



**HAL**  
open science

# Architecture et Conception de Rétines Silicium CMOS : Application à la mesure du flot optique

David Navarro

► **To cite this version:**

David Navarro. Architecture et Conception de Rétines Silicium CMOS : Application à la mesure du flot optique. Micro et nanotechnologies/Microélectronique. Université Montpellier II - Sciences et Techniques du Languedoc, 2003. Français. NNT : . tel-00164035

**HAL Id: tel-00164035**

**<https://theses.hal.science/tel-00164035>**

Submitted on 19 Jul 2007

**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.

ACADEMIE DE MONTPELLIER  
UNIVERSITE MONTPELLIER II  
- Sciences et Techniques du Languedoc -

**T H E S E**

pour obtenir le grade de  
**DOCTEUR DE L'UNIVERSITE MONTPELLIER II**

Discipline : **Electronique, Optronique et Systèmes**  
Formation Doctorale : **Systèmes Automatiques et Microélectroniques**  
Ecole Doctorale : **Informations Structure Systèmes**

présentée et soutenue publiquement  
par

**David NAVARRO**

Le 17 Octobre 2003

**Architecture et Conception de Rétines Silicium CMOS :**  
**Application à la mesure du flot optique**

---

**JURY**

M. Michel Robert	Professeur à l'Université de Montpellier II	Président du jury
M. Gaston Cambon	Professeur à l'Université de Montpellier II	Directeur de thèse
M. Guy Cathébras	Maître de Conférences à l'Université de Montpellier II	Encadrant de thèse
M. Patrick Garda	Professeur à l'Université Pierre et Marie Curie de Paris VI	Rapporteur
M. Yang Ni	Professeur à l'Institut National des Télécommunications d'Evry	Rapporteur
M. Bernard Lamalle	Professeur à l'Université de Bourgogne	Examineur
M. Ian O'Connor	Maître de Conférences à l'Ecole Centrale de Lyon	Membre invité

**A ma famille**, présente, future.



# Remerciements

En premier lieu, je voudrais remercier Michel Habib, directeur du LIRMM, et Gaston Cambon, ancien directeur, pour m'avoir accueilli au sein du laboratoire pour effectuer mon DEA et ma thèse.

Ensuite, mes remerciements vont aux membres du jury : Patrick Garda et Yang Ni, qui ont accepté d'être les rapporteurs de ma thèse, Bernard Lamalle, Ian O'Connor, ainsi que Michel Robert, président du jury.

Je tiens à remercier particulièrement Guy Cathébras, qui m'a encadré durant la thèse. Ses qualités pédagogiques, sa rigueur, sa patience et sa gentillesse en font une personne exceptionnellement agréable à côtoyer. J'exprime toute mon estime à Gaston Cambon, directeur de thèse, qui a eu la gentillesse de corriger mon rapport et de me conseiller lors des réunions et pour la soutenance. Merci aussi à Thierry Gil pour son aide sur les cartes Altera.

J'exprime ma gratitude à Bernadette Domenges, du LAMIP de Caen. Son intervention sur le circuit de test a permis d'effectuer des mesures intéressantes et cruciales.

J'adresse aussi mes remerciements à tous les permanents du département Microélectronique au LIRMM, grâce auxquels j'ai effectué une thèse dans de très bonnes conditions.

Enfin, j'adresse ma sympathie à tous les doctorants que j'ai eu l'occasion de croiser, Gilles, Serge, Jean Max, Arnaud, Vence, Régis, Pascal, et tous ceux que j'oublie !

Une pensée particulière pour ceux avec qui j'ai traversé ces trois - voire quatre - dernières années (par ordre d'apparition) : Xavier (merci pour tous les bons moments), Benaïssa (le mécano), Yannick, Benjamin et Mariane (sayonara), Jean Marc (bientôt docteur ...enfin !), et Renaud.

"Les passions sont les vents qui enflent les voiles du navire ; elles le submergent quelquefois, mais sans elles il ne pourrait voguer."(Voltaire)

Bonne lecture,



<b>INTRODUCTION GENERALE .....</b>	<b>1</b>
<b>Chapitre I : LES CAPTEURS D'IMAGES .....</b>	<b>5</b>
I. Principes de base .....	5
I-1 Rappels d'optique .....	5
I-2 Phototransduction.....	8
II. Critères de performance d'un capteur d'image .....	12
III. Capteurs CCD.....	14
IV. Capteurs CMOS .....	17
IV-1 Topologie des capteurs CMOS .....	17
IV-2 Photorécepteurs à base de photodiode .....	18
IV-3 Capteurs CMOS actif (APS).....	22
IV-3-1 Mode d'intégration du photocourant.....	22
IV-3-2 Mode courant.....	25
IV-3-3 Mode logarithmique .....	26
V. Défauts dans les capteurs CMOS actifs à intégration du photocourant et leur correction .....	28
V-1 Défauts dans les capteurs CMOS actifs .....	28
V-1-1 Le courant de noir.....	28
V-1-2 Le bruit spatial fixe (BSF ou FPN) .....	29
V-1-3 Le bruit temporel.....	29
V-2 Double échantillonnage corrélé (CDS) .....	33
V-3 Double échantillonnage non corrélé (NCDS, ou DS) .....	34
VI. Perturbations lumineuses .....	34
VII. Rétines CMOS .....	36
VII-1 Différents types de rétines .....	36
VII-1-1 Les rétines à traitement spatial .....	37
VII-1-2 Les rétines à traitement spatio-temporel .....	39
VII-2 Rétines dédiées à l'estimation de mouvement .....	40
VII-2-1 Rétines basées sur les filtres spatio-temporels .....	40
VII-2-2 Rétines basées sur le contraste .....	41

VII-2-3 Rétines basées sur des structures de calculs numériques .....	44
---	----

## **Chapitre II : ETUDE ET CONCEPTION DE CAPTEURS ET DE RETINES ..... 47**

I. Rétine d'intercorrélation.....	47
I-1 Présentation .....	47
I-2 Fonctionnement du pixel.....	48
I-3 Principe.....	50
I-4 Résultats .....	51
II. Imageur CMOS.....	53
II-1 Contexte .....	53
II-2 Structure et fonctionnement.....	54
II-3 Choix du montage de caractérisation.....	59
II-4 Résultats et mesures.....	62
II-4-1 Lumière ambiante.....	62
II-4-2 Mesure du courant d'obscurité ou courant de noir .....	66
II-4-3 Capture d'images .....	67
II-4-4 Mesures des disparités du capteur.....	71
III. Conclusion.....	74

## **Chapitre III : L'ESTIMATION DE MOUVEMENT ..... 77**

I. Principe.....	77
II. Equation du flot optique .....	79
III. Méthodes pour estimer un mouvement .....	81
III-1 Méthodes différentielles .....	81
III-2 Méthodes fréquentielles.....	82
III-3 Méthodes de mise en correspondance .....	83
IV. Présentation de la méthode retenue (LTS: "Local Transform and Search") .....	87
IV-1 Présentation .....	87



IV-2 Algorithme.....	88
IV-2-1 Census Transform.....	88
IV-2-2 Recherche des codes de la transformée.....	91
IV-2-3 Complexité algorithmique.....	92
IV-3 Caractéristiques de l'algorithme LTS .....	93
V. Conclusion .....	94

## **Chapitre IV : VALIDATION DE L'ALGORITHME LTS .** ..... **99**

I. Evaluation de méthodes d'estimation du mouvement .....	99
I-1 Programmation de la SAD .....	99
I-2 Programmation de la méthode retenue.....	103
I-3 Comparaison entre les algorithmes SAD et LTS de manière logicielle	104
I-4 Modification de l'algorithme LTS en vue de son implantation matérielle ..	106
.....	106
II. Implantation matérielle de la méthode retenue .....	107
II-1 Implantation dans un FPGA .....	107
II-2 Implantation dans un ASIC - capteur et traitement.....	108

## **Chapitre V : RETINE D'ESTIMATION DU MOUVEMENT** ..... **113**

I. Architecture globale du circuit.....	113
II. Matrice de pixels : photoréception et codage.....	104
II-1 Comparateur en tension .....	116
II-1-1 Structure du comparateur .....	116
II-1-2 Fonctionnement.....	118
II-1-3 Simulations.....	119
II-2 Comparateur en temps .....	120
II-2-1 Structure du comparateur .....	120
II-2-2 Fonctionnement.....	121
II-2-3 Simulations.....	122

III. Recherche de codes .....	122
IV. Implantation dans un ASIC et caractérisation .....	124
IV-1 Dessin des blocs comparateurs.....	125
IV-2 Dessin des blocs de recherche des codes .....	127
V. Mesures.....	128
V-1 Mise en œuvre .....	128
IV-1-1 Présentation.....	128
IV-1-2 Problèmes rencontrés.....	131
V-2 Transformées des images.....	132
V-2-1 Acquisition de codes .....	132
V-2-2 Caractéristiques des codes.....	134
V-2-3 Mesure de la phototension .....	135
V-3 Interprétation des résultats.....	137
V-4 Comparaison des comparateurs par la simulation.....	137
V-5 Conclusion sur les matrices de codage.....	138
<b>CONCLUSION GENERALE ET PERSPECTIVES .....</b>	<b>141</b>
<b>REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>145</b>
<b>ANNEXES .....</b>	<b>155</b>

# INTRODUCTION GENERALE

Longtemps après la description de la "camera obscura" par Léonard de Vinci en 1515, les premiers appareils permettant la capture d'images sont apparus au XIX<sup>ème</sup> siècle. Ces derniers font partie de la famille des appareils photographiques argentiques. Ils ont comme principe une réaction chimique connue depuis le moyen-âge : le noircissement du nitrate d'argent exposé à la lumière. La première photographie stabilisée sur papier, réalisée par le français Daguerre, date de 1816. Ce procédé a été développé et amélioré durant le XX<sup>ème</sup> siècle, grâce à l'évolution de la chimie, mais la capture d'image ne trouvera d'alternative novatrice que dans les années 1930 avec l'apparition des tubes et particulièrement avec l'évolution qu'a connu le marché des semiconducteurs entre 1946 et le début des années 1970.

La technologie à base de silicium est la technologie la plus utilisée en Microélectronique. Elle compose la quasi-totalité des circuits actuels. Grâce à ses bonnes propriétés photoélectriques, elle englobe aussi les capteurs d'images. Les technologies silicium - ayant des propriétés photoniques - les plus utilisées sont depuis leur début les technologies CMOS et CCD. En 1970, la technologie CCD, de par sa nature et sa structure, se prête mieux à la conception de capteurs d'images que la technologie MOS, alors freinée par de gros pixels. Cependant, le marché de la technologie a beaucoup évolué entre 1970 et 1990, et la CMOS, dominante, a connu un essor particulier. Cet essor s'est traduit par une évolution de cette technologie, devenant de plus en plus fine. La réduction de la taille des transistors a permis de réaliser des circuits de plus en plus complexes, et de s'intéresser à nouveau aux capteurs d'images CMOS. Ceux-ci sont en effet meilleur marché compte tenu de la disponibilité et du volume de production du CMOS, et offrent enfin des pixels de taille comparable aux CCD. A la fin des années 1990, ces capteurs sont commercialement viables, et inondent le marché lors de l'avènement du multimédia.

Un des avantages de ces capteurs CMOS est qu'ils ne nécessitent pas de chaîne de production strictement dédiée. En effet, on peut utiliser les chaînes standard CMOS pour leur production (quelques paramètres du procédé de fabrication sont toutefois ajustés). Il est ainsi possible de créer des capteurs d'images, des périphériques de traitement ou des processeurs avec la même technologie. La philosophie des rétines CMOS est justement d'allier

phototransduction et traitement de l'image dans un même circuit : un système sur puce.

Les recherches dans ce domaine ont connu un fort regain d'intérêt dès les années 1990 lors du développement du multimédia. En effet, la thématique du traitement d'images se heurte fatalement à des masses de données très lourdes à gérer et donc à des calculs très importants.

Dans cette thèse, qui se veut prospective dans ce domaine – puisque pionnière au LIRMM -, nous allons détailler les architectures de capteurs d'images et de rétines CMOS que nous avons conçus et caractérisés. Le laboratoire a en effet travaillé en collaboration avec le LE2I du Creusot, mais aucun thème de recherche sur les rétines n'existait. En 1992, un capteur passif (PPS) de 16x16 pixels a été réalisé en technologie ES2 1,5  $\mu\text{m}$ . La caractérisation a révélé un problème d'éblouissement entre les pixels, causé par la conduction sous le seuil des transistors de sélection. La thèse de Eric Goujou [Gou-95] a aussi mené à des travaux communs. Ensuite, l'étude et la réalisation d'une rétine d'intercorrélation a fait l'objet de travaux de 1995 à 1999. Cette étude a mené à un brevet [Gor-99], et à un premier prototype menant à des publications dans les conférences [Lam-00]. Un deuxième prototype a été réalisé en 2000, publié dans [Voo-01a] [Cat-02] et dans une revue [Voo-01b].

Ma thèse a commencé en 2000 au LIRMM, avec pour but de découvrir les connaissances fondamentales des capteurs actifs, et d'étudier et concevoir une rétine effectuant une estimation du mouvement. En décembre 2002, Fabrice Gensolen a commencé sa thèse CIFRE (ST-Microelectronics) au laboratoire, dans le but de proposer des structures d'imageurs ou de rétines facilitant le calcul de l'estimation du mouvement.

Dans le premier chapitre, nous présentons le domaine des capteurs d'images. Nous expliquons la structure des capteurs à technologie CCD, pour comprendre ses caractéristiques et ses avantages qui en font une référence. Ensuite, nous détaillons les types de capteurs CMOS les plus utilisés, puis les rétines CMOS. Ces rétines ont de nombreuses applications possibles, comme le filtrage, la stéréovision, ou l'estimation du mouvement, application à laquelle nous nous sommes intéressés.

Comme nous l'avons vu, outre quelques collaborations anciennes, le thème de recherche portant sur la conception des capteurs d'images et de rétines est relativement nouveau au LIRMM. En conséquence, nous avons dû découvrir ce domaine. Le chapitre II présente donc l'étude, la conception et la caractérisation de circuits que nous avons menées, afin d'acquérir

des connaissances fondamentales.

Dans le troisième chapitre, nous présentons l'estimation du mouvement, qui constitue en soi un thème de recherche à part entière. De nombreuses solutions informatiques existent, mais même en ayant recours à des implantations matérielles, elles posent toujours des problèmes de temps de calcul, de surface, ou de consommation. C'est à cette application que nous nous sommes intéressés pour concevoir une rétine.

Le quatrième chapitre est destiné à l'étude de l'algorithme que nous avons choisi d'implanter. Nous avons donc programmé des algorithmes d'estimation du mouvement, et étudié leurs comportements face aux distorsions des images. Nous avons aussi évalué des implantations de type standard pour cet algorithme.

La conception de notre rétine d'estimation du mouvement est présentée dans le chapitre V. Nous y présentons sa structure, son fonctionnement, ainsi que les mesures effectuées.



# Chapitre I :

## LES CAPTEURS D'IMAGES

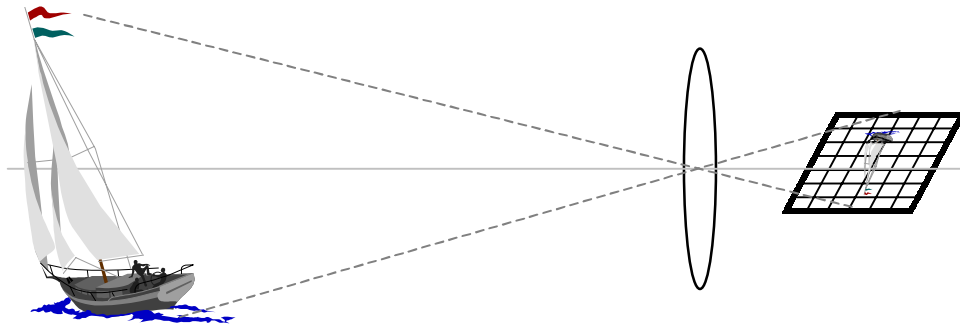
---

Ce chapitre veut faire l'état de l'art des capteurs d'images à semiconducteurs. Pour cela, nous introduisons ce domaine en présentant les principes de base (optiques et physiques). Les capteurs d'image se regroupent en deux catégories : les capteurs d'images standards (imageurs en technologie CCD ou CMOS) et les capteurs d'images "intelligents" (rétines en technologie CMOS). Dans un premier temps, nous expliquons la structure des capteurs CCD, référence dans le domaine des capteurs d'images, puis nous détaillons la structure des capteurs CMOS. Nous présentons aussi les problèmes qui nuisent à la qualité des images (défauts du capteur CMOS (bruit, sensibilité, diaphonie) et non linéarités de la lumière). Enfin, nous recensons les différents types de traitements effectués dans des rétines CMOS.

### I. Principes de base

#### I-1 Rappels d'optique

Les capteurs d'images qui nous intéressent ici se composent d'un circuit intégré générant une information électrique à partir de la quantité de lumière (photons) reçue. Le principe optique est le même pour tous les types de capteurs (y compris non électroniques) : une "cible" est placée derrière un objectif, à une distance adéquate pour que l'image s'y projette nettement. Pour les capteurs électroniques, la cible est composée d'une puce en silicium sur laquelle est disposée une matrice de pixels. Le capteur transforme alors l'image optique en image électronique (figure I.1).

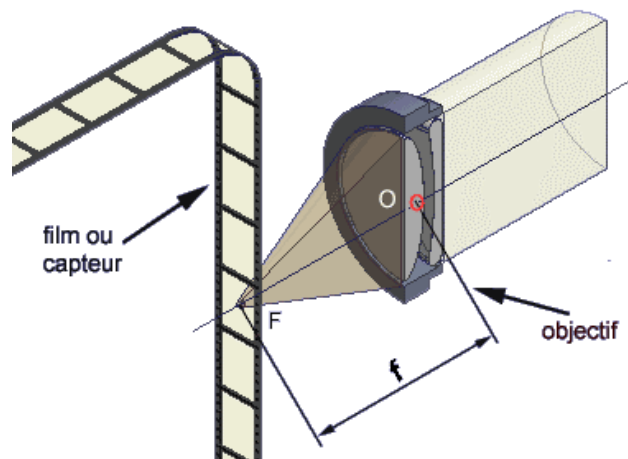


**figure I.1 : Projection d'une scène sur le capteur correctement placé**

Comme dans tout système optique, nous devons régler les paramètres de l'objectif pour projeter convenablement l'image sur la puce. Nous allons passer en revue ces différents paramètres (parfois interdépendants).

- Choix de la focale

La longueur (ou distance) focale ( $f$ ) est définie par la distance entre le foyer  $F$  (point où convergent les rayons lumineux) et le centre optique  $O$  (centre de l'objectif). Le foyer est le point de convergence de rayons lumineux arrivant sur l'objectif parallèlement à l'axe optique. Le choix de la focale est fait en fonction du champ de vision désiré et de la taille du capteur.



**figure I.2 : Distance focale d'un objectif [Pro-03]**

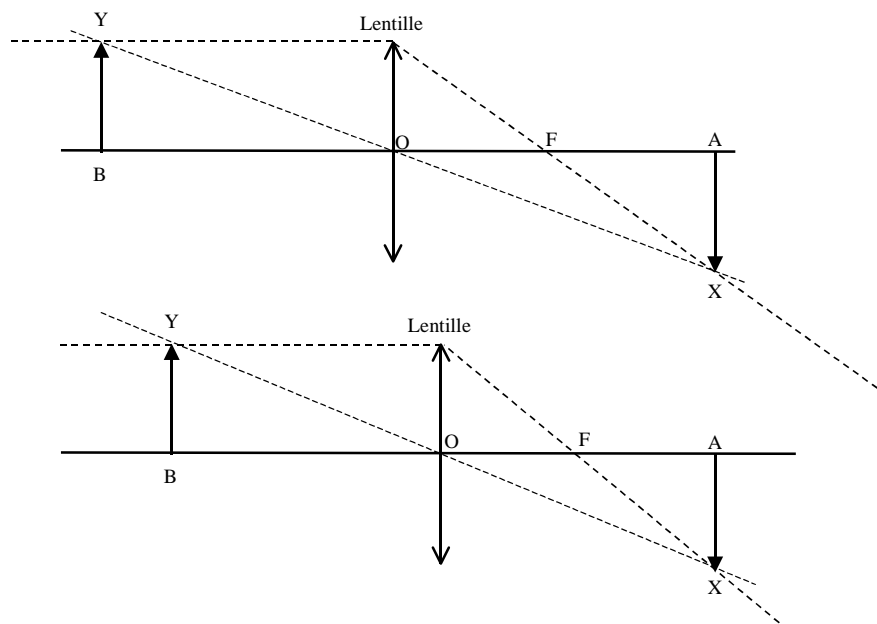
- Réglage de la netteté

Le réglage de la netteté (ou mise au point) permet d'ajuster le centre optique  $O$  de l'objectif par rapport à la distance de l'objet observé. Lorsque ce paramètre est correctement réglé, l'image est nette. La figure I.3 illustre la relation de géométrie optique :



$\frac{1}{OF} = \frac{1}{OA} + \frac{1}{OB}$ . Cette relation montre le lien entre la distance OB de l'objet à l'objectif et la distance OA de la lentille au capteur.

Pour une taille de capteur donnée, nous voyons qu'un mouvement de la lentille permet de projeter convenablement l'objet BY.



**figure I.3 : Optique géométrique et réglage de la netteté**

La netteté peut être altérée par un capteur électronique. En effet, il contient des couches d'oxyde qui entraînent une diffraction de la lumière. Ces paramètres sont à prendre en compte lors de la conception.

- Ouverture du diaphragme

Le diaphragme est une ouverture souvent circulaire dont le diamètre se règle, afin d'augmenter ou de diminuer la quantité de lumière qui entre dans l'appareil. L'éclairement est proportionnel à la surface de cette ouverture,  $\pi D^2/4$ . En photométrie, l'éclairement diminue avec le carré de la distance. Comme cette distance est la focale  $f$  ( $= OF$  sur la figure I.3), l'éclairement est proportionnel au rapport  $K.(D / f)^2$ . L'ouverture du diaphragme est définie par un chiffre, provenant du rapport de la focale de l'objectif par l'ouverture en millimètres du diaphragme ( $f / D$ ). Les ouvertures des diaphragmes sont normalisées, et ont pour valeurs une suite géométrique de raison  $\sqrt{2}$  (1 - 1,4 - 2 - 2,8 - 4 - 5,6 - 8 - 11 - 16 - 22 - 32 - 44 - etc ...), l'éclairement reçu sur le plan focal double à chaque changement d'ouverture.

Le diaphragme modifie aussi la profondeur de champ. Plus l'ouverture du diaphragme est petite et plus la profondeur de champ augmente. La profondeur de champ est la zone de netteté devant et derrière le sujet sur lequel la mise au point est faite.

- Temps d'exposition – ou vitesse d'obturation

Le temps d'exposition représente le temps nécessaire pour que la cible sensible à la lumière acquière la scène observée. Ce temps est bien sûr lié à la quantité de lumière reçue. En effet, le produit de l'intensité de la lumière moyenne avec le temps d'exposition doit être globalement constant et calibré par rapport à l'élément photosensible. Pour les systèmes argentiques, le temps d'exposition correspond au temps d'ouverture d'un obturateur mécanique, laissant le réactif du film (anciennement le nitrate d'argent) réagir avec la lumière. Plus la réaction chimique est longue, plus les photons cassent les liaisons du réactif, noircissant ainsi le film.

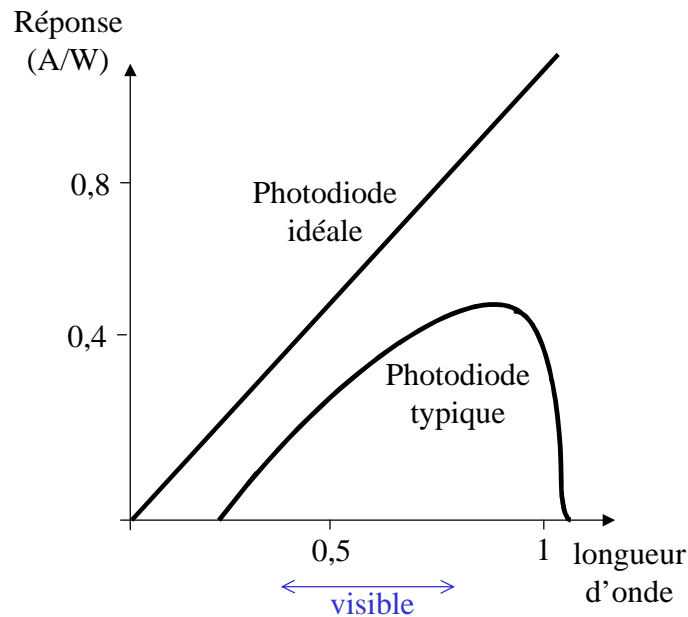
Les capteurs d'images standards n'ont pas toujours d'obturateur mécanique. C'est dans ce cas le temps de validation des circuits de conversion de l'information lumineuse qui modifie la valeur de l'information électrique. Le principe physique de ces capteurs est détaillé dans la section suivante.

## **I-2 Phototransduction**

Le phénomène de transduction est possible grâce à l'effet photoélectrique naturel du silicium. La figure I.4 montre que la réponse spectrale d'une photodiode silicium correspond aux spectres autour du visible, ce qui rend possible la réalisation de capteurs d'images à base de silicium.

Ce dernier est en effet sensible à une grande plage de longueur d'onde, couvrant le domaine visible ( $\lambda = 390$  à  $760$  nm) et le proche infrarouge ( $\lambda = 760$  nm à  $1,1$   $\mu\text{m}$ ).

Notons que la réponse spectrale est directement proportionnelle à l'efficacité quantique pour une longueur d'onde donnée [Joh-92]. L'efficacité quantique est le rapport entre le nombre d'électrons collectés et le nombre de photons incidents. Elle est liée à la structure du photorécepteur (profondeur de la jonction). Une photodiode idéale a une efficacité quantique qui vaut 1, représentée par une droite de coefficient directeur unitaire dans la figure I.4.



**figure I.4 : Réponse spectrale d'une photodiode silicium [Joh-92]**

Le silicium pur est un semiconducteur intrinsèque (composé d'un seul type d'atomes). L'énergie suffisante pour arracher un électron périphérique à son atome est de 1,12 eV. Pour augmenter la conductivité du matériau, on dope le semiconducteur, qui devient alors extrinsèque, puisque composé d'autres atomes (bore, gallium, phosphore, arsenic ...). Ces atomes, appelés impuretés ou dopants, composent le dopage, qui rend le matériau (extrinsèque) plus conducteur. Selon le dopant utilisé, le rapport entre les populations N et P est changé. Par exemple, un atome de bore (3 électrons sur la couche périphérique) - remplaçant un atome de silicium (4 électrons sur la couche périphérique) - entraîne la présence d'un trou qui attire les électrons. Le bore est appelé accepteur d'électrons. Le dopage ainsi effectué crée une zone de type P. De la même manière, un atome de phosphore (5 électrons sur la couche périphérique) joue le rôle de donneur d'électrons pour créer une zone de type N. Ces électrons et ces trous sont libres (ils peuvent se déplacer).

En mettant ces deux types (N et P) en contact, les électrons et les trous peuvent interagir. Des charges se recombinent, créant ainsi une zone de déplétion (ou zone de charge d'espace). On crée ainsi une jonction PN, caractérisée par une barrière électrostatique entre les zones P et N.

Comme le montre la figure I.5, l'effet de phototransduction se traduit par l'interaction des photons de la lumière et du semiconducteur extrinsèque. Ces photons apportent une énergie permettant de libérer des électrons périphériques.

L'énergie apporté par un photon vaut :

$$E = h.c / \lambda \quad (1.1.)$$

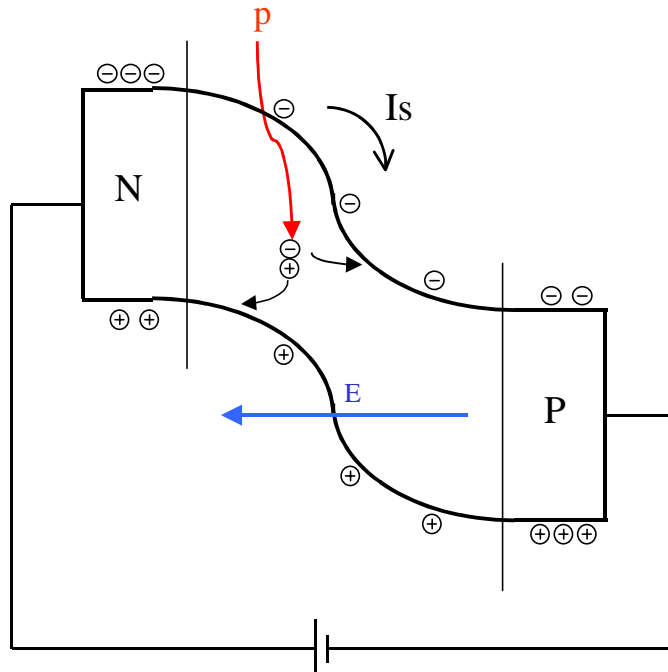
$h$  : constante de Plank ( $6,6. 10^{-34}$  J.s)

$c$  : célérité de la lumière

$\lambda$  : longueur d'onde.

L'énergie apportée est donc inversement proportionnelle à la longueur d'onde du rayonnement lumineux. En prenant comme énergie du photon 1,12 eV ( $1,8.10^{-19}$  J), la longueur d'onde correspondant au minimum d'énergie à apporter pour libérer un électron dans le silicium - est de  $\lambda = 1,1 \mu\text{m}$ . Cette valeur fixe donc la limite de sensibilité du silicium (figure I.4).

La séparation des électrons et des trous se fait dans la zone de charge d'espace des semiconducteurs extrinsèques (figure I.5). La jonction PN est alors polarisée en inverse.

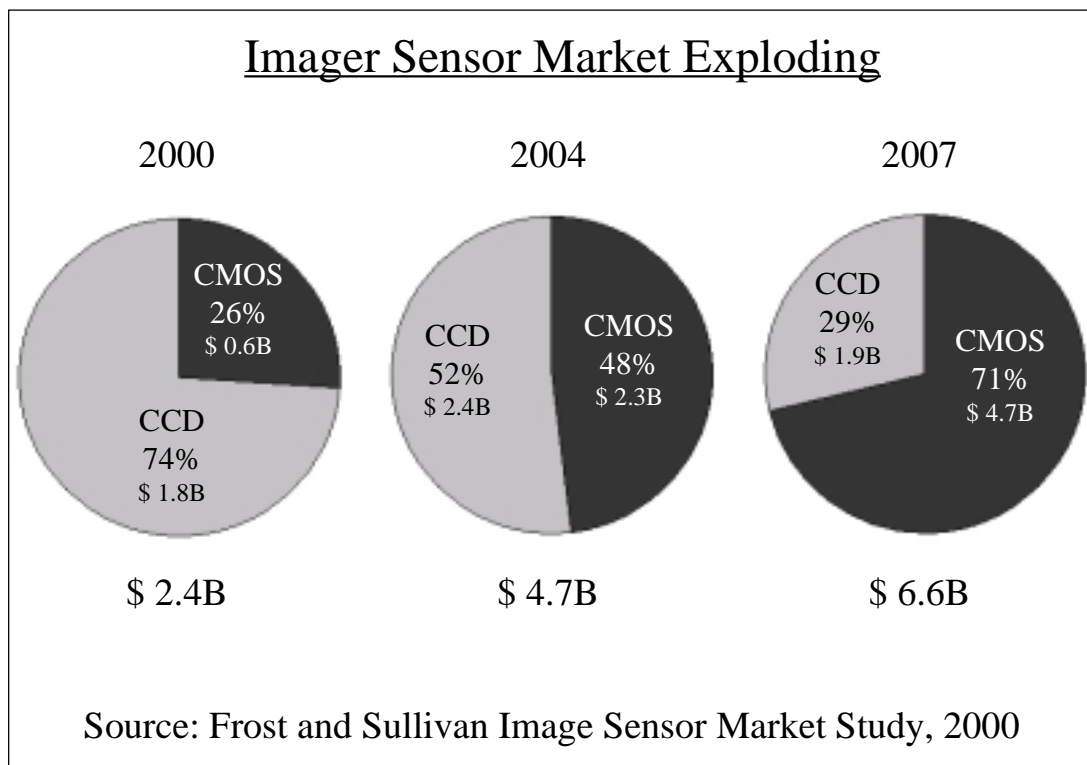


**figure I.5 : Séparation des paires électrons – trous dans une jonction polarisée en inverse**

Cette région est soumise à un champ électrique  $E$ . La valeur de ce champ électrique est d'autant plus élevée que la tension de polarisation de la jonction est grande. Les électrons et les trous sont alors séparés par le champ électrique  $E$  présent dans la zone. Ce phénomène entraîne des flux de charges électriques, proportionnels à l'information lumineuse.

La phototransduction – conversion du flux de photons en flux d'électrons - a un rendement inférieur à 1, à cause du rendement quantique (formule 1.2 dans la section suivante) et des recombinaisons d'électrons libres avec des trous.

L'industrie utilise principalement deux technologies silicium pour fabriquer des capteurs d'images : la technologie CCD et la technologie CMOS. La technologie CCD, que nous détaillons plus bas, est très utilisée; elle fait office de référence. La technologie CMOS (bien que plus ancienne et très utilisée par ailleurs) n'a réussi à trouver sa place sur le marché des capteurs d'images que vers la fin des années 90, lors de l'avènement du multimédia.



**figure I.6 : Prédiction de l'évolution du marché des capteurs d'images**

La figure I.6 montre la prédiction de l'évolution du marché dans le domaine des capteurs d'images. Ce marché est en pleine expansion puisque le chiffre d'affaire des ventes doit augmenter de 2,4 à 6,6 milliards de dollars en 7 ans ; on constate aussi que les capteurs CMOS devraient connaître une très forte expansion. Le marché des capteurs CCD resterait globalement constant, avec une légère augmentation en 2004. Cette augmentation temporaire pourrait correspondre à la vente de produits CCD destinés à des applications éligibles par des produits CMOS, lesquels prendront leur place ensuite. L'évolution de ce marché correspond au développement de produits multimédia créant ainsi une demande nouvelle dans le marché des capteurs.

Dans la section suivante, nous présentons les critères de performance des capteurs d'images, puis nous présentons ces deux technologies en donnant leurs caractéristiques principales.

## II. Critères de performance d'un capteur d'image

Les performances d'un capteur d'images se regroupent en différentes catégories [Goy-02]. Ces catégories sont les paramètres géométriques (résolution, taux de remplissage) [Che-00], électriques (dynamique, bruit, vitesse de lecture) [Mey-01] et physiques (rendement quantique, réponse spectrale) [Car-02] [Cat-00].

- La résolution

La résolution d'un capteur donne le nombre de points de l'image qui en sera issue. Elle s'exprime en nombre de pixels. Plus la résolution est élevée, plus le confort visuel est grand (les détails sont bien définis), mais plus le capteur est coûteux. Bien que certains appareils, comme les caméras ou les appareils photographiques numériques, offrent des résolutions toujours plus élevées, il faut préciser qu'il existe bon nombre d'applications ne nécessitant pas une très forte résolution. C'est le cas par exemple des caméras de surveillance ou des webcams, dont la priorité est un compromis entre la qualité d'image et le prix. De plus, une image à grande résolution entraîne une masse de données d'autant plus importante à traiter.

- Le facteur (ou taux) de remplissage

Il représente la configuration spatiale à l'intérieur du capteur, c'est à dire le rapport entre l'aire totale du pixel et l'aire photosensible. La différence entre ces deux surfaces est celle occupée par le routage, ou par les parties électroniques servant au transport ou au traitement de l'information. Un taux de remplissage peu élevé peut caractériser une sensibilité faible (l'aire photosensible est petite) ou un pixel de taille importante (beaucoup d'électronique est intégrée).

- La dynamique

La dynamique est la plage de luminosité dans laquelle le capteur est capable de mesurer correctement l'information. Elle est en fait comprise entre les deux niveaux extrêmes (du signal) que peut traiter le capteur. Le premier est le niveau de saturation du capteur, le second est le niveau du bruit. Les capteurs d'images ont donc des problèmes de sensibilité en lumière faible (faible flux), et des problèmes de saturation en lumière forte. Des capteurs à conversion logarithmique existent [Fil-00] [Kav-00] pour augmenter la dynamique (évitant la saturation), mais au prix d'un bruit plus important et d'une précision plus faible.

- Le temps de lecture

Le temps de lecture peut être important selon l'application envisagée. Il est ici considéré comme le temps mis entre l'instant de déclenchement de la capture d'image et l'instant où l'image est complètement sortie du capteur. Il existe des applications dans lesquelles la rapidité est primordiale, on parle ainsi de vidéo rapide [Bou-96]. Notons toutefois que la résolution et la vitesse de lecture sont deux paramètres inversement proportionnels. Ainsi, la lecture d'une image de résolution VGA (640 x 480 pixels) à 25 images/s entraîne un débit de 7,6 Mbits/s, et 153 Mbits/s à 500 images/s. Le débit de 153 Mbits/s permet aussi la lecture d'une matrice de 2400 x 2400 pixels à 25 images/s.

- Le bruit spatial fixe ("BSF" ou "FPN")

Un des défauts rencontrés est la dispersion qui peut exister entre différentes zones d'un même capteur. Ces dispersions, indépendantes du temps, proviennent de non uniformités rencontrées lors du processus de fabrication. Comme elles ne sont pas prévisibles avant la fabrication du circuit, il faut donc recourir à des méthodes de compensation pour pouvoir les atténuer *a posteriori*. Pour cela, des techniques de calibration permettent de compenser ces défauts. Les techniques de réduction du bruit spatial fixe sont détaillées plus loin.

- Le rendement quantique et la réponse spectrale

Le rendement quantique exprime le nombre de photons qui génèrent une paire électron-trou. La sensibilité de la structure est liée à ce paramètre (ainsi qu'au taux de remplissage). Le rendement effectif d'un pixel est en fait égal à :

$$\eta_{\text{eff}} = \eta \cdot F_r \quad (1.2.)$$

$\eta$  : rendement quantique

$F_r$  : facteur de remplissage du pixel

Comme l'explique J. Goy dans sa thèse [Goy-02], le rendement quantique  $\eta$  dans les capteurs CMOS dépend de la profondeur de la jonction séparant les paires électrons-trous et du type de photodiode choisi. Une jonction plus profonde permet en effet d'améliorer le rendement quantique pour les grandes longueurs d'onde ( $\lambda > 600$  nm) – de faible énergie (voir équation 1.1).

Le rendement quantique est donné pour une longueur d'onde précise. La représentation du rendement quantique en fonction de la longueur d'onde est la réponse spectrale.

La réponse spectrale est la gamme de longueur d'onde pour laquelle la structure est sensible. Pour le silicium, elle s'étend de 400 nm (pour le violet) à 1100 nm (pour le proche infrarouge) - voir figure I.4. Précisons que la sensibilité n'est pas uniforme sur la plage de longueur d'onde. Ainsi, selon la longueur d'onde de la lumière renvoyée par l'objet qu'observe un capteur en niveaux de gris, cet objet paraîtra plus ou moins lumineux.

Nous allons présenter les deux technologies CCD et CMOS utilisées pour l'imagerie. Deux phases caractérisent l'acquisition d'une image par un capteur électronique : la génération des paires électrons-trous par les photons (phototransduction), puis la collection des électrons. De manière générale, en CCD comme en CMOS, la phase de phototransduction peut être considérée comme identique. La différence entre ces deux technologies réside dans le mode de transfert (ou de lecture) des données vers la sortie du circuit. Les capteurs CCD ont la particularité de transférer l'information, sous forme de charges analogiques, de pixel en pixel. Les capteurs CMOS classiques transmettent des tensions ligne par ligne à travers des colonnes.

### **III. Capteurs CCD**

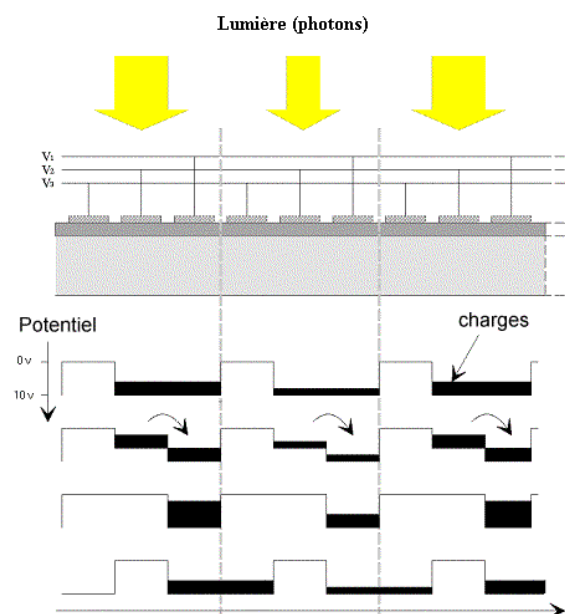
Les capteurs de type CCD, Charge Coupled Device, ont été les précurseurs dans le domaine de l'acquisition d'images à base de circuit intégré. Le principe de transfert de charges caractérisant cette technologie date de 1970 (laboratoire Bell) [Gou-95], et les premiers appareils à base de capteur CCD ont été commercialisés par Sony en 1981. Dans un premier temps réservés aux professionnels de l'image, ils ont peu à peu remplacé les caméras amateur (à film argentique) dans les années 80, puis les appareils photographiques avec l'avènement du numérique dans les années 90. Grâce à ses propriétés très intéressantes et donc à son développement, la qualité des procédés CCD a atteint une grande maturité, et place ce type de capteurs comme référence dans l'imagerie [Jan-02]. Nous en donnons donc le principe de fonctionnement.

La lumière incidente crée dans chaque zone photosensible une quantité de charge proportionnelle à la quantité de lumière reçue. Dans un capteur de type CCD, chaque pixel est composé d'une zone photosensible (photoMOS ou photodiode) connectée à des capacités MOS. Une fois la phase d'acquisition terminée (après un temps d'accumulation des photocharges), il faut véhiculer ces informations vers l'extérieur du capteur. Pour cela, les



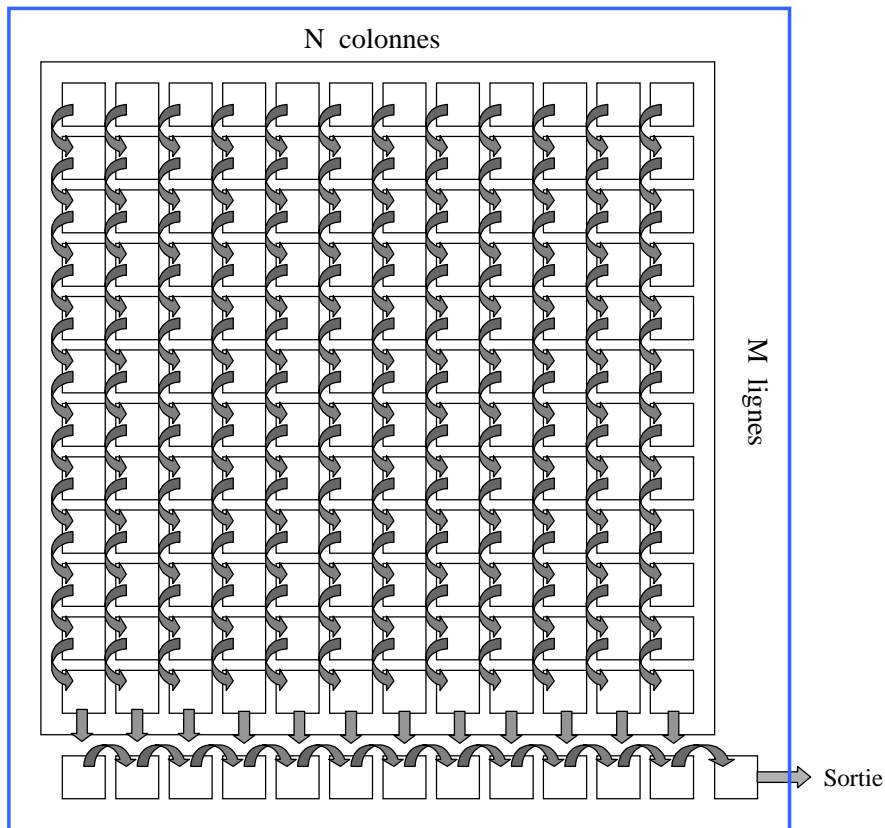
charges vont être transférées de capacité en capacité et ainsi de pixel en pixel jusqu'à la sortie, en utilisant un registre à transfert de charges analogiques CCD.

Chaque pixel est donc composé d'une partie photosensible et d'un registre à transfert de charges analogiques (opacifié). Comme le montre la figure I.7, chaque pixel comprend 3 capacités. Pour un registre en ligne, tous les capacités des pixels sont branchées de la même façon, de manière à opérer en parallèle. L'excitation des grilles des capacités MOS permet de les charger et de les décharger, afin de véhiculer des charges de capacité en capacité et de pixel en pixel. Ce séquençage est cependant complexe et nécessite différents niveaux de tension. Des circuits périphériques dédiés au contrôle composent alors le système de caméra.



**figure I.7 : Transfert de charges**

La figure I.8 illustre le fonctionnement (exemple le plus basique) nécessaire pour évacuer les informations hors du capteur. Les charges sont décalées de manière série à travers les capacités des lignes et des colonnes de la matrice. Pour une matrice de M par N pixels, il en résulte un nombre total de  $N(M+1)$  décalages.



**figure I.8 : Architecture d'un capteur d'images CCD**

Pour obtenir un tel fonctionnement, le circuit est piloté par un circuit de commande générant plusieurs horloges avec différentes tensions. En effet, les transferts (charges et décharges) sont réalisés en appliquant des tensions élevées et convenablement agencées sur les capacités.

En sortie de la matrice, les charges sont transférées sur une capacité de faible valeur. La tension ainsi obtenue est ensuite amplifiée à l'extérieur du capteur pour attaquer un convertisseur analogique numérique (généralement de résolution 8 à 15 bits).

Les capteurs CCD font partie de la vie courante, puisqu'on les retrouve aussi bien dans les caméscopes d'ancienne génération que dans les derniers modèles haut de gamme de caméras et d'appareils photographiques numériques, en passant par les caméras de surveillance ou encore les WebCams. Bien qu'ils offrent des caractéristiques intéressantes, ils imposent tout de même certaines contraintes. Nous allons donner un aperçu de leurs caractéristiques.

Les circuits CCD ont la particularité d'offrir des performances optiques et électriques très intéressantes, notamment leur qualité d'image et leur bruit relativement faible. Ces qualités sont le fruit d'une technologie mûre qui a été modifiée et dédiée pour cet usage unique, mais de ce fait coûteuse.

Cependant, bien qu'optimisée, les transferts de charges entraînent des pertes (de l'ordre de  $10^{-5}$  par transfert). Ce défaut limite la taille des capteurs par rapport à une qualité définie dans le cahier des charges. En effet, le rendement du transfert analogique est lié au temps mis pour effectuer ce transfert. Aussi, tous ces décalages de charges se traduisent par une consommation électrique et un temps de lecture élevés. En effet, nous avons vu qu'il faut effectuer  $N(M+1)$  décalages pour obtenir l'information complète d'une image. Aussi, la sérialisation des données impose une utilisation stricte (lecture complète de l'image) et ne permet pas de se focaliser sur une zone précise de l'image de manière optimale. En effet, même s'il est possible de sous échantillonner la matrice de manière régulière (ligne / colonne), il est impossible d'accéder aléatoirement à des pixels. Il n'existe pas de mécanisme d'adressage. Ceci soulève directement un autre problème : le débit d'information est assez limité, et il n'est pas possible de l'augmenter artificiellement de façon notable. Les transferts de charges maximum sont classiquement cadencés à 40 MHz, ce qui entraîne des débits d'images à la cadence vidéo.

Finalement, ces composants CCD ne sont pas pourvus de zones de diffusions N et P adéquates à une électronique de traitement au sein même du capteur. Cette technologie rend donc quasi-impossible (ou du moins difficile et coûteuse) l'intégration de tels capteurs dans un système sur puce.

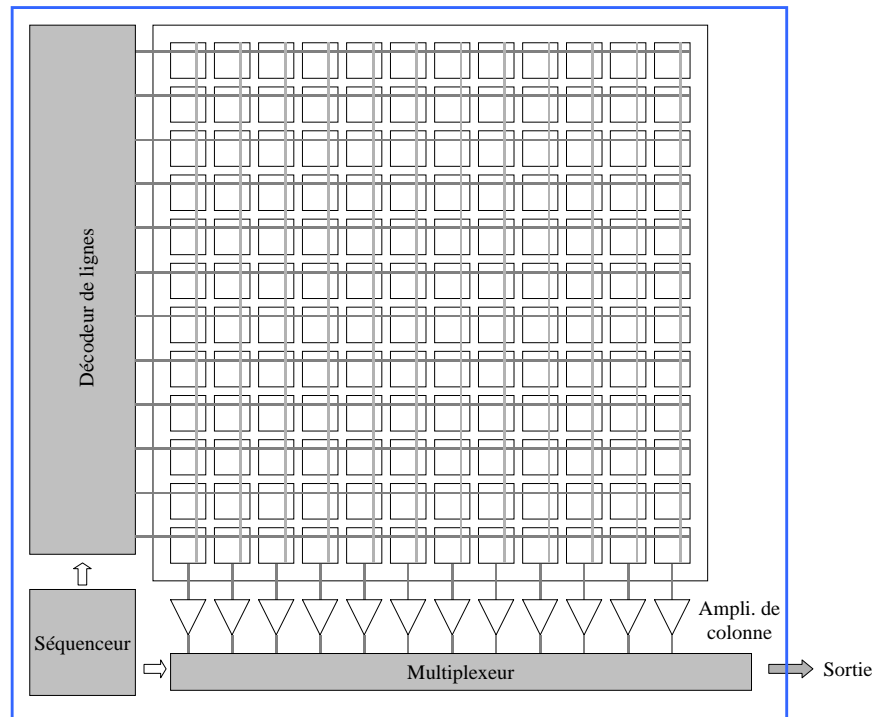
## IV. Capteurs CMOS

Bien que cette technologie date de 1963 (la longueur du canal était alors de  $10\mu\text{m}$ ) [Wan-63] et que l'effet photoélectrique des circuits CMOS soit connu et étudié depuis longtemps, il n'a été exploité industriellement que récemment. Ceci est dû principalement à la "légendaire" grande taille des pixels CMOS (ne permettant, compte tenu du coût, que des matrices de faible résolution), et à leur faible rapport signal / bruit.

La forte densité d'intégration, liée à l'arrivée des technologies sub-microniques, a permis de mieux maîtriser les technologies et d'atteindre des pixels de quelques micromètres de côté, de taille équivalente aux pixels des capteurs CCD, et donc des résolutions intéressantes de plusieurs millions de pixels. Beaucoup de personnes se sont intéressées à cette alternative, par exemple [Fil-00] [Tia-01b], car les capteurs d'images fabriqués en technologie CMOS ouvrent des horizons et des perspectives de conception nouvelles [Fos-97].

## IV-1 Topologie des capteurs CMOS

Les capteurs d'images CMOS, ou imageurs CMOS, ont une architecture générale présentée figure I.9.



**figure I.9 : Architecture d'un capteur d'images CMOS**

Le principe général est le même que pour tout capteur d'image : l'information optique est convertie en information électrique, qui est acheminée vers l'extérieur du composant. La structure de tels circuits est par contre spécifique, car la cohabitation des photorécepteurs et de transistors performants au sein d'un même système permet une intégration globale. Un séquenceur, via un décodeur de ligne, commande une ligne parmi les  $M$  lignes de la matrice, permettant la lecture de celle-ci. Les données de la matrice sont donc accessibles aléatoirement, par adressage [Mey-98b]. Des amplificateurs de colonne et un système de multiplexage améliorent la dynamique et aiguillent les données vers la sortie.

Différents types de capteurs CMOS existent (photodiode, phot grille, phototransistor), mais la photodiode est le photorécepteur présentant le meilleur compromis [Cav-98][Goy-92]. En effet, les photogrilles et phototransistors sont composés d'une couche de polysilicium, qui effectue un filtrage et opacifie la surface photosensible. Nous nous intéressons donc aux capteurs CMOS actifs à base de photodiode, les plus courants [Ogi-97].

## IV-2 Photorécepteurs à base de photodiode

La lumière incidente génère des paires électrons-trous, donc un photocourant. La figure I.10 montre que cet apport en courant a pour effet de décaler la caractéristique statique courant / tension de la diode. Lorsque le photocourant est nul, la caractéristique est celle d'une diode classique. C'est en fait la zone de déplétion de la jonction PN qui permet de collecter les charges libérées par les photons. Toute diode de ce type réalisée dans un circuit à base de silicium est sensible à une gamme de longueur d'onde donnée (fonction de la profondeur de la jonction).

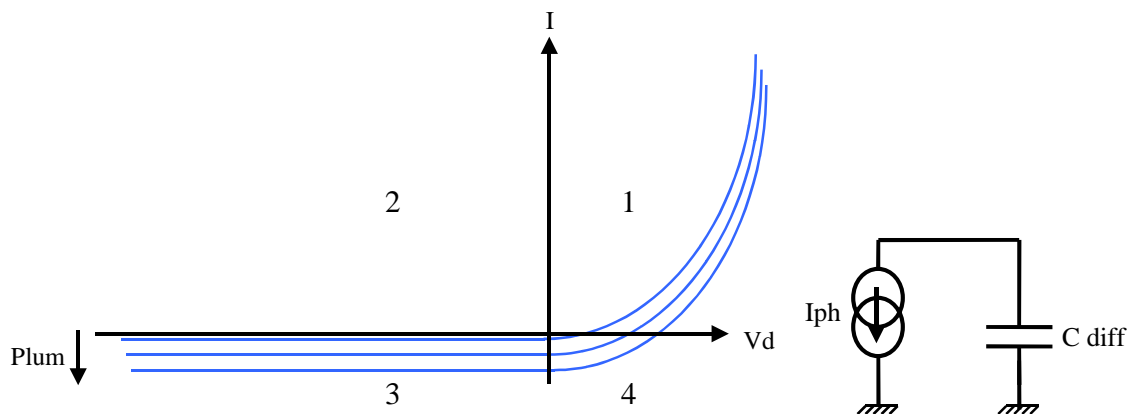


figure I.10 : Caractéristique d'une diode et schéma équivalent dans le 3<sup>ème</sup> quadrant

La caractéristique de la photodiode est :

$$I = I_s \cdot (e^{\frac{q V_d}{n k T}} - 1) - I_{ph} \quad (1.3.)$$

Où I est le courant qui traverse la photodiode et  $I_{ph}$  le photocourant.

Compte tenu des caractéristiques de la photodiode, deux modes de fonctionnement sont possibles, en polarisant la diode dans le troisième ou le quatrième quadrant.

En polarisation directe (dans le quatrième quadrant), la photodiode fonctionne dans le mode photovoltaïque. Dans ce cas, on s'intéresse à la tension  $V_d$  aux bornes de la photodiode, qui a une relation logarithmique avec le courant ( $V_d = \frac{n k T}{q} \cdot \ln \frac{I + I_s + I_{ph}}{I_s}$ ). Généralement, la photodiode est connectée à un transistor suiveur (donc on peut négliger I), la relation peut alors s'écrire :

$$V_d = \frac{n k T}{q} \cdot \ln \frac{I_s + I_{ph}}{I_s} \quad (1.4.)$$

Ce mode permet une lecture directe de l'information en tension, cependant le système RC composé de la résistance dynamique et de la capacité de la diode limite la rapidité de lecture. Aussi, les variations du photocourant peuvent entraîner des variations très faibles de tension. Nous expliquons ce mode de fonctionnement plus loin.

La polarisation dans le troisième quadrant est intéressante pour l'allure constante du courant. Dans ce cas,  $V_d$  est négatif, l'équation 1.3 peut alors s'écrire  $I = -I_s - I_{ph}$ . La photodiode est alors assimilable à un générateur de courant commandé par la puissance lumineuse. Le schéma équivalent est composé d'un générateur de courant en parallèle avec la capacité de jonction entre la diffusion et le substrat (figure I.10). Cette capacité dépend des paramètres  $C_j$  et  $C_{jsw}$  de la diffusion, ainsi que de ses dimensions. En considérant  $A$  l'aire de la diffusion et  $P$  son périmètre, la capacité est calculée par la relation :

$$C = A.C_j + P.C_{jsw}. \quad (1.5.)$$

$A$  : Aire de la diffusion

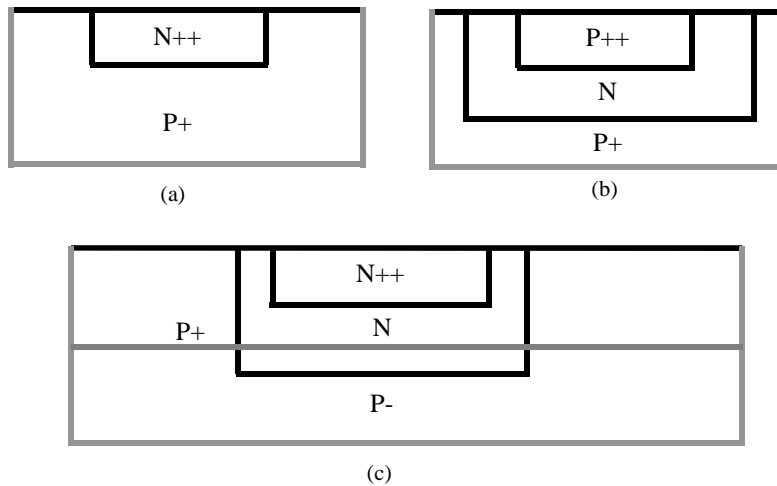
$P$  : Périmètre de la diffusion

$C_j$  : Capacité surfacique (à 0v)

$C_{jsw}$  : Capacité périmétrique (de bords) (à 0v)

Ces capacités  $C_j$  et  $C_{jsw}$  sont à calculer pour une polarisation donnée (voir équation 1.6)

Les photodiodes les plus courantes sont faites avec les diodes à jonction et les diodes à puits N (figure I.11). Les diodes à jonction PN sont de type N (a) : diffusion N dans substrat P ( $N_{diff} / P_{sub}$ ), ou de type P (b) : diffusion P dans puits N ( $P_{diff} / N_{well}$ ). En technologie CMOS 0.35 $\mu$ m, les profondeurs de jonctions sont de 200 nm. Les diodes N ont un meilleur rendement que les diodes P, en plus de contraintes de dessin moins importantes [Gou-95]. En effet, elles ne nécessitent pas de puits N qui imposent des distances minimales importantes. Les diodes de type N sont dans un substrat P faiblement dopé, alors que les diodes de type P sont dans un puits N plus dopé. Or, la zone de charges d'espace se développe d'autant plus que le semi-conducteur est faiblement dopé, et la capacité de jonction est plus faible. Une zone de charges d'espace plus grande permet statistiquement une plus grande collecte des charges libérées par les interactions entre les photons et les électrons périphériques. Les diodes de type N sont donc plus intéressantes.



**figure I.11 : Types de photodiodes: diffusion N (a), diffusion P (b), puits N (c)**

Une autre catégorie de diode a été développée à l'IMEC, en Belgique [Die-97a]: la diode à puits (c). La société Fillfactory [Fil-00] commercialise en effet des capteurs de ce type. Son principe est de créer une diffusion N dans un puits N, afin de collecter les charges, en profondeur, isolées grâce à la barrière électrostatique créée par la zone P+. Cette méthode est intéressante dans la mesure où elle offre un taux de remplissage bien supérieur aux méthodes classiques. Les photodiodes à puits N ont l'avantage d'offrir une zone photosensible importante, permettant d'atteindre un taux de remplissage proche de 100%. Ce taux de remplissage important vient du fait que la zone de charges d'espace se fait dans une zone très faiblement dopée (P-), donc elle est très développée. De plus, la capacité de jonction est faible. Cependant, du fait de la profondeur de la zone de collection plus importante, seules des longueurs d'ondes plus grandes généreront des paires électrons-trous dans cette zone, atténuant les composantes bleues de l'image. Les auteurs [Die-97a] proposent aussi de créer un passage à travers la zone P+ en dessinant simplement un implant de puits N, formant ainsi une structure similaire aux CCD ou aux photoMOS.

Notons que la technologie CMOS n'est pas prévue à l'origine pour fabriquer des capteurs d'images, cependant elle offre des propriétés intéressantes. Des études ont montré que les capteurs CMOS passifs souffrent d'un fort bruit [Cou-00] et de diaphonie entre les pixels (effets d'éblouissements) [Gam-98]. Les capteurs passifs se sont donc peu développés, nous ne considérons par la suite que les capteurs de type actif. Les capteurs actifs sont caractérisés (par rapport aux capteurs passifs) par une amplification dans le pixel.

### IV-3 Capteurs CMOS actif (APS)

Les capteurs CMOS actifs ("Active Pixel Sensor") comptent plusieurs structures. Nous détaillons dans cette partie le mode d'intégration du photocourant, le mode courant, et le mode logarithmique. Ces modes de fonctionnement sont les plus courants dans les capteurs CMOS actuels. Ils utilisent tous une photodiode, mais ils ont des structures électroniques qui lisent l'information différemment.

#### IV-3-1 Mode d'intégration du photocourant

Le principe de fonctionnement d'une photodiode en mode tension (ou intégration du photocourant) est basé sur le phénomène de charge et de décharge de la capacité équivalente. Le schéma de la figure I.12 montre la structure de base d'un tel montage. Le capteur actif le plus classique compte trois transistors. Ces transistors sont de type N afin d'optimiser la surface, évitant ainsi les contraintes spatiales liées à un puits N. Un unique puits P est nécessaire sur toute la matrice. La surface de l'électronique est ainsi minimale, offrant un rapport surface sensible / surface totale (le taux de remplissage) maximal. La sortie de la photodiode est reliée au transistor de Reset.

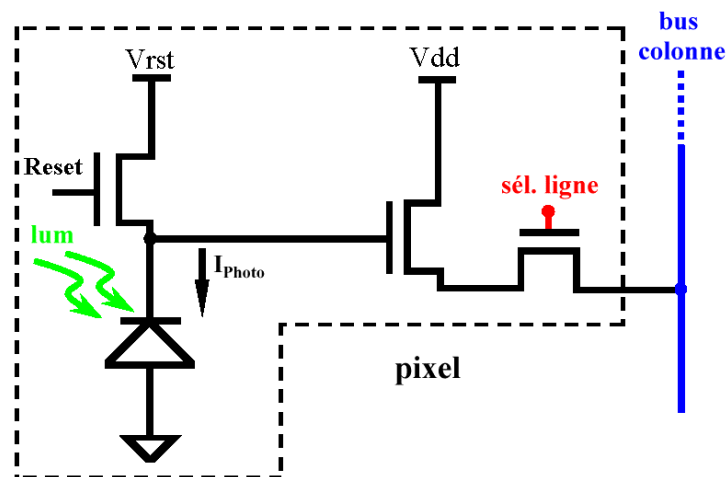


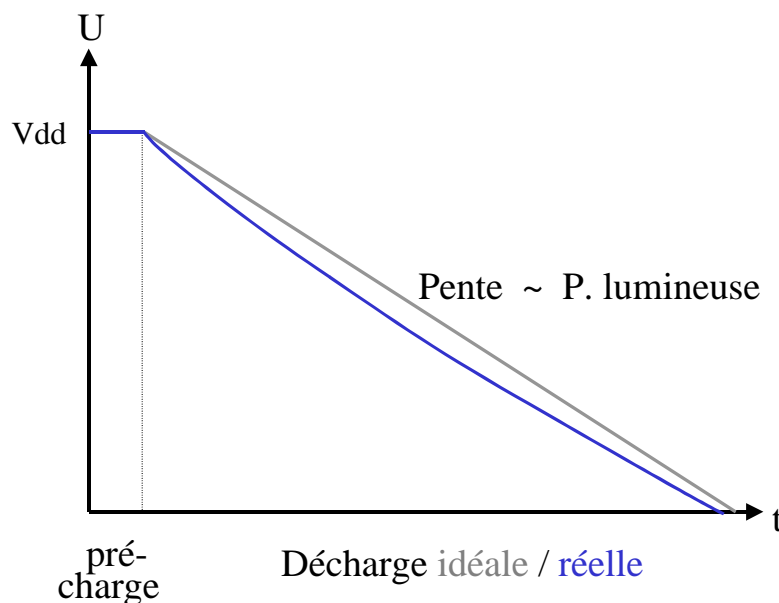
figure I.12 : Montage de la structure en mode d'intégration du photocourant

Le transistor de Reset sert à précharger la diode à une tension connue, proche de  $V_{rst}$ . Ce transistor doit être rapide, puisque le temps du reset dépend du courant qui le traverse pour charger la photodiode. Comme il est de type N, la tension de précharge est dégradée. La tension aux bornes de la photodiode est en effet de  $V_{dd} - V_t$  (avec  $V_{rst} = V_{dd}$ ). Si la tension



de l'entrée "Reset" est à  $V_{dd}$ , le transistor est en régime de saturation ( $V_{gs} = V_{ds}$ ). Pour éviter ce phénomène, une solution consiste à appliquer une tension plus élevée sur la grille ( $> V_{dd}$ ), ou à prendre  $V_{rst} < V_{dd}$ , de manière à ce que le transistor soit toujours en régime triode. Une troisième solution consiste en l'utilisation d'un transistor de type P, mais il impose alors un puits N qui nécessite des distances de séparation élevées - afin de respecter les règles technologiques du fondeur. Une dispersion sur les caractéristiques des transistors de Reset entraîne un temps de précharge qui peut sensiblement varier d'un pixel à l'autre. Une erreur est possible si ce temps est trop court, causant un offset.

Lorsque la phase d'initialisation est terminée, si nous considérons que le potentiel aux bornes de la photodiode est  $V_{dd}$ , la décharge de la capacité de la photodiode se fait proportionnellement à l'intégrale de la valeur du photocourant. On néglige ici la faible capacité de grille du transistor suiveur. La capacité se décharge à courant constant si la lumière est constante. 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 (figure I.13). Plus le photocourant est important, plus la décharge est rapide. En réalité, l'allure de la décharge est légèrement différente. Les phases d'initialisation et de décharge sont présentées figure I.13.



**figure I.13 : Précharge puis décharge de la photodiode**

Cette non-linéarité est due à des paramètres technologiques ( $V_j$  - potentiel électrostatique de la jonction PN - et  $M$  - coefficient de gradation de la jonction-) introduisant une variation de la capacité de la diode fonction de la tension  $V_d$  à ses bornes. Elle peut se modéliser par la

relation [Per-94]:

$$C = \frac{C_j}{\left[1 - \frac{V_d}{V_j}\right]^M} \quad (1.6.)$$

C<sub>j</sub>: capacité de la jonction non polarisée (à 0v)

V<sub>d</sub>: tension de polarisation de la photodiode

V<sub>j</sub>: potentiel électrostatique de la jonction PN

M: coefficient de gradation de la jonction. 0,3 (jonction abrupte) < M < 0,5.

L'ordre de grandeur de la capacité d'une photodiode est de 100 fF. Cette valeur est purement indicative, puisqu'elle dépend de nombreux paramètres, comme le type de la photodiode, sa taille, la technologie utilisée et la tension de polarisation de la jonction.

Un transistor suiveur sert d'adaptation d'impédance (entre la photodiode et la colonne), et attaque un transistor de sélection de ligne. Ce transistor doit avoir les mêmes propriétés dans tous les pixels, la dispersion entre les transistors suiveurs est souvent à l'origine de défauts que nous étudierons plus tard.

Comme nous avons vu précédemment, toutes les colonnes d'une ligne sont adressées en même temps, l'information analogique de chaque pixel sélectionné dans la ligne circule alors sur un bus colonne. En bas de la matrice, des amplificateurs de colonne reçoivent l'information électrique d'une ligne entière. Ces amplificateurs de colonne génèrent le défaut prédominant dans ce type de structure : des erreurs de gains dues aux dispersions créent le bruit de colonne que nous verrons plus loin.

Un des avantages du mode à intégration du photocourant est que la plage dynamique de l'information électrique est connue et reste fixe : elle varie ici de la valeur de la précharge, tension connue, à 0. De plus, travailler en mode tension permet de contourner les problèmes liés aux faibles photocourants mis en jeu. Enfin, des techniques de réduction du bruit sont facilement utilisables (ces techniques sont expliquées plus loin). L'inconvénient du mode à intégration est que la dynamique de mesure est limitée. Aussi, un temps d'exposition, contrôlé par l'entrée "Reset", doit être choisi en fonction des conditions lumineuses.

Dans les capteurs actifs classiques de ce type, les informations analogiques transitent ensuite par un amplificateur de colonne, un dispositif de suppression du bruit, puis un convertisseur analogique-numérique.

### IV-3-2 Mode courant

Le fonctionnement en mode courant exploite directement le photocourant issu de la photodiode. Ce mode est moins utilisé que le mode d'intégration du photocourant et trouve sa place lorsque la lecture en continu de l'information est nécessaire. Il nécessite un convoyeur de courant ou un miroir de courant, qui ont pour fonction de transmettre le photocourant issu de la photodiode. Compte tenu des faibles courants mis en jeu, ces structures ont une grande impédance d'entrée et fournissent une information électrique (tension ou courant) qui peut être traitée. Les miroirs de courant peuvent être des miroirs simple, cascode, ou cascode grande excursion, parfois asymétriques pour amplifier le signal. Le montage de base d'un miroir simple composé de transistors P se trouve figure I.14.

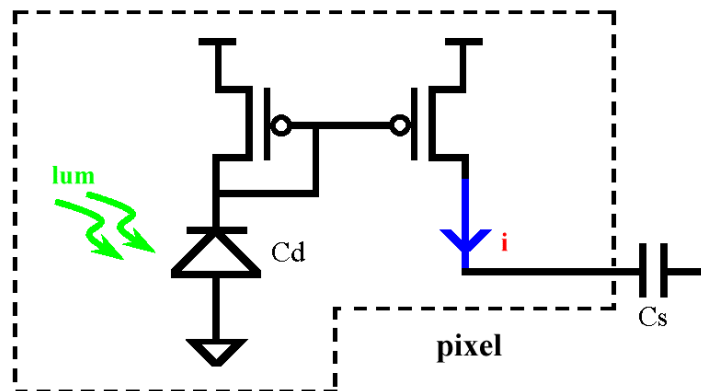


figure I.14 : Montage de base pour un fonctionnement en mode courant

Les convoyeurs peuvent avoir des structures plus complexes, comme par exemple un étage d'entrée boot-strap, dont une étude est en cours au laboratoire.

Ce mode demande beaucoup d'attention lors de la conception à cause du faible courant débité. En effet, les paramètres technologiques entraînent des photocourants de l'ordre de quelques dizaines de picoampères sur un système composé d'un objectif et d'une scène à éclairage ambiant. Le photocourant est une image linéaire de la puissance lumineuse. De ce fait, la plage dynamique de ce courant est très grande et la conception d'un système pouvant traiter des ordres de grandeurs comprenant plusieurs décades s'avère difficile. Aussi, nous ne pouvons pas mesurer des valeurs si faibles sans dégradation, des structures d'amplification sont donc nécessaires. A cause du taux de remplissage et du bruit, il n'est pas évident d'amplifier ces courants de manière très significative. Enfin, le bruit spatial fixe est compliqué à supprimer en mode courant.

Le principal avantage de ce mode de fonctionnement est la consommation. En effet, la tension d'alimentation n'intervient pas, elle peut donc être diminuée [Cou-00]. Aussi, les courants mis en jeu étant très petits, les charges et décharges des différents nœuds n'entraîneront pas de pics de puissance importants. Bien que le mode courant permette une lecture directe de l'information, il faut préciser que les faibles courants associés aux capacités à charger ou à décharger forment des systèmes qui ont un temps de réponse non nul. Les cadences vidéo élevées se heurtent à ce problème de temps de lecture.

### **IV-3-3 Mode logarithmique**

Un capteur fonctionnant en mode logarithmique est intéressant car, par la nature de sa réponse, il réalise des mesures dans une plage d'éclairement importante. Dans sa thèse sur la caractérisation d'une cellule neuronale, P. Perron donne la dynamique des conditions lumineuses naturelles variant de  $10^{-1}$  lux à  $10^5$  lux [Per-94]. L'œil humain est ainsi sensible à une dynamique d'environ  $10^5$  lux, soit 100 dB.

Le mode logarithmique présente cependant des inconvénients. Du fait de son comportement logarithmique, les informations en sortie ont une dynamique très petite dans la plupart des scènes observées, ce qui cause des problèmes de quantification et de rapport signal sur bruit. En effet, ce comportement est utile dans les scènes très fortement contrastées, mais le signal a une dynamique faible dans les scènes à contraste moyen. Au niveau du pixel, le transistor de Reset de la structure à intégration du photocourant est connecté pour fonctionner comme une diode. Les deux autres transistors ont les mêmes rôles (un suiveur et un sélecteur). Le principal défaut du mode logarithmique provient du transistor fonctionnant en diode. Comme le transistor est en régime d'inversion faible (mode photopile), ses variations de la tension de seuil (dus au procédé de fabrication) causent un grand bruit spatial, qu'il convient de corriger. Généralement, ces compensations se font a posteriori, en dehors du capteur. Cependant, [Ni-01] propose une structure de compensation du bruit spatial dans le circuit. Un exemple de schéma d'un pixel logarithmique est détaillé figure I.15.

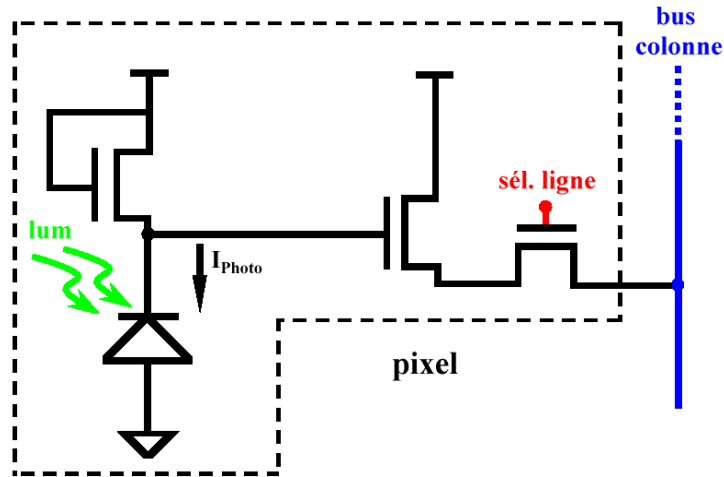


figure I.15 : Montage d'un pixel en mode logarithmique

Le transistor dont la grille est reliée au drain (connecté en diode) a les tensions  $V_{gs}$  et  $V_{ds}$  égales. Le schéma équivalent de la structure est cette diode en série avec le générateur de courant constitué par la photodiode. En considérant  $V_{tr}$  la tension à ses bornes, le courant traversant ce transistor monté en diode est alors de la forme :

$$I = I_s \left( e^{\frac{q(V_{tr})}{nkT}} - 1 \right) \quad (1.7.)$$

La tension aux bornes de ce transistor est donc proportionnelle de manière logarithmique avec le courant qui le traverse :

$$V_{tr} = \frac{nkT}{q} \cdot \ln \frac{I}{I_s} \quad (1.8.)$$

Enfin, la tension de sortie aux bornes de la photodiode est :

$$V_{ph} = V_{dd} - \frac{nkT}{q} \cdot \ln \frac{I}{I_s} \quad (1.9.)$$

## V. Défauts dans les capteurs CMOS actifs à intégration du photocourant et leur correction

### V-1 Défauts dans les capteurs CMOS actifs

Les défauts dans les capteurs CMOS sont issus de problèmes technologiques rencontrés pendant la fabrication du circuit. Ils se caractérisent typiquement par une dispersion du comportement des pixels. Les défauts sont le courant de noir et le bruit. Le courant de noir est une conséquence d'un défaut de la photodiode. H. Tian a étudié le bruit dans les capteurs CMOS pendant sa thèse [Tia-00] [Tia-01b]; il en ressort que le bruit est constitué d'un courant de noir, d'une composante spatiale (dominante) et d'une composante temporelle.

#### V-1-1 Le courant de noir

Le courant de noir est dû aux fuites dans la photodiode, qui se décharge même dans l'obscurité. Cette décharge se fait par le substrat à travers la diode diffusion-substrat. Ce courant est modélisé par un générateur de courant supplémentaire qui débite un courant constant. Son effet est de limiter la sensibilité de la photodiode. La figure I.16 montre ce générateur de courant constant ( $I_{fuite}$ ) en parallèle avec le schéma équivalent vu plus haut.

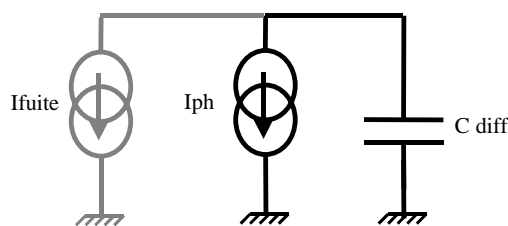


figure I.16 : Schéma équivalent de la photodiode avec son courant de fuite

Considérons un dispositif avec un éclairage (fort) de 500 lux sur des photodiodes de  $900\mu\text{m}^2$  en technologie  $0,6\mu\text{m}$ . L'ordre de grandeur du photocourant ( $I_{ph}$ ) est de  $50\text{nA}$ . Le courant de noir mesuré avec le même capteur est de  $8\text{ fA}$ . Dans ces conditions, le courant de noir crée une erreur de  $0,016\%$ . En considérant que l'application utilise 256 niveaux de gris, le pas de quantification est de  $195\text{ fA}$ . Les fuites représentent alors une erreur de  $1/24\text{ LSB}$ . Cette erreur est relativement faible. Le courant de fuite cause surtout des problèmes dans les applications

faible flux (comme la visée stellaire).

### V-1-2 Le bruit spatial fixe (BSF ou FPN)

Le bruit spatial fixe (BSF ou FPN : Fixed Pattern Noise), provient des dispersions du processus rencontrées lors de la fabrication des composants. Des éléments identiques, comme les pixels d'un capteur, ont finalement des caractéristiques électriques sensiblement différentes. Ce sont en fait les non uniformités durant les étapes de fabrication qui sont à l'origine du bruit spatial. La valeur du bruit n'est pas prévisible, elle peut être atténuée par des méthodes empiriques lors du choix des dimensions des transistors. Ce bruit est surtout présent dans les transistors suiveurs des pixels et particulièrement dans les amplificateurs de colonne, ces éléments sont donc des points critiques auxquels il convient de prêter attention lors de la conception. L'équipe de G. Meynants a déterminé l'amplitude de ce bruit dans une structure d'APS à intégration du photocourant [Mey-01]. Il représente une variation de 0,2 % sur toute la matrice. En considérant une dynamique de 3v, il représente un écart type de 6mV crête à crête.

Comme le BSF est fixe par rapport au temps, des structures de réduction du bruit existent. Elles ont pour principe d'effectuer une lecture différentielle pour s'affranchir de l'offset qui cause l'erreur. Les techniques de réduction du BSF sont exposées plus loin.

### V-1-3 Le bruit temporel

Ce bruit est composé de différents types de bruits, principalement le bruit de grenaille, le bruit en 1/f et le bruit thermique. Les bruits dominants sont les bruits de grenaille et thermique. On note cependant que les bruits temporels sont négligeables par rapport au bruit spatial. Des équipes ont étudié les différents bruits, comme le montre la publication IEEE [Deg-00]. En s'appuyant sur cette référence et [Joh-96], nous allons indiquer les différents bruits temporels présents dans les capteurs CMOS de type APS 3 transistors : les bruits de grenaille, en 1/f, thermique, et l'offset.

- **Le bruit de grenaille** (ou "shot noise") résulte de la fluctuation du nombre de porteurs sous l'action d'un champ électrique. C'est une fluctuation de type poissonienne. Ce bruit est donc directement lié au courant électrique auquel il se superpose. Il est surtout présent dans les diodes, et a pour densité spectrale :

$$X^2(f) = 2 q I \quad (1.10.)$$

Unité :  $A^2 / Hz$

q : charge élémentaire ( $1,6 \cdot 10^{-19} C$ )

I : courant moyen

- **Le bruit en  $1/f$**  (appelé aussi bruit rose, de scintillement, ou "flicker noise") est un bruit dont la densité spectrale est inversement proportionnelle à la fréquence. Il est lié aux défauts du semiconducteur (fluctuations des taux de recombinaison des paires électrons trous, piégeage des porteurs dans les défauts de l'oxyde du transistor), et dépend donc de paramètres technologiques. Pour un transistor de dimensions W,L, il est donné par la relation :

$$X^2(f) = \frac{K}{W.L.C_{ox}.f} \quad (1.11.)$$

Unité :  $V^2 / Hz$

K : coefficient empirique lié à la technologie

W, L : dimensions du transistor

$C_{ox}$  : Capacité de l'oxyde mince

f : fréquence

- **Le bruit thermique** vient de l'agitation thermique des électrons dans un matériau, qui génère une tension observable, même en l'absence de champ électrique. Il est lié au potentiel thermique  $U_t$  donné par la technologie ( $U_t = \frac{kT}{q}$ ). Le bruit thermique est présent dans les résistances, la tension d'entrée équivalente du bruit vaut [Joh-96] :

$$V^2(f) = 4. KTR \quad \text{unité : } V^2 / Hz \quad (1.12.)$$

Et le courant équivalent:

$$I^2(f) = \frac{4.kT}{R} \quad \text{unité : } A^2 / Hz \quad (1.13.)$$

k : constante de Boltzman ( $1,38 \cdot 10^{-23} J/K$ )

T : Température (K)

R : Valeur de la résistance

h : constante de Plank ( $6,6 \cdot 10^{-34} J.s$ )



Dans une structure de type RC, la capacité ne génère pas de bruit, mais peut stocker le bruit thermique généré par la résistance. En considérant un système du premier ordre résistance bruitée et capacité, la bande passante de bruit de ce filtre passe-bas de fréquence de coupure

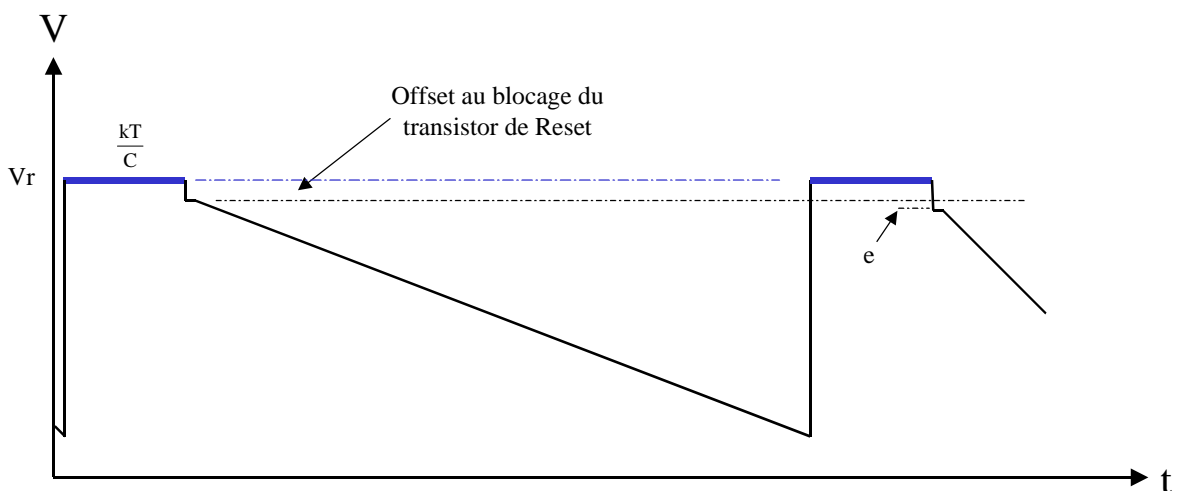
$$f_0 = \frac{1}{2\delta \cdot RC} \text{ vaut } \frac{\delta \cdot f_c}{2} \text{ [Joh-96]}. \text{ Le bruit de la résistance est un bruit blanc, le bruit en sortie}$$

$$\text{de ce système vaut donc: } V^2(f) = \frac{\delta}{2} \frac{1}{2\delta \cdot RC} \cdot 4 \cdot KTR = \frac{KT}{C} \quad (1.14.)$$

- **Le bruit de Reset** s'explique par deux phénomènes :

L'utilisation d'un transistor de type N comme transistor de Reset – dans les pixels actifs APS classiques 3 transistors - entraîne un offset négatif sur la tension mesurée aux bornes de la photodiode. Cet offset apparaît lors de la commutation du transistor de Reset (de '1' à '0'), il est dû aux injections de charges provenant du canal de ce transistor.

Lorsque le transistor de Reset est fermé, le système RC (transistor et capacité de la photodiode) engendre un bruit  $\frac{KT}{C}$ . Le transistor devient en effet sensible aux variations dues au bruit thermique à la fin de la précharge, car il est en mode de conduction sous le seuil. Ce bruit (aléatoire) entraîne une erreur lors de l'échantillonnage de la tension de précharge  $V_r$ . La figure I.17 illustre ces phénomènes d'offset et de bruit.



**figure I.17 : Défaits du mode en intégration du photocourant**

Ici, e représente l'écart de tension dû au bruit de Reset. A cause des bruits spatiaux et temporels, des différences de réponse existent entre les différents pixels, spatialement, mais aussi pour un même pixel entre deux acquisitions, temporellement.

Ce bruit se caractérise par des imperfections sur l'image. La figure I.18 illustre une image d'origine et sa dégradation par un bruit de type pixel et de type colonne.



**figure I.18 Image d'origine et avec du bruit : de type pixel et de type colonne**

Le bruit de type pixel provient des dispersions inter-pixels (photodiode, transistors de Reset et suiveur). Le bruit de type colonne provient des dispersions des amplificateurs de colonne, qui ont des gains différents. Le bruit de type colonne est (légèrement) plus important que le bruit de type pixel dans les capteurs de type APS 3 transistors : il engendre une dispersion de 2%, soit 5 niveaux de gris sur 256 [Goy-92]. Notons aussi que le bruit colonne, à cause de sa structure, est très observable, alors que le bruit de type pixel est peu détectable par l'œil.

Dans des imageurs où la conversion analogique / numérique se fait dans le pixel, le bruit est uniquement de type pixel. Ces imageurs sont appelés DPS : "Digital Pixel Sensor". [Kle-01] propose un DPS avec conversion analogique / numérique 8 bits dans le pixel, qui a un bruit spatial relativement faible (0,8 % du signal).

Pour atténuer ces bruits dans les imageurs standards, il existe des structures électroniques de réduction du bruit. Elles consistent en un procédé de double mesure (double échantillonnage), qui permet de ne s'intéresser qu'à la décharge effective de la phototension. Les conditions initiales (tension  $V_r$  qui fluctue) n'influent donc pas sur le résultat. La réduction du bruit peut être effectuée de deux manières : par CDS ("Correlated Double sampling" : Double Echantillonnage Corrélé) ou par NCDS ("Non Correlated Double Sampling" : Double Echantillonnage Non Corrélé).

## V-2 Double échantillonnage corrélé (CDS)

Le double échantillonnage corrélé, ou "Correlated Double Sampling", est une mesure de deux tensions sur une même exposition du pixel [Gam-98] [Fuj-00]. Les deux échantillons sont corrélés, puisque le bruit et l'offset perturbent de manière identique la décharge. La figure I.19 indique ces deux niveaux de tension, A et B. A est la tension effective aux bornes de la photodiode à la fin du Reset, qui a pour valeur nominale  $V_r$  et comme valeur réelle  $V_r - \text{offset}$ .

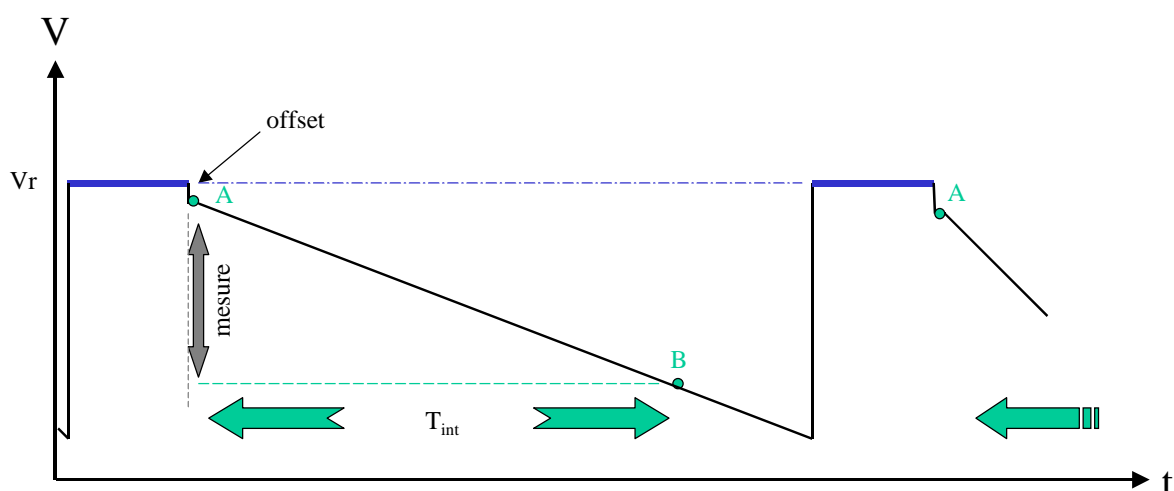


figure I.19 : Mesures utilisant le double échantillonnage corrélé (CDS)

De plus, compte tenu des instants de mesure, le CDS élimine le bruit en  $kT/C$  présent lors de la précharge de la photodiode. Après le temps d'exposition (ou d'intégration)  $T_{int}$ , la deuxième tension, B, est relevée ; la mesure de la décharge est alors définie comme  $V_{dech} = A - B$ .

Cette mesure est la plus précise, mais elle pose des problèmes de mise en œuvre. En effet, il faut mémoriser une information analogique sur un temps assez long (plusieurs millisecondes), ce qui ne peut pas se faire sans perte. Certains travaux font état de mémorisation dans le pixel, comme l'équipe de A. Simoni [Moi-97a] avec des capacités commutées, mais avec des pertes lors des transferts de charges, ou comme [Ni-02], mais avec des temps de mémorisation brefs. La solution serait donc d'utiliser un obturateur mécanique ou de mémoriser toute une trame ; mais cela entraîne un coût bien plus élevé. En pratique, le CDS est donc utilisé principalement dans les pixels actifs à photogrille.

### V-3 Double échantillonnage non corrélé (NCDS, ou DS)

Pour pallier les problèmes rencontrés pour mettre en œuvre le CDS, une alternative existe : le double échantillonnage non corrélé ("Non Correlated Double Sampling" ou "Double Sampling"). Ce dernier est utilisé pour les pixels actifs à photodiode trois transistors classiques (figure I.12). Cette technique a pour principe la mesure de la phototension sur deux acquisitions. Concrètement, la mesure au bout du temps d'intégration se fait en B, et celle de la valeur effective du Reset en A (figure I.20). L'erreur causée par un offset non constant (dû au bruit KTC) perturbe donc la mesure ; c'est pour cela que le NCDS est moins précis.

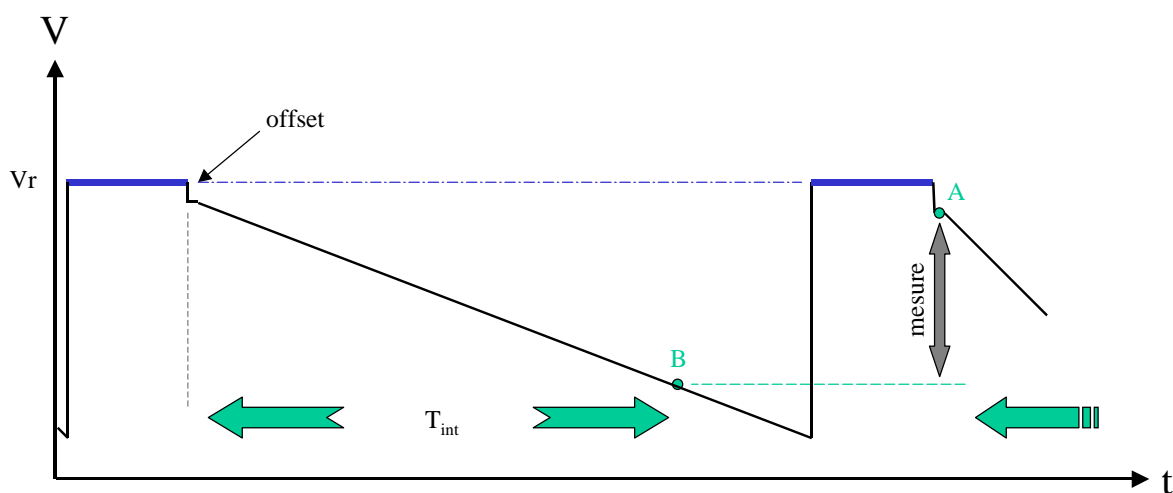


figure I.20 : Mesures donnant le double échantillonnage non corrélé (NCDS)

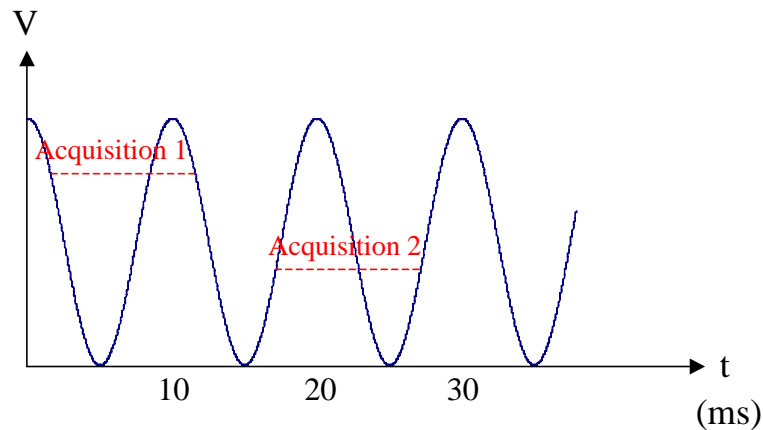
Le double échantillonnage non corrélé permet une mise en œuvre bien plus aisée, surtout compte tenu du temps de mémorisation plus court (par rapport au CDS).

La société Sarnoff [Sar-03] commercialise par exemple un capteur d'image CMOS à photodiode qui, après réduction du bruit, n'entraîne que 0,16 % (données constructeur) d'erreur sur les images (soit 0,4 niveau de gris sur une dynamique de 256 niveaux).

## VI. Perturbations lumineuses

Les conditions lumineuses peuvent être elles-mêmes à la source de perturbations du signal utile. En effet, les scènes réelles offrent des quantités de discontinuités optiques, comme des réflexions non uniformes de certains objets, des contrastes très forts posant des problèmes de saturation, ou des éclaircissements discontinus.

représentent le phénomène optique le plus courant. Ils se caractérisent par un changement dans le champ d'éclairage (nuage, personne, objet), ou par une variation de la source lumineuse elle-même (lever / coucher de soleil, éruption solaire, éclairage artificiel). En prenant comme exemple l'éclairage artificiel présent dans une pièce, celui-ci est une sinusoïde de fréquence 100 Hz, puisque l'intensité lumineuse est proportionnelle au carré du courant traversant le néon.



**figure I.21 : Acquisitions sous un éclairage artificiel**

Un "bruit" à 100 Hz peut donc perturber la mesure, la figure I.21 montre l'allure du signal lumineux et deux acquisitions d'images (temps d'intégration). Pour éviter ces perturbations, il faut que le temps d'intégration soit un multiple de la période de l'éclairage (de 10 ms), comme les acquisitions 1 et 2 de la figure. L'intégration se fait donc sur un nombre entier de périodes, la quantité de lumière reçue est alors la même. Ce paramètre est à prendre en compte quand on met en œuvre un capteur d'images.

Aussi, les capteurs fonctionnant en intégration en mode ligne ont forcément un bruit lumineux. En effet, comme les imageurs standards CMOS ne mémorisent pas l'image, ils acquièrent l'image ligne par ligne. Ce fonctionnement entraîne que l'image et/ou les conditions lumineuses peuvent changer entre deux lignes acquises. Classiquement, un effet saccadé peu apparaître lors de mouvements de la caméra.

## VII. Rétines CMOS

Comme le résume Thierry Bernard dans son article [Ber-97b], "Analyser l'image là où elle est acquise pour n'en retenir et en transmettre qu'un extrait pertinent pour la tâche de vision en cours, tel est le principe des rétines artificielles".

Une rétine CMOS est donc un capteur "intelligent", dans lequel phototransduction et traitement de l'image cohabitent. Les rétines sont des composants qui combinent les propriétés photoélectriques de la technologie CMOS avec ses caractéristiques électroniques. Nous avons vu précédemment que la coexistence de ces caractéristiques aboutit à la création de matrices adressables, permettant d'accéder aléatoirement ("random access") aux informations de l'image brute. De tels circuits permettent aussi de traiter directement (et au plus tôt) l'information lumineuse, en modifiant le signal utile qui sera transmis. Ce pré-traitement de l'image a pour but d'alléger le calcul à effectuer par la suite. Le système de vision, classiquement composé d'un capteur et de périphériques de traitement des données s'intègre ainsi dans un unique composant, afin de rendre les calculs plus rapides, et de s'affranchir des problèmes de bande passante et de consommation d'un système sur carte.

Le concept de rétine silicium est apparu dans les années 1980. [Gar-88] propose en effet une rétine silicium basée sur des calculs entre des images binaires. Carver Mead propose aussi un modèle de rétine silicium effectuant un filtrage de l'image observée [Mea-89]. Cette réalisation est inspirée des découvertes alors récentes en neurobiologie, et se basent sur le système de vision de primates. Le nom associé à cette catégorie de composants provient de ces premiers types de circuits bio-inspirés. De manière générale, tout circuit de vision générant une information (traitée) autre que l'image brute acquise est appelé rétine.

### VII-1 Différents types de rétines

Il existe de très nombreuses applications d'imagerie reposant sur l'utilisation de rétines CMOS. Nous ne donnerons pas une liste exhaustive de ces travaux, mais une présentation générale en s'appuyant sur quelques exemples significatifs. Ces diverses applications peuvent se classer en deux catégories que nous allons détailler: les applications à traitement spatial et les applications à traitement spatio-temporel.

## VII-1-1 Les rétines à traitement spatial

Les applications spatiales concernent les traitements de type focal se limitant aux informations issues d'une unique acquisition.

Nous pouvons les regrouper de cette manière :

- **Adaptation aux conditions lumineuses.**

Les rétines adaptatives ont comme particularité d'adapter la réponse aux conditions lumineuses. Cette adaptation peut se faire de deux manières : en faisant varier le temps d'exposition (adaptation temporelle), ou en ajustant la valeur des pixels (traitement spatial). Nous ne considérons pas ici les travaux afférents aux imageurs APS dont le temps d'exposition est automatiquement réglé en fonction de l'illumination de la scène, par exemple [Gou-95], [Ham-01]. [Loo-01] propose aussi un capteur logarithmique dont la compensation analogique du signal est configurable.

Les rétines adaptatives ont une conception différente, qui est inspirée de la biologie, et effectuent un traitement spatial (inter-pixels). Les recherches en neurologie ont permis de modéliser le fonctionnement de la rétine biologique. Les rétines se basent sur ces modèles, introduisant une interaction forte des pixels, permettant une réponse adaptée à une illumination moyenne. La première réalisation de ce type [Mea-89] repose sur l'utilisation de réseaux effectuant un lissage à travers un maillage résistif inter-pixels.

- **Caractéristiques et calculs**

L'extraction de caractéristiques dans l'image, comme des contours ou des coins (rencontre de deux contours de directions distinctes) ont été implantés dans des rétines CMOS. Les modèles basés sur la détection de contours utilisent la propriété lumineuse de ces endroits caractéristiques (le gradient est fort). Les rétines effectuant ce pré-traitement sont en quasi-totalité basées sur l'utilisation de filtres gaussiens, comme [Ni-93] qui utilise des filtres de type capacitifs. Plusieurs approches existent, comme l'implantation de filtres analogiques basés sur des maillages de type bio-inspirés [Mea-89] (effectuant un lissage), ou des approches reposant sur des opérations inter-pixels (multiplications, soustractions, ...) analogiques ou numériques [Fun-97] [Bar-02] [Tab-01].

Les travaux présentés dans [Ind-99] proposent un circuit neuromorphique détectant les gradients forts dans une image (et leur suivi). Beaucoup de ces implantations sont basées sur une détection de contours en vue de leur suivi, et donc se classent dans la catégorie des rétines

à traitement spatio-temporel.

La position d'un objet peut être extraite d'une scène, les coordonnées donnent alors une information qui peut être traitée par la suite [Nez-00]. La scène observée est cependant relativement simple, et se limite à un objet de fort contraste sur un fond uniforme.

Des calculs spatiaux sur l'image sont aussi possibles, comme des moyennes, des rotations, des filtres de déformation, ou tout autre calcul spatial. Ceux-ci peuvent être mis en œuvre par des processeurs de type SIMD ("Single Instruction Multiple Data"), de granularité fine, traitant l'image au niveau du pixel de manière parallèle [Ber-97a] [Gen-99]. L'inconvénient de ces capteurs réside dans leur pilotage complexe, dû au séquençement des instructions, nécessitant des périphériques externes. Une autre rétine programmable [Eti-99b] effectue un calcul analogique dont la fonction et le voisinage considéré sont paramétrables. Les applications sont toutefois limitées à un filtrage ou à une convolution dans un voisinage variant de 3x3 à 11x11 pixels.

- **Modulation spatiale de la résolution**

Des rétines proposent une structure de capteur avec une résolution non classique. Principalement deux types de rétines proposent des approches originales : les rétines à résolution programmable et les rétines à résolution fovéale.

Les capteurs à résolution programmable reposent sur un procédé de sous-échantillonnage consistant à regrouper plusieurs pixels pour qu'ils ne donnent qu'une unique information. La fréquence spatiale, donc la résolution, s'en trouve changée [Kem-96] [Cou-00]. L'intérêt de la résolution programmable est que la quantité d'information peut être réduite, en s'intéressant particulièrement à une zone d'intérêt où la résolution sera maximale.

Les rétines dont la résolution est fovéale s'inspirent de la biologie. Le système de vision biologique est caractérisé par une résolution élevée au centre de l'image et faible sur le pourtour. Les rétines CMOS de ce type reprennent un agencement log-polaire, avec une grande résolution au centre, et diminuant logarithmiquement autour [San-00]. Ce type de résolution permet le suivi d'objets, en plaçant les éléments importants de la scène au centre (résolution élevée). Les détails autour de ce centre nécessitent alors moins de données à gérer. L'inconvénient de cette structure est que la résolution log-polaire est fixe. Aussi, cet agencement en cercles engendre une perte de place au niveau des pixels, donc une sensibilité plus faible.



## VII-1-2 Les rétines à traitement spatio-temporel

Les applications spatio-temporelles sont généralement plus complexes, car un paramètre supplémentaire influe sur l'image : le temps. Les calculs se font en fait sur plusieurs images (séquences), introduisant des problèmes dus aux discontinuités. Les traitements de telles rétines s'appliquent sur des systèmes multidimensionnels (spatial - X et Y- et temporel), ils sont donc complexes. Nous verrons que face à cette complexité, la quasi-totalité des rétines ne s'appuient pas sur une intégration numérique d'algorithmes. Cette solution se heurte en effet à une masse de calcul itératifs importante difficilement réalisable en des temps satisfaisants, souvent à cause des nombreux accès mémoires externes requis.

La difficulté première est due au stockage d'image(s) que nécessitent ces calculs. Ce paramètre a tendance à rendre l'intégration consommatrice en surface et compliquée à réaliser. De ce fait, la surface des pixels constituant ces rétines est toujours bien plus importante que celle de pixels d'un imageur classique. Les principales applications visées sont la reconnaissance et le suivi d'éléments ("tracking"), le calcul inter-images (différence, corrélation), ou la mesure de mouvement.

Un suivi monodimensionnel (dans un axe spatial) d'éléments de l'image a été implanté par [Ind-99]. Il s'agit de suivre le contraste le plus fort, en utilisant une structure du type "Winner Take All". Un fort et unique contraste peut donc être suivi sur un axe de translation, mais cette réalisation ne permet pas de traiter des séquences complexes contenant plusieurs informations importantes. Un autre champ d'application, comme [Ber-00], est le suivi de particules générant un stimulus très contrasté similaire.

Un calcul de corrélation, implanté dans une rétine analogique, a été présenté dans la revue OE-Magazine [Voo-01b]. L'architecture du pixel est basée sur l'aiguillage des photocourants de pixels classés dans une catégorie. Cette réalisation a été étudiée pendant la thèse, elle est détaillée dans le chapitre II.

Des applications de différence entre deux images successives ont été développées [Ma-99] [Ni-02]. Ces techniques peuvent par exemple alléger une compression ultérieure basée sur la différence inter-images, la reconnaissance de formes, ou constituer la base d'un calcul de stéréovision ou d'estimation du mouvement.

La stéréovision et l'estimation de mouvement sont deux applications spatio-temporelles très intéressantes et utiles dans le sens où elles donnent une information de haut niveau qui

permet de reconstruire un environnement 3D ou de prendre des décisions par rapport à l'évolution de l'environnement. La stéréovision consiste en l'extraction des informations combinées de deux capteurs distincts. Ces deux capteurs ayant un angle de vue différent, transmettent au système des informations susceptibles de calculer une information de profondeur [Mey-98b]. Ce type d'application est complexe à réaliser à cause des calculs nécessaires et de sa mise en œuvre. L'alignement des capteurs doit en effet être précis et un calibrage du système de vision s'impose. Il est cependant très utilisé en robotique compte tenu de l'importance de l'information qu'il donne.

Nous détaillons les études portant sur la conception de rétines dédiées à l'estimation de mouvement dans la prochaine partie, puisque c'est dans cette catégorie de rétines que s'insèrent nos travaux.

## **VII-2 Rétines dédiées à l'estimation de mouvement**

Les rétines dédiées à l'estimation du mouvement donnent l'information du déplacement qui peut être représenté par un champ de vecteurs. Bien que les calculs soient tous basés sur les caractéristiques spatio-temporelles des images, à savoir l'évolution des illuminations des pixels au cours du temps, ces circuits peuvent se regrouper en trois catégories selon qu'ils sont basés sur des filtres, sur les contrastes, ou qu'ils s'appuient sur des calculs numériques.

### **VII-2-1 Rétines basées sur les filtres spatio-temporels**

Ces circuits constituent les rétines bio-inspirées, axées autour d'une structure commune : l'utilisation d'interactions au niveau des pixels. Ils utilisent le gradient des pixels à travers des réseaux résistifs [Sto-98], ou des réseaux de filtres [Moi-97a] [Hig-00] [Sic-99]. Ces implantations intègrent généralement soit un algorithme (existant), soit la résolution d'équations différentielles sous forme de calculs analogiques.

La structure générale de ces rétines est montré dans la figure I.22. Un réseau interconnecte les pixels à travers les transconductances  $G$ , modélisant des résistances et/ou des capacités.

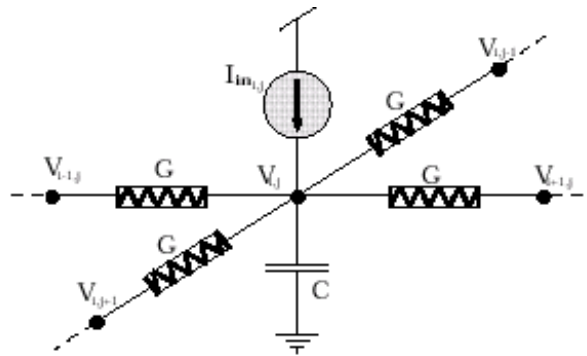


figure I.22 : Exemple de réseau d'interconnexion des pixels

Par exemple, dans la thèse de Gilles Sicard [Sic-99] ("De la biologie au silicium: une rétine bio-inspirée analogique pour un capteur de vision intelligent adaptatif"), les transconductances sont composées de cellules RC, dans lesquelles la valeur des résistances est ajustée pour changer la fréquence de coupure des filtres. Il est ainsi possible de rehausser les contours de l'image. Cependant, l'utilisation de capacités et le grand nombre de transistors nécessaires donne aux pixels un taux de remplissage faible (7 %).

Cependant, les informations issues de ces rétines bio-inspirées sont des réponses à des filtres, donc analogiques, qu'il faut pouvoir stocker et/ou traiter de manière parallèle. La conception demande donc des blocs supplémentaires (mémoires analogiques, convertisseurs) permettant une interface plus aisée. Pour ces raisons de mise en œuvre, les circuits de ce type rencontrés dans la littérature ont toujours des résolutions faibles, par exemple 7x7 pixels [Sto-98]. Certains de ces circuits bio-inspirés donnent une information de mouvement globale plus simple à mesurer [Hig-00]. De manière plus générale, les rétines basées sur les réseaux de filtres sont complexes à réaliser dès qu'on prend en considération l'aspect parallèle analogique des sorties. Pratiquement, il n'existe pas de circuit de ce type traitant des images fortement texturées à une résolution suffisante, où des objets en mouvement constitueraient les centres d'intérêts.

## VII-2-2 Rétines basées sur le contraste

Une autre catégorie de rétines estimant le mouvement utilise des éléments caractéristiques de l'image : les contrastes, caractérisés par un fort gradient. Deux approches existent : le suivi des fronts de contraste ou le suivi d'impulsions générées par des gradients forts.

Le principe du suivi des fronts de contraste est le suivant (figure I.23). En considérant un front de contraste se déplaçant horizontalement de gauche à droite (ce qui est limitatif), celui-

ci génère le même stimulus sur les structures identiques A, B et C, et peut donc être reconnu.

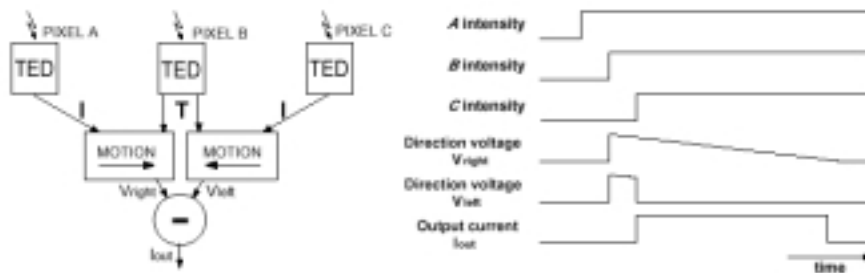


figure I.23 : Exemple de suivi d'un front de contraste [Hig-98]

Ce stimulus est alors cherché dans une direction qui est celle du mouvement. Ce type de mesure est intuitif et facilement compréhensible en considérant une scène "minimaliste" composée d'un objet se déplaçant sur un fond uniforme, comme dans cet exemple. En effet, dans la figure I.23, les intensités des pixels A et B restent constantes avant et après le passage de ce contraste, ce qui représente le suivi le plus simple à réaliser. Cependant, cette mesure perd son efficacité et ne trouve plus sa place lorsque l'information scénique est très riche. Une forte texture donne en effet de nombreux déclenchements de ces structures, entraînant une grande difficulté d'interprétation. Aussi, l'hypothèse d'illumination constante est faite, à ceci vient s'ajouter le fait qu'une telle structure demande une surface conséquente pour effectuer une estimation en deux dimensions. [Hig-98] propose en effet une telle architecture en prenant en compte les mouvements horizontaux et verticaux, mais pas les mouvements diagonaux.

L'autre approche basée sur le contraste consiste à extraire des impulsions à partir des variations du gradient de l'image. La plupart des réalisations utilisent le photorécepteur adaptatif développé par Tobi Delbrück [Kra-97] [Deu-98b]. Le schéma de principe est illustré figure I.24.

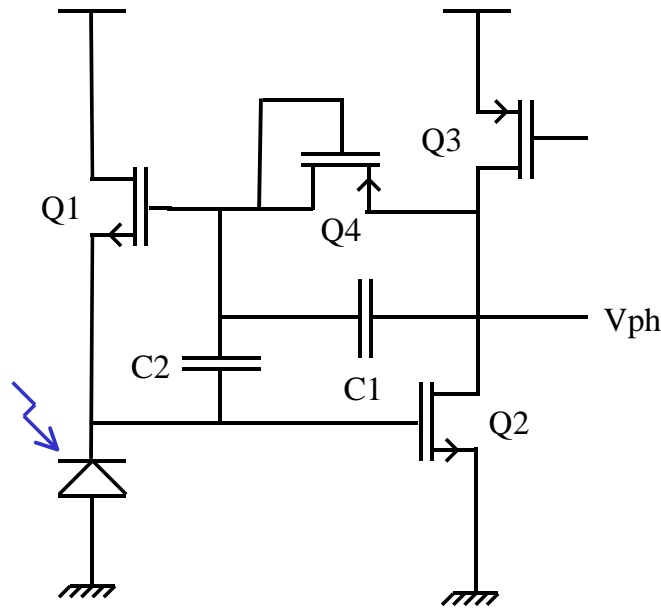


figure I.24 : Schéma du récepteur adaptatif de Delbrück [Kra-97]

Un bloc de contre réaction, composé des éléments C1, C2, Q3 et Q4, donne un gain adaptatif à cette structure. Ainsi, les signaux statiques sont peu amplifiés, seuls les signaux à forte variation le sont. Cette structure a une propriété de filtrage des stimuli qui donne une réponse aux variations de l'illumination, ce qui permet d'extraire les hautes fréquences, caractéristiques de l'image. Cette tension issue du photorécepteur adaptatif est ensuite mise sous forme d'impulsion, l'image texturée est donc convertie en contours (hautes fréquences sous forme d'impulsions) dont le suivi est plus simple. La structure fonctionne en temps continu, la tension de sortie  $V_{ph}$  donne donc un signal à tout instant, couvrant une grande dynamique (6 décades).

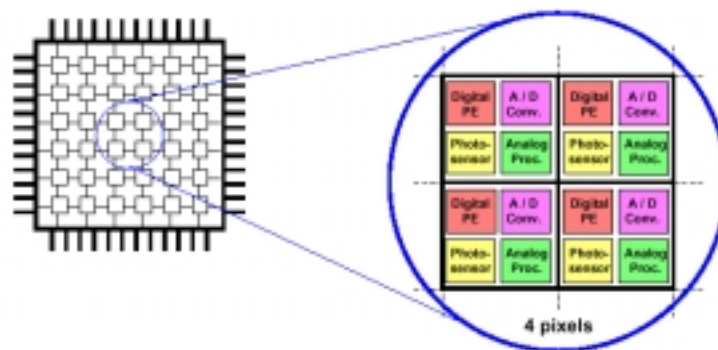
Cette structure a cependant quelques inconvénients : la nature adaptative des informations crée l'effet d'une égalisation des gradients qui peut causer des erreurs lors de la reconnaissance, et la surface occupée par les deux capacités n'est pas négligeable. Aussi, le nombre d'éléments de la structure (transistors et capacités) en fait une solution coûteuse en surface. Enfin, les détecteurs sont orientés, ce qui ne leur confère qu'une sensibilité deux dimensions (orthogonale).

[Jia-99] montre une autre méthode basée sur l'utilisation d'impulsions. Ces impulsions sont générées à l'aide d'un circuit de mise en forme connecté aux photodétecteurs : des transistors BJT parasites. Ces impulsions, de même amplitude et de même durée, sont décalées temporellement. Elles sont ensuite corrélées à travers une structure logique orientant les impulsions sur une des quatre sorties (gauche, droite, haut, bas). Le train d'impulsions sur une de ces sorties indique le sens et la vitesse du mouvement. La faiblesse de cette méthode

vient de la normalisation des impulsions, ce qui empêche de reconnaître un contraste par sa valeur et donc peut entraîner des erreurs d'interprétation. De plus, les pixels présentés sont imposants par leur taille (100 $\mu$ m de côté), ce qui rend un tel circuit coûteux.

### VII-2-3 Rétines basées sur des structures de calculs numériques

Les rétines de ce type sont intéressantes car elles peuvent effectuer des calculs de différentes sortes sur les images. En effet, la plupart de ces réalisations reposent sur des machines parallèles intégrant des éléments de calcul dans les pixels. Nous nous focalisons sur les rétines, nous ne détailleront donc pas ici les circuits du type [Row-97], [Tur-00] ou [Ari-02] qui proposent un calcul de mouvement intégré dans un ASIC ou un FPGA, qu'il faut associer à un capteur et de la mémoire externe. Plusieurs de ces rétines, associant phototransduction, unité arithmétique et logique, mémoire et registres ont été réalisées [Kom-97]. L'exemple le plus intéressant du point de vue de la résolution et des résultats est la rétine PVLSAR [Ber-97a] [Pai-99]. Un processeur élémentaire est situé dans chaque photosite, et effectue des calculs booléens (AND, OR, XOR et copie) sur des images binaires (figure I.25). Cependant, l'accumulation des informations au cours du temps permet de travailler sur des niveaux de gris, en extrapolant des seuils à partir des différents temps d'exposition.



**figure I.25 : Schéma de principe de quatre photosites de PVLSAR**

Le photosite de cette rétine est composé d'un photorécepteur, d'un convertisseur analogique numérique et de fonctionnalités de calcul. Comme nous l'avons dit précédemment, le défaut de ce type de structures SIMD est qu'il est nécessaire de les piloter avec des séquences de commande extérieures, ce qui impose donc l'utilisation d'un séquenceur externe. De plus, le nombre d'instructions devient très conséquent lorsque des opérations complexes sont réalisées.







## Chapitre II :

# ETUDE ET CONCEPTION DE CAPTEURS ET DE RETINES

---

Dans ce chapitre, nous présentons le travail que nous avons réalisé afin de découvrir le domaine des capteurs d'images. Le thème de recherche relatif à mon travail, portant sur les capteurs d'images CMOS, est relativement nouveau au sein du laboratoire, nous avons donc conçu et étudié des circuits pour acquérir des connaissances sur les imageurs et les rétines CMOS. En effet, bien que les publications apportent une connaissance théorique certaine, elles développent généralement une étude dans sa globalité. Concrètement, rien ne remplace l'expérience qu'on peut acquérir en concevant un circuit et en effectuant des mesures. Cette approche pratique permet de mesurer des paramètres sur des systèmes simples, avant de concevoir des systèmes plus complexes. Aussi, les différentes publications apportent chacune des connaissances ponctuelles, difficiles à transposer sur un cas précis. Ce chapitre présente les différents circuits que nous avons étudiés et conçus, ainsi que les résultats obtenus.

## I. Rétine d'intercorrélation

### I-1 Présentation

Une collaboration, portant sur la conception et la mise en œuvre de rétines CMOS, existe depuis 1994 entre le laboratoire LE2I du Creusot et le département Microélectronique du LIRMM. Celle-ci a donné lieu à l'étude et la conception d'un dispositif d'intercorrélation entre deux images et le dépôt d'un brevet [Gor-99], puis la réalisation en technologie 1.2 $\mu$ m d'un premier prototype ayant une image de référence fixe, déterminée à la fabrication du circuit [Lam-00]. Comme les résultats étaient intéressants, ce projet a abouti à la conception d'un second prototype permettant de programmer l'image de référence. Le calcul de la corrélation entre deux images se fait ainsi : une image de référence est mémorisée, ensuite les variations entre l'image courante et l'image mémorisée se lit sur les sorties analogiques du circuit.

A mon arrivée en thèse, la cellule de base, c'est à dire le pixel, était déjà réalisé ; il

restait à créer le système de bufférisation des signaux de commande, et les différents blocs de réglage d'offset en entrée et d'amplification des signaux de sortie. Ce circuit a été conçu avec une technologie mixte (double polysilicium, trois métaux) 0.6  $\mu\text{m}$  de AMS (Austria Mikro Systems). La réalisation et la caractérisation de ce circuit ont donné lieu à des articles [Voo-01a] [Voo-01b] [Cat-02], et à un prix remporté lors de la conférence DATE'02.

## I-2 Fonctionnement du pixel

Nous avons étudié le principe de fonctionnement de ce circuit pour pouvoir achever sa conception. Ce circuit utilise directement le photocourant, ce qui permet un temps de réponse faible et une intégration simple des opérations à effectuer, à travers des différences, des moyennes et des comparaisons de courants. Le montage a aussi l'avantage d'être adaptatif à la lumière ambiante dans la mesure où il effectue des calculs reposant sur des différences et des moyennes entre une image mémorisée et une image observée. Ce circuit n'a pas besoin d'être calibré si il change de scène, il ne nécessite que la mémorisation de l'image de référence.

Chaque photosite est composé d'un photorécepteur, d'un comparateur, d'un point mémoire, et d'un système d'aiguillage du courant. La figure III.1 schématise un pixel, avec ses différents blocs.

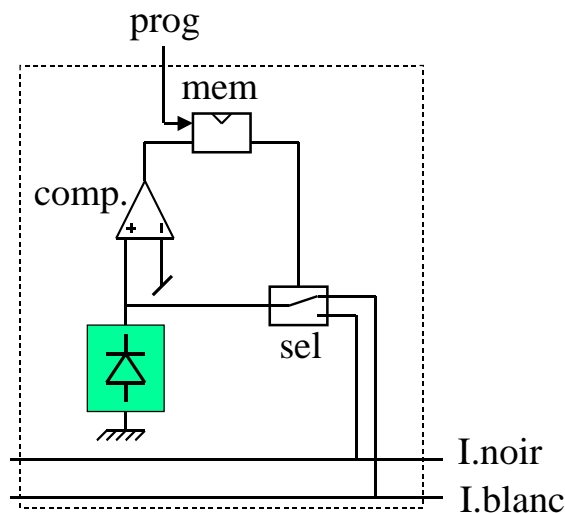


figure III.1 : Diagramme d'un pixel

La photodiode est symbolisée par le symbole de la diode, qui est connectée à un comparateur 'comp' qui binarise la différence entre la luminosité du pixel courant et la moyenne de tous les pixels de la rétine. Ce bit de configuration est stocké dans un point mémoire 'mem', qui commande le système d'aiguillage 'Sel'. Ce point mémoire est déclenché par un signal externe 'prog'. La figure III.2 montre une vue schématique du pixel détaillant les blocs de la figure III.1.

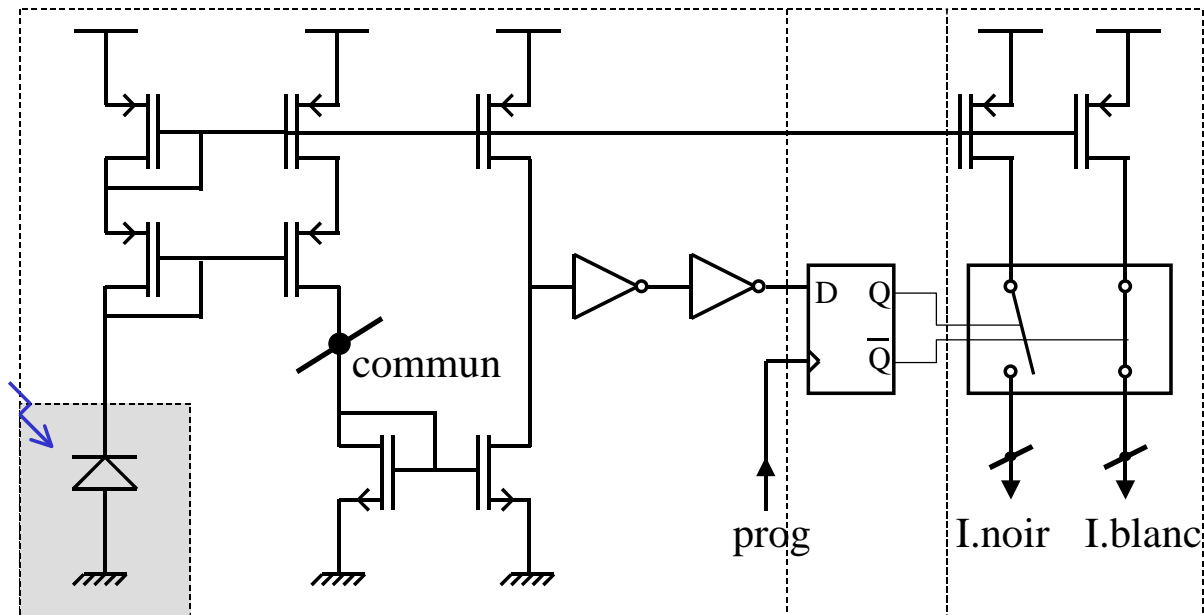


figure III.2 : Schéma d'un pixel

Le photocourant généré par la photodiode est dupliqué par un miroir de courant en quatre copies. La première copie contribue à la somme des courants de tous les pixels, grâce à un point commun. La deuxième copie sert à effectuer la différence entre le photocourant et la moyenne des photocourants : la somme de ces photocourants divisée par leur nombre. Le résultat de cette différence impose un 1 ou un 0 logique sur l'entrée D de la bascule, cette valeur est alors mémorisée par le signal prog. Les deux dernières copies du photocourant servent à la contribution de celui-ci dans le courant I.noir ou I.blanc, par addition avec les courants des autres pixels.

Le principe de ce montage est donc de classer les pixels plus lumineux que la moyenne dans la catégorie "pixels blancs" et les pixels plus sombres dans la catégorie 'pixels noirs'. L'aiguillage connecte tous les pixels de chaque catégorie ensemble, ce qui a pour effet que les pixels de la première catégorie contribuent au courant noir I.noir et les autres au courant blanc I.blanc. Ces courants sont en fait la somme des photocourants de chaque composante.

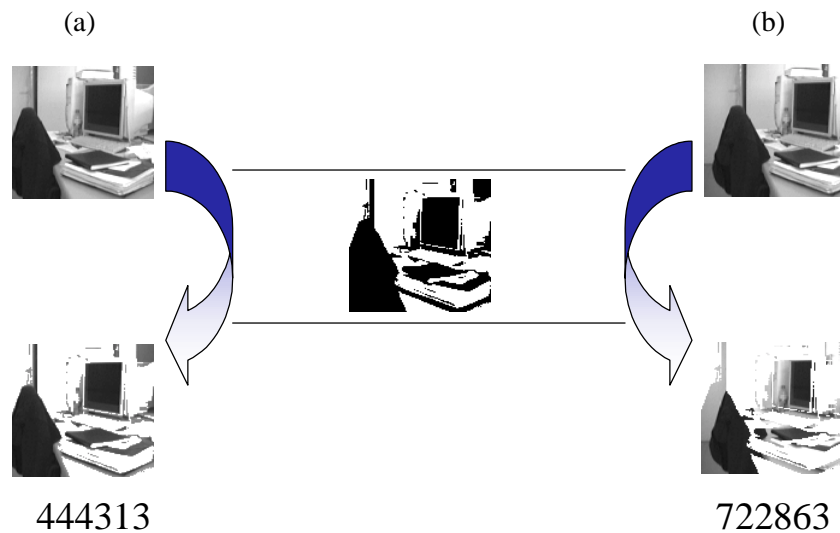
### I-3 Principe

Après mémorisation de l'image de référence, une image observée est caractérisée à tout instant par ses deux valeurs analogiques I.blanc et I.noir. Lorsque cette image change de position ou de forme, des pixels classés noirs deviennent blancs (et vice-versa), donc les intensités des courants changent. Le principe de fonctionnement est illustré par les images de la figure III.3. L'image (a) est antérieure à l'image (b) lors d'une rotation de la caméra de gauche à droite.



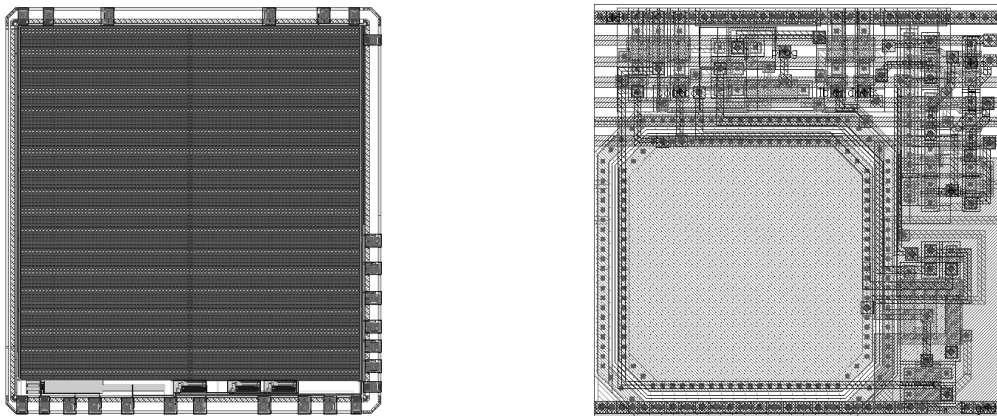
**figure III.3 : Images en niveaux de gris et leur binarisation**

Les images acquises, en niveaux de gris sont stockées de manière binaire. Dans cet exemple, chaque image binaire compte un nombre différent de pixels blancs ou noirs: l'image binaire (a) a 6347 pixels blancs et 6635 pixels noirs, l'image binaire (b) a 6328 pixels blancs, et 6654 pixels noirs. Cet exemple montre qu'une simple mesure des pixels noirs et blancs ne s'avère pas suffisante, les valeurs sont en effet assez proches, de l'ordre de 0,1 %, ce qui est source d'erreurs. C'est pour cette raison que lors de la mémorisation de l'image de référence, les pixels sont associés à une catégorie, blanc ou noir. En considérant l'image (a) en niveaux de gris comme image de référence, l'image binaire (a) représente alors un masque de calcul. La somme des niveaux de gris de l'image (a) à travers son masque binaire vaut 444313 pour la composante noire, et l'image en niveaux de gris (b) donne 722863 à travers ce même masque (figure III.4). La différence permettant le discernement des scènes s'élève alors à environ 60 %. Cela s'explique par le fait que les pixels blancs ont naturellement une valeur supérieure (comprise entre la moyenne et 255) à celle des pixels noirs (comprise entre 0 et la moyenne). En effet, on remarque sur l'image masquée de (b) que des parties plus lumineuses (comme l'arrière de la chaise) sont désormais plus lumineuses, ce qui explique que le niveau de la composante noire augmente. De la même manière, la valeur de la composante blanche s'en trouve diminuée.



**figure III.4 : Calcul de la composante noire sur les deux images (a) et (b)**

Cette rétine a une résolution de 100x100 pixels. Chaque pixel mesure 50 $\mu$ m de côté, donc occupe une surface de 2500 $\mu$ m<sup>2</sup>, et a un facteur de remplissage de 36%. Le circuit total, figure III.5, occupe une surface de 31 mm<sup>2</sup>.



**figure III.5 : Dessin des masques de la rétine d'intercorrélation et d'un pixel de la rétine**

## I-4 Résultats

Les mesures, principalement effectuées par l'équipe de l'Université de Bourgogne, au Creusot, ont montré que le circuit est capable de reconnaître des objets ou leur position - au préalable mémorisées - avec une précision d'environ 1mm sur une surface image de 50cm de côté. L'application directe qui peut en découler est le tri. Ce circuit est par exemple capable

de différencier un billet de banque d'un autre, ou le bon positionnement d'un objet sur une chaîne de production. Il a aussi l'avantage de pouvoir être détourné de sa fonctionnalité initiale. Grâce à sa structure, il est en effet possible de faire des mesures de photocourant intéressantes. Ainsi, nous nous sommes intéressés à des mesures de l'ordre de grandeur des courants en sortie, qui reflètent la valeur du photocourant. En effet, pour cette technologie AMS 0,6 $\mu$ m, ces géométries de photodiodes et notre système optique, nous n'étions pas en mesure de connaître précisément la valeur des courants. Comme ce circuit permet de connaître la somme des photocourants mémorisés, l'ordre de grandeur du photocourant a été déterminé. En effet, le photocourant total, issu de la somme de tous les photocourants unitaires, est mesurable. Le photocourant unitaire (d'un pixel) est alors facile à déterminer. Pour un éclairage fort de 60000 lux, il est de 100nA sans objectif, et de l'ordre de 100 pA avec notre objectif de focale 10 mm réglé à l'ouverture maximale 1,9. Ces mesures sont nécessaires, car les outils de conception n'intègrent pas les caractéristiques liées à la phototransduction. En effet, les modèles des composants sont ceux d'une technologie standard, qui n'est pas a priori faite pour fabriquer des capteurs ou des rétines.

Compte tenu des premiers résultats encourageants de la rétine en phase de reconnaissance et du travail nécessaire à sa mise en œuvre, à sa caractérisation, ainsi qu'à l'étude des améliorations de ce circuit, une thèse s'est orientée sur ce projet à l'université de Bourgogne.

Nous nous sommes aperçus, lors de l'utilisation de la rétine d'intercorrélation, des problèmes liés à la caractérisation d'une rétine. Ces problèmes sont des paramètres à prendre en compte lors de la conception de rétines futures. Premièrement, il est important de pouvoir observer l'image originelle que le capteur reçoit, ce que la rétine d'intercorrélation ne fait pas. En effet, une rétine donne une information sur la scène observée, mais le fonctionnement de celle-ci ne peut être correctement validé que si on est à même de comparer le traitement effectué sur l'image brute et l'image traitée par la rétine. Aussi, la mise en œuvre en est facilitée, autant par la mise en place du système optique (par exemple la focalisation), que par la vérification du bon fonctionnement du circuit. Aussi, l'expérience nous a montré que, face à une rétine, il vaut mieux respecter quelques règles de "dessin en vue de la caractérisation", comme par exemple de nombreux points d'accessibilité et d'observabilité, au prix de quelques plots et/ou des multiplexeurs et démultiplexeurs. Une rétine est réalisée pour valider une structure à travers la mesure. Son fonctionnement, et ses éventuelles dérives, doivent pouvoir être correctement mesurés. Face à la mise en œuvre du circuit, il ne faut pas chercher une

simple validation fonctionnelle, mais une caractérisation (par exemple de la plage de fonctionnement, de la sensibilité ...). Pour permettre ces mesures, certaines parties du circuit doivent être conçues dans cet esprit.

Nous avons aussi tiré des conclusions par rapport à la structure du circuit. En effet, le pixel composé de miroirs de courant a des inconvénients. Par exemple, nous avons constaté que les transistors de sortie sont déséquilibrés par rapport aux transistors de l'étage cascode. La copie du courant n'est donc pas précise. Aussi, l'amplification des photocourants à travers les miroirs est très limitée, puisque directement liée à la taille des transistors. Le mode courant n'est donc pas intéressant avec cette structure.

Pour compléter l'étude de la rétine d'intercorrélation, la conception et la caractérisation d'un imageur se sont imposés. Nous avons donc décidé de concevoir un imageur CMOS de résolution moyenne - en mode d'intégration du photocourant. Cette résolution doit nous permettre de travailler sur de réelles images, et d'avoir un aperçu des dispersions existantes.

## **II. Imageur CMOS**

### **II-1 Contexte**

La rétine d'intercorrélation précédente nous a permis de valider une hypothèse de fonctionnement, de confirmer l'ordre de grandeur des photocourants mis en jeu, mais il donne une information globale qu'il a fallu compléter par une étude plus précise.

Nous avons donc conçu un circuit composé d'un nombre de pixels plus important, en pouvant accéder à l'information contenue dans chaque pixel. De cette manière, nous pouvons nous rendre compte des dispersions existantes entre les pixels d'un même circuit, et par la même occasion avoir une image du bruit spatial fixe dont les publications font état (voir chapitre I), mais sur lequel n'avions pas de mesures, et pas de connaissances suffisantes. Notons cependant que sa structure n'est pas celle d'un capteur APS 3 transistors classique, qui nécessite une structure de réduction du bruit spatial. Sa structure est choisie en vue de concevoir des rétines, donc pour générer de faibles dispersions. Pour cela, il comprend des transistors de type N et P, et mémorise l'image entière de manière simultanée dans tous les pixels.

Ce circuit a été réalisé en technologie mixte 0.6  $\mu\text{m}$  de AMS, et a comme principe d'acquérir une image en pouvant la mémoriser de manière binaire dans les pixels. Il fait donc

partie des imageurs digitaux (DPS : "Digital Pixel Sensor") à intégration.

## II-2 Structure et fonctionnement

Pour le choix des géométries du pixel et de la photodiode, nous nous sommes inspirés de la bibliographie et d'imageurs provenant de l'industrie. En effet, ST-Microelectronics propose à la vente des capteurs CMOS de technologie  $0,8\ \mu\text{m}$  dont les pixels mesurent  $12\ \mu\text{m}$  de côté, nous avons donc conçu des photodiodes mesurant environ  $10\ \mu\text{m}$  de côté en technologie  $0,6\ \mu\text{m}$ . Les caractéristiques des imageurs industriels ne sont qu'une indication puisque les technologies ne sont pas les mêmes, cependant, elles ont confirmé les ordres de grandeur donnés dans les publications [Che-00]. Les photodiodes de  $30\ \mu\text{m}$  de côté utilisées dans le circuit d'intercorrélation sont en effet sur-dimensionnées pour un fonctionnement en mode d'intégration. La figure III.6 montre le schéma simplifié au niveau portes du pixel de cet imageur, et le dessin des masques (layout) correspondant.

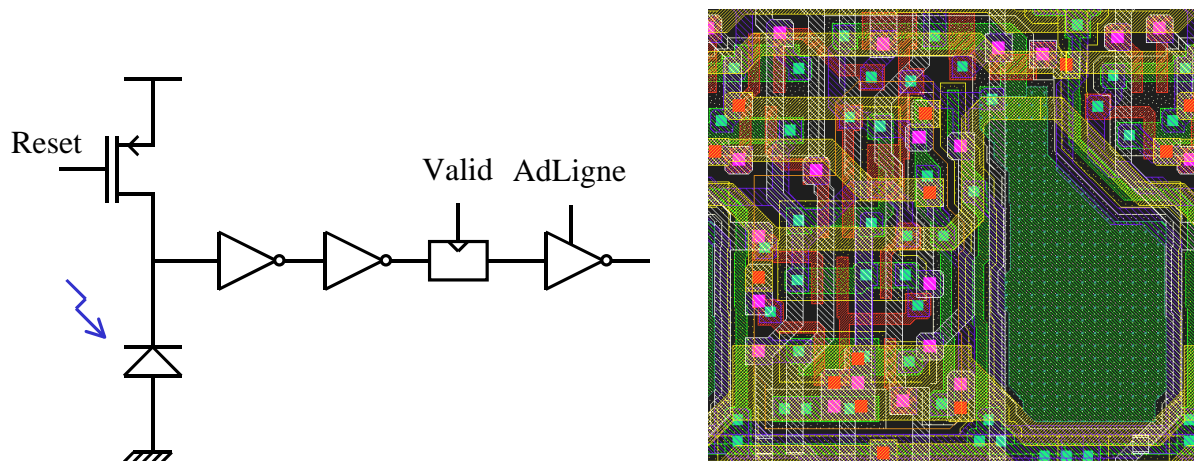


figure III.6 : Schéma de principe du photosite de l'imageur et sa vue "layout"

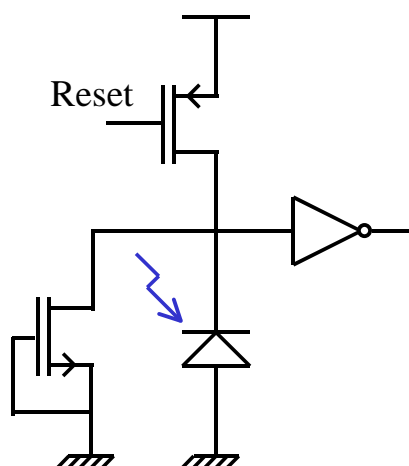
Le pixel est principalement composé d'une chaîne d'inverseurs qui convertissent la tension de la photodiode en valeurs binaires '1' ou '0'. Un transistor précharge la photodiode, nous l'avons choisi de type P pour garder une grande dynamique, et éviter l'offset négatif - présent dans les structures APS classiques - lors du blocage d'un transistor N. Un des avantages de ce circuit est qu'il n'y a pas d'amplificateur de colonne pouvant entraîner un bruit de type colonne sur l'image. Le principal défaut est un bruit de type pixel, puisque la conversion se fait dans le pixel et que les inverseurs peuvent avoir un seuil différent d'un pixel à un autre. Cette chaîne d'inverseurs a été dimensionnée pour avoir une consommation la plus faible possible, et pour pouvoir charger convenablement la capacité de colonne en sortie. En



effet, la pente de la décharge de la photodiode peut être lente, si la luminosité est faible, ce qui peut entraîner un courant de court circuit – à travers les transistors N et P des inverseurs – d’une certaine durée. Ceci explique que le premier transistor de la chaîne ait un W faible et un L grand, puisque le courant de court circuit est proportionnel à  $W/L$ . Les autres transistors composent une chaîne d'inverseurs, dont les tailles ( $W/L$ ) sont successivement plus fortes pour pouvoir piloter le nœud de la colonne, à forte composante capacitive.

Le fonctionnement du circuit est le suivant : tout d’abord, pour initialiser le circuit, les photodiodes sont préchargées en polarisation inverse à  $V_{dd}$  à l’aide des transistors de reset. Cet état correspond à l’état d’initialisation des photodiodes. Ces transistors sont ensuite bloqués, et la capacité des photodiodes est déchargée par le photocourant (voir chapitre I), nous assistons alors à une décharge proportionnelle à l’information lumineuse. Après un temps d’exposition donné, les pixels les plus éclairés ont une tension aux bornes de la photodiode faible, donc un ‘1’ en sortie du premier inverseur, alors que les pixels plus sombres donneront un ‘0’. A l’aide du signal "Valid", ces informations binaires peuvent être stockées dans un point mémoire statique composé de deux inverseurs. Enfin, un inverseur 3 états, commandé par le signal "ADLigne", peut fournir l’information sur la colonne lorsqu’une ligne est sélectionnée.

Un des inconvénients du mode à intégration du photocourant est qu’une fois la capacité déchargée, rien n’empêche la photodiode de se polariser en direct. Des injections de porteurs peuvent alors se faire dans le substrat, entraînant un parasitage inter-pixels. Cet effet est connu [Fuj-00], il produit un éblouissement d’un pixel sur les autres ("blooming"). Pour éviter ce phénomène, un transistor doit être placé en parallèle avec la photodiode, figure III.7.



**figure III.7 : Photosite avec diode de protection**

Ce transistor de protection a la fonction d'une diode, et limite la tension de la cathode de la

photodiode, en ce sens qu'il se met à conduire lorsque sa tension  $V_{gs}$  devient positive (en conduction sous le seuil). Le photocourant traverse alors cette diode.

Pour pouvoir caractériser plus précisément le pixel, nous avons remarqué que la structure d'anti-éblouissement pouvait être modifiée et utilisée pour recopier un courant. En effet, dans un miroir de courant, le transistor recopiant le courant a la même tension grille – source que le transistor imposant le courant à copier. En bloquant ces transistors, nous créons alors un dispositif d'anti-éblouissement; ils se mettent à conduire sous le seuil si la phototension devient négative. En polarisant convenablement ces transistors, nous pouvons imposer un courant externe. Cela permet d'injecter sur les photodiodes un courant pour émuler un éclairage lorsque le circuit est dans le noir. Comme ce courant est de l'ordre de grandeur du photocourant, de quelques dizaines de pico-ampères, nous avons placé un miroir de courant cascode composé de deux transistors en série, afin d'obtenir une grande résistance de sortie. La figure III.8 montre la structure du pixel avec un miroir cascode classique.

La matrice de cet imageur est composée de 213x160 pixels, c'est à dire 34080 pixels, soit le neuvième de la résolution d'un affichage VGA (640/3 x 480/3). La matrice de pixels est associée à un décodeur de lignes, et à une machine d'état possédant différentes fonctionnalités, comme un fonctionnement interne ou contrôlé de l'extérieur, un timer programmable, et un contrôleur VGA.

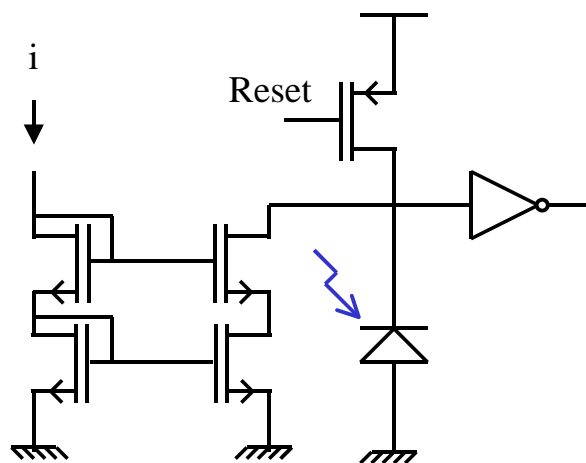
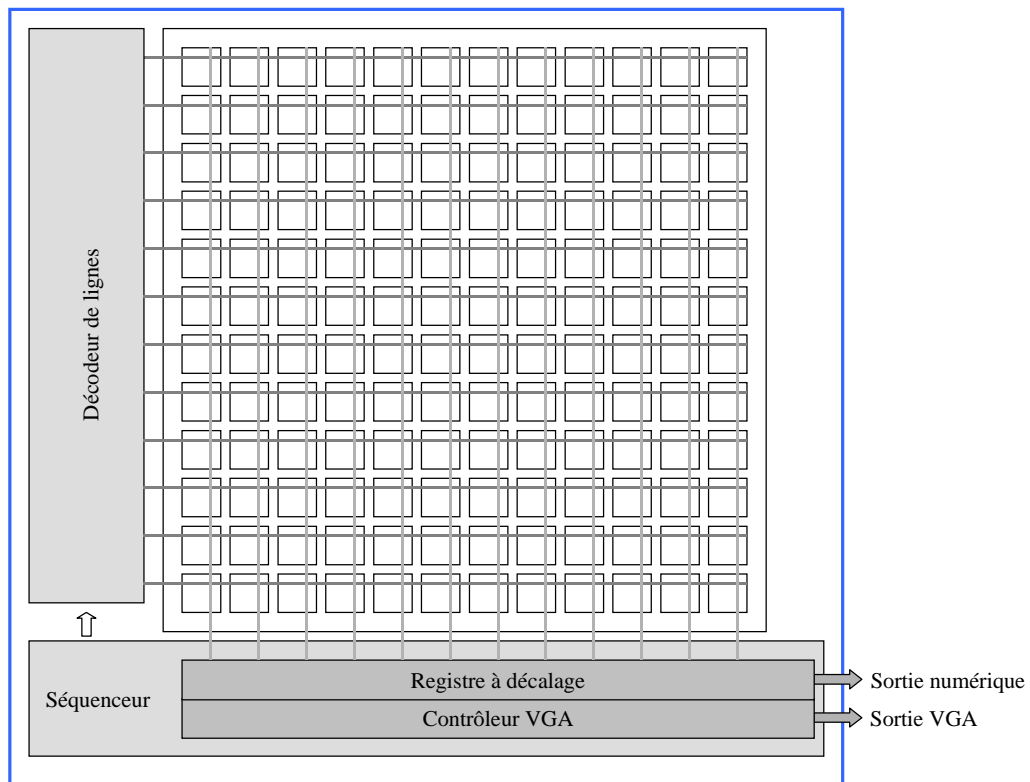


figure III.8 : Photosite avec miroir cascode

Il est possible d'utiliser différents modes de sortie des informations, soit une sortie composée d'un registre à décalage de 213 bascules, soit d'une sortie compatible aux normes d'affichage VGA. Cette information est composée de trois sorties : la synchronisation des lignes (synchronisation horizontale), la synchronisation des trames (synchronisation

verticale), et enfin la sortie qui code l'information de luminosité. Il faut relier les fils des trois composantes (rouge, vert et bleu) d'un écran pour observer les pixels noirs ou blancs. Cette réalisation montre un des avantages des capteurs en technologie CMOS standard : nous avons développé un contrôleur VGA en VHDL et nous l'avons intégré dans le capteur, obtenant un système sur puce compact. Le circuit se connecte directement à un écran d'ordinateur, ce qui nous a permis une mise en œuvre rapide (vérification de fonctionnalité globale, réglage du banc optique). Dans ce mode de fonctionnement, le circuit est totalement autonome et ne nécessite qu'une horloge extérieure cadencée à 25 MHz. La figure III.9 schématise les blocs qui composent le circuit, et la figure III.10 le dessin des masques de cet imageur. La partie appelée séquenceur gère le décodage de ligne, et l'aiguillage des données vers la sortie du circuit.



**figure III.9 : Schéma bloc du circuit**

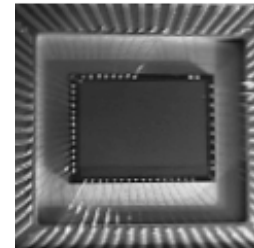


**figure III.10 : Microphotographie de pixels et vue layout du circuit imageur**

Le bloc numérique est une machine d'état, regroupant le séquenceur, le registre à décalage, le contrôleur VGA et un timer programmable a été décrit en VHDL, synthétisé par AMBIT et placé-routé par Silicon Ensemble de Cadence. Il a pour fonction de générer les signaux de Reset, de mémorisation dans les pixels - après un temps d'exposition paramétrable de l'extérieur par des switchs - et de balayer successivement les lignes afin de charger le registre à décalage. Cette ligne est ensuite aiguillée vers la sortie.

La figure III.11 résume les principales caractéristiques de l'imageur CMOS.

Procédé de fabrication	AMS 0.6 $\mu$ m mixte
Taille du circuit	30 mm <sup>2</sup>
Résolution - Nb de pixels	213x160 - 34080
Taille du pixel	28 $\mu$ m x 28 $\mu$ m
Surface de la photodiode	123 $\mu$ m <sup>2</sup>
Taux de remplissage	16 %
Nombres de portes du bloc numérique	3104
Consommation (sous 5v)	150 mW

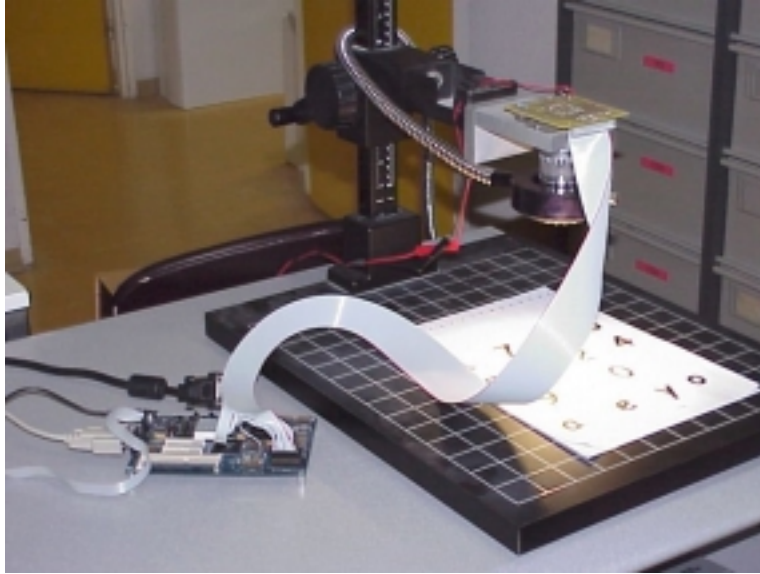


**figure III.11 : Caractéristiques du circuit imageur et sa microphotographie**

### II-3 Choix du montage de caractérisation

Après avoir vérifié la fonctionnalité du séquenceur dans ce capteur, les mesures ont nécessité une mise en œuvre plus complexe. Un banc d'optique sur pied a été utilisé, avec un éclairage artificiel dédié et un objectif de focale 10 mm (l'ouverture pouvant varier de 1,9 à

22). Pour caractériser le circuit, celui-ci est configuré dans un mode de fonctionnement entièrement externe. Des périphériques peuvent alors piloter toutes les entrées du circuit, permettant un grand nombre de mesures. La figure III.12 montre le montage expérimental mis en œuvre pour caractériser l'imageur.



**figure III.12 : Montage expérimental de l'imageur**

Cette partie présente les différents périphériques que nous avons utilisés, en donnant leurs limites. Ces périphériques sont des équipements courants. Aussi, cette étude montre au lecteur comment choisir directement un montage correspondant à son circuit et à ses mesures. Nous avons mis en œuvre le circuit de 3 manières différentes: avec 2 cartes de prototypage Altera, et avec une carte d'acquisition de PC. Notre étude porte sur ces 3 cartes permettant le contrôle du circuit et l'acquisition des images.

**- Carte Altera UP1 avec le logiciel MaxPlus+2.**

Pour utiliser la carte de prototypage Altera UP1, nous avons programmé un FPGA FLEX 10K20 (de 20 000 portes) en VHDL comportemental, en utilisant le logiciel MaxPlus+2 fourni avec cette carte. Il s'est avéré en effet que les fonctions primitives proposées en vues schématiques ne sont pas suffisantes pour contrôler le circuit, c'est pourquoi nous avons développé des macro-blocs en VHDL. Cependant, ce langage n'est pas bien adapté pour cette utilisation. En effet, le contrôle d'un circuit ne nécessite qu'un séquençement d'instructions en respectant un timing précis, alors que la nature du VHDL est de traiter les instructions de manière concurrentes. De nombreuses boucles et processus interconnectés sont nécessaires pour finalement effectuer un contrôle simple.

Aussi, bien que le nombre d'entrées-sorties accessibles soit important, la transmission des données pose un problème. En effet, l'acquisition des données dans le FPGA se fait aisément, mais pour des problèmes de taille mémoire et pour étudier ces données, il convient de les transférer vers un PC. Nous avons alors développé un protocole de communication entre la carte UP1 et le PC, en programmant en VHDL (dans le FPGA) un protocole d'envoi des données sur 8 bits, et en Visual C (sur PC) un protocole de réception des données par le port parallèle du PC. Ce mode de transmission est lourd à mettre en œuvre en VHDL (voir ci-dessus), et la vitesse de transfert est limitée. En effet, la fréquence de travail de la carte Altera est de 25 MHz, et comme l'acquisition des données (une image : 34080 pixels) provenant de l'imageur se fait de manière série, elle nécessite 1,3 ms. Ensuite, lors du transfert des données de la carte au PC, le port parallèle est limité à quelques centaines de Kbits/s. En considérant une vitesse du port de 300 Kbits/s et un envoi des données par octets, 14 ms sont nécessaires. Ces limites, associées à la faible taille mémoire disponible dans le FPGA, ne permettent pas d'acquérir des images en niveaux de gris en des temps raisonnables.

Il s'est aussi avéré que l'utilisation de différents systèmes entraîne des problèmes de synchronisation ; et la mise en œuvre est peu souple compte tenu des différents logiciels devant cohabiter. Notons que MaxPlus+2 utilise lui aussi le port parallèle pour programmer le FPGA. Enfin, la carte UP1 est compatible TTL, et se destine aux circuits (imageurs) ayant des entrées-sorties logiques compatibles TTL (logique 5v).

Pour régler ces problèmes de taille mémoire et de communication, nous avons ensuite utilisé une carte d'acquisition PC.

#### **- Carte d'acquisition PC ADLink 7200 32 voies numériques**

Cette carte, installée sur un port PCI du PC, permet une mise en œuvre bien plus aisée sur plusieurs plans. Ses 32 voies d'entrées-sorties sont suffisantes pour adresser convenablement un circuit. Aussi, la programmation est simple en utilisant les routines propriétaires ADLink. Ces routines sont utilisables dans des programmes en Visual Basic, Visual C++ ou C/C++ sous DOS. L'image transite alors à travers la carte et est stockée directement dans la mémoire du PC.

Les limites de ce type de solution résident principalement dans la vitesse de fonctionnement. En effet, la carte utilisée a une fréquence de travail de 2 MHz; en pratique, la fréquence de fonctionnement est proche de 1 MHz. Cette fréquence de travail ne permet pas de piloter les circuits assez vite pour certaines caractérisations. Il s'est aussi avéré que le timer intégré à cette carte a une période imprécise (non constante), qui ne le dédie pas à ce type de mesures.

En effet, un timer imprécis entraîne des temps d'exposition mal maîtrisés.

Il existe cependant des cartes d'acquisition (numériques et/ou analogiques) plus performantes, mais leur coût est généralement très élevé. Notons aussi que les cartes d'acquisitions numériques sont généralement compatibles TTL, et sont donc destinées aux circuits ayant des entrées-sorties 0 / 5v.

#### - Carte Altera Excalibur NIOS et Quartus II

La carte Excalibur NIOS est, à notre sens, un périphérique qui convient bien au contrôle de circuits. Un FPGA APEX 20K200 (de 200 000 portes), connecté à une horloge de 33 MHz, est – entre autres - connecté à de la mémoire externe, un timer, et un port de communication série. A l'aide du logiciel Quartus II, ce FPGA peut être programmé avec un macro-bloc propriétaire d'Altera, le processeur NIOS. Un processeur 32 bits est alors utilisable à partir d'un haut niveau de programmation : le langage C. L'intérêt de cette carte est que les accès mémoires et la communication série avec le PC sont gérés par le processeur sans que l'utilisateur ait à les programmer.

L'utilisateur écrit un programme en C - incluant des fonctions spéciales comme les accès aux broches d'entrées-sorties ou la gestion de timing – qu'il compile à travers un "shell" dans une fenêtre de commande. Une commande d'exécution envoie ensuite le fichier généré dans le processeur, qui l'exécute. L'association de la programmation de haut niveau et de la gestion automatique des périphériques en font une solution performante et très adéquate aux contrôle de circuits. Le timer, réglable à la  $\mu$ s près, est précis, et les entrées - sorties délivrent les signaux à une fréquence maximale d'environ 15 MHz. L'apprentissage du logiciel Quartus II et du fonctionnement du NIOS sont des étapes d'apprentissage obligatoires, mais il en ressort que cette solution est la plus rapide à mettre en œuvre et la plus souple à utiliser.

Le seul inconvénient que nous avons observé est le brochage des entrées - sorties: ce dernier est entremêlée de broches réservées (GND, horloges ...), le rendant peu portable et limitant les 74 broches disponibles à 40 entrées - sorties.

## II-4 Résultats et mesures

Dans cette partie, nous détaillons des mesures que nous avons menées sur les différents blocs de l'imageur. Nous avons utilisé pour cela un oscilloscope numérique et un analyseur de paramètres (HP 4156A). Cet analyseur de paramètres s'avère un appareil très utile pour

effectuer des mesures précises, qu'elles soient en tension ou en courant.

#### II-4-1 Lumière ambiante

Pour estimer le photocourant, nous avons effectué des mesures en utilisant la structure présentée figure III.13 composant tous les pixels de la matrice. Pour effectuer cette mesure, nous imposons un courant constant de l'extérieur, recopié dans chaque pixel à travers des miroirs de courant, puis nous faisons une acquisition binaire classique : Précharge des photodiodes, décharge par le courant imposé (le circuit est dans le noir), puis lecture des données. En s'intéressant à l'instant de commutation des pixels, nous obtenons le temps de décharge.

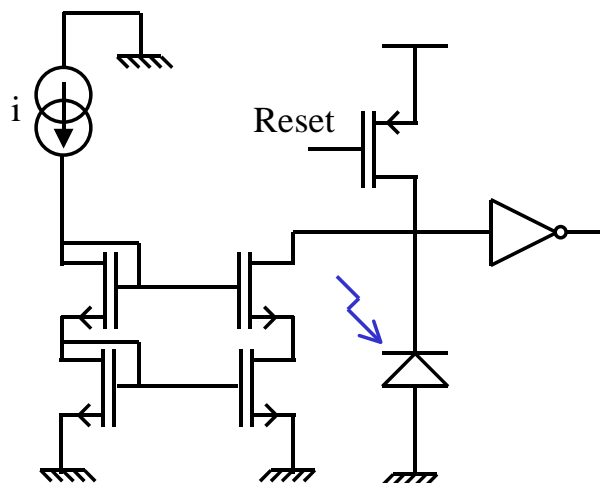


figure III.13 : Montage de mesure de la décharge d'une photodiode à courant contrôlé

Nous avons ainsi imposé un courant extérieur de 50 nA, pour obtenir une demi-décharge en 5 ms, ce qui correspond au temps d'exposition moyen pour une acquisition d'image binaire simple. Ce résultat corrobore nos hypothèses, puisqu'en simulation "corner" (prenant en compte les dispersions liées au procédé de fabrication) ce courant extérieur entraîne un photocourant variant de 40 à 90 pA. En considérant le photocourant lors de la simulation en valeur typique (70 pA), nous pouvons calculer la valeur de la capacité de la photodiode :

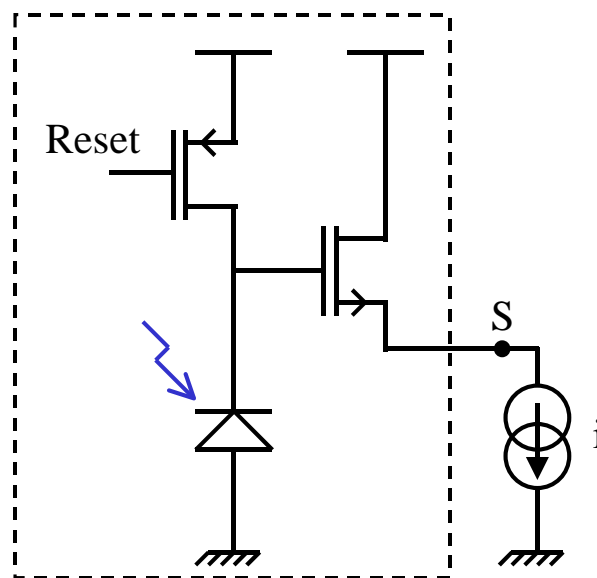
$C = I \Delta t / \Delta U = 140 \text{ fF}$ . Notons que cette valeur est assortie d'une erreur potentielle non négligeable de 57 %, compte tenu des variations du courant imposé dans le pixel et des variations de seuil des inverseurs. La valeur de la capacité mesurée se trouve donc comprise entre 77 fF et 177 fF (valeur typique: 140 fF). Notons cependant que cette valeur est en fait celle de la capacité de la photodiode ajoutée à la capacité du premier inverseur. Pour plus de précision, nous avons pris en compte la capacité caractérisant l'étage de sortie de la



photodiode (composé de la chaîne d'inverseurs). Pour cela, nous avons simulé une rampe de courant attaquant cet étage (pour estimer sa capacité). Les simulations donnent pour la capacité de sortie  $C_s = 17$  fF, nous estimons donc la valeur de la capacité de la photodiode (en perdant de la précision à cause de la valeur simulée), entre 60 et 160 fF, typiquement 120 fF.

En théorie, la capacité de la photodiode est calculée par la relation donnée dans le chapitre I :  $C = A.C_j + P.C_{jsw}$ . La photodiode a une aire de  $120 \mu\text{m}^2$  et un périmètre de  $53,4 \mu\text{m}$ . Compte tenu des paramètres du fondeur et de la tension de polarisation de la jonction (prise à la valeur moyenne de la décharge entre  $V_{dd}$  et  $V_{dd}/2 : 3,75\text{v}$ ),  $C_j = 508.10^{-12}$  F/m<sup>2</sup> et  $C_{jsw} = 452.10^{-6}$  F/m, la capacité théorique vaut donc  $C = 87$  fF. Comme cette valeur est comprise dans la plage de valeurs issues des mesures (entre 60 et 160 fF), nous considérons avoir validé la mesure.

Des cellules de test nous ont permis de faire des mesures directes des décharges des photodiodes, permettant de confirmer les ordres de grandeur des photocourants mesurés avec le circuit d'intercorrélation. Ces cellules de test sont en fait composées d'une photodiode et de 2 transistors; ils ont la structure classique des pixels APS (figure III.14) - sans le transistor de sélection.



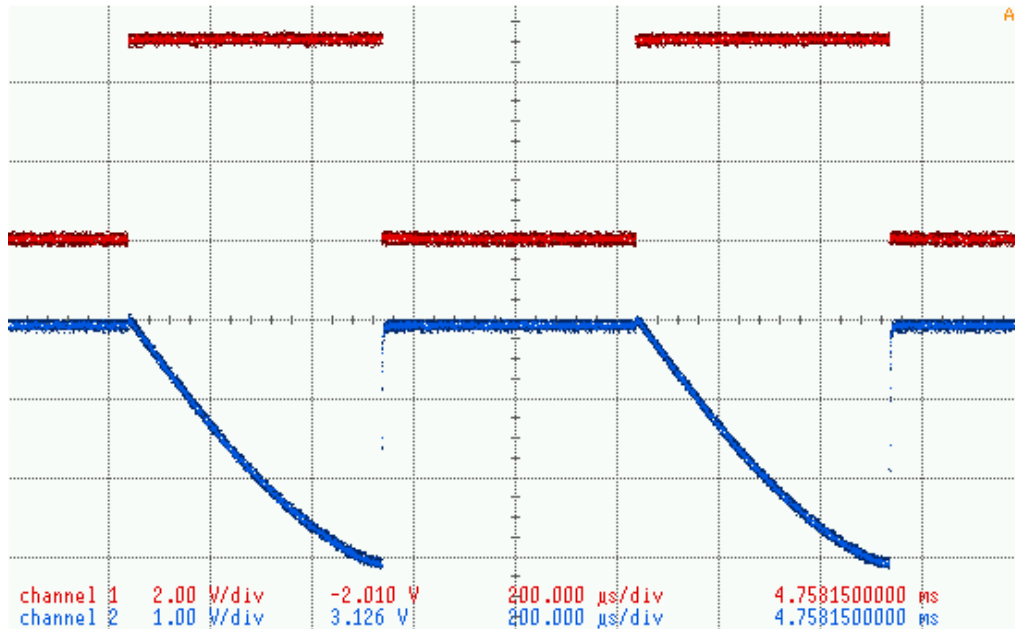
**figure III.14 : Montage de mesure de la décharge d'une photodiode**

L'ensemble transistor de précharge - photodiode est le même dans ce montage et dans les pixels de la matrice (mesure précédente), la valeur de la capacité équivalente est donc connue. Cependant, la capacité de l'étage suivant est différent : la photodiode est connectée à un suiveur. Nous avons effectué la même simulation que pour l'inverseur, estimant ainsi la

capacité du suiveur :  $C_s = 7 \text{ fF}$ . La capacité totale (photodiode et suiveur) varie donc entre 70 et 170 fF, et vaut typiquement 130 fF.

Alors que nous trouvions intéressant de caractériser différents montages grâce à des cellules de test, nous avons tout de même conclu qu'il était nécessaire de concevoir les mêmes pixels dans la matrice et les structures de test, en insérant simplement des points d'observabilité et de contrôlabilité. De plus, l'accès direct à un miroir de courant et à la sortie du pixel aurait permis des mesures plus approfondies du miroir et de la réponse de la photodiode. Cela nous aurait permis d'obtenir des résultats plus concrets, entièrement issus de mesures corrélées. En effet, nous devons maintenant faire des approximations pour recouper les résultats des deux montages, nous perdons ainsi de la précision.

Connaissant la valeur de la capacité, nous pouvons estimer le photocourant en observant le signal au point S (figure III.14). Ce nœud est extérieur au circuit, ainsi que la polarisation du transistor suiveur, assurée par l'analyseur de paramètres. La figure III.15 montre la décharge d'une photodiode à lumière constante, sous un éclairage de 40000 lux (sans objectif).



**figure III.15 : Décharge mesurée d'une photodiode**

L'allure de cette décharge – assez proche du linéaire - correspond bien à celle que nous avons prévue (voir chapitre I). La non-linéarité de la décharge présentée est de 22 %. La décharge de la capacité de la photodiode est pour cet exemple, de  $\Delta V = 3,1 \text{ v}$  en  $\Delta t = 520 \text{ μs}$ . En

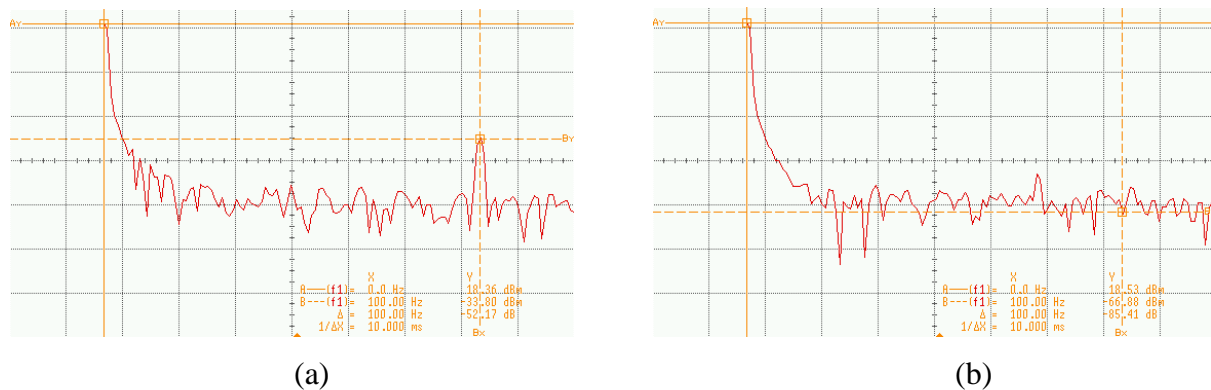
considérant cette décharge linéaire, nous pouvons modéliser la décharge de la capacité à courant constant :  $U = U_0 - I \Delta t / C$ .

Pour calculer le photocourant de la photodiode, nous utilisons la valeur typique de capacité calculée plus haut ( $C = 130 \text{ fF}$ ). En considérant que cette capacité reste constante au cours de la décharge, on peut estimer le photocourant mis en jeu:  $I = C \cdot \Delta V / \Delta t = 0,77 \text{ nA}$ .

En réintroduisant cette valeur de photocourant ( $0,77 \text{ nA}$ ) en simulation (on fixe simplement le générateur de courant équivalent à cette valeur), le temps de décharge (de  $3,1\text{v}$  à  $0\text{v}$ ) relevé est de  $160 \mu\text{s}$ .

Le résultat présente une erreur importante, d'un rapport de 3. Nous estimons que cette erreur provient principalement des approximations sur la valeur de la capacité de la photodiode et, dans une moindre mesure, des imprécisions dues aux capacités parasites de mesure. De la même manière, nous avons mesuré le photocourant à luminosité maximale : il est de  $100 \text{ nA}$  avec un éclairage de  $60000 \text{ lux}$  (halogène de  $50\text{W}$  éclairant le circuit à  $10 \text{ cm}$  sans objectif).

Nous avons aussi relevé la transformée de Fourier des charges et décharges des photodiodes de la figure précédente.



**figure III.16 : FFT des décharges de photodiodes**

Nous avons mené ces mesures avec une lumière continue dans les mêmes conditions que précédemment. La mesure (a) a été faite avec la lumière continue parasitée par les néons de la salle de mesure. Bien que leur contribution soit faible, nous observons leur influence. Le pic mesuré à  $100\text{Hz}$  est en effet dominant par rapport aux autres parasites. Par rapport au signal continu – utile –, il a une atténuation de  $50\text{dB}$ . La mesure (b) a été faite avec le même dispositif, mais avec un éclairage purement continu. Le pic à  $100\text{Hz}$  n'est plus présent, et le niveau moyen des parasites est de  $-80\text{dB}$  par rapport au signal utile.

#### II-4-2 Mesure du courant d'obscurité ou courant de noir

Le courant d'obscurité, courant de décharge des photodiodes lorsque celles-ci sont dans l'obscurité, constitue une des faiblesses des capteurs CMOS. Nous avons vérifié l'ordre de grandeur de ce photocourant, dû aux fuites de la capacité de la photodiode. H. Tian [Tia-01b] estime ce courant à 2,3 fA pour des photodiodes de puits (puits N dans substrat P) de  $57 \mu\text{m}^2$  dans une technologie CMOS 0,35  $\mu\text{m}$ . Pour effectuer la mesure, nous avons placé le capteur dans l'obscurité et observé l'instant de basculement des photodiodes. Nous avons mesuré un temps de décharge (entre  $V_{dd}$  et  $V_{dd}/2$ ) dans le noir de 92 s. Compte tenu de la valeur de la capacité connue, le courant vaut:  $I = C \cdot \Delta U / \Delta t = 3,4 \text{ fA}$ . Il faut prendre en considération l'erreur sur la valeur de la capacité (57%) et l'incertitude sur le temps de décharge (environ 5 %).

#### II-4-3 Capture d'images

Pour caractériser l'imageur, nous avons utilisé une plaquette de prototypage Excalibur-NIOS d'Altera, comprenant un FPGA APEX 20K200. L'environnement permet de programmer le comportement en langage C à travers un processeur NIOS 32 bits - programmé dans le FPGA. Le séquençement du programme est illustré dans la figure III.17 ci-dessous.

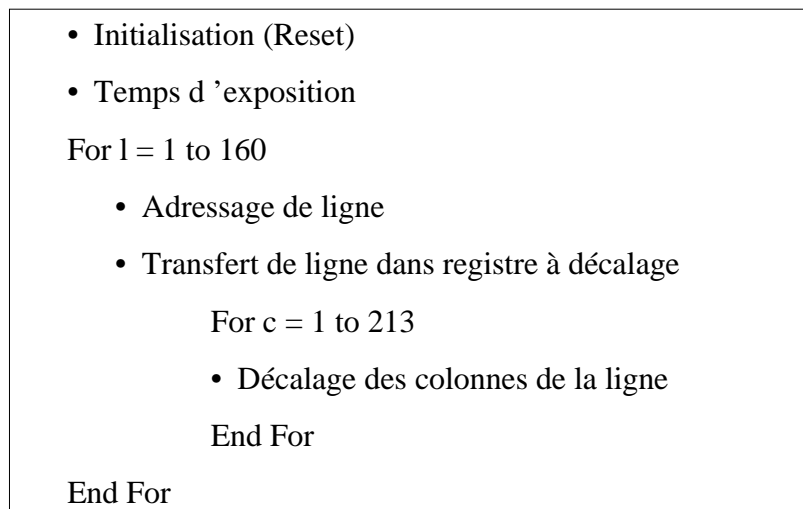
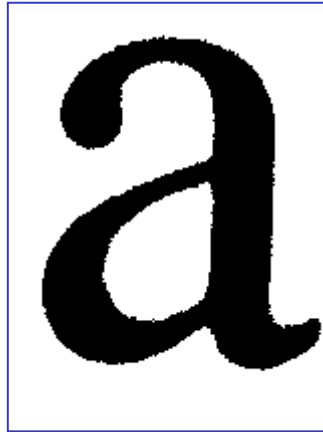


figure III.17 : Séquençement des commandes pour acquérir une image binaire

Après un reset initialisant les photodiodes, le temps d'acquisition est respecté. Ensuite, l'adresse d'une ligne est validée et le contenu de cette ligne entre dans un registre à décalage (voir figure III.9, p.57). Enfin, une horloge extérieure cadence ce décalage. L'avantage de ce registre à décalage est le gain en terme de plots de sortie, il évite ainsi de réaliser un

multiplexage à 213 entrées, nécessitant 10 plots (6 bits d'adresse pour une sortie sur 4 bits). Aussi, ce registre est doté de 4 sorties réparties sur le mot de 213 bits à extraire, pour limiter le nombre de décalages à 54. Nous avons suivi ce protocole pour extraire des images du circuit, dont voici un exemple :

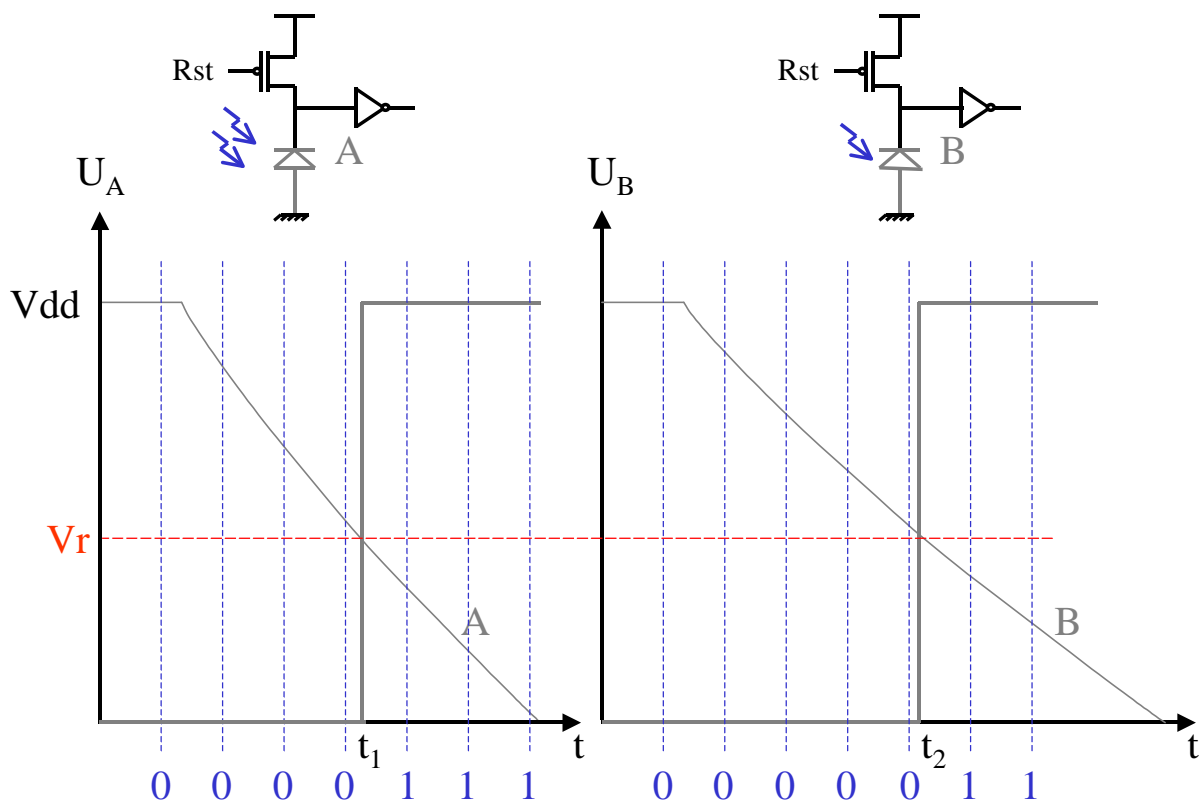


**figure III.18 : Image binaire acquise par l'imageur**

L'inconvénient la structure de cet imageur est que le temps de décalage dans le registre limite la cadence d'extraction des informations. Notons qu'un imageur de 213 x 160 pixels qui génère une image à une cadence de 30 images par seconde impose un flot de donnée de 1 Mbits/s sur 2 niveaux (image binaire). Même à une fréquence de décalage de 33 MHz, 1 ms est nécessaire pour extraire une image binaire de 213x160 pixels, soit 34080 bits. Grâce au découpage du registre à décalage en sous-registres de 54 points mémoires, le temps nécessaire décroît à 250 ns. Ce temps n'est pas négligeable, il pose des problèmes lorsque l'imageur est utilisé pour acquérir des images en niveaux de gris.

Nous avons présenté l'acquisition d'images binaires, nous allons exposer maintenant son utilisation pour acquérir des images en niveaux de gris. La conversion analogique digitale a une dynamique faible (1 bit), qui peut être augmentée en la transposant temporellement. Bien entendu, un calcul est ensuite nécessaire afin de reconstituer l'image en niveaux de gris, mais ce calcul se limite à une accumulation de résultats dans un tableau. Le principe de cette mesure est soit d'échantillonner plusieurs fois chaque pixel pendant une même acquisition, soit d'acquérir plusieurs images en utilisant des temps d'exposition croissants. La figure III.19 montre, à travers la décharge de deux pixels, le principe du multi-échantillonnage temporel. Le pixel A est plus éclairé que le pixel B, sa décharge est alors plus rapide. L'inverseur, jouant un rôle assimilable à celui d'un comparateur à  $V_{dd} / 2$ , fournit le passage de '0' à '1' à l'instant

$t_1$ , avant celui du pixel B (à  $t_2 = t_1 + dt$ ). En échantillonnant les valeurs des pixels aux mêmes instants, la mesure consiste à attacher aux pixels un code température. Pour l'exemple de la figure ci-dessous, nous notons: 0000111 pour le pixel A, et 0000011 pour le pixel B. Cette mesure permet donc l'acquisition de 7 niveaux de gris: 0000000, 0000001, 0000011, 0000111, 0001111, 0011111, 0111111, 1111111. Pour acquérir une image en 256 niveaux de gris, il faut échantillonner 256 fois la décharge des pixels, c'est à dire acquérir 256 images binaires successivement et accumuler les valeurs pour obtenir une dynamique de 256 niveaux. Chaque pixel vaut '0' au moment du Reset, puis passe à '1' à un moment, ce qui donne ce code température. Cependant, les échantillonnages effectués autour de la tension de seuil des inverseurs ( $V_r$ ) peuvent engendrer des erreurs. En effet, si un pixel a une réponse fortement bruitée, son code peut alors être 0001011. Cependant, comme le niveau de gris résulte de l'accumulation des images binaires, c'est à dire du nombre de bits à '1' dans le mot, sa réponse sera identique au code 0000111. Les erreurs temporelles sont donc lissées par ce mode d'acquisition, car tous les bits sont de même poids. De la même manière si un code vaut 0010111 au lieu de 0001111, aucune erreur ne sera notée.



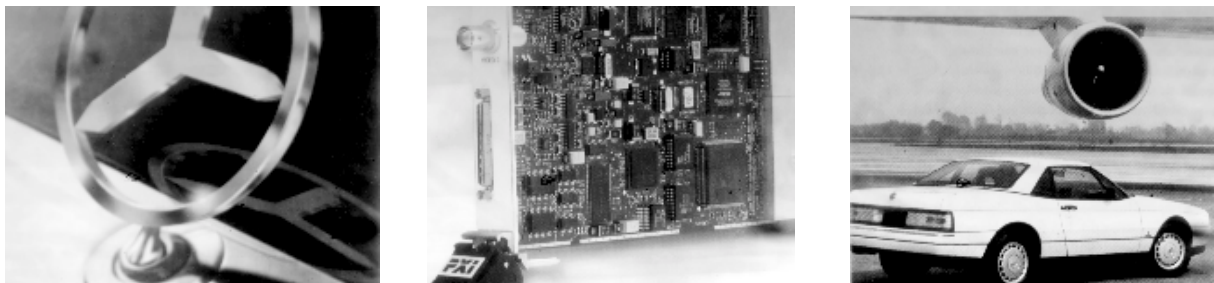
**figure III.19 : Multi-échantillonnage temporel d'une décharge**

Concrètement, à chaque instant d'échantillonnage, le point mémoire du pixel (figure III.6)

stocke un '0' ou un '1' selon la quantité de lumière reçue.

En considérant une décharge de durée 10 ms en échantillonnant 256 fois une même décharge, le système de mesure doit être apte à mesurer un échantillon toutes les 40  $\mu$ s. Cependant, cette cadence ne peut pas être atteinte avec la carte Excalibur NIOS. En effet, le processeur reçoit une horloge à 33 MHz, mais utilise plusieurs cycles machines pour exécuter les lignes de code C. Il en résulte une fréquence de décalage de 1,4 MHz (ou 5,6 MHz en utilisant les 4 bits de sortie), ce qui représente 24 ms (ou 8ms) pour décaler une image binaire. Ce temps est parfaitement incompatible avec des temps d'exposition qui ont le même ordre de grandeur. Même si nous envisagions de recourir au VHDL ou à l'assembleur NIOS, les temps d'accès mémoires restent prédominants. En effet, la mémoire externe est nécessaire pour stocker l'image, mais les accès mémoires prennent beaucoup trop de temps: 12 ms pour le stockage d'une image binaire. Cela vient du fait qu'il faut utiliser le NIOS pour communiquer avec la mémoire externe.

La solution est donc d'acquérir successivement toutes les images, en fixant un temps d'exposition précis pour chacune. Dans ces conditions, un grand nombre d'images peuvent être acquises, puisque les seules limitations sont la taille de la mémoire externe et le temps consacré à cette tâche. Ainsi, nous avons capturé des images de plus de 2500 niveaux de gris, mais le pouvoir de discrimination de l'œil ne fait plus la différence avec une image de dynamique standardisée de 256 niveaux de gris. Nous avons aussi acquis des images en 256 niveaux de gris. En utilisant un programme en C à travers le NIOS, cette acquisition nécessite 10 secondes (36ms x 256). Nous savions de toute façon qu'en transposant le problème de la dynamique dans l'espace temporel, nous serions limités à l'acquisition d'images statiques. La figure III.20 montre des images – de magazine - acquises de cette manière (en 256 niveaux de gris).



**figure III.20 : Images en niveaux de gris reconstitués acquises par l'imageur**

Bien que ce mode de fonctionnement soit limitatif, il faut en effet que la scène observée soit

statique, il offre un avantage certain pour la caractérisation d'une rétine. En effet, sa conception est très simple, et ne nécessite pas de conception analogique poussée : la conversion analogique-numérique se fait très tôt et la dynamique peut être adaptée. Le temps de lecture est alors de 130ns/pixel avec la carte Excalibur. La contrepartie de ce montage est qu'il nécessite une plus grande surface qu'une structure APS 3 transistors. Notons aussi que ce type de mesure entraîne une erreur, car l'image est composée de différentes expositions non corrélées. Toutefois, le bruit doit être limité dans ce type de structure, du fait de l'utilisation d'un transistor P comme transistor de Reset et du fait de la conversion analogique – numérique dans le pixel, évitant par exemple le phénomène de colonne.

Il serait possible d'effectuer cette lecture sans mémorisation dans le pixel de deux manières. La première consiste à lire rapidement l'image. Le temps de lecture correspondrait alors à l'erreur qu'on tolère. Cette solution n'est cependant envisageable que pour les petites matrices. La deuxième repose sur une acquisition pixel par pixel correctement temporisée. Dans ce cas, le temps de capture d'une image sera long.

#### **II-4-4 Mesures des dispersions du capteur**

Nous avons réalisé une étude statistique sur les images acquises pour mesurer les dispersions présentes dans le capteur. Nous avons mesuré la dispersion spatiale et temporelle du circuit. Compte tenu de la structure du pixel, ces défauts proviennent des variations des sensibilités des photodiodes et des seuils des inverseurs.

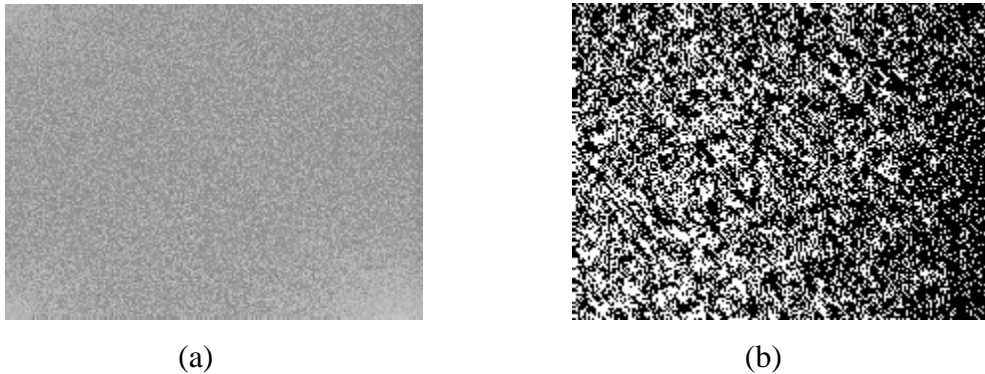
- **Dispersion spatiale**

Sur une image uniforme de luminosité constante, nous avons acquis un grand nombre d'images, en calibrant précisément le temps d'exposition avec la luminosité de la scène, de sorte que le nombre de pixels blancs soit égal au nombre de pixels noirs, c'est à dire  $34080 / 2 = 17040$  pixels de chaque "couleur". Si le capteur et la lumière étaient idéaux, ce classement ne pourrait théoriquement pas être fait, puisque tous les pixels répondraient de la même manière, et donneraient le même niveau de gris. La réponse serait donc uniforme sur toute la matrice.

Pour réaliser cette mesure, les problèmes optiques ne doivent pas influencer sur les résultats. Pour cela, nous avons réalisé les mesures avec une source de lumière continue, puis en émulant un photocourant grâce au dispositif présenté figure III.8. Les mesures réalisées avec une lumière continue offrent en effet un dégradé dû à la non-homogénéité de la lumière. La non homogénéité de l'image ne représente pas les défauts du capteur, mais les défauts de tout



le système de mesure. Nous notons aussi que la dispersion que nous mesurons à l'aide des miroirs de courants est celle des pixels, mais aussi celle des miroirs eux-mêmes. Des erreurs résiduelles sont donc toujours mesurées. Deux images de la dispersion spatiale du capteur sont données figure III.21.



**figure III.21 : Dispersions des pixels de l'imageur**

En faisant une acquisition avec un temps d'exposition classique, choisi en fonction du courant imposé dans le circuit, nous observons une dispersion de type pixel (a). Nous avons mesuré la dynamique de cette dispersion. Elle est de 24 niveaux de gris à travers la matrice, soit de 9,3 % sur 256 niveaux. Notons qu'un imageur standard APS (en technologie CMOS standard) a un bruit s'élevant à 1,7% (3 niveaux sur 256) [Tia-01b], mais l'image n'est pas acquise simultanément sur tous les pixels.

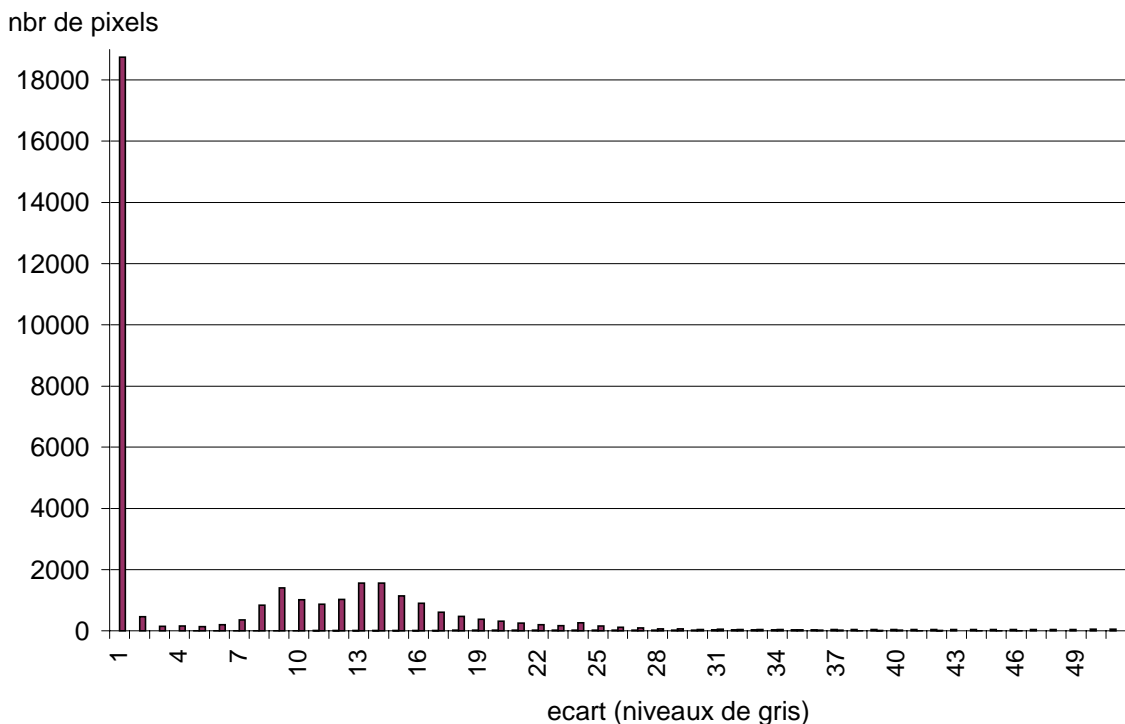
L'image binaire (b) est une mesure plus précise du bruit spatial. Elle est prise avec un temps d'exposition correspondant exactement à l'instant de basculement des points mémoires des pixels. A l'instant de mémorisation des valeurs, les phototensions sont très proches de la tension de seuil. Les très faibles différences de tension au niveau des pixels sont donc mesurées. Nous pensions d'abord mesurer les variations entre les seuils des inverseurs, représentant la dispersion du process lors de la fabrication. Cependant, nous nous sommes aperçus que, lors de la conception, nous avons placé les buffers du signal "Valid" (voir figure III.6, p 54) sur un seul côté, à gauche de la matrice. La propagation de ce signal le long des lignes correspond à la dispersion horizontale que nous observons figure III.21. Elle est cependant très faible, inférieure à un niveau de gris pour une exposition classique de plusieurs millisecondes. Nous ne pouvons cependant pas mesurer le temps de transit à cause de la lecture série à travers le registre à décalage. Cette observation est à prendre en compte pour les futurs circuits, notamment si leur taille est importante. Une prise en compte du temps de propagation et ainsi la conception d'une bufférisation réfléchie est préférable dans ce cas.



- **Dispersion temporelle**

Nous avons mené une étude statistique sur la variation des pixels pour une image fixe. Cette mesure a été menée sur des images réelles, afin de mieux caractériser le système complet face aux variations temporelles, en vue d'une application spatio-temporelle. Cette mesure a porté sur l'acquisition de 100 images identiques. Afin d'obtenir une image de référence, nous avons calculé la moyenne de ces 100 résultats. Ensuite, pour chaque pixel, nous avons retenu, parmi les 100 valeurs, la variation maximale. Notons aussi que cette mesure repose sur des images complètement décorréelées (accumulation d'images binaires). Nous avons donc préféré faire une étude du pire cas, car nous pensons que le résultat reflète plus les problèmes auxquels nous pouvons être confrontés lors d'acquisitions d'images en vue de l'estimation d'un mouvement.

Les résultats sont représentés sur l'histogramme de la figure III.22.



**figure III.22 : Histogramme de la dispersion temporelle de l'imageur**

Les résultats présentés sont relatifs à une mesure réalisée avec une lumière continue bruitée par une lumière artificielle à 100 Hz. L'étude porte sur 100 images de 34080 pixels (213x160). Statistiquement, la majorité (56 %) des pixels est invariante dans le temps. La variation maximale est de 50 niveaux, la quasi-totalité des pixels (96 %) a une erreur inférieure à 25 niveaux de gris. Nous avons constaté que le type de lumière influe beaucoup sur la précision des résultats. La figure III.23 montre que la répartition de l'erreur (l'écart en

niveaux de gris) est globalement homogène entre 10 et 22 niveaux de gris.

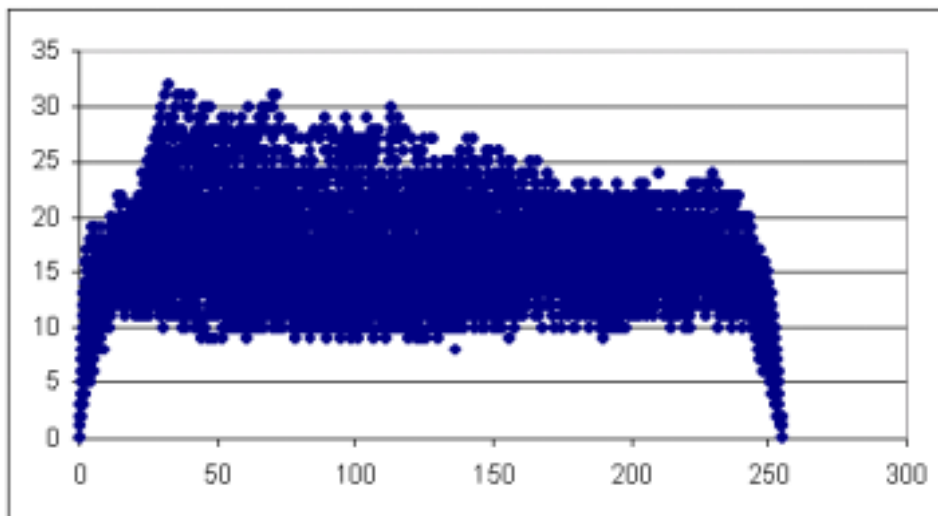


figure III.23 : répartition de la valeur de l'erreur en fonction du niveau de gris

Nous notons cependant que l'erreur est légèrement supérieure pour les niveaux de gris faibles, correspondant à des décharges (de la photoension) longues. Pour une tension de seuil donnée, une décharge faible entraîne en effet un instant de franchissement - et donc un changement d'état des inverseurs - plus sujet à des variations.

### III. Conclusion

Après avoir étudié et finalisé une rétine d'intercorrélation, nous avons remarqué qu'il est important d'avoir une observabilité et une contrôlabilité maximales dans un tel circuit. Aussi, une rétine doit pouvoir fournir l'image originale qui est traitée pour pouvoir valider le traitement effectué. Enfin, elle doit avoir une structure causant le moins de dispersions possibles dans l'image. Afin d'éviter une distorsion de l'image, une rétine doit donc acquérir l'image de manière simultanée sur tous les pixels. Notons aussi que les rétines ne comprennent généralement pas de structure de réduction du bruit spatial.

Pour caractériser une telle structure de capteur actif, nous avons réalisé un imageur binaire (DPS), qui peut aussi acquérir des images en niveaux de gris. Il s'avère que les erreurs (mesurées) dues au bruit dans cet imageur sont de 9,3 %, mais permettent de mémoriser l'image à un instant précis.





# Chapitre III :

## L'ESTIMATION DE MOUVEMENT

---

Ce chapitre présente l'estimation du mouvement. Cette application est complexe et représente un thème de recherche à part entière dans différentes disciplines. Comme nous nous intéressons plus particulièrement au calcul du mouvement dans une rétine, il convient donc de dédier un chapitre à la présentation de cette application. Dans une première partie, nous expliquons son principe et l'équation mathématique modélisant le mouvement. Nous présentons ensuite les méthodes existantes, qui se classent en trois familles: les méthodes différentielles, les méthodes fréquentielles, et les méthodes de mise en correspondance. Ensuite, nous présentons la méthode à laquelle nous nous intéressons. Pour cela, nous détaillons son algorithme, puis nous donnons ses caractéristiques.

### I. Principe

L'estimation de mouvement est une application fondamentale et complexe dans le traitement d'une séquence d'images. Elle a un grand nombre d'utilisations possibles, elle est par exemple utilisée lors de la compression de séquences vidéo utilisant une méthode de prédiction (ex : Mpeg 4). La stabilisation d'images requiert aussi une estimation de mouvement globale. Le mouvement des objets par rapport à la caméra permet aussi, dans les applications de robotique mobile, de recalculer une représentation de l'espace constituant l'environnement du robot.

Une séquence est composée d'images en niveaux de gris, caractérisées à tout instant par leur luminance  $I(x,y,t)$ , représentant le niveau d'éclairement du pixel de coordonnées  $(x,y)$  au temps  $t$ . L'intensité de chaque pixel est donc une image de la puissance lumineuse réellement reçue. En fait, le capteur reçoit sur chaque pixel un éclairement caractérisé par une puissance lumineuse (unité: Watt ou Lumen), et restitue un courant (Ampère).

L'estimation de mouvement est en fait un calcul effectué sur une scène tridimensionnelle (3D), projetée sur le plan du capteur (bidimensionnel – 2D). Le champ de vitesse mesuré est appelé le flot optique.

Pour tout point réel  $(X,Y,Z,T)$  de la scène, ayant une vitesse  $\vec{V}$  et se projetant sur l'image en  $(x,y,t)$ , le flot optique informe d'un mouvement  $\vec{v}$  en  $(x,y)$ , formé des deux composantes orthogonales  $dx$  et  $dy$ .

Couramment, l'estimation de mouvement est basée sur l'hypothèse de conservation de la quantité de lumière de tous les points de l'image. Ainsi le pixel projeté en  $(x,y)$  voit sa luminance se conserver :

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (3.1.)$$

Cette formule exprime le déplacement du pixel caractérisé par son éclairement  $I(x,y,t)$ . Cet éclairement (de même amplitude  $I(x,y,t)$ ) peut se retrouver après un instant  $dt$  et un déplacement de  $(dx,dy)$ .

Cependant, ces conditions ne sont pas respectées dans le cas de scènes réelles. La lumière et le contraste subissent souvent des variations brutales. Ces perturbations sont par exemple des occlusions, des réflexions ou le scintillement à 100 Hz des éclairages artificiels. Pour tenir compte de ces variations, on complète la formule ci-dessus avec des termes correctifs :

$$I(x, y, t) = K(x, y, t) \cdot I(x + dx, y + dy, t + dt) + \delta I(x, y, t + dt) \quad (3.2.)$$

$K(x, y, t)$  et  $\delta I(x, y, t + dt)$  représentent des composantes supplémentaires prenant en compte les discontinuités, inconnues qu'il faut déterminer.

Toutefois, les modélisations complexes du mouvement nécessitent un grand nombre de paramètres et de variables, et conduisent à des opérations complexes et des solutions peu stables. En pratique, l'équation de la conservation de la luminance (3.1) est la plus utilisée.



## II. Equation du flot optique

L'équation du flot optique modélise le mouvement d'un pixel dans une séquence d'images. Les paramètres spatiaux et temporels sont alors pris en compte. Cette équation repose sur l'hypothèse d'illumination constante de l'image acquise. Elle provient donc de l'équation (3.1), en considérant une fréquence d'échantillonnage élevée [Hor-81]. Pour résoudre cette équation (3.1), on la développe en série de Taylor - de type :  $y(t + \delta t) = y(t) +$

$\delta t \cdot \frac{dy}{dt} + \frac{\delta t^2}{2} \frac{d^2y}{dt^2} + \dots$  -, on obtient (au premier ordre):

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + \frac{\delta I}{\delta x} dx + \frac{\delta I}{\delta y} dy + \frac{\delta I}{\delta t} dt + R \quad (3.3.)$$

R est le reste de Taylor représentant les ordres 2 et supérieurs des dérivées. En utilisant (3.3) dans (3.1) et en négligeant le reste R, on obtient l'équation :

$$\frac{\delta I}{\delta x} dx + \frac{\delta I}{\delta y} dy + \frac{\delta I}{\delta t} dt = 0 \quad (3.4.)$$

En divisant par dt et en posant que  $\frac{dx}{dt} = u$  et  $\frac{dy}{dt} = v$ , on obtient l'équation du flot

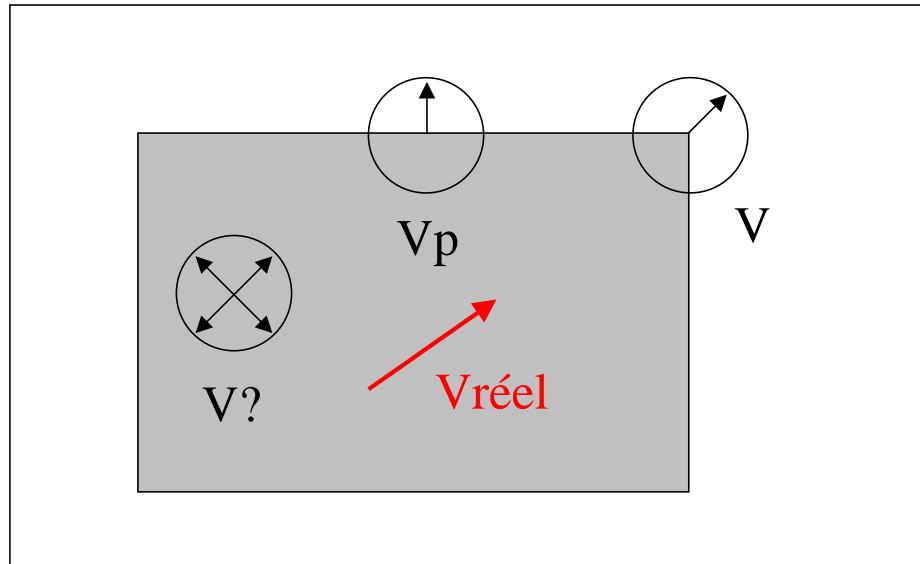
optique :

$$\frac{\delta I}{\delta x} u + \frac{\delta I}{\delta y} v + \frac{\delta I}{\delta t} = 0 \quad (3.5.)$$

$\frac{\delta I}{\delta x}$  et  $\frac{\delta I}{\delta y}$  sont les variations spatiales (gradients),  $\frac{\delta I}{\delta t}$  exprime les variations de la luminosité dans le temps. Cette équation donne la relation du mouvement formé des 2 composantes orthogonales u et v, et d'une illumination restant constante dans un domaine tridimensionnel (x,y,t : spatio-temporel), tel que l'intensité d'un pixel (x,y,t) donné se retrouve un instant dt plus tard à une autre position (x',y',t').

Comme l'explique C. Bernard dans sa thèse sur le flot optique et l'interpolation irrégulière [Ber-99], la résolution de cette équation du flot optique est un problème mal posé, puisque cette unique équation ne permet pas de trouver les deux inconnues u et v. Le système

est sous-déterminé. Seule la composante du mouvement perpendiculaire aux contours peut être estimée, ce problème est connu sous le nom de problème d'ouverture ("aperture problem") [Bea-95]. La figure II.1 en donne un exemple.



**figure II.1 : Illustration du problème d'ouverture**

Dans cette zone homogène dont le mouvement réel est de direction diagonale, le mouvement ne peut pas être déterminé – il est noté  $V?$  – car le gradient spatial (composé de  $\frac{\delta I}{\delta x}$  et  $\frac{\delta I}{\delta y}$ ) est nul. En effet, en faisant abstraction des pixels au bord de cette zone, tout point  $(x,y)$  a la même illumination  $I(x,y)$  que ses voisins. Sur un contour, le mouvement déterminé -  $V_p$  - est perpendiculaire à ce contour, car le gradient n'est orienté que dans une seule direction ( $\frac{\delta I}{\delta y}$ ). Ceci est dû à l'homogénéité des zones en mouvement, faisant que, localement, un déplacement diagonal ne se démarque pas d'un mouvement d'angle différent : la composante en abscisse du gradient est nulle. Seule la combinaison d'informations issues de différents contours permet de calculer le mouvement réel ( $V = V_{réel}$ ). Dans ce dernier cas, l'estimation du mouvement n'est possible que par l'utilisation de paramètres issus d'un voisinage spatial. Nous remarquons que la mesure du flot optique est en fait effectuée sur la projection du gradient. La solution est alors de considérer la moyenne du mouvement sur ce voisinage. Cependant, en plus de la conservation de la luminance, une hypothèse de continuité spatiale du flot optique est donc nécessaire pour déterminer le mouvement. Le cas le plus fréquent de discontinuité spatiale du flot optique se situe sur les zones de transition séparant des régions

en mouvement dans l'image. L'équation du flot optique (3.5) ne permet pas de déterminer le mouvement dans de tels cas.

### III. Méthodes pour estimer un mouvement

Pour obtenir une information locale de mouvement, les différentes méthodes nécessitent la prise en compte d'un voisinage spatial ou spatio-temporel. Les différentes méthodes permettant d'estimer un mouvement se classent en trois familles, selon que le mouvement est calculé à partir de dérivées spatio-temporelles de la luminosité, de la transformée de Fourier de la luminosité, ou de régions ou éléments caractéristiques de l'image. Nous ne présentons pas ici des techniques récentes nécessitant un traitement très haut niveau sur l'image, comme celles basées sur la logique floue, généralement peu adaptées aux systèmes embarqués.

#### III-1 Méthodes différentielles

Ces méthodes tendent à minimiser la DFD ("Displaced Frame Difference") exprimée en (3.1). Elles s'appuient sur l'équation du flot optique (3.5), et utilisent donc des dérivées spatio-temporelles de la luminance, ou des images filtrées. Elles ont donc une contrainte et un problème inhérents à cette équation : elles nécessitent une illumination constante, et connaissent le problème d'ouverture. Elles ont cependant l'avantage que l'équation du flot optique est linéaire avec le vecteur vitesse, expliquant le nombre important de travaux portant sur ces méthodes [Bar-94].

La famille des méthodes différentielles se décompose en deux catégories : les méthodes globales et les méthodes locales [Bea-95]. Les méthodes globales utilisent un terme global pour calculer un champ de vecteur vitesse sur l'image entière, alors que les méthodes locales reposent sur l'association de vecteurs vitesse présents dans un voisinage pour estimer le mouvement réel dans ce voisinage. Pour s'affranchir du problème d'ouverture, ces méthodes sont améliorées avec l'adjonction de termes de régularisation à l'équation du flot optique.

- La méthode de régularisation dédiée aux méthodes globales a en premier été proposée par [Hor-81]. Elle consiste à ajouter une contrainte de lissage reposant sur l'hypothèse que des vecteurs vitesse voisins, correspondant à des structures de l'image, doivent être proches. D'autres travaux ont été menés pour compléter ou reformuler ces termes de régularisation (champs de Markov, estimateurs robustes),

afin de prendre en compte les discontinuités du mouvement.

- Afin de régulariser les méthodes locales, Lucas et Kanade [Luc-81] ont proposé un modèle local permettant des calculs dans des fenêtrages. Cette méthode s'appuie sur l'hypothèse que le flot optique est constant localement.

Ces deux exemples sont les plus anciens, et les plus connus, mais d'autres méthodes ont été développées plus récemment, apportant d'autres approches pour lever les indéterminations. [Com-01] en fait l'inventaire. Nous ne parlons pas de ces méthodes car elles offrent des solutions mathématiques peu destinées à l'intégration. Une fois implantés de manière informatique, ces traitements évolués souffrent aussi de temps de calcul très importants. Nous notons enfin que ces méthodes différentielles reposent sur un grand nombre d'hypothèses, et lorsque les images sont par exemple trop fortement bruitées, la précision des mesures est dégradée.

### III-2 Méthodes fréquentielles

Le principe de ces méthodes est de transposer l'image dans le domaine des fréquences. Pour cela, l'image passe dans un banc de filtres spatio-temporels. En considérant  $I(k,\omega)$  la transformée de Fourier de  $I(s,t)$ , un signal se déplaçant peut être caractérisé par :

$$TF(I(s, t - \tau)) = I(k,\omega) \cdot \exp(-j\omega\tau) \quad (3.6.)$$

$k$  est la fréquence spatiale de la position  $s$  (composée de  $x$  et  $y$ ) et  $\omega$  est la fréquence temporelle.  $I(k,\omega)$  est la transformée de Fourier de  $I(s,t)$ , l'exponentielle est le terme de phase. Ces formules expriment qu'une information de luminosité doit se retrouver après un temps de retard. Elles sont en fait la transposition dans le domaine fréquentiel de l'équation (3.1). Elles ne prennent pas en compte les éventuelles dégradations de l'image au cours du temps. Les mouvements sont donc à l'origine supposés constants dans le domaine spatial et dans le domaine temporel. Deux approches existent, suivant que l'information issue des filtres repose sur l'énergie ou la phase.

Les méthodes basées sur l'énergie, comme celle de Adelson et Bergen [Bea-95], utilisent l'amplitude mesurée en sortie des filtres. Cette valeur donne l'orientation spatio-temporelle, ce qui est équivalent au mouvement. L'inconvénient de cette méthode est que l'amplitude est liée

aux contrastes de l'image, et pose des problèmes lorsqu'ils sont faibles.

Les méthodes basées sur la phase, comme par exemple celle de Fleet et Jepson, s'appuient sur le terme de retard de l'équation (3.6). Le mouvement se calcule en mesurant le signe et l'amplitude de la phase. L'avantage de la phase du signal est qu'elle est indépendante de l'amplitude des stimuli en entrée des filtres, elle est donc plus utile.

Comme nous l'avons vu, l'information de mouvement est extraite à l'aide de filtres passe-bande spatio-temporels orientés, dédiés à des gammes de vitesses. Plus le nombre de filtres est important, plus la mesure pourra être précise. Adelson et Bergen utilisent au minimum 6 filtres spatio-temporels. Le temps de calcul lié au nombre de filtres nécessaires est la principale limitation. Ceci entraîne forcément une implémentation non optimale, et le phénomène de lissage entraîne des problèmes sur les mesures de séquences vidéo rapides. Cette approche est généralement basée sur les caractéristiques neuromorphiques de systèmes de vision simples (insectes, primates), elle est dite bio-inspirée.

### III-3 Méthodes de mise en correspondance

Les méthodes de mise en correspondance déterminent le déplacement qui apparie le mieux des éléments d'images de la séquence entre deux instants consécutifs. Ces éléments peuvent être par exemple des régions (pixels, blocs de pixels, formes ...) ou des éléments caractéristiques de l'image (contours, coins ...). L'appariement effectué se base en général sur un paramètre à minimiser ou à maximiser (selon le cas) entre les éléments des différentes images successives. Cette contrainte se présente sous la forme d'une corrélation ou d'une distance entre ces éléments. Classiquement, les calculs s'effectuent sur des fenêtrages de  $N \times N$  pixels déplacés d'une distance  $(u,v)$  entre les deux images. De nombreuses méthodes ont été développées, les plus classiques peuvent se classer en deux familles [Gia-00] : les méthodes de mesure de distance et de mesure de corrélation. Dans les équations ci-dessous,  $\sum_F$  représente la sommation effectuée sur un fenêtrage  $F$ . Le mouvement spatial est considéré ici de la position  $(x,y)$  à la position  $(x' = x + u, y' = y + v)$ .

- Mesure de distance :

Les méthodes de mesure de distance s'appuient sur les mesures SAD, ZSAD, SSD, et ZSSD définies par les formules suivantes.

SAD	<i>somme des valeurs absolues des différences</i>	$SAD = \sum_F   I_1(x, y) - I_2(x', y')  $	(3.7.)
ZSAD	<i>somme des valeurs absolues des différences moyennes</i>	$ZSAD = \sum_F   [ I_1(x, y) - \bar{I}_1 ] - [ I_2(x', y') - \bar{I}_2 ]  $	(3.8.)
SSD	<i>somme des différences au carré</i>	$SSD = \sum_F [ I_1(x, y) - I_2(x', y') ]^2$	(3.9.)
ZSSD	<i>somme des différences moyennes au carré</i>	$ZSSD = \sum_F [ (I_1(x, y) - \bar{I}_1) - (I_2(x', y') - \bar{I}_2) ]^2$	(3.10.)

Les techniques SAD et SSD sont les plus simples et ont donc un coût en calcul moins important. Elles reposent sur la comparaison simple des luminances entre les pixels des deux images, ce qui leur confère une forte sensibilité aux images bruitées. Dans son article comparant des méthodes de corrélation, [Mar-95], J. Martin précise qu'elles sont généralement préférables aux techniques de corrélation standard (face au bruit), notamment la SSD. Les méthodes SAD et SSD sont connues pour leur faible complexité, et sont souvent prises comme références pour illustrer la performance d'algorithmes. La méthode SAD compare les valeurs d'un pixel de l'image 1 originale ( $I_1$ ) avec les pixels de l'image 2 dans un voisinage F. Les techniques ZSAD et ZSSD introduisent un terme de moyennage ( $\bar{I}_1$  et  $\bar{I}_2$ ). Cette régularisation permet une mesure pondérée par rapport à la moyenne de la luminosité de l'image. Ce terme correctif limite les erreurs lors de changements globaux de luminosité sur l'image. Cette amélioration se fait cependant au détriment de la complexité et donc du temps de calcul. Aussi, ce type de mesure de distance perd de son efficacité dans les scènes très texturées, pour lesquelles les écarts entre pixels voisins restent faibles.

- Mesure de corrélation :

NCC	<i>corrélation normalisée</i>	$NCC = \frac{\sum_F [ I_1(x, y) \cdot I_2(x', y') ]^2}{\sqrt{\sum_F I_1^2(x, y) \cdot \sum_F I_2^2(x, y)}} \quad (3.11.)$
ZNCC	<i>corrélation normalisée moyenne</i>	$NCC = \frac{\sum_F [ [ I_1(x, y) - \bar{I}_1 ] [ I_2(x', y') - \bar{I}_2 ] ]^2}{\sqrt{\sum_F [ I_1(x, y) - \bar{I}_1 ] \sum_F [ I_2(x', y') - \bar{I}_2 ]}} \quad (3.12.)$

Il existe un calcul de corrélation simple (appelé "Cross-Correlation" et ayant pour valeur :  $CC = \sum_F I_1(x, y) \cdot I_2(x', y')$ ), mais il est connu pour être trop sensible au bruit et au type d'image, il n'est donc pas utilisé. Les mesures de corrélation, caractérisées par le produit des luminosités, se présentent donc sous deux formes : NCC et ZNCC. Ces calculs normalisés prennent en compte les différences de luminance entre les blocs, permettant une grande robustesse, au prix d'un calcul plus coûteux.

Ces méthodes de mise en correspondance (par distance ou corrélation) n'utilisent pas l'équation du flot optique, le vecteur mouvement (calculé par mesure de distance ou de corrélation) est défini comme la distance entre deux éléments caractérisés par leur positions spatio-temporelles  $(x, y, t)$  et  $(x', y', t')$ , avec  $t' = t + dt$ . Le rapport entre la distance spatiale et le temps  $dt$  donne la vitesse. En considérant une fenêtre temporelle  $dt$ , le vecteur vitesse s'écrit:

$$V_{dt} = (x', y', t') - (x, y, t) \quad (3.13.)$$

La figure II.2 illustre cette équation. Le vecteur  $V_{dt}$  relie l'élément de coordonnées spatiales  $(x = 3, y = 6)$  de l'image à l'instant  $t$  à l'élément  $(x' = 4, y' = 4)$  de l'image à l'instant  $t'$  suivant. Le vecteur vitesse a pour composantes:  $\frac{x'-x}{t'-t}, \frac{y'-y}{t'-t}$ .

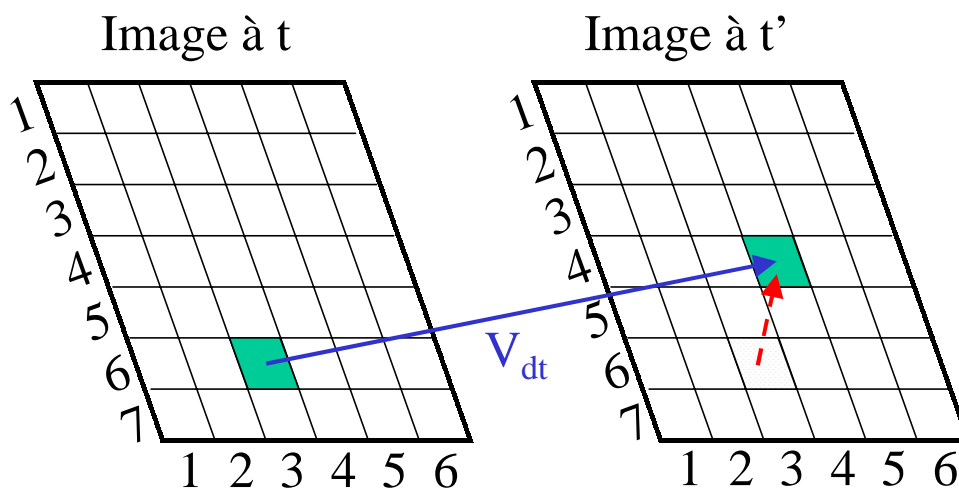


figure II.2 : Déplacement d'un élément à travers deux images successives

Une zone de recherche est parcourue pour retrouver l'élément dans la deuxième image. Pour éviter des temps de recherche prohibitifs, ce fenêtrage doit être le plus petit possible. Cependant, un petit fenêtrage ne mesure que les petits mouvements ; toutefois, ce défaut peut se compenser en effectuant un calcul extrêmement rapide. La fréquence d'échantillonnage est alors augmentée. L'inconvénient de ce type de méthode est la discrétisation de l'image, qui entraîne des imprécisions sur ces algorithmes ; ils ne permettent pas d'avoir une résolution "sub-pixellique".

Certaines de ces méthodes de mise en correspondance ont l'avantage d'être robustes face aux conditions lumineuses, et ne nécessitent pas un grand nombre d'images pour effectuer les calculs. Ces méthodes sont en fait très utilisées, par exemple dans les algorithmes de compression vidéo de type MPEG. Cependant, elles sont généralement coûteuses en temps de calcul et prennent mal en compte les grands déplacements, c'est pourquoi des variantes basées sur une approche hiérarchique ou sur une transformée spatiale sont apparues.

Nous allons détailler la méthode à laquelle nous nous intéressons. Elle appartient à la famille des algorithmes basés sur la mise en correspondance de blocs de pixels comprenant une transformée spatiale de l'image.

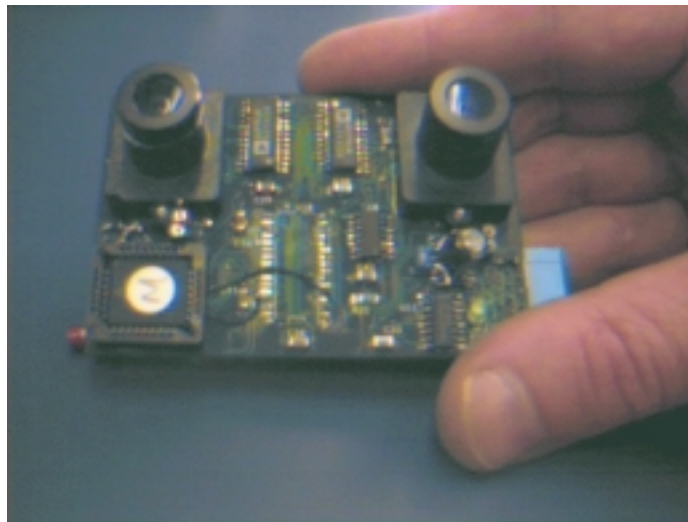


## IV. Présentation de la méthode retenue (LTS: "Local Transform and Search")

### IV-1 Présentation

Les méthodes de mise en correspondance de blocs de pixels basées sur un codage préalable de l'image ont la propriété de ne pas travailler directement sur la luminance des pixels, mais sur une information issue de ces luminances. Certes, si le codage est complexe, ces solutions peuvent être très coûteuses en calcul ; celle que nous présentons ici a l'avantage de concilier robustesse et simplicité de calcul [Ban-97].

Cette méthode est utilisée en robotique pour différentes applications : la stéréovision et l'estimation de mouvements [Zab-94]. La stéréovision est un dispositif composé de deux capteurs d'images distincts, comme le montre la figure II.3 issue des travaux présentés dans [Kon-98a].



**figure II.3 : Système de vision stéréoscopique**

Nous voyons sur ce système que les deux capteurs sont espacés, ils acquièrent donc deux images légèrement différentes. Un de ces deux capteurs observe la scène vue de gauche, l'autre de droite. La combinaison des deux vues donne à l'image finale une information de profondeur. La méthode de mise en correspondance sert précisément à rassembler les deux images, de gauche et de droite, au niveau du pixel. La profondeur ainsi calculée permet une représentation de la scène en trois dimensions ou l'estimation de distances d'objets.

L'estimation de mouvements est basée sur le même calcul inter-images. Cependant, ces images ne sont plus espacées spatialement, mais temporellement. En effet, un seul capteur est nécessaire, il acquiert des images à des instants différents. Le calcul résultant donne le déplacement effectué par chaque pixel entre deux instants d'acquisition d'images : c'est le flot optique. La représentation la plus courante du flot optique est une image composée uniquement de vecteurs montrant le mouvement des pixels.

## **IV-2 Algorithme**

Compte tenu de sa structure, que nous détaillons ici, nous appelons cet algorithme LTS ("Local Transform and Search"). L'algorithme LTS est robuste : il utilise les valeurs de luminosité de plusieurs pixels, ce qui limite les erreurs locales, et les calculs sont basés sur des différences de luminosité, ce qui évite les problèmes lors de changement de luminosité ambiante. Les hypothèses de la luminosité localement ou globalement constantes ne sont pas nécessaires ici. L'algorithme LTS est basé sur une phase de transformation de l'image, la "Census Transform"[Zab-94], qui transpose l'image en une matrice de codes. Une deuxième phase est destinée à rechercher la similitude entre les codes calculés.

### **IV-2-1 Census Transform**

Dans cette partie, nous nous efforçons d'expliquer la transformée telle qu'elle est implantée de manière informatique ou numérique. Dans une première phase, chaque pixel est comparé à son voisinage, cette comparaison donne une valeur binaire.

Après l'acquisition d'une image, l'illumination de chaque pixel est caractérisée par une valeur entière : son niveau de gris. Classiquement pour les traitements d'images de ce type, les niveaux sont codés sur 8 bits, représentant 256 valeurs différentes. Elles s'étendent donc de 0 (pour le noir) à 255 (pour le blanc). La figure II.4 ci dessous représente les niveaux de gris d'une image.

	0	1	2	3
0	0	114	14	90
1	226	90	119	169
2	36	169	252	14
3	90	0	226	119

**figure II.4 : Valeurs de la luminosité des pixels d'une image**

Ensuite, la luminosité de chaque pixel de la matrice est comparée aux luminosités de ses 8 voisins. Pour illustrer ces comparaisons, nous allons nous focaliser sur le pixel de coordonnées (1,1) – de valeur 90 - et de son voisinage 3x3 (figure II.4). Il va être codé par rapport aux autres pixels composant son voisinage. Une simple comparaison de la valeur des niveaux de gris est faite. Si un pixel voisin est supérieur au pixel central, il est codé avec un '1', sinon, il est codé par un '0'. Sur l'exemple de la figure II.4, la comparaison donne donc :

1	0	1
0	X	0
1	0	0

**figure II.5 : Codes liés aux différences des luminosités des pixels**

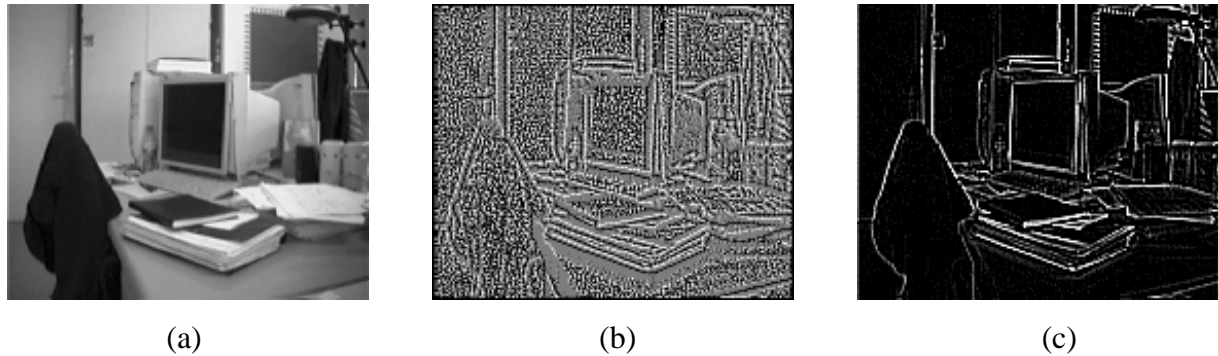
Ces comparaisons représentent en fait la direction – ou le signe - du gradient entre le pixel central et les pixels voisins (dans la fenêtre considérée).

Après ces comparaisons, il est caractérisé par la valeur binaire donnée par ce calcul. En notant arbitrairement de gauche à droite et de bas en haut, la valeur est 10100100 en binaire. Chaque pixel de la matrice aura un code semblable (sur 8 bits), résultant du même calcul avec son voisinage propre. Nous pouvons souligner que ces comparaisons donnent à la méthode LTS sa très grande robustesse quant aux variations de luminosité. Certes, l'aspect itératif est à considérer, même s'il reste linéaire par rapport à la taille de la matrice. Pour une matrice de M x N pixels, le nombre de comparaisons est de 8.M.N.

C'est de ces comparaisons successives que vient le nom de cette transformée "Census

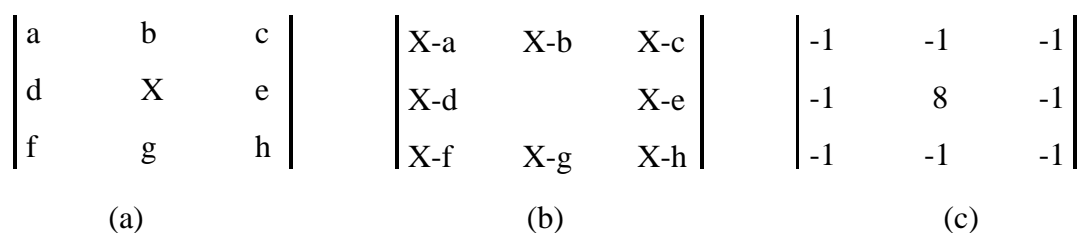
Transform" : on effectue en effet un recensement des pixels par rapport à leur valeur, qui est supérieure ou inférieure à la valeur du pixel central.

Cette transformée spatiale de l'image produit un codage local de la texture, qui est illustré figure II.6.



**figure II.6 : Image 'I' en niveaux de gris (a), image de la transformée de 'I' (b) – le niveau de gris représente le nombre de '1' dans les codes - , 'I' filtrée par un passe-haut classique**

La transformée spatiale de l'image a des similitudes avec une détection de contours. Ceci est dû à la phase de codage composée de comparaisons inter-pixels. En considérant la figure II.7, le voisinage 3x3 du pixel X (a) effectuée au cours de la transformation une série de comparaisons (b), qui a pour but de lui donner une information du type  $I_n = X-a + X-b + X-c + X-d + X-e + X-f + X-g + X-h$  qui est équivalent à  $I_n' = 8X - a - b - c - d - e - f - g - h$  donnée par un filtre passe-haut classique - figure II.6 (c) et figure II.7 (c).



**figure II.7 : Voisinage du pixel X (a), étape de la transformée (b), filtre passe-haut (c)**

L'image de la transformée est issue des codes de chaque pixel. Pour chaque pixel, le nombre N de bits à '1' dans le code définit le niveau de gris pour représenter l'image (b). Les niveaux de gris I sont donc donnés par :  $I = N \times 31$ . La plage des niveaux s'étend donc de 0 à 248 (8 x 31). Un pixel ayant une luminosité supérieure à celle de ses voisins est donc clair.

Après ce filtrage (b), on ne conserve que le signe des comparaisons (voir figure II.5). Le seuil qui différencie les niveaux est ainsi changé, puisqu'il ne reste qu'une information binaire

indiquant si le pixel voisin est éclairé ou moins éclairé que le pixel central. Cette information binaire confère plus de robustesse au codage. En effet, pour chaque pixel, il existe une plage de variation qui ne change pas le code.

Pour une application de stéréovision, le même processus est utilisé pour transformer les images issues des deux capteurs. Si nous souhaitons réaliser une détection de mouvement, ces images sont acquises puis codées de manière séquentielle. Ces valeurs sont alors mémorisées. Une phase de recherche de similarité entre les codes de ces deux images suit l'étape de codage.

#### IV-2-2 Recherche des codes de la transformée

Cette deuxième phase consiste à faire la correspondance entre les éléments des deux images stockées. Le principe est qu'un code d'un pixel de l'image  $t$  peut être associé à un code de l'image  $t'$ . Pour associer les pixels, la distance minimale entre les codes donne la probabilité la plus grande du mouvement réel.

Une fenêtre de recherche est sélectionnée dans l'image  $t'$ , et on calcule la différence booléenne entre le pixel de référence (au temps  $t$ ) et tous les pixels de la fenêtre (au temps  $t'$ ). Cette étape est schématisée par la figure II.8. Le pixel est successivement comparé à tous les pixels de la fenêtre, la distance la plus faible détermine le vecteur mouvement  $V_{dt}$ . Si le temps d'acquisition entre les deux images est assez bref par rapport au mouvement observé, le pixel de l'image à l'instant  $t_1$  pourra être retrouvé dans un voisinage correspondant à l'instant  $t_2$ .

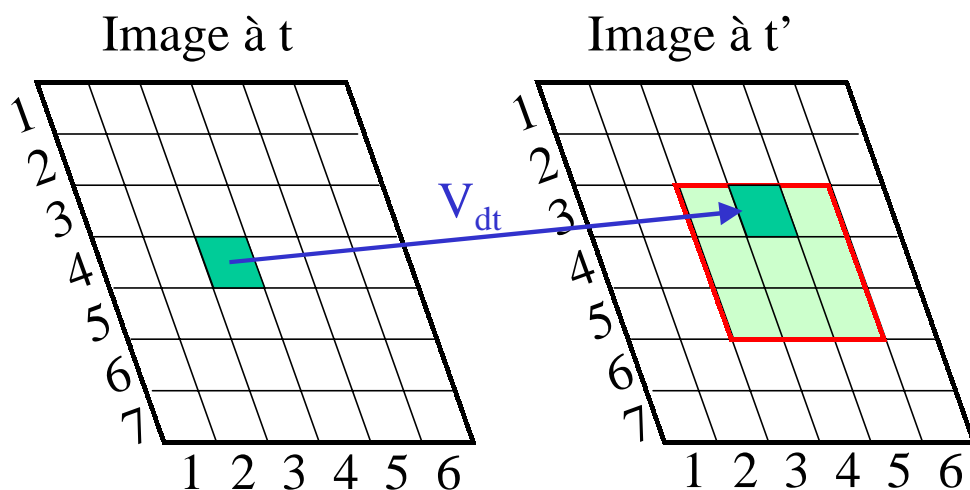
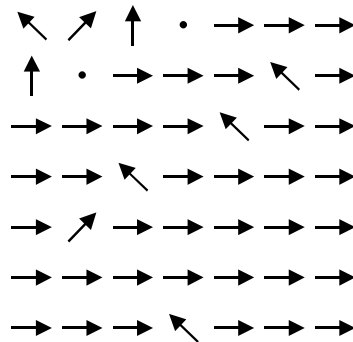


figure II.8 : Recherche d'un code de l'image à  $t$  dans un fenêtrage de l'image à  $t'$ .

C'est la distance de Hamming la plus faible qui indique la nouvelle position du pixel de référence. Classiquement, cette distance, qui consiste en la différence bit à bit des codes, donne le nombre de positions binaires de valeurs différentes séparant deux mots d'un code donné. Si nous comparons deux mots A et B de N bits, le résultat est égal au nombre de bits différents (parmi les N) dans les mots A et B.

Cette recherche de distance de Hamming minimale est effectuée pour tous les pixels de la matrice, ce qui donne un champ de vecteur dense (flot optique). La figure II.9 donne un exemple de champ de vecteur dense. Les points représentent des mouvements nuls.



**figure II.9 : Champ de vecteurs d'une séquence d'images.**

Nous associons à chaque pixel un vecteur indiquant son déplacement. Le mouvement illustré est un mouvement de translation vers la droite. Quelques erreurs sont présentes, nous donnons une étude plus approfondie de la méthode dans le chapitre IV.

### IV-2-3 Complexité algorithmique

Nous avons vu que la complexité du codage (dans le voisinage 3x3) est de 8.MN. Pour une recherche dans un voisinage de 3x3 pixels, 9 comparaisons logiques (sur 8 bits) par pixel sont nécessaires, ce qui donne 9.M.N comparaisons, suivies de la recherche du minimum. Pour un voisinage 5x5, la complexité devient 25.M.N. Ces deux voisinages sont les plus utilisés pour effectuer la recherche.

La complexité de l'algorithme croît linéairement avec le nombre de pixels de la matrice. Il nécessite, pour un voisinage 3x3, 33.M.N calculs (composés d'opérations binaires et de nombreuses conditions), et nécessite autant d'accès mémoire. Pour une image 256x256, chaque champ de vecteur calculé représente 2,1 millions d'opérations, en considérant un flot de données à la cadence vidéo de 25 images par secondes, cela représente 54 Mops (millions d'opérations par seconde), et pour 100 images par seconde, 216 Mops. Nous voyons ici les problèmes afférents aux applications de vidéo rapide (suivi de soudage laser, de projectiles

...), où les cadences de données à traiter deviennent importantes. Un flot de données (et de calcul) conséquent entraîne une puissance consommée à prendre en compte, particulièrement dans un contexte de système embarqué où la consommation est une contrainte forte.

### IV-3 Caractéristiques de l'algorithme LTS

Cet algorithme a été présenté par [Zab-94] comme une méthode robuste dans les cas où l'environnement est bruité. Comme cet algorithme est constitué d'une transformée codant l'image suivi par une recherche de ces codes, nous la nommons LTS ("Local Transform and Search"). Dans cet article, l'auteur compare cette technique à des méthodes standard de mesure de distance (SSD) et de corrélation (NCC) pour un calcul dans un environnement bruité. Cette comparaison est intéressante dans la mesure où elle compare l'efficacité des techniques mettant en jeu des calculs d'ordre assez simples. L'axe des ordonnées est relatif, il indique la pertinence des résultats – les vecteurs mouvement sont justes -, et l'axe  $n$  représente la fenêtre de calcul considérée – de  $n$  pixels de côté. Il en ressort que le résultat des techniques proposées dans l'article, notées 'Param', est très intéressant même pour un voisinage  $n$  petit ( $n = 3$ ). Cela permet de concilier rapidité et efficacité.

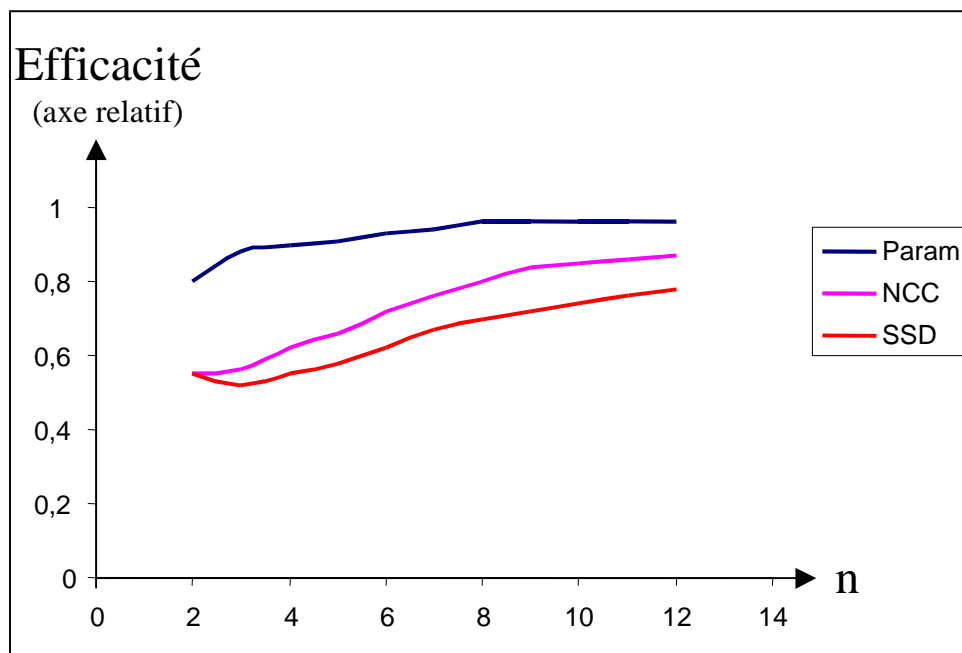


figure II.10 : Comparaison de techniques de mises en correspondance.  $n$  indique la taille de la fenêtre de calcul.

D'autres auteurs, [Ban-97], [Com-01], [Sme-02] et [Kon-98a], ont étudié cet algorithme. [Woo-97] l'a implanté dans un système à base de FPGA. Ces travaux rapportent que l'algorithme est efficace. Une étude récente [Sme-02] utilise un voisinage 3x3 pour calculer la transformée de Census, qu'il compare à sa nouvelle méthode basée sur les ondelettes, nommée "Ranklets". La tableau ci-dessous est extrait de sa publication et montre les différences entre ces méthodes.

Méthode	Erreur
Ranklets	16,7 %
Census	20,3 %

**figure II.11 : Comparaison des techniques Ranklet et Census [Sme-02].**

Cette étude montre que la méthode LTS à laquelle nous nous intéressons est toujours compétitive, malgré la simplicité relative de ses calculs. Certes, d'autres algorithmes [Com-01] offrent de meilleurs résultats, mais les temps de calculs et les ressources matérielles mis en jeu ne sont pas réalistes.

La méthode LTS a donc de bons résultats, mais nécessite un grand nombre d'itérations. Compte tenu de ses traitements locaux, sur des voisinages 3x3 et 5x5, et de la nature booléenne de ses calculs, cette méthode peut bien s'intégrer dans un système de calcul matériel. De plus, il est possible de paralléliser les calculs dans une structure de calcul matricielle ou linéaire. Nous comptons donc étudier des structures de calcul électronique (numériques ou analogiques) afin d'implanter cet algorithme.

## V. Conclusion

L'étude des méthodes d'estimation du mouvement est un domaine vaste, puisque ces méthodes sont étudiées en Informatique, en Robotique, et en Microélectronique. Il en ressort que les méthodes les plus performantes ont des temps de calcul prohibitifs, et qu'un compromis vitesse / efficacité s'impose. Dans cette optique, nous nous sommes intéressés à une famille d'algorithmes (mise en correspondance de blocs) qui est très utilisée en informatique, car performantes, mais qui nécessite des calculs complexes et donc longs. De ce fait, peu de rétines CMOS intègrent un tel calcul, particulièrement sur des images réelles. Nous avons cependant étudié un algorithme de cette famille qui a une approche bas niveau dans l'image, permettant une implantation matérielle prometteuse.



Nous présentons dans le chapitre IV nos implantations informatiques de différentes méthodes d'estimation du mouvement. Nous y détaillons donc une étude plus concrète de l'algorithme LTS.







# Chapitre IV :

## VALIDATION DE L'ALGORITHME LTS

---

Dans ce chapitre, nous détaillons les validations que nous avons effectuées. Ces validations portent sur la programmation de l'algorithme d'estimation du mouvement LTS, et sur son implantation matérielle standard.

### I. Evaluation de méthodes d'estimation du mouvement

Cette partie présente les validations que nous avons menées à travers des simulations. Ces simulations ont pour but de valider la robustesse de l'algorithme d'estimation du mouvement présenté au chapitre II. Nous avons effectué les premières simulations avec MATLAB, puis, compte tenu du temps d'exécution important avec ce langage interprété, nous avons développé des programmes en langage C. Ces programmes effectuent un calcul d'estimation du mouvement par corrélation SAD et par l'algorithme LTS [Zab-94] que nous avons étudié.

#### I-1 Programmation de la SAD

La SAD, somme des valeurs absolues des différences, présentée dans le chapitre III, a comme formule :  $SAD = \sum_F | I_1(x, y) - I_2(x', y') |$ . Elle représente donc la distance à minimiser entre les niveaux de gris des pixels de l'image de référence et les niveaux de gris des pixels de la seconde image considérée. Lorsque cette distance est nulle, le mouvement d'un pixel d'une image à l'autre est trouvé. En terme de complexité calculatoire, ce calcul est le plus simple à réaliser. Cependant, même s'il n'est pas réellement utilisé dans des applications, il fait office de référence pour comparer différentes méthodes.

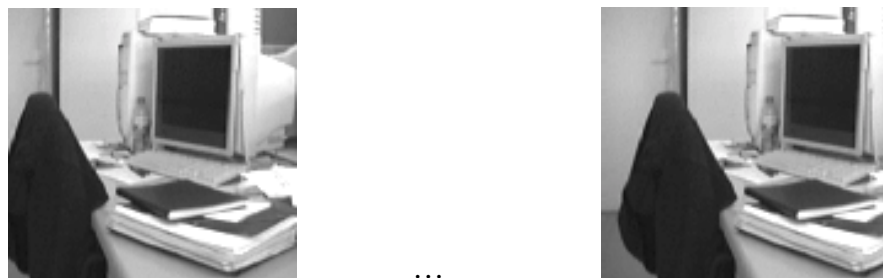
L'implantation de cet algorithme est donc simple, et se limite à effectuer, sur un fenêtrage considéré, une différence de deux entiers combinée à la mémorisation du meilleur résultat à chaque itération. La complexité de calcul est proportionnelle à la résolution des images et à la taille de la fenêtre considérée.

La structure du programme est illustrée figure IV.1. La SAD est le minimum de la différence (en valeur absolue) du niveau de gris du pixel de l'image 1 ( $I_1$ ) et du niveau des pixels  $I_2$  d'un voisinage (u,v) autour de la position de I1 dans l'image 2.

```
For l = 1 to N
  For c = 1 to M
    For {voisinage [u,v]}
      SAD(l,c) = min [ Abs (  $I_1(l,c) - I_2(l,c)[u,v]$  ) ]
    end For
  end For
end For
```

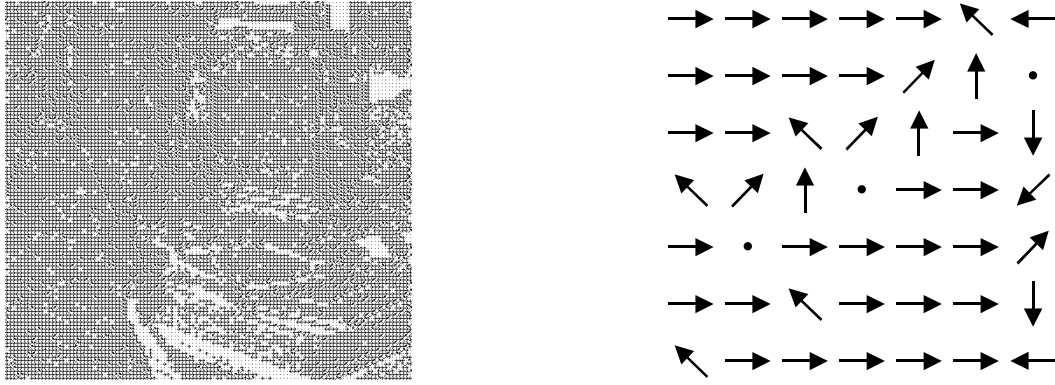
**figure IV.1 : Algorithme du calcul de la SAD**

Pour la séquence d'images de résolution 256x256, représentée par la figure IV.2, et une recherche dans un voisinage 3x3, le temps de calcul est typiquement de l'ordre de 0,3 seconde sur une station SPARC Ultra10 dotée d'un processeur cadencé à 400 MHz.



**figure IV.2 : Images 256x256 utilisées pour l'estimation du mouvement**

La figure IV.3 illustre la représentation du mouvement par un champ de vecteurs dense pour un voisinage de recherche de 3x3 pixels.



**figure IV.3 : Champ de vecteurs dense illustrant le mouvement calculé par SAD (sur l'image entière et en détail sur une zone de 7x7 pixels de cette image)**

Pour des raisons de lisibilité, la partie de droite représente un zoom d'une partie du champ de vecteurs complet. Le taux d'erreur, que nous définissons par le rapport entre le nombre de vecteurs représentant un mouvement faux et le nombre de vecteurs total vaut dans cet exemple:  $T = 40 \%$ .

Nous avons ensuite développé le programme utilisant la méthode LTS que nous étudions.

## I-2 Programmation de la méthode retenue

L'algorithme LTS que nous avons retenu est composé de deux phases : le codage des pixels, puis la recherche de la distance minimale entre ces codes. En langage C, plusieurs fonctions sont nécessaires afin d'effectuer ces tâches, et ce programme s'avère très itératif, donc long. La structure du programme est donnée dans la figure IV.4. L'image considérée est de résolution  $N \times M$  pixels, et les calculs du codage de la texture et les recherches des codes se font respectivement sur les voisinages  $[u, v]$  et  $[u', v']$ . Le résultat est, pour chaque pixel, le mouvement  $Mvt[u_i, v_c]$  tel que la distance de Hamming entre les codes est minimale pour cette position  $[u_i, v_c]$  du voisinage  $[u', v']$ .

```

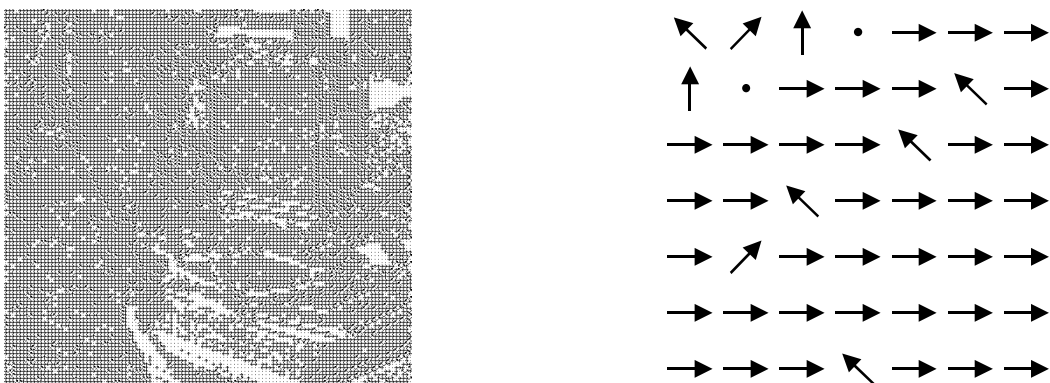
For l = 1 to N
  For c = 1 to M
    For {voisinage [u,v]}
      codes1 = census( I1 (l,c) [u,v] )
      codes2 = census( I2 (l,c) [u,v] )
    end For
  end For
end For

For l = 1 to N
  For c = 1 to M
    For {voisinage [u',v']}
      Mvt (l,c) = [ul,vl] ∇ min [ codes1 (l,c) HAM codes2 (l,c)[u',v'] ]
    end For
  end For
end For

```

**figure IV.4 : Algorithme de calcul de la méthode LTS**

Pour effectuer l'estimation de mouvement sur une séquence d'images de résolution 256x256 pixels, avec une recherche dans un voisinage 3x3, le temps d'exécution du programme est typiquement de l'ordre de 2 secondes, le taux d'erreur reste très dépendant des images.

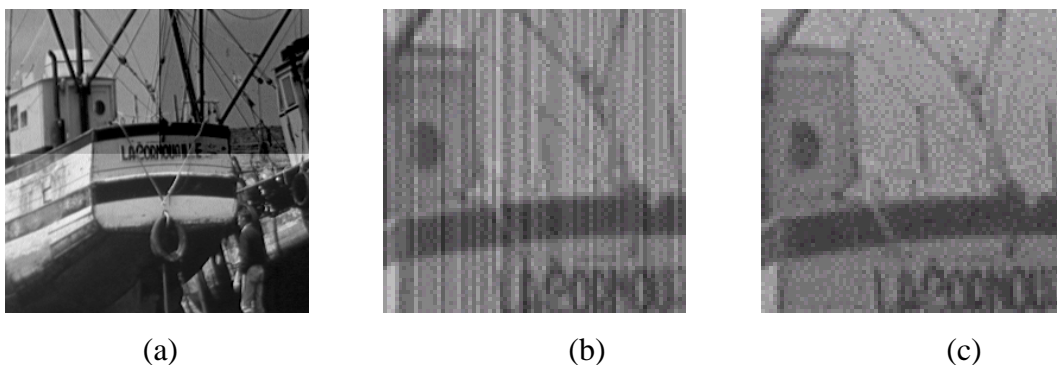


**figure IV.5 : Champ de vecteurs illustrant le mouvement calculé par la méthode LTS**



La figure IV.5 illustre la même représentation du mouvement que la figure IV.3. Le mouvement original de la caméra va de gauche à droite. Nous pouvons constater sur cet exemple que les vecteurs en direction de la droite sont plus nombreux que dans la figure IV.5. Une comparaison plus précise des méthodes SAD et LTS est présentée plus loin.

Nous avons aussi testé l'algorithme LTS par rapport aux changements de luminosité et par rapport aux bruits de l'image induits par les défauts des capteurs. Les types d'images utilisés sont représentés dans la figure IV.6. Il s'agit pour la séquence (a) de changements de luminosité, représentant d'éventuels mouvements d'ombre dans une scène. La même image non bruitée sert d'origine pour calculer le mouvement. La séquence (b) – composée de deux images - a un bruit de type "colonne", typiquement rencontré dans les imageurs à pixel actif 3 transistors sans système de réduction du bruit (CDS); et la séquence (c) a un bruit de type pixel, typiquement rencontré dans les structures d'imageurs à amplification "in-pixel". Pour ces deux bruits, l'image semble être perçue à travers des barreaux ou un grillage fixes. Dans ces séquences (b) et (c), les deux images de la séquence présentent donc ce bruit, ce qui rend plus complexe l'appariement des pixels codés.



**figure IV.6 : Extraits des séquences bruitées utilisées pour caractériser l'algorithme**

Les résultats sont résumés dans la figure IV.7. Il apparaît que le codage de la texture de l'image rend l'algorithme peu sensible aux variations de luminosités ("Var Lum") partielles ou globales. Les erreurs supplémentaires relevées correspondent en fait aux frontières des régions dont la luminosité change. Cependant, nous notons qu'en présence de bruit correspondant aux défauts de capteur d'images standards, le nombre de vecteurs erronés est important. Une telle erreur est caractérisée par un offset aléatoire présent sur tous les pixels des deux images de la séquence. Ces variations changent les codes de la transformée, qui ne sont plus retrouvés efficacement lors de la recherche.

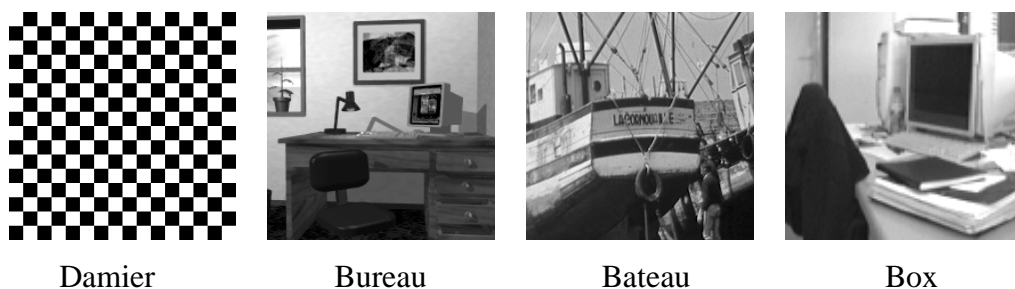
Bruit	Aucun	Var Lum	Colonne	Pixel
Taux vecteurs justes	86 %	80 %	45 %	46 %

**figure IV.7 : Caractérisation de l'algorithme avec les images bruitées**

Une rétine ne possède généralement pas de technique de réduction du bruit. Pour éviter de telles erreurs, elle doit donc avoir une structure adaptée, évitant le plus possible les sources de bruit.

### **I-3 Comparaison entre les algorithmes SAD et LTS de manière logicielle**

Pour comparer efficacement les deux méthodes, nous avons programmé ces dernières pour calculer l'estimation du mouvement sur plusieurs séquences d'images. La figure IV.8 présente ces séquences.



**figure IV.8 : Présentation des séquences d'images**

Les séquences "Damier", "Bureau", et "Bateau" connaissent un mouvement de translation, la séquence "Box" subit une rotation de la caméra.

Le tableau de la figure IV.9 résume les résultats obtenus, en indiquant, pour chaque séquence, la résolution des images, le temps de calcul nécessaire et le taux d'erreur relevé. Ce taux d'erreur est égal au rapport entre le nombre de vecteurs faux et le nombre total de vecteurs.

	SAD	LTS	SAD	LTS	SAD	LTS	SAD	LTS
Séquence	Damier		Bureau		Bateau		Box	
Résolution	256x256		200x200		256x256		118x110	
Tps calcul	0,3 s	2,1 s	0,2 s	1,3 s	0,3 s	2,1 s	0,06 s	0,5 s
Tx erreur	94 %	78 %	72 %	45 %	55 %	14 %	40 %	30 %

**figure IV.9 : Tableau comparatif de la SAD et de la méthode LTS**

Ces séquences présentent des caractéristiques propres à tester efficacement les algorithmes. La séquence "Damier" a la particularité de présenter beaucoup de zones homogènes, dans lesquelles le mouvement n'est pas perçu. En effet, ce dernier n'est perçu que sur les transitions noir / blanc. Le codage de la texture à ces endroits permet à la méthode LTS de présenter un meilleur résultat. La séquence "Bureau" est une image de synthèse présentant aussi des zones homogènes. Dans ces séquences, les résultats justes sont faibles, il faut considérer que les algorithmes donnent beaucoup de vecteurs nuls dans ces zones homogènes, mais ont un bon rendement au niveau des hautes fréquences de l'image. Les séquences "Bateau" et "Box" présentent des scènes plus réelles dans lesquelles les résultats sont meilleurs. La scène "Box" offre des performances légèrement moins bonnes, sûrement à cause de sa particularité : elle provient d'un capteur CMOS grand public couleur (WebCam).

Il apparaît que la méthode LTS est effectivement plus performante que le calcul de distance simple. Nous notons que même la SAD présente un temps de calcul important, de 0,3 seconde pour des images de résolution 256x256. Cette méthode ne permet pas de calculer un champ de vecteurs à la cadence vidéo. La méthode LTS souffre elle aussi d'un temps de calcul élevé. La solution logicielle n'est donc pas une solution acceptable pour la plupart des applications. De plus, la phase d'acquisition des images dans un système informatique – non prise en compte dans ces simulations - est un facteur très limitant, puisque généralement long : il consiste en une acquisition par un port de communication et en un rangement des données en mémoire. Enfin, cette solution informatique constitue un système qui a de grosses contraintes de géométrie et trouve difficilement sa place dans un système embarqué. La solution matérielle s'impose alors.

## I-4 Modification de l'algorithme LTS en vue de son implantation matérielle

En étudiant l'implantation matérielle de l'algorithme LTS, nous avons remarqué que la recherche d'une valeur scalaire minimum parmi N possibilités nécessite une surface importante. En effet, dans une architecture optimisée en temps de calcul, les soustractions et comparaisons parallèles, puis le stockage des valeurs, nécessitent un nombre de portes logiques élevé. Nous avons constaté que la phase de recherche des codes n'est pas optimisée. Nous avons donc modifié l'algorithme original, afin d'obtenir une implantation plus compacte, sans altérer de manière significative son efficacité.

Lors de la recherche des codes (la deuxième phase de l'algorithme), l'algorithme original recherche la distance de Hamming minimale. En fait, un pixel codé dans la première image doit se retrouver dans la deuxième; en théorie, ces codes sont donc identiques (leur distance est alors nulle). Nous nous sommes cependant rendus compte que la quasi-totalité des vecteurs calculés provient d'une distance de Hamming nulle entre les codes. Même des vecteurs erronés, ne représentant pas le mouvement réel, ont souvent une distance nulle entre leurs codes. Dans ce cas, l'erreur provient en fait de la phase précédente, lors du codage de l'image.

Nous avons aussi observé que le fait de déterminer un vecteur mouvement à partir du code le plus proche, c'est à dire de l'erreur minimale, entraîne très souvent une mauvaise estimation. En effet, la quasi-totalité des vecteurs provenant d'une distance de Hamming non nulle indique un mouvement faux. En conséquence, nous avons modifié l'algorithme pour qu'il effectue une recherche de distance strictement nulle. Nous avons comparé les efficacités des méthodes originale et modifiée, les résultats sont résumés dans la figure IV.10.

Séquence	Méthode LTS originale		Méthode LTS modifiée	
	Nb vecteurs justes	%	Nb vecteurs justes	%
Damier	50550	22 %	50548	21,9 %
Bureau	21715	55 %	21715	55 %
Bateau	52150	86 %	52478	87 %
Box	8554	70%	8554	70 %

**figure IV.10 : Tableau comparatif des méthodes LTS originale et modifiée**

Pour les séquences précédentes, le mouvement est calculé avec les deux versions de la méthode LTS. Les pourcentages de vecteurs justes n'étant pas assez significatifs, nous donnons aussi le nombre de ces vecteurs. Le tableau indique bien que la modification de la recherche n'influence quasiment pas les résultats.

Cette amélioration permet cependant une implantation matérielle plus optimisée, nous allons l'étudier en considérant deux implantations : dans un FPGA et dans un ASIC.

## **II. Implantation matérielle de la méthode retenue**

Afin de palier aux problèmes de temps de transit et de traitement des informations, l'implantation matérielle est – pour l'instant - la seule alternative pour estimer le mouvement en des temps courts. En effet, un système sur carte composé du capteur et d'une unité de calcul constitue une solution bien plus intéressante qu'une implantation logicielle du point de vue du temps de calcul. Aussi, ce système sur carte comprend quelques composants et ne nécessite pas un système informatique évolué et encombrant.

### **II-1 Implantation dans un FPGA**

Pour l'implantation de l'algorithme dans un FPGA, nous avons retranscrit ce dernier en langage VHDL, afin d'avoir une implantation présentant de bonnes performances. Ce programme en VHDL a la même fonctionnalité que le programme en langage C (figure IV.4), il permet cependant de paralléliser des calculs. Pour effectuer l'estimation de mouvement, nous avons programmé un FPGA, grâce à la plaquette Excalibur d'Altera. Ce FPGA APEX 20K200, de 200.000 portes équivalentes, réalisé en technologie 0.18  $\mu\text{m}$ , est cadencé à une fréquence d'horloge de 33 MHz. Sa taille n'a cependant pas permis d'implanter une version de l'algorithme optimisée en vitesse, car la surface est limitée. Ce sont en fait les éléments mémoires qui occupent trop de place. L'image considérée ici est de résolution 100x100 pixels. Nous avons estimé que la version optimisée en vitesse nécessitait cependant 30302 cycles, soit un temps de calcul de 6ms. La version optimisée en surface occupe 30% du FPGA. Cependant, le calcul complet de l'algorithme, pour deux images de 100x100 pixels, requiert 4320050 cycles, soit 131ms à une fréquence de 33 MHz. Le site d'Altera ([www.altera.com](http://www.altera.com)) permet de faire une estimation de la consommation du FPGA, pour une configuration donnée. Pour la configuration correspondant au code VHDL de la méthode LTS, la consommation de l'APEX est de 141mW. En considérant un système sur carte, la consommation moyenne d'un

imageur standard (par exemple le VV5301 de ST-Microelectronics) est de 75mW. Enfin, un circuit mémoire, comme le M68AW031 (256 Kbits) de ST-Microelectronics, consomme en moyenne 20mW. Le système complet (capteur, FPGA, mémoire) consomme donc 226mW et se limite à une cadence de 7 images/seconde. Notons qu'une cadence proche de la cadence vidéo pourrait être atteinte en implantant cet algorithme en faisant un compromis vitesse/surface.

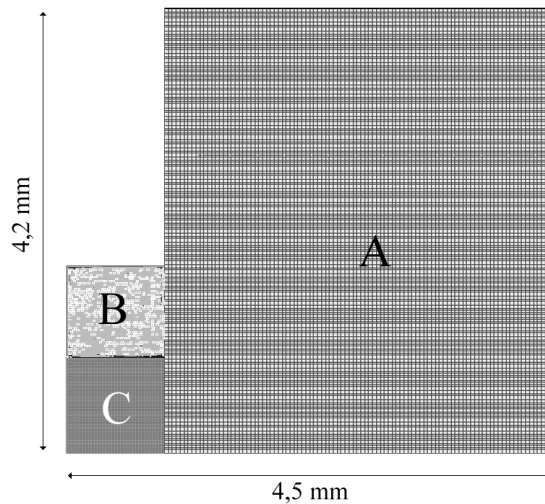
Le tableau ci-dessous résume les principales caractéristiques de cette implantation.

Implantation de l'algorithme LTS				Estimation du système sur carte				
Optim.	Surface	Cycles	Temps		FPGA	Capteur	Mémoire	<b>Système</b>
Vitesse	> 100 %	30302	6 ms	Conso.	141 mW	75 mW	20 mW	<b>236 mW</b>
Surface	30 %	4,3 M	131 ms	Cadence	7 im/s	25 im/s	1400 im/s	<b>7 im/s</b>

**figure IV.11 : Tableau récapitulatif de l'implantation avec un FPGA APEX 20K200**

## II-2 Implantation dans un ASIC – capteur et traitement

Nous avons aussi synthétisé, à partir d'une description VHDL, le bloc de calcul complet pour effectuer l'estimation de mouvement. En considérant une solution matérielle numérique, dans un Système sur Puce (SoC), ce bloc prend place aux côtés de la matrice de pixels. Une structure d'imageur standard se connecte, dans un même circuit, à un élément de calcul et à de la mémoire. Le problème principal de cette solution réside dans la mémorisation des valeurs des pixels. En effet, il est coûteux (en terme de surface), d'intégrer un convertisseur analogique / numérique et deux mémoires de 8 bits dans chaque pixel. [Kle-01] propose un capteur intéressant doté d'une mémoire 8 bits au sein des pixels. Pour l'estimation de mouvement, cette mémoire doit cependant être doublée (deux images sont nécessaires pour le calcul). D'un autre côté, la mémorisation de valeurs analogiques n'est pas une solution optimale, puisque les capacités de stockage sont toujours soumises à des fuites. La solution serait donc de mémoriser les valeurs à l'extérieur de la matrice, dans un plan mémoire dédié (figure IV.12).



**figure IV.12 : Structure d'imageur APS (C), une mémoire (A) et l'élément de calcul (B)**

Dans cette figure, la matrice photosensible (c) a une résolution de 100x100 pixels, soit une taille de 0,81 mm<sup>2</sup>. Les pixels de cette matrice sont des pixels actifs 3 transistors standards, mesurant 9µm de côté en technologie AMS 0.35µm. La zone B est un bloc numérique effectuant le calcul complet de l'algorithme LTS. Il intègre la version VHDL optimisée pour la vitesse de traitement (6 ms), et synthétisée avec AMBIT de Cadence. Cet élément logique comprend 4400 portes logiques, et occupe 0,8 mm<sup>2</sup>. Le bloc le plus important (a) est une mémoire statique standard de 160Kbits, pouvant stocker 2 images de 100x100 pixels en 8 bits (256 niveaux de gris). Ce système sur puce occupe une surface totale d'environ 18 mm<sup>2</sup>. Si nous comparons ce système complet à une rétine effectuant un traitement parallèle, la taille ramenée de chaque pixel de cette rétine serait de 43µm x 43µm, ce qui est une taille raisonnable pour un pixel de rétine comprenant du traitement. Bien entendu, la taille de ces pixels dépend beaucoup du traitement d'images considéré, mais la faisabilité est chiffrée. Notons cependant que l'utilisation de mémoires d'autres types, DRAM ou MRAM (magnétiques), pourraient changer les surfaces indiquées. Leur conception est cependant plus complexe, mais le plan mémoire pourrait représenter alors 3 mm<sup>2</sup>.

Le système proposé en figure IV.12 ne peut pas acquérir des images permettant de calculer correctement les mouvements. En effet, la fréquence de lecture moyenne des pixels dans un imageur standard est de 15 MHz, pour fournir un débit de 25 images/s. Ce mode de fonctionnement est basé sur un système de balayage des lignes ("Partial Scan" ou "Interlaced Scan"), qui consiste à acquérir l'image ligne par ligne en les transférant successivement vers les éléments de mémorisation. Ce mode a l'avantage de ne nécessiter qu'une bande passante limitée et de ne pas générer de dérive de l'image. En effet, si on acquérait une image simultanément, il faudrait lire très rapidement tous les pixels pour ne pas causer une distorsion

de l'image.

Cependant, l'acquisition de telles images n'est pas utilisable pour calculer une estimation de mouvements, car elle découpe l'image en une série de lignes. La scène peut changer pendant un tel mode d'acquisition, et chaque ligne peut alors échantillonner le mouvement. Ainsi, un trait vertical se déplaçant horizontalement pourrait apparaître en oblique dans l'image. Il faut donc un système qui puisse acquérir l'image entière instantanément, mais des problèmes de stockage de données ou de vitesse de lecture des pixels apparaissent.

Il faut alors se tourner vers d'autres solutions. Celle que nous retenons est la conception d'une rétine CMOS basée sur la structure de l'imageur présenté dans le chapitre III. Cette dernière doit en effet stocker tous les pixels de l'image simultanément et effectuer un calcul parallèle ou matriciel pour traiter l'information rapidement.







# Chapitre V :

## RETINE D'ESTIMATION DU MOUVEMENT

Dans ce chapitre, nous présentons l'implantation matérielle de la méthode LTS dans une rétine CMOS : circuit spécifique (ASIC) qui capte les informations lumineuses, les convertit en information électrique, et effectue un traitement de l'image. Nous détaillons l'architecture globale de cette rétine, les différents blocs qui la composent, puis les mesures que nous avons faites. Enfin, nous dégagons des perspectives par rapport à ce travail et plus généralement par rapport aux rétines destinées à l'estimation du mouvement.

### I. Architecture globale du circuit

L'architecture globale du circuit comprend trois blocs : la zone matricielle de photoréception et de conversion analogique - numérique, la zone de calcul numérique, et un séquenceur. Le traitement d'images, à savoir l'estimation de mouvement, se fait dans les deux premier blocs, le troisième gère le déroulement et l'enchaînement des calculs. La figure V.1 illustre la disposition ("floorplanning") du circuit.

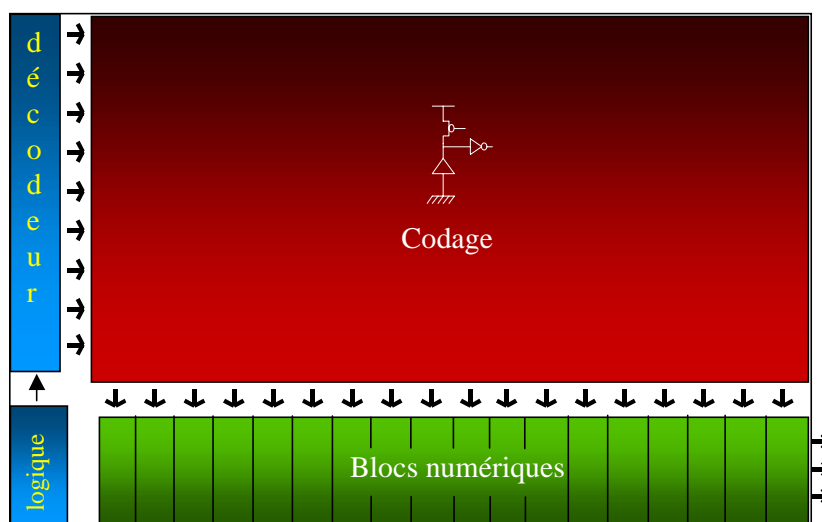


figure V.1 : Disposition des éléments dans le circuit

L'algorithme LTS, présenté dans le chapitre II, se déroule en deux étapes : un codage

des pixels, puis une recherche de ces codes. Les deux blocs présentés effectuent chacun une de ces deux étapes. La partie marquée "Codage" dans la figure ci-dessus est composée des éléments photosensibles, les photodiodes, et d'une structure de calcul électronique effectuant la comparaison des niveaux de gris de l'image, donc leur codage. Ce bloc a une structure de calcul matriciel, par nature très performante. La partie inférieure de la figure représente des éléments de calcul numériques, qui recherchent l'appariement entre les codes calculés précédemment. Ces éléments sont conçus pour effectuer des calculs sur une ligne entière, conservant ainsi la force de calcul parallèle des structures matricielles telles que celles rencontrées dans les capteurs d'images (et a fortiori dans les rétines).

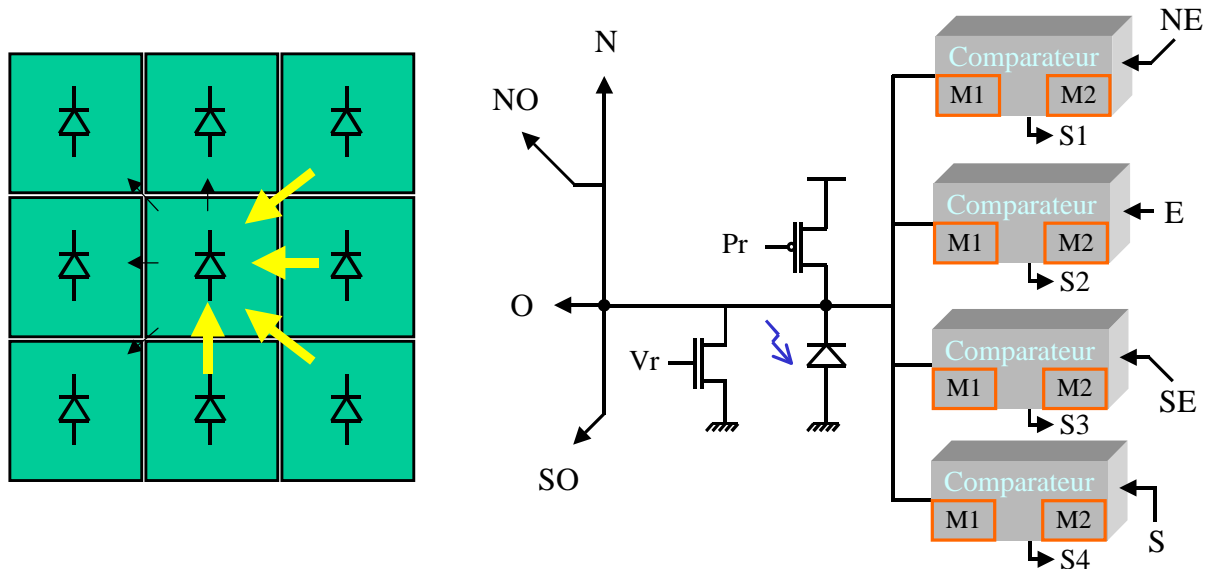
## II. Matrice de pixels : photoréception et codage

La matrice effectue la phototransduction et le codage de la texture de l'image. Comme nous l'avons montré dans le chapitre I, une photodiode effectue la transduction photons – électrons dans chaque pixel. L'information lumineuse est donc transposée en une valeur analogique dans chaque pixel. Ensuite, dans chaque pixel, la valeur est comparée à celle des pixels voisins, et le résultat du signe de cette comparaison est stocké dans des points mémoires.

Nous utilisons le mode d'intégration du photocourant, car nos premières simulations ont montré que l'ordre de grandeur des photocourants (mode courant) et les réponses de type logarithmique causaient des problèmes de sensibilité dans les structures de comparaison. En mode courant, par exemple, si on considère un photocourant - correspondant au stimulus lumineux maximum – de 50 pA, la résolution de 256 niveaux de gris entraîne un courant "élémentaire" de l'ordre de 200 fA. Compte tenu des paramètres technologiques, nous ne pouvons pas comparer de telles valeurs de courant. En effet, l'ordre de grandeur des courants de fuite (quelques fA), la taille improbable de miroirs de courants amplificateurs par 100 ou 1000, ou plus généralement la disparité inter-pixels de structures amplificatrices à grand gain sont autant de facteurs limitants.

En mode d'intégration du photocourant, la capacité de la photodiode est initialisée à un potentiel, puis se décharge. Cette décharge est d'autant plus rapide que la quantité de lumière reçue est importante. Comme le montre la figure V.2, le principe du codage inter-pixels est de comparer les tensions des différents pixels, et d'en obtenir un codage de la texture à travers des comparateurs électroniques. Ainsi, chaque pixel est comparé à ses huit voisins directs. Pour des raisons de surface, et compte tenu des symétries, le nombre de comparateurs est

réduit à quatre. En effet, comme les comparaisons se font sur chaque pixel, la moitié des comparaisons est redondante. Par exemple, le résultat de la comparaison entre le pixel A et le pixel B est le complément de la comparaison entre le pixel B et le pixel A. Un simple aiguillage, couplé à une inversion, est à mettre en place lors de la lecture des valeurs.



**figure V.2 : Comparaisons d'un pixel avec son voisinage et structure d'un pixel**

Le pixel est composé de la zone photosensible, de la photodiode, et de quatre comparateurs. Pour coder les deux images acquises, chaque comparateur est capable de stocker deux codes, dans deux points mémoires M1 et M2. La photodiode est couplée à un transistor de protection, qui limite la décharge de la photodiode; son rôle est expliqué en détail dans le chapitre III. Afin de garder des explications et des figures simples, nous considérons par la suite la comparaison de deux pixels. Soit à comparer deux pixels A et B, A étant moins éclairé que B, les décharges relatives à ces deux pixels ont l'allure donnée dans la figure V.3.

Nous constatons alors qu'il existe deux moyens pour comparer ces décharges : soit en tension, soit dans le temps.

La comparaison en tension a pour principe de comparer, à un instant donné  $t_c$ , les deux tensions analogiques décroissantes. Nous utilisons pour cela un comparateur déclenché, qui, à partir d'une faible différence de tension, génère une valeur binaire ('0' ou '1') suivant quelle tension est supérieure.

La comparaison en temps consiste à mémoriser l'instant de franchissement d'une tension de seuil  $V_c$ . La première pente qui franchit le seuil provient forcément du pixel qui a reçu le plus de lumière.

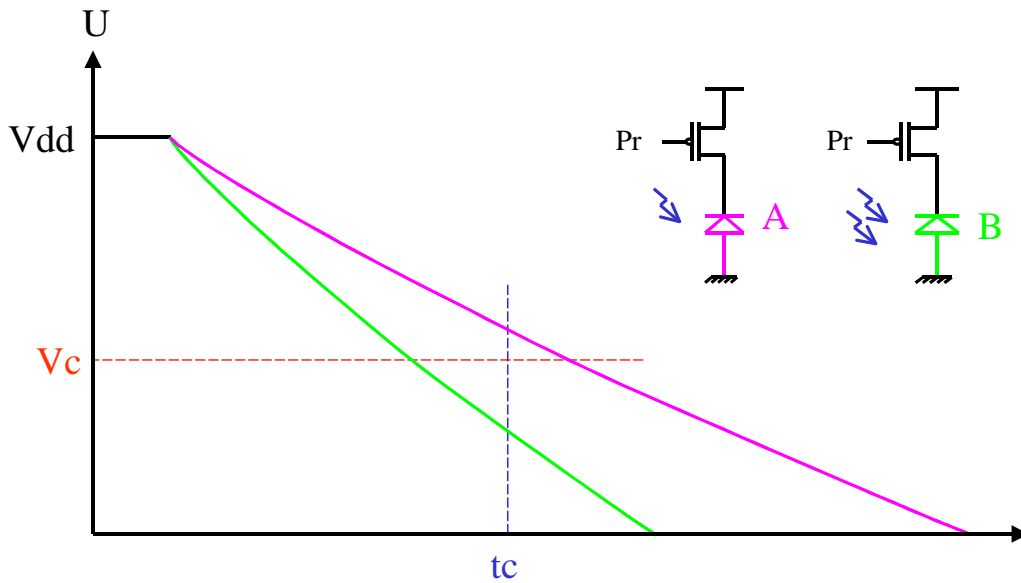


figure V.3 : Décharge des pixels A et B à comparer

Nous allons détailler ces deux structures de comparateurs, en indiquant leur structure, leur fonctionnement, et les simulations correspondantes.

## II-1 Comparateur en tension

### II-1-1 Structure du comparateur

Le comparateur en tension est de type déclenché. Il reçoit deux entrées analogiques à comparer, un signal de commande, et génère un résultat numérique sous forme d'un bit. Le schéma de principe de cette structure est représenté dans la figure V.4. Il s'inspire des structures de comparateurs [Yuk-85] [Raz-92] rapides et sensibles, que nous avons modifiées pour qu'elles soient plus adaptées à notre application. Un point mémoire statique – composé de deux inverseurs - est au centre de ce comparateur. Comme le montre le schéma de principe de la figure V.4, deux suiveurs de tension P forcent le point mémoire à '1' ou à '0' suivant les valeurs des tensions à leurs grilles.

Les tensions présentes sur les photodiodes fixent l'état initial du point mémoire statique. Le grand gain de cette structure amplifie la différence de tension, et le point mémoire converge et bascule dans un état stable. Compte tenu de l'instabilité des points mémoires lors du déclenchement, il est peu probable que l'état métastable perdure, particulièrement avec des entrées non statiques.

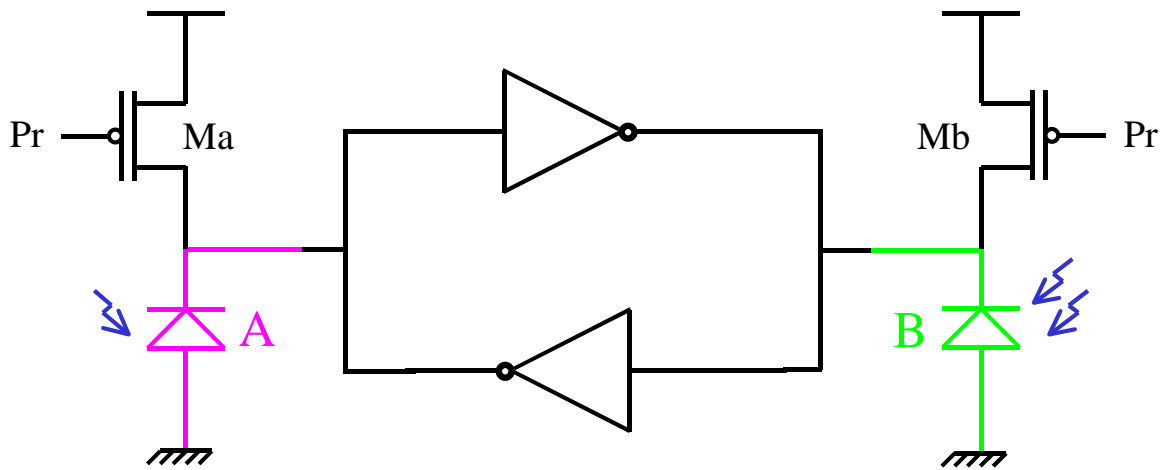


figure V.4 : Schéma de principe du comparateur en tension

Un schématique électrique plus précis du comparateur - connecté à deux pixels - est donné dans la figure V.5. Le point mémoire est composé des transistors M2, M3, M7 et M8. Les transistors M6 et M9 servent à l'initialisation de la structure. La sortie de ce comparateur est soit simple (S1), soit différentielle (S1,S2), puisque S1 et S2 sont complémentaires.

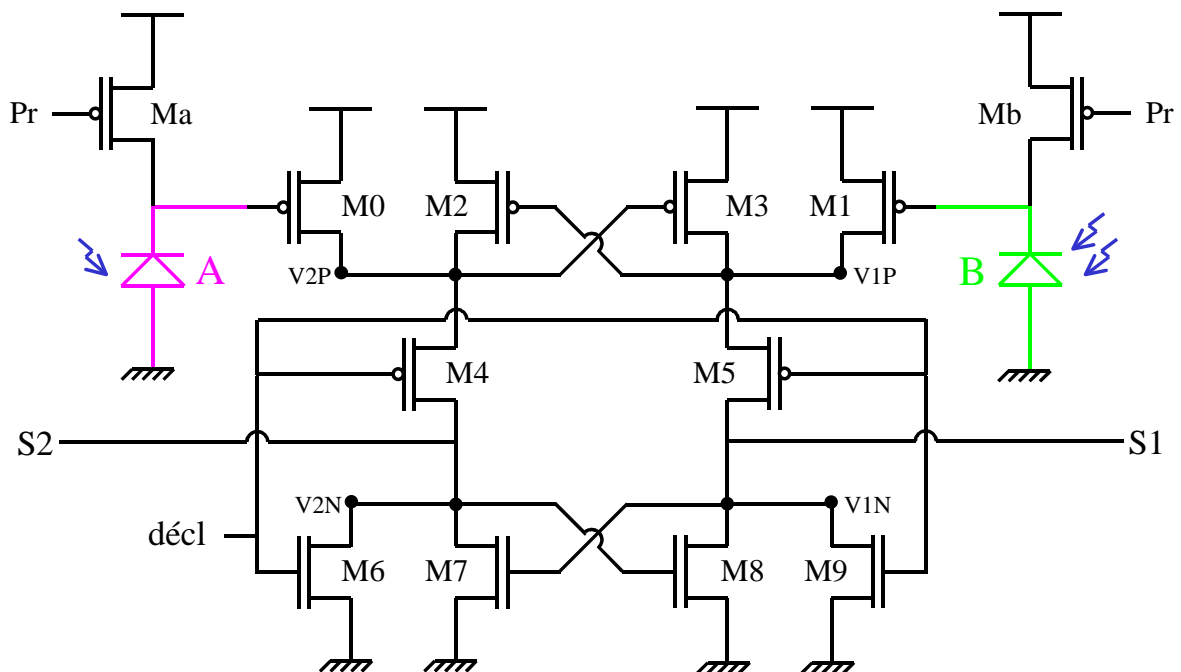


figure V.5 : Schéma du comparateur en tension

En réalité, deux transistors supplémentaires sont connectés en parallèle avec les transistors M0 et M1, pour pouvoir re-initialiser les transistors M2 et M3 au même potentiel (ils sont alors bloqués). Ils ne sont pas présentés pour des raisons de lisibilité. Pour caractériser

correctement les comparateurs de cette matrice, une structure d'imageur APS à 3 transistors est connectée en parallèle avec le transistor M0. Une image sous forme analogique est donc disponible en même temps que le codage, pour valider ce dernier. Nous allons détailler le fonctionnement de cette structure.

## II-1-2 Fonctionnement

Dans une phase d'initialisation, les transistors Ma et Mb préchargent les photodiodes à Vdd. Pour cela, le signal "Pr" est mis à '0' pendant un temps correspondant à la charge des capacités des photodiodes par le courant que les transistors Ma et Mb peuvent transmettre. Ensuite, la quantité de lumière impose des décharges plus ou moins prononcées. L'entrée "décl" est le signal qui déclenche la comparaison à l'instant  $t_c$  (figure V.3). Quand "décl" est à l'état haut, les transistors M6 et M9 sont conducteurs, tandis que M4 et M5 sont bloqués. Même si M0 et M1 sont (faiblement) conducteurs, on a  $V_{1P} = V_{2P} = V_{dd}$  et  $V_{1N} = V_{2N} = 0$ . En conséquence, la sortie est dans l'état haut car l'entrée de l'inverseur est forcée à 0. Les capacités continuent à se décharger au cours du temps. Celles-ci sont en parallèle avec les capacités de grille des transistors M0 et M1, mais ces dernières sont faibles. Au moment où "décl" passe à l'état bas, les transistors M4 et M5 (de type P) deviennent conducteurs, tandis que M6 et M9 (de type N) se bloquent. La mise en conduction de M4 et M5 provoque simultanément une diminution de  $V_{1P}$  et  $V_{2P}$  et une augmentation de  $V_{1N}$  et  $V_{2N}$ . Cependant, du fait du câblage des transistors M2 et M3, une diminution de  $V_{1P}$  devrait entraîner une augmentation de  $V_{2P}$  et réciproquement. Supposons que le pixel A soit moins éclairé que le pixel B, à l'instant de comparaison  $t_c$  nous obtenons  $V_{2P} > V_{1P}$ . M1 fournit alors plus de courant que M0. Cette petite différence de tensions ( $V_{2P} - V_{1P}$ ) suffit cependant à déséquilibrer les courants fournis par les transistors M2 et M3, amplifiant le phénomène  $V_{2P} > V_{1P}$ . Le basculement de la structure est alors amorcé et converge vers un niveau logique (0 ou Vdd). Quand l'état s'est établi, les faibles variations qui peuvent subvenir en entrée du comparateur ne le perturbent plus.

A cause de sa nature (amplification de faibles différences), la structure doit être parfaitement équilibrée. Des capacités différentes sur chacun des nœuds internes du point mémoire conduiraient en effet à un déséquilibre des tensions  $V_{1P}$  et  $V_{2P}$ , donc un résultat potentiellement erroné. C'est pourquoi nous avons placé en sortie deux inverseurs identiques. Un soin particulier doit être donné à ce type de structure lors du dessin des masques, afin de les rendre le plus symétrique possible.

Des étages de sortie ou des capacités de colonnes différentes peuvent en effet exister. Le



résultat est donc disponible sur une sortie simple, ou sur une sortie différentielle redondante.

Le résultat obtenu en sortie du comparateur – sur 1 bit, ou 2 bits complémentaires – correspond à l'un des 8 bits représentant le résultat du codage d'un pixel.

### II-1-3 Simulations

Nous avons simulé le fonctionnement de ce comparateur, à deux niveaux : un comparateur et ses deux pixels, et un pixel complet composé d'une photodiode et de 8 comparateurs (4 comparateurs par image acquise) connectés aux pixels voisins. Ces voisins sont aussi connectés à des comparateurs pour étudier précisément tous les effets, cela représente un système lourd à simuler. Alors que les simulations d'un seul comparateur sont prévisibles, avec même une précision improbable au femtoampère, nous avons mis à jour des problèmes lors de la simulation d'un pixel complet : des couplages capacitifs entre les différents blocs limitent la précision de la structure. Ces couplages ont une influence très limitée, cette conclusion provient de nos hypothèses. La figure ci-dessous représente le chronogramme observé en simulations.

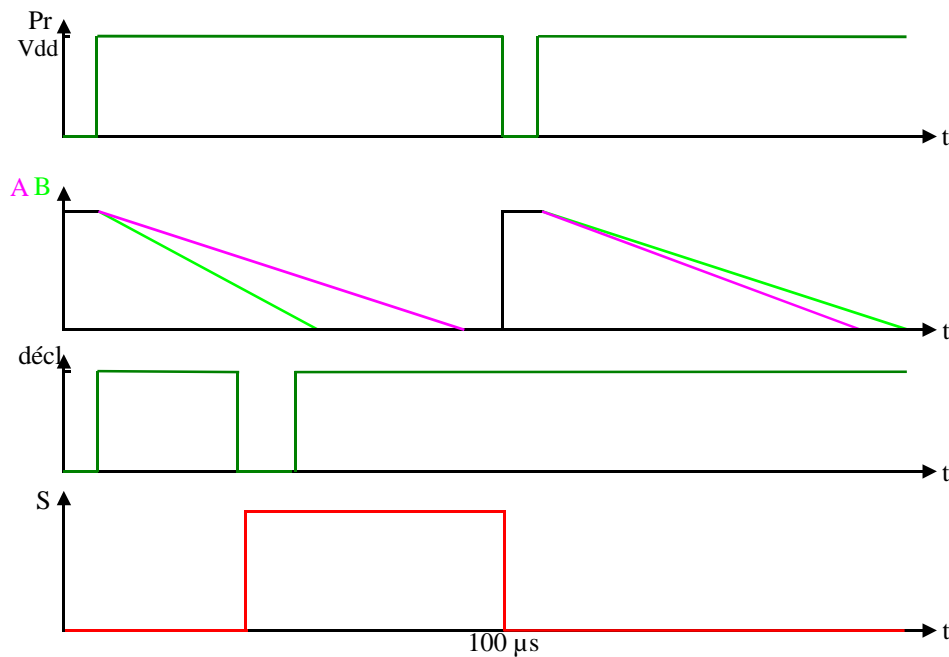


figure V.6 : Chronogramme des simulations

Les simulations que nous avons effectuées sur un pixel complet (en vue schématique et en vue extraite prenant en compte les capacités parasites) ont donné une précision de 50 μV, correspondant à une différence des photocourants de 0,1 pA. Nous définissons la précision comme la différence de photocourant minimale que la structure peut comparer. En effectuant un calcul par rapport à un photocourant maximal de 50 pA, nous trouvons que cette structure

permet de comparer 500 niveaux de gris. Cette résolution est acceptable, puisque supérieure aux 256 niveaux utilisés dans l'algorithme original.

Nous avons aussi mesuré, par la simulation, le temps de commutation de cette structure, il est de 940 ps. Ce temps correspond au délai entre le déclenchement du comparateur et l'établissement du niveau logique en sortie.

## II-2 Comparateur en temps

### II-2-1 Structure du comparateur

Le comparateur en temps est basé sur le principe de la détection du premier front montant arrivant dans la structure de comparaison. Il s'appuie sur la structure d'une bascule RS, en amont de laquelle une série d'inverseurs met en forme le signal analogique. La figure V.7 illustre le schéma d'un comparateur entre deux pixels.

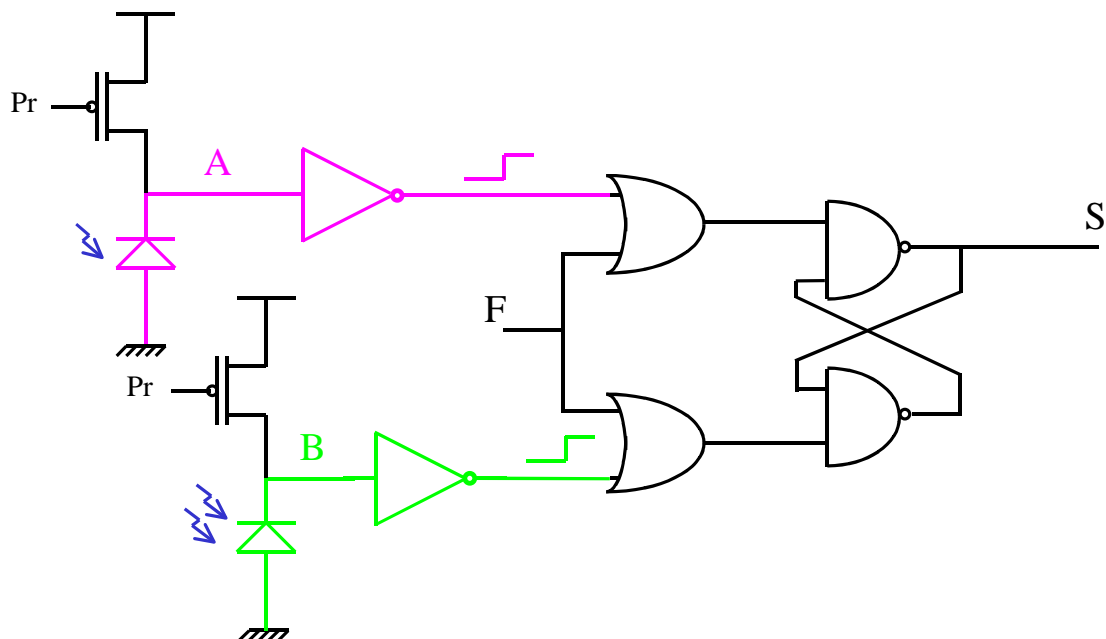


figure V.7 : Schéma du comparateur en temps

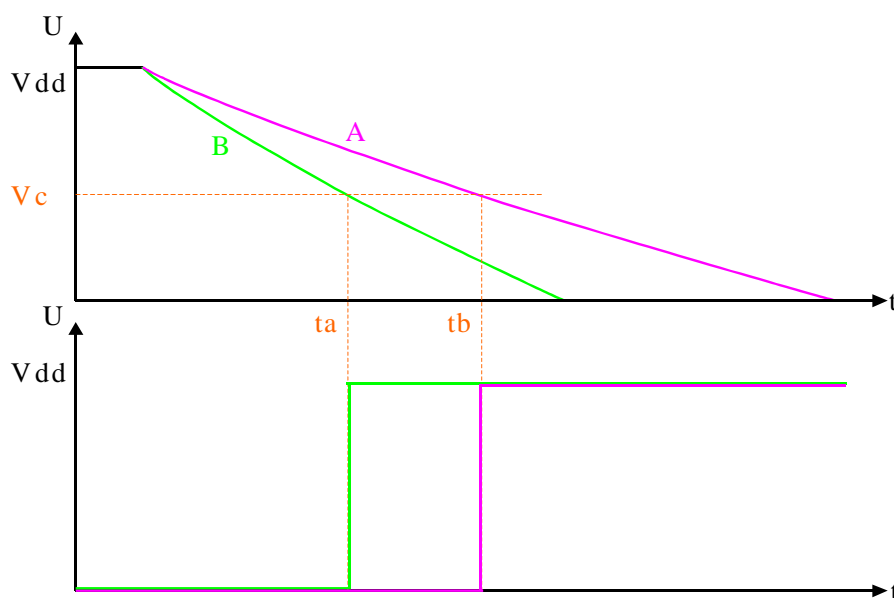
La conversion analogique – numérique se fait donc à travers des inverseurs, le comparateur ne mémorise que l'information contenant le signe de la différence des luminosités. Cette mémorisation se fait sous forme binaire: '0' ou '1' selon quel front s'est présenté en premier dans la bascule, qui a naturellement la propriété de mémoire. La propriété de cette bascule est aussi que le deuxième front entrant ne perturbe pas le résultat établi par le premier. Un signal de forçage "F", peut inhiber les valeurs provenant des pixels. Ce signal permet de bloquer – et isoler - la valeur dans le point mémoire de la structure. Un autre

comparateur en parallèle peut alors acquérir la deuxième image. Comme dans la structure de comparaison précédente, les entrées "Pr" servent à initialiser les photodiodes à Vdd.

L'étape de calcul numérique très itératif est transposée ici en une commutation de portes dans chaque pixel. Ces bascules engendrent un calcul électronique simple se faisant de manière matricielle (parallèle à la fois en X et en Y), donc très rapides. Nous avons aussi placé dans le pixel une structure d'imageur, permettant de capturer les images qui sont codées. Pour cela, nous nous sommes inspirés de la structure de l'imageur présenté dans le chapitre II. Des inverseurs commandés, connectés après la chaîne d'inverseurs, aiguillent l'image binaire vers la sortie du circuit. Nous pouvons cependant effectuer le seuillage temporel expliqué dans ce chapitre II pour acquérir des images en niveaux de gris.

## II-2-2 Fonctionnement

Chaque chaîne d'inverseurs reçoit un signal assimilable à une pente descendante (dont la pente traduit l'éclairement du pixel), et la binarise. Comme la chaîne se compose de trois inverseurs, la pente descendante entraîne un front montant à un instant donné. Le seuil de conversion  $V_c$  de cette structure est donc celle d'un inverseur, proche de  $V_{dd}/2$ . Comme le montre la figure V.8, la différence de pente est ainsi transposée en deux fronts montants décalés d'un temps  $\Delta t = t_b - t_a$ .



**figure V.8 : Mise en forme et conversion des signaux analogiques**

La précision de ce comparateur est alors donnée par sa rapidité. Il doit en effet être à même de discriminer deux signaux qui peuvent être très proches. Cependant, compte tenu de la faible complexité de la structure RS, et donc de son faible temps de réponse, une bonne

sensibilité est prévisible.

### **II-2-3 Simulations**

Nous avons mené des simulations sur cette structure avec le même protocole que la structure précédente : un comparateur et ses deux pixels, puis un pixel complet comparé à ses huit voisins. Dès les premières simulations, cette structure s'est montrée plus précise que le comparateur déclenché. Du fait de sa structure numérique et de son verrouillage, ce comparateur est peu sensible aux fluctuations des entrées. Nous avons mené des simulations sur différentes vues (schématiques et extraites) en prenant compte des disparités possibles (analyses "Corners").

Cette structure a une précision de 0,05 pA. En considérant une luminosité maximale de 50 pA, le comparateur traite des images en 1000 niveaux de gris ; la précision est donc très bonne. Le temps de commutation que nous avons relevé est de 48 ps.

## **III. Recherche de codes**

Du point de vue matériel, la recherche des codes est un processus itératif à  $N$  étapes,  $N$  étant la taille du voisinage considéré. Si elle se fait sur un voisinage  $3 \times 3$ , elle nécessite 3 itérations (une par ligne). Cette structure permet de conserver le parallélisme présent à la sortie de la matrice de codage. Elle traite tous les pixels d'une ligne simultanément. Afin de rechercher les codes dans un fenêtrage, les éléments de calcul élémentaires que nous avons développés sont capables de communiquer avec leurs voisins. Le nombre de voisins à connecter est choisi dès le départ, puisqu'il fixe la densité des interconnexions et la complexité de ces éléments. En accord avec les résultats montrés dans le chapitre III (efficacité de l'algorithme en fonction du voisinage considéré), nous avons fixé ce voisinage à  $3 \times 3$  pixels. La recherche des codes se décompose en plusieurs phases : la lecture des codes des pixel dans les deux images, puis la recherche de leur correspondance à travers les distances de Hamming.

```

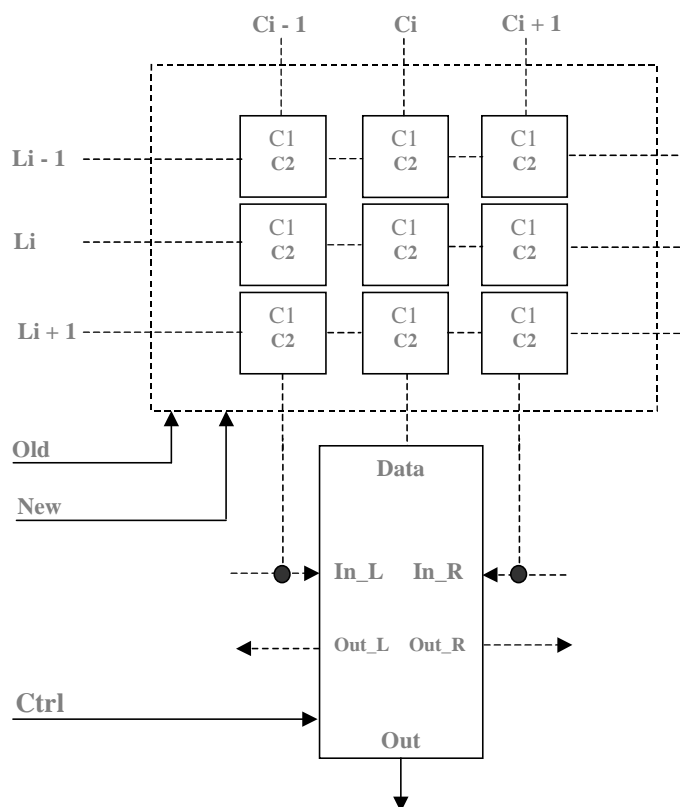
For l = 1 to N
  im1 = lecture_ligne(l)
  For {lignes_voisines [l-1 , l+1 , l]}
    im2 = lecture_ligne(v)
    calcul_min(im1,im2)
  end For
end For

```

**figure V.9 : Processus de calcul matériel pour la recherche des codes**

La figure ci-dessus explique les itérations nécessaires pour effectuer le calcul matériel. Pour chaque ligne de codes de la première image, nous balayons les 3 lignes de l'image 2, en dessus (l-1), au dessous (l+1) et à la même ligne. Le fait de finir par la même ligne privilégie le mouvement nul si d'autres mouvements ont le même score.

L'avantage est de traiter le calcul au niveau des lignes. La complexité – et donc le temps - de calcul est proportionnelle au nombre de lignes de l'image. Pour une recherche dans un voisinage 3x3 dans une image comportant N lignes, 3.N itérations sont nécessaires. La figure V.10 montre l'architecture liée à ces blocs parallèles.



**figure V.10 : Élément de calcul pour effectuer la recherche des codes**

Nous pouvons voir à travers les sorties Out\_L, Out\_R et les entrées In\_L et In\_R le principe de la communication horizontale permettant d'effectuer le calcul par ligne. Quand une ligne est adressée dans la matrice, les codes C1 ou C2 sont accessibles, selon que Old ou New est actif. Nous utilisons deux signaux à cause de la structure des comparateurs dans les pixels. Pour l'exemple de la figure ci-dessus, dans une première phase, le code C1 de la ligne centrale (Li) est stockée dans les éléments de calcul. Ensuite, une recherche se lance à chaque lecture d'une des trois lignes de codes C2. Comme chaque élément est connecté à un pixel, la communication latérale sert à l'élément de la colonne Ci pour lire les codes des colonnes Ci-1 et Ci+1.

La taille des vecteurs en entrée ("Data") est de 8 bits, le résultat ("Out") est sur 3 bits, indiquant une des 9 directions possibles. Les entrées de contrôle "Old" et "New" proviennent d'un séquenceur, sélectionnant l'accès à une des deux images de la séquence. Une fois le calcul d'une ligne effectué, les résultats de celle-ci seront dirigés vers la sortie du circuit, sur un bus de trois bits.

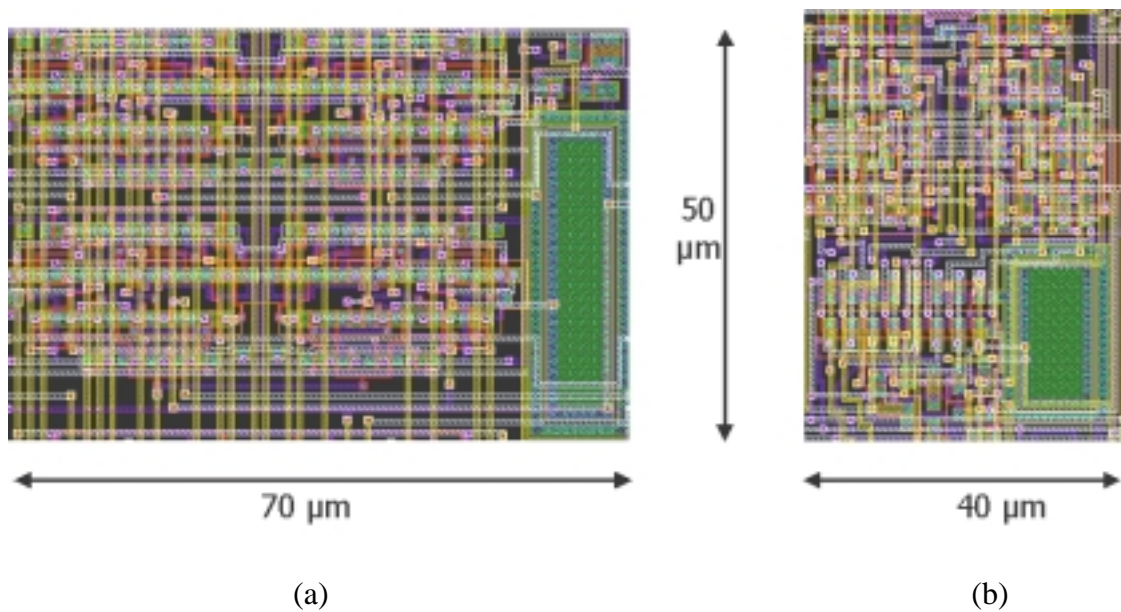
Le temps de calcul de cette structure est de 14,4 ns pour une ligne (recherche d'un code C1 parmi trois lignes de codes C2). En considérant le temps de codage des pixels de 1 ns, une image complète de 100x100 pixels nécessite donc un temps de calcul total de 1,441  $\mu$ s. Si la rétine a un débit de données similaire à un celui d'un imageur standard, par exemple le VV5301 de ST-Microelectronics (17,7 Mbits/s), elle pourra transférer 30 images par secondes (soit 33 ms par image). En considérant que le temps d'acquisition de deux images est de 66 ms, le temps d'acquisition et de calcul de la rétine est de 66,0014 ms.

## **IV. Implantation dans un ASIC et caractérisation**

Comme il n'est pas possible de modéliser tous les paramètres optiques et de phototransduction en simulation, la validation par la mesure s'impose dans de tels systèmes. De plus, une simulation au niveau système n'est pas concevable avec les outils informatiques actuels, à cause du temps de calcul nécessaire. Pour valider et caractériser le calcul d'estimation du mouvement, nous avons donc dessiné le "layout" (dessin des masques) des différentes structures dans la technologie AMS 0,35  $\mu$ m. Compte tenu des caractères analogiques et numériques des deux parties (calcul des codes et recherche de ces derniers à travers les images), nous les avons conçus différemment.

### **IV-1 Dessin des blocs comparateurs**

Les blocs de codage constituent une partie des pixels, l'autre partie étant la photodiode. Ceux-ci ont été dessinés au niveau transistor avec une attention particulière quant à leur disposition afin d'obtenir une surface minimale. Les deux versions de pixels (correspondant aux deux types de comparateurs), sont présentés dans la figure V.11: le bloc (a) est le comparateur en tension déclenché, le bloc (b) est le comparateur en temps à bascule RS.



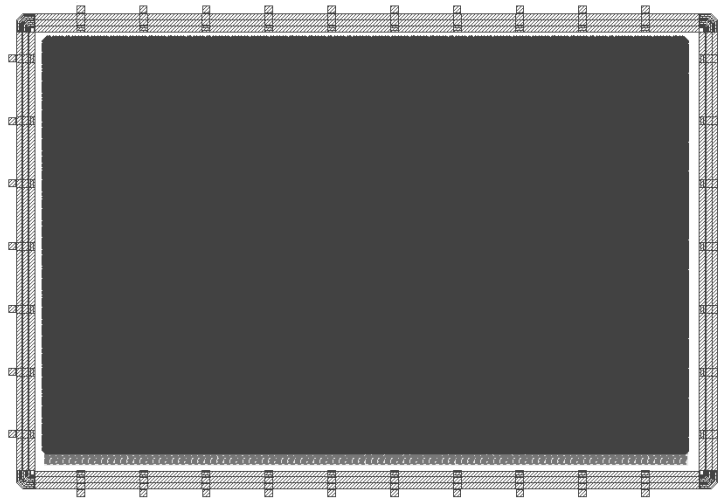
**figure V.11 : Dessin des masques des deux structures de comparateurs**

Les taux de remplissage ("fill factor") de ces deux pixels (a) et (b) sont respectivement de 5,8 % et 5 %, pour des surfaces totales respectives de  $3419 \mu\text{m}^2$  et  $2470 \mu\text{m}^2$ .

Nous avons conçu ces pixels de manière appliquée, mais non optimale en terme de surface, les surfaces données sont donc indicatives pour la faisabilité d'un tel système. De plus, les pixels et les photodiodes sont rectangulaires (et non carrés), ce qui cause une sensibilité spatiale du capteur différente dans les deux axes. Ces paramètres sont à prendre en compte pour des capteurs de résolution importante. Cependant comme ces pixels sont destinés à simplement valider notre intégration sur des matrices de test (de faible résolution), nous avons omis ces règles de conception d'un capteur.

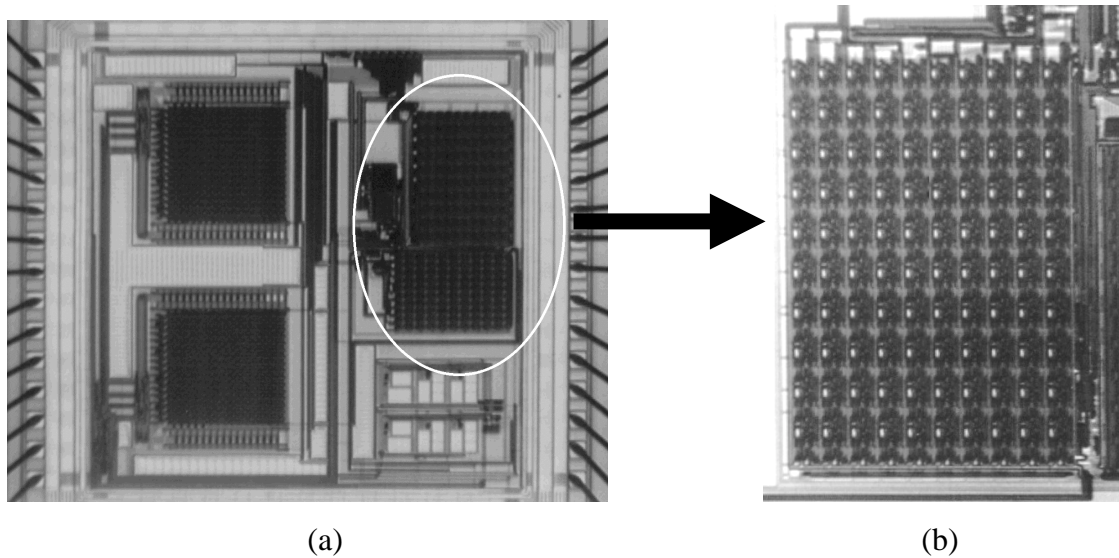
Pour réaliser une matrice de  $100 \times 100$  pixels, la surface nécessaire pour une telle matrice est de (respectivement)  $34 \text{ mm}^2$  et  $24 \text{ mm}^2$ , et pour une matrice de  $256 \times 256$  pixels  $223 \text{ mm}^2$  et  $161 \text{ mm}^2$ . La figure V.12 montre le dessin (layout) d'un circuit contenant une matrice de  $100 \times 100$  pixels et les 100 blocs numériques de calcul (qui s'aboutent en bas de la matrice). Ces blocs sont détaillés plus bas. Nous avons considéré pour cet exemple la solution la plus

coûteuse, c'est à dire utilisant des pixels à comparateurs en tension.



**figure V.12 : Rétine contenant une matrice de 100x100 pixels**

Compte tenu du coût de la fabrication d'un circuit, nous avons en fait dessiné des matrices de test dans un circuit multi-projets. Leur résolution est donc très faible et permet uniquement de valider les structures analogiques présentées. La validation doit en effet être vérifiée au niveau de chaque pixel, les deux matrices sont de 100 pixels - 10x10 – (figure V.13).



**figure V.13 : Microphotographie du circuit et d'une des deux cellules test de 10x10 pixels**

La consommation moyenne de la matrice de test est faible, de l'ordre de  $500\mu\text{W}$  (simulée), à la cadence vidéo de 25 images/s. Une matrice de 100x100 pixels consommerait donc 50mW.



## IV-2 Dessin des blocs de recherche des codes

Les blocs de recherche sont des éléments de calcul qui finalisent le calcul en associant les pixels. Comme le séquencement est assez complexe à réaliser au niveau portes, nous avons utilisé le langage VHDL pour décrire le comportement de tels blocs. Nous avons écrit ce code afin d'obtenir le bloc le plus petit possible. Comme ces éléments de calcul viennent s'ajouter en bas de la matrice, ils ont leur largeur fixée par la taille des pixels. La figure V.14 montre le dessin des masques (vue "layout") d'un élément de calcul.

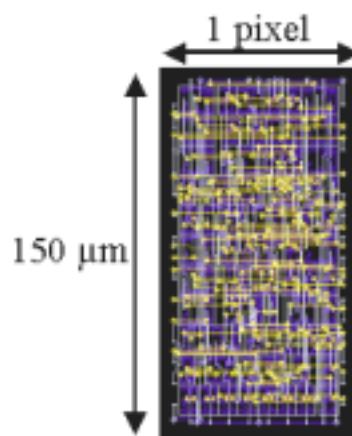


figure V.14 : Dessin des masques d'un élément de calcul

Nous avons utilisé les outils Cadence pour concevoir ces éléments de calcul : la synthèse avec AMBIT, et le placement-routage avec Silicon Ensemble. Chaque bloc comporte 62 portes logiques et a une surface unitaire de  $11000 \mu\text{m}^2$ . Pour des matrices de  $100 \times 100$  pixels - comme de  $256 \times 256$  pixels -, ils représentent environ 3 % de la surface totale du capteur. Si nous avons gardé l'algorithme LTS original (voir chapitre IV), chaque bloc compterait 1013 portes logiques, ce qui représente une solution coûteuse en surface.

Comme le fonctionnement des blocs numériques est sûr dans la technologie  $0,35 \mu\text{m}$  utilisée, nous avons jugé utile de ne faire fabriquer que les matrices effectuant le codage. En effet, c'est en fait la matrice, regroupant phototransduction, conversion analogique / numérique et calcul des codes avec le voisinage, l'élément le plus critique. Ceci apporte aussi plus de simplicité, de contrôlabilité et d'observabilité pour les caractérisations.

## V. Mesures

Dans cette partie, nous présentons la mise en œuvre du circuit, les mesures effectuées, puis nous interprétons les résultats. Enfin, nous proposons une comparaison des deux comparateurs utilisés.

### V-1 Mise en œuvre

#### V-1-1 Présentation

Nous avons mis en œuvre le circuit de test avec la plaquette Altera Excalibur-NIOS et le banc optique déjà utilisés pour le circuit imageur. Pour des raisons de coût et des contraintes de surface minimale, le circuit fabriqué est un circuit multi-projets. Il est composé de différents blocs (figure V.15), nous avons adressé le bloc B2 à l'aide d'entrées démultiplexées numériques (Num) et analogiques (Ana) et de sorties multiplexées. Le nombre de broches du circuit est limité, le circuit a donc des plots d'entrée et de sortie communs.

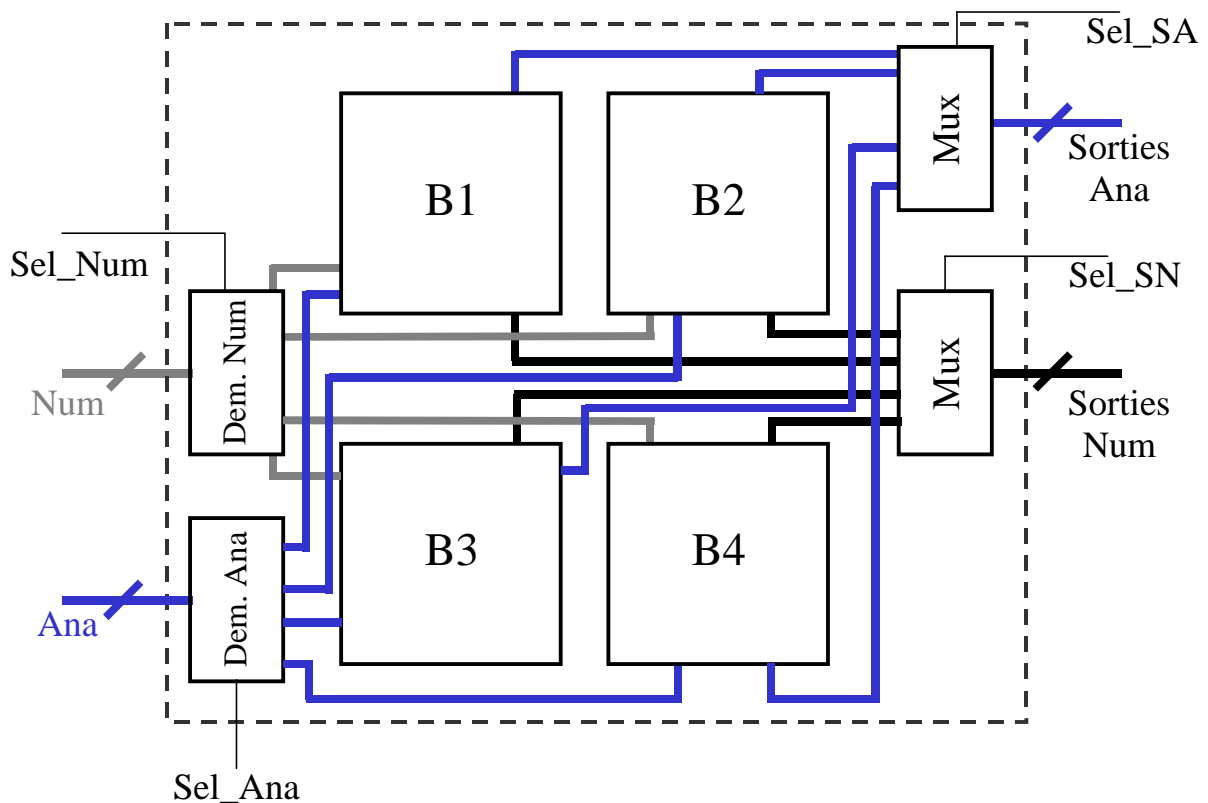
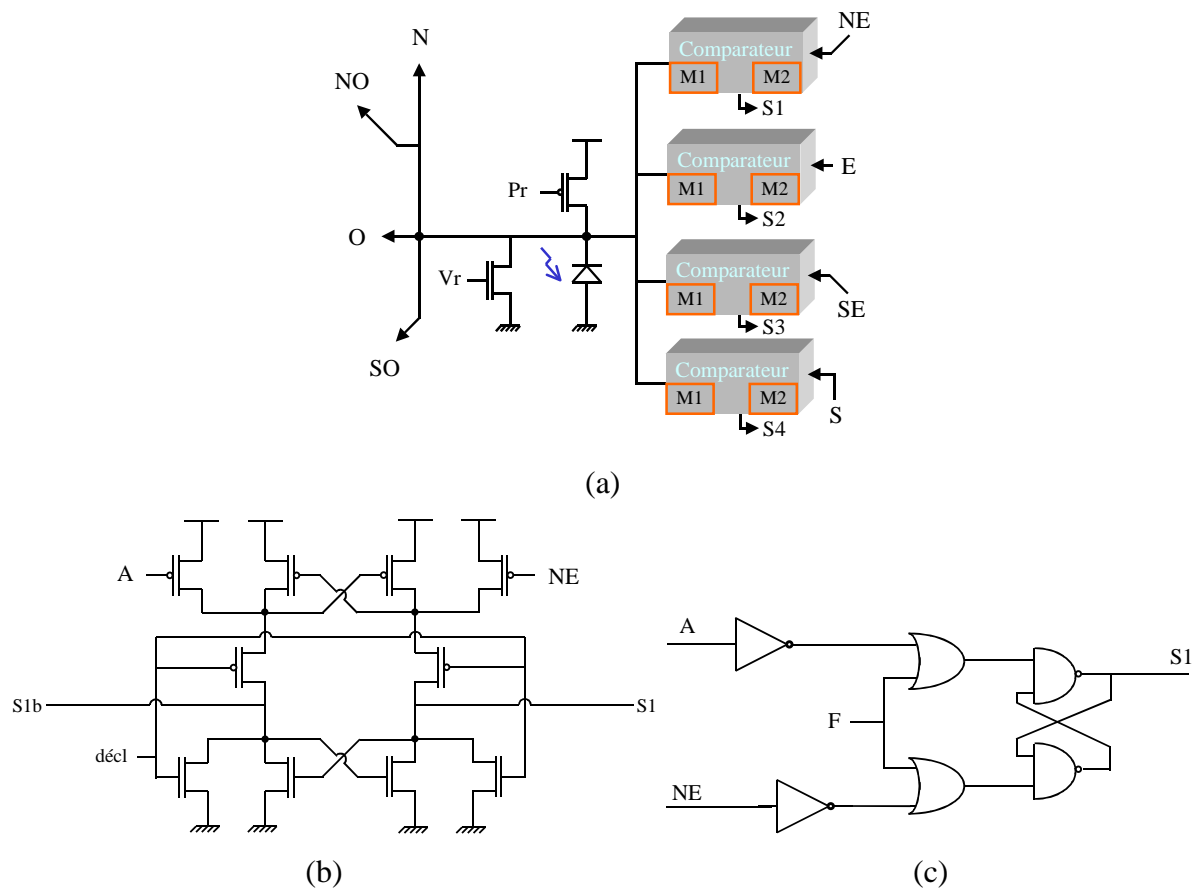


figure V.15 : Schéma du circuit complet multi-projets

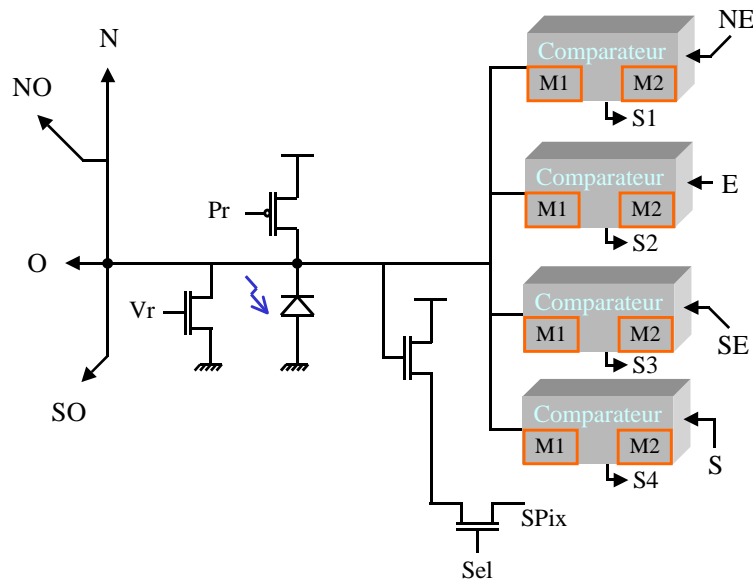
Les mesures ont porté sur l'acquisition d'images en niveaux de gris, et la lecture des codes des transformées associés à ces images. Nos mesures ont porté sur les deux matrices de codage : composées des comparateurs en tension ou en temps. Nous rappelons la structure du pixel (a), ainsi que la structure des comparateurs qui le composent dans la figure V.16. (b) est le comparateur déclenché, (c) le comparateur en temps.



**figure V.16 : Schéma du pixel et détail des comparateurs**

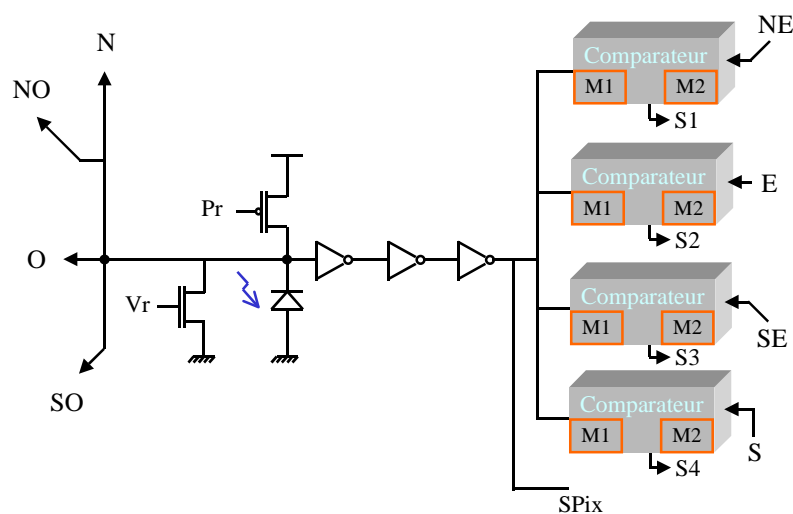
En entrée, nous appliquons les signaux de précharge des photodiodes ("Pr"), de déclenchement de la comparaison (pour le comparateur déclenché : "décl"), puis, après la comparaison, nous récupérons les valeurs des codes. La lecture des codes se fait sur 4 bits (S1 à S4) en adressant une ligne et en sélectionnant un pixel de cette ligne à l'aide de décodeurs. Afin de caractériser ces codes, les pixels incluent, en parallèle des comparateurs, une structure permettant de lire l'image. Dix plots (analogiques) sont dédiés à cette fonction, afin de lire l'image ligne par ligne (résolution de la matrice : 10x10 pixels). Chaque matrice génère en fait une image de nature différente, selon que la structure du comparateur est analogique

(comparateur en tension déclenché) ou numérique (comparateurs en temps à bascule RS). Les comparateurs en tension ont donc une structure de type APS 3 transistors (figure V.17), permettant de relever le signal analogique des décharges des photodiodes en sortie. Nous avons ajouté deux transistors (suiveur et sélection) dans le pixel.



**figure V.17 : Structure de type APS 3 transistors pour lire l'image dans les pixels à comparateurs déclenchés**

Les comparateurs en temps produisent une image binaire - de type DPS 1 bit -, le signal provient en effet de la chaîne d'inverseurs dans le pixel (figure V.18).



**figure V.18 : Structure de type DPS pour lire l'image dans les pixels à comparateurs en temps**

En utilisant la méthode d'échantillonnage temporel présentée dans le chapitre III, nous pouvons acquérir des images en niveaux de gris. L'adressage se faisant ligne par ligne, nous avons jugé inutile d'intégrer des points mémoires dans les pixels. En effet, comme la matrice a une résolution faible, il est possible d'échantillonner la valeur de tous les pixels (successivement) pendant leur décharge sans qu'elle ne cause une erreur importante. Le temps de lecture des 10 lignes est rapide : 10  $\mu$ s à 1 M échantillons / s, et génère une erreur de 0,1 % sur la mesure (décharges de phototensions de 10 millisecondes pour cet exemple).

### V-1-2 Problèmes rencontrés

Lors des mesures, nous avons constaté deux erreurs de conception qui limitent la caractérisation.

La première erreur concerne la matrice de comparateurs en temps. L'étage de sortie des comparateurs présente un court-circuit, comme le montre la figure V.19.

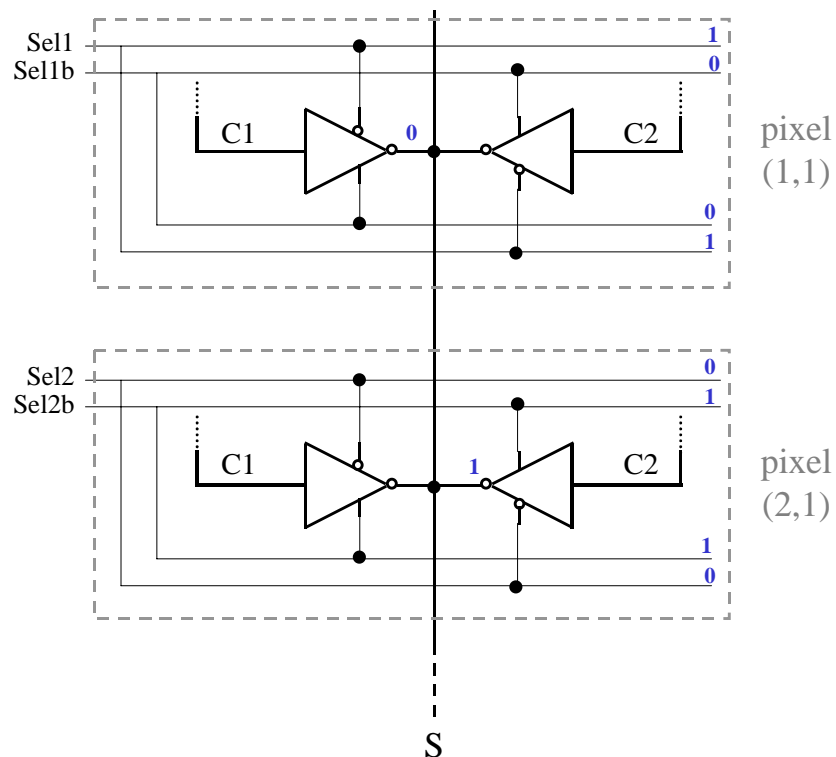


figure V.19 : Court-circuit dans l'étage de sortie des comparateurs en temps

Cet exemple considère deux pixels d'une colonne donnée (la colonne 1), sur deux lignes successives (lignes 1 et 2). C1 et C2 sont les deux codes mémorisés dans le pixel, Sel1 et

Sel1b sont deux entrées complémentaires de sélection de ligne, et S le bus colonne véhiculant l'information binaire. Nous rappelons que la matrice a comme résolution 10x10 pixels.

Lorsqu'on sélectionne la ligne 1, le mot binaire "1000000000" est appliqué sur les entrées Sel1 à Sel10 (et le complément "0111111111" sur les entrées Selb). Finalement, le bit C1 du pixel (1<sup>ère</sup> ligne, 1<sup>ère</sup> colonne) est en conflit avec les bits C2 de toutes les autres lignes.

Le mode d'adressage ne permet pas de sélectionner de lignes afin d'effectuer la lecture. Le codage peut s'effectuer, mais nous ne pouvons donc pas le relever. Il n'est pas possible d'envisager la réparation de ce bloc, compte tenu du trop grand nombre de modifications à apporter (4 par pixel).

La deuxième erreur concerne la matrice de comparateurs déclenchés. Nous avons placé un suiveur de tension en sortie des différents blocs multiplexés dans le circuit, afin de les isoler. Malheureusement, les suiveurs des pixels dans la matrice sont connectés sur ces suiveurs, ils ne sont donc plus polarisés. Grâce à de bonnes relations avec le laboratoire LAMIP de Caen, qui dispose d'un FIB (microscope à faisceau d'ions), ce dernier a accepté de tenter de réparer cette erreur. Compte tenu de la difficulté de cette intervention, peu de sorties sont effectivement fonctionnelles, mais nous avons pu effectuer des mesures très intéressantes, que nous détaillons plus loin. Nous ne pouvons cependant pas extraire assez d'informations sur l'image pour caractériser le codage. Nous devons utiliser une matrice pour acquérir l'image, et l'autre pour acquérir le codage. Nous avons défini les zones sensibles des matrices en lisant l'image ou en interprétant le code.

Puisque cette matrice n'est conçue qu'à des fins de caractérisation (par rapport à une étude de faisabilité) nous pensons a posteriori qu'il était plus judicieux de ne placer dans le circuit fabriqué qu'un comparateur de codes par pixel. Le pixel complet aurait pu nous servir pour faire simplement une estimation de la surface réelle du pixel final. Ces conclusions complètent notre expérience dans le domaine.

## **V-2 Transformées des images**

### **V-2-1 Acquisition de codes**

Notre première mesure a consisté à définir les zones sensibles des deux matrices, nous avons ensuite fait l'acquisition d'images et de codes.

Les premiers stimuli lumineux que nous avons appliqués correspondent à un contraste

fort séparant deux zones homogènes. Les codes correspondants à cette mesure et leur codage sont illustrés dans la figure V.20.

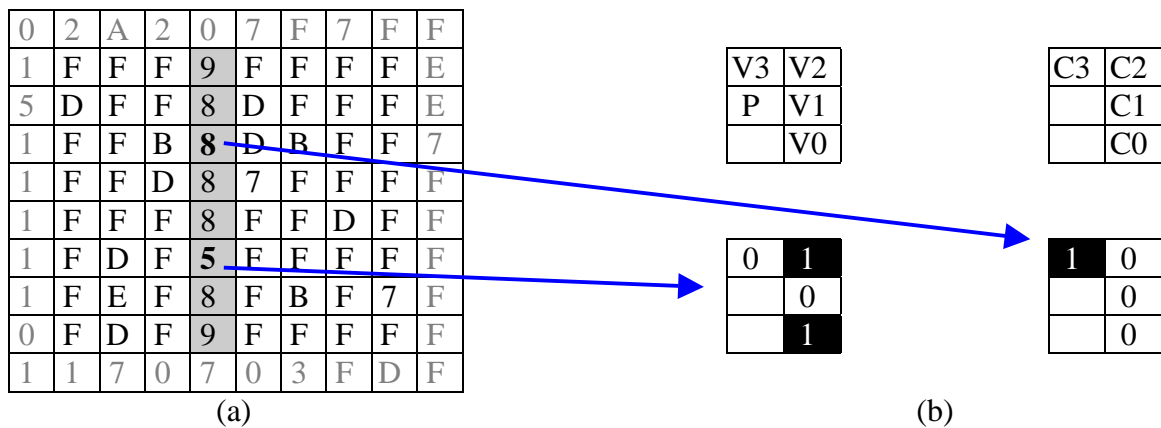


figure V.20 : Codes d'une image ayant un contraste fort (a) et masques de comparaison (b)

Pour lire ces codes, nous devons nous intéresser au masque de codage présenté en (b) dans la figure ci-dessus. Soit à comparer un pixel P avec ses quatre voisins (V0 V1 V2 et V3), le signe de la différence des niveaux de gris est codé en '1' ou en '0' binaire composant le code hexadécimal montré en figure V.20 (a). Si la luminosité de P est inférieure ou égale à celle d'un voisin V, le code vaut '1', sinon '0'. Le bit de poids fort (C3) est attribué au résultat de la comparaison de P et de V3 et ainsi de suite.

Dans la partie (a) de la figure, les codes de la périphérie sont grisés car ils n'ont pas de signification, à cause de l'effet de bord causant le déséquilibre des comparateurs. Nous observons les deux zones homogènes à gauche et à droite d'une zone de transition - qui est le fort contraste. Nous constatons dans cet exemple 10 erreurs de comparaison (en fait 10 bits faux) dans les zones homogènes. Nous attribuons ces erreurs au basculement "naturel" dû aux capacités différentes sur les nœuds des comparateurs, lorsque les stimuli sont identiques. Le code indiquant le contraste est "8", nous remarquons donc que 75 % des codes sont justes.

Avec cette mesure, nous avons pu déterminer quel est le contraste minimum que la matrice peut coder. Il s'avère que la structure n'est sensible qu'à une variation de 19 % de la luminosité (50 niveaux de gris sur la dynamique de 256 valeurs) . Une explication de ce phénomène est donnée plus loin, en V-3.

### V-2-2 Caractéristiques des codes

La figure V.21 montre une image acquise par ce circuit de test. Pour bien caractériser la précision du comparateur, cette image a une dynamique de 30 niveaux de gris. Les valeurs sont ramenées sur une échelle allant de 0 à 256 pour plus de clarté.

0	0	0	0	0	32	144	184	200	208
0	0	0	0	0	112	144	184	184	208
0	0	0	0	120	160	160	176	176	184
0	0	0	80	152	192	192	192	192	200
0	0	48	112	160	200	200	200	208	208
0	0	48	112	160	208	208	208	208	216
0	0	0	72	160	208	208	208	216	216
0	0	0	0	0	160	208	208	216	216
0	0	0	0	56	160	192	208	216	216
0	0	0	0	0	144	184	200	208	216

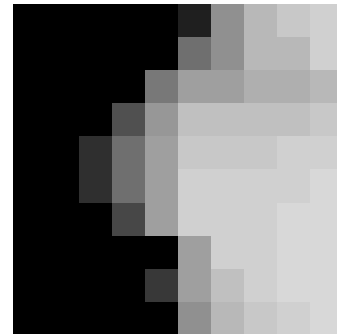


figure V.21 : Acquisition d'une image en 30 niveaux de gris

La figure V.22 montre la transformée mesurée et la transformée théorique que nous aurions du obtenir. Les codes de la périphérie ne sont pas représentés sur la figure ci-dessous.

	F	F	E	8	8	8	B	1	
	F	E	C	C	E	8	2	0	
	E	C	8	8	E	E	E	C	
	C	8	8	8	E	E	C	E	
	9	8	8	8	F	F	E	C	
	B	1	1	9	F	F	C	F	
	F	B	2	0	0	F	8	F	
	F	F	D	8	8	0	9	F	

(a)

	F	F	E	C	F	8	F	F	
	F	E	C	C	F	8	F	F	
	F	C	F	8	F	F	F	F	
	F	8	E	8	F	F	F	F	
	F	9	B	1	F	F	F	F	
	F	9	B	1	F	F	F	F	
	F	B	3	0	0	F	F	F	
	F	F	F	8	F	F	F	F	

(b)

figure V.22 : Transformée de l'image théorique (a) et mesurée (b)

Les résultats grisés dans le tableau de mesures (b) montrent les erreurs commises lors du codage. Cette mesure confirme que la variation minimum que peuvent détecter les comparateurs en tension est d'environ 50 niveaux de gris. En dessous de ce seuil, le codage ne reflète pas la texture, puisqu'il la considère homogène. La sensibilité des comparateurs est donc de 19%.

Nous avons mesuré la dispersion temporelle en sortie des comparateurs. Pour cela, nous



avons acquis un grand nombre de fois (600) la transformée d'une image statique. Une telle mesure devrait donner 600 fois le même codage de l'image. Pour que cette mesure ait un sens, nous l'avons effectuée en respectant la faible sensibilité des comparateurs. Nous avons ainsi acquis une image à fort contraste. Il s'avère que les structures sont relativement stables, puisqu'elles génèrent environ 0,3 % de fluctuation au cours du temps. 99,7% des codes restent donc inchangés. Nous estimons que les fluctuations proviennent de comparateurs dont les entrées sont sensiblement identiques, comme nous l'avons remarqué sur les zones homogènes de la figure V.20 (a). Nous avons aussi testé le codage puis la lecture des codes plusieurs fois. La lecture des résultats dans les comparateurs ne les perturbe pas.

Nous avons relevé le temps de codage (commutation) des comparateurs en observant à l'oscilloscope les stimuli envoyés et les réponses du circuit. Pour cela, les décodeurs en sortie du circuit sont figés dans un état permettant la lecture directe du code d'un pixel. Ensuite, nous acquerrons les codes. Nous nous sommes intéressés à la mesure du temps séparant les deux fronts entre le signal "décl" (voir figure V.5) et le code en sortie. Ce temps est de 600 ns, il est bien plus important que le temps issu de la simulation. Nous avons alors pensé que des couplages capacitifs pourraient ralentir la structure, nous avons effectué une mesure de la phototension afin de vérifier cette hypothèse (V-2-3).

La figure V.23 résume les caractéristiques des comparateurs déclenchés.

Sensibilité		Justesse des codes	Stabilité des codes dans le temps	Temps de calcul analogique
Nx gris	Pourcentage			
50 ( / 256)	19 %	50 %	99,3 %	600 ns

**figure V.23 : Caractéristiques des codes des comparateurs déclenchés**

### V-2-3 Mesure de la phototension

Nous avons pu mesurer la phototension grâce à la réparation du circuit, et celle-ci s'est avérée très intéressante comme le montre la figure V.24.

La courbe (1) de l'oscillogramme est le signal "décl" de déclenchement de la comparaison, et (2) est la sortie analogique du pixel (point SPix de la figure V.26).

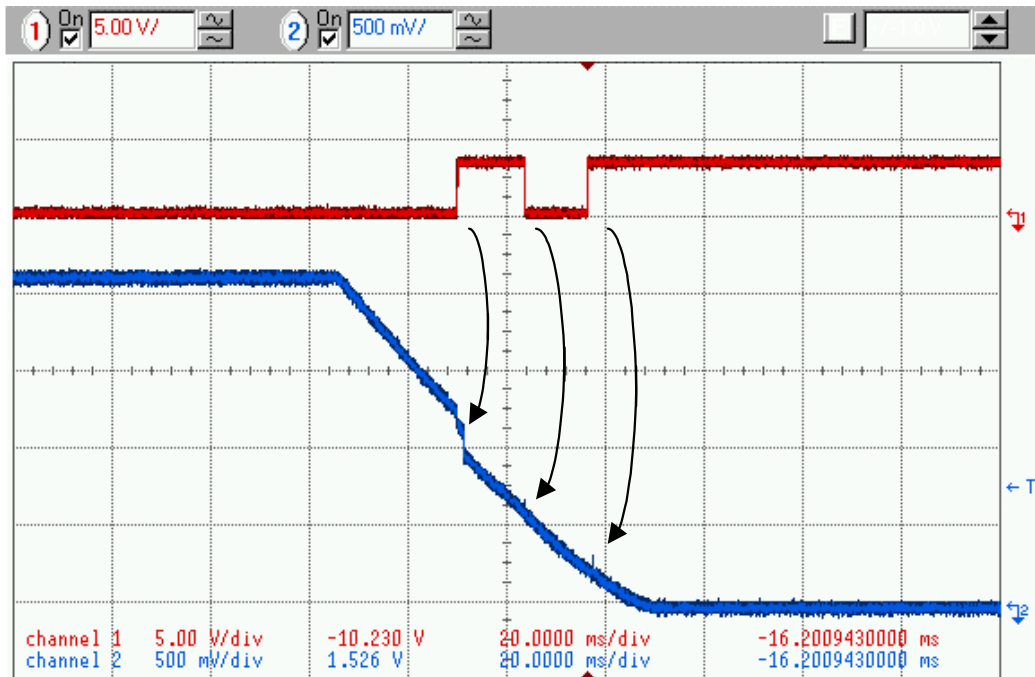


figure V.24 : Mesure du couplage capacitif dans le pixel

Nous constatons, après l'initialisation de la photodiode par le signal "Pr", la décharge de la tension Spix (2), qui a une allure classique de décharge de photodiode à travers un transistor suiveur. Au moment du déclenchement du comparateur, un effet parasite perturbe le signal, après quoi la décharge de la phototension continue.

L'offset généré a une amplitude de 230 mV, c'est à dire 9 % de la dynamique du signal et dure 1 ms pour une exposition en conditions normales.

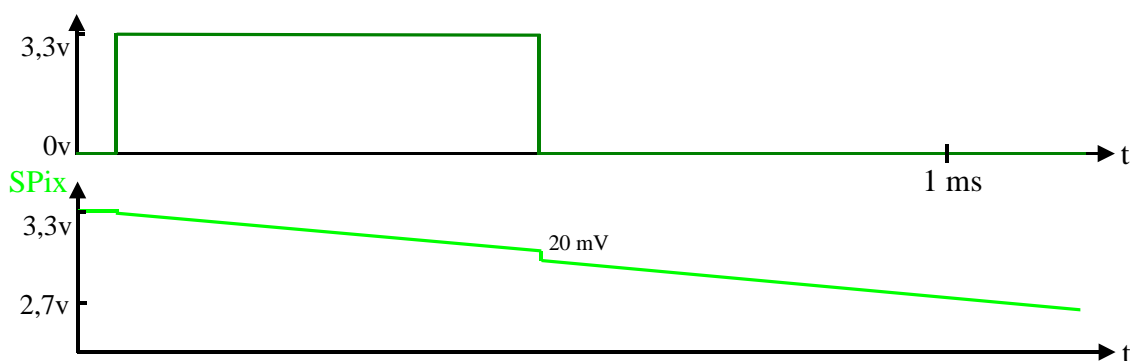


figure V.25 : Simulation du couplage capacitif dans le pixel

Cet effet de couplage existe en simulation, mais la valeur de l'offset est alors inférieure à la valeur mesurée (20 mV en simulation "corners", pire cas). Nous en concluons donc que le comportement de cette structure est peu prévisible, et qu'elle n'est pas bien adaptée à notre application.

### V-3 Interprétation des résultats

Sur la mesure précédente (figure V.24), le deuxième front montant du signal "décl" génère peu de perturbations, ce qui montre que ce n'est pas la propagation du signal "décl" qui perturbe la tension de sortie. L'effet parasite provient en effet du basculement des comparateurs.

Au niveau du pixel, les couplages capacitifs se font en fait au niveau des transistors suiveurs de type P, comme le schématise la figure V.26. Le basculement de chaque comparateur influe donc sur les comparateurs connectés à même photodiode (choc en retour), entraînant un retard et une perte de sensibilité de la structure. En effet, les comparateurs les plus lents à commuter sont ceux dont les entrées sont proches.

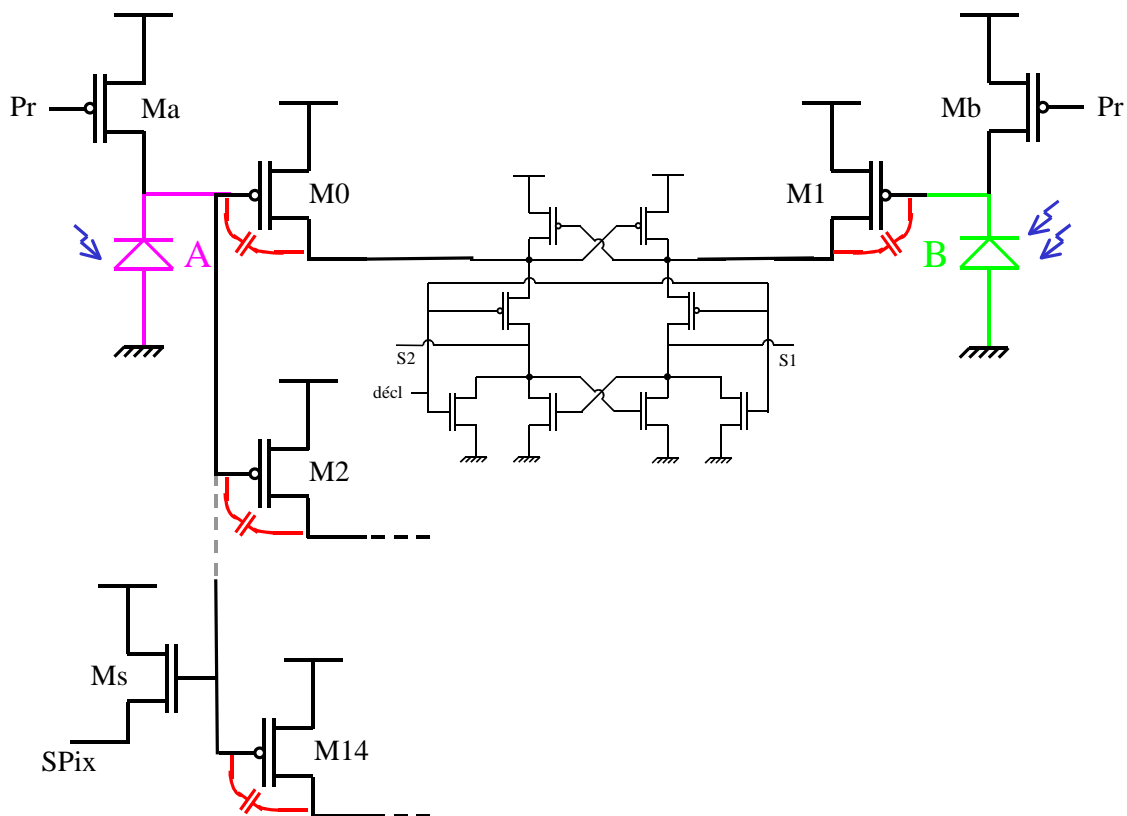


figure V.26 : Capacités parasites dans les pixels

## V-4 Comparaison des comparateurs par la simulation

Ne pouvant effectuer de mesures sur les comparateurs en temps, nous les avons comparés par la simulation, en s'appuyant sur les mesures effectuées.

	Temps de calcul		Précision		Couplages	
	mesure	simulation	mesure	simulation	mesure	simulation
Comparateurs déclenchés	600 ns	940 ps	50 nx gris	500 nx gris	230 mV	20 mV
Comparateurs en temps	X	48 ps	X	1000 nx gris	X	0 V

figure V.27 : Comparaison des performances des comparateurs

Le tableau ci-dessus résume que le comparateur en temps a des performances simulées meilleures. Ce comportement plus juste provient de sa conversion analogique / numérique et de sa structure numérique de comparaison. La conversion en numérique évite en effet les phénomènes de couplages, ils sont inférieurs au microvolt au niveau de la photodiode (en simulation), et inexistant à la sortie de la chaîne d'inverseurs (en mesure). Cette mesure est possible malgré le non fonctionnement du codage.

De plus, nous avons dès le départ estimé que les comparateurs déclenchés (analogiques) reposent sur une structure plus sensible aux parasites. Cette hypothèse repose sur le fait que les comparateurs ont un grand gain amplifiant de faibles différences, d'autant plus sensibles à de petites variations. Cette sensibilité tend à décorrélérer les valeurs simulées et mesurées sur ces comparateurs déclenchés. Comme les comparateurs en temps ne présentent pas ce couplage et ont des performances simulées meilleures, nous concluons qu'ils constituent une solution intéressante.

## V-5 Conclusion sur les matrices de codage

Nous avons évalué l'implantation de la "Census Transform" qui compose la phase de codage de l'algorithme LTS. Il s'avère que le codage produit près de 50 % d'erreur avec les comparateurs déclenchés, sur des images de faible dynamique, ce qui engendrerait un champ de vecteur dense peu fiable. Cependant, les comparateurs en temps constituent une solution

plus précise et plus robuste, compte tenu de la binarisation de l'information et de leur structure plus "numérique". Malheureusement, des erreurs de conception nous ont empêché de chiffrer leur précision. Ces mesures seraient d'autant plus intéressantes que nous avons caractérisé un imageur - basé sur la même structure de conversion binaire dans le pixel - qui donne des résultats encourageants.



# CONCLUSION GENERALE ET PERSPECTIVES

## I. Conclusion générale

Les capteurs en technologie CMOS ont créé un nouveau marché correspondant à l'avènement du multimédia à la fin des années 1990, ils sont des composants très utilisés de nos jours. Leurs caractéristiques optiques, ainsi que leur coût, les ont en effet destinés au marché "grand public". L'association de l'électronique et de l'optique permet aussi d'intégrer des fonctionnalités d'acquisition et de traitement sur le même substrat, ce qui confère de la compacité au système de vision. Cette intégration fait l'objet de recherches afin de proposer des solutions offrant de meilleures caractéristiques que les systèmes classiques, composés d'un capteur standard et de périphériques effectuant le traitement.

Cette thèse prospective au laboratoire porte sur l'étude et la conception de capteurs d'images et de rétines CMOS. Pour découvrir le domaine des circuits de vision, nous avons conçu et caractérisé un capteur d'images. Sa structure n'est pas celle d'un capteur standard APS, car ils souffrent d'un bruit de colonne (dominant) forçant l'utilisation de structures de réduction de bruit. Ce capteur est en fait un DPS ("Digital Pixel Sensor"), qui a la propriété de convertir numériquement l'information dans le pixel. Contrairement à ce DPS, pour des raisons de bande passante et de surface, les capteurs standards actifs APS acquièrent les images ligne par ligne, ce qui est source d'erreurs lorsque la scène n'est pas statique. Le capteur a été réalisé dans le but de caractériser une structure générant moins de perturbations, dont l'acquisition est simultanée sur tous les pixels. L'utilisation d'une structure ayant ces propriétés est importante pour une rétine d'estimation du mouvement. Classiquement, les rétines n'ont en effet pas de techniques de réduction de bruit.

Nous avons ensuite conçu une rétine qui acquiert l'image de manière synchrone sur tous les pixels, et en effectue un traitement. Ce traitement est une estimation de mouvement, calculée à partir d'une méthode robuste et efficace, mais qui n'a jamais été implantée dans une rétine. Cette méthode a pour principe la mise en correspondance de pixels à travers deux

images. Elle se compose de deux phases, un codage des pixels des deux images et une recherche de ces codes. Alors que l'implantation de manière classique requiert plusieurs composants (capteur et périphériques de traitement), une consommation importante et un calcul long, nous proposons un système composé d'un unique "capteur intelligent" apportant de meilleures performances dans ces domaines. Cette rétine comprend des pixels comportant une électronique de traitement dans le plan focal. Dès l'acquisition de l'image, les valeurs analogiques présentes sur les photodiodes sont comparées, et le résultat des comparaisons est stocké dans les pixels. Cette étape correspond à la phase de codage des pixels. Ensuite, un bloc de calcul numérique parallèle effectue la phase de recherche des codes. Ce bloc de calcul prend place en bas de la matrice de réception-codage, et traite les codes lors de leur transfert vers la sortie de la rétine.

Nous avons conçu et fait fabriquer un circuit de test comprenant la matrice de photoréception et de codage. Les comparateurs analogiques des pixels constituent en effet le point sensible de cette rétine. A cause d'une erreur de conception, nous n'avons malheureusement pu tester qu'une des deux structures implantées. Les mesures ont montré que d'importants couplages capacitifs perturbent la précision estimée.

Après avoir étudié un circuit imageur et une rétine dédiée à l'estimation du mouvement, nous pouvons tirer des conclusions quant à la caractérisation de circuits. Le fait de concevoir plusieurs circuits ne laisse pas le temps de caractériser pleinement un circuit et de prendre assez de recul pour en tirer tous les résultats et toutes les conclusions. En effet, même à l'issue de cette thèse, nous avons encore des idées de mesures originales avec l'imageur DPS. C'est en fait l'accumulation de ces mesures qui est constructive. Nous estimons que le temps de caractérisation d'un circuit de ce type est bien supérieur à celui requis pour son étude et sa conception.

## **II. Perspectives**

Suite à ce travail, plusieurs perspectives apparaissent. La première concerne la rétine que nous avons conçue. La sensibilité des comparateurs déclenchés est limitée à cause de couplages capacitifs. Nous voudrions tester la sensibilité des comparateurs en temps, leurs structures numériques promettent en effet un codage bien plus précis. Cependant, face à une trop grande précision, nous nous attendons à rencontrer d'autres problèmes. En effet, les simulations indiquent que la sensibilité peut être inférieure à un niveau de gris. Compte tenu de la



dispersion présente (même dans les imageurs), nous pensons qu'une sensibilité trop grande peut entraîner des erreurs lors de calculs électroniques (par exemple les zones homogènes sont captées comme texturées). Si les mesures confirment cette hypothèse, une solution consisterait à concevoir des comparateurs utilisant une logique ternaire [Mar-97]. Cette logique permettrait de classer les comparaisons des pixels A et B en trois classes : A est supérieur à B, A est inférieur à B, le résultat est dans la zone d'incertitude. Aussi, un seuil de comparaison programmable est une solution pour s'affranchir de niveaux de perturbation trop forts. Une autre solution consisterait à intégrer un système de réduction du bruit fixe (double échantillonnage) dans les comparateurs. Notons que les rétines ne proposent généralement pas de tels systèmes, et effectuent des calculs électroniques sur des pixels fortement bruités.

La deuxième perspective porte sur l'utilisation du flot optique en robotique. Le circuit que nous avons conçu génère en effet un flot optique dense, c'est à dire un vecteur mouvement par pixel. Cependant, cette information brute et (trop) riche n'est pas l'information couramment utilisée en robotique pour interpréter la scène. En effet, un des intérêts du flot optique est qu'il donne une information utile sur les objets en mouvement. Une phase de régularisation du flot dense est alors nécessaire, il reste à chiffrer sa nature et sa complexité.

La troisième perspective que nous dégagons de ce travail concerne les circuits dédiés à l'estimation du mouvement. Nous pensons qu'une rétine manque crucialement de souplesse, principalement à cause de sa nature spécifique (ASIC), et de son utilisation figée. En plus, la caractérisation est souvent difficile à cause de problèmes de contrôlabilité et d'observabilité. Nous avons aussi constaté que les capteurs standards souffrent de dispersions non négligeables, et que les rétines n'ont pas de structure de réduction du bruit. L'influence du bruit doit donc être minimisée a priori lors de la conception. Pour ces raisons, nous pensons qu'il est plus intéressant de s'orienter vers des circuits qui se situent entre l'imageur et la rétine. Ces circuits sont en fait des imageurs dédiés à une application, qui requièrent des images ayant des caractéristiques précises. Par exemple, pour l'estimation du mouvement, il est judicieux d'étudier un imageur faisant des acquisitions rapides, simultanées sur toute l'image, et peu bruités.

Les seuls avantages des rétines de nos jours sont leur rapidité de calcul et leur place occupée (système sur puce). Cependant, les périphériques de traitement (FPGA ou DSP) ont des puissances de calcul de plus en plus importantes et une consommation de plus en plus faible. Cette évolution pousse les rétines à effectuer des traitements de plus en plus haut niveau sur

l'image pour rester une alternative intéressante. Cependant, ceci va à l'encontre de l'approche "pixel" des rétines, qui ne peuvent que très mal traiter l'image dans sa globalité. Nous pensons donc que les rétines de pré-traitement de l'image (ASIC) peuvent laisser place à un circuit imageur dédié à une fonctionnalité spécifique (ASI: "Application Specific Imager"). La conception d'imageurs dédiés à une application est en effet une perspective intéressante. Le traitement peut être effectué par des périphériques reprogrammables, souples d'utilisation et très performants.

## REFERENCES BIBLIOGRAPHIQUES

[Abc-03]

*AbcElectronique*

<http://www.abcelectronique.com>

[Ari-02]

*Miguel Arias-Estrada, Eduardo Rodríguez-Palacios*

**An FPGA Co-processor for Real-Time Visual Tracking**

Field Programmable Logic and Applications, Montpellier, 2002

[Ban-97]

*J. Banks*

**Non-Parametric Techniques For Fast And Robust Stereo Matching**

IEEE TENCON '97, Brisbane, Australia, December 1997

[Bar-94]

*J.L. Barron*

**Performance of Optical Flow Techniques**

International Journal of Computer Vision, 1994

[Bar-02]

*M. Barbaro, P.Y. Burgi, A. Mortara, P. Nussbaum, F. Heitger*

**A 100x100 Pixel Silicon Retina for Gradient Extraction With Steering Filter Capabilities and Temporal Output Coding**

IEEE Journal of Solid-State Circuits, vol. 37, N° 2, feb 2002

[Bea-95]

*S.S. Beauchemin*

**The Computation of Optical Flow**

ACM Computing Surveys, Vol 27, No 3, pp 433-467, 1995

[Ber-97a]

*T. M. Bernard*

**PVLSAR2.2 Monolithic Sensor**

Nov 97

[Ber-97b]

*T. M. Bernard*

**Rétines artificielles: Quelle intelligence au sein du pixel ?**

Calculateurs Parallèles, special issue on {FPGAs and Smart Sensors}, Hermès, pp77-108, mars 97

[Ber-99]

*C. Bernard*

**Ondelettes et problèmes mal posés : la mesure du flot optique et l'interpolation irrégulière**

Ecole Polytechnique de Paris, Novembre 1999

[Ber-00]

*J.D. Berst et al.*

**A Monolithic Active Pixel Sensor for Charged Particle Tracking and Imaging using Standard VLSI CMOS Technology**

PIXEL 2000, Genova, June 2000

[Bou-96]

*F. Bouffault*

**Conception et réalisation de chaînes d'acquisitions vidéo rapides**

Thèse, Université de Nice-Sophia Antipolis, 1996

[Car-02]

*P. B. Catrysse, B. A. Wandell*

**Optical efficiency of image sensor pixels**

Journal Optical Society of America Vol. 19, No. 8 August 2002, pp1610-1620

[Cat-00]

*P. Catrysse, X. Liu, A. El Gamal*

**QE Reduction due to Pixel Vignetting in CMOS Image Sensors**

Proc. SPIE Vol. 3965 Electronic Imaging '2000 conference, California, January 2000

[Cat-02]

*G. Cathebras, D. Navarro, O. Aubreton, B. Bellach, P. Gorria, B. Lamalle, L. Lew Yan Voon*

**A Continuous Time Pattern Recognition Retina**

IEEE European Solid-State Circuits Conference, Septembre 2002, Italie, pp 719-722

[Cav-98]

*C. Cavadore*

**Conception et caractérisation de capteurs d'images à pixel actifs CMOS-APS**

Thèse SUPAERO - THOMSON-TCS 1998

[Che-00]

*T. Chen, P.B. Catrysse, A. El Gamal, B.A. Wandell*

**How Small Should Pixel Size Be ?**

EE 370 Information Systems Seminar Winter 2000

[Com-01]

*F. Comby*

**Estimation du mouvement apparent majoritaire dans une séquence d'images vidéo par accumulation de votes bimodaux sur un histogramme approché**

Thèse, Université de Montpellier II, Déc 2001

[Cou-00]

*J. Coulombe*

**Capteur d'images intégré en mode courant et à résolution variable**

Mémoire, Ecole Polytechnique de Montréal, Nov 2000

[Deg-00]

*Y. Degerli, F. Lavernhe, P. Magnan, Jean A. Farré*

**Analysis and Reduction of Signal Readout Circuitry Temporal Noise in CMOS Image Sensors for Low-Light Levels**

IEEE TRANS. ON ELECTRON DEVICES, VOL. 47, NO. 5, pp 949-962, MAY 2000

[Deu-98b]

*R.A. Deutchmann, C. Koch*

**Compact Analog VLSI 2-D Velocity Sensor**

IEEE International Conference on Intelligent Vehicles 1998

[Die-97a]

*B. Dierickx, G. Meynants, D. Scheffer*

**Near 100% Fill Factor Standard CMOS Active Pixel**

IEEE CCD & AIS workshop, page p. P1, Jun 1997

[Eti-99b]

*R. Etienne-Cummings*

**A High Density Focal-Plane Image Processing Array**

Conf. Information Sciences and Systems, Baltimore, MD, 1999

[Fil-00]

*Fillfactory*

**Fuga1000 : 11.5 mm (type 2/3 “) Logarithmic, Non-integrating, Random- Accessible CMOS Area Image Sensor**

<http://www.fillfactory.com>

[Fos-97]

*E.R. Fossum*

**CMOS Image Sensors: Electronic Camera-On-A-Chip**

IEEE Trans. On Electronic Devices, Vol 44, N° 10, oct 97

[Fuj-00]

*I.L. Fujimori, C.C. Wang, C.G. Sodini*

**A 256x256 CMOS Differential Passive Pixel Imager with FPN Reduction Techniques**

IEEE Journal of Solid-State Circuits, Vol 35 december 2000

[Fun-97]

*E. Funatsu, Y. Nitta, Y. Miyake, T. Toyoda, J. Ohta, K. Kyuma*

**An Artificial Retina Chip with Current-Mode Focal Plane Image Processing Functions**

IEEE Transactions on Electron Devices, Vol. 44, N°. 10, pp 1777-1782, October 1997

[Gam-98]

*A. El Gamal, B. Fowler, H. Min, X. Liu*

**Modeling and Estimation of FPN Components in CMOS Image Sensors**

*SPIE*, vol. 3301, pp. 168-177, April 1998

[Gar-88]

*P. Garda, A. Reichart, H. Rodriguez, F. Devos, B. Zavidovique*

**Une rétine électronique automate cellulaire**

Traitement du Signal, vol. 5, pp. 435-449, 1988

[Gen-99]

*A. Gentile, J. L. Cruz-Rivera, D. Scott Wills*

**Real-Time Image Processing on a Focal Plane SIMD Array**

Workshop on Parallel and Distributed Real-Time Systems, 1999

[Gia-00]

*A. Giachetti*

**Matching Techniques to compute Image Motion**

Image and Vision Computing, pp 245-258, Italy, march 2000

[Gor-99]

*P. Gorria, B. Lamalle, G. Cathebras*

**Dispositif d'intercorrélation d'une image**

Brevet n° 99 00008, 4 juin 1999

[Gou-95]

*Eric Goujou*

**Etude & Réalisation d'un capteur d'images à temps d'exposition variable en CMOS**

Thèse, Université de bourgogne, Juillet 1995

[Goy-92]

*J. Goy*

**Etude, Conception, et réalisation d'un capteur d'image APS en technologie standard CMOS pour des applications faible flu de type viseur d'étoiles**

Thèse Institut National Polytechnique de Grenoble, mars 1992

[Ham-01]

*T. Hamamoto, K. Aizawa*

**A Computational Image Sensor with Adaptive Pixel-Based Integration Time**

IEEE Journal of Solid-State Circuits, Vol 36 April 2001

[Hig-98]

*C.M. Higgins*

**An Integrated Vision Sensor for the Computation of Optical Flow Singular Points**

Advances in Neural Information Processing Systems 11, MIT Press, Denver, CO, 1998

[Hig-00]

*C. M. Higgins, S. Korrapati*

**An Analog VLSI Motion Energy Sensor Based on the Adelson-Bergen Algorithm**

ICSC Symp. On Biologically-Inspired Systems, 12-15 dec 2000

[Hor-81]

*B. K. P. Horn, B. G. Schunk*

**Determining Optical Flow**

Artificial Intelligence, vol. 17, pp. 185 203, 1981

[Ind-99]

*G. Indiveri*

**Neuromorphic Analog VLSI Sensor for Visual Tracking : Circuits and Application Examples**

IEEE Trans. On Circuits and Systems II : Analog and Digital Signal Processing, Vol46, N° 11, Nov. 1999.

[Jan-02]

*J. Janesick*

**Dueling detectors: CCD or CMOS ?**

Spie OE-Magazine, pp30-33, February 2002

[Jia-99]

*H.C. Jiang and C.Y. Wu*

**A 2-D Velocity- and Direction-Selective Sensor with BJT-Based Silicon Retina and Temporal Zero-Crossing Detector**

IEEE JOURNAL OF SOLID-STATE CIRCUITS, VOL. 34, NO. 2, FEBRUARY 1999

[Joh-92]

*John M. Senior*

**Optical Fiber Communications – Principles and Practice**

Prentice/Hall International, ISBN 0-13-638222-3, 1992

[Joh-96]

*D. A. Johns, K. Martin*

**Analog Integrated Circuit Design**

Ed Wiley, ISBN 0-471-14448-7

[Kav-97]

*S. Kavadias, B. Dierickx, D. Scheffer, A. Alaerts, D. Uwaerts, J. Bogaerts*

**A Logarithmic Response CMOS Image Sensor with On-Chip Calibration**

IEEE Journal of Solid-State Circuits, VOL. 35, NO. 8, pp 1146-1152, August 2000

[Kem-96]

*S. Kemeny, B. Pain, R. Panicacci, L. Matthies, and E.R. Fossum*

**Programmable Multiresolution CMOS Active Pixel Sensor**

SPIE Vol. 2654 Solid State Sensor Arrays and CCD Cameras 1996

[Kle-01]

*S. Kleinfelder, S. Lim, X. Liu, A. El Gamal*

**A 10000 Frames/s CMOS Digital Pixel Sensor**

IEEE Journal of Solid-State Circuits, Vol 36, N° 12, Dec 2001

[Kom-97]

*T. Komuro, I. Ishii, M. Ishikawa*

**Vision Chip Architecture Using General-Purpose Processing Elements for 1ms Vision System**

4th IEEE Workshop on Computer Architecture for Machine Perception, pp.276-279, Cambridge, USA, October 1997

[Kon-98a]

*K. Konolige*

**Small Vision Systems**

<http://www.ai.sri.com/~konolige/svs/svm.htm>

[Kra-97]

*J. Kramer, R. Sarpeshkar, C. Koch*

**Pulse-Based Analog Velocity Sensors**

IEEE Trans. On Circuits and Systems II : Analog and Digital Signal Processing, Vol 44, pp 86-101, 1997

[Lam-00]

*B. Lamalle, P. Gorria, L.F.C Lex Yan Voon and G. Cathebras*

**The Study and Design of a Programmable Silicon Retina for Real-Time Pattern Recognition**

SPIE/EOS Symposium on Applied Photonics - EUROPTO'2000, Optical Diagnostics for Industrial Applications, Vol. 4076, pp. 225-234, Glasgow, Scotland, UK, 22-25 May 2000.

[Loo-01]

*M. Loose, K. Meier, J. Schemmel*

**A self-Calibrating Single-Chip CMOS Camera with Logarithmic Response**

IEEE Journal of Solid-State Circuits, Vol 36 April 2001

[Luc-81]

*B.D. Lucas, T. Kanade*

**An Iterative Image Registration Technique with an Application to Stereo Vision**

Proceedings of Imaging Understanding Workshop, pp. 121-130 (1981)

[Ma-99]

*S. Y. Ma, L. G. Chen*

**A Single Chip CMOS APS Camera with Direct Frame Difference Output**

IEEE Journal Solid-State Circuits, vol. 34, pp. 1415-1418, Oct 1999

[Mar-95]

*J. Martin, J. L. Crowley*

**Experimental Comparison of Correlation Techniques**

IAS-4, International Conference on Intelligent Autonomous Systems, Karlsruhe, March 1995.

[Mar-97]

*R. Mariani, F. Pessolano, R. Saletti*

**A new CMOS Ternary Logic Design for Low-Power Low-Voltage Circuits**

7<sup>th</sup> Seventh International Workshop PATMOS, Belgique Septembre 1997

[Mea-89]

*C. Mead*

**Analog VLSI and Neural Systems**

Addison-Wesley Publishing Company - ISBN: 0-201-05992-4



[Mey-98b]

*G. Meynants, B. Dierickx, D. Scheffer, J. Vlummens*

**A Wide Dynamic Range CMOS Stereo Camera**

Advanced Microsystems for Automotive Applications March 1998, Germany

[Mey-01]

*G. Meynants, B. Dierickx, D. Uwaerts, J. Bogaerts*

**Fixed pattern noise suppression by a differential readout chain for a radiation tolerant image sensor**

2001 IEEE Workshop on CCDs and AISs

[Moi-97a]

*A. Moini*

**Vision Chip or Seeing Silicon**

<http://www.eleceng.adelaide.edu.au/Groups/GAAS/Bugeye/visionchips>

[Nez-00]

*T. Nezuka, M. Hoshino, M. Ikeda, K. Asada*

**A Position Detection Sensor for 3-D Measurement**

26th European Solid-State Circuits Conference, Stockholm, Sweden, 19-21 September 2000

[Ni-93]

*Y. Ni, Y.M. Zhu, F. Devos, B. Arion*

**Yet Another Analog 2D Gaussian Convolver**

ISCAS93, Chicago, 3-6 May, 1993, pp. 192-5, Vol. 1

[Ni-01]

*Y. Ni, K. Matou*

**A CMOS Log Image Sensor with on-chip FPN Compensation**

28<sup>th</sup> European Solid-State Circuit Conference, 2001, Villach, Austria

[Ni-02]

*Y. Ni, X.L. Yan*

**CMOS Active Differential Imaging Device with Single in-pixel Analog Memory**

28<sup>th</sup> European Solid-State Circuit Conference, 2002, Firenze, Italy

[Ogi-97]

*W. Ogiers*

**Survey of CMOS imagers, issue 1.0**

IMEC report P60280-MS-RP-002, August 1997

[Pai-99]

*F. Paillet D. Mercier T. M. Bernard*

**Second Generation Programmable Artificial Retina**

IEEE ASIC/SOC Conf., pp.304-309, Sept. 1999

[Per-94]

*O. Perron*

**Rétine en silicium pour la compression d'images**

Mémoire Diplôme d'ingénieur CNAM, 1994

[Pro-03]

*Production Sextant*

<http://www.sextantvideo.com>

[Raz-92]

*B. Razavi, B.A. Wooley*

**Design Techniques for High-Speed, High-Resolution Comparators**

IEEE Journal of Solid-State Circuits, Vol 27, N° 12, dec 92

[Row-97]

*T. Röwekamp, L. Peters*

**A Real-Time Smart Sensor System for Visual Motion Estimation**

European Design and Test Conference, pp 613, 1997

[San-00]

*G. Sandini, P. Questa, D. Scheffer, B. Dierickx, A. Mannucci*

**A Retina-Like CMOS Sensor And Its Applications**

IEEE SAM Workshop, March 16-17, 2000, Cambridge, USA

[Sar-03]

*Sarnoff Corporation*

[http://www.sarnoff.com/products\\_services/imagers/cmos\\_imager.asp](http://www.sarnoff.com/products_services/imagers/cmos_imager.asp)

[Sic-99]

*G. Sicard*

**De la biologie au silicium: une rétine bio-inspirée analogique pour un capteur de vision "intelligent" adaptatif**

Thèse, INPG, 1999

[Sme-02]

*F. Smeraldi*

**Ranklets: orientation selective non-parametric features applied to face detection**

Proceedings of SSAB'02, pp 1–4, Lund (Sweden), March 2002

[Sto-98]

*A. Stocker, R. Douglas*

**Computation of Smooth Optical Flow in a Feedback Connected Analog Network**

Advances in Neural Information Processing Systems 11, 1998, denver, USA

[Tab-01]

*M. Tabet, R. Hornsey*

**Dual Mode Active Pixel Sensor with Focal Plane Edge Detection**

IEEE Workshop on CCDs and Advanced Image Sensors, Crystal Bay, Nevada, June 2001

[Tia-00]

*H. Tian*

**Noise Analysis in CMOS Image Sensors**

Ph.D Thesis, Stanford University, august 2000

[Tia-01b]

*H. Tian, B. Fowler, A. El Gamal*

**Analysis of temporal noise in cmos photodiode active pixel sensor**

IEEE Journal of Solid-State Circuits, Vol 36 January 2001

[Tur-00]

*E. Turenne, M. Arias-Estrada, D. Poussart*

**Motion Computation using a smart sensor**

2000

[Voo-01a]

*L.F.C Lew Yan Voon, B. Lamalle, P. Gorria, G. Cathebras, B. Bellach and D. Navarro*

**Real-Time Pattern Recognition Retina in CMOS Technology**

International Conference on Quality Control by Artificial Vision pp. 238-242, May 2001.

[Voo-01b]

*L.F.C Lew Yan Voon, B. Lamalle, P. Gorria, G. Cathebras, B. Bellach, F. Meriaudeau, D. Navarro*

**Optical Correlation allows silicon retinas to recognize patterns stored in memory**

Revue SPIE Optical Engineering Magazine, N°7, juillet 2001

[Wan-63]

*Frank Wanlass – Fairchild Semiconductor*

**Low Stand-By Power Complementary Field Effect Circuitry**

U.S. Patent # 3,356,858, 1963 – 1967

[Yuk-85]

*A. Yukawa*

**A CMOS 8-Bit High-Speed A/D Converter IC**

IEEE Journal of Solid-State Circuits, Vol 20, N° 3, June 1985

[Zab-94]

*R. Zabih, J. Woodfill*

**Non-Parametric Local transforms For Computing Visual Correspondance**

3rd European Conference on Computer Vision, pages 151-158, 1994



# ANNEXE

Schéma du pixel de la rétine d'intercorrélation

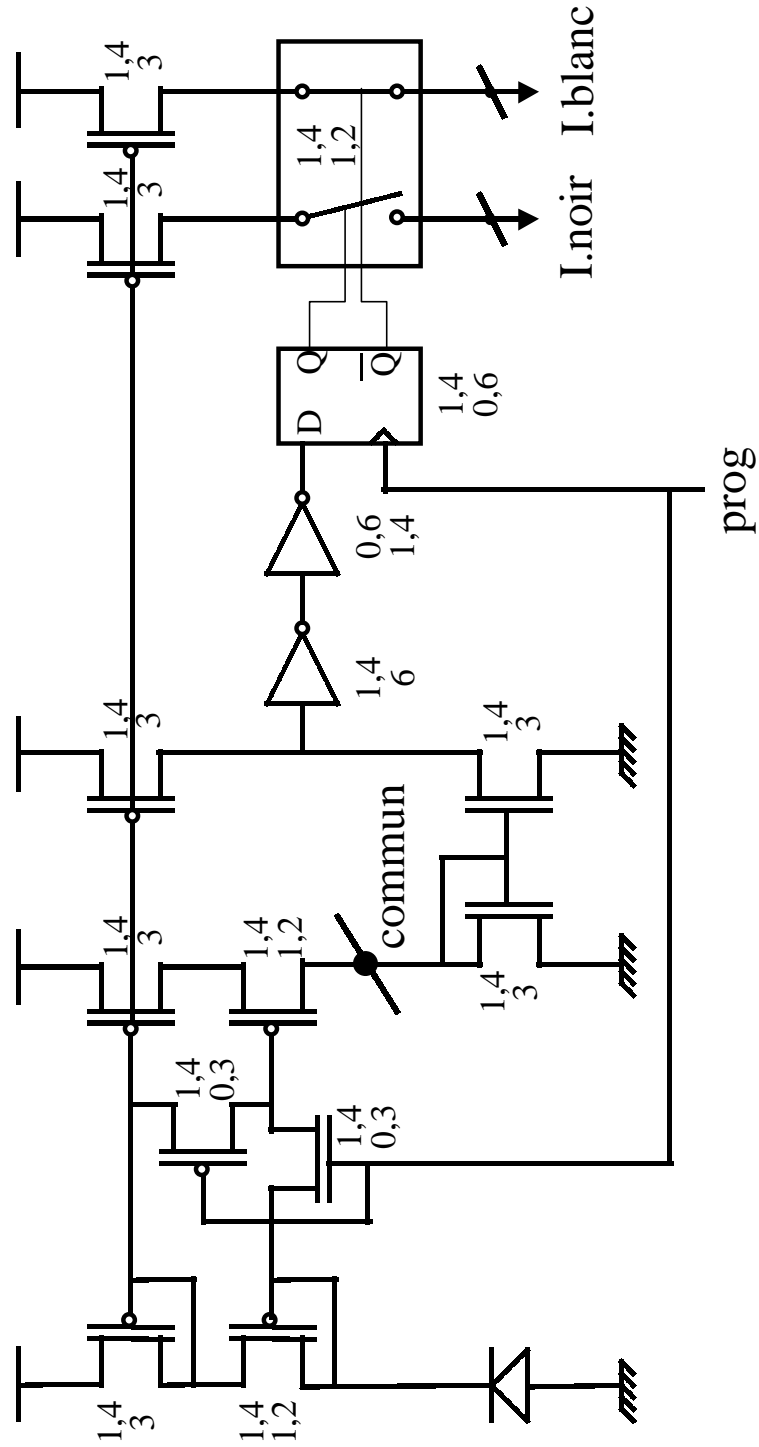
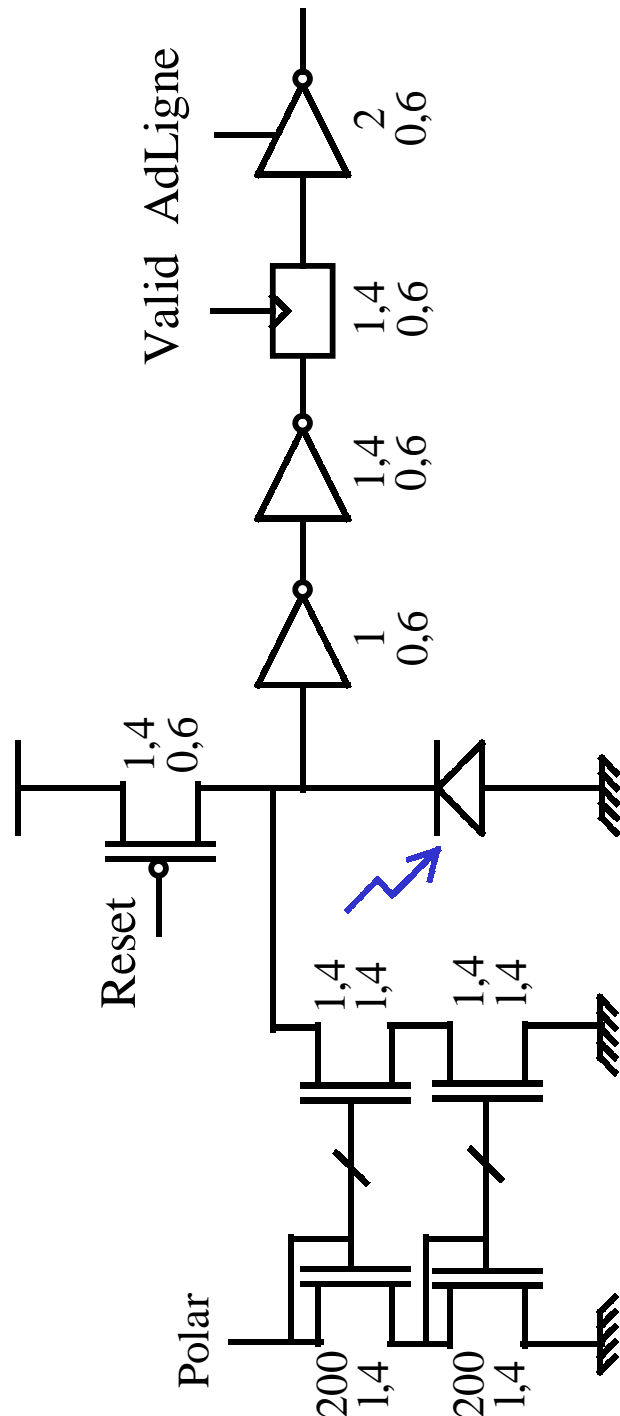




Schéma du pixel de l'imageur DPS



## Extrait du code VHDL : contrôleur VGA (partie logique de l'imageur DPS)

```

library ieee;
use ieee.std_logic_1164.all;
use ieee.numeric_std.all;

entity total is
  port (
    Reset,h,ResetX,ValidX,TypeReset,TypeValide,
    Ld,Sck,VADLext,cpt_ad,RgbSynExt,ADLSynExt,
    VsynExt: in std_logic;
    outB : in std_logic_vector(212 downto 0);
    Sout: out std_logic_vector(3 downto 0);
    AdLig : out unsigned(7 downto 0);
    hsync,vsync,rgb,ValidMat,ValidMatB,ResetMat,Sel:
  out std_logic );
end total;

architecture comportement of total is
  signal rstmat:boolean;
  signal Vsyn:std_logic;
  signal SVsyn:std_logic;
  signal Hsyn:std_logic;
  signal temp : std_logic_vector(212 downto 0);
  signal clock : std_logic;
  signal RgbSync : std_logic;
  signal SRgbSync : std_logic;
  signal Video : std_logic;
  signal cptr:unsigned(7 downto 0);
  signal entree : std_logic_vector(212 downto 0);

begin
  entree<=not(outB);
  --- VGA ---
  hvsyn : process(h,Reset)
    variable cptr0      : integer ;
    variable cptr1      : integer ;
    variable cptr2      : integer ;
    variable cptrentree : integer ;
    variable ligne      : std_logic ;
    variable trame      : std_logic ;
  begin
    if Reset='0' then
      cptr0:=0; cptr1:=0; cptr2:=0;
      hsync<='0'; Hsyn<='0';
      ligne:='0'; trame:='0';
      vsync<='1'; Vsyn<='1';
      rgb<='0'; Video<='0'; RgbSync<='0';
    elsif (h='1' and h'event) then
      cptr1:=cptr1 + 1;
      if (cptr0=0) then
        vsync <= '0';
        Vsyn<='0';
        cptr0:=1;
      end if;
      if (ligne='0' and cptr1=94) then
        hsync<='1';
        Hsyn<='1';
        cptr1:=0;
        ligne:='1';
      end if;
      if (ligne='1' and cptr1=700) then
        hsync<='0';
        Hsyn<='0';
        cptr1:=0;
        ligne:='0';
        cptr2:=cptr2 + 1;
        cptrentree:=212;
      end if;
      if (trame='0' and cptr2=2) then
        vsync<='1';
        Vsyn<='1';
        cptr2:=0;
        trame:='1';
      end if;
      if (trame='1' and cptr2=517) then
        vsync<='0';
        Vsyn<='0';
        cptr2:=0;
        trame:='0';
      end if;
      if (cptr2>31 and cptr2<192) then
        if (ligne='1' and cptr1=20) then
          Video<='1';end if;
        if (ligne='1' and cptr1=680) then
          Video<='0';end if;
        if (ligne='1' and cptr1=1) then
          RgbSync<='1';end if;
        if (ligne='1' and cptr1=680) then
          RgbSync<='0';end if;
        else
          Video<='0';
          RgbSync<='0';
        end if;
        if (cptr2>31 and cptr2<191) then
          if (ligne='1' and cptr1>46 and cptr1<677)
            then
              if (cptrentree>-1) then
                rgb<=entree(cptrentree);
                cptrentree:=cptrentree-1;
              else
                rgb<='0';
                end if;--RgbSync<='0';end if;
            end if;
          end if;
        end if;
      end if;
    end process;
  end process;

```



Schéma du comparateur déclenché (rétine d'estimation du mouvement)

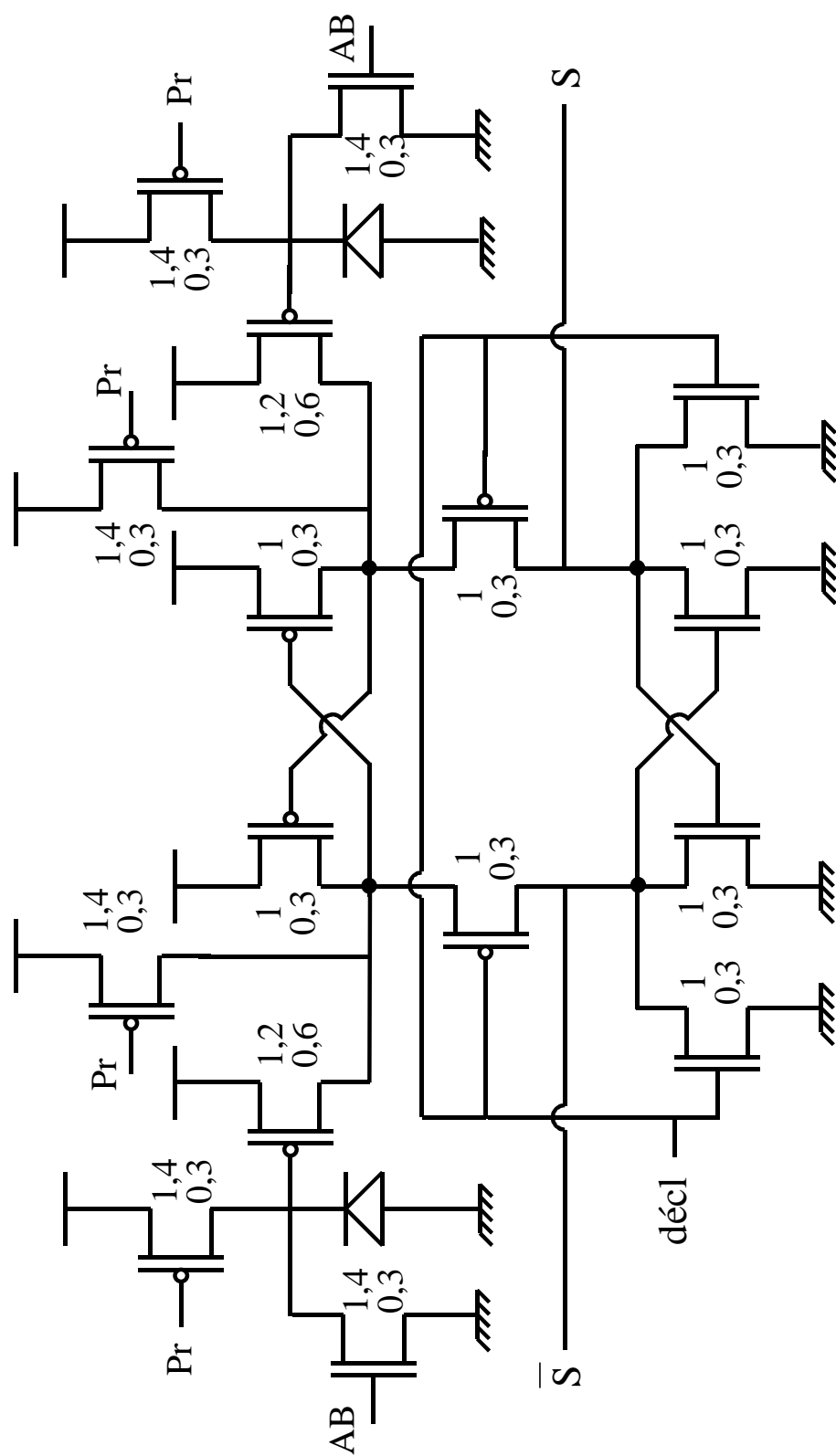
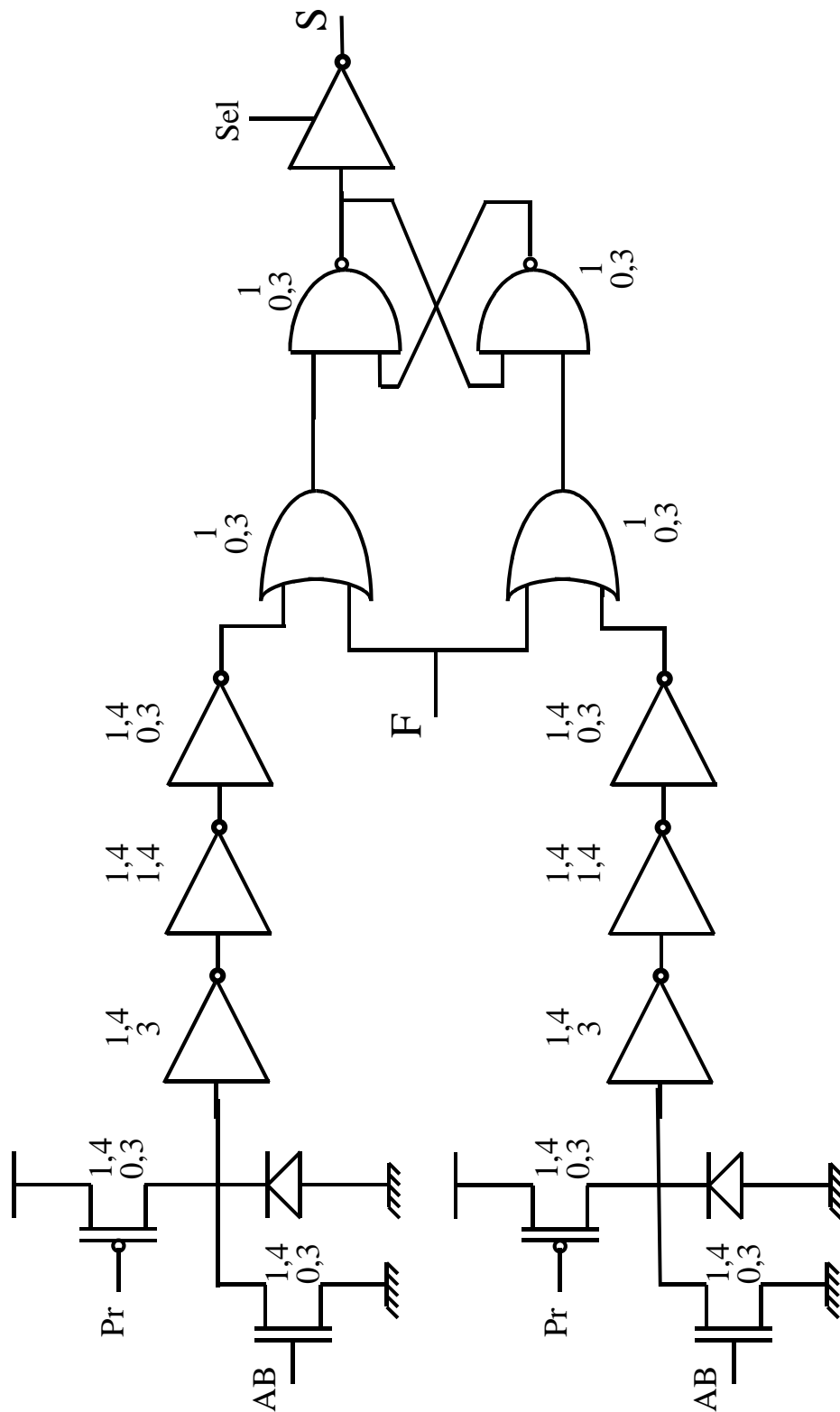


Schéma du comparateur en temps (rétine d'estimation du mouvement)





# **ARCHITECTURE ET CONCEPTION DE RETINES SILICIUM CMOS :**

## **APPLICATION A LA MESURE DU FLOT OPTIQUE**

Thèse de Doctorat, 17 Octobre 2003, David NAVARRO

**RESUME :** Le développement des technologies sub-microniques a permis un regain d'intérêt pour les capteurs d'images CMOS, qui inondent aujourd'hui le marché des capteurs. 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 consiste à associer, dans un même circuit – une rétine CMOS –, les photocapteurs et des fonctions de pré-traitement de l'image, permettant ainsi de répartir et d'optimiser le traitement. Ces rétines ont des performances en vitesse, en intégration et en consommation meilleures que les solutions classiques (capteurs puis traitements logiciels et/ou matériels). Cette thèse porte plus précisément sur l'intégration d'un algorithme d'estimation du mouvement en transposant le calcul numérique fortement itératif en une structure de calcul électronique. Après avoir réalisé un circuit permettant d'acquérir des connaissances dans le domaine des capteurs d'images CMOS, nous avons conçu un circuit de vision estimant le mouvement. Cette estimation de mouvement est basée sur une méthode robuste de mise en correspondance de blocs de pixels, comprenant une phase de pré-codage des pixels suivi d'une recherche de ce codage dans une fenêtre de destination potentielle. Cette approche est novatrice car elle propose une rétine CMOS pouvant traiter (électroniquement) des scènes fortement texturées, et à luminosité changeante, en s'appuyant sur une méthode jusqu'alors réservée aux approches numériques (FPGA, DSP) ou logicielles.

**MOTS CLES:** Rétine, CMOS, image, mouvement, estimation, flot optique, algorithme

---

## **ARCHITECTURE AND DESIGN OF CMOS SILICON RETINAS :**

### **APPLICATION TO OPTICAL FLOW MEASUREMENT**

**ABSTRACT :** The enhancement of sub-micronic technologies made CMOS image sensors an interesting alternative. Vision machines are traditionally based on architectures linked to an image sensor. The proposed approach in this work deals with the integration of photodetecting elements plus image processing circuits in a single chip. These smart sensors, also called retinas, enable to make the vision system both compact and optimised. The processing is indeed speeded up, and power consumption reduced, compared to classical solutions (sensor, then hardware or software processing).

More precisely, this thesis deals with the integration of a motion estimation algorithm, using electronic functions for computations, instead of a strong iterative numeric computation. After finishing a circuit for a correlation-based retina project, we first designed a standard sensor (an imager) to acquire knowledge in the image sensors field. Then, We designed a retina performing these computations in an electronic way. That motion estimation algorithm is based on a robust block matching technique, composed of a pixel coding step, then a search step of these codes in a potential destination window. That approach is new because it can treat textured scenes with changing luminosity, moreover we integrate a kind of algorithm that was only used in software or digital approaches (FPGA, DSP).

**KEY WORDS:** Retina, CMOS, image, movement, estimation, optical flow, algorithm