h\nu = h\frac{c}{\lambda}$, dans la gamme de longeur d'onde λ=[400 ;1100] nm, pénètrent la zone de charge d'espace et dissocient les paires électrons-trous, contribuant ainsi au courant de transit. Mais tous les photons incidents ne génèrent pas de photo-porteurs, on définit alors le rendement quantique en fonction de la longeur d'onde $\eta(\lambda)$. D'après l'expression du courant photonique 2.4, vue dans la section 2.3.2, le rendement quantique se déduit simplement par la relation inverse en introduisant la réponse spectrale R (en A.W^{-1}) relativement à la surface active du photodétecteur, il vient :

$$\eta(\lambda) = \frac{I_{ph}}{P_{opt}} \frac{h\ c}{\lambda\ q} = R\ \frac{h\ c}{\lambda\ q} \tag{2.14}$$

Les propriétés d'absorption et de réflexion des diffusions sont à l'origine de ce phénomène. Le rendement quantique dépend donc uniquement de la réponse spectrale du photodétecteur. En effet, le photocourant généré par la photodiode dépend de l'éclairement incident du pixel, de la surface du pixel et de la longueur d'onde. De plus, la surface active de la photodiode n'occupe qu'une partie de la surface totale du pixel, c'est ainsi que l'on définit le taux de remplissage α. L'efficacité quantique est le rendement quantique ramené à la surface du pixel, son expression est donnée par la relation 2.15.

$$QE(\lambda) = \eta(\lambda) \times \alpha \tag{2.15}$$

En technologie CMOS, le rendement quantique est de l'ordre de 70%. Le taux de remplissage est fonction de l'application, il va de 10% à 80%. Dans notre application, le taux de remplissage est de 10% et la réponse spectrale est mesurée dans la plage de longueur d'onde allant de 400 à 1100 nm.

2.7.2 Facteur de conversion

Dans un pixel à intégration du photocourant, en première approximation, on suppose le courant de transit constant (éclairement constant durant la phase d'intégration). Le potentiel cathodique subit une décroissance linéaire. La fonction de transfert qui lie l'intensité du courant de transit avec le potentiel cathodique est une transconductance que l'on nomme facteur de conversion. Il s'exprime généralement en $\mu V/e^-$. A partir de la relation 2.12 qui régit l'évolution du potentiel cathodique, avec les hypothèses précitées, il vient :

$$C_{int}\ dv(t) = I_{ph}\ dt = dq \tag{2.16}$$

D'où le facteur de conversion C_{vf} (en microvolt par électron injecté) obtenu en substituant la différentielle dq par la charge élémentaire q :

$$C_{vf} = \frac{q}{C_{int}} \tag{2.17}$$

La valeur de la capacité d'intégration de la photodiode détermine à elle seule le gain de conversion [Mag03].

2.7.3 Courant d'obscurité

Lorsque la photodiode est polarisée en inverse, des charges thermiques circulent à travers la jonction de la photodiode et contribuent de fait au courant d'obscurité. Il est mesuré en l'absence de toute illumination et dépend fortement de la température. Dans le cas du silicium, le courant d'obscurité double tous les 7°. Sa distribution spatiale en intensité obéit à une distribution gaussienne centrée sur la moyenne du courant thermique. La densité du courant d'obscurité est ainsi donnée en fonction de la température [Sze81] :

$$J = A\,T^{3/2}\,e^{-\frac{Vg\,q}{2\,k\,T}} \tag{2.18}$$

Où A est une constante exprimée dans les unités du système international, Vg est la tension de gap ($\approx 0,7\,V$ pour le silicium). Sa valeur dépend du procédé de fabrication et s'étend de 300 à 1600 $pA.cm^{-2}$.

Cette grandeur a des composantes géométriques (périmétrique et surfacique), on peut donc la minimiser dans le procédé standard de fabrication en agissant sur la géométrie de la photodiode, c'est ce qui a été privilégié dans ce travail de thèse.

2.7.4 Dynamique

La dynamique caractérise la capacité du pixel à mesurer des éclairements d'une certaine amplitude. La figure 3.11 représente les dynamiques de plusieurs capteurs d'images superposées à celle de l'œil humain.

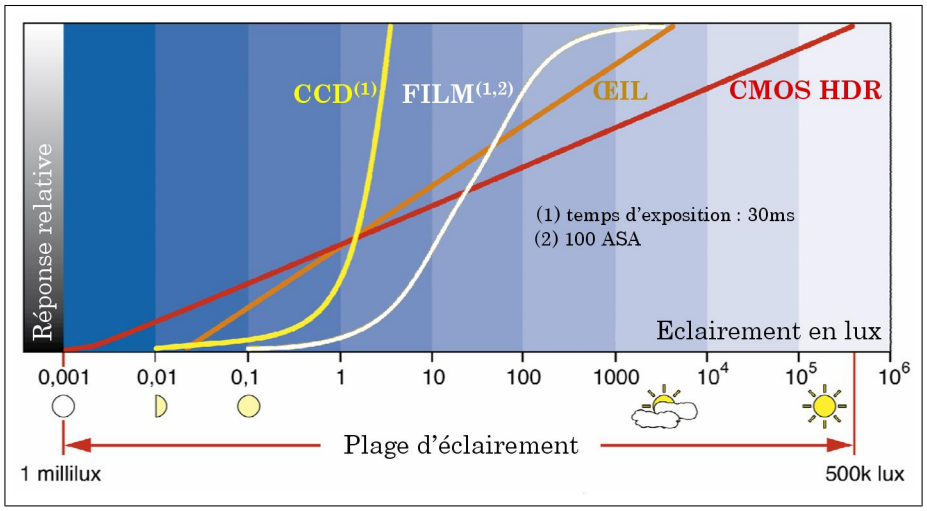


FIG. 2.17 – Réponses caractéristiques de différents capteurs

La dynamique est définie par le rapport de l'éclairement induisant une amplitude maximale (saturation aux forts éclairements) sur l'éclairement induisant une amplitude minimales (niveau de bruit aux faibles éclairements) en sortie du pixel. Son expression est donnée par la relation 2.19.

$$Dyn_{dB} = 20\ log_{10}\left(\frac{\Delta E\left(V_{max}\right)}{\Delta E\left(V_{min}\right)}\right) \tag{2.19}$$

La technologie CMOS supplante potentiellement la technologie CCD, notamment avec le pixel à conversion logarithmique [LMS01] qui permet d'atteindre une dynamique de plus de huit décades d'éclairement. Dans notre capteur, nous agissons sur le temps d'intégration d'un étage APS pour augmenter la dynamique.

Une autre grandeur caractéristique à prendre en considération dans la conception de capteurs d'images en général, et plus particulièrement dans le cas de la technologie CMOS, c'est le bruit. La section qui suit présente les sources de bruit dans les APS.

2.8 Le bruit dans les imageurs CMOS de type APS

Des différences de réponses existent entre les différents pixels, spatialement, mais aussi pour un même pixel entre deux acquisitions, temporellement. Ces variations aléatoires sont dues à des sources de bruit et se caractérisent par des imperfections sur l'image. De nombreux travaux démontrent que les sources de bruit dominantes dans un imageur CMOS sont décorrélées du signal utile [SDM93, Sei97, PTV98, DLMF00]. On peut donc supprimer les effets des sources de bruit par différence de niveaux de réponse durant l'intégration.

Par définition, le bruit est un signal électronique aléatoire (au sens des variables) qui se superpose au signal utile, si on suppose que le bruit et le signal utile sont décorrélés, ce qui se vérifie expérimentalement. Ce bruit est donc considéré comme une variable aléatoire ou non déterministe. C'est pour cette raison que l'on ne peut pas calculer directement sa puissance avec l'outil mathématique classique de Fourier (transformée de Fourier). On utilise donc l'intermédiaire de la fonction d'autocorrélation. D'après le théorème de Wiener-Kintchine, la densité spectrale de puissance d'un processus aléatoire, dont la valeur moyenne et la fonction d'autocorrélation sont invariantes dans le temps (stationnarité au sens large), est égale à la transformée de Fourier

de sa fonction d'autocorrélation (cf. relation 2.20 où * est la conjugaison complexe). Elle représente la répartition de la puissance de la variable aléatoire sur l'axe des fréquences.

$$\forall j\omega \in \mathbb{C}, \quad \Gamma(j\omega) = \int_{-\infty}^{+\infty} \left(\int_{-\infty}^{+\infty} x^*(t)x(t+\tau)dt \right) e^{-j\omega\tau} d\tau \tag{2.20}$$

Il existe cinq sources de bruits temporels dans une diode à jonction PN, polarisée en inverse : le bruit de grenaille, le bruit en 1/f, le bruit de génération/recombinaison, le bruit d'avalanche et le bruit thermique de lecture (thermal/reset noise). Ces cinq sources de bruit sont décorrélées, on peux donc appliquer le théorème de superposition et sommer les densités spectrales de puissance.

2.8.1 Le bruit de grenaille (shot noise)

Ce bruit est généré par la variation du nombre de photo-porteurs sous l'effet d'un champ électrique, dans la zone de charge d'espace. Si ce champ est fixe, la distribution suit la loi de Poisson [Sei97]. Ce bruit est donc proportionnel au courant auquel il se superpose [Deg00, Nav03]. Ce même courant est la somme du courant d'obscurité de la jonction PN (I_{dc}) et du photocourant (I_{ph}). La densité spectrale de puissance unilatérale de ce bruit est donnée par la relation 2.21 :

$$\forall f \in \mathbb{R}^+, \; DSP_{shot}(f) = 2\ q\ \ (I_{dc} + I_{ph}) \tag{2.21}$$

La densité spectrale de puissance est constante sur la bande de fréquence considérée, c'est un bruit blanc et additif. Le courant d'obscurité contribue fortement à l'écart type de ce bruit, il est donc important de minimiser ce courant d'obscurité (cf. section 2.7.3)

2.8.2 Le bruit en 1/f (flicker noise)

Aussi appelé bruit rose ou de scintillement, ce bruit a une densité spectrale de puissance qui varie en 1/f. Prépondérant dans les basses fréquences (≈100Hz), ce bruit contribue aux fluctuations du courant qui traverse la jonction PN [KOA91, VLR94, HKHC99, KSGA00]. Il est dû aux défauts du composant semiconducteur et dépend des paramètres du procédé de fabrication. Sa densité spectrale de puissance est donnée par la relation 2.22.

$$\forall f \in \mathbb{R}^+,\ DSP_{1/f}(f) = K_m \frac{I^a}{f^b} \tag{2.22}$$

Où K_m est une constante de construction, I est le courant continu total circulant à travers la jonction PN. Les constantes a et b sont des paramètres expérimentaux qui dépendent du procédé de fabrication.

Cette source de bruit peut être négligée dans le cadre de notre étude, d'autant plus dans le cas de faibles temps d'intégration.

2.8.3 Le bruit de génération - recombinaison

Le bruit de génération - recombinaison contribue aux variations du nombre de photo-porteurs traversant la jonction PN. Toutes les paires électrons-trous dissociées, sous l'effet de la lumière, ne génèrent pas toutes un photo-porteur susceptible de traverser la jonction PN. Certains de ces photo-porteurs se recombinent aléatoirement [Mat01]. La densité spectrale de puissance est donnée par la relation 2.23.

$$\forall f \in \mathbb{R}^+,\ DSP_{gr}(f) = K_{gr} \frac{4\ q\ I_d}{1 + (2\ \pi\ f\ \tau_{gr})^2} \tag{2.23}$$

Où K_{gr} est un paramètre expérimental qui dépend de la technologie et de la géométrie, τ_{gr} est la durée de vie des photo-porteurs, qui dépend de la concentration de dopage et I_d est le courant total traversant la jonction. Dans le cas des capteurs CMOS de type APS ce bruit est négligeable, surtout pour de faibles temps d'intégration.

2.8.4 Le bruit d'avalanche

Le bruit d'avalanche n'existe que si la jonction PN est fortement polarisée (au delà de 7 V). Il contribue à la fluctuation du facteur de multiplication des porteurs de charge [Amb98]. L'expression de la densité spectrale de puissance du courant de bruit d'avalanche est donnée par la relation 2.24.

$$\forall f \in \mathbb{R}^+,\ DSP_{av}(f) = 2\ q\ \overline{M}^2 I_d \tag{2.24}$$

Où I_d est le courant continu total traversant la jonction, et $\overline{M}$ est la moyenne du facteur de multiplication. Dans le cas de la technologie CMOS 0,35 μm, les différences de potentiels ne dépassent pas 3,3 V, donc cette source de bruit est négligée.

2.8.5 Bruit thermique et de lecture (thermal/reset noise)

Considérons le premier étage d'un pixel APS constitué d'une photodiode de type N et d'un transistor de préchargement de type NMOS (cf. figure 2.6). Pendant la phase de pré-chargement, le transistor "Reset" est fermé et la photodiode est polarisée en inverse. La photodiode est alors équivalente à une capacité C (la capacité de conversion au nœud flotant) et le transistor à une résistance R (R_{DSON}). La résistance génère un bruit thermique d'écart type σ modulé par la fonction de transfert H du circuit RC. La densité spectrale de puissance unilatérale de bruit se calcule de la façon suivante, avec k la constante de Boltzman et T la température du canal drain-source du transistor :

$$\begin{aligned}
DSP_{th} &= \int_0^{+\infty} (\sigma\,|H(f)|)^2\,df \\
DSP_{th} &= \int_0^{+\infty} \left(\sqrt{4\ k\ T\ R}\left|\frac{1}{1+j\ R\ C\ 2\ \pi\ f}\right|\right)^2 df \\
DSP_{th} &= \int_0^{+\infty} \frac{4\ k\ T\ R}{1+(R\ C\ 2\ \pi\ f)^2} df \\
DSP_{th} &= \frac{2\ k\ T}{\pi\ C}\ \int_0^{+\infty} \frac{du}{1+u^2} \\
DSP_{th} &= \frac{2\ k\ T}{\pi\ C}\ [\arctan(u)]_0^{+\infty} = \frac{k\ T}{C}
\end{aligned} \tag{2.25}$$

A la fin du préchargement (fin de l'impulsion de reset), la capacité de conversion est chargée par un transistor en zone de faible inversion. Il est démontré [SDM93] qu'à l'équilibre thermodynamique, la densité spectrale de puissance de la tension de bruit est toujours égale à $\frac{k\ T}{C}$, en considérant les potentiels aux nœuds drain et grille fixes.

La présence de ce bruit induit une variation du niveau de référence (niveau d'offset, voir section 2.4.2) à chaque nouvelle intégration. Il est prédominant dans le cas précis des pixels APS avec de faibles temps d'intégration. Sa densité spectrale de puissance est constante, c'est un bruit blanc, additif et décorrélé du signal utile.

A partir de la relation 2.16, l'écart type de tension de bruit vient s'ajouter à la différentielle ΔV et comme ce bruit blanc est décorrélé du signal utile, il vient :

$$C_{int}\ \Delta V + \sqrt{k\ T\ C_{int}} = I_{ph}\ \Delta T = \Delta Q \tag{2.26}$$

Cette relation met en lumière le fait que la contribution du bruit thermique de lecture est proportionnelle à la racine carré de la capacité C_{int}, vue au nœud flotant de la photodiode. De ce point de vue, il est important de minimiser la capacité parasite, qui ne contribue pas au photocourant. Deux solutions existent pour diminuer cette capacité parasite :

1. Les fondeurs (fabricants de circuits intégrés) ont mis au point des procédés technologiques, dits "Opto Process", dédiés à la conception d'imageurs CMOS qui permettent d'optimiser la conception de la photodiode verticale de type N, notamment par l'ajout d'une diffusion N faiblement dopée où la zone de charge d'espace peut s'étendre d'avantage et contribue à augmenter le gain de conversion. Toutefois, de tels procédés spécifiques engendrent un surcoût non négligeable.
2. Pour palier cet inconvénient il existe une deuxième solution qui consiste à optimiser la conception de la photodiode et notamment sa géométrie, ce qui permet d'augmenter sensiblement le gain de conversion.

C'est précisément ce dernier aspect qui a été privilégié dans ce travail de thèse.

2.8.6 Bruit spatial fixe (BSF)

Le bruit spatial fixe ou BSF (De l'anglais FPN pour Fixe Pattern Noise) contribue aux différences de réponses qui subsistent spatialement entre les différents pixels (cf. figure 2.18). Il est invariant dans le temps et dû principalement aux non uniformités des caractéristiques physiques des éléments qui constituent le pixel. C'est un bruit de construction, induit par le procédé même de fabrication. Cela implique par exemple, que deux transistors qui se trouvent à deux endroits distincts de la matrice de pixels, peuvent avoir leurs tensions de seuil différentes de quelques mV (pour une référence de 0,5 V).

La disparité des tensions de seuil des transistors est la source de bruit fixe dominante [Sol96, PTV98, BSPS97, DLMF00]. Si l'imageur CMOS, de type APS, dispose d'amplificateurs de colonnes (architecture classique), alors le bruit spatial engendré par la disparité des caractéristiques des amplificateurs, contribue au bruit fixe de colonnes. Ce bruit fixe de colonnes trouve son origine dans les non uniformités de gain et d'offset des amplificateurs, ce qui entraîne une disparité des réponses en l'absence de signal. Dans ce cas, le bruit fixe de colonne est prépondérant sur le bruit fixe de pixel. Comme le montre la figure 2.18(b), des bandes verticales apparaissent sur l'image.

La firme autrichienne AMS (AustriaMicroSystem) effectue des tests, sous pointes, de leurs structures électroniques sur trois Wafers à raison de cinq

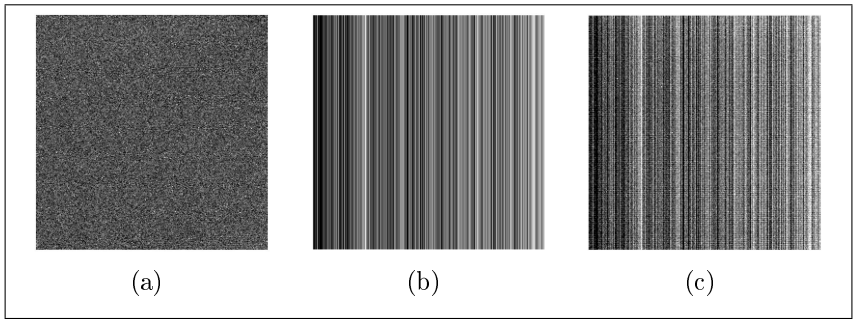


FIG. 2.18 – (a) Bruit spatial fixe de pixel , (b) bruit spatial fixe de colonne , (c) bruit spatial fixe total

sites par Wafer soit 15 mesures pour chacun des paramètres des structures de tests. Les structures de tests sont des composants CMOS courants : Transistors NMOS, PMOS, résistances et condensateurs. Intéressons nous, par exemple, à la tension de seuil d'un transistor NMOS. le paramètre est VTO10X10N, c'est à dire la tension de seuil d'un transistor NMOS dont la largeur de canal et la largeur de grille sont chacune de 10 μm. En technologie 0,35 μm, les résultats expérimentaux en sortie de chaîne de fabrication donnent un écart type de 0,011978 V pour une moyenne de 0,462933 V soit 2,5873% de la valeur moyenne, pour le seul paramètre VTO10X10N. Pour conclure sommairement sur ce résultat, en fonction de la complexité de l'électronique du pixel et éventuellement de l'amplificateur de colonne, le bruit spatial fixe est de l'ordre de quelques %, en technologie CMOS AMS 0,35 μm.

Le bruit spatial fixe et le bruit thermique de lecture ($\frac{k\,T}{C}$) sont les deux sources principales de bruit dans notre application. La solution la plus efficace pour compenser ces sources de bruit est la mise en œuvre de la technique CDS.

2.9 Compensation du bruit dans les imageurs CMOS à intégration

Pour compenser le bruit de lecture et d'éventuelles autres sources de bruit (bruit de grenaille, en 1/f...), une technique dite CDS (de l'anglais Correlated Double Sampling ou double échantillonnage corrélé) est particulièrement efficace. Elle consiste à réaliser deux lectures successives d'un même pixel pendant la lecture d'une trame (cf. figure 2.19).

Lors de l'ouverture du transistor de Reset (cf. figure 2.6, section 2.4.2),

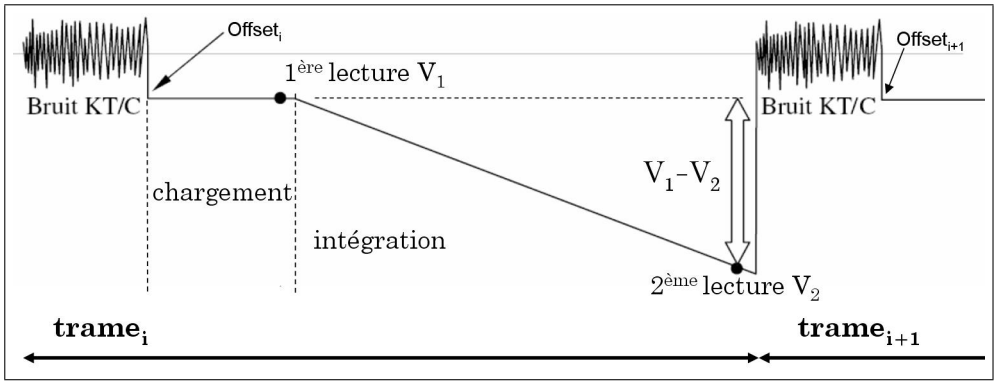


FIG. 2.19 – Technique de réduction du bruit dans les pixels APS

le potentiel cathodique subit une chute de tension, que l'on appelle tension d'offset. Cette tension d'offset, caractéristique du bruit de lecture de remise à zéro du pixel, a une distribution aléatoire. Le bruit de lecture est décorrélé du signal utile. C'est pourquoi la différence de deux potentiels, lue à deux instants différents pendant la phase d'intégration, élimine les effets des sources de bruits indésirables [MKG+97]. Néanmoins, il faut stocker le résultat de la première lecture durant le temps d'intégration.

2.10 Conclusion

Dans ce deuxième chapitre, nous avons traité différents aspects de la constitution des capteurs d'images CMOS, les différentes technologies de pixels, de CAN mais aussi les caractéristique physiques du pixel APS. Nous avons insisté sur les caractéristiques physiques et les modèles analytiques de la photodiode verticale de type N car nous utilisons cette technologie photodétecteur dans notre réalisation en apportant une réflexion supplémentaire sur sa géométrie. Cette étude est justement l'une des parties du chapitre suivant.

En conclusion de ces deux premiers chapitres, il y a ainsi, au sein même du pixel, un compromis à faire sur l'implémentation entre des structures purement numériques (processeur numérique associé à un CAN) et des structures analogiques. Cependant, il est important de comprendre qu'il n'y a pas une frontière physique mais seulement conceptuelle entre les deux approches. En effet, on bascule déjà dans le "monde numérique" dès lors qu'on réalise, avec des fonctions analogiques, une comparaison ($\lambda \leq \alpha$) entre deux signaux à valeurs réelles et bornées (limitées en amplitude par la tension d'alimentation par exemple). Une solution relativement satisfaisante serait de ne considé-

rer que les paliers (en potentiel par exemple) des signaux considérés ou plus exactement le codage multi-niveaux (signaux M-aires) comme cela se fait en transmissions numériques en bande de base sur des canaux peu bruités. En effet, nous avons vu que le bruit fixe, dû aux non uniformités des transistors, est relativement important ainsi que le bruit thermique. On imagine bien que ce phénomène s'amplifie avec la complexité des traitements implémentés au sein du pixel et c'est pourquoi, considérer le codage du signal en plusieurs niveaux peut permettre de minimiser les effets de ces sources de bruits indésirables.

Le chapitre suivant porte sur la conception de notre rétine, qui est d'abord inspirée de la technologie APS (à intégration du photocourant), mais avec une architecture encore inédite de pixel analogique. Nous nous sommes fixés une taille de pixel de 35 μm dans un technologie standard CMOS 0,35 μm. Notre pixel intègre des fonctionnalités dédiées à l'imagerie rapide, aux traitements d'images linéaires spatiaux (détecteurs de contours) et spatio-temporels (détection de mouvements).

Chapitre 3

Architecture globale et algorithmes embarqués sur HISIC

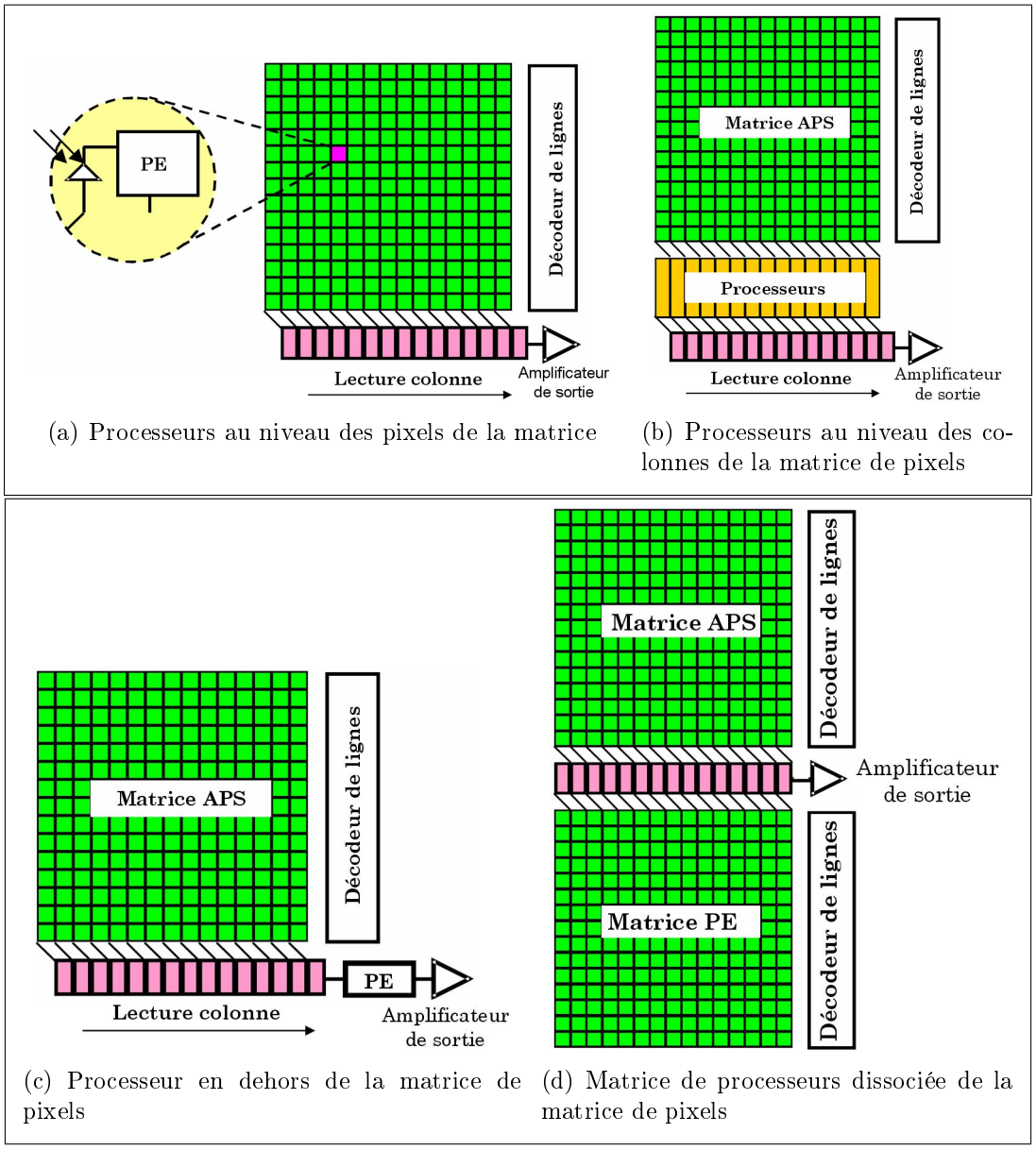
3.1 Introduction

Nous avons vu dans les précédents chapitres qu'une rétine artificielle CMOS est un capteur d'images intégrant des fonctions de traitement de bas niveau sur l'image visant à raréfier les données "pixels" avant d'être déplacées vers une unité de calcul de plus haut niveau placée en dehors de la matrice de pixels sur le même Wafer (approche ViSOC) ou sur le même PCB (Printed Circuit Board) ou circuit imprimé. Ces fonctions peuvent jouer un rôle déterminant dans un système global de vision artificielle. Nous nous sommes intéressés à la conception d'un système de vision embarqué où l'unité de calcul haut niveau est un circuit numérique programmable de type FPGA interfacé au circuit HISIC.

Notre pixel intègre des fonctionnalités dédiées à l'imagerie rapide, aux traitements d'images linéaires spatiaux (détecteurs de contours) et spatio-temporels (détection de mouvements). A l'issue de ces travaux de thèse, nous avons conçu un système de vision artificielle dédié à l'imagerie rapide et aux traitements d'images linéaires. Plus exactement nous avons réalisé un prototype pré-industriel de caméra embarquée pouvant réaliser des acquisitions à plus de 10 000 images par seconde, et des traitements spatiaux et/ou spatio-temporel de bas niveau à près de 2000 images par seconde. Dans ce chapitre sont successivement décrites l'architecture globale et les possibilités algorithmiques du système.

3.2 Architecture globale du système de vision artificielle

Le premier chapitre fait état d'une partie des solutions existantes en matière de rétines CMOS programmables. Parmi les solutions, on distingue quatre catégories (cf. figure 3.1).



(a) Processeurs au niveau des pixels de la matrice

(b) Processeurs au niveau des colonnes de la matrice de pixels

(c) Processeur en dehors de la matrice de pixels

(d) Matrice de processeurs dissociée de la matrice de pixels

FIG. 3.1 – Architectures de rétines CMOS programmables

Nous nous plaçons dans la première catégorie où le processeur élémentaire est implanté au sein du pixel. De cette façon, la rétine est équivalente à une machine SIMD massivement parallèle où les éléments de traitement sont au plus près du photodétecteur. La figure 3.2 représente le schéma bloc du circuit. Son architecture contient de trois blocs principaux, comme beaucoup de capteurs d'images décrits dans la littérature.

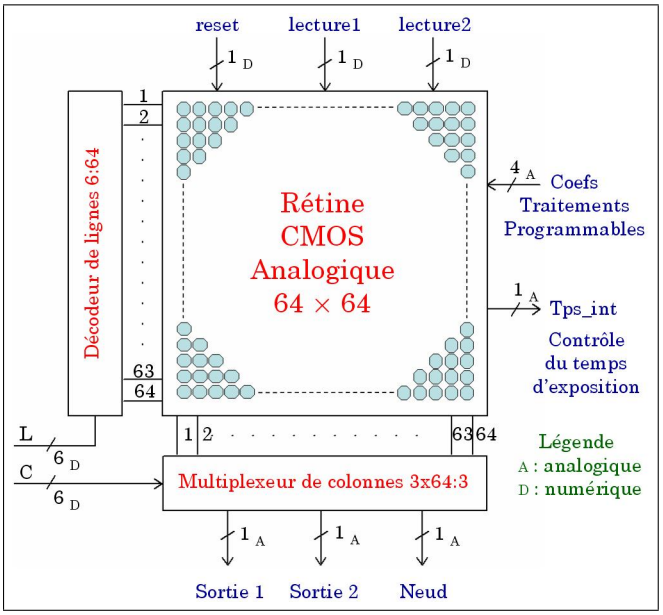


FIG. 3.2 – Schéma bloc du circuit HISIC

D'abord, la matrice de 64×64 pixels est placée au centre du diagramme. Puis, placés autour de la matrice, le décodeur de lignes et le multiplexeur de colonnes, assurent ainsi l'accès aléatoire aux pixels de la matrice. Le décodeur et le multiplexeur sont pilotés par les signaux numériques "L" et "C" sur 6 bits chacun. Trois bus analogiques asynchrones sont multiplexés vers la sortie de la rétine : Le premier bus "Nœud" est dédié aux traitements analogiques et les deux autres à l'imagerie rapide.

Sept signaux, "lecture1", "lecture2", "reset" et les quatre coefficients (coef_i, $i \in \{1;4\}$), sont globaux à toute la matrice. Un changement de niveau de tension sur ces signaux a un impact sur tous les pixels de la matrice. Les quatre coefficients sont donc, à chaque instant, les mêmes pour tous les pixels de la matrice et fixent les valeurs de pondérations des combinaisons linéaires avec le proche voisinage 2×2 du pixel correspondant.

Comme le montre la figure 3.3, chaque pixel du circuit contient une photodiode, un étage APS, deux structures MA^2M (Mémoires Analogiques, Amplificateur et Multiplexeur) et une UA^2 (Unité Arithmétique Analogique).

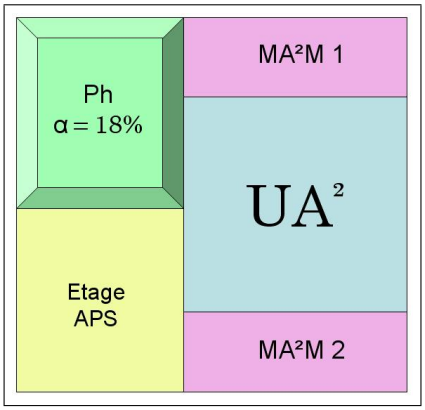


FIG. 3.3 – Schéma bloc du pixel d'HISIC

Le photodétecteur utilisé est une photodiode verticale de type N. La justification de ce choix technologique est faite dans le chapitre suivant. La surface des pixels est de 35×35 μm^2 et le taux de remplissage α est de 18%. La transduction éclairement-signal (étage APS) au niveau du pixel est réalisée par intégration du photocourant comme dans le cas des pixels actifs.

Une unité arithmétique analogique (UA^2) est capable de calculer le résultat d'une combinaison linéaire de quatre pixels connexes. L'opération "combinaison linéaire des quatre pixels voisins" permet d'envisager de nombreuses possibilités algorithmiques, comme les traitements spatiaux. La pondération de la combinaison linéaire étant reconfigurable, les perspectives d'intégration d'algorithmes sont grandes, comme le montrent les paragraphes qui suivent.

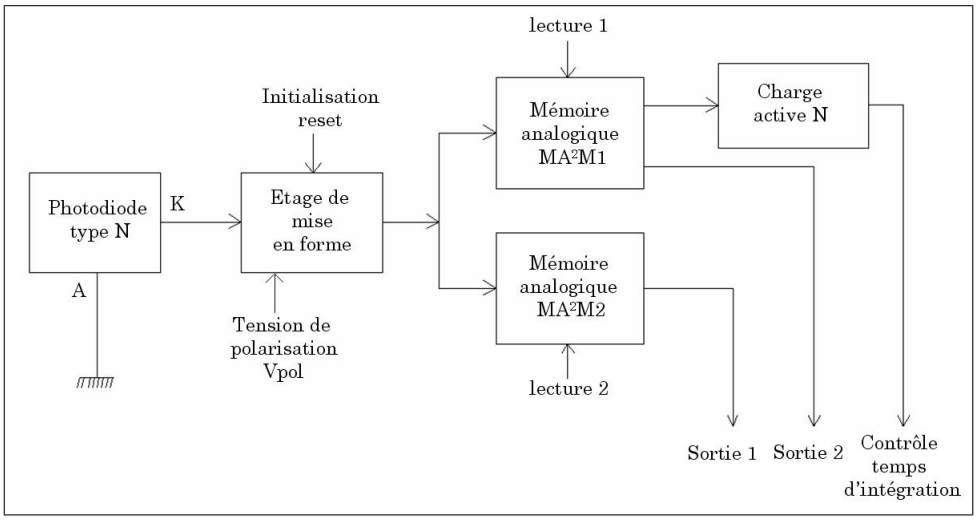


FIG. 3.4 – Synoptique du pixel dédié à l'imagerie rapide

La figure 3.4 illustre le synoptique du premier étage du pixel. L'étage de mise en forme assure la transduction éclairement - tension au plus près du photodétecteur. Deux structures dites MA^{2}M (Mémoire Analogique Amplificateur et Multiplexeur) permettent de mémoriser et amplifier le signal issu de ce premier étage. Avec ces deux structures MA^{2}M, la trame d'acquisition peut être stockée dans la première mémoire, à l'issue d'une période d'exposition, en parallèle avec la lecture de la trame précédemment acquise dans la deuxième mémoire. Cette stratégie est courante en programmation, elle porte le nom de "Ping-Pong" (cf. figure 3.5).

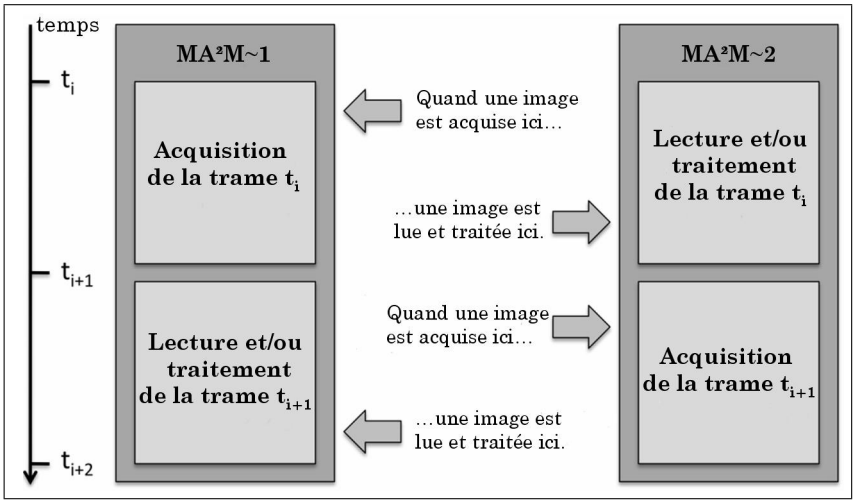


FIG. 3.5 – Parallélisme de tâches entre séquences d'acquisitions et lectures séquentielles (stratégie Ping-Pong)

Cette stratégie a plusieurs avantages :

1. La vitesse d'acquisition est doublée. En effet, durant le temps d'exposition, l'une des deux mémoires est lue, en l'occurrence celle qui contient l'information éclairement du pixel de la trame précédente. L'autre mémoire est écrite à l'issue d'un temps d'intégration qui vaut au maximum le temps de lecture d'une trame.
2. L'acquisition d'images est dissociée du traitement d'images, ainsi la vitesse d'exécution de l'unité de calcul analogique (voir la section qui suit, pour les détails concernant cette structure) est toujours maximale.
3. L'image de la trame précédemment acquise est toujours disponible sans attendre le temps d'intégration nécessaire à l'obtention d'une nouvelle image.

Le signal "reset", sur le schéma bloc de la rétine (cf. figure 3.2), précharge toutes les photodiodes de la matrice, afin de diminuer les distorsions

présentes lors d'une commande ligne par ligne. Au niveau de potentiel 0V (appliqué sur la grille du transistor de reset), toutes les photodiodes de la matrice sont polarisées en inverse et au niveau de potentiel $+V_{DD}$. L'intégration du photocourant fait décroître le potentiel cathodique, proportionnellement à l'éclairement incident du pixel. C'est alors que les signaux "lecture1" et "lecture2" pilotent les cellules MA^2M1 et MA^2M2, parallèlement dans tous les pixels de la matrice. Si le signal "lecture1" ou "lecture2", à l'issue du préchargement, est à l'état haut alors le niveau de tension, image de l'éclairement du pixel, est recopié dans la cellule MA^2M1 ou MA^2M2.

Le signal appelé "Tps_int" est une des sorties analogiques de la rétine, qui informe sur l'éclairement moyen incident au plan focal image du capteur. De cette façon, si l'éclairement moyen est trop bas, le temps d'intégration peut être augmenté, à contrario si l'éclairement moyen est trop haut, le temps d'intégration peut être diminué. Ce temps d'intégration est le temps qui sépare la fin de la période de préchargement fixée par le signal "reset" et l'instant où le niveau de potentiel est stocké dans la cellule MA^2M1 ou MA^2M2. Cette architecture permet de contrôler automatiquement le temps d'intégration en fonction de l'éclairement incident moyen au plan focal image et ainsi d'augmenter considérablement la dynamique du capteur.

Avant de détailler l'architecture microélectronique des structures électroniques évoquées ci-avant, intéressons nous aux algorithmes pouvant être exécutés sur notre rétine.

3.3 Traitements et algorithmes embarqués

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 :

1. La première catégorie inclut les traitements dits "intra-pixel" dont le but principal est l'amélioration de la qualité de l'image. Un exemple typique est le pixel actif (APS) [YPPS+97, YPB03] (cf. figure 3.6(a)).
2. La seconde catégorie correspond aux traitements dits "inter-pixels" dans lesquels les traitements sont basés sur des voisinages de pixels pour non seulement acquérir l'image mais aussi dans le but d'exécuter des pré-traitements sur l'image. Les transistors, qui sont placés autour du détecteur photoélectrique, au niveau du pixel, peuvent être vus comme un processeur à part entière de traitement du signal ayant pour

but d'augmenter les fonctionnalités du capteur. Ces processeurs analogiques permettent typiquement de réaliser des calculs sur plusieurs pixels.

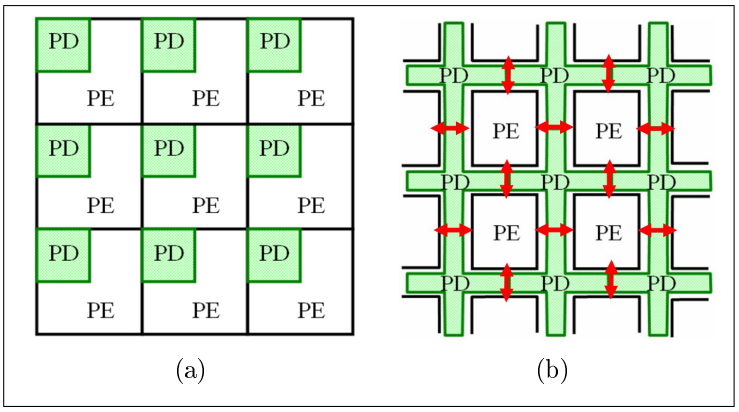


FIG. 3.6 – Photosite avec processeur : (a) intra-pixel, (b) inter-pixel

Notre travail se place dans cette deuxième catégorie, les architectures distribuées en topologie 4-connexe, car notre objectif est l'implémentation au niveau du pixel de traitements d'images reconfigurables et/ou programmables utilisant un voisinage local comme par exemple les gradients spatiaux. 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 (cf. figure 3.6(b)), afin de limiter d'une part les interconnexions entre processeurs élémentaires et d'autre part de réfléchir à une forme optimale de photodiode. Notre objectif principal est l'exécution de divers traitements d'images in situ basés sur des voisinages locaux, tels que les gradients spatiaux comme par exemple les opérateurs Sobel et Laplacien.

Au premier abord, les principales applications de notre capteur s'articulent autour de l'extraction d'éléments caractéristiques d'une scène en mouvement, la détection rapide de contour et la pré-analyse de mouvements sans marqueur.

3.3.1 Gradient spatial

Notre structure de site de traitement connexe à quatre photo-détecteurs se prête très bien à la mesure du gradient spatial (cf. figure 3.7).

Une image numérique en niveaux de gris peut être vue comme la discrétisation d'une fonction continue à deux dimensions. Par exemple, un relief

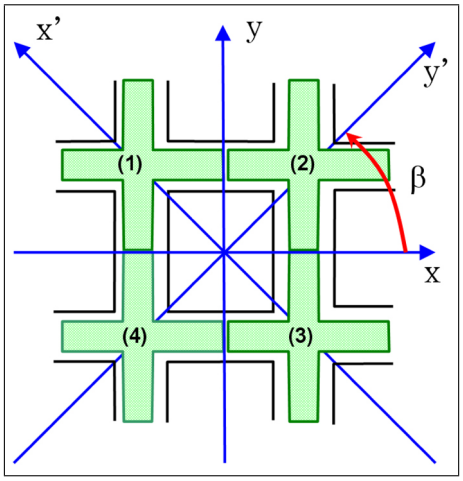


FIG. 3.7 – Mesure du gradient spatial

montagneux peut être décrit par une telle application de $\mathbb{R}^2$ vers $\mathbb{R}$, dont l'espace de départ serait les coordonnées au sol et l'espace d'arrivée l'altitude du relief. L'idée principale pour évaluer un gradient spatial sur une image [BBM+02] est basée sur la définition de la dérivé première partielle, dans la direction du vecteur contour $\vec{\xi}$, d'une application linéaire à deux dimensions exprimée comme suit :

$$\frac{\partial V(x,y)}{\partial \vec{\xi}} = \frac{\partial V(x,y)}{\partial x'} cos(\beta) + \frac{\partial V(x,y)}{\partial y'} sin(\beta) \tag{3.1}$$

Où β est l'angle formé avec la direction du vecteur contour $\vec{\xi}$. Une discrétisation de l'équation 3.1, au niveau du pixel, selon la figure 3.7, donnerait :

$$\frac{\partial V}{\partial \vec{\xi}} = (V_2 - V_4)\, cos(\beta) + (V_1 - V_3)\, sin(\beta) \tag{3.2}$$

Où V_i, $i \in \{1;4\}$ est l'image de l'éclairement du pixel i, *i.e.*, le signal issu de la photodiode, mis en forme et amplifié. De cette façon, la dérivée locale dans la direction du vecteur contour $\vec{\xi}$ est la combinaison linéaire, dans une base orthonormée, des dérivées premières discrétisées dans les directions x' et y'. Les contributions des pixels du voisinage peuvent être dissociées pour donner l'équation 3.3.

$$P = V_1\, sin(\beta) + V_2\, cos(\beta) - V_3\, sin(\beta) - V_4\, cos(\beta) \tag{3.3}$$

Pour réaliser l'opération P, il faut quatre multiplieurs quatre quadrants, car l'espace d'arrivée de l'application linéaire est dans $\mathbb{R}$. Les fonctions cosinus et sinus sont à valeurs dans $\mathbb{R}$, donc positives ou négatives ainsi que le résultat de l'application P.

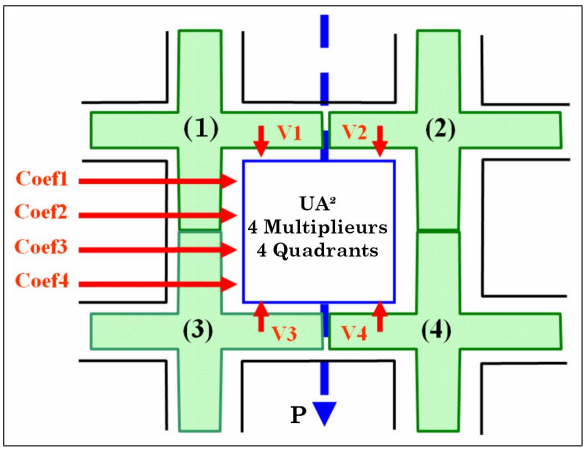


FIG. 3.8 – Implémentation d'une Unité Arithmétique Analogique ou UA^2, composée de quatre multiplieurs quatre quadrants, au niveau du pixel

Ainsi, la fonction P donnée par la relation 3.3 peut être implémentée (cf. figure 3.8), à l'aide de quatres multiplieurs quatre quadrants constituant une unité arithmétique analogique UA^2. Cette architecture d'UA^2 est détaillée dans le chapitre suivant. Cette opération P demande l'équivalent de quatre multiplications signées sur une architecture externe à la matrice de pixels. Ici, l'opération est réalisée au niveau du pixel et cela pour tous les pixels de la matrice (de résolution 64×64) soit l'équivalent de 16384 multiplieurs signés travaillant en parallèle.

En conséquence, le traitement mis en œuvre au niveau du pixel réalise une combinaison linéaire de quatre pixels adjacents, pondérée par quatre coefficients ($\text{coef}_i, i \in \{1;4\}$). Ainsi pour évaluer P, donnée par la relation 3.3, les coefficients forment une matrice de rotation définie par l'angle β, dans un espace vectoriel euclidien orienté réel de dimension 2 :

$$\begin{pmatrix} coef1 & coef2 \\ coef4 & coef3 \end{pmatrix} = \begin{pmatrix} sin(\beta) & cos(\beta) \\ -cos(\beta) & -sin(\beta) \end{pmatrix} \tag{3.4}$$

Les gradients horizontaux et verticaux peuvent être simplement évalués en fixant respectivement la valeur de β à 0 et à $\pi/2$ rad.

3.3.2 Opérateur Sobel

La structure de notre architecture est également bien adaptée à divers algorithmes de convolution basés sur des voisinages de pixels. Par exemple, l'évaluation d'un opérateur convolutif de dimension 3×3, avec notre circuit, est directement obtenue dans une base orthonormée (cf. figure 3.9(a). Afin d'exécuter cette opération mathématique, un masque 3×3 est appliqué à l'image, suivant la figure 3.9(b). a_i, $i \in \{1;9\}$ sont les coefficients du masque et $(k), k \in \{1;4\}$, les sites de traitements où sont réalisés les opérations de type combinaison linéaire.

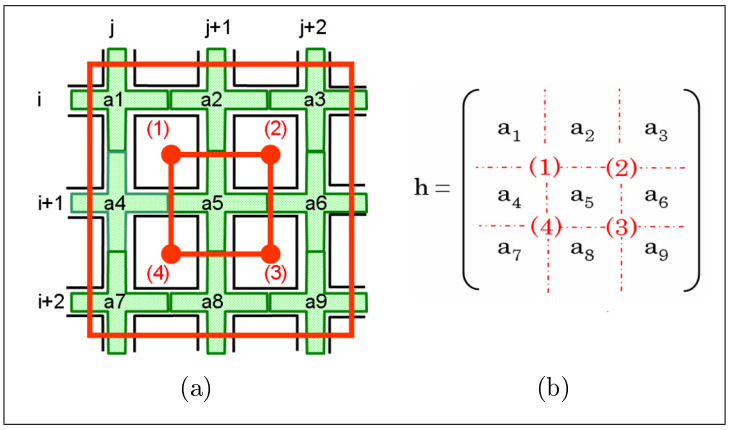


FIG. 3.9 – (a) Voisinage de pixels, (b) masque 3×3 appliqué localement autour de quatre éléments de calcul

Pour exécuter l'algorithme de Sobel, dans une base orthonormée, il est nécessaire de construire deux matrices h_1 et h_2 de dimension 3×3 :

$$h_1 = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad h_2 = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \tag{3.5}$$

Concrètement, quatre masques de dimension 2×2 sont successivement appliqués à tous les processeurs élémentaires de la rétine. D'après l'équation 3.5, l'exécution de l'algorithme de Sobel dans les quatre processeurs élémentaires k, $k \in \{1;4\}$ (cf. figure 3.9(a)) doit se traduire par une série d'opérations élémentaires :

$$h_1 : \begin{array}{l} I_{11} = -(I_1 + I_4) \\ I_{12} = +(I_3 + I_6) \\ I_{13} = +(I_6 + I_9) \\ I_{14} = -(I_4 + I_7) \end{array} \qquad h_2 : \begin{array}{l} I_{21} = -(I_1 + I_2) \\ I_{22} = -(I_2 + I_3) \\ I_{23} = +(I_8 + I_9) \\ I_{24} = +(I_7 + I_8) \end{array} \tag{3.6}$$

Où les valeurs I_{1k} et I_{2k} proviennent des processeurs élémentaires (k). Puis, de ces opérations élémentaires sont déduites les amplitudes suivant l'axe vertical ($I_{h1} = I_{11}+I_{12}+I_{13}+I_{14}$) et l'axe horisontal ($I_{h2} = I_{21}+I_{22}+I_{23}+I_{24}$). Ainsi, l'évaluation d'un tel gradient peut être réalisée à l'issue de quatre cycles rétiniens définis comme l'exécution d'une opération élémentaire. A chaque nouveau cycle, les coefficients de l'UA^2 sont reconfigurés aux valeurs du masque élémentaire 2×2 qui traduit une des opérations élémentaires I_{1k} ou I_{2k}.

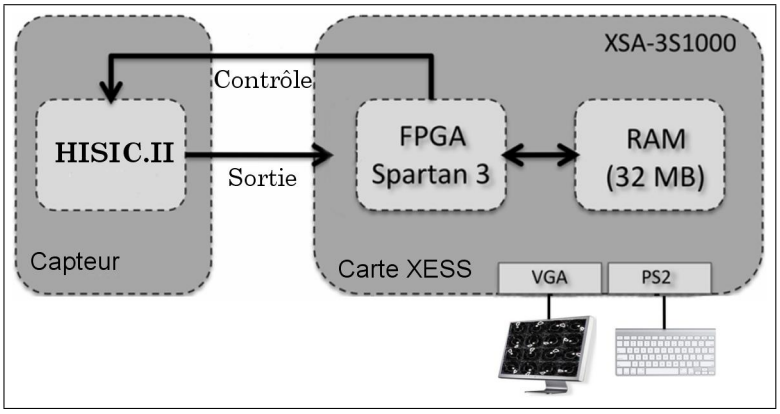


FIG. 3.10 – Schéma bloc du système d'acquisition

A l'issue de ces cycles rétiniens, les coefficients de la matrice sont configurés quatre fois, soit deux configurations par direction, à l'aide d'un FPGA placé physiquement en dehors du circuit (cf. figure 3.10). Les résultats sont accumulés dans une mémoire externe à la matrice de pixel et cela à chaque nouvelle configuration. A la fin des cycles rétiniens, les valeurs stockées dans la mémoire externe sont directement exploitables, ce qui constitue déjà un pré-traitement spatial de l'image entre deux trames. En effet, dans notre architecture, tous ces cycles rétiniens peuvent être réalisés à l'issue d'une seule acquisition de l'image constituant une trame. Entre deux trames, il est possible de réaliser plusieurs cycles, 400 au maximum et ces cycles ont une durée de 100 μs, soit 40 Mpixel/s en résolution 64×64.

Par construction, les coefficients appliqués sont les mêmes dans toute la matrice, ce qui permet de paralléliser les calculs. De cette façon, notre rétine est une matrice de processeurs analogiques SIMD massivement parallèle. De plus, à l'issue de la phase d'acquisition, des structures à mémoires analogiques (MA^{2}M 1 et MA^{2}M 2 sur la figure 3.3) conservent l'information "éclairement incident" dans chaque pixel. Il est donc possible de reconfigurer plusieurs fois les coefficients appliqués entre deux trames et donc d'exécuter plusieurs opérations de base 4-connexes.

A l'issue de la phase d'acquisition, dans une première lecture trame, pour évaluer par exemple la moitié gauche de l'amplitude verticale du gradient de Sobel $I_{11} + I_{14}$, les coefficients de l'Unité Arithmétique Analogique doivent être configurés comme suit :

$$\begin{pmatrix} coef1 & coef2 \\ coef3 & coef4 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} \tag{3.7}$$

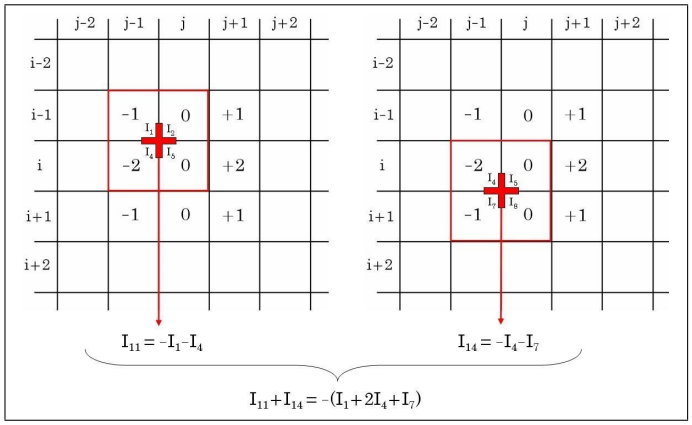
Il s'agit là de la première configuration de l'unité de calcul analogique. Ensuite, la moitié droite de l'amplitude verticale du gradient de Sobel $I_{12} + I_{13}$, s'obtient en fixant les coefficients aux valeurs suivantes :

$$\begin{pmatrix} coef1 & coef2 \\ coef3 & coef4 \end{pmatrix} = \begin{pmatrix} 0 & +1 \\ 0 & +1 \end{pmatrix} \tag{3.8}$$

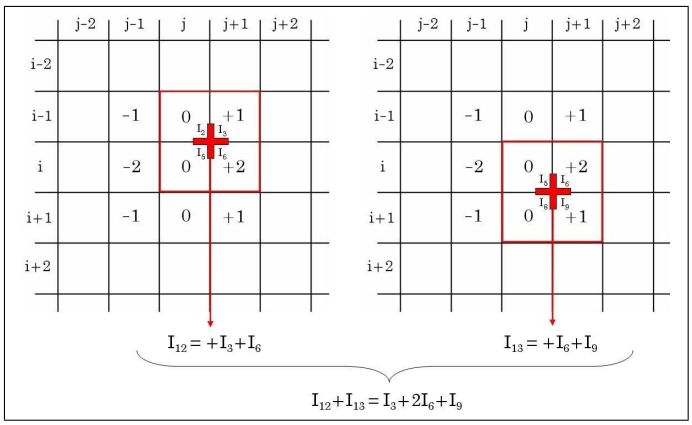
Pour poursuivre l'évaluation du gradient, il faut reconfigurer les coefficients en conséquence et accumuler les résultats. Il suffit d'accumuler les valeurs mémorisées, à l'aide d'un processeur externe, afin d'évaluer les amplitudes du gradient. Ainsi, les unités de calcul implantées au sein du pixel allègent considérablement le travail du processeur externe (un FPGA dans notre système) (cf. figure 3.10). Son rôle se limite à accumuler et à stocker les valeurs issues des unités de traitements.

La figure 3.11 montre comment, à l'issue de la phase d'acquisition, les macro opérations I_{11} et I_{14} sont évaluées lors de la première configuration. Puis les valeurs I_{12} et I_{13} sont évaluées lors de la deuxième configuration, dans le but de calculer l'amplitude verticale du gradient de Sobel. L'algorithme à appliquer est le suivant :

1. Faire l'acquisition d'une image (100 μs).
2. Fixer les coefficients aux quatre valeurs données par la relation 3.7.
3. Adresser séquentiellement toute l'image et lire les résultats issus des UA^2 (100 μs soit 40Mpixels/s en 64×64).
4. Accumuler les valeurs lues dans des mémoires placées à l'extérieur de la rétine (pas de MAC).
5. Changer les coefficients par les quatre valeurs données par la relation 3.8.
6. Adresser séquentiellement toute l'image et lire les résultats issus des UA^2 (100 μs par configuration et relecture).
7. Accumuler les valeurs lues dans des mémoires.
8. A cette dernière étape, les mémoires contiennent les valeurs de gradient directement exploitables.



(a) Première configuration et lecture trame



(b) Deuxième configuration et relecture

FIG. 3.11 – Séquence de reconfiguration dynamique pour la composante verticale du filtre de Sobel

En théorie, on peut extraire les contours d'une scène, avec cette méthode en seulement 300 μs, c'est-à-dire à 3330 images/s. L'opérateur Laplacien se traite de la même façon, avec quatre configurations dynamiques successives. En suivant cette ligne stratégique, il est possible d'augmenter significativement la vitesse des traitements d'images, l'objectif étant d'alléger le travail du processeur externe et surtout de diminuer le temps d'exécution de l'algorithme.

3.3.3 Détecteur du second ordre : Laplacien

Une détection de contours, basée sur la dérivée du second ordre, comme l'opérateur Laplacien, peut être implémentée dans notre architecture. Dans

ce cas, c'est le passage par zéro du signal traité qui donne la position du contour recherché. Contrairement au gradient de Sobel, le Laplacien est un scalaire, il ne nous informe pas sur la direction du contour mais uniquement sur sa position. En générale, on utilise un des deux Laplaciens suivant :

$$\triangle_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}) \qquad \triangle_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{3.9}$$

En raisonnant de la même façon que précédemment, de ces deux masques 3×3, on extrait les séries d'opérations suivantes :

$$\triangle_1 : \begin{array}{l} I_{11} = I_4 - I_5 \\ I_{12} = I_2 - I_5 \\ I_{13} = I_8 - I_5 \\ I_{14} = I_6 - I_5 \end{array} \qquad \triangle_2 : \begin{array}{l} I_{21} = 0,5(I_1 + I_2) - I_5 \\ I_{22} = 0,5(I_3 + I_6) - I_5 \\ I_{23} = 0,5(I_4 + I_7) - I_5 \\ I_{24} = 0,5(I_8 + I_9) - I_5 \end{array} \tag{3.10}$$

On obtient ainsi, I$\triangle_1 = I_{11}+I_{12}+I_{13}+I_{14}$ et I$\triangle_2 = 2\times(I_{21}+I_{22}+I_{23}+I_{24})$, les amplitudes discrétisées des deux Laplaciens présentés ici. Chacune de ces opérations I$\triangle_1$ et I$\triangle_2$ nécessitent seulement quatre configurations de l'unité de calcul. Pour le deuxième Laplacien, les coefficients sont des signaux multi-niveaux 4 aires (coef$\in \{-1; -0,5; +0,5; +1\}$). Les quatre configurations successives, nécessaires à l'exécution de l'opérateur I$\triangle_1$, sont récapitulées ci-après :

$$\begin{array}{ll} I_{11} : \begin{pmatrix} 0 & 0 \\ 1 & -1 \end{pmatrix} & I_{12} : \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \\ I_{13} : \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} & I_{14} : \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix} \end{array} \tag{3.11}$$

Les quatre configurations successives, qui permettent l'exécution des macro-opérations de l'opérateur I$\triangle_2$, sont récapitulées ici même :

$$\begin{array}{ll} I_{21} : \begin{pmatrix} 0.5 & 0.5 \\ 0 & -1 \end{pmatrix} & I_{22} : \begin{pmatrix} 0 & 0.5 \\ -1 & 0.5 \end{pmatrix} \\ I_{23} : \begin{pmatrix} 0.5 & -1 \\ 0.5 & 0 \end{pmatrix} & I_{24} : \begin{pmatrix} -1 & 0 \\ 0.5 & 0.5 \end{pmatrix} \end{array} \tag{3.12}$$

A l'issue de ces quatre configurations, les résultats sont accumulés dans une mémoire externe (une SDRAM dans notre système) via un circuit programmable de type FPGA. Le FPGA n'effectue qu'une accumulation des

résultats dans une "matrice résultat". En fonction de la position du point courant dans le masque de base 2×2 (cf. figure 3.12) il faut, à chaque cycle rétinien, ajouter 1 ou -1 à l'adresse de ligne ou de colonne de la matrice résultat externe à la matrice de pixel.

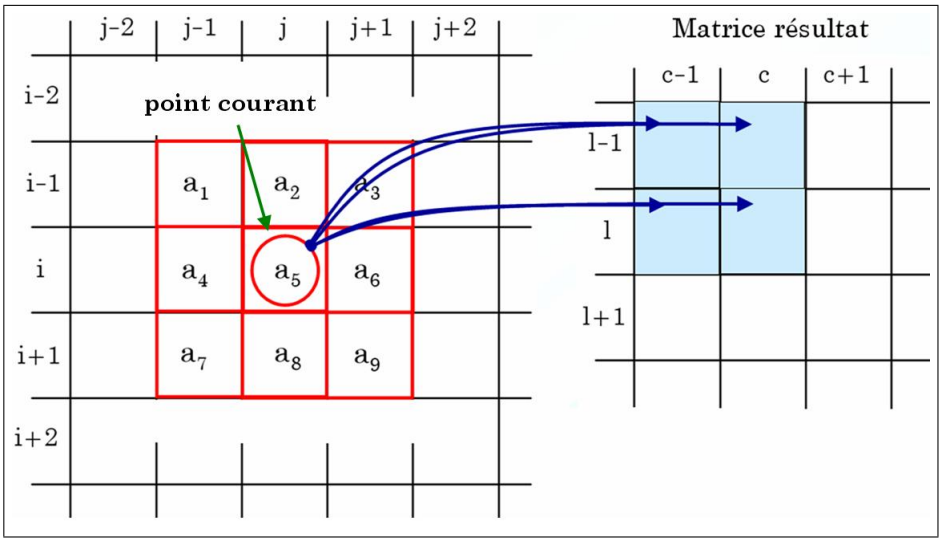


FIG. 3.12 – Position du point courant dans le masque de base 2×2 dans le cas d'une convolution de dimension 3×3

Par exemple, pour la macro-opération I_{11}, le point courant se situe en bas à droite du masque, pour la macro-opération I_{14}, le point courant se situe en haut à gauche du masque. Le point courant n'ayant pas la même position à chaque configuration des UA^2, il est nécessaire d'ajuster l'adressage de la matrice résultat vis à vis de l'adressage de la matrice de pixels. De plus, un facteur 2 est à appliquer aux résultats des accumulations des macro-opérations du deuxième Laplacien. L'algorithme à appliquer est donc le suivant :

1. Faire l'acquisition d'une image (100 μs).
2. Fixer les coefficients aux valeurs de I_{11} données par la relation 3.11,
3. Adresser séquentiellement toute l'image et lire les résultats issus des UA^2 (100 μs).
4. Accumuler les valeurs lues dans des mémoires statiques (pas de MAC), en veillant à ajuster l'adresse de la matrice résultat en fonction de la position du point courant.
5. Changer les coefficients par les quatre valeurs de I_{12}.

6. Adresser séquentiellement toute l'image et lire les résultats issus des UA^2 (100 μs).
7. Reprendre les étapes 3 et 4 avec les masques de base I_{13} et I_{14} (2×100 μs).

Avec cette technique on peut extraire des masques de convolution quelconques de dimension 3×3 en seulement 500 μs, c'est-à-dire à une cadence de 2000 images/s.

3.3.4 Généralisation

Le principal traitement réalisé sur le plan focal est une combinaison linéaire de quatre pixels connexes. Le pixel courant, associé à un certain coefficient ainsi que les trois pixels immédiatement voisins associés chacun à un coefficient, constituent les huit signaux d'entrée d'une Unité Arithmétique Analogique.

Cette UA^2 est implantée au sein de chaque pixel, ce qui représente, pour une rétine de résolution 64×64, une architecture de 16384 multiplieurs travaillant en parallèle. Il faut raisonnablement 100 μs, temps de conversion inclu, pour appliquer à l'image acquise un masque de taille 2×2 soit 16384 multiplications signées sur 8 bits en 100 μs (6 ns par multiplication). Extraire de simples gradients spatiaux ne représente qu'une petite partie des possibilités de notre rétine.

3.3.4.1 Filtres spatiaux

Dans les précédentes sections, nous nous sommes centrés sur l'application de masques de convolution simples 2×2 et 3×3. Dans le cas des masques de dimension 2×2, les coefficients sont configurés une seule fois. Dans le cas de masques de dimension 3×3, il faut plusieurs cycles rétiniens à l'issue desquels l'unité de calcul analogique est reconfigurée au moins deux fois en fonction des symétries de l'opérateur. Si le masque de dimension 3×3 est quelconque, il faut quatre cycles rétiniens (cf. figure 3.13) pour calculer la convolution.

A chaque cycle, le résultat de la convolution simple 2×2 est transmis au FPGA qui accumule le résultat avec le contenu d'un registre tampon lui même placé dans une matrice résultat. Cependant, les coefficients du masque à calculer ne sont pas rigoureusement les coefficients qui doivent être appliqués à la rétine. En effet, les accumulations se recouvrent deux fois sur les coefficients a_2, a_4, a_6, a_8 et quatre fois sur le coefficient a_5. Il faut donc modifier les coefficients appliqués à la rétine par rapport aux coefficients du masque

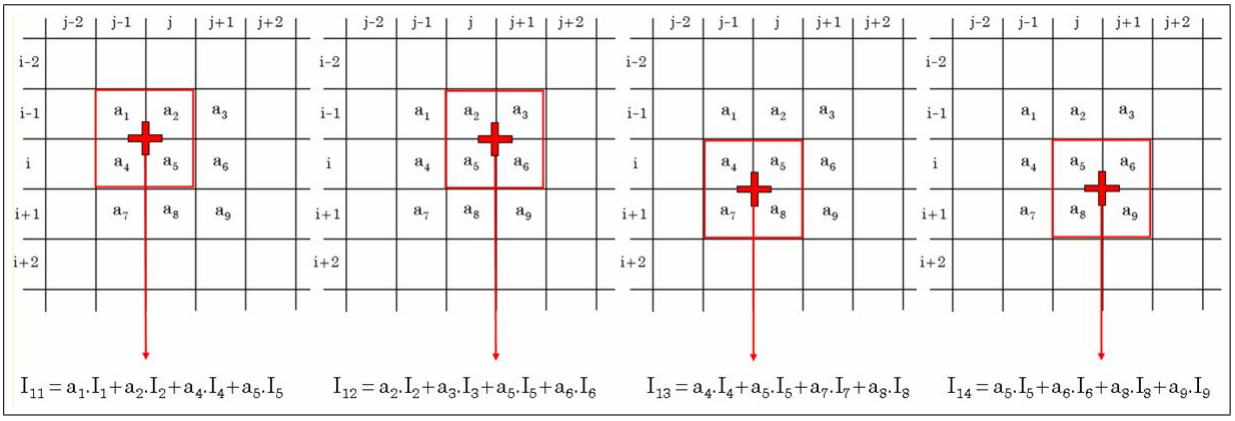


FIG. 3.13 – Calcul d'un masque quelconque de dimension 3×3

original, en appliquant un facteur 1/2 sur les coefficients a_2, a_4, a_6, a_8 et 1/4 sur le coefficient central a_5.

On peut très bien envisager l'application des masques de tailles plus importantes, sans augmenter par ailleurs les ressources externes à la rétine. L'expérience a montré que l'on n'est limité que par la durée de rétention d'information des mémoires analogiques. Ces mémoires maintiennent à l'entrée de l'unité de calcul analogique, l'information éclairement du pixel sous la forme d'un niveau de potentiel. Ainsi la valeur d'acquisition (signal analogique proportionnel à l'éclairement du pixel) est maintenue sur le port d'entrée du multiplieur pendant plus de 40 ms. Il est possible de changer la valeur des coefficients sans avoir à réaliser une nouvelle acquisition. Il est possible de reconfigurer 400 fois l'unité arithmétique (100 μs par cycle rétinien). Notre rétine permettrait d'appliquer un masque de convolution de dimension 20×20 avec un pas de un à la cadence vidéo (25 images par seconde, soit 40 ms par trame).

3.3.4.2 Application aux réseaux convolutifs

Une application de notre capteur serait l'implémentation du premier étage d'un réseau convolutif, celui de l'algorithme de détection de visage CFF(Convolutional Face Finder) [GD04]. Pour réaliser le premier étage de l'algorithme de détection de visage de type CFF, il est nécessaire d'appliquer 4 masques 6×6 avec un pas variable, constituant ainsi une pyramide à quatre niveaux. Le premier niveau utilise un pas de 2 alors que les trois suivants utilisent un pas 1,2 fois moindre. Au total, cela représente 16 convolutions spatiales de dimension 6×6.

Plus précisément, le premier étage consiste à accumuler $(R-4)^2/4$ valeurs, avec R la résolution du capteur. dans le cas d'une rétine 64×64, 900 valeurs sont donc accumulées entre chacune des reconfigurations. Neuf reconfigurations sont nécessaires pour appliquer le masque 6×6. Ces même masques sont appliqués avec un pas différent pour les trois autres niveaux.

La présence d'un convertisseur analogique numérique par colonne (voir chapitre suivant), permet d'accumuler les résultats d'une ligne entière de pixels en moins de 1,6 μs, soit l'équivalent de 256 MAC (Multiplication signée et Accumulation). En effet, il n'est plus nécessaire d'utiliser des multiplieurs arithmétiques numériques éliminant de fait l'intégration de multiplieurs matériels signés. Les ressources matérielles sont alors considérablement réduites. La seule tâche du processeur externe est d'accumuler les valeurs issues des unités arithmétiques analogiques. L'expression du temps total de calcul est donnée par la relation 3.13.

$$T_{total} = T_{acquisition} + \alpha \underbrace{\left(\frac{R-4}{2}\right)^2}_{1^{er}\ niveau} + \underbrace{\left(\frac{E\left(\frac{R}{1,2}\right)-4}{2}\right)^2}_{2^{e}\ niveau} + \underbrace{\left(\frac{E\left(\frac{E\left(\frac{R}{1,2}\right)}{1,2}\right)-4}{2}\right)^2}_{3^{e}\ niveau} + \underbrace{\left(\frac{E\left(\frac{E\left(\frac{E\left(\frac{R}{1,2}\right)}{1,2}\right)}{1,2}\right)-4}{2}\right)^2}_{4^{e}\ niveau} \tag{3.13}$$

Où $\alpha = (9 \times 4/R) \times 1,6\ \mu s$ et où E est la fonction partie entière. Pour une résolution de 64×64 et un temps d'intégration de 100 μs, le temps de calcul du premier étage du réseau convolutif avec cette méthode ne dépasse pas 4 ms, ce qui laisse 36 ms pour exécuter la suite de l'algorithme à la cadence vidéo. A titre de comparaison, la partie extraction de l'algorithme dure au moins 20 ms avec une architecture classique (capteur CMOS APS et processeurs). Les opérations de convolutions réalisées directement par la rétine représentent environ 80% du temps de calcul de la partie extraction de l'algorithme CFF. Pour résumer, l'algorithme à appliquer est le suivant :

1. Faire l'acquisition d'une image.
2. Fixer les coefficients aux quatre premières valeurs du masque 6x6.

3. Adresser séquentiellement toute l'image et lire les résultats issus des UA^2.
4. Accumuler les valeurs lues dans des mémoires statiques (pas de MAC), il y aura $(R\text{-}4)^2/4$ registres de 16 bits.
5. Changer les coefficients par les quatre valeurs suivantes.
6. Réitérer les opérations 3 et 4...

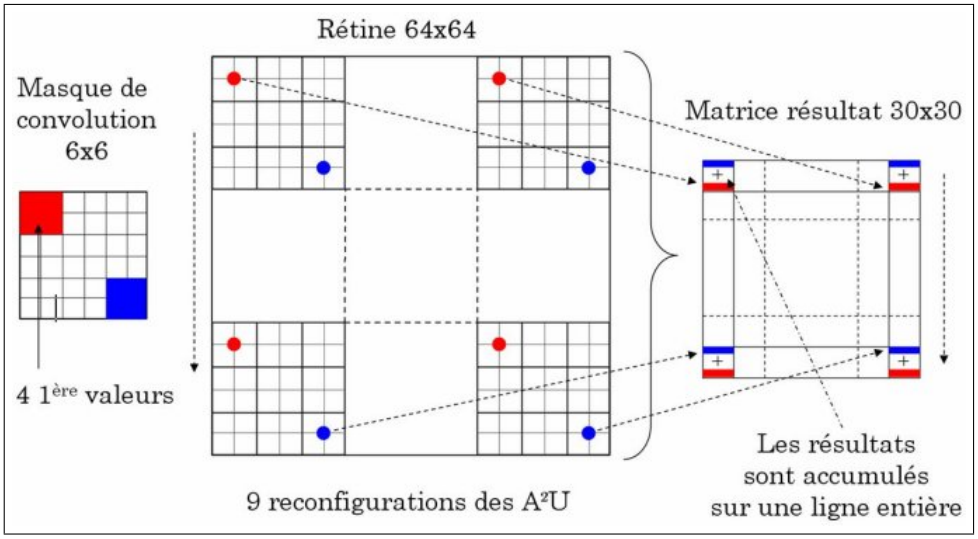


FIG. 3.14 – Stratégies de configuration appliquées au premier étage d'un réseau convolutif

La figure 3.14 montre comment il est possible de réaliser ce que l'on appellerait le premier étage d'un réseau convolutif, dédié à l'exécution de l'algorithme CFF sur HISIC. A gauche sur la figure, on trouve le masque de convolution avec ses quatre premières et ses quatre dernières valeurs (au sens où ces valeurs sont accumulées en dernier). Lorsque les coefficients sont configurés aux quatre premières valeurs du masque, les résultats issus des UA^2 sont lus et accumulés dans une matrice résultat (à droite sur la figure) de résolution 30×30, dans le cas où le pas est de 2 ((64-4)/2=30).

A moyen terme, l'objectif serait d'appliquer ce concept sur des capteurs de haute résolution en parcellant le capteur HDV en plusieurs macro-rétines de résolution 64×64 (cf. figure 3.15). Les perspectives de conception d'un capteur HDV de résolution 1440×1080 sont séduisantes. Le premier problème auquel on doit faire face lors de la conception d'un capteur rapide haute résolution est la gestion du flot de données. La solution la plus simple est de paralléliser, en ajoutant un niveau supplémentaire à la hiérarchie du capteur, entre le niveau pixel et le niveau colonne, le niveau macro-rétine.

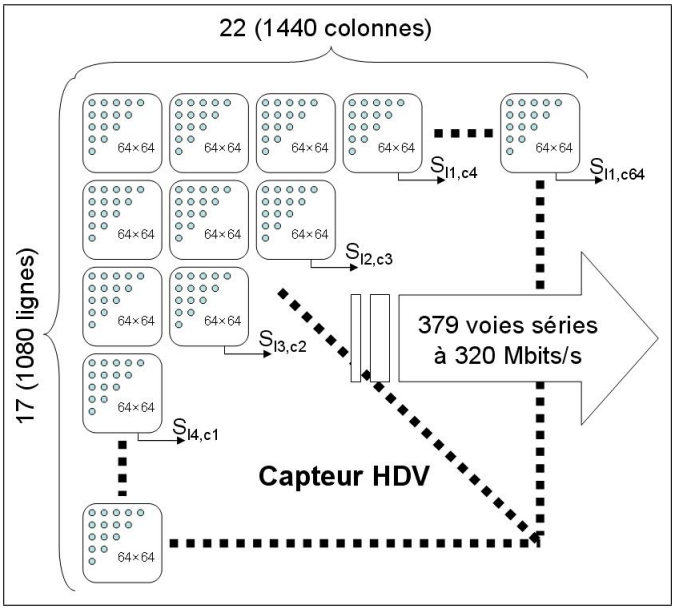


FIG. 3.15 – Synoptique d'un capteur HDV rapide et intelligent

Ainsi le capteur HDV serait composé de 22×17 macro-rétines. Toutes ces macro-rétines de résolution 64×64 seraient adressées en lignes et en colonnes par le même bus à 40 mégapixels par seconde. De cette façon, il n'y aurait que 379 sorties série à 320 mégabits par seconde (40 Mpixel/s×8 bits) à gérer, ce qui est tout à fait envisageable avec par exemple un ou plusieurs FPGA Virtex-5 XC5VLX50 de 220 à 560 ports entrées/sorties, dont la fréquence d'horloge principale peut atteindre 550 MHz.

3.3.5 Détection de mouvement

La théorie sur la morphologie temporelle a trouvé une application notamment dans la détection de mouvement. Cette partie du chapitre se propose d'appliquer un cas particulier des techniques qui découlent de cette théorie. Une moyenne récursive classique permet d'apprécier le mouvement d'un mobile [Ric06] avec une certaine robustesse et cela malgré la simplicité de l'algorithme spatio-temporel mis en œuvre.

Sur une séquence d'images, on calcule récursivement l'érosion, la dilatation et le gradient morphologique temporel. Le tableau 3.1 décrit l'algorithme utilisé pour ce détecteur de mouvement utilisant la moyenne récursive classique. L'UA^2 est utilisée dans ce cas précis pour calculer l'érosion, la dilatation et le gradient morphologique. Cependant, les com-

TAB. 3.1 – Algorithme morphologique spatio-temporel

Initialisation trame 0 :	
Pour chaque pixel :	$\begin{cases} DIL_0 <= PIX_0; \\ ERO_0 <= PIX_0; \end{cases}$
Pour chaque trame i :	
Dilatation :	$\begin{cases} \text{si } (DIL_{i-1} > PIX_i) \text{ alors} \\ DIL_i <= (PIX_i + DIL_{i-1})/2; \\ \text{sinon} \\ DIL_i <= PIX_i; \\ \text{fin si ;} \end{cases}$
Erosion :	$\begin{cases} \text{si } (ERO_{i-1} < PIX_i) \text{ alors} \\ ERO_i <= (PIX_i + ERO_{i-1})/2; \\ \text{sinon} \\ ERO_i <= PIX_i; \\ \text{fin si ;} \end{cases}$
Gradient morphologique :	{ Grad_i <= DIL_i - ERO_i ;

paraisons sont réalisées à l'aide du circuit FPGA. Les cellules MA^{2}M sont mises à contribution pour stocker les résultats (érosion et dilatation) de la trame précédente. Dans le dernier chapitre, nous verrons le résultat de l'implémentation de cet algorithme et une application à la détection du mouvement des parties mobiles lors du déploiement d'une tapette à souris.

3.4 Conclusion

Dans ce chapitre, nous avons vu dans l'ensemble ce qu'il est possible de faire avec notre rétine. Une UA^2 implémentée au sein du pixel peut réaliser une combinaison linéaire de quatre pixels connexes. Deux cellules MA^2M permettent, soit de faire de l'imagerie à haut cadence en dissociant phases d'acquisition et de lecture, soit d'exécuter des algorithmes récursifs temporels comme l'érosion et la dilatation temporelles présentées ci-avant. Il est donc possible en théorie de réaliser des traitements, à haute cadence (acquisition=100μs + 100μs/cycle rétinien), de type spatiaux (gradients, convolution) et spatio-temporels (gradient morphologique). Dans le chapitre suivant, nous détaillons la conception du pixel et du convertisseur analogique numérique implémenté dans la dernière version d'HISIC. Des choix technologiques aux limites physiques de notre architecture, nous verrons comment nous avons conçu et développé le photodétecteur, les structures de calcul (UA^2) et les cellules MA^2M.

Chapitre 4

Conception microélectronique d'HISIC

4.1 Introduction

DAns le chapitre précédent nous avons passé en revue les possibilités algorithmiques offertes par notre rétine. Ce nouveau chapitre est consacré à la conception de notre circuit. Nous utilisons la technologie CMOS pour pouvoir implémenter des fonctions électroniques particulières au sein même du pixel d'un capteur d'image. La technologie standard CMOS AMS 0,35 μm est beaucoup employée en France dans la recherche universitaire en microélectronique. Les différents projets des laboratoires sont regroupés et les dessins des masques centralisés au CMP (Circuits Multi-Projets) à Grenoble afin d'être gravés sur un même Wafer. Dans ce contexte universitaire, le coût de cette technologie ($\approx 650\, euro/mm^2$) est l'une des raisons de notre choix.

Dans le contexte industriel actuel, nous sommes contraints de limiter la taille du pixel conçu et cela quelque soit l'application visée (appareil photo numérique, téléphonie mobile, caméras rapides industrielles, etc.). Même si nous concevons un capteur dans une technologie sub-micronique (standard CMOS 0,35 μm), l'interpolation, "sur le papier", vers des technologies fortement sub-microniques (90, 45 nm) doit être faite avec beaucoup de précautions, d'autant plus que la spécificité des imageurs CMOS entraîne des problèmes liés à la diaphonie optique avec l'évolution des technologies. En effet, les pixels des capteurs d'images CMOS ne comportent par de séparations optiques latérales opaques, les parois latérales sont relativement transparentes et les rayons de lumière obliques peuvent balayer des pixels contigus.

Dans le but d'envisager une application industrielle de nos travaux, nous

devons fixer un dimensionnement de notre pixel. Compte tenu de l'évolution de la technologie, nous avons vu dans le premier chapitre que nous ne pouvions pas directement interpoler des travaux de conception vers une technologie fortement submicronique (transistors analogiques : 18 nm en 2013 et 13 nm en 2016) sans prendre des précautions, vu l'importance du courant tunnel sur le courant de saturation du transistor. En appliquant une marge de sécurité, on peut néanmoins estimer la surface du pixel dans ces technologies de demain. Nous nous somme fixées une taille de pixel de 35 μm de côté dans une technologie standard CMOS 0,35 μm, soit une diagonale de capteur HDV (1440×1080) probablement inférieure au centimètre dès 2013 en production de masse (taille de pixel inférieure à 4 μm en technologie CMOS analogique 18 nm).

HISIC intègre des traitements programmables au niveau même du pixel. C'est une machine massivement parallèle de 4096 processeurs analogiques arithmétiques interconnectés selon une grille 64×64 en topologie 4-connexe. Nous avons fait le choix de concevoir des traitements d'images de bas niveaux dans le plan focal (au plus près du pixel) à l'aide de fonctions exclusivement analogiques et cela pour deux raisons.

1. Les processeurs analogiques comportent relativement peu de transistors par rapport aux processeurs numériques [EGYF99] (cf. figure 4.1).
2. Le contrôle de la rétine ainsi créée est beaucoup plus simple que celui des rétines purement numériques (opcode, gestion des registres, etc.).

Ainsi nos recherches s'articulent autour de la conception de capteurs d'images CMOS rapides et intégrant des traitements d'images linéaires dans le plan focal et massivement parallèles. Dans ce chapitre, nous présentons l'architecture du circuit HISIC et les choix technologiques que nous avons faits pour mener à bien cette étude théorique. Les caractéristiques principales de l'architecture du pixel sont décrites dans la section 4.2 où une discussion sur la forme du photodétecteur débouche sur une simulation numérique. Puis, les détails sur les mémoires analogiques et l'unité de calcul y sont successivement décrits. La section 4.3 s'articule autour de la conception du convertisseur analogique numérique. Enfin, la section 4.4 présente l'architecture de la partie adressage du capteur.

4.2 Conception du pixel

La conception d'une rétine artificielle commence par une réflexion sur le pixel ou photosite. Le photo-détecteur, l'étage d'acquistion (transduction

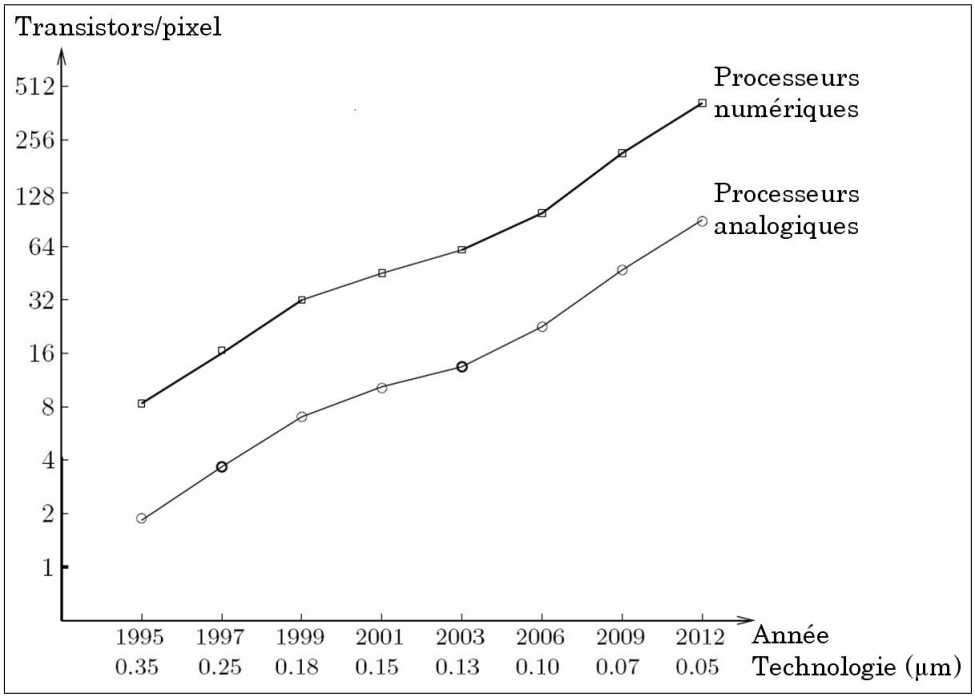


FIG. 4.1 – Courbe d'évolution du nombre de transistors par pixel dans les architectures distribuées dans le cas de processeurs analogiques et de processeurs numériques de 1995 à 2012, pour des pixels de 5 μm de côté et un facteur de remplissage de 30%

éclairement/signal), et les traitements implémentés, sont les éléments constitutifs du pixel qui sont successivement détaillés dans cette section.

4.2.1 Photo-détecteur

Le photo-détecteur assure la fonction de transduction éclairement - tension. Son choix est déterminant, autant en termes de densité d'intégration qu'en terme d'efficacité et de réponse spectrale.

4.2.1.1 Choix technologique

La photodiode verticale de type N est le choix technologique retenu en terme d'efficacité et de densité d'intégration. C'est à ce titre qu'elle est la plus populaire parmi les photodétecteurs [EGE05] en technologie CMOS utilisant le procédé standard de fabrication. Dans un procédé dédié, les choix technologiques sont plus variés mais le coût de fabrication est bien plus important.

Dans ce mémoire, nous ne nous intéresserons qu'au procédé standard CMOS AMS (AutriaMicroSystem) 0,35 μm (double-poly, quadruple-métal), dont la figure 4.2 donne la vue en coupe.

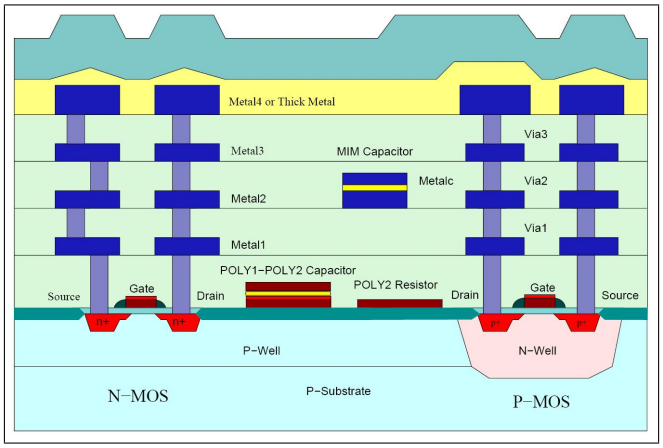


FIG. 4.2 – Vue en coupe de la technologie de fabrication CMOS AMS 0,35 μm (double-poly, quadruple-métal)

Les photodiodes verticales de type N sont constituées d'une diode à jonction PN, soit une diffusion fortement dopée d'atomes donneurs (cf. annexe A) dans un substrat faiblement dopé en accepteur (N_{diff}/P_{sub}). En technologie CMOS 0,35 μm, les profondeurs de jonctions sont de 200 nm, ce qui est tout à fait adapté pour les applications dans le visible [400 nm ; 1100 nm]. Ces photodiodes ont un meilleur rendement que les photodiodes de type P (diffusion P dans un puits N, P_{diff}/N_{well}) car le puits N, qui les constitue, est davantage dopé que le substrat P constituant l'anode des photodiodes de type N. De plus, les contraintes de dessin sont nettement moins importantes [Gou95]. En effet la présence d'un puits N impose des distances minimales conséquentes pour la conception du dessin des masques du pixel. Par exemple, entre un caisson N et une diffusion N, il faut respecter une distance minimale de 2,6 μm, d'après les données constructeurs (AustriaMicroSystem) en technologie 0,35 μm. Nous avons donc fait le choix d'implémenter des photodiodes verticales de type N dans notre photosite.

4.2.1.2 Réflexion sur la géométrie de la photodiode

Les contraintes spatiales d'interconnexion avec les pixels voisins ont une influence sur la géométrie du photo-détecteur, qui a elle-même une influence sur les performances de la photodiode. Le second travail est donc de trouver un compromis entre les contraintes liées à la connectique et les performances

induites par la géométrie du photo-détecteur. De plus, la géométrie choisie doit contribuer à satisfaire au mieux les contraintes liées aux interconnexions imposées par l'implémentation d'une unité arithmétique au sein même du pixel. C'est pourquoi, une structure particulière de photosite (cf. section 3.3, figure 3.6(b)) a été choisie. Le principal avantage de cette topologie est la minimisation des longueurs de métal d'interconnexions entre pixels adjacents avec l'unité de calcul, contribuant à un meilleur taux de remplissage et à un temps d'accès plus court en raison notamment des capacités de couplage minimisées.

Afin d'atteindre des performances intéressantes en terme de fréquence d'acquisition, un des éléments clés est le photo-capteur qui doit être précieusement conçu et optimisé. Nous avons choisi une de photodiode de type N pour faciliter la conception du pixel. Cette structure est la plus répandue dans les capteurs APS. De plus, les photodiodes de type N ont un très bon rendement quantique et sont particulièrement intéressantes en terme de densité d'intégration. Les paramètres critiques qui interviennent dans la conception d'une photodiode sont le courant d'obscurité et l'efficacité quantique ou la réponse spectrale [WSL+04]. La structure et la géométrie de la photodiode ont une influence déterminante sur les performances de l'imageur [SBYP02, SYP03].

Les travaux existants dans ce domaine convergent tous. Zhang et Chan démontrent en 1998 [ZC98] que les photodiodes verticales de type N, dans un procédé de fabrication standard CMOS, s'avèrent être les plus efficaces comparées aux autres types de photodiodes. Lee et Hornsey proposent en juin 2001 [LH01] une étude sur la réduction des capacités latérales de jonction de photodiodes de type N en vue d'augmenter le gain de conversion sur la base de résultats expérimentaux en technologie standard CMOS 0,35 μm. Shcherback présente en juin 2002 [SBYP02] une étude de l'influence de la géométrie d'une photodiode de type N sur son courant d'obscurité. Cette étude a permis d'établir un modèle mathématique du courant d'obscurité, basé sur des résultats expérimentaux, obtenus à partir d'un capteur APS de résolution 256×256 en technologie standard CMOS 0,5 μm. Shcherback et Yadid-Pecht présentent en janvier 2003 [SYP03] un modèle semi-analytique de photo-réponse basé sur le même capteur CMOS. Chung-Yu Wu présente en février 2004 [WSL+04] une étude comparative sur deux nouvelles structures de photodiodes présentant un faible courant d'obscurité (intégration présentant une pente de 30,6 mV/s avec une polarisation inverse de 2V) et une meilleure réponse spectrale, dans une technologie CMOS 0,35 μm mais non standard (Opto Process). Le procédé de fabrication a été modifié en introduisant des diffusions spécifiques pour la photo-transduction.

Dans notre circuit, au niveau du pixel, la photodiode est une structure

de type N, *i.e.*, une diffusion fortement dopée d'impuretés donneurs (Nd = $212\times10^{15}/\text{cm}^3$) sur un substrat de type P. La photodiode est polarisée en inverse (phase de pré-chargement voir section 4.2.2.2). Il se forme alors une zone de déplétion en porteurs libres (ou zone de charge d'espace), plus épaisse aux bords de la cathode qu'au centre de la photodiode. En effet, les photons incidents à la cathode dissocient plus de paires électron-trous là où la zone de déplétion est plus épaisse [LH01]. C'est pourquoi il n'est pas forcément optimal de choisir une forme carrée pour la réalisation du layout de la photodiode. Nous avons analysé et testé trois formes de photodiodes. La photodiode carrée, classiquement utilisée dans la littérature, la forme en croix qui serait optimale pour le tracé de l'unité de calcul, dans le but de faciliter les inter-connexions aux pixels voisins, et la forme quasi octogonale basée sur des géométries à 45° donc intégrable dans un procédé standard CMOS.

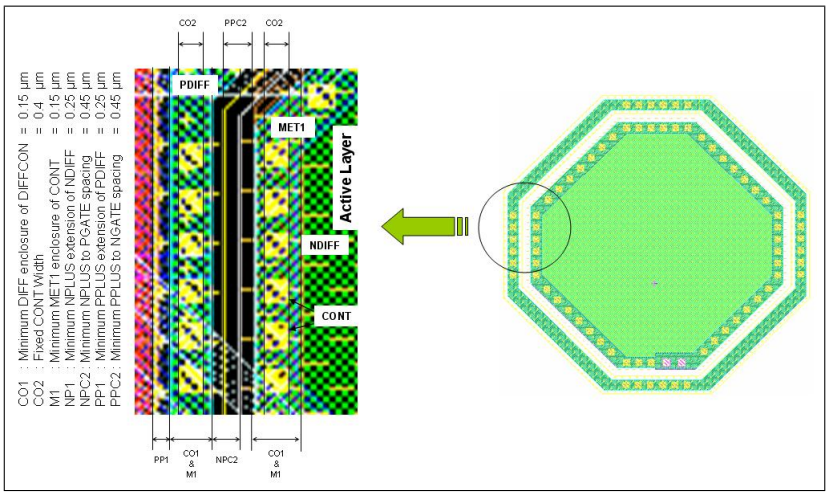


FIG. 4.3 – Règles de dessin appliquées à une photodiode de type N

Les règles de dessin, dans le procédé de fabrication CMOS AMS (AustriaMicroSystems) 0,35 μm, imposent une certaine largeur de diffusions N^+ et P^+ et de métal1, destinée exclusivement aux connexions de l'anode et de la cathode de la photodiode et ne contribuant pas au photo-courant. En effet, la couche de métal 1 est opaque à la lumière. Une grandeur, directement liée au procédé de fabrication, est arbitrairement notée β. Elle contribue dans une large mesure au taux de remplissage du pixel. Il faut donc y apporter une importance toute particulière. L'expression de la grandeur β est simple. Il s'agit de la somme des différentes contraintes de fabrication, elle est donnée par la relation 4.1, avec PP1, CO1, CO2 et PPC2 les règles de dessin reportées sur la figure 4.3.

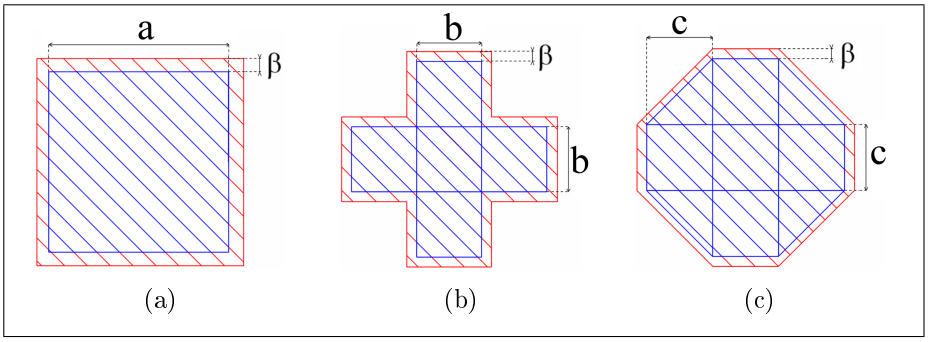


FIG. 4.4 – (a) forme carrée, (b) forme en croix, (c) forme octogonale

$$\beta = 2 \times PP1 + 4 \times CO1 + 2 \times CO2 + PPC2 = 2,35\ \mu m \tag{4.1}$$

Il s'agit bien ici de la valeur minimale de β, imposée par les contraintes de fabrication. Ainsi, la surface totale occupée par une photodiode est la somme d'une surface dite "active" (Sac), qui participe à la génération du photocourant, et d'une surface dite "passive" (Spa), qui ne sert qu'à connecter le dipôle à son électronique de contrôle. La figure 4.4 illustre les différentes formes probables de photodiodes. Nous aurions pu y inclure le disque mais une telle photodiode ne pourrait être réalisable par le fabricant de circuit. En effet, les fabricants imposent que les géométriques des diffusions soient orthogonales, voire à 45° ou à 135°. Néanmoins, pour l'instant en production de masse, le procédé de fabrication est quelque peu ralenti lors du calcul de diffusion à 45° ou à 135°. Pour chacune de ces trois formes, on peut distinguer la surface intérieure active (Sac) et la surface périmétrique passive (Spa). Notons "a" le coté intérieur de la forme carrée, "b", le paramètre géométrique principal de la forme en croix, et "c", le paramètre géométrique de la forme octogonale. On notera P, le périmètre intérieur de la forme géométrique. Sur la base de ces paramètres, on peut établir formellement les expressions des surfaces actives et passives pour les différentes formes géométriques proposées.

$$\text{structure carrée : } \begin{cases} Sac = a^2 \\ Spa \cong P \times \beta = 4\beta a \end{cases} \tag{4.2}$$

$$\text{structure en croix : } \begin{cases} Sac = 5b^2,\ b = \frac{a}{\sqrt{5}} \\ Spa \cong \frac{3}{\sqrt{5}}(4\beta a) \approx 1,342 \times (4\beta a) \end{cases} \tag{4.3}$$

$$\text{structure octogonale : } \begin{cases} Sac = 7c^2,\ c = \frac{a}{\sqrt{7}} \\ Spa \cong \frac{(1+\sqrt{2})}{\sqrt{7}}(4\beta a) \approx 0,912 \times (4\beta a) \end{cases} \tag{4.4}$$

De ces expressions formelles, nous pouvons déduire les graphes des surfaces passives des différentes formes géométriques relativement à la longueur du côté intérieur de la forme carrée (cf. figure 4.5).

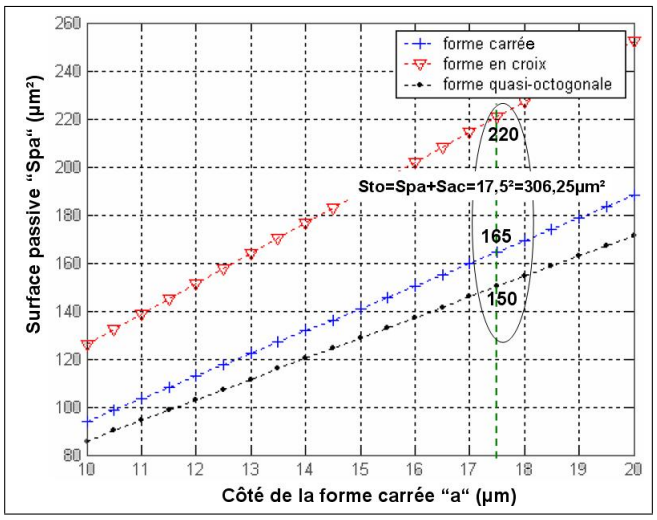


FIG. 4.5 – Graphe de la surface passive Spa pour les trois formes géométriques en fonction du paramètre a, la longueur du côté intérieur de la forme carrée

La conception du pixel est contrainte. D'une part, sa surface ne doit pas être trop importante, dans la technologie considérée, pour que le pixel puisse avoir un avenir industriel par interpolation vers des technologies fortement sub-microniques. D'autre part, le taux de remplissage doit être suffisant afin d'obtenir un rapport signal sur bruit satisfaisant pour de faibles temps d'intégrations.

Nous avons fait le choix de travailler en technologie standard CMOS 0,35 μm pour des raisons de coût de prototypage (650 euros/mm^2). La surface du pixel est fixée à 35×35=1225 μm^2. Les pixels de HISIC.I ont un taux de remplissage de 12,7% (*i.e.*, le rapport entre la surface active de la photodiode et la surface du pixel est de 12,7%). Avec une taille de pixel de 35×35 μm^2, la surface totale de la photodiode ne doit pas excéder $Sto = Spa + Sac = 306.25\ \mu m^2$, ce qui donne le paramètre géométrique de référence $a = 17.5\ \mu m$. D'après la figure 4.5, nous constatons que :

1. Une photodiode en forme de croix n'est pas intéressante en terme de densité d'intégration. En effet la surface "passive" représente plus des 2/3 de la surface "active".
2. Les photodiodes de forme carrée et octogonale ont leur surface passive du même ordre de grandeur (respectivement $165\mu m^2$ et $150\mu m^2$) tout en étant très inférieures à la surface passive de la forme en croix.

Pour conclure sommairement, une forme en croix est à exclure et une forme octogonale est intéressante en termes de densité d'intégration. Cette forme particulière est propice à la topologie 4-connexe (cf. figure 4.6). Dans cette topologie, les processeurs élémentaires sont interconnectés pour réaliser des traitements spatiaux locaux, comme le précise la section 3.3, dans le chapitre 3.

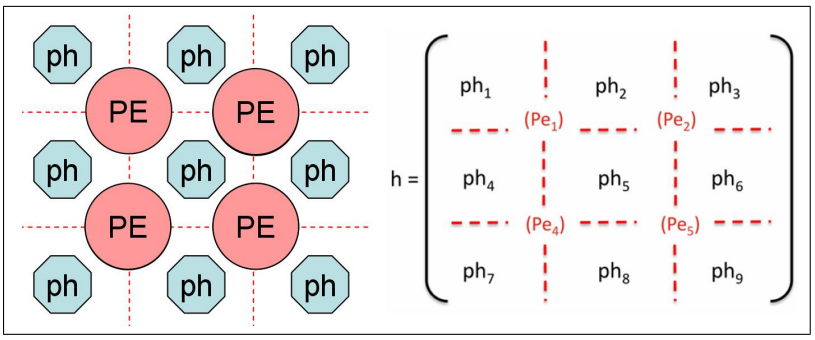


FIG. 4.6 – Topologie 4-connexe avec une photodiode octogonale

La figure 4.7 illustre l'arrangement des pixels et son repère orthonormé pour l'évaluation du gradient spatial, comme décrit précédemment dans ce mémoire (cf. chapitre 3, section 3.3.1), dans le cas de pixels octogonaux.

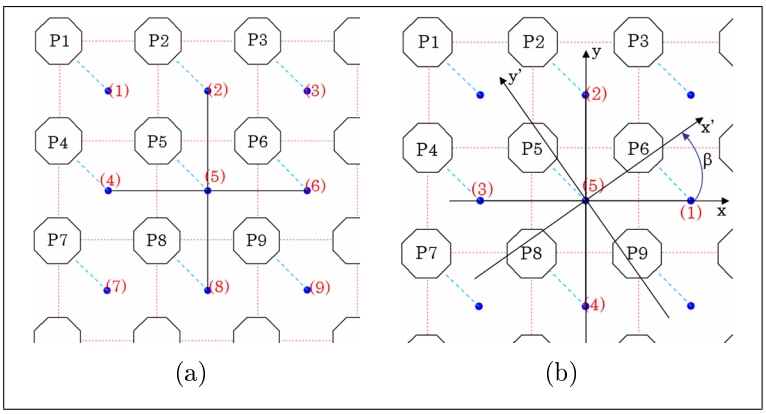


FIG. 4.7 – (a) Matrice de pixels basée sur une géométrie octogonale, (b) repère orthonormé pour l'évaluation du gradient spatial

4.2.1.3 Modèle théorique déduit des différences finies

La figure 4.8 illustre en trois dimensions la gravure d'une photodiode verticale de type N dans le procédé standard de fabrication CMOS. On peut

y voir la cathode (diffusion N^+, en vert), polarisée au potentiel $+V_{DD}$ et l'anode (diffusion P^+, en rouge), connectée au substrat (en orange) et au potentiel 0V. A l'interstice entre la cathode et le substrat, le champ électrique vertical crée une zone de charge d'espace (ZCE) majoritairement dans le substrat faiblement dopé. En regardant cette figure, on comprend bien que la différence de potentiel entre l'anode et la cathode est plus importante aux bords de la photodiode (là où sont physiquement connectées anode et cathode) qu'au centre.

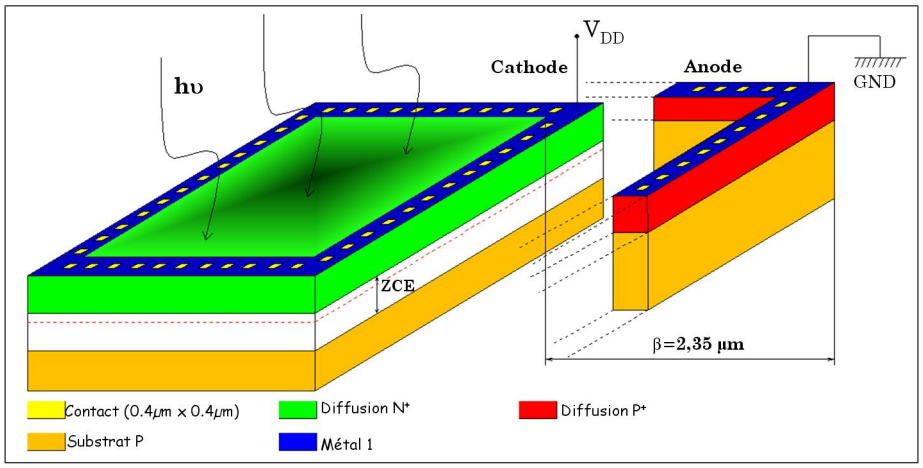


FIG. 4.8 – Vue en trois dimensions d'une photodiode verticale de type N

Dans une photodiode verticale de type N, polarisée en inverse, la zone de charge d'espace ne présente pas la même épaisseur en tout point de la surface de la cathode de la photodiode. En effet, cette zone de déplétion a tendance à être plus large aux bords, là où le potentiel de la cathode est proche de V_{DD}.

En utilisant la méthode des différences finies, notamment la résolution numérique du Laplacien et en définissant convenablement les conditions aux bords de la structure ($+V_{DD}$), on obtient des résultats qui corroborent les précédentes hypothèses. Outre l'effet photo-électrique dans le silicium et les phénomènes de diffusion en dehors de la ZCE (cf. chapitre 2, section 2.3.2), Le potentiel électrostatique dans la couche épitaxial a une distribution poissonienne. En appliquant la loi de Poisson (problème de Dirichlet $\Delta V=0$, dans la zone de charge d'espace) et en supposant le courant photonique constant et donc le régime établi, la résolution numérique du Laplacien permet d'apprécier l'épaisseur de la zone de charge d'espace dans une photodiode verticale de type N. Le Laplacien à deux dimensions $\partial^2/\partial x + \partial^2/\partial y$ est approché sur une grille de pas h_x et h_y par :

$$
\begin{aligned}
(\Delta_h u)_{i,j} &= h^{-2} \begin{pmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{pmatrix} u_{i,j} \\
&= \tfrac{1}{h_x^2} \left\{ u_{i+1,j} - 2u_{i,j} + u_{i-1,j} \right\} + \tfrac{1}{h_y^2} \left\{ u_{i,j+1} - 2u_{i,j} + u_{i,j-1} \right\}
\end{aligned}
\tag{4.5}
$$

Les différences finies montrent que le potentiel est plus élevé là où la diffusion N^+ est polarisée. La largeur de la zone de charge d'espace est fonction de la concentration de dopage, de la tension de polarisation inverse appliquée à la photodiode. Son expression est, en première approximation, proportionnelle à la racine carrée du potentiel cathodique. La figure 4.9 est une image de l'épaisseur normalisée de la zone de charge d'espace dans la photodiode étudiée.

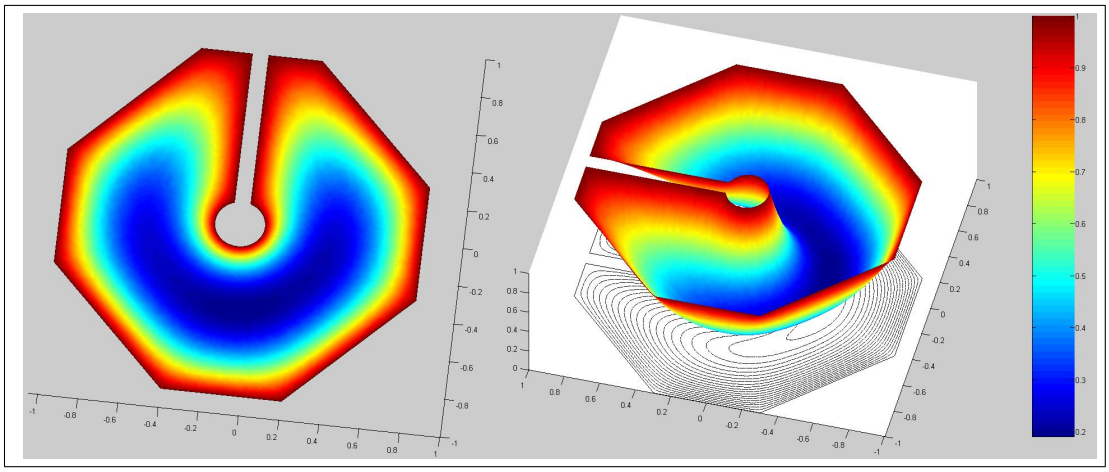


FIG. 4.9 – Distribution du potentiel électrostatique dans une photodiode verticale de type N

Cette simulation confirme une étude de Shcherback [SYP03] en 2003 sur l'influence de la géométrie de la photodiode sur le photo-courant. L'efficacité de la photodiode dépend effectivement du nombre de sommets et des angles aux sommets de la géométrie. Sur la figure 4.9, les couleurs chaudes (au voisinage des bords de la géométrie) représentent une densité de photo-porteurs plus importante aux bords et qui s'étend sur les sommets de la diffusion. Ces résultats montrent que l'efficacité d'une photodiode verticale de type N augmente avec le nombre de sommets de sa géométrie, des angles au sommets et de la polarisation de sa cathode.

En technologie standard CMOS, seuls des angles à 45°, 90° et 135° sont possibles. On ne pourrait donc pas imaginer se rapprocher de la forme circulaire en traçant par exemple une photodiode en forme de dodécagone avec des angles aux sommets de 150°, ce qui rendrait en théorie la photodiode plus efficace.

Ces premiers résultats ont été validés expérimentalement à partir de structures de test implémentées dans le premier circuit HISIC.I (voir chapitre 5) et d'autres simulations numériques ont permis de pousser encore plus loin les concepts de la polarisation cathodique d'une photodiode verticale de type N. Avec le calcul par différences finies, il est possible de simuler toutes sortes de polarisations de la cathode. Par exemple, il est possible d'étudier les performances de la photodiode avec des segments de bandes de métal qui divisent la surface active en plusieurs secteurs.

Pour conclure sommairement sur ce principe, une photodiode verticale de type N de forme octogonale, avec une polarisation de sa cathode au centre via un segment de métal, possède principalement trois avantages :

1. La surface dédiée à la polarisation (dite passive "Spa") est 10% inférieure à la surface passive d'une photodiode de forme carrée, ce qui est intéressant en terme de densité d'intégration.
2. La forme octogonale est basée sur des angles à 45°, donc intégrable dans un procédé de fabrication standard.
3. La géométrie octogonale possède 2 fois plus de sommets que la géométrie carrée et des angles au sommet 50% supérieurs, ce qui la rend plus efficace [SYP03].

La section qui suit présente l'étude et la conception des structures électroniques intégrées dans le pixel.

4.2.2 Électronique de lecture et de traitement au sein du pixel

Au plus près du photo-détecteur, se trouve l'électronique de lecture et de traitement qui se compose des blocs suivants :

1. L'étage de mise en forme qui réalise la transduction éclairement tension.
2. Les traitements qui sont exclusivement conçus à l'aide de fonctions analogiques.

En effet, nous avons décidé d'exclure, de notre étude, l'implantation de traitements numériques, qui exigent l'implémentation d'un CAN et d'un pro-

cesseur arithmétique booléen au sein du pixel. Le nombre de transistors, nécessaires pour réaliser cette implémentation, serait conséquent et le contrôle numérique global de la rétine serait complexe. Les traitements analogiques demandent moins de ressources matérielles et le contrôle global du capteur est beaucoup plus simple. Néanmoins, les structures analogiques doivent être soigneusement caractérisées pour limiter les sources d'erreurs.

L'étage de mise en forme assure la transduction éclairement - tension au plus près du photodétecteur. Son principe et son architecture sont décrits dans la section qui suit. Deux structures dites MA^2M (Mémoire Analogique Amplificateur et Multiplexeur) permettent de mémoriser et amplifier le signal issu du premier étage. Ces structures sont décrites dans la section 4.2.2.2. Puis le section 4.2.2.3 décrit la conception d'une unité arithmétique analogique reconfigurable utilisée pour exécuter des traitements spatiaux programmables.

4.2.2.1 Transduction éclairement/signal

La transduction éclairement/signal est le premier étage du pixel. Dans le second chapitre, nous avons fait état de l'art des différentes architectures de pixels. Nous nous plaçons dans la catégorie des pixels à intégration ou APS [EGE05] (cf. chapitre 2, section 2.4.2) où le signal utile, sous forme de potentiel, est lu à l'issue de l'intégration du photocourant. Dans la conception d'un pixel de type APS, il existe deux types d'interrupteurs commandés pour polariser une photodiode en inverse à une valeur V_d. Dans le cas de l'utilisation d'un transistor PMOS, ce dernier, monté en source commune, est commandé par le signal "reset" en logique négative. Le PMOS passe en régime de forte inversion, il présente donc une faible tension de déchet et la photodiode est polarisée avec une tension inverse Vd maximale. Dans l'autre cas, un transistor NMOS, monté en drain commun, est piloté par le signal "reset" en logique positive. Le NMOS, intéressant en terme de densité d'intégration, passe en régime d'inversion modéré ($\approx\ V_{THN} + 10\,mV$) et présente une tension de déchet de l'ordre du Volt (en technologie CMOS 0,35 μm, 3,3 V). Nous avons fait le choix d'implémenter un transistor PMOS, afin d'augmenter la dynamique de ce premier étage (cf. figure 4.10) et en conséquence d'augmenter le rapport signal sur bruit. De plus, la faible tension de déchet contribue à augmenter la tension de polarisation inverse, donc à diminuer la capacité parasite (cf : equation 2.1, chapitre 1). L'augmentation de la dynamique, conjuguée à la diminution de la capacité parasite, contribue à augmenter le rapport signal sur bruit (cf : equation 2.26, chapitre 1).

Le potentiel cathodique, issu de l'intégration du photocourant, est un signal analogique qui nous informe sur le niveau d'éclairement du pixel. Il

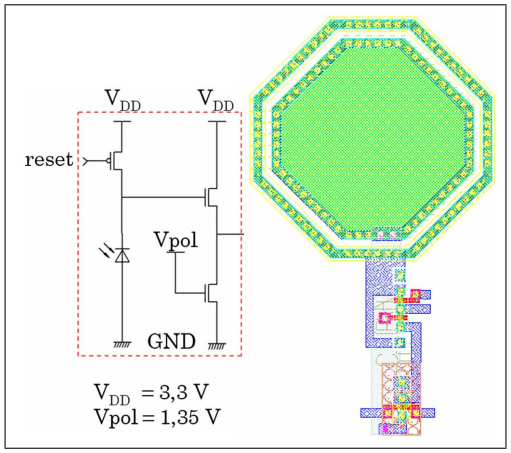


FIG. 4.10 – Schéma électronique du premier étage et son dessin des masques (à droite)

est donc important de se donner la possibilité d'accéder à cette grandeur sans perturber ses variations ou sa valeur. C'est pourquoi un étage suiveur est requis pour sa grande impédance d'entrée et un gain voisin de l'unité. Cet étage suiveur est un simple transistor NMOS aux dimensions minimales monté en drain commun. Les variations du potentiel cathodique sont relativement faibles surtout pour de faibles temps d'intégration, c'est pourquoi l'amplification de ce signal est réalisée au cœur même du pixel. Il est donc nécessaire de polariser les variations du potentiel cathodique autour du point statique de fonctionnement de l'amplificateur placé en aval de ce premier étage. Ce point statique est $V_{DD}/2$, *i.e.*, la moitié de la tension d'alimentation. Pour polariser le potentiel cathodique autour de $V_{DD}/2$, un transistor NMOS monté en source commune est intercalé entre la source du suiveur et la masse. Ce transistor est piloté par la tension de polarisation Vpol qui devient alors un paramètre de réglage remarquable. Cette tension de polarisation agit directement sur le niveau statique du potentiel cathodique autour duquel les faibles variations, qui traduisent l'image du niveau d'éclairement du pixel, sont ensuite amplifiées au cœur du pixel. Ce paramètre de réglage agit ainsi sur le contraste de l'image. Dans notre application, en l'état, nous le contrôlons manuellement mais à terme, l'idéal serait de réaliser un contrôle intelligent et automatique de ce paramètre en fonction d'autres paramètres caractéristiques extraits de la scène étudiée. Le chapitre suivant revient sur ce point et les perspectives d'un contrôle intelligent de ce paramètre.

4.2.2.2 Mémoires Analogiques, Amplificateurs et Multiplexeurs (MA^2M)

Ce capteur a initialement été conçu pour réaliser des captures d'images a 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 100μs, dans des conditions d'éclairement suffisantes ($\geq$3 000 lux). De plus, dans ce même intervalle temporel, une lecture séquentielle d'une trame précédemment acquise est tout a fait envisageable compte tenu de la résolution du capteur (4096 pixels à balayer en 100μs soit $\approx$ 40 mégapixels par seconde). Aussi, un compromis est à déterminer entre la vitesse de lecture séquentielle et la résolution du capteur.

Dans le but d'augmenter les possibilités algorithmiques de l'architecture, un des points clés est la possibilité, au niveau du pixel, de dissocier la phase d'acquisition et la phase de lecture [CBM+02]. En effet, il suffit pour cela de stocker, au niveau du pixel, la trame en cours d'acquisition et d'accéder, en lecture, à la trame précédemment acquise. Pour cela, il est nécessaire de disposer, au niveau du pixel, de deux mémoires analogiques qui sont, alternativement, accédées en lecture et en écriture de façon complémentaire.

Ainsi dans chaque pixel du circuit sont implémentées deux structures spécifiques MA^2M (Mémoires Analogiques, Amplificateurs et Multiplexeurs), incluant une mémoire analogique, un amplificateur et un multiplexeur (cf. figure 4.11).

Le circuit fonctionne avec une simple alimentation continue de 3,3V. Dans chaque pixel (cf. figure 4.11) le photocapteur est une photodiode de type N, préchargée par un transistor PMOS au potentiel V_{DD} via le signal "reset", signal actif à l'état bas (0V) et global à toute la matrice. Le potentiel cathodique est ensuite polarisé autour de $V_{DD}/2$ (*i.e.*, la moitié de la tension d'alimentation) à l'aide d'un étage suiveur constitué de deux transistors NMOS. Pour que le nœud "N", soit polarisé autour de $V_{DD}/2$, le transistor monté en source commune travaille en régime d'inversion et est piloté par une tension de polarisation stable (Vpol=1,35V). Cette première structure, constituée d'une photodiode et de trois transistors, constitue le premier étage du pixel. A ce premier étage, sont interconnectées au nœud "N" les deux cellules MA^2M, chacune constituée de deux commutateurs CMOS, d'un amplificateur et d'un condensateur de stockage.

Dans le but de minimiser le nombre de transistors, l'amplificateur est un inverseur CMOS utilisé en régime petits signaux autour de $V_{DD}/2$. Autour de ce potentiel sur la plage [0,65 V ; 2,65 V], cette structure se comporte comme un amplificateur inverseur linéaire de fort gain présentant une forte

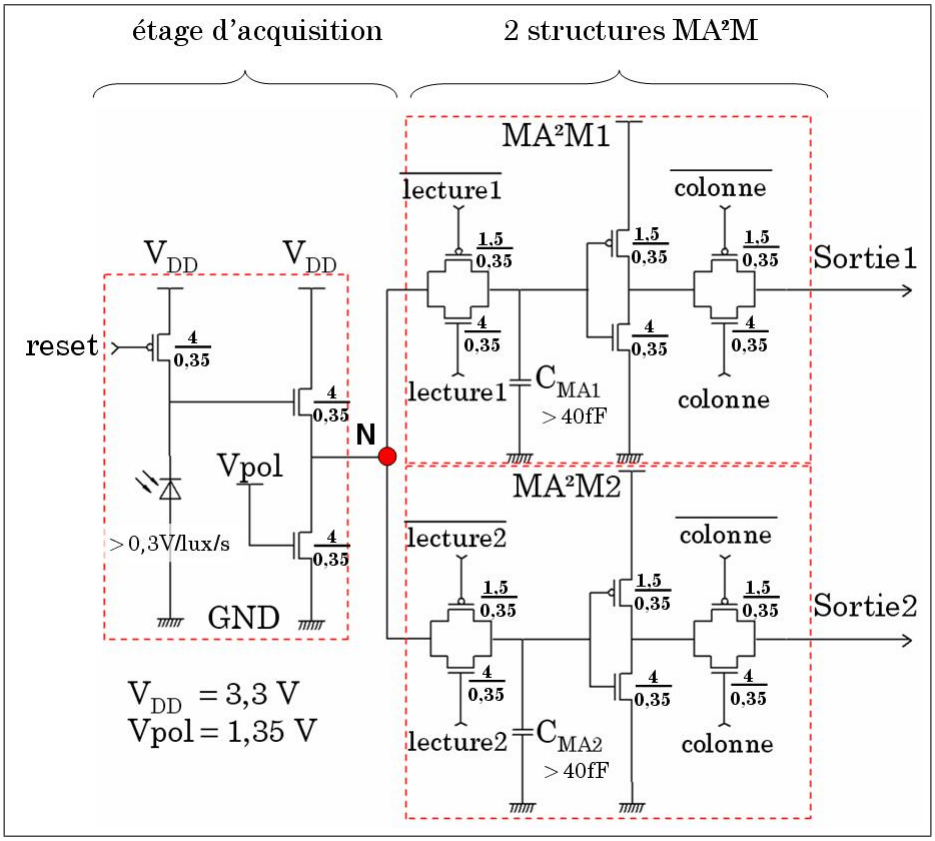


FIG. 4.11 – Schéma du premier étage du pixel et des deux structures MA^{2}M

impédance d'entrée et une faible impédance de sortie. Cet amplificateur présente peu de défauts de linéarité autour de $V_{DD}/2$ (inférieure à 3% d'après la simulation). La forte impédance d'entrée de l'amplificateur, permet au condensateur de stockage de conserver sa charge à l'issue d'une opération de lecture. Deux couches de poly-silicium (POLY1-POLY2) de plus de 50 μm^2 sont superposées pour former un condensateur de capacité supérieure à 40 fF. Une erreur inférieure à 2% a été mesurée sur la différence de potentiel au bornes du condensateur de stockage sur une période de 40 ms environ, ce qui est particulièrement intéressant pour des applications temps réel.

Le potentiel du nœud "N" est inversement proportionnel à la quantité de photons incidents à la cathode et peut être porté aux bornes de la capacité de stockage en agissant sur les signaux "lecture1" et "lecture2", signaux actifs à l'état haut (V_{DD}) et globaux à toute la matrice. Après amplification, les signaux résultants peuvent être redirigés vers les bus asynchrones de sortie "sortie1" et "sortie2" via un deuxième commutateur CMOS, en agissant sur le signal "colonne". Ce signal est actif à l'état haut et global à toute la

colonne à laquelle appartient le pixel considéré. Le circuit d'acquisition et les deux structures MA^{2}M comportent seulement 15 transistors. Les transistors NMOS sont aux dimensions minimales (W_N=0,4 μm) et la largeur de canal des transistors PMOS est W_P=1,5 μm pour compenser la faible mobilité des trous.

Avec cette électronique de contrôle, la rétine ainsi créée peut atteindre des cadences allant jusqu'à 10 000 images par seconde pour des scènes relativement bien éclairées ($\geq 3000\,lux$). Le chapitre 5 présente les résultats des mesures que nous avons réalisée sur cette implantation matérielle.

4.2.2.3 Unité Arithmétique Analogique (UA^2)

L'unité arithmétique analogique est le processeur élémentaire du pixel. Interconnectée à ses quatre proches voisins, cette structure est capable de réaliser une combinaison linéaire de quatre pixels connexes. Par définition, pour réaliser une combinaison linéaire il faut quatre multiplieurs quatre quadrants interconnectés.

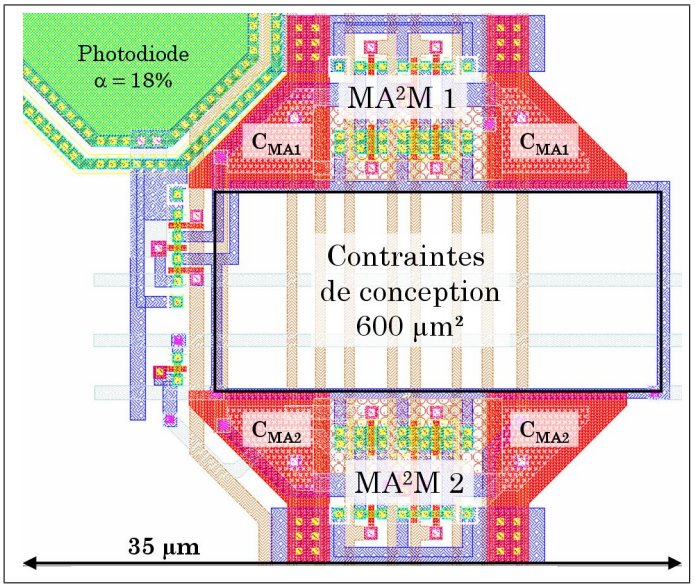
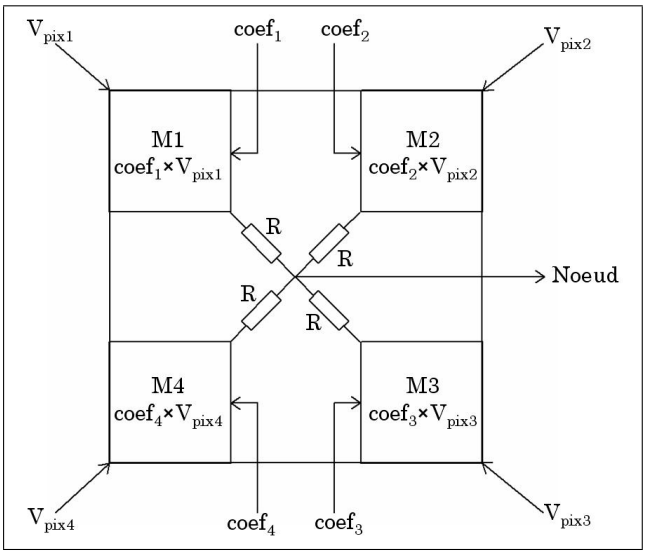


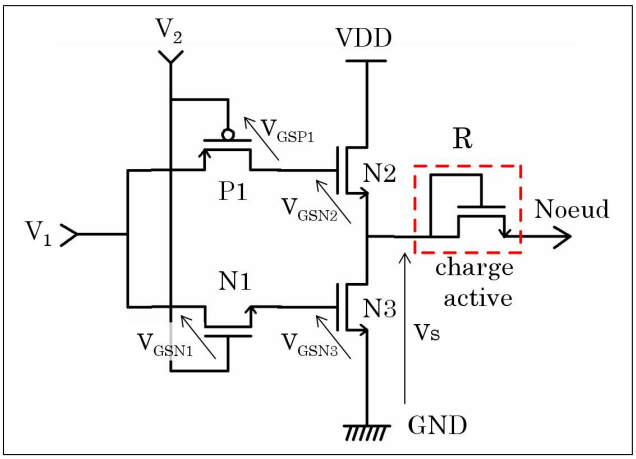
FIG. 4.12 – Layout du pixel avec le circuit d'acquisition et les deux cellules MA^{2}M

Nous nous sommes imposées une taille de pixel de 35 μm de côté en technologie standard CMOS 0,35 μm, soit une surface de 1225 μm^2. Compte tenu du taux de remplissage, qui est ramené à 10% dans la dernière version d'HISIC, et de la surface qu'occupent les 15 transistors des cellules MA^{2}M, il

ne reste plus que 600 μm^2 pour implémenter quatre multiplieurs quatre quadrants (cf. figure 4.12).



(a) Architecture globale



(b) Schéma d'un multiplieur

FIG. 4.13 – L'unité Arithmétique Analogique (UA^2)

Les architectures déjà existantes de multiplieurs quatre quadrants sont généralement basées sur la cellule de Gilbert, d'abord dans les années 70 avec la technologie bipolaire [Rya70] puis dans les années 90 avec la technologie CMOS [LH95]. Il existe aussi des structures basées sur des miroirs de courants [BBM+02]. Toutes ces architectures ont fait leurs preuves mais elles comportent un nombre de transistors bien trop élevé ($\approx$ 10 à 20 transistors par multiplieur) pour envisager d'implémenter au niveau du pixel quatre

multiplieurs quatre quadrants utilisant ces structures. Après avoir réalisé le dessin des masques de la photodiode, de l'étage d'acquisition et des cellules MA^{2}M, les contraintes de conception nous imposent de limiter le nombre de transistors à une vingtaine environ pour la conception de l'UA^2.

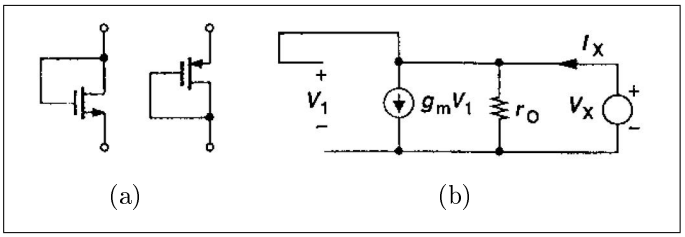


FIG. 4.14 – ((a)Transistors NMOS et PMOS connectés en charge active, (b) circuit équivalent petits signaux

La question est de savoir si nous avons besoin de multiplieurs performants. La réponse à cette question dépend de l'application. Pour exécuter un opérateur de type Sobel ou Laplacien, une précision de 25% suffit car il sagit de la mutiplication par un signal (un coefficient) à quatre niveaux $\{-1; -0,5; 0,5; 1\}$ (cf. chapitre 3). Nous avons alors imaginé, conçu et validé expérimentalement une nouvelle architecture de multiplieur quatre quadrants. L'UA^2 occupe la partie centrale du pixel et inclut quatre multiplieurs quatre quadrants (nommés M1, M2, M3 et M4) (cf. figure 4.13(a)). Le multiplieur compte seulement cinq transistors aux dimensions minimales (cf. figure 4.13(b)).

La sortie Nœud de l'UA^2 fournit le résultat d'une combinaison linéaire par sommation des courants, issus des multiplieurs, via des charges actives. En technologie CMOS, il est difficile de fabriquer des résistances de faible encombrement. Une solution simple et intéressante en terme de densité d'intégration, consiste à connecter la grille et le drain d'un transistor NMOS (cf. figure 4.14(a)). Dans cette configuration, le transistor se comporte comme une résistance en régime petits signaux (cf. figure 4.14(b)). En effet, pour obtenir l'impédance du dispositif, on écrit $V_1 = V_X$ et $I_X = V_X + gm\, V_X$. De ce fait, l'impédance est simplement égale à $(1/gm)\|ro \approx 1/gm$. Avec gm la transconductance du transistor et ro la résistance à vide du transistor.

Le courant résultant au Nœud de sortie est l'image de la fonction P vue dans le second chapitre. C'est une combinaison linéaire des quatre potentiels V_{pix1}, V_{pix2}, V_{pix3} et V_{pix4} pondérés par les quatre coefficients coef1, coef2, coef3 et coef4. Son expression est donnée par la relation :

$$P = coef_1 \times V_{pix1} + coef_2 \times V_{pix2} + coef_3 \times V_{pix3} + coef_4 \times V_{pix4} \qquad (4.6)$$

Les quatre coefficients sont les poids de la combinaison linéaire des quatre pixels connexes au site de traitement considéré. L'unité arithmétique analogique est donc entièrement configurable par ses quatre coefficients qui constituent ses paramètres de configuration lors d'un cycle rétinien (cf. chapitre 3).

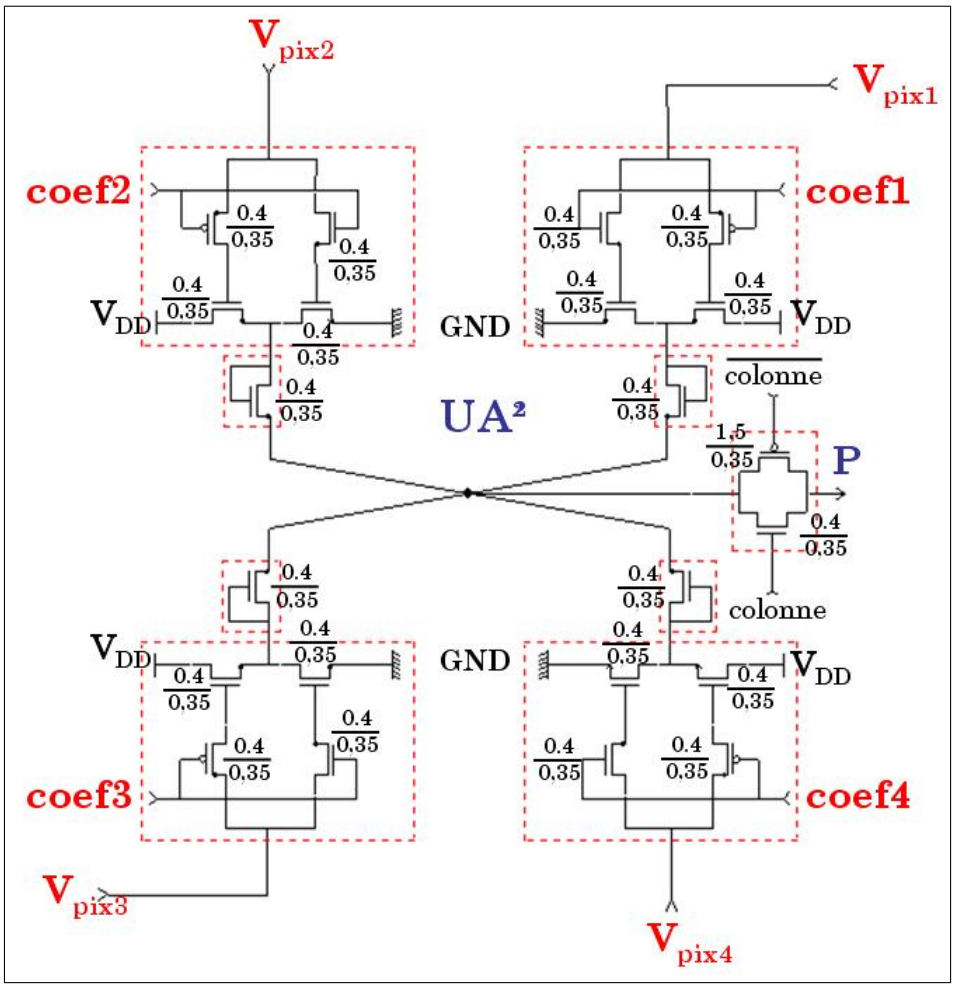


FIG. 4.15 – Schéma complet de l'Unité Arithmétique Analogique (UA^2) : 4 multipieurs 4 quadrants réalisant la combinaison linéaire de 4 pixels adjacents

Cinq transistors aux dimensions minimales suffisent pour réaliser le produit de deux signaux multiniveaux M aire (cf. théorie de la transmission de signal en bande de base à travers un canal bruité). Où le paramètre M définit le nombre de niveaux différentiables par la structure multiplicatrice. Avec un nombre aussi restreint de transistors, il faut émettre des réserves sur la valeur de M, d'autant plus que le résultat du produit est un signal à M^2 niveaux. Le schéma structurel complet compte seulement 22 transistors tous aux di-

mensions minimales (cf. figure 4.15). La sortie Nœud est multiplexée vers la sortie du pixel via un commutateur CMOS piloté par le signal "colonne".

En s'appuyant sur la figure 4.13(b), on démontre analytiquement la fonction de transfert du multiplieur au premier ordre à l'équilibre statique. Le courant de drain d'un transistor MOS (de type N ou P), selon l'analyse théorique connue de la région ohmique, a pour expression :

$$I_D = \mu C_{ox} \frac{W}{L} \left[(V_{GS} - V_{TH}) V_{DS} - \frac{1}{2} V_{DS}^2 \right] \tag{4.7}$$

Avec μ la mobilité des porteurs majoritaires dans le canal créé, L est la largeur du canal, W sa largeur, V_{TH} est la tension de seuil du transistor. C_{ox} est la capacité d'oxyde de grille. Les valeurs de tous ces paramètres sont listées dans le tableaux A.1, en annexe A. V_{GS} et V_{DS} sont respectivement les différences de potentiels Grille-Source et Drain-Source du transistor. Dans cette expression, si on suppose que les transistors opèrent en régime triode (cf. figure 4.16) avec une différence de potentielle Drain-Source faible, $V_{DS} \ll 2(V_{GS} - V_{TH})$, le courant de drain peut être approximé par :

$$I_D \approx \mu C_{ox} \frac{W}{L} (V_{GS} - V_{TH}) V_{DS} \tag{4.8}$$

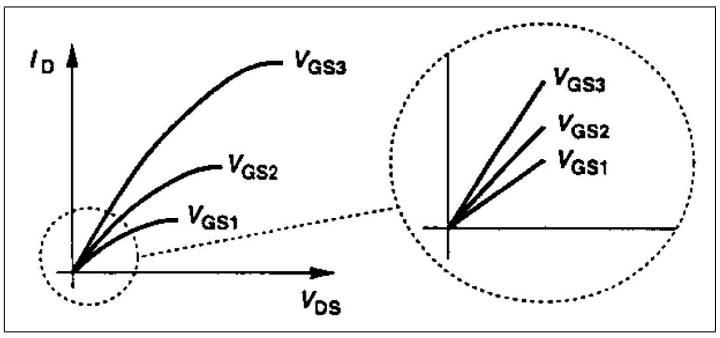


FIG. 4.16 – MOS en régime triode, I_D faible

Avec ces hypothèses, le courant de drain est alors une application linéaire de V_{DS}. Le canal ainsi créé, d'électrons ou de trous suivant la nature du transistor MOS de type N ou P, se comporte comme une résistance linéaire contrôlée par la tension V_{GS} (cf. figure 4.17) dont l'expression est donnée par :

$$R_{on} \approx \frac{1}{\mu C_{ox} \frac{W}{L} (V_{GS} - V_{TH})} = \frac{1}{\beta (V_{GS} - V_{TH})} \tag{4.9}$$

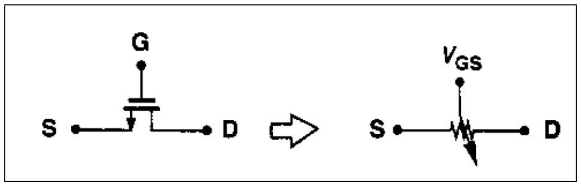


FIG. 4.17 – MOS contrôlé comme une résistance linéaire

D'après la figure 4.13(b), et dans le cadre des hypothèses pré-citées, il vient l'expression de la tension de sortie Vs :

$$V_S = \frac{R_{N3}V_{DD}}{R_{N2}+R_{N3}} = \frac{V_{DD}}{R_{N2}}\left(\frac{1}{\frac{1}{R_{N2}}+\frac{1}{R_{N3}}}\right)$$
$$\text{avec, } \begin{cases} R_{N2} = \frac{1}{\beta_N(V_{GSN2}-V_{THN})} \\ R_{N3} = \frac{1}{\beta_N(V_{GSN3}-V_{THN})} \end{cases} \tag{4.10}$$

Où R_{N2} et R_{N3} sont les résistances de canal des transistors N2 et N3. De manière identique, les résistances de canal des transistors P1 et N1 ont pour expression :

$$\begin{cases} R_{P1} = \frac{1}{\beta_P(V_{GSP1}-V_{THP})} & \text{avec, } V_{GSP1} = V_2 - V_S - V_{GSN2} \\ R_{N1} = \frac{1}{\beta_N(V_{GSN1}-V_{THN})} & \text{avec, } V_{GSN1} = V_1 - V_2 \end{cases} \tag{4.11}$$

A l'équilibre statique, supposons que les potentiels V_1 et V_2 équilibrent les deux branches formées par les transistors N1 et P1 (cf. figure 4.13(b)), on peut donc établir la relation suivante dans un domaine de fonctionnement limité à de faibles amplitudes des tensions V_1 et V_2 autour de $V_{DD}/2$:

$$\frac{V_{GSN3} - V_1}{R_{N1}} = -\frac{V_{GSN2} + V_S - V_1}{R_{P1}} \tag{4.12}$$

Des équations 4.10 et 4.12, on déduit le système suivant :

$$\begin{cases} V_S\left(V_{GSN2} + V_{GSN3} - 2V_{THN}\right) = V_{DD}\left(V_{GSN2} - V_{THN}\right) \\ \frac{\beta_N}{\beta_P}\left(V_{GSN3} - V_1\right)\left(V_1 - V_2 - V_{THN}\right) \\ \qquad = \left(V_1 - V_S - V_{GSN2}\right)\left(V_2 - V_S - V_{GSN2} - V_{THP}\right) \end{cases} \tag{4.13}$$

Où le rapport $\frac{\beta_N}{\beta_P} = \frac{\mu_N C_{ox}\frac{W_N}{L}}{\mu_P C_{ox}\frac{W_P}{L}} = \alpha \approx 2,94$ équivaut à un rapport de mobilités entre électrons et trous puisque tous les transistors sont aux dimensions

minimales (W_N=W_P=0,4μm). Supposons le point de repos à $V_{DD}/2$ et toujours le régime triode, $V_{GSN3} = V_{GSN2} = V_{GS} = V_{THN}$, on obtient alors le système d'équations suivant :

$$\begin{cases} V_S \cong \frac{V_{DD}}{2} \\ \alpha\,(V_{THN} - V_1)\,(V_1 - V_2 - V_{THN}) \\ \qquad\qquad = (V_1 - V_S - V_{THN})\,(V_2 - V_S - V_{THN} - V_{THP}) \end{cases} \tag{4.14}$$

Dans cette expression, on constate que les termes qui traduisent, soit un défaut d'isolation des ports d'entrée (termes en V_1^2, V_2^2, V_1^3, etc.), soit un produit d'intermodulation (termes en $V_1^2\,V_2$, $V_1\,V_2^2$, $V_1^2\,V_2^2$, etc.) sont pondérés par un facteur qui vaut au mieux 11% du coefficient de transfert. Autour du point de repos $V_{DD}/2$, le potentiel de sortie V_S subis donc des variations principalement dues au produit V_1V_2. Les contributions, au premier ordre, extraites de la deuxième égalité de la relation 4.14, sont données par l'équation suivante :

$$\alpha\,(V_1V_2) = (V_1V_2) + 2V_SV_{THN} + V_SV_{THP} \tag{4.15}$$

On peut établir l'équation du multiplieur, autour du point de repos $\frac{V_{DD}}{2}$:

$$V_S = K(V_1 - V_{DD}/2)(V_2 - V_{DD}/2) \quad avec, \quad K = \frac{\alpha - 1}{2V_{THN} + V_{THP}} \tag{4.16}$$

Le coefficient de transfert est maximum ($K \approx 8,07\,V^{-1}$) avec une largeur de canal du transistor PMOS égale à la valeur minimale W_P=0,4μm. Ce coefficient est prépondérant et minimise les effets des produits d'intermodulations et des défauts d'isolations du multiplieur. Néanmoins cette valeur théorique n'est valable que pour de faibles variations des tensions V_1 et V_2 autour de $V_{DD}/2$. L'expérience a montré que ce coefficient de transfert décroît à mesure que l'amplitude de V_1 et V_2 augmente, ce que nous pressentions compte tenu de l'expression analytique de V_S. Cette structure est capable de multiplier sans erreur le signal pixel par un signal (un coefficient) à huit niveaux soit une erreur relative inférieure à 12,5% ou encore une précision d'au moins 87,5%, ce qui est loin d'être exceptionelle comparée aux performances de la cellule de Gilbert en technologie standard CMOS [Gil68, KDVF02]. Néanmoins, cette UA^2 peut être utilisée pour appliquer des masques locaux sur huit niveaux (profondeur de 3 bits), au delà les performances se dégradent considérablement comparées à une architecture classique (capteur et calculateur séparés). Toutefois, une préscision de trois bits est suffisante dans un

grand nombre d'applications comme par exemple le filtre de Canny utilisé dans un cas particulier [DK97].

Dans le chapitre suivant, nous présentons les résultats des mesures expérimentales réalisées sur une structure de test en régime harmonique afin de quantifier les défauts de la structure (produits d'intermodulations et défauts d'isolations des ports d'entrée) et nous proposons de comparer les performances de notre capteur avec une architecture classique.

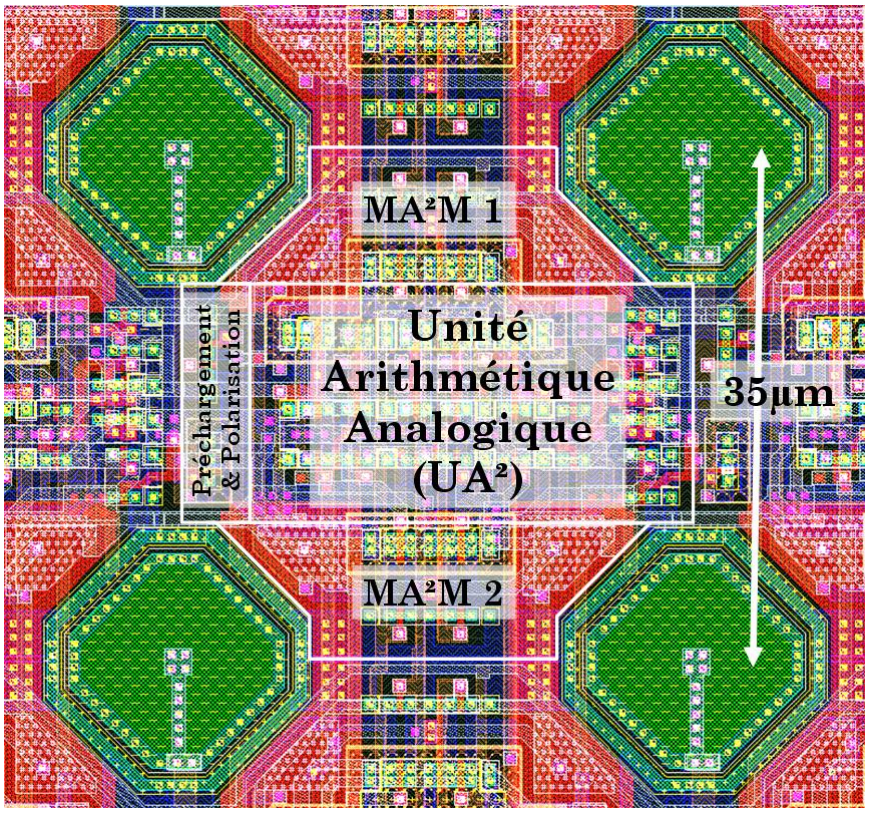


FIG. 4.18 – Layout d'un pixel et de ces quatre photodiodes connexes

Le dessin des masques du pixel (quadruple métal, double poly), avec ses quatre photodiodes connexes, est illustré figure 4.18. La symétrie de la structure est remarquable. Le centre de l'unité arithmétique analogique est équidistant des centres des photodiodes. Cela a une incidence directe sur le bruit spatial fixe du capteur qui est en l'occurrence moindre dans cette configuration. Dans cette topologie 4-connexe, il est plus facile de réaliser un routage efficace en limitant par exemple les capacités parasites de couplages entre les différentes UA^2 implémentées au sein des pixels. L'UA^2 avec ses 20 transistors représente un encombrement de seulement $600\,\mu m^2$ soit la moitié

de la surface totale du pixel (1225 μm^2) dont la taille est de seulement 35 μm de côté.

La section qui suit présente l'architecture du convertisseur analogique numérique intégré en amont du multiplexeur et répartie au niveau des colonnes du capteur.

4.3 Architecture du convertisseur analogique-numérique

La première version d'HISIC était dépourvue de convertisseur analogique numérique, la conversion se faisait à l'extérieur de la rétine. Les signaux analogiques des trois bus asynchrones traversait des longueurs de métal non négligeables et donc susceptibles au bruit rayonnant. Pour palier ces effets indésirables, nous avons conçu une structure de CAN rapide, de résolution 8 bits et basé sur le principe de l'approximation successive. Ce CAN se compose d'un comparateur par colonne de l'imageur et d'un CNA R/2R 8 bits. La stratégie de commande utilisée pour contrôler l'adressage des colonnes permet de doubler la vitesse de conversion par rapport à une commande en simple rampe.

Nous avons implémenté 64 comparateurs rapides en amont du multiplexeur (cf. figure 4.19). Cette façon de procéder a l'avantage de ne pas charger la partie résistive de l'impédance de la colonne et contribue ainsi à augmenter le rapport signal sur bruit. De plus, la conversion est monolithique, ce qui réduit considérablement les longueurs de métal qui propagent le signal analogique par rapport à un convertisseur hors substrat, sur PCB. Le dessin des masques du quart inférieur gauche de la rétine est illustré sur la figure 4.20.

Le CNA qui génère la valeur de référence (à comparer) est l'un des rares composants propriétaires qui a servit à la conception de notre rétine. Ce CNA est un réseau R/2R de la bibliothèque de composants AMS. Un des composants clés de ce convertisseur est le comparateur qui doit être rapide. En effet, nous souhaitons convertir 64 lignes de pixels en 100μs (10 000 im/s), soit 1,56μs par ligne. Or dans le cas le plus défavorable d'une commande simple rampe, il faut 256 cycles d'horloge pour convertir une ligne de pixels, soit une comparaison toutes les 6 ns. De plus nous nous imposons un seuil différentiel de détection inférieur à 1/2 LSB sur une plage de tension de 3,3 V, soit un seuil de $3,3\ V/2^9 \approx 6,5\ mV$.

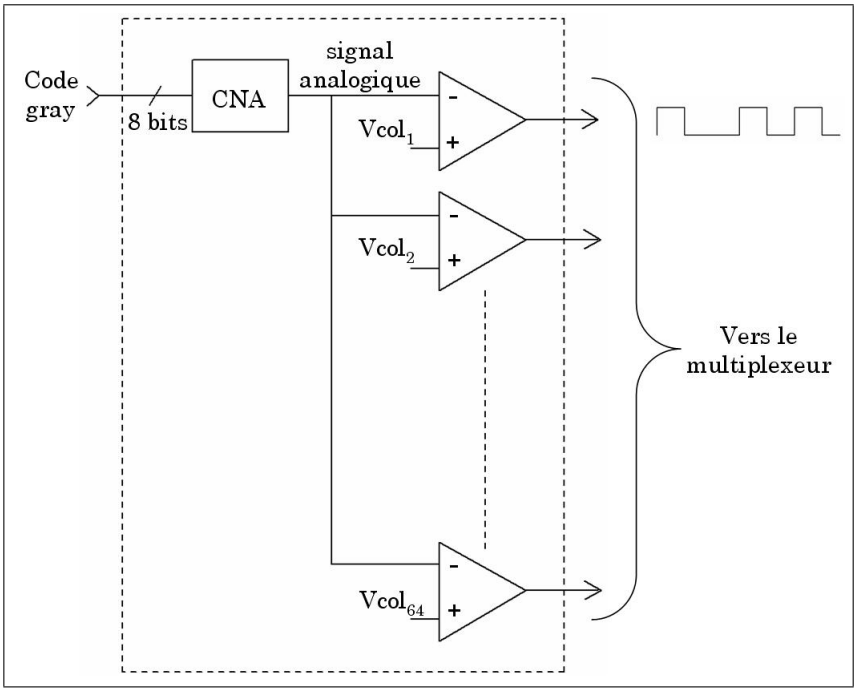


FIG. 4.19 – Synoptique de l'architecture de convertisseur

4.3.1 Architecture du comparateur

Un comparateur rapide et basse consommation [GM93, SSA98] est requis pour comparer le potentiel de référence avec le potentiel de la sortie du pixel adressé. L'architecture de comparateur proposée est composée de trois étages (cf. figure 4.21) : (a) l'amplificateur opérationnel à transconductance (amplificateur différentiel de courant), (b) la paire différentielle et (c) l'amplificateur inverseur.

D'après les données du constructeur AMS en technologie standard CMOS 0,35 μm (cf. annexe A), $\mu_N C_{ox}$=189$\mu A.V^{-2}$ et $\mu_P C_{ox}$=64$\mu A.V^{-2}$. Dans les sections qui suivent, nous exprimons le gain gain statique de chaque étage et nous proposons un dimensionnement. Dans le dimensionnement proposé, nous faisons le choix de fixer la largeur de grille des transistors à la valeur minimale de la technologie soit $L = 0,35\,\mu m$ et le paramètre géométrique variable est donc la largeur de canal $W \geq 0,4\mu m$.

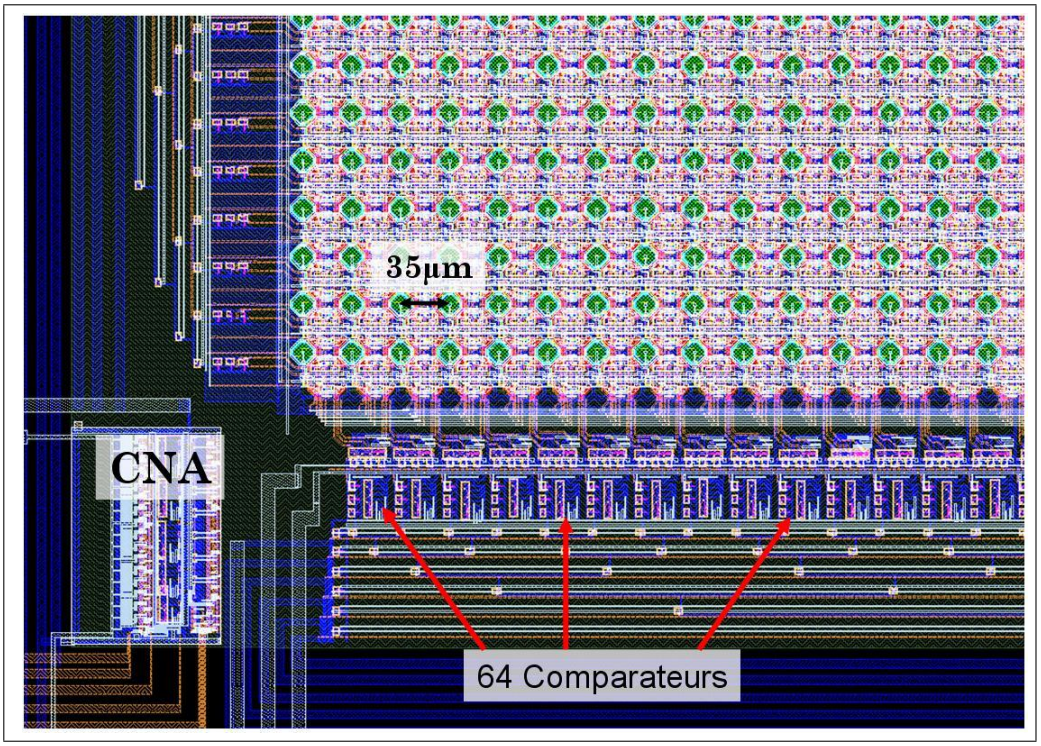


FIG. 4.20 – Dessin des masques en technologie CMOS 0,35μm du convertisseur analogique numérique

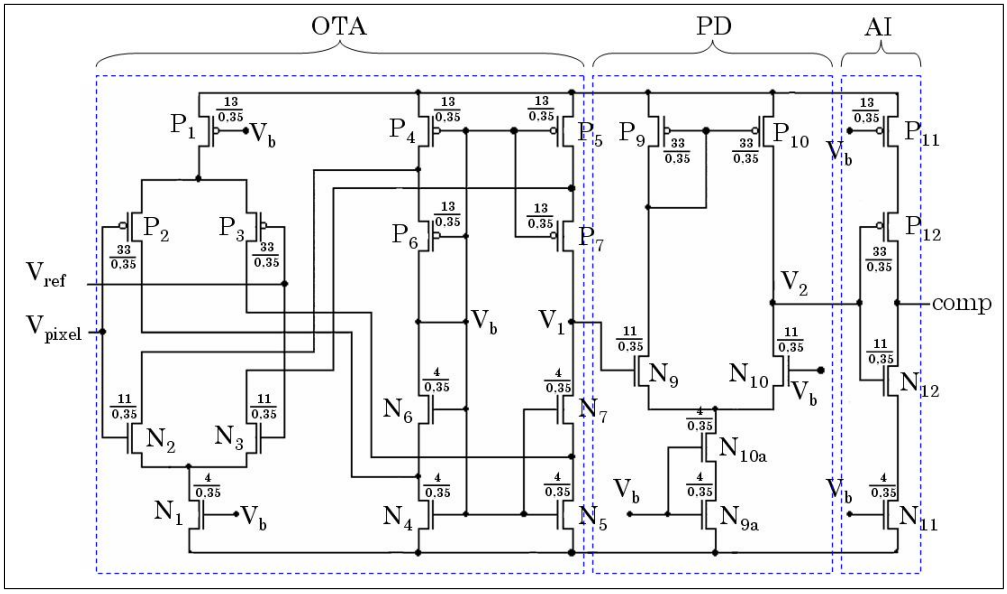


FIG. 4.21 – Schéma électronique du comparateur

4.3.1.1 Étage OTA : Amplificateur Opérationnel à Transconductance

Un amplificateur opérationnel à transconductance (OTA, sur la figure 4.21) avec une dynamique maximale de tension d'entrée de mode commun est présenté dans cette partie. Une auto-polarisation permet de réduire sensiblement les effets des variations de température et de tension d'alimentation [KPS+05]. Son architecture, remarquablement symétrique [BLB98], est composée des transistors N_1 à N_7 et P_1 à P_7. V_{pol} est la tension de polarisation du réseau cascodé composé des transistors P_4, P_6, N_4 et N_6. Le gain différentiel de l'OTA est donné par la relation :

$$Av_{OTA} = (gm_{N_2,N_3} + gm_{P_2,P_3}) \left(\frac{r_{out_N}\ r_{out_P}}{r_{out_N} + r_{out_P}} \right) \tag{4.17}$$

Où gm_N, gm_P sont les transconductances des transistors de la paire différentielle, de type N et P respectivement. Les parties résistives des impédances de sortie du réseau cascodé sont notées r_{out_N} et r_{out_P}. Cette structure cascode, en régime petits signaux, se replie sur elle même. On dit qu'elle est auto-cascodée et la partie résistive de son impédance de sortie, vue de la branche composée des transistors de type N (N self-cascode), est donnée par la relation :

$$r_{out,Nself-cascode} = [1+(gm_{N7,P7}+gmb_{N7,P7})\, ro_{N5,P5}]\, ro_{N7,P7}+ro_{N5,P5} \tag{4.18}$$

Où gmb est la transconductance de substrat du transistor et ro, la résistance à vide de sortie du transistor seul. Il est à noter que -gm.ro est le gain intrinsèque du transistor compris, en valeur absolue, entre 15 et 25 [Raz01], en technologie CMOS 0,35 μm. Dans le modèle petits signaux du transistor (cf. figure 4.22), la transconductance de substrat dépend de la différence de potentiel entre la source et le substrat. Il est important de toujours polariser le substrat au potentiel d'alimentation le plus faible. En technologie CMOS, on utilise une alimentation simple ($V_{DD} - V_{GND} = 3,3\,V$, en technologie CMOS 0,35 μm), le substrat est donc toujours polarisé au potentiel V_{GND} via une diffusion P^+ (P^+ vs. V_{GND} pour les NMOS et N^+ vs. V_{DD} dans un caisson N pour les PMOS). La position et le nombre de ces diffusions ont un impact direct sur la transconductance de substrat. L'OTA est utilisé en comparateur. Le critère de dimensionnement porte sur le temps de réponse. Il faut d'abord déterminer un dimensionnement qui optimise le temps de réponse puis calculer le gain statique de ce premier étage.

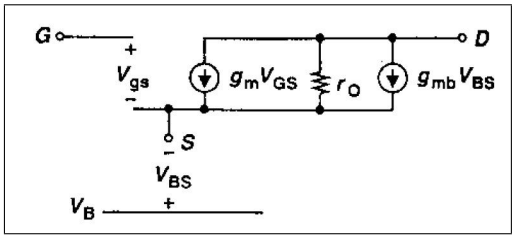


FIG. 4.22 – Modèle petits signaux avec prise en compte de l'effet de substrat

La tension V_b sert à fixer le courant de polarisation dans les paires différentielles d'entrées (P_1 à P_3 et N_1 à N_3) et dans les deux branches cascodes (P_4 à P_7 et N_4 à N_7). Les sources de courant fonctionnent dans leur zone de pincement et toute la difficulté posée par le dimensionnement du montage repose sur l'optimisation de la tension de saturation $V_{DSsat} = V_{GS} - V_{TH}$. L'OTA cascode replié a tendance à être bruyant ; un faible V_{DSsat} augmente la dynamique, mais un plus grand V_{DSsat} minimise l'encombrement de la source de courant et donc le bruit généré par elle-même. Une autres caractéristique est encore plus importante et concerne le temps de réponse. Il dépend de la capacité de charge de l'OTA. Les outils de simulation permettent alors de trouver un compromis en traçant l'évolution de capacité de charge de l'OTA en fonction de la tension V_{DSsat} (cf. figure 4.23).

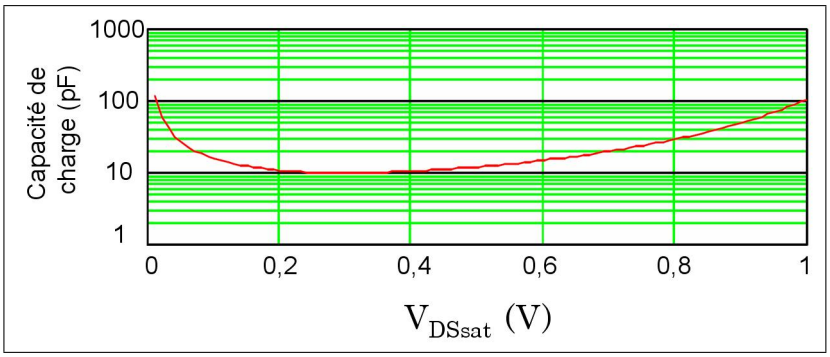


FIG. 4.23 – Capacité de charge de l'OTA en fonction de la tension de saturation

On constate que la capacité est minimale (10 pF) pour une tension $V_{DSsat} = 330\,mV$ soit $1/10^{me}$ de la tension d'alimentation. Par conséquent, nous fixons les différences de potentiels Drain-Source des sources de courant (N_1, N_4, N_5, P_1, P_4 et P_5) à 330 mV afin de limiter leur encombrement sans trop faire baisser la dynamique de cet étage tout en minimisant le temps de réponse. La valeur du courant de polarisation fixe à la fois le dimensionnement, la puissance consommée et le bruit. Un courant de polarisation fort est

défavorable à la densité d'intégration et à la consommation, mais un courant faible a un impact négatif sur la bande passante de l'OTA. Par conséquent, nous fixons le courant de polarisation à 1,5 mA, soit une puissance statique de 14×330 mV×1,5 mA≈7 mW, ce qui est aceptable. Connaissant les valeurs des tensions de pincement et des courants de polarisation, nous pouvons dimensionner les sources de courant via l'expression :

$$(W/L) = \left| \frac{2\, I_D}{\mu\, C_{ox}\, V_{DSsat}} \right| \tag{4.19}$$

On a donc $(W/L)_{P1,4,5,6,7}$=36, et $(W/L)_{N1,4,5,6,7}$=12. Puisque les valeurs minimales et maximales de la tension de sortie V_1 sont respectivement 0,6 V et 2,7 V, le niveau de sortie de mode commun est 1,65 V=$V_{DD}/2$. Les dimensions minimales de N_2 et N_3 dépendent du niveau d'entrée de mode commun $V_{GS3} + V_{DSsat} = V_{DD}/2$, soit V_{GS3}=1,65 V-0,33 V=1,32 V. Dans ce cas, la tension de pincement des transistors N_2 et N_3 est $Vp = V_{GS3} - V_{THN} = 1,32\,V - 0,46\,V = 0,86\,V$ et donc $(W/L)_{N2,3}$=32. En appliquant le même raisonnement sur P_2 et P_3, on a $(W/L)_{P2,3}$=94.

On peut dorénavant calculer le gain statique de l'étage à partir des relations 4.17 et 4.18. En première approximation, on peut négliger l'influence des transcondustances de substrat et calculer le gain petit signal des transistors en utilisant $gm = 2\,I_D/(V_{GS} - V_{TH})$ et $ro = 1/(\lambda\, I_D)$. Le gain de l'étage minimisant la capacité de charge de l'OTA est voisin de 15. Pour augmenter le gain total du comparateur, on ajoute une paire différentielle.

4.3.1.2 Étage PD : Paire Différentielle

Le deuxième étage du comparateur est une paire différentielle (PD, sur la figure 4.21), composée des transistors $N_{9,9a}$, $N_{10,10a}$, P_9 et P_{10}. Cet étage contribue à augmenter le gain différentiel total du comparateur au détriment du gain de mode commun et donc à réduire l'effet de l'offset. Pour un comparateur, la tension d'offset est le seuil différentiel de détection[1]. Le gain statique de la paire différentielle est donné par la relation :

$$Av_{PD} = gm_{N_9} \left(\frac{r_{out_{N_{10}}}\; r_{out_{P_{10}}}}{r_{out_{N_{10}}} + r_{out_{P_{10}}}} \right) \tag{4.20}$$

Où $r_{out_{N_{10}}}$ et $r_{out_{P_{10}}}$, sont les parties résistives des impédances de sortie du montage cascode. Elles ont les mêmes expressions que l'équation 4.18.

[1]Le seuil différentiel de détection est la tension différentielle d'entrée minimale qui induit une tension différentielle de sortie maximale.

Sans cet étage, le point de fonctionnement statique de l'amplificateur de sortie [N_{10} ; N_{11} ; P_{10} ; P_{11}] serait sensiblement affecté. Comme pour l'OTA, les sources de courant sont des transistors fonctionnant dans leur zone de pincement [N_{9a} ; N_{10a}] et le courant de polarisation est fixé à I_D=1,5 mA. Ce qui donne $(W/L)_{N9a,10a}$=12. Les autres transistors ont une plage de fonctionnement beaucoup plus étendue et il faut prendre en compte la modulation de largeur de canal. Il faut trouver un compromis qui ne pénalise pas le temps de monté du montage. Pour augmenter le gain, nous savons que $gm\,ro = \sqrt{2\,\mu C_{ox}(W/L)I_D}/(\lambda\,I_D)$ donc $gm\,ro$ varie avec $\sqrt{WL/I_D}$. Mais augmenter W fait augmenter la capacité de charge et donc diminue d'autant la bande passante du montage. Le compromis à faire est identique à l'étage OTA, on a donc $(W/L)_{N9,10}$=32 et $(W/L)_{P9,10}$=94, soit un gain voisin de 10.

4.3.1.3 Étage AI : Amplificateur Inverseur

Le dernier étage est un amplificateur inverseur (AI, sur la figure 4.21), composé des transistors N_{10}, N_{11}, P_{10} et P_{11}. Cette structure contribue à augmenter le gain statique et à convertir la sortie de la paire différentielle en un signal numérique standard. Son gain statique est donné par la relation :

$$Av_{AI} = (gm_{N_{12}} + gm_{P_{12}}) \left(\frac{r_{out_{N_{12}}}\ r_{out_{P_{12}}}}{r_{out_{N_{12}}} + r_{out_{P_{12}}}} \right) \tag{4.21}$$

Comme ci-dessus, le courant de polarisation est fixé à I_D=1,5 mA, ce qui donne $(W/L)_{N11}$=12 et $(W/L)_{P11}$=36. Pour les mêmes raisons que précedemment $(W/L)_{N12}$=32 et $(W/L)_{P12}$=94, soit un gain voisin de 15. Le gain total du comparateur est en théorie d'au moins 2200. Dans le chapitre suivant, nous présentons les résultats de mesures réalisées sur le comparateur et en particulier la mesure du temps de réponse. La section suivante présente la manière dont est contrôlée la valeur de référence et l'adressage du multiplexeur de colonnes.

4.3.2 Stratégie de commande

Il existe plusieurs stratégies de commande, la plus simple mais la plus lente étant la méthode simple rampe. Nous proposons une autre stratégie qui permet de doubler la vitesse de conversion sans ressources matérielles supplémentaires. Le principe de l'approximation successive est utilisé pour améliorer le contrôle de ce convertisseur.

On utilise cet algorithme convergent pour doubler la vitesse de conversion

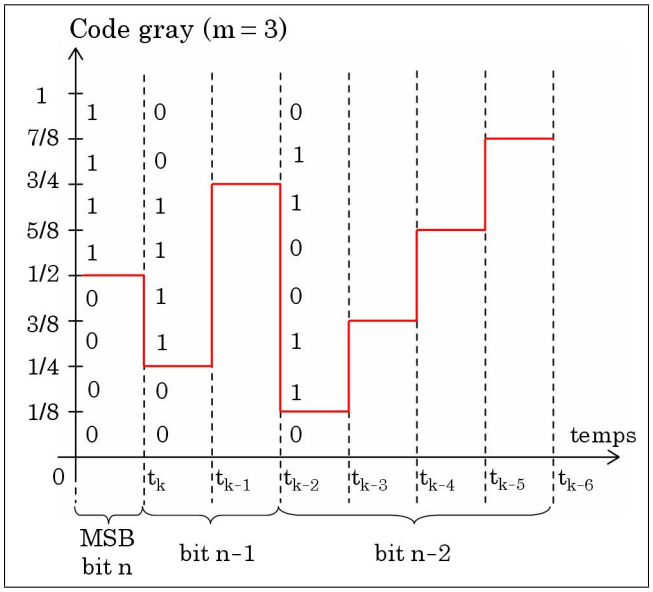


FIG. 4.24 – Signal appliqué en entrées des comparateurs

par rapport à une commande simple rampe. Une stratégie simple consiste à lire séquentiellement, une première fois, les résultats des comparaisons à la valeur de référence $V_{DD}/2$. Les résultats, qui sont stockés dans 64 registres de 8 bits, donnent directement la valeur du bit de poids fort. L'accès aléatoire aux colonnes est exploité dès la seconde lecture. En fonction de la valeur du bit, stockée à l'issue de la dernière lecture séquentielle, seules les colonnes concernées sont adressées et les résultats correspondants sont stockés en poids et valeurs dans les registres concernés. L'algorithme est physiquement implémenté dans le circuit FPGA interfacé à la rétine.

Un signal semblable à celui de la figure 4.24 est appliqué à l'entrée des comparateurs à l'aide du circuit FPGA. De cette façon, à mesure que le signal progresse, seules sont adressées une partie des 64 colonnes. Par exemple, pour une ligne donnée, si $t \in [0; t_k]$ tous les pixels de la ligne sont comparés à la valeur $V_{DD}/2$. Les résultats des comparaisons sont successivement stockés dans le bit de poids fort MSB=n (1 si la valeur du pixel est supérieure, 0 sinon) d'une mémoire RAM de 64 octets aux adresses correspondantes. Ensuite, si $t \in [t_k; t_{k-1}]$, seules les colonnes dont l'adresse correspond effectivement à un résultat négatif (0) sont physiquement adressées pour comparer leurs valeurs à $V_{DD}/4$. Les résultats des comparaisons sont successivement stockés dans le bit de poids n-1 de la RAM de 64 octets aux adresses correspondantes. si $t \in [t_{k-1}; t_{k-2}]$, seules les colonnes dont l'adresse correspond effectivement à un résultat positif (1) sont physiquement adressées pour comparer leurs va-

leurs à $3V_{DD}/4$. Les résultats des comparaisons sont successivement stockés dans le bit de poids n-1 de la RAM de 64 octets aux adresses correspondantes. De cette façon, les deux bits les plus significatifs de la conversion sont entièrement déterminés et permettent de savoir quelles sont les colonnes qui doivent être adressée dans les quatre étapes qui suivent pour déterminer le bit de poids n-2.

Globalement, ce système se comporte comme un système asservi. L'algorithme dichotomique, utilisé ici, permet de doubler la vitesse de conversion. En effet, il existe 2^n possibilités d'écrire un nombre binaire sur n bits et il faut autant de $(1)_2$ et de $(0)_2$ pour écrire tous ces nombres dans la base 2, soit $2^n \times 8/2$ $(1)_2$ et autant de $(0)_2$. Globalement, le nombre d'accès aux multiplexeurs est divisé par deux. Concrètement, il faut seulement 128 cycles d'horloge pour convertir une ligne entière de pixels, soit au plus $100\,\mu s/64$ lignes$/128$ cycles≈ 12 ns par comparaisons pour un capteur de 64 lignes à la cadence de 10 000 images/s. Ce concept peut, très facilement, s'appliquer à des imageurs de grandes tailles car les ressources mémoires nécessaires au niveau du circuit FPGA (BRAM) sont relativement faibles, une mémoire simple port de N octets suffit, avec N le nombre de colonnes de l'imageur.

4.4 Décodeur de lignes et multiplexeur de colonnes

Le décodeur de lignes et le multiplexeur de colonnes sont nécessaires pour réaliser un adressage X Y de la matrice. Cet adressage permet l'accès aléatoire aux pixels. La structure du décodeur et du multiplexeur est basée sur une architecture en cascade.

Le décodeur de ligne (cf. figure 4.25) est une fonction purement numérique, tous ces transistors sont aux dimensions minimales sauf pour les inverseurs dont le transistor PMOS a une largeur de canal de $1{,}5\,\mu$m. Les charges actives NMOS créent le zéro récessif nécessaire pour assurer la stabilité à l'entrée des triples inverseurs à droite du schéma.

Le multiplexeur de colonne (cf. figure 4.26) est aussi une fonction purement numérique, néanmoins un peu plus complexe que la structure du décodeur de ligne. La partie décodage d'adresse à droite du schéma est exactement la structure du décodeur de ligne. Pour simplifier la représentation, le commutateur est représenté de manière symbolique. Nous aurions pu imaginer le même décodeur d'adresses avec des portes ET mais la structure comporterait

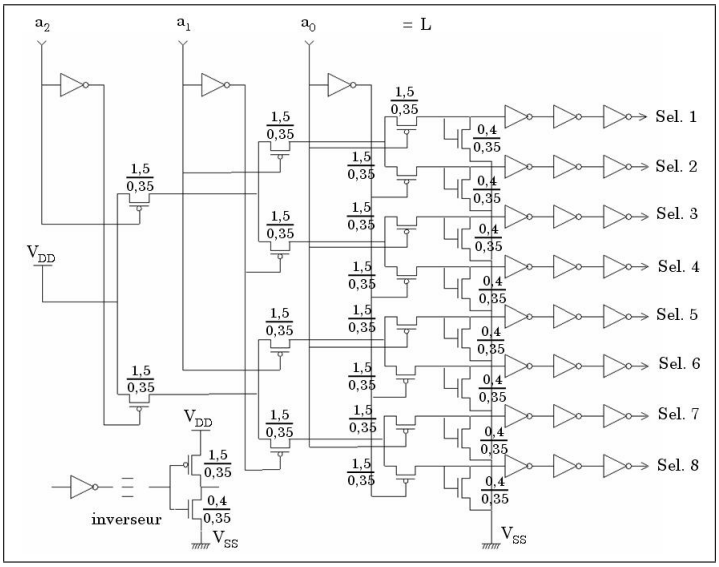


FIG. 4.25 – Architecture du décodeur de lignes

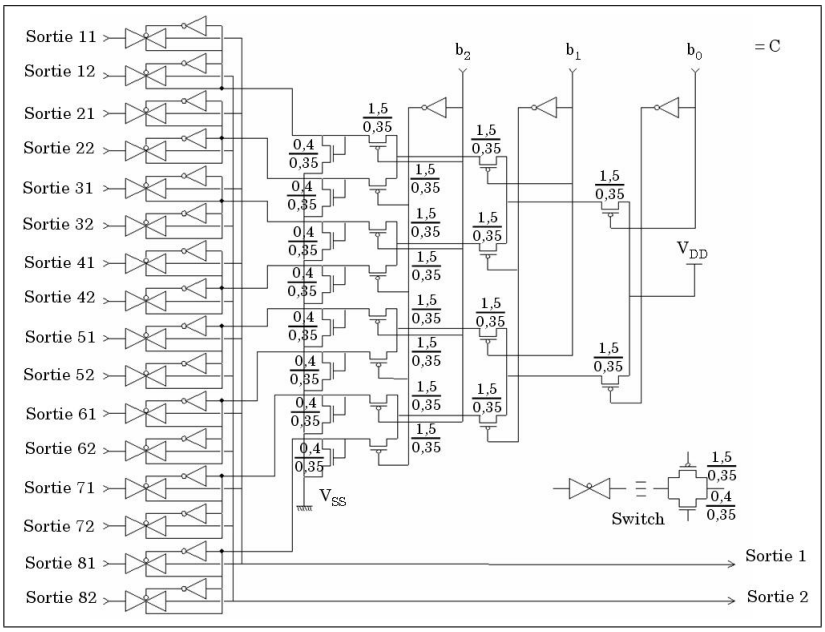


FIG. 4.26 – Architecture du multiplexeur de colonne

alors trois fois plus de transistors et consommerait d'autant plus d'énergie électrique.

4.5 Conclusion

HISIC présente une architecture de capteur d'images CMOS classique de type APS avec un convertisseur analogique numérique réparti au niveau des colonnes du capteur, un décodeur de lignes et un multiplexeur de colonnes. La particularité de notre capteur est la conception du pixel qui intègre à la fois des fonctions dédiées à l'imagerie rapide (cellules MA^2M) et des éléments de calculs élémentaires interconnectés en topologie 4-connexe (UA^2). Nous avons présenté le dimensionnement des structures électroniques de notre capteur, du photo-détecteur au CAN, en passant par le multiplieur à cinq transistors. La plupart des points abordés ont été validés analytiquement et des simulations numériques ont appuyé certains choix technologiques (photodiode). La stratégie que nous avons développée pour le contrôle du convertisseur permet de doubler la vitesse de conversion par rapport à un commande simple rampe.

Le chapitre suivant présente les résultats de mesures que nous avons faites sur les circuits HISIC, la plateforme expérimentale d'évaluation (HISIC.I), le système de vision HISIC.II ainsi que les résultats d'acquisitions et de traitements d'images linéaires.

Chapitre 5

Réalisations et résultats expérimentaux

5.1 Introduction

Le chapitre précédent décrit l'architecture de notre rétine CMOS qui intègre, au sein de chaque pixel (cf. figure 5.1), un étage APS, une UA^2 (Unité Arithmétique Analogique) et deux cellules MA^{2}M (Mémoire Analogique, Amplificateur et Multiplexeur). Ce chapitre présente les résultats expérimentaux et les mesures que nous avons réalisées sur nos circuits.

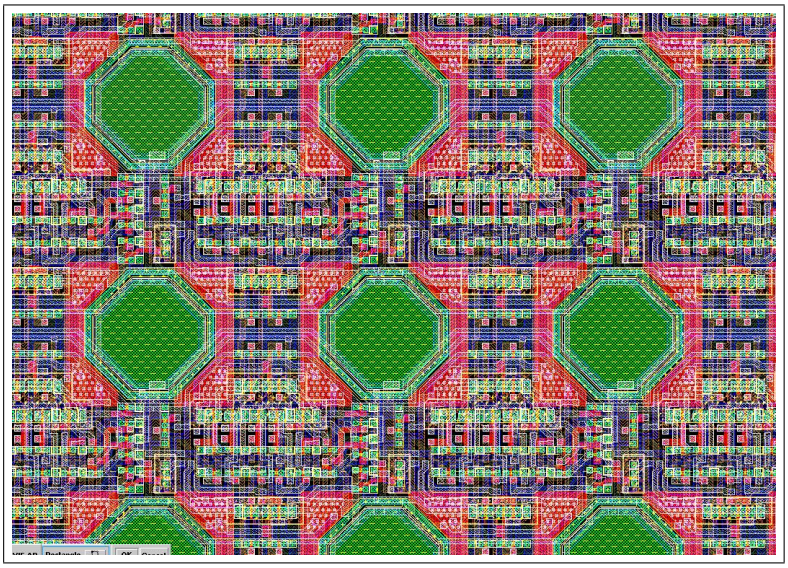


Fig. 5.1 – Dessin des masques de 6 pixels de HISIC.I

Nous attachons une grande importance à la réalisation matérielle des concepts que nous développons. Nous avons envoyé en fabrication deux prototypes de rétines CMOS. Chaque rétine de résolution 64×64 est fabriquée en technologie standard CMOS 0,35 μm. Le principal objectif de ce chapitre est de valider expérimentalement les concepts fondamentaux que nous avons développés durant cette thèse. Dans un premier temps nous décrivons la plate-forme expérimentale du premier circuit HISIC.I (cf. figure 5.2) et le système de vision HISIC.II. Puis une seconde section se consacre à la caractérisation du système à deux niveaux de l'architecture : le niveau capteur et le niveau pixel. Puis, une section se propose d'exposer des résultats de traitements d'images allant de la simple capture d'images haute cadence à la détection de mouvements en passant par le gradient spatial. Enfin, une comparaison entre les résultats de traitements obtenus à partir de notre rétine et les résultats obtenus à partir d'une unité de calcul numérique permet de conclure ce dernier chapitre.

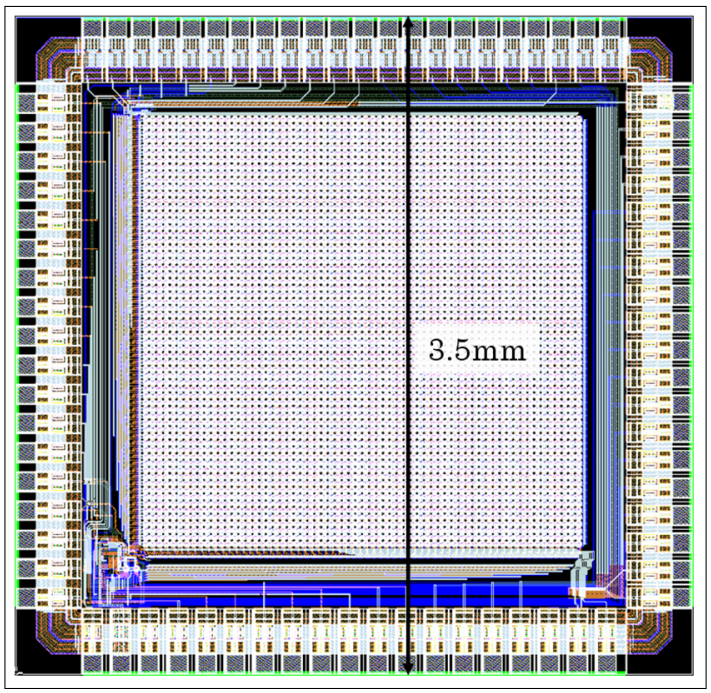


FIG. 5.2 – Dessin des masques du circuit HISIC.I

5.2 Projet HISIC

Le projet HISIC a pour ambition de réaliser un démonstrateur pré-industriel de caméra rapide et "intelligente". Le premier circuit a servi principalement à la caractérisation du pixel. Le second a permis la réalisation d'un prototype de système de vision dédié à l'imagerie rapide et aux traitements d'images linéaires.

Les circuits HISIC (High Speed Image Capture) sont des prototypes de rétines CMOS artificielles de résolution 64×64. La seconde version inclut un convertisseur analogique numérique. Chaque circuit HISIC comporte 160 000 transistors sur une surface de 3,675×3,775 mm^2. La première version contient des structures de test utilisées pour caractériser et comparer les performances de photodétecteurs.

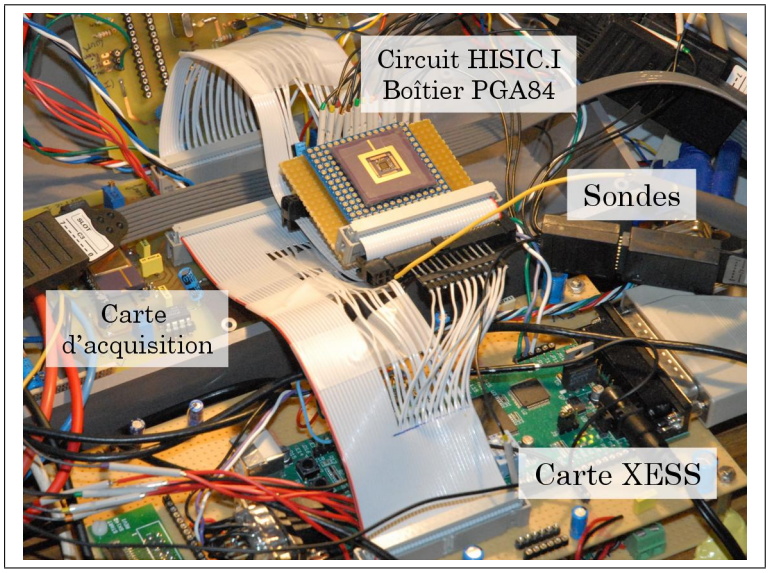
5.2.1 Plate-forme expérimentale d'instrumentation et de caractérisation du circuit HISIC.I

Ce premier run (fabrication du circuit) de la rétine a révélé des défauts technologiques auxquels il a fallu palier en réalisant les modifications qui s'imposent. En effet, dans cette rétine, nous avons pris le risque d'utiliser la couche de métal 4 pour réaliser le routage du pixel (38 transistors dans 906 μm^2), au lieu de l'utiliser pour protéger le circuit du rayonnement électromagnétique, ce qui a donc rendu susceptible le circuit aux perturbations électromagnétiques rayonnantes. De plus des problèmes liés au plan de masse dégradaient le fonctionnement des décodeurs d'adresses.

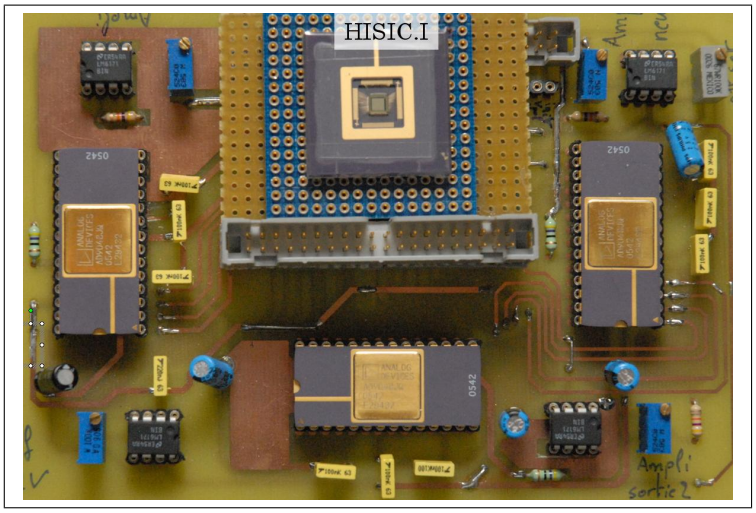
La plate-forme expérimentale, qui a servi à trouver les origines des dysfonctionnements de HISIC.I, est illustrée par la figure 5.3. La carte d'acquisition est composée de trois amplificateurs d'instrumentation à large bande (National Semiconductor LM6171, produit gain-bande de 100 MHz) et de trois convertisseurs analogiques numériques, d'Analog Devices (AD9048 à 40 Mo/s), pour assurer la conversion des bus asynchrones analogiques de HISIC.I.

5.2.1.1 Erreurs d'antenne

Nous avons corrigé des erreurs d'antenne le jour même du run de HISIC.I, à l'aide de diffusions N^+, de telle manière que le circuit puisse juste être envoyé en fabrication. Mais ces diffusions n'ont pas été disposées uniformément sur les lignes d'adresses, ce qui peut être à l'origine de certains dysfonctionnements. Les erreurs d'antenne ont des conséquences irrémédiables lors



(a) Plateforme expérimentale du circuit HISIC.I



(b) HISIC.I et sa carte d'acquisition

FIG. 5.3 – HISIC.I et son interface pour l'instrumentation

de la phase de dopage par implantation ionique qui consiste à accélérer des impuretés ionisées avec un champ électrique, afin de leur conférer l'énergie nécessaire pour rentrer dans le matériau à doper. Durant cette phase de fabrication, le potentiel de grille augmente de telle sorte que le champ dans la couche d'oxyde dépasse le champ disruptif. Cela a pour effet la destruction irrémédiable du transistor lors de la fabrication.

Ces erreurs peuvent être corrigées a l'aide de ponts (passage d'une couche

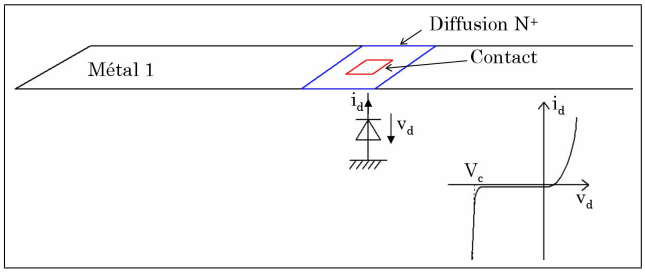


FIG. 5.4 – Correction des erreurs d'antenne

de métal à une autre de niveau différent, ex : métal1 - via1 - métal2) et de diffusion N^+. En particulier, des ponts sont systématiquement placés aux plus près de chacune des grilles des transistors de commande des décodeurs d'adresses (dans HISIC.II). Les diffusions N^+ servent à réaliser des diodes, dont l'anode est connectée au substrat et la cathode au métal considéré (cf. figure 5.4). Le courant de fuite de la diode permet d'éviter l'accumulation de charges lors de la phase d'implantation ionique dans le procédé de fabrication.

5.2.1.2 Effets d'antenne

Des effet d'antennes (différents des erreurs d'antenne) peuvent aussi être à l'origine de certains dysfonctionnements du circuit. Si une grande surface de métal, dans un circuit intégré, est connectée à une grille de transistor MOS, ce qui est souvent le cas dans la structure du décodeur d'adresse, alors toutes les sources de bruit (thermique, 1/f, grenaille, etc.) rayonnent au travers de cette section, comme le fait une antenne patch[1] dans un téléphone portable. Pour peu que le plan de masse ne soit pas suffisant pour découpler convenablement le circuit, ces effets d'antenne dégradent les performances du capteur.

Un défaut de plan de masse et l'utilisation de la dernière couche de métal (métal 4) pour réaliser une partie du routage du circuit, sont à l'origine des dysfonctionnements de HISIC.I. Cette dernière couche de métal joue à la fois le rôle de plan de masse et de réflecteur d'onde électromagnétique. Elle se comporte comme un miroir vis à vis du champ électromagnétique et rend moins susceptible le circuit aux perturbations électromagnétiques rayonnantes et conduites.

[1]Une antenne patch est une antenne plane de forme rectangulaire

5.2.1.3 Améliorations

Trois principales améliorations sont apportés à la conception de HISIC.II par rapport à HISIC.I :

1. Afin de réserver la couche de métal 4 à la réalisation d'un plan de masse, qui rende ainsi moins susceptible le système aux perturbations électromagnétiques conduites et rayonnées, la solution retenue est de diminuer la taille de la photodiode. En effet, en diminuant la taille de la photodiode, tout en maintenant les performances du photodétecteur, il reste plus de surface pour réaliser le routage du pixel. Le pixel fait 35 μm de côté avec un taux de remplissage de 12,7% soit une photodiode de 156 μm^2 de surface active et de 17,5 μm de côté. Avec une photodiode de 10 μm de côté, soit 70 μm^2 de surface active, et donc un taux de remplissage ramené a 6%, il reste en théorie suffisamment de surface pour router le pixel avec seulement 3 couches de métal.
2. Nous avons implémenté 2×64 bascules D synchronisées sur une même horloge externe au niveau du dernier étage du décodeur de lignes et du multiplexeur de colonnes afin de limiter la présence de perturbations conduites hautes fréquences.
3. Enfin, la partie résistive de l'impédance du multiplexeur analogique 64 vers 1 (R_{ON}) est trop élevée ($\approx$ 100 Ω), d'où une perte significative du signal à l'entrée du convertisseur analogique numérique. Le convertisseur se trouvait à l'exterieur du circuit HISIC.I, sur une carte d'acquisition (cf. figure 5.3(b)). Dans la deuxième version, nous avons implémenté un CAN au niveau des colonnes du capteur (cf. chapitre 4, section 4.3).

5.2.1.4 Exploitation

Cette plate-forme d'instrumentation nous a aussi permis de caractériser le pixel à l'aide de circuits tests placés dans le coin inférieur gauche de HISIC.I, en dehors de la matrice de pixels. A défaut de pouvoir mettre en œuvre la matrice de pixels, nous avons exploité les structures de test et cela pour deux raisons :

1. La caractérisation du pixel informe sur la quasi totalité des caractéristiques physiques du capteur, hormis sur le bruit spatial fixe.
2. Les résultats expérimentaux sur les structures de test permettent de valider les concepts théoriques mis en avant dans ces travaux et en particulier le choix d'une photodiode verticale de type N et de forme octogonale.

5.2.2 Système de vision artificielle HISIC.II

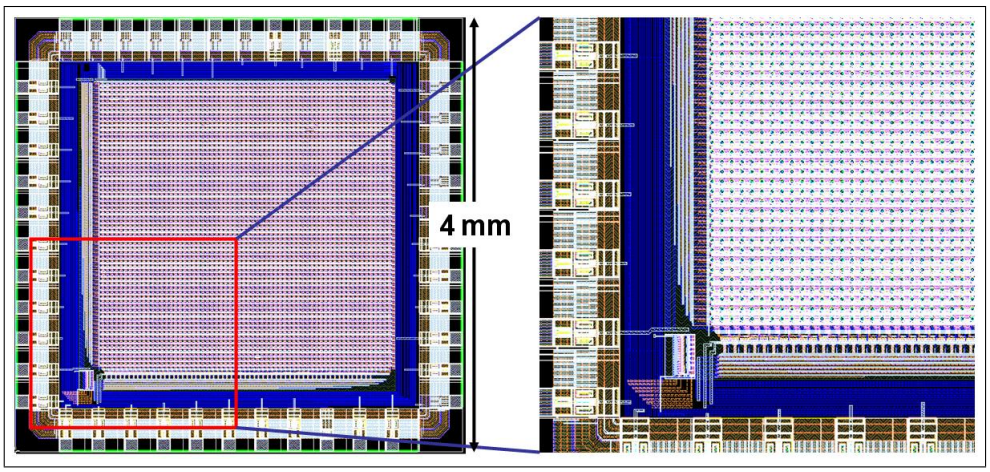


FIG. 5.5 – Dessin des masques de la rétine HISIC.II

Le dessin des masques du second circuit HISIC.II est illustré figure 5.5. Nous avons mis au point un prototype de caméra rapide et intégrant des traitements d'images, du moins une première étape avant d'envisager la réalisation d'un capteur HDV de vision artificielle. HISIC.II, ce sont 4096 processeurs arithmétiques analogiques interconnectés selon une grille 64×64 en topologie 4-connexe. Deux mémoires analogiques avec amplificateur sont implémentées dans chaque pixel de la matrice. Ce capteur, gravé en technologie CMOS 0,35 μm (double-poly, quadruple-métal), 3,3 V, permet de faire des acquisitions à haute cadence allant jusqu'à 10 000 images par seconde et d'exécuter des traitements locaux spatiaux et spatio-temporels sur les images acquises avec plusieurs niveaux d'itérations, jusqu'a 400 itérations 4-connexes à la cadence vidéo.

Ce système embarqué constitué de la rétine CMOS HISIC.II et d'une carte à FPGA est un premier résultat. La figure 5.6 est un synoptique de la plateforme expérimentale présentée avec la carte de contrôle et la rétine. Un port VGA permet d'afficher directement les résultats des séquences d'acquisition sur un moniteur.

Le circuit HISIC.II est interfacé à une carte industrielle à composant programmable (ref : XESS 3S1000). Cette carte propriétaire est constituée d'un FPGA, le Spartan-3 (1 million de portes, 100 MHz), d'une mémoire vive SDRAM[2] de 32 Mo, et d'une mémoire FLASH de 2 Mo pour stoker les don-

[2]Synchronous Dynamic Random Access Memory ou Mémoire Dynamique Synchrone à Accès Aléatoire

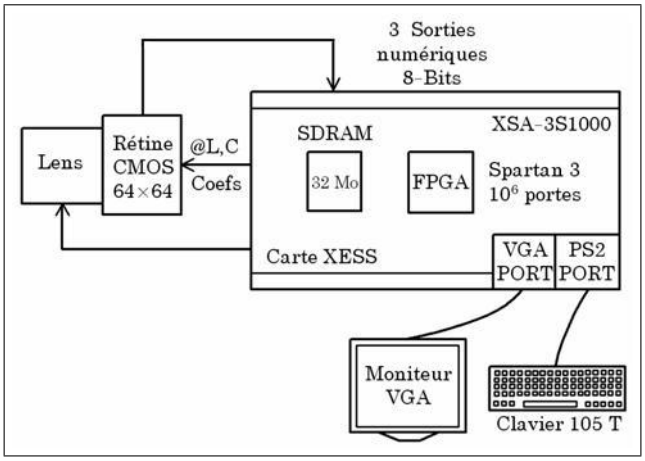


FIG. 5.6 – Schéma bloc de la platforme expérimentale

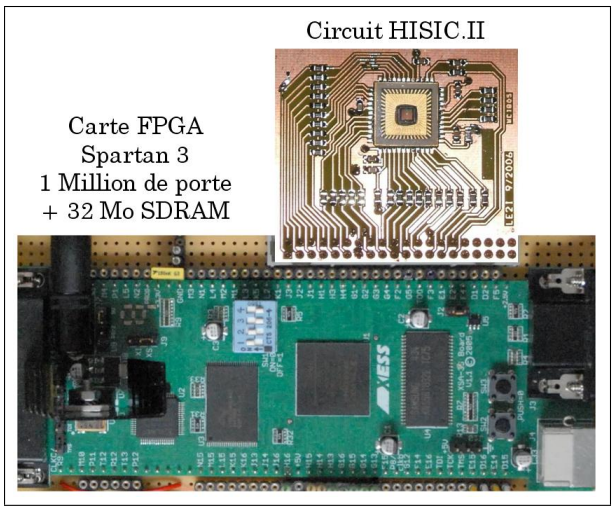


FIG. 5.7 – Illustration photographique de l'ensemble des circuits du système

nées des fichiers de configurations du FPGA. Une mémoire EEPROM[3] série I2C est également interfacée à la carte XESS, pour stocker les paramètres de calibration de la caméra. La figure 5.7 donne une illustration photographique de l'ensemble (sans la mémoire I2C). Les principales caractéristiques de HISIC.II sont regroupées dans le tableau 5.1.

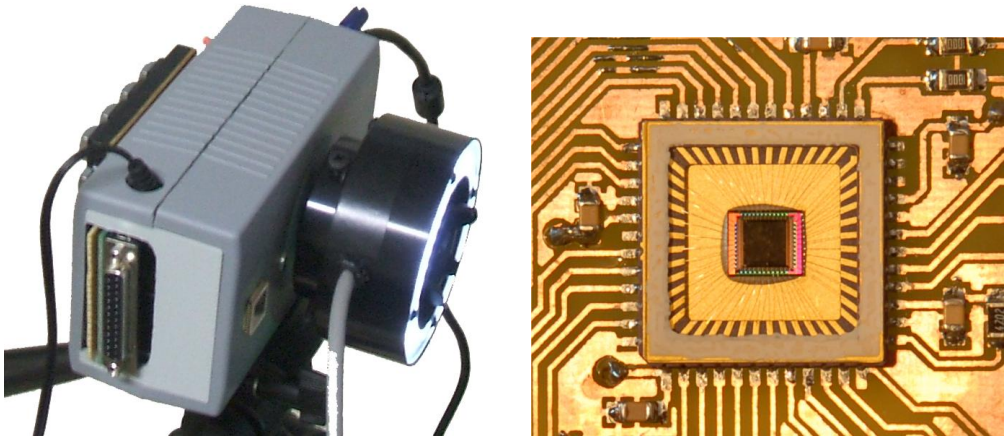
Le prototype expérimental est constitué du circuit HISIC.II, d'une carte à composant programmable, d'une optique à focale variable et d'une interface homme-machine. L'ensemble, mis sous boîtier plastique, est illustré, fi-

[3]Electrically Erasable Programmable Read-Only Memory ou mémoire morte effaçable électriquement et programmable

TAB. 5.1 – Caractéristiques du circuit HISIC.II

Technologie	0.35μm 2-poly 4-metal CMOS
Résolution	64×64
Surface	$\approx 11\ mm^2$
Nombre de transistors	160 000
Nombre de transistors / pixel	38
Tailles des pixels	35 μm $\times$ 35 μm
Taux de remplissage	10%
Consommation	110 mW (27 μW/pixel)
Tension d'alimentation	3.3 V

gure 5.8(a).



(a) (b)

FIG. 5.8 – (a) Illustration photographique du prototype de caméra rapide et intelligente, (b) Circuit HISIC.II, sous boîtier LCC48.

La figure 5.8(b) est une illustration photographique du circuit HISIC.II. La rétine est encapsulé sous boîtier CLCC 48 pour former un composant CMS, lui-même soudé sur une carte époxy pour rendre moins susceptible le circuit aux perturbations électromagnétiques rayonnantes, grâce au plan de masse de la carte et pour faciliter le blocage mécanique en position de l'optique sur le plan focale image, en vue de réaliser un démonstrateur de

système de vision artificielle.

5.3 Caractérisation du système

Nous avons réalisé des mesures sur les deux circuits HISIC et nous avons caractérisé complètement le capteur et son pixel, en termes de réponse spectrale, de sensibilité, de bruit temporel, de bruit fixe, de linéarité, de sensibilité, de consommation et de dynamique.

5.3.1 A l'échelle du pixel

A partir des structures de tests, placées dans le coin inférieur gauche de HISIC.I, nous avons évalué quantitativement le pixel du capteur et validé expérimentalement certains concepts théoriques développés plus en amont dans ce mémoire.

5.3.1.1 Validation expérimentale de la forme octogonale

L'objectif principal de la conception de HISIC.I était de pouvoir caractériser le plus précisément possible les structures électroniques du pixel qui font appel pour la plupart à des solutions innovantes et donc non éprouvées. Dans HISIC.I, nous avons implémenté des structures de test monolithiques, notamment deux pixels tests, dont un pixel constitué d'une photodiode carrée, et un pixel constitué d'une photodiode octogonale, dans le coin inférieur gauche de la rétine (cf. figure 5.9). Nous avons pu ainsi comparer les réponses spectrales afin de valider un choix technologique.

En première approximation, on peut considérer que l'intégration du photo-courant est linéaire. Ceci est surtout vrai pour de faibles temps d'intégration. On peut donc appliquer la relation : $\Delta V \approx (I_{ph}/C_{int})\,\Delta T$. La fraction rationnelle $(\Delta V/\Delta T)$ est mesurable à l'aide d'un simple oscilloscope. De plus, nous pouvons écrire que $(\Delta V/\Delta T) = (I_{ph}/C_{int})$ est une grandeur qui ne dépend que de la longueur d'onde λ. A partir des mesures de $(\Delta V/\Delta T)$ pour 15 longueurs d'onde dans l'intervalle [400 nm ; 1100 nm], on déduit les valeurs de photocourant à un éclairement donné.

Par exemple, considérons le cas simple de trois mesures réalisées aux longueurs d'onde λ_1, λ_2 et λ_3. Pour chacune de ces longueurs d'onde, on mesure les rapports $(\Delta V/\Delta T)$. Or on sait que $(\Delta V/\Delta T) = (I_{ph}/C_{int})$ et que la capacité de conversion au nœud flottant est invariante en fonction de la longueur d'onde. Par conséquent, il vient le système d'équations :

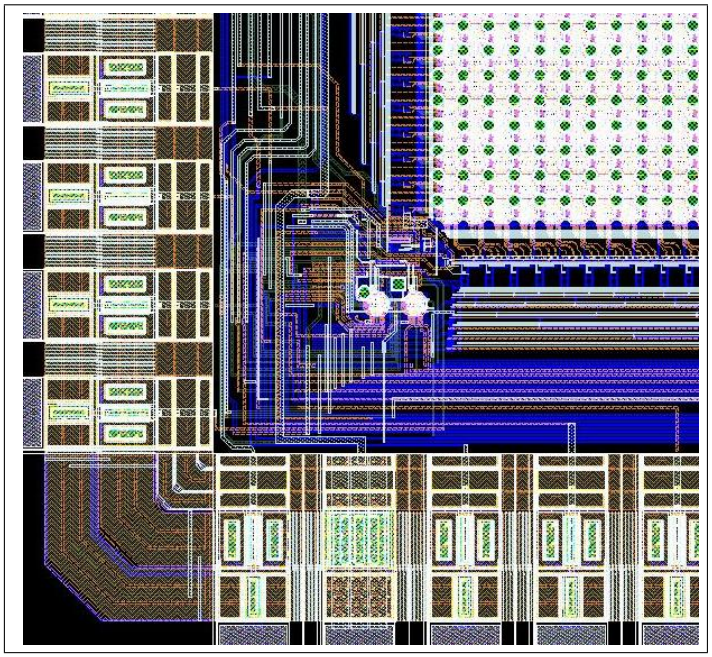


FIG. 5.9 – Dessin des masques des pixels tests de HISIC.I

$$\begin{cases} \frac{I_{ph}(\lambda_1)}{I_{ph}(\lambda_2)} = \alpha \\ \frac{I_{ph}(\lambda_2)}{I_{ph}(\lambda_3)} = \beta \\ \frac{I_{ph}(\lambda_1)}{I_{ph}(\lambda_3)} = \delta \end{cases} \tag{5.1}$$

Dans ce système, les inconnues sont $I_{ph}(\lambda_1)$, $I_{ph}(\lambda_2)$ et $I_{ph}(\lambda_3)$. Dans cet exemple, on substitue $I_{ph}(\lambda_1)$ de la première égalité dans la troisième afin d'exprimer $I_{ph}(\lambda_2)$ en fonction de $I_{ph}(\lambda_3)$ et de reporter cette expression dans la deuxième égalité pour finalement déduire la valeur de $I_{ph}(\lambda_3)$, puis celle de $I_{ph}(\lambda_2)$ et enfin celle de $I_{ph}(\lambda_1)$. Pour résoudre les systèmes de ce type comportant plus d'inconnues, 15 inconnues dans notre cas, on a avantage à utiliser un logiciel matriciel comme MATLAB©.

Les réponses spectrales des photodiodes carrées et octogonales sont données en A/W sur la figure 5.10. Les mesures ont été réalisées à l'aide d'un monochromateur à diodes électroluminescentes dans l'intervalle de longueurs d'onde allant de 400 nm à 1100 nm. L'expérience montre que la structure octogonale offre de meilleurs performances que la structure carrée et ce pour pratiquement toutes les longueurs d'onde.

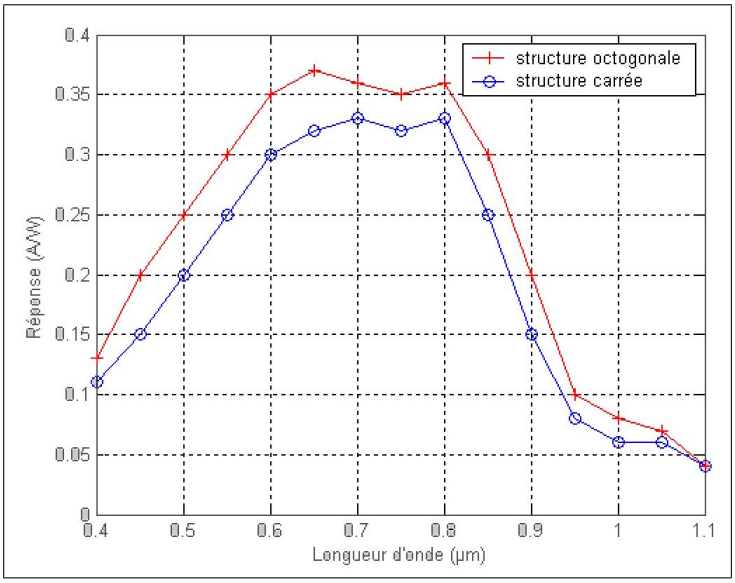


FIG. 5.10 – Réponse spectrale en A/W des photodiodes aux formes carrées et octogonales

De ces mesures, on peut déduire la capacité de conversion au nœud flottant ($C_{int} = I_{ph}\Delta T/\Delta V \approx 9,5\,fF$). Compte tenu de ces résultats et de l'étude théorique du chapitre 4, nous sommes en mesure de confirmer qu'une photodiode verticale de type N et de forme octogonale est plus efficace qu'une photodiode verticale de type N et de forme carrée à même surface active.

Nous avons aussi réalisé d'autres mesures comparatives à l'aide du circuit HISIC.II qui ont conduit à la conclusion suivante : polarisée en son centre à l'aide d'un segment de métal et de neuf contacts, une photodiode verticale de type N et de forme octogonale voit son efficacité augmenter de 10% par rapport à la même photodiode sans polarisation cathodique centrale.

5.3.1.2 Circuit d'acquisition et cellules MA^{2}M

Quinze transistors assurent les fonctions d'acquisition, de mémoires et d'amplifications au plus près du photo-détecteur. Nous avons caractérisé complètement l'étage d'acquisition et les deux cellules MA^{2}M à partir des mesures que nous avons faites sur le pixel test octogonal d'HISIC.I. La figure 5.11 illustre les résultats expérimentaux obtenus à l'aide d'un oscilloscope à mémoire (l'Agilent DSO5054A 2 GSs/s). On peut y voir évoluer, en fonction du temps, tous les signaux caractéristiques du pixel : "reset, lecture1, lecture2, sortie1, sortie2 et nœud". Le signal "sortie suiveur" est un signal

intermédiaire, il s'agit du potentiel de sortie de l'étage suiveur qui précède les cellules MA^{2}M (cf. chapitre 4, section 4.2.2.2, figure 4.11).

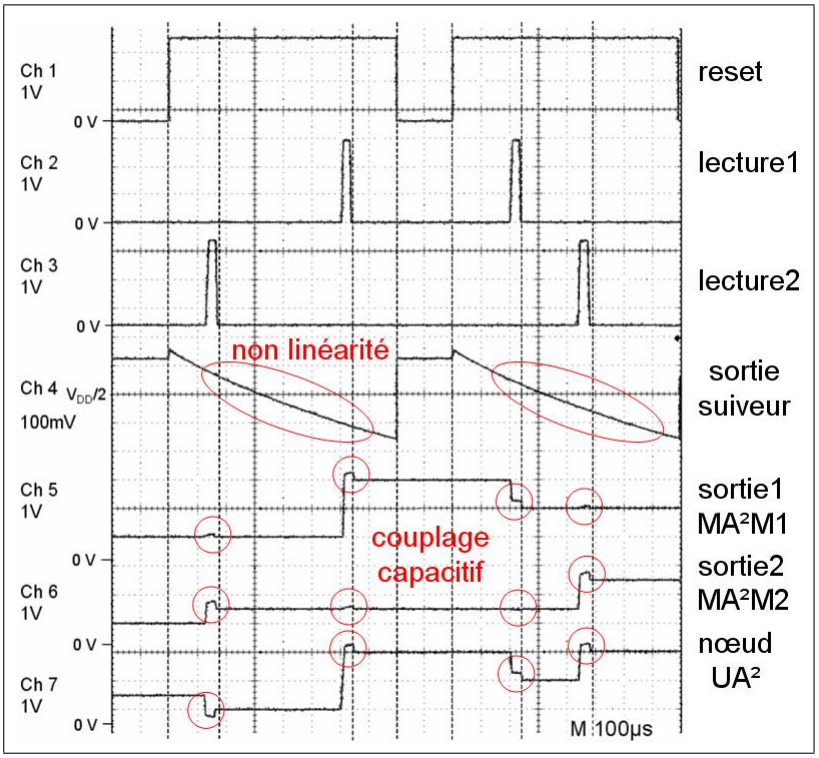


FIG. 5.11 – Graphe des signaux caractéristiques du pixel test octogonal (HISIC.I) avec un temps d'intégration de 100μs, pour une longueur d'onde de 680 nm et un éclairement de 5000 lux.

Le signal "reset", actif à l'état bas (V$_{GND}$=0V), sert à fixer le potentiel cathodique à approximativement $+V_{DD}$ (présence d'une faible tension de déchet) durant les périodes de pré-chargement et à ouvrir ce circuit de pré-chargement lors de la phase d'intégration du photo-courant. Les signaux "lecture1" et "lecture2", actifs à l'état haut (V_{DD}), sont mis à contribution, à l'issue du pré-chargement, pour recopier le niveau de tension du signal "sortie suiveur", image de l'éclairement incident du pixel, dans la cellule MA^{2}M1 (lecture1) ou MA^{2}M2 (lecture2). En agissant sur la tension de polarisation Vpol et sur le temps d'intégration (temps séparant la fin de la phase de pré-chargement et l'activation du signal "lecture1" ou "lecture2"), on peut avec précision fixer la plage de tension des signaux de sortie.

Dans l'expérience présentée, les signaux "lecture1" et "lecture2" sont des signaux numériques impulsionnels et activés durant la phase d'intégration du photo-courant. La tension de polarisation Vpol est fixée à 1,35 Volt. La

sortie "nœud", signal issu de l'unité arithmétique analogique, est la différence entre le signal "sortie2" et le signal "sortie1". Ce signal est obtenu en fixant les coefficients aux valeurs suivantes : $coef1 = V_{DD}, coef2 = 0V$ et $coef3 = coef4 = V_{DD}/2$. Ces premiers résultats permettent d'apprécier le comportement dynamique de l'UA[2].

On observe des effets de couplage sur les sorties (représentés par de petits cercles rouge sur la figure) très probablement dûs à la présence de capacités parasites, elles mêmes induites par le routage du circuit. En effet, lorsque deux segments de métal de niveaux différents se croisent, ils sont couplés par une capacité dite de couplage. Cette capacité de couplage a deux composantes géométriques, une composante surfacique et une composante périmétrique (cf. figure 5.12). D'après les données du constructeur AMS, en technologie CMOS 0,35 μm, la capacité surfacique Cs métal2-métal1 est typiquement de 0,036 fF.μm^{-2} et la capacité périmétrique Cp est de 0,048 fF.μm^{-1}. Dans cet exemple, la capacité parasite entre deux segments de métal1 et metal2, qui se croisent perpendiculairement, est Ct=Cp×2,2 μm+Cs×0,03 μm^2 ≈0,1 fF. Les effets de ces capacités parasites, combinés aux capacités grille-canal des transistors, apparaissent durant les changements de niveau sur les signaux "lecture1" et "lecture2" et ne sont pas gênants sur un plan strictement fonctionnel.

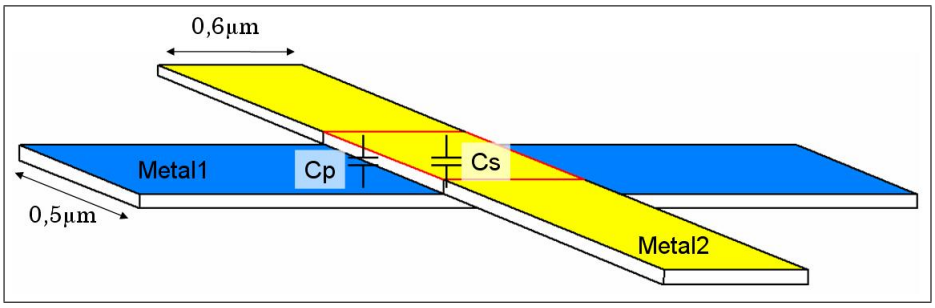


FIG. 5.12 – Couplage capacitif entre deux segments de métal de niveau différent et qui se croisent.

D'après le graphe 5.11, l'intégration présente une pente de 200 mV pour 100 μs, avec un éclairement de 5000 lux au plan focal image soit une sensibilité S donnée par l'expression :

$$S = \frac{\Delta V}{E\,\Delta T} = 0,4\,V.lux^{-1}.s^{-1} \tag{5.2}$$

De plus cette intégration n'est pas tout à fait linéaire, même à 10 000 images/s. Les disparités de linéarité sur les sorties "sortie1" et "sortie2" sont

dues aux défauts de linéarité de l'étage suiveur (représentées par des ellipses rouge sur la figure) et au défauts de linéarité des amplificateurs. Ces disparités de linéarité sont de 4,3% sur une plage de tension de 2 V (1 V RMS de part et d'autre de $V_{DD}/2$ soit $[0.65V; 2,65V]$). Le gain de conversion est de 17 μV/e$^-$ RMS ($C_{int} \approx 9,5\,fF$), obtenu grâce à une photodiode de forme octogonale et polarisée en son centre (voir second chapitre, section 4.2.1.3). Le tableau 5.2 compare ces résultats avec les capteurs de Yang, Yadid-Pecht et Sugiyama. Ces capteurs sont tous réalisés en technologie CMOS et dans chacun de ces capteurs, le photodétecteur est une photodiode de type N.

TAB. 5.2 – Comparaisons avec quelques capteurs d'images

Capteur	HISIC.II	[YEGFT99a]	[YPB03]	[STT+05]
Technologie (μm)	0,35	0,35	0,5	0,6
Taille des pixels (μm^2)	35×35	10,5×10,5	14,4×14,4	20×20
Transistors/pixel	38 T	6 T	4 T	5 T
Taux de remplissage	10%	29%	37%	44%
Gain de conversion (μV/e$^-$)	17	5,1	12	10,7
Sensibilité (V.lux^{-1}.s^{-1})	0,4	0,15	0,32	0,25

Le gain en tension de l'étage amplificateur de la structure MA^{2}M est $Av = 12$ sur la plage $[0.65V; 2,65V]$. Enfin nous avons évalué la dynamique maximale du capteur avec des temps d'intégration allant de 100 μs à 40 ms. Il est sensible à plus de 3 décades d'éclairement et plus précisément sa dynamique est de 68 dB.

La consommation d'énergie électrique est de 27 μW/pixel, ce qui donne 110 mW pour la matrice de pixel. Le convertisseur analogique numérique est assez gourmand en énergie : 8 mW/comparateur soit 512 mW. La puissance dissipée par le convertisseur est prépondérante devant celle dissipée par la matrice de pixel. La puissance totale est en première approximation donnée par l'équation suivante :

$$\forall\, n \in \mathbb{N}, \quad P_{tot} = n^2\, P_{pix} + n\, P_{comp} \tag{5.3}$$

Cette équation met en lumière le fait qu'à un certain rang n, le terme $n^2\, P_{pix}$ devient prépondérant devant le terme $n\, P_{comp}$. Il vient le rang $n \geq P_{comp}/P_{pix} \approx 296$ à partir duquel l'énergie consommée par la matrice est supérieure à l'énergie consommée par le convertisseur. Cela signifie que cette

architecture a globalement un bilan d'énergie qui s'équilibre entre la matrice de pixels et le convertisseur pour des résolutions de capteurs à partir du format CIF (352×288).

5.3.1.3 Unité arithmétique analogique, multiplieurs quatre quadrants

D'après la figure 5.11, le temps de réponse de l'UA^2 est de l'ordre de grandeur du temps de réponse des cellules MA^{2}M, soit environ 500 ns. Le signal "nœud" représente la différence arithmétique entre les signaux "sortie1" et "sortie2". Le résultat est centré sur $V_{DD}/2$ avec une erreur relative de 11,5% sur la gamme de tension [0,65 V ; 2,65 V]. L'amplitude du signal V_i issue du pixel $_i$ est limitée à la plage de tension [0,65 V ; 2,65 V], ce qui est possible en agissant sur la tension de polarisation V_{pol} et sur le temps d'intégration de la structure MA^{2}M. Vue l'erreur relative sur la sortie, notre mutiplieur distingue seulement 8 niveaux (3 bits) centrés sur $V_{DD}/2$. Si le signal "nœud" est supérieur à $V_{DD}/2$, cela signifie que le résultat est positif, à contrario, que le résultat est négatif. Il faut ensuite interpréter les résultats issu de la conversion analogique numérique pour traduire mathématiquement l'opération "différence". Pour cela, il faut introduire un bit de signe qui est le résultat de la comparaison du signal "nœud" avec $V_{DD}/2$.

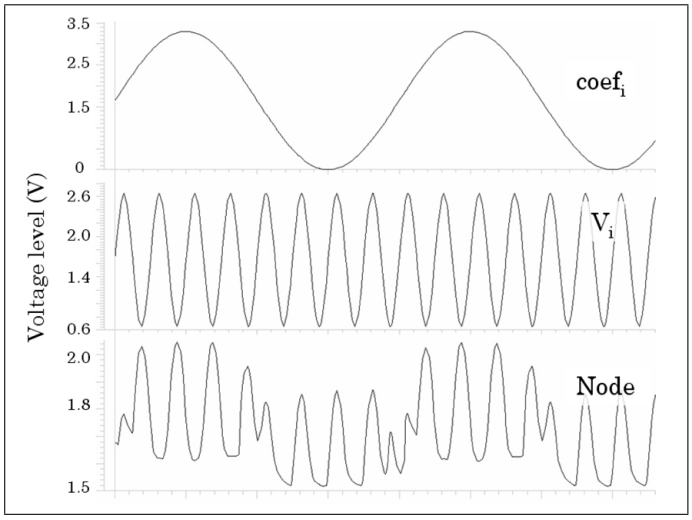
L'étude classique harmonique permet de mettre en exergue les défauts de notre multiplieur. La figure 5.13 illustre les résultats des mesures que nous avons réalisées sur la structure multiplicatrice élémentaire de l'UA^2 (cf. chapitre 4, section 4.2.2.3), excitée en régime harmonique. Des signaux sinusoïdaux, de fréquences de l'ordre du kHz (100μs/configuration), sont appliqués en entrée de la structure et nous mesurons les caractéristiques spectrales du port de sortie. Les stimulis harmoniques ont les caractéristiques suivantes :

$$\forall_i \in \{1;4\}, \begin{cases} coef_i = A.cos(2\pi f_1) \text{ avec, } f_1 = 2.5kHz \\ V_i = B.cos(2\pi f_2) \text{ avec, } f_2 = 20kHz \end{cases} \tag{5.4}$$

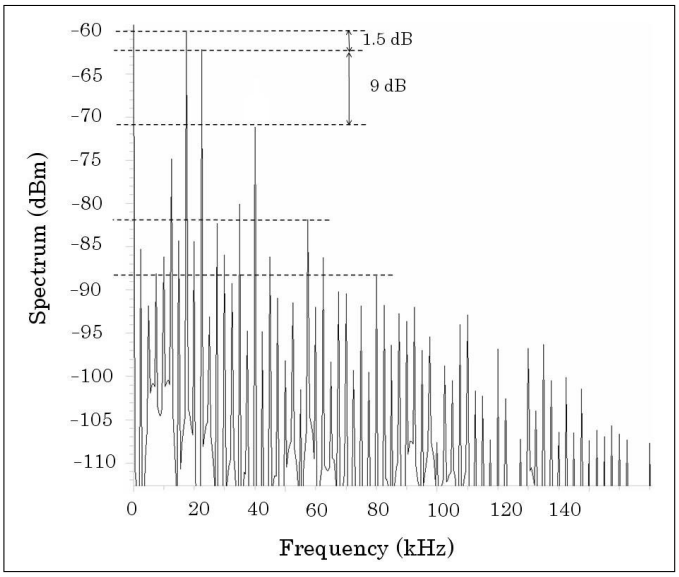
Dans le cas d'une structure idéale, les variations du niveau de signal du port de sortie seraient régies par l'équation :

$$P = \frac{A.B}{2}[cos(2\pi(f_2 - f_1)) + cos(2\pi(f_2 + f_1))] \tag{5.5}$$

Or le spectre du signal de sortie, représenté sur la figure 5.13(b), comporte de nombreuses imperfections. On retrouve bien les deux raies principales autour de la porteuse (f_2=20 kHz) de fréquences f_2-f_1=17.5 kHz et f_2+f_1=22.5 kHz, mais la structure présente plusieurs défauts d'isolation des



(a) Multiplication de signaux sinusoïdaux



(b) Spectre en fréquence du signal de sortie

FIG. 5.13 – Résultats expérimentaux du multiplieur quatre quadrants

ports d'entrée qui se traduisent par la présence de raies aux fréquences multiples des fréquences des signaux d'entrée. Notamment, la présence d'une raie à la fréquence $2\times f_2$=40 kHz, traduit la présence d'un défaut d'isolation d'ordre 2. Ces défauts d'isolation et produits d'intermodulation sont de plus faibles amplitudes que celles des deux principales raies (cf. figure 5.13(b)). En effet, la raie à 40 kHz a un niveau de 9 dB au dessous de la plus faible des deux principales raies, soit une atténuation en puissance dans un rap-

port de 8. Par conséquent, les défauts d'intermodulation et d'isolation des ports d'entrée de la structure traduisent l'erreur relative sur niveaux évoquée ci-avant et complète notre approche expérimentale.

5.3.2 A l'échelle du capteur

A l'échelle du capteur, on peut quantifier avec précision l'amplitude de l'écart type du bruit spatial fixe de la matrice et apprécier les performances du convertisseur analogique numérique. Le tableau 5.3 récapitule les différentes mesures que nous avons réalisées sur les rétines HISIC.

TAB. 5.3 – Récapitulatif des mesures expérimentales sur le pixel test octogonal (HISIC.I) et sur la rétine HISIC.II, avec un temps d'intégration de 100 μs, pour une longueur d'onde de 680 nm et un éclairement de 5000 lux.

Gain de conversion	17 μV/e$^-$ RMS
Sensibilité	0,4 $V.lux^{-1}.s^{-1}$
Bruit spatial fixe (noir)	19,2 mV RMS
Bruit temporel	8,2 mV RMS
Disparité de linéarité	4,3%
Erreur relative UA^2	11,5%
Gain en tension	12
Dynamique maximale	68 dB

Les mesures du gain de conversion, de la sensibilité, de linéarité, de gain en tension et de dynamique sont réalisées à l'aide du pixel test octogonal du circuit HISIC.I (cf. section 5.3.1.3). Les mesures de bruit spatial fixe et de bruit temporel sont réalisées à l'aide du deuxième circuit.

5.3.2.1 Bruit spatial fixe et bruit temporel

La première caractéristique physique à mesurer à l'échelle du capteur est l'amplitude de l'écart type du bruit spatial fixe, généralement dominant dans les capteurs de type APS. Le bruit spatial fixe dépend des non uniformités des paramètres des transistors (tension de seuil, courant de saturation, etc.). Dans un capteur d'image dont le photodétecteur est une photodiode, ce bruit fixe trouve aussi sa source au travers des disparités des courants d'obscurité, c'est pourquoi on le mesure aussi en l'absence de lumière.

Pour obtenir le bruit spatial fixe, on mesure l'écart type du niveau de tension sur la première sortie analogique (sortie1) avec une image de noir. La valeur efficace du bruit fixe d'une image de noir est mesurée sur HISIC.II à partir d'une acquisition en l'absence de lumière avec un temps d'intégration de 100 μs. La mesure, en sortie de l'amplificateur (gain=12), donne 19,2 mV RMS (au sens de la valeur efficace du signal étudié) avec une polarisation inverse de 3,27V. En effet, la tension de déchet du transistor PMOS est de 0,03 V et la tension d'alimentation de 3,3 V, donc la photodiode est polarisée en inverse avec une tension V_d=3,27 V. Cette performance est comparable au capteur d'Alexander [KN03] qui présente un écart type de 2 mV avant amplification et dans la même technologie.

Le niveau de bruit temporel est mesuré après avoir stocké une image de blanc en SDRAM. L'image de blanc stockée en SDRAM, sous forme d'une matrice 64×64 codée sur 8 bits, est la cartographie du bruit spatial fixe du capteur. Cette matrice de bruit spatial fixe est ensuite retranchée aux valeurs des pixels correspondants. Or les sources de bruit fixe et temporel sont décorrélées du signal utile, donc en retranchant le bruit fixe, il reste le bruit additif temporel qui contribue à des variations aléatoires sur les valeurs du signal utile. Nous avons donc mesuré les contributions de deuxième ordre (écart type) des sources de bruit temporel. L'écart type du niveau de tension sur l'une des sorties analogiques (après amplification) de HISIC.II est de 8,2 mV RMS. En théorie, d'après la valeur de la capacité de conversion au nœud flottant et d'après la valeur du gain de l'amplificateur, le bruit de reset est de $G\sqrt{kT/C} = 7,7\,mV$. Le différentiel est de 0,5 mV, soit 6% de la valeur mesurée. Outre l'erreur engendrée par la mesure elle-même, le différentiel s'explique aussi par la présence d'autres sources de bruit (grenaille, 1/f, génération/recombinaison, cf. annexe **??**) et par le bruit thermique généré par les cellules MA^2M.

Ces sources de bruit peuvent être compensées directement au niveau du pixel à l'aide des cellules MA^2M en utilisant la technique CDS.

5.3.2.2 Compensation du bruit par CDS

La méthode de lecture à double échantillonnage corrélé (CDS) a trois objectifs.

1. Supprimer le bruit thermique du noeud réalisant la conversion charge tension dans le pixel (souvent dénommée "diffusion flottante").
2. Supprimer le bruit spatial fixe dit FPN (Fixe Pattern Noise) dû aux non-uniformités des transistors et des amplificateurs internes aux pixels.

3. Réduire le bruit basse fréquence engendré par les transistors MOS utilisés pour la lecture ou le transfert du signal électrique.

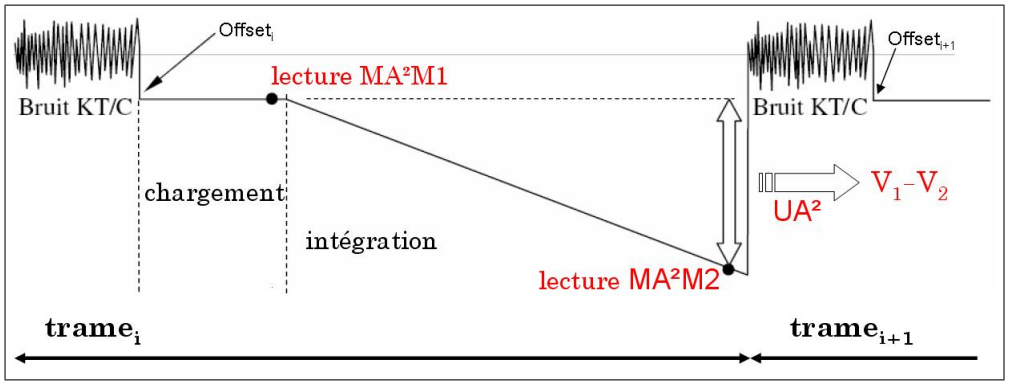
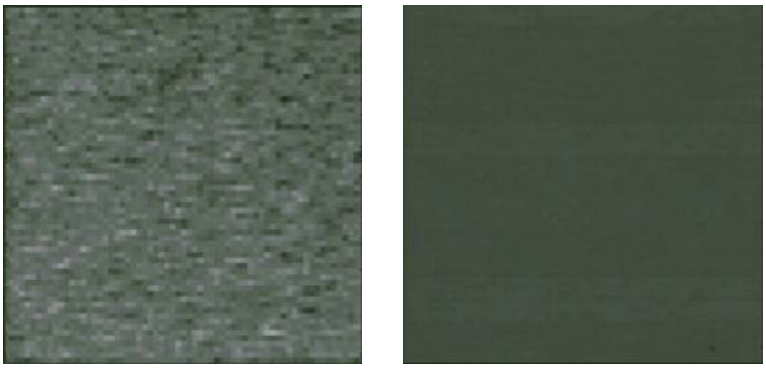


FIG. 5.14 – Correction du bruit à l'aide des cellules MA^{2}M et de l'UA^2

Les cellules MA^{2}M, implantées au sein du pixel, permettent de mettre en oeuvre très facilement ce procédé (cf. figure 5.14). Une première lecture du pixel, à l'issue du pré-chargement, donne le niveau de potentiel V1 et une seconde à la fin de l'intégration donne le niveau V2. La différence arithmétique entre les deux niveaux de tension V_1 et V_2 compense les effets des sources de bruits fixes et temporels. Cette opération est réalisée par l'unité arithmétique analogique (UA^2) implantée au sein du pixel.



(a) sans correction (b) avec CDS

FIG. 5.15 – Acquisition d'image en l'absence de lumière avec un temps d'intégration de 100 μs, (a) image de bruit fixe et temporel sans correction, (b) image corrigée par CDS.

Le double échantillonnage corrélé est une technique simple, qui ne nécessite pas plusieurs intégrations pour améliorer considérablement la qualité de

l'image acquise. La figure 5.15 montre l'effet de cette technique sur l'acquisition d'une image en l'absence de lumière avec un temps d'intégration de 100 μs. La valeur efficace, après correction, du bruit spatial fixe est alors de 1,13 mV RMS et celle du bruit temporel de 376 μV RMS (cf. figure 5.15(b)). Avec cette technique, les sources de bruits temporels, principalement le bruit de lecture $\sqrt{kT/C}$ qui se superpose au potentiel de pré-chargement (voir chapitre 1, section 2.8.5), et de bruit fixe sont compensées d'un facteur 17. Néanmoins, la mise en œuvre du CDS utilise toutes les ressources mémoires du pixel et ne permet pas d'appliquer la stratégie Ping-Pong pour optimiser les acquisitions à haute cadence.

5.3.2.3 Convertisseur analogique numérique

Nous avons implémenté un convertisseur analogique numérique dans le circuit HISIC.II au moyen de 64 comparateurs, d'un CNA 8 bits et d'un contrôle numérique externe à la rétine (cf. chapitre 4, section 4.3). L'entrée commune des comparateurs est connectée à la sortie du réseau R/2R (CNA) qui fixe le niveau de référence à comparer en parallèle avec les niveaux de sortie d'une ligne de 64 pixels.

L'expérience consiste à mesurer le temps de réponse d'un des comparateurs de la rétine, supposés tous appariés. Le plan focal est éclairé uniformément à l'aide d'une DEL blanche haute luminosité. Les adresses du décodeur de lignes et du multiplexeur de colonnes sont maintenues à une valeur fixe arbitraire. Le temps d'intégration est variable et vaut au minimum 100 μs. Pour cette expérience, le FPGA est programmé de telle sorte qu'il incrémente un registre sur 8 bits toutes les 12 ns. Ce registre est affecté à un port parallèle 8 bits lui-même connecté au CNA de la rétine. Le CNA produit en sortie un signal rampe de période 3 μs. On observe en concordance de temps le signal en sortie du comparateur et le signal analogique de sortie de la rétine à l'aide d'un oscilloscope à mémoire pour mesurer le temps de réponse. La figure 5.16 donne les résultats de mesures sur le temps de réponse du comparateur en fonction de la tension différentielle d'entrée.

Le temps de réponse est inférieur à 4 ns (250 millions de comparaisons par seconde) pour une tension différentielle d'entrée supérieure à 1/2 LSB, soit $3,3\,V/2^9 \approx 6,5\,mV$. La puissance dissipée par la structure est alors de 8 mW sous une tension d'alimentation de 3,3 V. Ces résultats sont strictement inclus dans le gabarit que nous nous étions fixés au départ : un temps de réponse d'au plus 6 ns pour un seuil différentiel de détection inférieur à 1/2 LSB de V_{DD}. Pour caractériser complètement le convertisseur analogique numérique, il reste à mesurer les erreurs différentielles (DNL) et intégrales (INL).

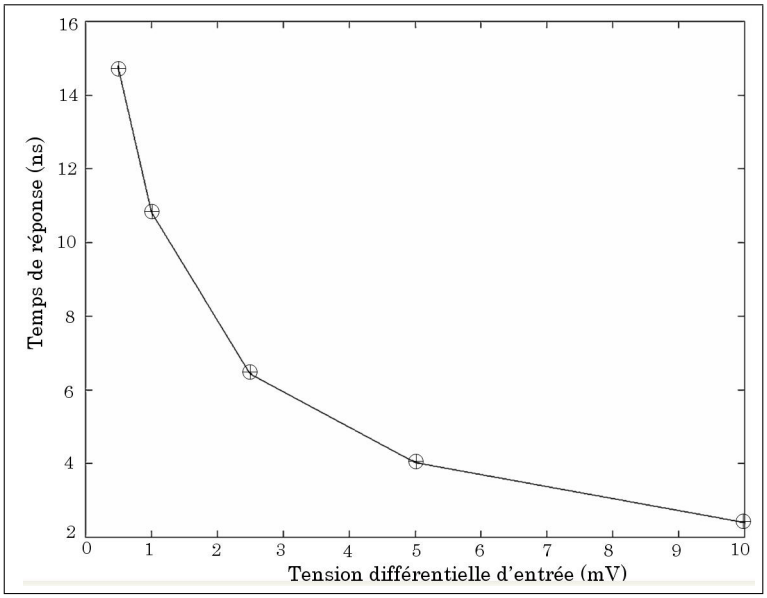


FIG. 5.16 – Résultats expérimentaux sur le temps de réponse du comparateur en fonction de la tension différentielle d'entrée

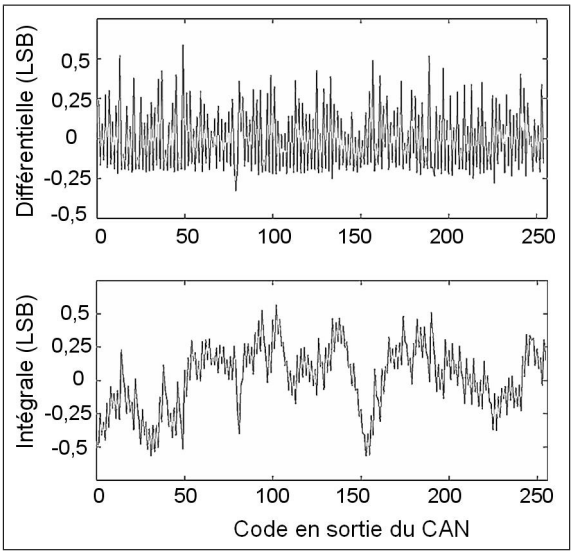


FIG. 5.17 – Erreurs différentielles et intégrales du convertisseur en fraction de LSB, en fonction de la sortie codée sur 8 bits.

La non-linéarité différentielle est la différence maximale entre deux seuils consécutifs sur la sortie numérique par rapport à l'entrée analogique. Pour mesurer cette erreur, le plan focal est éclairé uniformément à l'aide d'une DEL blanche à fort éclairement (50 lm). Nous fixons la tension de polarisation Vpol

à une valeur telle que la sortie analogique ait un niveau de tension proche de $V_{DD}/2$. Un registre de 16 bits, interne au FPGA, contient la valeur du temps d'intégration pour faire varier l'entrée analogique du CAN de 0 V à 3,3 V. Pour mesurer cette erreur, on peut utiliser une sonde numérique pour visualiser l'état du CAN. On mesure alors la différence sur l'entrée analogique entre chaque basculement du comparateur.

La non-linéarité intégrale (INL) est la différence maximale entre les seuils effectifs et théoriques sur la sortie numérique par rapport à l'entrée analogique. Les conditions expérimentales sont les mêmes que précédemment. Nous agissons sur le temps d'intégration pour faire varier le potentiel de sortie (entrée analogique du CAN) de 0 V à 3,3 V à l'aide du circuit FPGA. La valeur théorique est naturellement fournie par le CNA de la rétine. Tous les signaux étant synchronisés, on mesure en concordance de temps, à l'aide d'un oscilloscope numérique, la différence entre le potentiel du CNA et le potentiel de sortie analogique de la rétine.

Les mesures indiquent que les erreurs différentielles (DNL) et intégrales (INL) sont inférieures à 1/2 LSB en valeur absolue. La figure 5.17 illustre les graphes de mesures des erreurs différentielles et intégrales du convertisseur analogique numérique, implanté dans HISIC.II, en fonction de la sortie codée sur 8 bits.

TAB. 5.4 – Comparaisons de capteurs d'images

Capteur	HISIC.II	[CRVG+04]	[ASIS04]	[CMC+06]
Technologie (μm)	0,35	0,35	0,18	0,5
Résolution	64×64	128×128	640×480	90×90
DNL (LSB)	0,48	0,45	0,38	0,4
INL (LSB)	0,48	0,48	0,42	0,45

Le tableau 5.4 compare ces résultats avec les capteurs de Cembreno, Acosta et Chi qui intègrent tous trois un CAN au niveau des colonnes du capteur en technologie standard CMOS. La section qui suit présente des résultats d'acquisition et de traitement sur les images acquises.

5.4 Traitements d'images embarqués et exemples d'applications

Le système de vision artificielle HISIC permet d'exécuter des traitements d'images embarqués. Les possibilités de reconfigurations dynamiques de l'UA^2 conjuguées à la vitesse d'acquisition donnent au système HISIC la possibilité de réaliser des traitements d'images spatiaux et spatio-temporels.

5.4.1 Capture d'images à haute cadence

Les cadences atteintes par HISIC sont de l'ordre de la dizaine de milliers d'images par seconde. Cette partie du mémoire se propose d'exposer des résultats d'acquisitions à haute cadence.

5.4.1.1 Texture d'une goutte de lait

Nous avons choisi l'exemple d'une goutte de lait car le coefficient de viscosité est supérieur à celui de l'eau (ou du vin). En synthèse d'image, une texture est une région dans une image numérique qui a des caractéristiques homogènes. Ces caractéristiques sont par exemple un motif basique qui se répète ou des caractéristiques fréquentielles. Le film de la chute d'une goutte de lait est une série d'images peu texturées.

FIG. 5.18 – Plate-forme expérimentale d'acquisition d'une goutte de lait

Pour filmer une goutte de lait, le protocole expérimental est difficile à mettre en œuvre car il faut déterminer les valeurs des paramètres de l'expérience qui permettront d'apprécier la texture de l'objet filmé (tension de

polarisation, inclinaison du plan d'éclairement). Il faut aussi prévoir un système de goutte à goutte à débit variable. La figure 5.18 illustre la plate-forme expérimentale qui a permis d'obtenir ces résultats.

HISIC peut réaliser des captures d'images à très haute cadence, comme cette goutte de lait filmée à 2000 images par seconde (cf. figure 5.19). Le temps d'intégration est fixé à 500 μs. L'intégration et la lecture de la trame sont séparées à l'aide des deux cellules MA^{2}M pour que l'intégration dure toute la période de la trame, soit $1/2000^{eme}$ de seconde. L'éclairement au plan focal objet est de 2250 lux et le plan d'éclairement est incliné de 30° par rapport à la normale du plan objet pour mettre en contraste les ombres formées par les vagues.

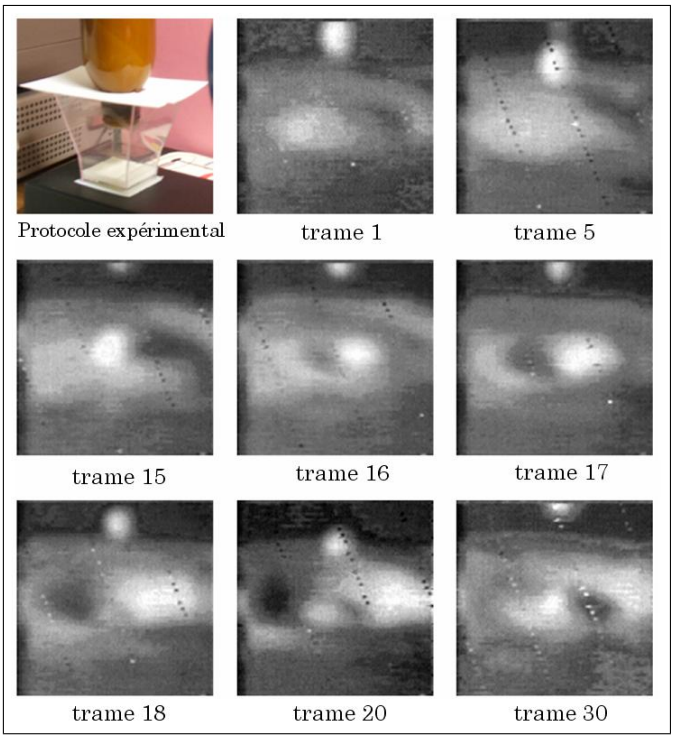


FIG. 5.19 – Résultats expérimentaux d'acquisitions haute cadence d'une goutte de lait

Dans cette expérience, nous n'avons pas mis en oeuvre le CDS pour ne pas utiliser toutes les ressources mémoires du pixel. Néanmoins, une image de blanc (FPN) est préalablement stockée en SDRAM sur la carte XESS. Le FPGA fait la différence entre l'image acquise et l'image de référence pour éliminer le bruit spatial fixe. Les traînées diagonales qui apparaissent sur certaines images sont dûes à des erreurs d'accès en mémoire SDRAM. Le

script de description VHDL (langage de description matériel) du contrôleur SDRAM est en cours de correction. Ces premiers résultats d'imagerie rapide sont encourageants malgré la faible résolution du capteur (64×64).

5.4.1.2 Ventilateur

La vitesse de rotation d'un ventilateur pour processeur d'ordinateur (cf. figure 5.20) est de 3750 tours par minute soit 1/2 tour en 8 ms (16×500 μs). Le paramètre de réglage Vpol, la tension de polarisation de l'étage suiveur (cf. chapitre 3, section 4.2.2.1, figure 4.10), influe directement sur le contraste de l'image (cf. figure 5.21). Cette tension de polarisation permet de positionner le point de fonctionnement de l'amplificateur inverseur des cellules MA^{2}M. Le nœud flottant, où sont inter-connectées les cellules MA^{2}M, a un niveau de potentiel égal à $V_{DD}/2$ pour une tension de polarisation Vpol=1,35 V. Si la tension de polarisation est supérieure à 1,35 V, le point de fonctionnement de l'inverseur est alors inférieur à $V_{DD}/2$ et le signal de sortie a tendance à saturer positivement.

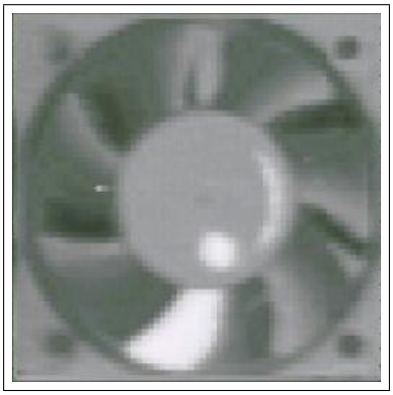
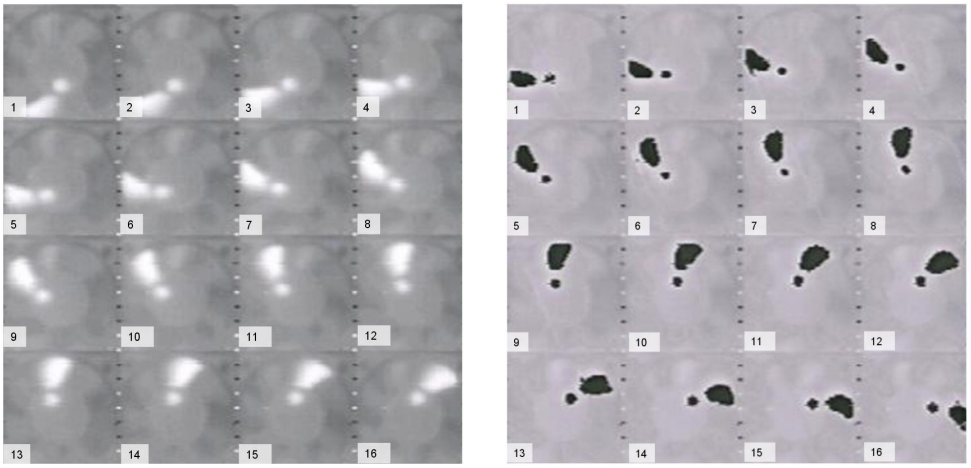


FIG. 5.20 – Image du ventilateur à l'arrêt

Pour accentuer visuellement le phénomène, un marqueur est dessiné en blanc sur le rotor (point blanc) et une des pales. Le paramètre Vpol permet de mettre en évidence le marqueur, notamment lorsque Vpol=1,65 V. Dans cette expérience, l'éclairement au plan focal objet est de 2000 lux et le plan d'éclairement est parallèle au plan objet.

Sur la figure 5.21(b), la tension de polarisation est telle que le point statique de l'inverseur est en deçà de $V_{DD}/2$ et donc l'image est mieux contrastée, ce qui permet de mettre en évidence un élément caractéristique de la scène : la pale. Les marqueurs dessinés sur le rotor et sur une des pales apparaissent en noir sur la figure, ce qui est un préalable pour d'éventuels



(a) Vpol=1,35 V (b) Vpol=1,65 V

FIG. 5.21 – 16 images consécutives d'un ventilateur en rotation à 3750 tr/min, avec un temps d'intégration de 500 μs.

traitements spatiaux sur les images.

5.4.1.3 Chenillard à DEL

Nous avons fabriqué un chenillard à diodes électroluminescentes (cf. figure 5.22) pour évaluer avec précision la vitesse du capteur. Un microcontrôleur (PIC 18F2620) est utilisé pour piloter les 12 DEL de ce chenillard. Une seule DEL est allumée à la fois pendant 133 μs ($\pm$ 5 ps) et la rotation s'effectue dans le sens des aiguilles d'une montre.

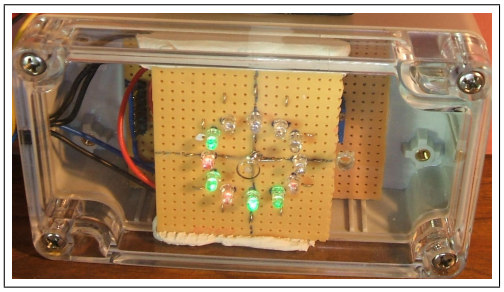


FIG. 5.22 – Chenillard à 12 DEL

Le temps d'intégration est fixé à 100 μs par le biais du FPGA. Les deux cellules MA^{2}M sont utilisées pour dissocier les phases d'acquisition et de lecture. La figure 5.23 illustre le résultat d'acquisition de 16 images consécutives.

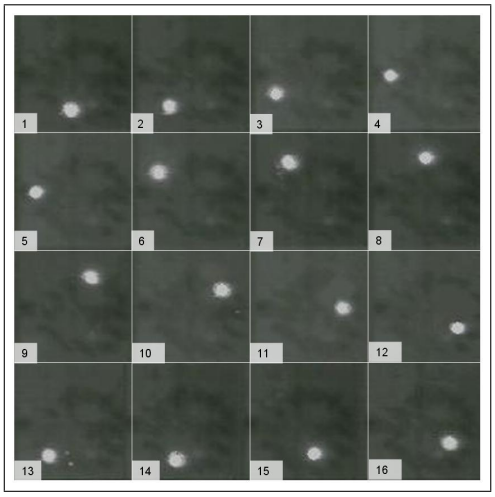


FIG. 5.23 – 16 images consécutives du chenillard, avec un temps d'intégration de 100 μs

Sur cette série d'image, on peut voir une rotation complète du chenillard soit 12 DELs allumées successivement en 16 trames. La vitesse d'acquisition est donnée par la relation :

$$F_{trame} = \frac{16\,_{trames}}{12\,_{DELs} \times 133.10^{-6}\,s} = 10\,025\,images/s \tag{5.6}$$

Comme nous l'avions prévu, la fréquence trame de notre rétine atteint 10 000 images/s.

5.4.1.4 Tapette à souris

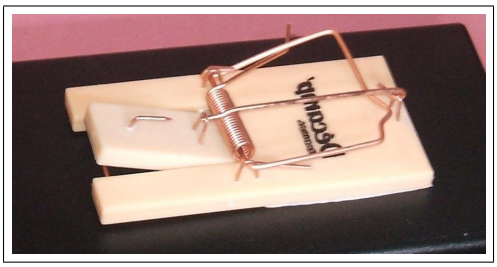


FIG. 5.24 – Image de la tapette à souris à l'arrêt

Nous utilisons une tapette à souris (cf. figure 5.24) pour apprécier un phénomène rapide. Cette tapette se déploie en 8 ms et le phénomène est filmé à 2000 images par seconde (cf. figure 5.25). L'éclairement au plan focal est

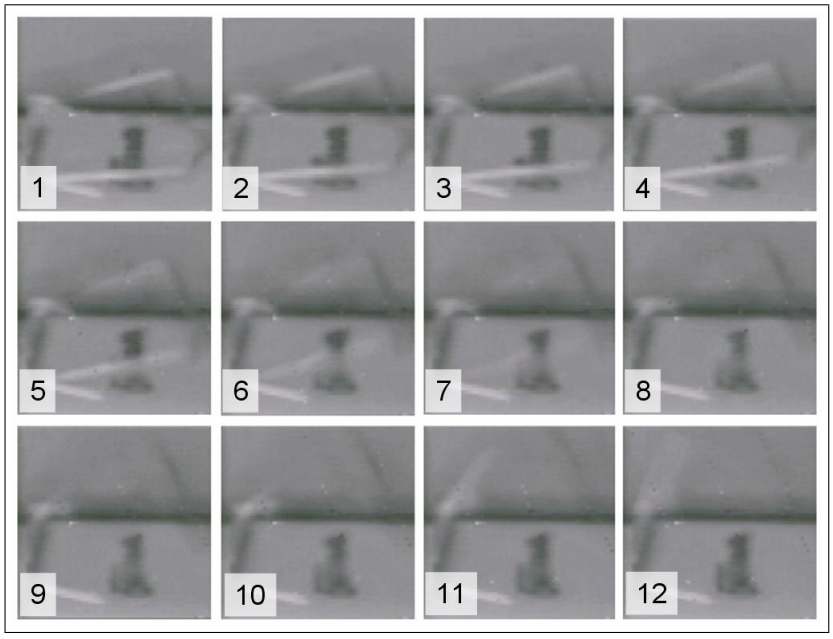


FIG. 5.25 – Déploiement de la tapette filmée à 2000 im/s

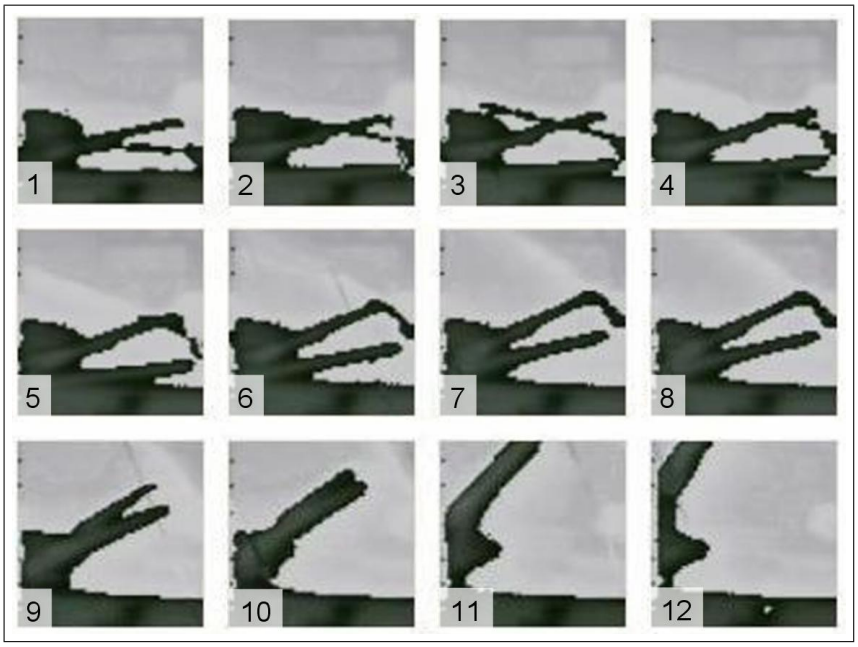


FIG. 5.26 – 12 images consécutives du déploiement de la tapette avec saturation et un temps d'intégration de 500 μs.

de 500 lux et la tension de polarisation Vpol est fixée à 1,35 Volt pour réduire le contraste et diminuer les phénomènes de saturation sur l'image. De cette façon, on peut apprécier la texture de l'objet, néanmoins les formes et les arrêtes de l'objet sont mal définies à cause du faible niveau d'éclairement.

Pour mettre en évidence les caractéristiques de la scène qui nous intéressent, en l'occurrence l'objet en mouvement (les parties mobiles), on augmente l'éclairement de 500 à 2500 lux, on change la tension de polarisation Vpol=1,65 Volt et l'orientation du plan focal image par rapport à la normale du plan d'éclairement. La figure 5.26 illustre les résultats des acquisitions. Les parties mobiles sont mises en évidence, ce qui permet l'application d'éventuels filtres spatiaux pour, par exemple, en extraire les contours.

5.4.2 Seuillage

L'opération de binarisation est intéressante pour extraire des éléments caractéristiques d'une scène, notamment en présence de marqueurs réfléchissants ou de phénomènes lumineux rapides. Un simple seuil permet de mettre en évidence un élément caractéristique de la scène. Dans le cas du ventilateur, la pale est marquée en blanc, la figure 5.27 est le résultat de l'opération obtenue avec une tension de polarisation Vpol fixée à 1,65 Volt. Seul le bit le plus significatif est alors pris en compte pour l'affichage des résultats par le biais du circuit FPGA.

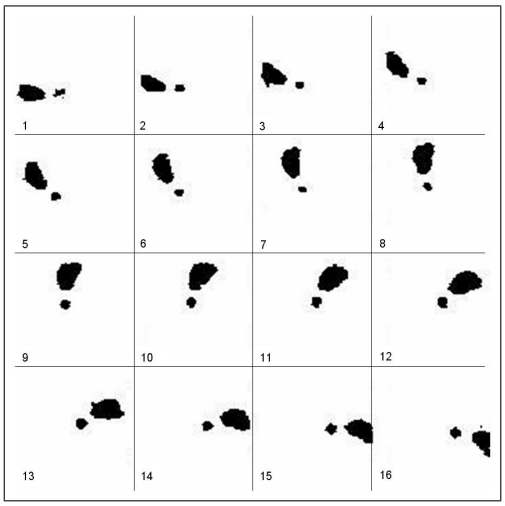


FIG. 5.27 – 16 images consécutives du ventilateur en rotation à 3750 tr/min, avec un temps d'intégration de 500 μs et une opération de seuillage sur les images.

Sur cette série d'images, on ne voit que le marqueur : un point et une pale. Dans cette expérience, le seuil est fixé par la tension de polarisation Vpol.

5.4.3 Gradients spatiaux

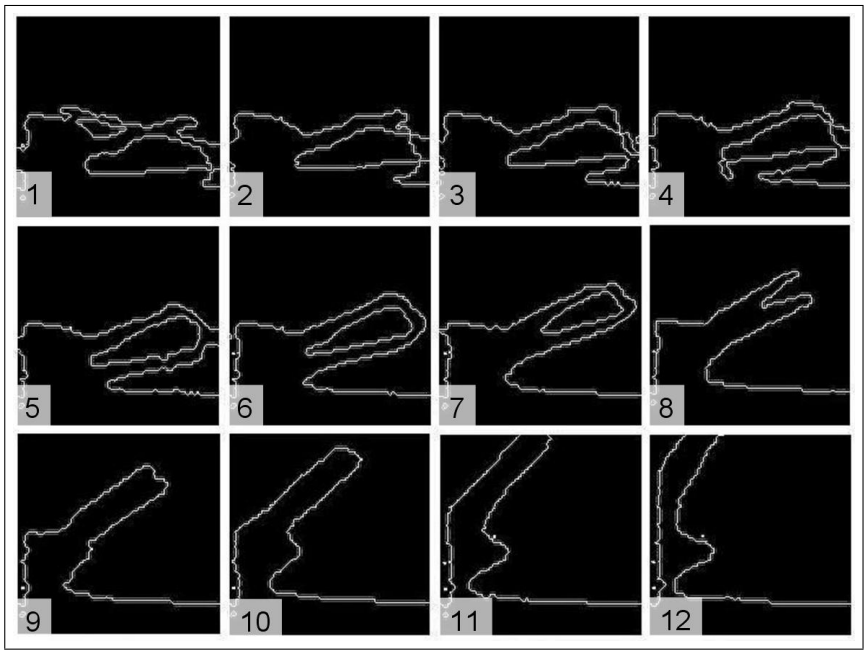


FIG. 5.28 – 12 images consécutives de la tapette à souris, avec un temps d'intégration de 100 μs et une opération de type Laplacien sur les images, soit au total 500 μs par trame.

Les contours de la tapette à souris (cf. figure 5.28) sont obtenus à la cadence de 2000 trames par seconde, et cela grâce à quatre configurations dynamiques successives ($4\times100\,\mu$s) de l'unité arithmétique analogique et un temps d'intégration de seulement 100μs (cf. chapitre 3, section 3.3.3). Un filtre Laplacien, dans sa forme la plus simple, est appliqué sur les images acquises à l'aide de l'UA[2]. L'éclairement au plan focal objet est de 2400 lux et le plan d'éclairement est incliné de 45° par rapport à la normale du plan objet. Les images résultats (filtrées) sont seuillées, la tension de polarisation est alors fixée à 1,65 Volt. La flexibilité apportée par le paramètre Vpol a un avantage déterminant sur l'impact des opérateurs convolutifs sur les images ainsi traitées. Une image très contrastée voire seuillée est plus facile à traiter qu'une image peu texturée et faiblement contrastée.

5.4.4 Gradient morphologique temporel

Le passage de l'algorithme (cf. chapitre 3, section 5.4.4) à l'implémentation sur HISIC ne pose pas de problèmes techniques liés aux calculs entiers car ceux-ci sont réalisés en analogique puis quantifiés sur 8 bits. Dans la plupart des architectures de rétines programmables, le processeur implémenté est numérique et les calculs se font sur des entiers, ce qui multiplie les erreurs liées à la quantification.

La figure 5.29 montre le résultat de l'implémentation de l'algorithme dans notre rétine sur le déploiement de la tapette à souris. Le temps d'intégration est de 100 μs et quatre configurations et relectures de l'UA^2 sont nécessaires, soit au total 500 μs par trame.

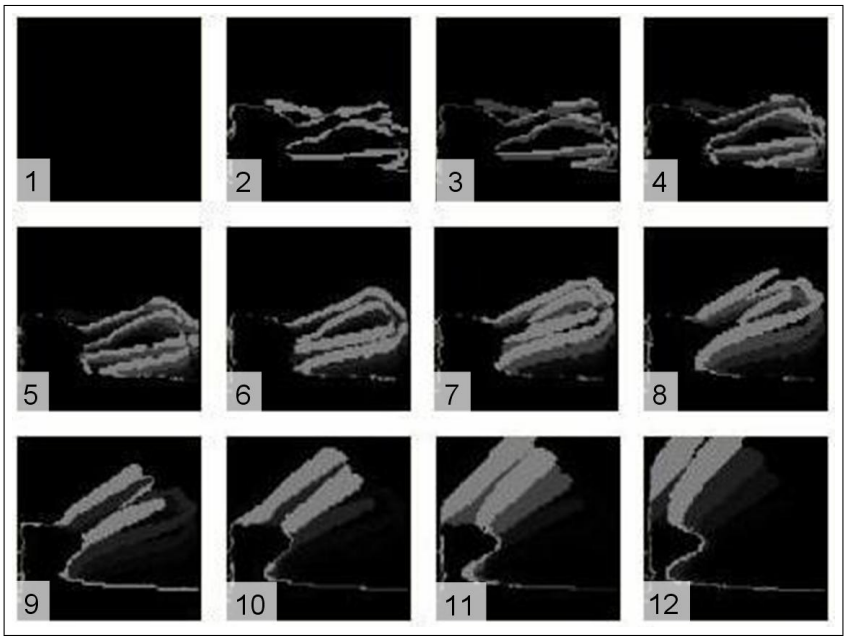


FIG. 5.29 – 12 images consécutives de la tapette à souris, avec un temps d'intégration de 100 μs et une opération de morphologie temporelle sur les images, soit au total 500 μs par trame.

Cet algorithme permet d'identifier les parties mobiles et d'analyser leurs mouvements. Seules les parties mobiles sont visibles et la récursivité permet d'apprécier le déplacement de la partie mobile entre deux trames d'acquisitions. Sur les images résultats, les parties les plus claires sont les plus récentes, elles représentent les parties mobiles dans leur position lors de la dernière acquisition. Les parties plus sombres correspondent aux positions des parties mobiles lors des trames précédemment acquises. Le dégradé, sur les images

en niveau de gris, représente les différentes positions successives des parties mobiles. On peut donc étudier avec précision la cinématique de l'objet en mouvement (position, vitesse, accélération, etc.).

5.4.5 Comparaison avec une architecture classique

Afin d'estimer les performances de l'UA^2, nous proposons de comparer les résultats de traitements d'images, issus de notre rétine, avec les résultats issus d'une unité de calcul numérique externe à la rétine. Les opérateurs utilisés sont de type Sobel horizontal, vertical et de type Laplacien.

Dans tous les cas, le traitement d'image est réalisé par l'UA^2 puis par une unité de calcul numérique externe à la rétine. On mesure la vraisemblance entre l'image traitée par la rétine et l'image traitée par une unité de calcul externe. Pour mesurer la vraisenblance, nous calculons dans chaque cas le coefficient de corrélation donné par la relation :

$$r = \frac{\sum_{j=1}^{N}\sum_{i=1}^{N}(R(i,j)-\overline{R})(P(i,j)-\overline{P})}{\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{N}(R(i,j)-\overline{R})^2}\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{N}(P(i,j)-\overline{P})^2}} \tag{5.7}$$

Où R est le l'image résultat obtenue par l'UA^2 sur la rétine, P est l'image résultat obtenue par le biais d'une unité de calcul numérique. $\overline{P}$ et $\overline{R}$ sont respectivement les moyennes des matrices P et R. N est le nombre de lignes et de colonnes de la matrice (N=64). Le tableau 5.5 récapitule les résultats de mesures.

TAB. 5.5 – Coefficient de corrélation entre les images traitées par la rétine et les images traitées par une unité de calcul numérique externe

Opérateur	coefficient de corrélation (r)
Sobel vertical	0.928
Sobel horizontal	0.930
Laplacian	0.939

Le coefficient de corrélation est un bon indicateur de performances de l'unité arithmétique analogique. En effet, ce coefficient est en moyenne de 93,2%. La vraisemblance est plus importante avec le Laplacien grâce à la parfaite symétrie de cet opérateur. Globalement, l'unité arithmétique analogique

offre de bonnes performances, comparées à une unité de calcul numérique externe à la rétine. Ce léger défaut de linéarite est palié par une vitesse de calcul supérieure à ce que l'on obtienderait avec une unité de calcul externe à la matrice.

5.5 Conclusion

Dans ce chapitre, nous avons caractérisé les deux circuits HISIC en termes de réponse spectrale, de sensibilité, de bruit, de dynamique et de consommation électrique. L'implantation de la technique CDS au sein même du pixel permet de compenser les sources de bruit d'un facteur 17. Néanmoins, le CDS utilise toutes les ressources mémoires du pixel, ce qui ne permet pas d'exécuter des traitements spatiaux à haute cadence et encore moins des traitements spatio-temporels. C'est pourquoi nous avons proposé de retrancher aux images acquises, une image fixe de référence (blanc de référence) stockée en SDRAM afin de compenser le bruit fixe. Nous avons évalué la vitesse du capteur à près de 10 000 images par seconde et validé l'implémentation de gradients spatiaux et spatio-temporels à près de 2000 images par seconde. Les opérateurs gradients permettent d'extraire les contours d'éléments caractéristiques de la scène comme la pale d'un ventilateur marquée en blanc. Le détecteur de mouvement basé sur le gradient morphologique temporel permet de mettre en évidence les parties mobiles d'une scène en mouvement rapide.

Ces premiers résultats sont encouragents et peuvent avoir des prolongements intéressants dans la réalisation d'un capteur de résolution HDV rapide et intégrant des traitements reconfigurables.

Conclusion et perspectives

Dans ce mémoire, nous avons présenté un prototype de système embarqué dédié à la vision artificielle. Les résultats que nous avons présenté sont encourageants et ouvrent la voie vers une possible réalisation à terme d'un capteur HDV rapide et "intelligent". Notre caméra permet de ralentir le temps 400 fois et de réaliser des traitements d'images linéaires. Un port VGA permet d'afficher les résultats des séquences d'acquisitions stockées en mémoire SDRAM. La rétine CMOS HISIC est interfacée à une carte constituée d'un circuit programmable FPGA (un Spartan-3, 1 million de portes), d'une mémoire FLASH (2Mo) utilisée pour stocker les fichiers de configuration du FPGA (bitstream files), d'une mémoire SDRAM (32Mo) pour stocker les séquences d'acquisitions et d'une EEPROM I^2C, pour garder en mémoire les paramètres de calibration de la caméra.

Nous avons fabriqué, testé et validé expérimentalement un capteur d'images monolithique, en technologie standard CMOS 0,35 μm. Chacun des 4096 pixels d'HISIC contient 38 transistors et implémente des fonctions d'amplifications, de mémoires et de traitements linéaires en topologie 4-connexes. Deux structures à mémoire MA^2M (voir section 4.2.2.2) permettent de dissocier les phases d'acquisitions et de lectures et dès lors d'atteindre des cadences allant jusqu'à 10 000 images par seconde. Une architecture originale de multiplieurs quatre quadrants à 5 transistors est implémentée au sein même du pixel, formant une unité arithmétique analogique UA^2 (voir section 4.2.2.3). Cette structure permet de réaliser des post-traitements linéaires sur la trame acquise. Ces traitements sont basés sur une combinaison linéaire 4-connexes et, par reconfiguration successives, peuvent s'étendre à l'application de filtres spatiaux complexes. Grâce aux mémoires implémentées, il est possible de reconfigurer plusieurs fois l'UA^2 et alors de réaliser des traitements plus complexes que des opérateurs de type Sobel ou Laplacien. La fréquence trame peut atteindre près de 2000 images/s avec un traitement simple, comme une convolution de dimension 3×3. En effet, si le temps d'intégration est fixé à 100μs, et que la période de lecture d'une trame est aussi de 100μs, comme il faut nécessairement quatre configurations de l'UA^2, alors seulement 500μs

sont nécessaires pour traiter l'image.

Les traitements d'images possibles par reconfigurations successives à la cadence vidéo (40ms par trame), soit 400 reconfigurations entre deux trames, offrent de larges perspectives dans la réalisation possible de capteurs "intelligents" performants. Un travail doit être envisagé sur le contrôle automatique de la tension de polarisation et ceci dans un premier temps, de manière simple sur le FPGA, en faisant évoluer le paramètre jusqu'à détection du contour d'un objet filmé à l'aide de l'opérateur Laplacien. En perspective, la partie numérique pourra être implantée sur un même ViSOC.

Nous avons validé expérimentalement des concepts fondamentaux au cours de ce travail de thèse à l'aide d'un capteur expérimental de résolution 64×64. Le circuit n'est qu'une preuve que le concept est réalisable. La prochaine étape de nos recherches s'oriente sur la conception d'un capteur basé sur le même principe, au format HDV (1440×1080), en technologie moderne CMOS 90 nm et intégrant des traitements linéaires complexes à la cadence vidéo.

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 processeurs externes. L'implémentation d'un algorithme de reconnaissance de visages basé sur des réseaux de neurones [YP03] est une autre ouverture possible. On peut aussi étendre le concept à l'implémentation d'un algorithme de reconnaissance et de suivi d'objets en temps réel et cela sur des images de haute définition. Il serait aussi envisageable d'implémenter un langage de programmation dédié à la cible HISIC.

Des fonctions non linéaires pourraient être ajoutées au sein du pixel. Par exemple, des comparateurs permettraient de déterminer des minima ou maxima locaux sur un voisinage de pixels afin d'exécuter des opérations particulières de morphologies mathématiques à haute cadence. Une autre caractéristique commune à l'ensemble des exemples de traitements non linéaires est leur formalisation sous forme mathématique à l'aide d'équations aux dérivées partielles (EDP) non linéaires. Le projet final serait la conception et la réalisation d'un système embarqué programmable entièrement dédié aux traitements non linéaires des images. Ce système sera programmable dans le sens où il permettra de réaliser différents types de traitements basés sur des EDP différentes.

A plus long terme, ces recherches pourraient avoir des prolongements intéressants en terme de validation in situ d'applications en traitement d'images non linéaires qui sont pour l'heure uniquement modélisées et simulées infor-

matiquement. De plus, sur le plan de l'optimisation des ressources matérielles, la reconfiguration dynamique des paramètres des EDP sur la rétine (entraînant donc une modification des traitements à la volée) permettrait de mettre au point des chaînes très complexes de traitement d'images à des cadences élevées.

Bibliographie

[AB95] A.G. Andreou and K.A. Boahen. A 590,000 transistor 48,000 pixel, contrast sensitive, edge enhancing, CMOS imager-silicon retina. In *16th Conf. Advanced Redearch in VLSI*, pages 225–240, 1995.

[ABLYV+04] O. Aubreton, B. Bellach, L.F.C. Lew Yan Voon, B. Lamalle, P. Gorria, and G. Cathébras. Retina for pattern matching in standard 0.6 μm CMOS technology. *Journal of Electronic Imaging*, 13(3) :559–569, July 2004.

[Amb98] A. Ambrozy. *Electronic Noise.* New York, 1998.

[ASIS04] P. Acosta-Serafini, M. Ichiro, and C. Sodini. A 1/3" VGA Linear Wide Dynamic Range CMOS Image Sensor Implementing a Predictive Multiple Sampling Algorithm With Overlapping Integration Intervals. *IEEE Journal of Solid-State Circuits*, 39(9) :1487–1496, September 2004.

[AW96] C.H. Aw and B Wooley. A 128x128 pixel standard CMOS image sensor with electronic shutter. *IEEE Journal of Solid State Circuits*, 31(12) :1922–1930, 1996.

[BB06] N. Burrus and T. Bernard. Adaptive vision leveraging digital retinas : Extracting meaningful segments. In *ACIVS*, 2006.

[BBM+02] 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*, 37(2) :160–172, February 2002.

[Bel03] B. Bellach. *Capteurs d'image : Application à l'observation et à la mesure de vitesse de phoénomènes lumineux rapides.* PhD thesis, Université de Bourgogne, 2003.

[BGE03] W Bidermann, Abbas El Gamal, and S. Ewedemi. A 0,18 μm high dynamic range NTSC/PAL imaging system-on-chip with embedded DRAM frame buffer. *ISSC Tech. Dig.*, pages 212–213, 2003.

[BLB98] R. J. Baker, Harry W. Li, and D.E. Boyce. *CMOS Circuit Design, Layout, and Simulation.* Wiley-IEEE, 1998.

[BLLYVC03] B. Bellach, B. Lamalle, L.F.C. Lew Yan Voon, and G. Cathébras. A CMOS image sensor design for speed determination of fast moving

luminous objects. In *Society of Manufacturing Engineers (SME) Technical paper*, volume MV03-254, pages pp. 120–128, USA, 2003.

[BSPS97] J. Bastos, M. Steyaert, A. Pergoot, and W. Sansen. Mismatch characterization of submicron MOS transistor. *Analog Integrated Circuits and Signal Processing*, 12 :95–106, February 1997.

[BZD93] T. Bernard, B. Zavidovique, and F. Devos. A programmable artificial retina. *IEEE Journal of Solid-State Circuits*, 28 :789–798, 1993.

[CADS90] K. Chen, M. Afghani, E. Danielsson, and C. Svensson. PASIC : A Processeur-A/D Converter-Sensor Integrated Circuit. *IEEE Int. Symposium On Circuits and Systems*, pages 1705–1708, 1990.

[CBM+02] G. Chapinal, S.A. Bota, M. Moreno, J. Palacin, and A. Herms. A 128 × 128 CMOS Image Sensor With Analog Memory for Synchronous Image Capture. *IEEE Sensors Journal*, 2(2) :120–127, April 2002.

[Cha05] D. Chanemougame. *Conception et fabrication de nouvelles architectures CMOS et étude du transport dans les canaux de conduction ultra minces obtenus avec la technologie SON.* PhD thesis, Institut National des Sciences Appliquées de Lyon, 2005.

[Cho00] K. et al. Cho. A 1.2v Micropower CMOS Active Pixel Image Sensor for Portable Applications. volume 43, pages 114–115, 2000.

[CMC+06] Y.M. Chi, U. Mallik, E. Choi, M. Clapp, G. Gauwenberghs, and R. Etienne-Cummings. CMOS pixel-level ADC with change detection. In *Proceedings of the International Symposium on Circuits and Systems (ISCAS)*, pages 1647–1650, May 2006.

[CRVG+04] G.L. Cembrano, A. Rodriguez-Vazquez, R.C. Galan, F. Jimenez-Garrido, S. Espejo, and R. Dominguez-Castro. A 1000 FPS at 128 × 128 Vision Processor With 8-Bit Digitized I/O. *IEEE Journal of Solid-State Circuits*, 39(7) :1044–1055, July 2004.

[DC06] P. Dudek and S. Carey. General-purpose 128×128 SIMD processor array with integrated image sensor. *Electronic Letters*, 42 :678–679, January 2006.

[Deg00] Y. Degerli. *Etude, modélisation des bruits et conception des circuits de lecture dans les capteurs d'image à pixels actifs CMOS.* PhD thesis, SUPAERO, octobre 2000.

[Den95] P.B. et al. Denyer. Intelligent CMOS imaging. In *Charge-Coupled Devices and Solid State Optical Sensors IV-Proceedings of the SPIE Electronic Imaging Conference*, volume 2415, pages 85–91, San Jose, CA, USA, February 1995.

[DH00] P. Dudek and P.J. Hicks. A CMOS general-purpose sampled-data analog processing element. *IEEE Transactions on Analog and Digital Signal Processing*, 47 :467–473, May 2000.

[DK97] D. Demigny and T. Kamlé. A discrete expression of Cann's criteria for step edge detector. Performances evaluation. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 11(19) :1199–1211., 1997.

[DL04] T. Delbrück and S.C. Liu. A silicon early visual system as a model animal. *Vision Research*, 44 :2083–2089, 2004.

[DLMF00] Y. Degerly, P. Laverne, P. Magnan, and J.A. Farré. Analysis and Reduction of Signal Readout Circuitry Temporal Noise in CMOS Image. *IEEE Transactions Electron Devices*, 47(5) :949–962, 2000.

[Dud05] P. Dudek. Implementation of simd vision chip with 128×128 array of analogue processing elements. In *the International Symposium on Circuits and Systems (ISCAS)*, volume 5, pages 5806–5809, 2005.

[ECKC01] R. Etienne-Cummings, Z. Kalayjian, and D. Cai. A programmable focal-plane mimd image processor chip. *IEEE Journal of Solid-State Circuits*, 36(1) :64–73, 2001.

[EGE05] A. El Gamal and H. Eltoukhy. Cmos Image Sensor. *IEEE Circuits and Devices Magazine*, 2005.

[EGYF99] A. El Gamal, D. Yang, and B. Fowler. Pixel level processing – Why, What and How ? In *Proceedings of the SPIE Electronic Imaging '99 conference*, volume 3650, pages 2–13, January 1999.

[ESA95] J.-E. Eklund, C. Svensson, and A. Astrom. Near Sensor Image Processing, A VLSI Realization. *Solid-State Sensors and Actuators*, 1(1) :126–129, June 1995.

[FEGY94] B. Fowler, A. El Gamal, and D.X.D. Yang. A CMOS area image sensor with pixel-level A/D conversion. *IEEE Int. Solid-State Circuits Tech.Dig.*, pages 137–138, 1994.

[FEGY98] B. Fowler, A. El Gamal, and D.X.D. Yang. Techniques for pixel-level analog-to-digital conversion. In *Proceedings of SPIE, Infrared Readout Electronics IV*, volume 3360, pages 2–12, Orlando, April 1998.

[FNM+97] E. Funatsu, Y. Nitta, Y. Miyake, T. Toyoda, J. Ohta, and K. Kuyma. An Artificial Retina Chip with Current-Mode Focal Plane Image Processing Functions. *IEEE Transactions on Electron Devices*, 44(10) :1777–1782, October 1997.

[Fos93] E.R. Fossum. Active pixel sensors : Are CCDs dinosaurs ? *International Society for Optical Engineering (SPIE)*, 1900 :2–14, 1993.

[Fos97] E.R. Fossum. CMOS Image Sensor : Electronic Camera On A CHIP. *IEEE Transactions on Electron Devices*, 44(10) :1689–1698, October 1997.

[Fow95] B Fowler. *CMOS area image sensors with pixel level A/D conversion.* PhD thesis, Stanford University, 1995.

[Fuj00] L. et al. Fujimori. A 256x256 CMOS Differential Passive Pixel Imager with FPN Reduction Techniques. In *Proceedings of the 2000 International Solid State Circuits Conference*, volume 43, pages 106–107, February 2000.

[FYG95] B. Fowler, D.X.D. Yang, and El Gamal. CMOS FPA with Multiplexed Pixel Level ADC. In *At 1995 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, Dana Point, CA, April 20-22 1995.

[GD04] C. Garcia and M. Delakis. Convolutional Face Finder : A Neural Architecture for Fast and Robust Face Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11), November 2004.

[Gil68] B. Gilbert. A precise four-quadrant multiplier with subnanosecond response. *IEEE Journal of Solid-State Circuits*, 3 :365–373, 1968.

[GM93] P. Gray and R. Mayer. *Analysis and Design of Analog Integrated Circuits*, chapter 6. USA, 1993.

[Gou95] E. Goujou. *Etude et réalisation d'un capteur d'image à temps d'exposition variable en CMOS.* PhD thesis, Université de Bourgogne, 1995.

[GPPH05] D. Ginhac, E. Prasetyo, M. Paindavoine, and B. Heyrman. Principles of a CMOS sensor dedicated to face tracking and recognition. In *IEEE CAMP05 International Workshop on Computer Architecture for Machine Perception*, July 2005.

[HA01] T. Hamamoto and K. Aizawa. A Computational Image Sensor with Adaptive Pixel-Based Integration Time. *IEEE Journal of Solid-State Circuits*, 36(4) :580–585, April 2001.

[HAB+05] A. Harton, M. Ahmed, A. Beuhler, F. Castro, L. Dawson, B. Herold, G. Kujawa, K. Lee, R. Mareachen, and T. Scaminaci. High dynamic range CMOS image sensor with pixel level ADC and in-situ image enhancement. In *Sensors and Camera Systems for Scientific and Industrial Applications VI. Proceedings of the SPIE*, volume 5677, pages 67–77, Mar 2005.

[HH03] Y. Huang and R. Horsney. Current-mode CMOS image sensor using lateral bipolar phototransistors. *IEEE Transactions on Electron Devices*, 50(12) :2570–2573, 2003.

[HKHC99] K.K. Hung, P.K. Ko, C. Hu, and Y.C. Cheng. A Unified Model for the Fliker Noise in Metal Oxide Semiconductor Field Effect. *IEEE Transaction Electron Devices*, 37 :654–665, 1999.

[HST+00] D. Handoko, Kawahito S, Y. Takokoro, M. Kumahara, and A. Matsuzawa. A CMOS image sensor for local-plane motion vector estimation. In *Symposium of VLSI Circuits*, volume 3650, pages 28–29, June 2000.

[JISF05] C. Jansson, P. Ingelhag, C. Svensson, and R. Forchheimer. An addressable 256×256 photodiode image sensor array with an 8-bit digital output. *Analog Integrated Circuits and Signal Processing*, 4(1) :37–49, 2005.

[KDMS99] J.O. Klein, A. Dupret, A. Moutault, and A. Share. Vers une nouvelle génération de rétines programmables. In *GRETSI'99*, Nantes, France, Septembre 1999.

[KDVF02] T. Kervyn, N. Donckers, M. Verleysen, and D. Flandre. Multiplieur analogique en technologie soi pour le décodage de turbo-codes. In *Colloque sur le Traitement Analogique de l'Information, du Signal et ses Applications*, 2002.

[KHE+86] R. Koch, B. Heise, F. Eckbauer, E. Engelhardt, J. Fisher, and F. Parzefall. A 12-bit Sigma-Delta Analog-to-Digital Converter with a 15 mhz Clock Rate. *IEEE Journal of Solid-State Circuits*, 21(6), December 1986.

[KIIY03] T. Komuro, I. Ishii, M. Ishikawa, and A. Yoshida. A digital vision chip specialized for high-speed target tracking. *IEEE Transactions on Electron Devices*, 50(1) :191–199, 2003.

[KKI04] T. Komuro, S. Kagami, and M. Ishikawa. A dynamically reconfigurable SIMD processor for a vision chip. *IEEE Journal of Solid-State Circuits*, 39(1) :265–268, 2004.

[KL95] C. Koch and H. Li. VISION CHIPS - Implementing Vision Algorithms with Analog VLSI Circuits. In *IEEE Computer Society Press*, Los Alamitos, CA, 1995.

[KLLEG01a] S. Kleinfelder, S. Lim, X. Liu, and A. El Gamal. A 10 000 Frames/s CMOS Digital Pixel Sensor. *IEEE Journal of Solid-State Circuits*, 36(12) :2049–2059, December 2001.

[KLLEG01b] S. Kleinfelder, S. Lim, X. Liu, and A. El Gamal. A 10,000 Frames/s 0.18 μm CMOS Digital Pixel Sensor with Pixel-Level Memory. In *Proceedings of the 2001 IEEE International Solid-State Circuits Conference*, pages 88–89, San Francisco, CA, February 2001.

[KN03] A. Krymsky and T. Niarong. A 9-V/Lux 5000-Frames/s 512 x 512 CMOS Sensor. *IEEE Transactions on Electron Devices*, 50(1) :136–143, January 2003.

[KOA91] S Kebeya, B. Orsal, and R. Alabedra. *Noise in Superlattice Avalanche Photodiodes.* Noise in Physical Systems and 1/f Fluctuations, 1991.

[KPS+05] G. Kottaras, N.P. Paschalidis, E.T. Sarris, N. Stamatopoulos, K. Karadamoglou, and V. Paschalidis. A 10-bit, Low Power, Successive Approximation, Digitally Auto-Zeroed CMOS ADC Core for the NASA TRIO Smart Sensor System on a Chip. *Analog Integrated Circuits and Signal Processing*, 42 :113–128, 2005.

[KRB+05] L.J. Kozlowski, G. Rossi, L. Blanquart, R. Marchesini, Y. Huang, G. Chow, J. Richardson, and D. Standley. Pixel Noise Suppression via SoC Management of Target Reset in a 1920 $\times$ 1080 CMOS Image Sensor. *IEEE Journal of Solid-State Circuits*, 40(12) :2766–2776, December 2005.

[KSGA00] E.A.M Klumperink, L.J. Sander, Gierkink, and P. Armoud. Reducing MOSFET 1/f Noise and Power Consumption by Switched Biasing. *IEEE Journal of Solid State Circuits*, 35(7) :994–1001, 2000.

[KVBA+99] A. Krymski, D. Van Blerkom, A. Andersson, N. Block, B. Mansoorian, and E.R. Fossum. A High Speed, 500 Frames/s, 1024×1024 CMOS Active Pixel Sensor. *Symposium on VLSI Circuits*, pages 137–138, 1999.

[LEG01a] S. Lim and A. El Gamal. Integrating Image Capture and Processing – Beyond Single Chip Digital Camera. In *Proceedings of the SPIE Electronic Imaging '2001 conference*, volume 4306, San Jose, CA, january 2001.

[LEG01b] 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*, volume 3, pages 1841–1844, 2001.

[LEG01c] X.Q. Liu and A. El Gamal. Photocurrent estimation from multiple non-destructive samples in a CMOS image sensor. In *Proceedings of the SPIE Electronic Imaging '2001 conference*, volume 4306, San Jose, CA, january 2001.

[LH95] S.I. Liu and Y.S Hwang. CMOS Squarer and Four-Quadrant Multiplier. *IEEE Transactions on Circuits and Systems-I :Fundamental Theory and Applications*, 42(2) :119–122, Feb 1995.

[LH01] J.S. Lee and R.I. Hornsey. CMOS Photodiodes with Substrate Openings for Higher Conversion Gain in Active Pixel Sensor. In *IEEE Workshop on CCDs and Advanced Image Sensors*, Crystal Bay, Nevada, June 2001.

[Lim03] S. H. Lim. *Video Processing Applications of High Speed CMOS Image Sensors*. PhD thesis, Stanford university, March 2003.

[Lit01] D. Litwiller. CCD vs. CMOS : Facts and Fiction. *Photonics Spectra*, pages 154–158, January 2001.

[LMJM05] L. Lindgren, J. Melander, R. Johansson, and B. Möller. A Multiresolution 100-GOPS 4-Gpixels/s Programmable Smart Vision Sensor for Multisense Imaging. *IEEE Journal of Solid-State Circuits*, 40(6) :1350–1359, June 2005.

[LMS01] M. Loose, K. Meier, and J. Schemmel. A Self-Calibrating Single-Chip CMOS Camera with Logarithmic Response. *IEEE Journal Of Solid-State Circuits*, 36, April 2001.

[LSB+98] M. Loinaz, K. Singh, A. Blanksby, D. Inglis, K. Azadet, and B. Ackland. A 200mv 3.3v CMOS Color Camera IC Producing 352 × 288 24-b Video at 30 Frames/s. *IEEE Journal of Solid-State Circuits*, 33(12) :2092–2103, 1998.

[MA99] T.G. Morris and Al. A column-based processing array for high-speed digital image processing. In *Conference on Advanced Research in VLSI*, pages 42–56, Atlanta, GA, 1999.

[Mag03] P. Magnan. *Detection of visible photons in CCD and CMOS : A comparative view*, volume 504, pages 199–212. Elsevier, May 2003.

[Man00] A. Manzanera. *Vision Artificielle Retinienne*. PhD thesis, Ecole Nationale Superieure des Telecommunications, 2000.

[Mat01] H. Mathieu. *Physique des semiconducteurs et des composants électronique*. Paris, 2001.

[Mea89] C.A. Mead. *Analog VLSI and Neural Systems*. Reading, Massachussets, 1989.

[Mey02] J.U. Meyer. Retina implant a bioMEMS challenge. *Sensors and Actuators*, pages 1–9, 2002.

[MKG+97] S.K. Mendis, S.E. Kemeny, R.C. Gee, B. Pain, C.O. Staller, Q. Kim, and E.R. Fossum. CMOS Active pixel image sensors for highly integrated imaging systems. *IEEE Journal of Solid-State Circuits*, 32(2) :187–197, 1997.

[MM88] C. Mead and M.A. Mahowald. A silicon model of early visual processing. *Neural Networks*, 1 :91–97, 1988.

[Mos06] R. Mosqueron. *Conception et réalisation d'une caméra rapide haute résolution avec traitements d'images embarqués*. PhD thesis, Université de Bourgogne, décembre 2006.

[Nav03] D. Navarro. *Architecture et conseption de rétines silicium CMOS : Application à la mesure du flot optique*. PhD thesis, Université de Montpellier II, octobre 2003.

[NHIA00] T. Nezuka, M. Hoshino, M. Ikeda, and K. Asada. A Position Detection Sensor for 3-D Measurement. In *26th European Solid-State Circuits Conference*, Stockholm, Sweden, September 2000.

[Ni01] Y. Ni. Element photoélectrique à très grande dynamique de fonctionnement. Technical report, Brevet N°0101001, 2001.

[NZDA93] Y. Ni, Y.M. Zhu, F. Devos, and K. Arion. Yet Another Analog 2D Gaussian Convoluer. In *ISCAS93*, volume 1, pages 192–195, Chicago, May 1993.

[Pai01] F. Paillet. *Intégration et évaluation de rétines artificielles numériques programmables de hautes performances*. PhD thesis, Université P. et M. Curie, 2001.

[Par08] C. Parks. Multiple output charge-coupled devices. Brevet d'invention PCT/US2007/015641, Eastman KODAK Company, 343 State Street, Rochester, New York 14650-2201 (US), January 2008.

[PKK+03] D.-S. Park, J.-H. Kim, H.-S. Kim, J.-H. Park, J.-K. Shin, and M. Lee. A foveated-structure CMOS retina chip for edge detection with local ligh adaptation. *Sensor and Actuators*, 108(1) :75–80, May 2003.

[PMB99] F. Paillet, D. Mercier, and T.M. Bernard. Second generation programmable artificial retina. In *IEEE ASIC/SOC*, pages 304–309, 1999.

[PTV98] M.J.M. Pelgrom, H.P. Tuinhout, and M. Vertregt. Transistor matching in analog CMOS applications. In *IEEE Int. Electron Devices Meeting Tech. Digest*, pages 915–918, San Francisco, 1998.

[Raz01] B. Razavi. *Design of Analog CMOS Integrated Circuits.* McGraw Hill, University of California, Los Angeles, 2001.

[Ric06] J. Richefeu. *Détection et analyse du mouvement sur système de vision à base de rétine numérique.* PhD thesis, Université de Paris 6, décembre 2006.

[RVB98] N. Ranganathan, N. Vijaykrishnan, and N. Bhavanishankar. a Linear Array Processor with dynamic frequency clocking for image processing applications. *IEEE Transactions*, 8(8, Issue 4, Aug 1998 Page(s) :435 - 445) :435–445, August 1998.

[RVEDC+96] A. Rodriguez-Vazquez, S. Espejo, R. Dominguez-Castro, R. Carmona, and E. Roca. Mixed-signal CNN array chips for image processing. In *SPIE*, volume 2950, pages 218–229, 1996.

[RVLCC+04] A. Rodriguez-Vasquez, G. Linan-Cembrano, L. Carranza, E. Roca-Moreno, R. Carmona, F. Jimenez-Garrido, R. Dominguez-Castro, and S. Meana. ACE16k : the third generation of mixed-signal SIMD-CNN ACE chips toward VSoCs. *IEEE Transactions on Circuits and Systems I : Regular Papers*, 51(5) :851–863, 2004.

[Rya70] C.R. Ryan. Applications of a four-quadrant multiplier. *IEEE Journal of Solid-State Circuits*, 5(1) :45–48, Feb 1970.

[SBYP02] I. Shcherback, A. Belenky, and O Yadid-Pecht. Empirical dark current modeling for complementary metal oxide semiconductor active pixel sensor. *Optical Engineering*, 41(6) :1216–1219, June 2002.

[SDM93] R. Sarpeshkar, T. Delbrük, and C.A. Mead. White noise in MOS transistor and resistors. *IEEE Circuits and Devices Magazine*, pages 23–29, November 1993.

[Sei97] P. Seitz. Image sensing with maximum sensitivity using industrial CMOS technology. In Proceedings of the SPIE, editor, *The International Society for Optical Engineering*, volume 3099, pages 22–33, 1997.

[Sei00] P. Seitz. Solid-State Image Sensing. *Handbook of computer Vision and Applications*, 1 :165–222, 2000.

[SHO04] R. Singh, D.L. Harame, and M.M. Oprysko. *SILICON GERMANIUM : Technology, Modeling, and Design.* IEEE Press, Canada, 2004.

[SHT+98] S. Smith, J. Hurwitz, M. Torrie, D. Baxter, A. Holmes, M. Panaghiston, R. Henderson, A. Murrayn, S. Anderson, and P. Denyer. A single-chip 306x244-pixel CMOS NTSC video camera. In *ISSCC Digest of technical papers*, pages 170–171, San Fransisco, CA, 1998.

[SKH+05] 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*, 40(5) :1147–1156, May 2005.

[Sol96] Solhusvik. *Etude et conception de capteurs d'images à pixels actifs et de l'électronique de traitement associée en vue d'application faible flux.* PhD thesis, SUPAERO, 1996.

[SQS+00] G. Sandini, P. Questa, D. Scheffer, B. Dierickx, and A. Mannucci. A Retina-like CMOS Sensor and its Applications. In *Proceeding of : 1st IEEE SAM Workshop*, Cambridge, USA, March 16-17 2000.

[SSA98] E. Sanchez-Sinencio and A. Andreou. *Low-Voltage/Low Power Integrated Circuits and Systems*, chapter 9, pages 278–282. USA, 1998.

[Sta91] D.L. Standley. An object position and orientation ic with embedded image. *IEEE Journal Of Solid-State Circuits*, 26 :1853–1859, December 1991.

[STT+05] 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*, 40 :2816–2823, 2005.

[SYP03] I. Shcherback and O Yadid-Pecht. Photoresponse Analysis and Pixel Shape Optimization for CMOS Active Pixel Sensors. *IEEE Transactions on Electron Devices*, 50(1) :12–18, January 2003.

[Sze81] S.M. Sze. *Physics of semiconductor Devices.* 1981.

[TH01] M. Tabet and R. Hornsey. CMOS image sensor camera with focal plane edge detection. In *Canadian Conference on Electrical and Computer Engineering*, volume vol. 2, pages pp. 1129–1133, May 13.16 2001.

[VLR94] L.K.J. Vandamme, X. Li, and D. Rigaud. 1/f Noise in MOS Devices, Mobility or Number Fluctuation ? *IEEE Transaction Electron Devices*, 41 :1936–1944, 1994.

[WC04] 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*, 4(4) :510–518, August 2004.

[WCD+99] B. Wandell, P. Catrysse, J. DiCarlo, D. Yang, and A. El Gamal. Multiple Capture Single Image Architecture with a CMOS Sensor. In *Proceedings of the International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives*, pages 11–17, Chiba, Japan, October 21-22 1999.

[Won96] H.S. Wong. CMOS active pixel image sensors fabricated using a 1.8v 0.25 μm CMOS technology. In *Proceedings of International Electron Devices Meeting*, pages 915–918, San Francisco, 1996.

[WS85] J. Wiley and Sons. *Semiconductor devices : Physics and technologiy.* 1985.

[WSL+04] C. Wu, Y. Shih, J. Lan, C. Hsieh, C. Huang, and J. Lu. Design, optimization, and performance analysis of new photodiode structures for CMOS active-pixel-sensor (aps) imager applications,. *IEEE Sensors Journal*, 4(1) :135–144, February 2004.

[YEGFT99a] D. Yang, A. El Gamal, B. Fowler, and H. Tian. A 640 x 512 CMOS Image Sensor with Ultra Wide Dynamix Range Floating-Point Pixel-Level ADC. *IEEE Journal of Solid-State Circuits*, 34 :1821–1834, December 1999.

[YEGFT99b] D. Yang, A. El Gamal, B. Fowler, and H. Tian. A 640x512 CMOS Image Sensor with Ultra Wide Dynamic Range Floating-Point Pixel-Level ADC. In *Proceedings of the 1999 IEEE International Solid-State Circuits Conference*, pages 308–309, San Francisco, CA, February 1999.

[YFEG96] D. Yang, B. Fowler, and A. El Gamal. A 128x128 Pixel CMOS Area Image Sensor with Multiplexed Pixel Level A/D Conversion. In *IEEE 1996 Custom Integrated Circuits Conference*, San Diego, CA, May 1996.

[YFEG98] D. Yang, B. Fowler, and A. El Gamal. A Nyquist Rate Pixel Level ADC for CMOS Image Sensors. In *Proceedings of IEEE 1998 Custom Integrated Circuits Conference*, pages 237–240, Santa Clara, CA, May 1998.

[YFEG99] D. Yang, B. Fowler, and A. El Gamal. A Nyquist-Rate Pixel-Level ADC for CMOS Image Sensors. *IEEE Journal of Solid-State Circuits*, 34(3) :348–356, March 1999.

[YP03] F. Yang and M. Paindavoine. Implementation of a RBF neural network on embedded systems : Real time face tracking and identity verification. *IEEE Transactions on Neural Networks*, 14(5) :1162–1175, september 2003.

[YPB03] O. Yadid-Pecht and A. Belenky. In-Pixel Autoexposure CMOS APS. *IEEE Journal of Solid-State Circuits*, 38(8) :1425–1428, August 2003.

[YPEC04] O Yadid-Pecht and R. Etienne-Cummings. *CMOS Imagers : From Phototransduction to Image Processing.* Kluwer Academic Publishers, 2004.

[YPPS+97] O. Yadid-Pecht, B. Pain, C. Staller, C. Clark, and E. Fossum. CMOS Active Pixel Sensor Star Tracker with Regional Electronic Shutter. *IEEE Journal of Solid-State Circuits*, 32(2) :285–288, February 1997.

[ZC98] W. Zhang and M. Chan. Properties and design optimization of photodiodes available in a current CMOS technology. In *IEDM*, pages 22–25, 1998.

Publications

Revue d'audience internationale avec comité de lecture

[1] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, and Barthélémy Heyrman, "A 10 000 fps CMOS Sensor with Massively Parallel Image Processing", IEEE Journal of Solid-State Circuits, 43(3) :706-717, March 2008. (FI=2,008 en 2006)

Communications à des conférences internationales avec actes (5)

[2] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, "A Multi-Processing 10 000 frames/s CMOS Image Sensor", Workshop on Design and Architectures for Signal and Image Processing, DASIP 2007, November 27, 2007, Grenoble, France.

[3] Michel Paindavoine, Julien Dubois, Romuald Mosqueron, Barthélémy Heyrman, **Jérôme Dubois**, Dominique Ginhac, "High speed camera with embedded image processing", 6th International Workshop on Embedded System, September 6-7, 2007, Vaasa, Finland.

[4] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, "VLSI Design of a High-Speed CMOS Image Sensor with in-situ 2D Programmable Processing", EUSIPCO 2006, September 8, 2006, Florence, ITALY.

[5] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, "A single-chip 10 000 frames/s CMOS sensor with in-situ 2D programmable image processing", IEEE International Workshop on CAMPS 2006, 18 September 2006, Montreal, Quebec, Canada.

[6] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, "Design of a 10 000 Frames/s CMOS sensor with in-situ image processing ", Proceedings of ReCoSoC06, 3 July 2006, Montpellier, France.

Communications à des conférences nationales avec actes (1)

[7] **Jérôme Dubois**, Dominique Ginhac, Michel Paindavoine, "Un Capteur d'Images Reconfigurable dédié à l'Imagerie Rapide, aux Traitements d'Images Linéaires et Réseaux Convolutifs ", 8ème colloque sur le Traitement Analogique de l'Information, du Signal et ses Applications, TAISA 2007, 18 Octobre 2007, Lyon, France.

Glossaire

AMS : AutriaMicroSystem est une firme autrichienne spécialisée dans la fabrication de circuits intégrés.

AOT : Amplificateur Opérationnel à Transconductance, c'est un amplificateur différentiel de courant.

Bonding : Les fils de bonding sont des micro-fils de connexions entre le circuit intégré et le boîtier.

Bridge : Un bridge, en technologie CMOS, c'est le passage d'une couche de métal à une autre de niveaux différents, ex : métal1 - via1 - métal2.

CAN : Convertisseur Analogique Numérique.

CCD : De l'anglais Charge-Coupled Devices ou Circuit à Couplage de Charge.

CDS : De l'anglais Correlated Double Sampling ou double échantillonnages corrélés.

CFF : De l'anglais Convolutional Face Finder, ou algorithme convolutif de détection de visage.

CMOS : De l'anglais Complementary Metal-Oxide-Semiconductor, c'est une technique de fabrication de composant électrique faible consommation.

CNN : De l'anglais Cellular Nonlinear Network, c'est une cellule d'un réseau de neurone.

DEL : Une DEL ou diode électroluminescente produit un rayonnement monochromatique.

DNL : Non-linéarité différentielle ou différence maximale entre deux seuils

consécutifs sur la sortie numérique par rapport à l'entrée analogique.

EEPROM : De l'anglais Electrically Erasable Programmable Read-Only Memory ou mémoire morte effaçable électriquement et programmable.

FLASH : La mémoire FLASH possède les caractéristiques d'une mémoire vive mais dont les données ne disparaissent pas lors d'une mise hors tension.

FPGA : De l'anglais, Fiel Programmable Gate Array, comprenez, réseau de portes programmables in-situ, c'est le successeur des CPLD pour Complex Programmable Logic Device, comprenez, circuit logique programmable complexe.

FPN : De l'anglais Fixe Pattern Noise ou du français BSF pour Bruit Spatial Fixe.

Glinch : Un glinch est un pic de signal qui à des répercutions sur une gamme de fréquences importante, sur plusieurs centaines de mégahertz.

HDV : De l'anglais High-Definition Video, définie un format vidéo haute résolution 1280×720 et 1440×1080.

I2C : De l'anglais Inter Integrated Circuit Bus, c'est le nom du bus historique, développé par Philips pour les applications de domotique et d'électronique domestique au début des années 1980, notamment pour permettre de relier facilement à un microprocesseur les différents circuits d'une télévision moderne.

INL : Non-linéarité intégrale (INL), c'est la différence maximale entre les seuils effectifs et théoriques sur la sortie numérique par rapport à l'entrée analogique.

Layout : Ou mask design, ne pas confondre avec le stick diagram, le dessin de masque représente les géométries planaires des couches successives du procédé de fabrication et le stick diagram n'en est qu'une grossière représentation.

LSTP : De l'anglais Low Stand-by Power, est un transistor qui privilégie une faible consommation tout en maintenant un courant de saturation assez élevé.

Offset : L'offset est le différence qui existe entre la réponse du pixel et une valeur de calibration.

PCB : De l'anglais Printed Circuit Board ou du français circuit imprimé.

RMS : De l'anglais Root Mean Squared Voltage ou racine carrée de l'intégrale

temporelle du carré du signal ou encore valeur efficace du signal étudié.

RVB : Système RVB : Rouge (680 nm), Vert (546,1 nm) et Bleu (435,8 nm).

SDRAM : De l'anglais Synchronous Dynamic Random Access Memory ou Mémoire Dynamique Synchrone à Accès Aléatoire.

SIMD : De l'anglais Single Instruction Multiple Data, désigne un mode de fonctionnement des ordinateurs dotés de plusieurs unités de calcul qui exécutent de manière massivement parallèle un ensemble d'instructions sur des données différentes.

Switch : Un switch CMOS est un interrupteur électronique commandable à l'aide de signaux numériques.

USB2 : Protocole de communication série pouvant atteindre un débit maximum théorique de 480 Mbits/s.

ViSOC : De l'anglais Vision System On Chip ou système de vision sur puce.

VHDL : De l'anglais Very High Speed Integrated Circuit Hardware Description Language, est un langage de description matériel destiné à décrire le comportement et/ou l'architecture d'un système électronique numérique.

Wafer : Wafer est le mot anglais qui désigne une tranche, c'est-à-dire un disque assez fin de matériau semi-conducteur, comme le silicium. Il sert de support à la fabrication de micro-structures par des techniques telles que le dopage, la gravure, la déposition d'autres matériaux et la photolithographie.

WAT : De l'anglais Wafer Acceptance Test ou test d'acceptance du Wafer.

Annexe A

Paramètres et grandeurs physiques

TAB. A.1 – Paramètres et grandeurs physiques en technologie CMOS AMS 0,35 μm

Constante de Planck	$h = 6,626.10^{-34}\,J.s$
Constante de Boltzmann	$k = 1,3806.10^{-23}\,J.K^{-1}$
Charge élémentaire	$q = 1,602.10^{-19}\,C$
Célérité de la lumière	$c = 299\,792\,458\,m.s^{-1}$
Longueur d'onde dans le rouge	$\lambda_R = 680\,nm$
Longueur d'onde dans le vert	$\lambda_V = 546,1\,nm$
Longueur d'onde dans le bleu	$\lambda_B = 435,8\,nm$
Énergie de gap du silicium	$Eg = 1,12\,eV$
Permittivité diélectrique du dioxyde de silicium	$\epsilon_{SiO2} \approx 4{,}4$
Permittivité absolue	$\epsilon_0 = \frac{1}{36\pi.10^9}\,F.m^{-1}$
Mobilité des électrons	$\mu_N = 370\,cm^2.V^{-1}.s^{-1}$
Mobilité des trous	$\mu_P = 126\,cm^2.V^{-1}.s^{-1}$
Concentration d'atomes accepteurs (diffusion P^+)	$N_A = 101.10^{15}\,cm^{-3}$
Concentration d'atomes donneurs (diffusion N^+)	$N_D = 212.10^{15}\,cm^{-3}$
Concentration intrinsèque du silicium	$N_i = 145.10^8\,cm^{-3}$
Largeur minimale de grille d'un transistor	$L = 0,35\,\mu m$
Largeur minimale de canal d'un transistor	$W = 0,4\,\mu m$
Épaisseur de la couche de dioxyde de silicium	$T_{ox} = 7,6\,nm$
Capacité surfacique d'oxyde de grille	$C_{ox} = 5,119\,fF.\mu m^{-2}$
Coefficient de modulation de longueur du canal N	$\lambda_N = 0,05\,V^{-1}$
Coefficient de modulation de longueur du canal P	$\lambda_P = 0,15\,V^{-1}$
Tension de seuil d'un transistor NMOS	$V_{THN} = 0,46\,V$
Tension de seuil d'un transistor PMOS	$V_{THP} = -0,68\,V$

Annexe B

Carte XESS 3S1000 et pilote SVGA

B.1 Description du système

La carte XESS 3S1000 est une carte électronique programmable dédiée au contrôle numérique. Elle comporte un FPGA[1] (Spartan-3 1 million de portes), une mémoire SDRAM de 32 Mo, un CPLD et une mémoire Flash[2] de 2 Mo. Une mémoire EEPROM[3] série I2C[4] est également interfacée à la carte XESS, pour y stocker les paramètres de calibration de la caméra. La figure B.2 illustre l'intérieur du système de vision, on peut y voir, à gauche la carte époxy sur laquelle est soudée la rétine HISIC et à droite la carte XESS et la mémoire EEPROM série d'un mégaoctet.

B.2 Le pilote SVGA

Un port VGA est conçu pour visualiser les résultats et en particulier pour réaliser un démonstrateur de système de vision embarquée qui ne nécessite pas de liaison vers une unité de traitement externe.

[1]De l'anglais, Field Programmable Gate Array, comprenez, réseau de portes programmables in-situ, c'est le successeur des CPLD pour Complex Programmable Logic Device, comprenez, circuit logique programmable complexe.

[2]La mémoire FLASH possède les caractéristiques d'une mémoire vive mais dont les données ne disparaissent pas lors d'une mise hors tension.

[3]De l'anglais Electrically Erasable Programmable Read-Only Memory ou mémoire morte effaçable électriquement et programmable.

[4]De l'anglais Inter Integrated Circuit Bus, c'est le nom du bus historique, développé par Philips pour les applications de domotique et d'électronique domestique au début des années 1980, notamment pour permettre de relier facilement à un microprocesseur les différents circuits d'une télévision moderne.

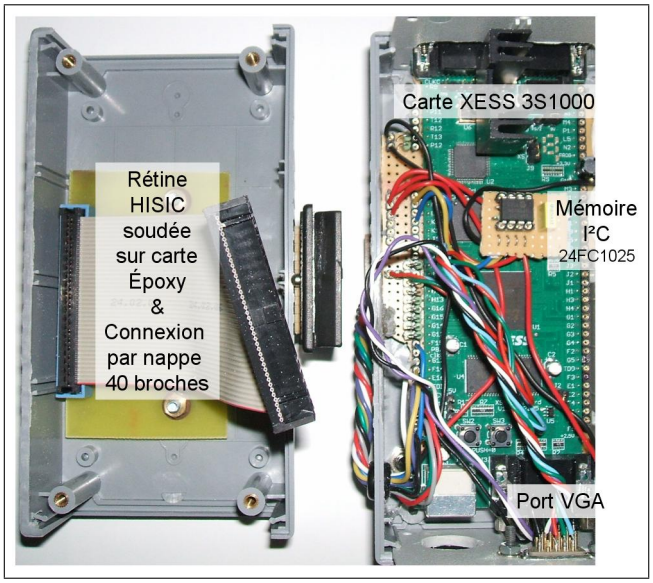


FIG. B.1 – Illustration photographique de l'intérieur du système de vision

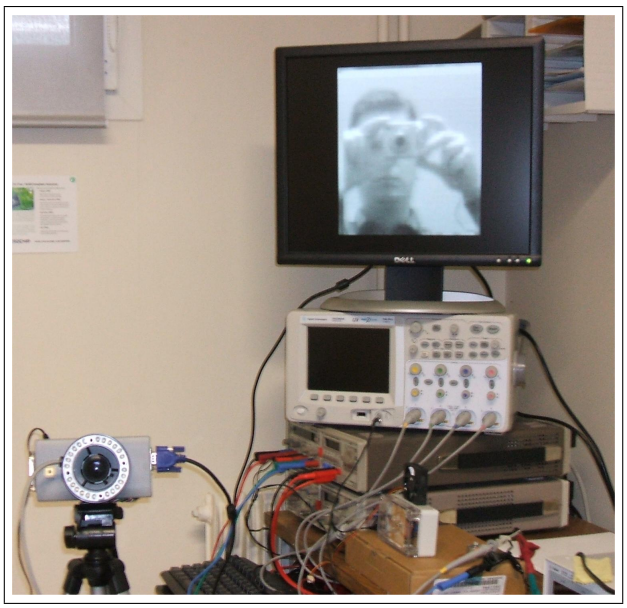


FIG. B.2 – Illustration photographique de l'utilisation du système de vision

Un réseau pondéré de résistances suffit pour réaliser un convertisseur numérique-analogique. La rétine fournie en sortie des valeurs de niveaux de gris sur 8 bits, il faut donc réaliser un contrôleur VGA avec un seul port en sortie, et de profondeur 8 bits. Il faut aussi interconnecter les trois entrées analogiques "red", "green" et "blue" du connecteur VGA pour ne former qu'une seule entrée

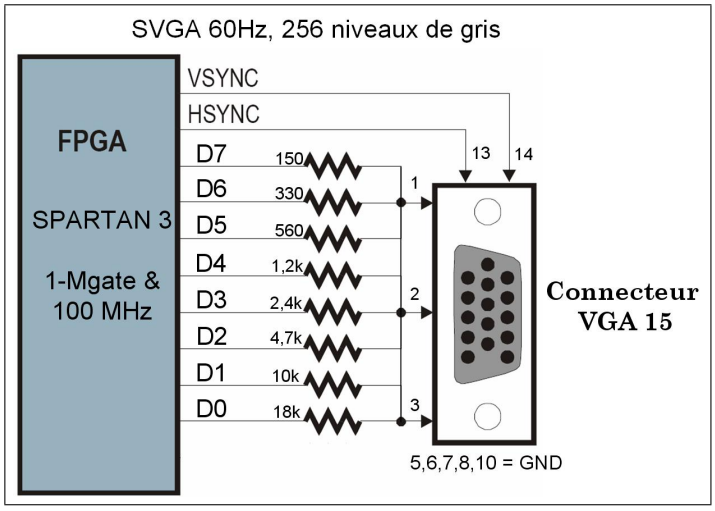


FIG. B.3 – Schéma de connexion du port VGA

d'impédance 25Ω (75Ω/3). La sortie du convertisseur doit être adaptée en impédance à l'entrée de l'afficheur VGA (25Ω). On note R la résistance connectée au bit de poids fort D7 (voir figure B.3), R_{VGA} l'impédance d'entrée du moniteur VGA ($R_{VGA} = 25\Omega$), V_{DD} la tension d'alimentation de la carte XESS (V_{DD}=3,3 V) et V_{VGA} l'amplitude maximale de tension à appliquer en entrée du moniteur VGA ($V_{VGA} \leq 0,85\,V$). L'adaptation d'impédance est alors vérifiée par la relation :

$$R \geq R_{VGA}\frac{(V_{DD} - V_{VGA})}{V_{VGA}}\sum_{n=0}^{7}\frac{1}{2^n} \approx 144\Omega \tag{B.1}$$

Parmi les valeurs de résistances de la série E12, on trouve la valeur 150Ω puis 330Ω ≈ 2 × 150Ω, etc. On construit ainsi le port VGA 256 niveaux de gris.

Le FPGA est un Spartan 3 cadencé à la fréquence d'horloge maître de 100 MHz, ce qui suffit à contrôler en continue un VGA à une résolution SVGA soit 800×600 à 60 Hz de fréquence synchro trames.

Résumé

Les travaux que nous présentons dans ce mémoire portent sur la conception, le test et la réalisation de capteurs d'images monolithiques CMOS rapides et "intelligents" : le principe, les performances, les limites et les perspectives sont le corps de ce mémoire. L'implémentation matérielle d'un système de vision programmable en est l'articulation centrale. Nous avons mis au point une plate-forme expérimentale pour l'instrumentation et l'évaluation des opérateurs rétiniens. Après un état de l'art sur l'imagerie rapide et sur les capteurs CMOS, la deuxième partie de ce mémoire est consacrée à l'étude et à la conception du pixel du capteur d'images. Nous avons conçu deux circuits HISIC.I et HISIC.II en technologie standard CMOS 0,35 μm (double-poly, quadruple-metal). Le premier a permis de déterminer un nouveau modèle de photo-détecteur, et le second de réaliser un prototype de caméra embarquée dédiée à l'imagerie rapide et aux traitements d'images linéaire. HISIC intègre des traitements programmables au niveau même du pixel. C'est une machine massivement parallèle de 4096 processeurs analogiques arithmétiques interconnectés selon une grille 64×64 en topologie 4-connexe. Enfin, la dernière partie du mémoire s'articule autour de la validation expérimentale du capteur, tous les résultats et procédures expérimentales y sont regroupés.

Mots clés : Système embarqué, imagerie rapide, traitements d'images sur plan focal, parallélisme massif, technologie standard CMOS, reconfigurabilité, programmabilité, vision artificielle.

Abstract

Our work presented in this thesis focuses on the design, testing and implementation of monolithics CMOS image smart sensors : The principle, performance and limitations. The hardware implementation of a vision smart system is the central link. HISIC is High Speed Image Capture with processing at pixel level. An experimental platform for instrumentation and evaluation of retina operators was conducted during this thesis. After a state of the smart sensors and CMOS retinas, the second part is dedicated to the study and design of the pixel image sensor HISIC. Two circuits were realized in CMOS technology. The first identified a new type of photo-detector, and the second, to create a prototype embedded camera dedicate to the speed imaging and image processing. It integrates processing programmable at pixel level. It is a massively parallel of 4096 analogue arithmetic processors inter-connected using a 64×64 topology 4-related. then the last part focuses on experimental validation of the sensor. All results and experimental procedures are represented. Finally an overview of the opportunities offered by the system are referred to conclude on this work.

Keywords : Embedded system, speed imaging, image processing on focal plane, massive parallel, standard CMOS technology, reconfigurability, programmability, artificial vision.