

Plus perçant que l'aigle – plus sensible que le hibou – plus perspicace que l'homme
Capteurs d'images semi-conducteurs dépassant toute limite



Leçon inaugurale de
Peter Seitz
Institut de Microtechnique, Université de Neuchâtel

Et Dieu dit: Qu'il ait la lumière; et la lumière fut.

Et les animaux de la terre développèrent le sens de la vue, et ils tirèrent profit de la lumière...

Depuis un milliard d'années, la nature a créé des systèmes de vision d'une énorme variété et aux performances absolument stupéfiantes. Le principal but de ces systèmes de vision naturelle est de former des images de l'environnement. Mais l'homme ne voulait pas seulement créer des images, il désirait aussi les enregistrer. Ainsi le film et les caméras furent inventés.

Au début des années 40, tout cela changea de manière dramatique, lorsque le physicien Russell Ohl découvrit que le silicium était extrêmement sensible à la lumière. D'ailleurs il fit ces observations en dépit des ordres donnés par ses supérieurs au laboratoire de recherche Bell de ne plus s'occuper du silicium; selon eux, cette occupation ne mènerait à rien, l'utilité du silicium étant considérée comme trop limitée.

Et ainsi a commencé un développement qui dure aujourd'hui encore et s'accélère même : Le développement de capteurs d'images à base de matériaux semi-conducteurs, d'abord avec le but d'approcher les performances des systèmes de vision naturelle. Entre temps, l'homme a réussi à créer des rétines artificielles, des capteurs d'images qui dépassent les performances des systèmes de vision naturelle. Les seules limites qui restent sont celles des lois fondamentales de la physique et de notre imagination.

Voilà donc l'enjeu de mon discours : La révolution des capteurs d'images à semi-conducteur.

Le silicium, le semi-conducteur sur lequel se base la microélectronique et notre entière société informatisée, n'est pas seulement capable d'effectuer des calculs – le silicium est aussi un détecteur de lumière extraordinairement efficace, en particulier dans la partie visible du spectre électromagnétique. Quelle aubaine ! Les capteurs d'images à semi-conducteur peuvent donc profiter directement du progrès incessant de la microélectronique. En effet, pendant presque un demi-siècle, ce progrès a vu un développement exponentiel, si prévisible qu'on parle de la « loi » de Moore : Chaque 18 mois, le nombre de transistors sur un circuit intégré se double.

Actuellement, le chiffre énorme de 300 milliards de transistors est fabriqué *chaque seconde* ! En même temps, la taille des transistors était réduite, et aujourd'hui une dizaine de transistors trouvent place sur une surface de 1 micromètre carré ! Pour cette raison il est devenu possible de fournir chaque capteur photonique individuel (on parle de « pixels ») avec de la fonctionnalité électronique. Ces « smart pixels » permettent donc de mettre à disposition des fonctions sur mesure : des formes aléatoires des parties photosensibles, des circuits analogiques et numériques, ainsi que des couches conductrices et structurées pour la création de champs électriques dans le semi-conducteur, avec lesquelles le transport des charges photo-générées peut être optimisé.

Pendant les dernières 15 années, des groupes de recherche se sont spécialisés dans le développement et l'application de familles de « smart pixels ». Auparavant, la fonctionnalité optoélectronique disponible était très limitée : Un pixel ne pouvait effectuer que quatre fonctions : (1) La conversion de la lumière incidente en un nombre équivalent de charges photo-générées à l'intérieur du semi-conducteur, (2) l'intégration et le stockage de ces charges, (3) l'accès séquentiel à l'information représentée par ces charges, et (4) l'amplification des signaux correspondants pour leur lecture et transmission.

Grâce aux efforts des chercheurs dans le monde entier, une grande gamme de « smart pixels » a été développée et appliquée, ce qui a permis une extension énorme de la fonctionnalité optoélectronique d'un capteur d'images. Chaque nouveau module de fonctionnalité de smart pixels était typiquement le résultat d'une ou plusieurs thèses, et l'ensemble représente une vraie « boîte à outils » modulaire pour la réalisation sur mesure de capteurs d'images optimisés pour des applications particulières.

La partie principale de mon discours est consacrée à la présentation de six exemples de smart pixels et leurs applications, des développements importants qui ont été faits ici à Neuchâtel, en étroite collaboration entre l'Institut de microtechnique (IMT) et le Centre suisse d'électronique et de microtechnique SA (CSEM) :

Un premier défi est lié à la grande variabilité dans la dynamique photométrique d'une scène naturelle. Tandis que nos yeux sont parfaitement adaptés à cette variabilité et nous permettent de percevoir une dynamique de plus de six ordres de grandeur, des caméras numériques conventionnelles n'approchent guère une dynamique de quatre ordres de grandeurs. Une des raisons pour la performance stupéfiante de nos yeux est leur réponse non linéaire, qui contraste celle, hautement linéaire, des capteurs d'images semi-conducteurs conventionnels. Le défi scientifique consiste donc en la conception d'un smart pixel avec réponse non linéaire, en utilisant qu'un petit nombre de composants électroniques dans chaque pixel. La solution inattendue, trouvée par un thésard de l'IMT, ne comprend qu'un transistor additionnel mais opéré d'une manière non conventionnelle. La fonction de ce transistor est de veiller sur le niveau d'illumination dans chaque pixel. Lorsque l'illumination locale est faible, la réponse du pixel reste linéaire, est la sensibilité du pixel reste inchangée. Par contre, au cas où ce niveau dépasse une certaine valeur – celle-là est même programmable – la réponse devient doucement logarithmique, manifestant une sensibilité typique de 30-40 mV par décade de lumière incidente. La dynamique de ce pixel, proprement dénommé « LinLog » dépasse de loin la performance de nos yeux : Une valeur de presque dix ordres de grandeur, 1 :10'000'000'000, a été mesurée, ce qui représente à ce jour le record du monde. Par conséquent, l'équipe de recherche responsable de ce succès a décidé de créer sa propre entreprise startup, Photonfocus SA à Lachen (SZ), qui

emploi actuellement plus de 20 collaborateurs. Plusieurs distinctions internationales étaient décernées à cette innovation, notamment le « Vision Award 2000 » et le « Photonics Circle of Excellence Award 2002 ».

Le défi suivant concerne la vitesse à laquelle notre système visuel peut acquérir des images. Sous de bonnes conditions d'illumination, cette vitesse correspond à un taux d'acquisition d'environ 60-70 images par seconde. Peut-on dépasser cette valeur avec une rétine artificielle ? En effet, cela ne pose aucun problème, en utilisant des modules de notre « boîte à outils » qui sont organisés et opérés de manière hautement parallèle et concomitante: dans l'arrangement matriciel des pixels, chaque colonne a son propre amplificateur qui fonctionne de façon indépendante des autres et simultanément. Il est donc possible de lire un nombre aléatoire de colonnes en même temps, en élevant la vitesse effective de lecture. Un bon exemple de cette technique est représenté par un capteur d'images en couleur qui a été développé pour l'entreprise suisse Weinberger AG, avec une résolution de presque un million de pixels et une vitesse d'acquisition et de lecture qui approche 1.7 milliard de pixels par seconde. Pendant un certain temps, ceci représentait le record du monde ! La caméra numérique ultra-rapide, construite à base de ce capteur, compte toujours parmi les plus rapides du monde, et elle représente actuellement le produit phare de l'entreprise Weinberger AG.

Bien que le système visuel de l'homme n'est pas le plus sensible que connaisse la nature, sa performance actuelle est quand même impressionnante : La petite quantité de 10 photons arrivant simultanément est déjà suffisante pour évoquer l'impression de lumière. Evidemment la limite physique inférieure est la détection d'un seul photon. Que faut-il pour approcher cette limite physique avec un capteur d'images à semi-conducteur ? La bonne surprise est que le silicium est extrêmement efficace pour convertir des photons incidents en électrons : les pertes de réflexions à la surface du capteur sont de quelques pourcents seulement. A l'intérieur du semi-conducteur, l'interaction entre les photons et le matériau est encore meilleure : la quasi-totalité des photons est convertie en paires de charges photo-générées. Le transport de ces charges de la profondeur du semi-conducteur à sa surface est aussi très efficace. En l'absence de champs électriques, les charges sont transportées par le mécanisme de la diffusion thermique, subissant des pertes négligeables sous condition que la distance à traverser reste inférieure à quelques dizaines de micromètres (la longueur de diffusion). Par contre, les champs électriques sont capables de transporter des charges sur des distances mesurées en millimètres, sans perte importante. Le vrai problème ne se pose donc pas dans le domaine physique, mais plutôt dans le domaine électronique : La limite actuelle de la détection des petits signaux de charges est affectée par le bruit thermique – dit « bruit de Johnson » – dans le canal des transistors. C'est donc le premier transistor dans n'importe quel circuit de détection qui limite la sensibilité avec laquelle un paquet de charges photo-générées peut être mesuré. Ce bruit fondamental correspond à un nombre de charges de bruit à l'entrée d'un transistor, nombre qui est proportionnel à la capacité effective vue par la grille du transistor, et à la racine de la température absolue fois la largeur de bande du circuit. Par conséquent, nous n'avons donc que ces trois paramètres de base pour influencer la sensibilité d'un circuit de détection optimisé. Dans la littérature scientifique et technique, on a recensé un grand nombre d'approches pour réduire ce bruit effectif et pour augmenter la sensibilité d'un capteur d'images. Dans notre travail nous avons poursuivi deux approches complémentaires :

Dans un premier temps nous avons utilisé un circuit de type miroir de courant, mais avec la modification d'une source de tension programmable entre les deux grilles des transistors.

Comme les transistors fonctionnent en mode d'inversion faible (« sub-threshold ») le gain d'amplification de ce circuit innovateur est typiquement une décade par 30-40 mV de différence de tension. Les résultats pratiques confirment qu'il est possible d'acquérir des images avec un bruit très faible et une sensibilité qui correspond à moins d'un photon. Malheureusement, ce résultat a été atteint en réduisant la largeur de bande effective. Bien que le bruit soit vraiment réduit au-dessous d'un photon, nous devons payer pour ce résultat excellent avec un temps de réaction réduit de notre capteur d'images ultra-sensible ; un temps qui peut approcher les centaines de millisecondes, ce qui n'est plus admissible pour un nombre important d'applications. Néanmoins, cette innovation d'un capteur d'images à semi-conducteur disposant d'une amplification programmable et d'une sensibilité de détection d'un seul photon s'est vue décerner la médaille d'argent de l'Ecole polytechnique fédérale de Zurich.

Pour surmonter cette complication de largeur de bande limitée, nous avons poursuivi une autre approche, en utilisant le phénomène physique de l'amplification par l'intermédiaire de l'effet avalanche. Dans ce cadre le défi technique est la conception des pixels avec la stabilisation individuelle de la tension de multiplication, une tension qui peut monter à une centaine de Volts, avec une précision de quelques dizaines de millivolts seulement. En dépit de ces obstacles techniques, nous avons abouti et nous pouvons démontrer le fonctionnement d'un capteur d'images avec une amplification programmable du signal photo-généré dans chaque pixel, avec un gain entre 1 et plusieurs milles. Malgré ce succès réjouissant, notre circuit a rencontré une limitation imprévue, celle du bruit additionnel de la multiplication, un « excess noise factor » de la technologie semi-conducteur utilisée, qui limite l'application de notre circuit de manière importante. Heureusement, la technologie du silicium a fait de grands progrès entre temps, et des chercheurs de l'EPFL vient de faire la démonstration d'un circuit similaire, en utilisant l'effet avalanche, avec lequel il est finalement devenu possible de détecter des photons individuels.

Comme notre monde est tridimensionnel, la plupart des animaux disposent du sens visuel en trois dimensions, la perception des distances locales dans une scène. Le mécanisme pour la vision en 3D est la stéréographie : La scène est vue de deux ou de plusieurs positions distinctes, et notre système visuel peut construire une représentation tridimensionnelle à base de cette information. Bien que cette partie de notre sens visuel nous semble fonctionner sans effort, un nombre énorme de calculs est en fait nécessaire, en utilisant des méthodes algorithmiques qui ne sont pas encore connus en détail. Pour cette raison, aucune solution technique basée sur la stéréographie n'existe actuellement pour la vision 3D qui pourrait approcher la performance de la vision tridimensionnelle dans la nature. Par contre, nous avons les moyens techniques pour l'exploitation d'un effet naturel qui nécessite une rapidité de mesure si grande qu'elle n'est pas disponible dans des systèmes biologiques : La vitesse de propagation de la lumière dans l'air qui a une valeur de presque 300'000 km/sec. Pendant une nano-seconde (un milliardième de seconde) la lumière parcourt donc une distance de 30 cm. Une approche évidente pour la vision tridimensionnelle serait donc la réalisation de pixels capables de mesurer d'une manière extrêmement exacte le temps d'arrivée d'une impulsion de lumière. Le développement d'un tel capteur nous a pris une dizaine d'années, mais nous y sommes parvenus : En utilisant une technologie semi-conducteur standard, nous avons pu réaliser des pixels qui sont capables de mesurer le décalage en temps des signaux photoniques avec une précision de 20 picosecondes. A base de ce capteur unique, une caméra 3D complète a été réalisée, dénommée SwissRangerTM : Une matrice de diodes électroluminescentes (LED) est modulée à une fréquence de quelques dizaines de MHz. Ceci produit un champ de lumière modulée qui illumine la scène, et cette

lumière est reflétée de manière diffusive par les différents objets dans la scène. Une petite partie de cette lumière peut être saisie par l'objectif de la caméra, et elle est focalisée sur la surface de notre capteur. Chaque pixel effectue la démodulation de la lumière incidente locale, et le temps de décalage entre la lumière émise et la lumière détectée est mesuré. Puisque nous connaissons la vitesse de la lumière, il est élémentaire de calculer la distance locale pour ce pixel à partir de la mesure du décalage en temps. Ainsi, nous obtenons une carte tridimensionnelle de la scène : Pour chaque pixel, la distance, le coefficient de réflexion et le niveau d'illumination locales sont connus, et ceci avec un taux de répétition équivalent à la période vidéo, c'est à dire 25 ou 30 images complètes par seconde.

Quelles sont les limites physiques de cette technique optoélectronique d'imagerie 3D ? La limite ultérieure est bien sûr imposée par la nature quantique de la lumière qui nous empêche de savoir avec certitude de combien de photons un paquet de photons consiste. On ne peut que déterminer une valeur moyenne du flux de photons. Dans le meilleur des cas (lumière non-corrélée), ce bruit photonique est proportionnel à la racine de la valeur moyenne des photons détectés, et il limite la précision avec laquelle la distance peut être mesurée à base de cette méthode de la mesure du temps de vol de la lumière modulée. Le résultat le plus important de nos efforts de recherche est que la précision de notre caméra 3D est actuellement limitée par le bruit photonique, pour une grande partie des conditions opérationnelles ; ce fait a pu être corroboré avec presque 100'000 mesures individuelles.

Notre caméra 3D miniaturisée représente le premier dispositif pratique, sans aucun élément qui bouge, pour la vision tridimensionnelle. Elle est à la base de plusieurs projets de développement industriels avec le but de la fabrication des caméras 3D moins coûteuses pour une multitude d'applications avec un énorme potentiel de marchés. C'est aussi pour cette raison que notre effort de recherche fondamentale et appliquée pendant une dizaine d'années a été décerné avec plusieurs distinctions internationales, notamment le grand prix « Information Society Technology Award 2004 » de la Commission Européenne, trois prix « Swiss Technology Award 2004 », la distinction « Literati Award of Excellence » et le prix Omega pour la meilleure thèse dans le domaine de la microtechnique.

Notre sens de la vue nous permet de voir même à travers des objets, sous condition qu'ils sont transparents et clairs. S'il s'agit d'un objet diffusant, nous ne percevons que des ombres et de vagues silhouettes. C'est dans ce cas précis que l'optoélectronique moderne a une solution unique à offrir pour l'imagerie tridimensionnelle à travers des objets diffusants : La tomographie à basse cohérence OCT (de l'anglais « Optical Coherence Tomography »). Cette technique n'est rien d'autre qu'une variante de l'interférométrie à lumière blanche. La conception d'un système OCT se base donc sur un interféromètre du type Twyman-Green, avec une source de lumière à basse cohérence (typiquement une diode super-luminescente SLD), un miroir demi transparent au centre, un miroir de référence qu'on peut déplacer dans la direction de l'axe optique, un objet dont la lumière diffusée est acquise en réflexion, et un objectif pour focaliser la lumière sur la surface d'un capteur d'images. Le défi scientifique est que le signal optoélectronique en fonction de la position du miroir de référence consiste en quelques petites franges d'interférence, superposées sur un signal géant (« DC offset »). La tâche du capteur optoélectronique est de détecter l'enveloppe de cette modulation à faible amplitude, sans tenir compte de cet offset géant. Notre contribution à cette méthode déjà connue est la réalisation d'un capteur d'images qui est capable de non seulement acquérir les signaux optiques par chaque pixel, mais d'effectuer le traitement de signal complet par un « smart pixel », disposant de circuits optimisés qui

fonctionnent simultanément et indépendamment des autres. Plusieurs milliers de ces « smart pixels » trouvent place sur un circuit intégré perfectionné pour la tâche d'imagerie OCT. Chaque pixel est responsable des fonctions suivantes : Soustraction d'un offset moyen, démodulation des franges d'interférences (p.ex. par rectification et filtrage basse bande), extraction de l'enveloppe des franges d'interférences par un filtre basse bande, et lecture des signaux obtenus, en séquence où de manière aléatoire. En utilisant ce parallélisme formidable dans l'acquisition et le traitement des données, il devient possible d'acquérir des images tridimensionnelles du volume d'un objet dans une fraction d'une seconde, et le taux d'acquisition de ces images 3D OCT peut approcher la fréquence vidéo, c'est à dire 25-30 Hz. Comme c'était le cas pour la caméra 3D SwissRanger, la performance de notre capteur d'images OCT s'établit également proche de la limite du bruit photonique. En effet, la limite pratique n'est plus octroyée par l'intensité de la lumière incidente, mais plutôt par la dissipation du courant nécessaire pour compenser les offsets dans les différents pixels. Si la dissipation totale d'un capteur d'images dépasse environ 1 W, sa performance risque de dégrader de manière importante et inadmissible.

Nous pronostiquons un grand nombre d'applications de cette nouvelle technologie, en particulier dans les domaines de la médecine (dermatologie, ophtalmologie, dentisterie, micro-chirurgie, etc.), de la biologie (microscopie, manipulation des cellules vivantes en 3D, etc.) et de la biométrie (acquisition des propriétés 3D caractéristiques à une personne).

Reste le défi ultime, dans un domaine où la nature règne toujours en maître : La réalisation d'une rétine artificielle capable de percevoir son environnement – la puce en silicium dotée du sens de la vue (« seeing chip »). A l'évidence, nous ne possédons pas aujourd'hui d'algorithmes de la vision qui puissent être profitablement implémentés dans le silicium ; les mécanismes de la vision naturelle ne sont pas encore éclaircis suffisamment dans le but de les utiliser pour la conception des rétines artificielles. De plus, le coût de développement de ce type de capteurs est encore assez élevé, dépassant aisément plusieurs centaines de milliers d'euros ! Cette démarche ne serait justifiable qu'économiquement pour autant que cette puce de la vision soit utilisée dans des applications à forts volumes (l'automobile, par exemple), ou comme une volonté électronique universelle. En dépit de ces obstacles, la communauté internationale des chercheurs a fait des progrès dans des applications bien ciblées. Bien que la puce dotée du sens de la vue générique nous échappera pour longtemps encore, ils existent déjà plusieurs exemples de systèmes de vision capable de percevoir une partie restreinte de l'environnement qui nous entoure. Un premier exemple est une seule puce qui intègre une caméra numérique complète ainsi que les circuits numériques pour la reconnaissance des empreintes digitales. Un tel circuit intégré pourrait remplacer un jour nos clés par des caractéristiques biométriques qui nous distinguent tous. Un deuxième exemple, également avec des applications dans la biométrie, serait un circuit intégré – un « seeing chip » – capable de reconnaître des êtres humains à base de leurs visages. Les travaux fondamentaux dans ce domaine sont d'ailleurs faits à l'institut de microtechnique de notre université par le groupe de recherche du professeur Farine.

Pas à pas, nous apprenons donc l'art de reconnaître les différents objets dans une scène, et nous nous approchons de la réalisation d'un circuit intégré qui puisse acquérir des images et comprendre leur contenu, une puce capable de percevoir le monde.

Mesdames et Messieurs, ce fut un plaisir pour moi de vous faire prendre part à la fascination exercée par cette activité de R&D, de créer des systèmes de vision artificielle, dépassant les performances des rétines naturelles. Et peut-être commencez-vous aussi à partager mon sentiment de fierté que notre université, en étroite collaboration avec le CSEM, a contribué et contribue toujours d'une manière importante aux développements dans ce domaine de la photonique, domaine qui présente un très fort potentiel de valorisation industrielle. Nous faisons ainsi partie des leaders au niveau mondial, et nous sommes déterminés de le rester pour longtemps encore:

Régner en maître sur la vision artificielle.