
Capteur d'images à grande dynamique et compression intégrée pour technologie 3D

Fadoua Guezzi-Messaoud¹, Antoine Dupret¹, Arnaud Peizerat¹,
Yves Blanchard²

1 Laboratoire de Circuits Intégrés et Intelligents pour l'Imagerie

CEA-LETI-DACLE, Campus Minatec

17, Rue des Martyrs F-38054 Grenoble Cedex

fadoua.guezzi@gmail.com, antoine.dupret@cea.fr, arnaud.peizerat@cea.fr

2. Université Paris-Est, Département Systèmes, Électronique

ESIEE Paris, Cité Descartes F-93160 Noisy Le Grand

y.blanchard@esiee.fr

RÉSUMÉ. A travers l'exploitation de l'intégration 3D, nous proposons d'améliorer la dynamique des capteurs d'images actuels et dépasser 120 dB. La technique utilisée se base sur l'adaptation du temps d'intégration par groupe de pixels en rétroagissant sur l'ensemble des pixels via les interconnexions verticales. L'architecture des pixels classiques, 3T ou 4T, n'est pas modifiée ce qui permet de bénéficier des hautes performances des imageurs classiques tout en rajoutant la grande dynamique. L'augmentation du nombre de bits pour représenter l'image grande dynamique est absorbée par une compression à deux niveaux, employée afin de réduire le flux de données en sortie du circuit tout en gardant la grande dynamique du signal.

ABSTRACT. This paper presents a new High Dynamic Range (HDR) Image Sensor architecture that uses the capabilities of Three-Dimensional Integrated Circuit (3D IC) to attain a range of 120 dB without modifying the classic pixel architecture (3T, 4T). The integration time is evaluated by group of pixels on a stack IC and feedback by vertical interconnections to the sensor. A two level compression is then applied on the pixel groups to reduce the output data rate while keeping the high dynamic range.

MOTS-CLÉS : 3D Intégration, capteur d'images, HDR, compression, DCT, codage flottant.

KEYWORDS: 3D-IC, image sensor, high dynamic range, compression, DCT, floating point coding.

DOI:10.3166/TS.30.343-365 © 2013 Lavoisier

Extended Abstract

High Dynamic Range Imager

A High Dynamic Range Imager is able to capture a frame with at the same time details in the light areas and in the dark areas. To obtain this result a dynamic of 120 dB, corresponding to 20 bits of linear coding is required by pixel, compared to 8-10 bits for classical imager.

Numerous work have dealt with the extension of the dynamic range as a classical pixel architecture does not allow this dynamic range because of saturation and signal to noise ration limitations. An ideal solution would be to control for each individual pixel the integration time of the photo current to adapt it to the illumination. However most of those works imply a deep modification of the pixel architecture and an important increase of its area.

This paper proposes a solution for the design of a HDR imager based on a 3D technology integration, where integrated circuits (IC) are stacked together and connected directly by Through Silicon Via (TSV). It is proposed to have a classical imager with 3T or 4T pixel architecture implemented on the first IC of the stack, while a second IC is used to do signal processing and control. On this processing level optimum integration times are evaluated in parallel for group of pixels, called here macro pixels. These information are fed back at the pixel level by the TSV connecting the two circuits. Furthermore as the quantity of data for HDR coding becomes important, a multi-level approach for a lossy compression scheme in each macro pixel is evaluated to reduce the data output throughout.

Macro pixel specificities

A macro pixel is a group of adjacent pixels that are going to have the same integration time. The macro pixel size is a compromise between the number of available TSV and the degradation brought down by not having individual pixel controls. The increase in the pixel dynamic is brought by a logarithmic variation of the exposure time, producing a pixel value in floating format with the Exponent corresponding to the integration time and the Mantissa to the conversion of the pixel value at the end of integration (Bermak and Kitchen, 2006).

$$Pixel_Valeur = M \cdot 2^E$$

One originality of this work is to have a common exponent for all the pixels belonging to the same macro pixel giving the advantages of the larger macro pixel area to implement the control electronics and a first level of data compression, but with the cost of a slight degradation of the PSNR as the macro pixel size increase.

In order to reduce the quantity of transmitted data a second level of compression of the mantissa is studied at the macro pixel level. Based on the Discrete Cosine Transform (DCT) it is tailored to the architecture specific needs, with a reduced available area and an imager partitioned into macro pixels. The DCT is not done

classically on an 8 by 8 pixels matrix but directly on the macro pixel size and the compression factor is obtained solely through quantification of the high frequency coefficients as a Variable Length Coder (VLC) would not fit in the available area. The study shows that this brutal quantification of the mantissa coefficients has a negative impact on the PSNR but that the HDR quality is kept and that the image can still be used for post-processing.

Exposure time choice architecture

The exposure time corresponds to the exponent of the pixel value and it is the same for all the pixels of the same macro pixel. For the choice of the common exponent, between different possibilities, it was found that the best PSNR was obtained by taking the highest exponent of all the pixels of the same group, corresponding to the exponent of the pixel receiving the maximum illumination. This information is obtained from the 32×32 pixels of the macro pixel by an analog Winner Take All that will bring the highest pixel value to a comparator used by the control logic to stop the integration before saturation and then and start the analog to digital conversion. An analog WTA requires very few supplementary transistors and the classical pixel architecture can be kept. However the WTA gain variation introduces nonlinearities that can produce an erroneous evaluation of the integration time. This can be compensated by a two phase approaches with first a rough approximation of the integration time and then a fine tuning of it leading to a correct evaluation.

Results and conclusion

A High Dynamic Range imager architecture has been described. It uses the vertical integration of 3D IC stacking to implement on a first level an imager using a classical 3T or 4T transistors pixel and on a second level all the additional control logic. HDR coding is based on a floating point value notation. A macro pixel organization, with a common exponent for different pixels allows the sharing of the processing logic between those pixels, giving enough area on the control level to implement processing for HDR acquisition and compression at the macro pixel level.

The various parameters that have an influence on the proposed architecture performance and output quality were evaluated with high dynamic range images. It was determined that a 32×32 macro pixel size with pixel mantissa having between 8-10 bits and a common exponent on 4 bits was offering a good tradeoff for implementation between area, compression and PSNR.

The WTA, crucial element for the evaluation of the exposure time, was designed on a double oxide CMOS 32 nm. It was necessary to compensate the nonlinearities introduced by the WTA with a two phases approaches to have correct values on all the input dynamic range.

The complexity of the final 32×32 DCT was reduced by using on chip transform coefficients coded with only 3 bits, the image being reconstructed off chip by inverse transform of the mantissa and multiplication by the macro pixel exponents. A compression ratio of 75% can be obtained while keeping a PSNR of about 40dB. Blocks artifacts are visible but the image keeps enough information for post processing. Higher compression ratios can be obtained, but at 95%, the result image is at a limit of being usable with a PSNR of 30dB.

The different optimization of the architecture allow to obtain a usable image with less than two bits/ pixel compared to the original 20 bits/pixel of an HDR image.

1. Introduction

La poursuite de l'intégration de fonctions toujours plus complexes au sein d'un même circuit constitue un des principaux enjeux de la microélectronique. L'intégration tridimensionnelle par empilement de circuits (3D stacking) constitue une voie prometteuse pour y parvenir. Elle permet notamment de dépasser certaines limitations atteintes par les circuits actuels, plus particulièrement dans les circuits pour lesquels les données sont distribuées et qui nécessitent des bandes passantes importantes. Les travaux présentés dans cet article visent à exploiter cette technologie émergente dans le cadre des capteurs d'images et dépasser la preuve de concept présentée par : Suntharalingam et Berger, 2005 ; Magnan *et al.*, 2009. Il s'agit ici d'apporter une réponse à la limitation de la dynamique des capteurs d'images, tout en gardant une architecture classique des pixels et en adressant le problème de la réduction de la quantité de données à transmettre.

L'extension de la dynamique des capteurs d'images a fait l'objet de nombreux travaux pour les circuits classiques 2D (Bermak et Kitchen, 2006 ; Young et El-Gamal, 1999 ; Spivak *et al.*, 2009 ; Akahane *et al.*, 2009 ; Mase *et al.*, 2005 ; Rhee et Joo, 2003 ; Rhee et Park, 2009), plus récemment pour de futurs circuits 3D (Guezzi *et al.*, 2011 ; 2010). Cependant, la majorité de ces techniques modifie l'architecture du pixel et, en particulier, rajoute de l'électronique. Par conséquent, la dynamique est étendue au détriment d'autres caractéristiques qui influent sur les performances, par exemple, le facteur de remplissage.

Les performances des imageurs étant fortement liées à l'architecture et au dessin de masque des pixels, toute nouvelle étude de celui-ci est coûteuse. Dans ce papier, nous présentons une nouvelle architecture d'imageur permettant d'étendre la dynamique à plus de 120 dB sans avoir à modifier l'architecture classique (3T ou 4T) des pixels (Guezzi *et al.*, 2011). Cette architecture tire parti de l'émergence de technologies d'interconnexions verticales denses, *Through Silicon Via (TSV)*, pour adapter localement le temps d'intégration par groupe de pixels. La capture d'une image à grande dynamique (HDR) avec un PSNR (*Peak Signal to Noise Ratio*) élevé induit une augmentation du flux de données en sortie du circuit. L'architecture HDR est donc couplée à un dispositif à deux niveaux de compression des données. Un premier niveau de compression est réalisé par un codage de la valeur du pixel avec

une notation flottante originale, le deuxième par une transformée en cosinus discrète (DCT, *Discrete Cosine Transform*).

L'article commence par exposer une vue globale de l'architecture 3D proposée et de son principe de fonctionnement, puis sont présentées la technique d'extension de dynamique qui assure également le premier niveau de compression, pour terminer avec le 2^e niveau de compression. Les principaux résultats obtenus sont détaillés.

2. Architecture du capteur d'images 3D HDR

Un capteur d'images HDR est un capteur capable de restituer à la fois les détails des zones faiblement et fortement éclairées de la scène capturée. Ceci requiert une dynamique d'environ 120 dB équivalente à 20 bits par pixel en codage linéaire. Compte tenu des flux photoniques des scènes naturelles, seule la technique d'intégration temporelle des photo-courants pendant le temps d'exposition permet d'atteindre des valeurs de SNR supérieures à 30 dB (Vatteroni *et al.*, 2007). Si le même temps d'intégration était appliqué pour toute l'image et puisque la scène naturelle peut à la fois contenir des zones sombres et très éclairées, il faudrait disposer de pixels capables de stocker plusieurs millions d'électrons tout en ayant des niveaux de bruits de l'ordre d'un électron. Les dimensions des pixels, nécessairement réduites (typiquement 1,5 à 7,5 μm de côté) dès lors que des résolutions de l'ordre du million de pixels sont envisagées, réduisent la capacité de stockage à quelques milliers d'électrons. En d'autres termes, les architectures classiques de pixels conduisent à la saturation d'un certain nombre de pixels ou à des signaux noyés dans le bruit pour les pixels n'ayant pas eu un temps d'exposition suffisamment long.

Idéalement, il faudrait pouvoir ajuster le temps d'exposition pour chaque pixel et disposer de suffisamment de surface dans le pixel pour mémoriser son temps d'exposition. Comme nous l'avons déjà mentionné, cela implique d'augmenter la surface du pixel et de modifier profondément son architecture. Pour pallier le manque de surface disponible au niveau pixel, le circuit proposé dans ce papier exploite l'intégration 3D de 2 circuits intégrés.

Le principe de fonctionnement se base alors sur une rétroaction sur le pixel ou groupe de pixels voisins. En pratique, la rétroaction vise à adapter le temps d'intégration par groupe de pixels, notés dans la suite macro-pixel. Le temps d'intégration est alors déterminé pour optimiser le temps d'exposition pour les pixels d'un macro-pixel : il s'agit ainsi d'éviter de saturer des pixels du fait de temps d'exposition trop longs ou d'avoir des niveaux de signal trop faibles du fait de temps d'exposition trop courts. Par conséquent, dans une même image, plusieurs temps d'intégration sont utilisés avec une granularité égale à celle des macro-pixels. La dimension des macro-pixels résulte d'un compromis entre la densité des TSV et de la dégradation du PSNR. Ce dernier est déterminé expérimentalement comme nous le verrons dans la section 3.

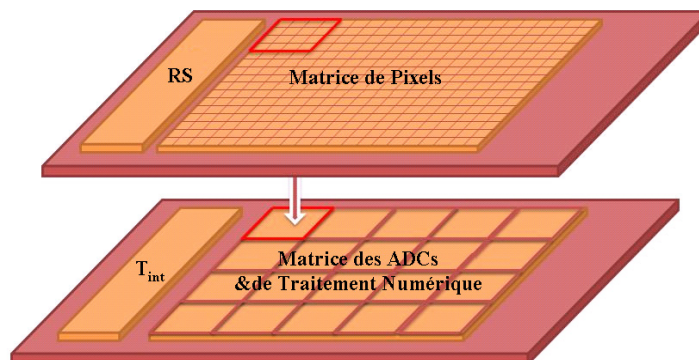


Figure 1. Architecture du capteur d'images 3D HDR

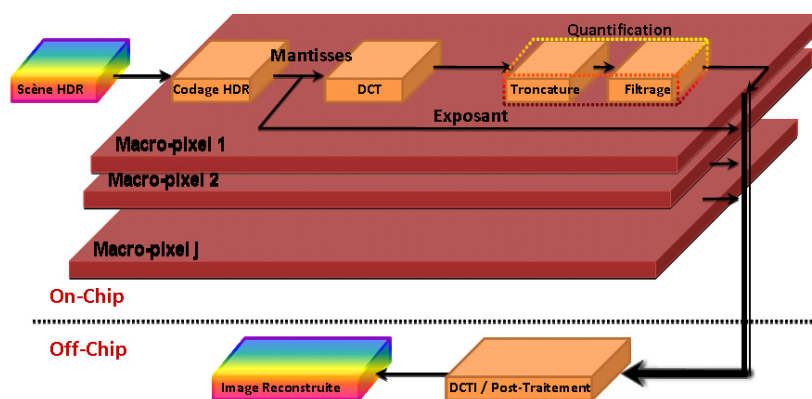


Figure 2. Schéma blocs du flot de compression

Une vue globale du circuit 3D est donnée en figure 1. Le circuit est composé de deux couches superposées verticalement et interconnectées par des TSV. La première couche reprend l'architecture d'un capteur d'images 2D classique comportant des pixels (3T ou 4T). Chaque pixel comprend une photodiode et un circuit de lecture. L'information issue de chaque pixel de la première couche est alors une tension analogique transmise à travers les interconnexions verticales à la 2^e couche. Le traitement effectué sur la 2^e couche comporte deux étapes. La première étape est l'évaluation du temps d'intégration adapté pour chaque macropixel, et dans un deuxième temps, la génération du signal de commande permettant d'agir localement sur le macro-pixel. Le traitement revient à effectuer un codage flottant de l'information appelé codage Mante-Exposant (M-E). Ce traitement sera explicité dans la section 3. Ce codage réalise une première diminution de la quantité de données qui serait obtenu par un codage grande dynamique linéaire. L'information étant numérisée dans chaque macropixel, on profite de cette étape pour implémenter,

sur la deuxième couche, une compression de données à base de DCT. L'objectif étant de faire un traitement parallèle par macropixel, il est proposé de faire une compression par simple quantification en sortie de la DCT sans codeur entropique. Le principe sera détaillé dans la section 5. La figure 2 illustre le traitement proposé et explicite le contenu de chaque élément de la deuxième couche du circuit 3D. Dans les sections suivantes, le principe du codage HDR et l'architecture proposée sont expliquées.

3. Codage HDR : principe et architecture

3.1. Principe

Les imageurs CMOS bas coût grands formats reposent sur des pixels d'architectures 3T ou 4T afin de concilier compacité et fort facteur de remplissage. Le mode de fonctionnement de tels pixels est l'auto-intégration dans la photodiode après chaque phase de remise à zéro. En effet, parmi l'ensemble des architectures de pixels, l'accumulation des charges résultant de la phototransduction permet d'obtenir le niveau de bruit le plus faible. Or, la quantité maximale de charges qu'il est possible de stocker dans le pixel est directement proportionnelle à la capacité de la diode ainsi qu'à la tension maximale aux bornes de la photodiode. La dynamique du pixel est donc donnée par :

$$DR_{dB} = 10 \log \left(\frac{Q_{Max}}{Q_{noise}} \right) \text{ Avec : } Q_{Max} = C \Delta V_{Max} \quad (1)$$

Il existe de nombreuses techniques permettant d'obtenir une quantité équivalente de charges stockées supérieure à celle permise par la capacité de la diode. Notre objectif est de proposer une architecture HDR qui préserve l'architecture pixel (3T ou 4T) permettant ainsi de capitaliser sur les études complexes requises pour la conception de ces pixels.

Parmi les différentes techniques existantes pour augmenter la dynamique, la variation logarithmique du temps d'exposition est l'une des solutions compatibles avec des architectures 3T ou 4T sous réserve que le contrôle du temps d'exposition soit effectué hors du pixel (Bermak et Kitchen, 2006). Un tel procédé est équivalent à un codage flottant de la valeur numérique du pixel de type Mantisse-Exposant.

L'adaptation du temps d'intégration, en fonction de l'éclairement reçu, est effectuée pour chaque macropixel. Le temps d'intégration évalué pour chaque macropixel est choisi parmi un ensemble de valeurs prédéfinies évoluant en puissance de 2 et dont le nombre dépend de la résolution choisie pour l'exposant. Par exemple, si l'exposant est codé sur 4 bits, le temps d'intégration peut prendre l'un des 16 intervalles de temps.

La mantisse représente la valeur numérique du signal analogique issu du pixel à la fin de la période d'intégration. Ceci permet d'écrire la valeur du pixel sous la forme suivante :

$$\text{Pixel_Valeur} = M \cdot 2^E \quad (2)$$

où E représente la valeur d'exposant et M la valeur de la mantisse.

Si Neq représente le nombre de bits nécessaires pour un codage linéaire du signal dans toute sa dynamique, il existe plusieurs couples de nombres de bits d'exposants, Ne , et de nombre de bits de mantisses, Nm , couvrant la même plage de valeurs que Neq bits linéaire. Le nombre de bits de la mantisse détermine le SNR (*Signal to Noise Ratio*) de quantification. Pour un bruit de quantification fixé, Nm est fixé et l'on détermine Ne par la relation suivante :

$$Ne = \lceil \log_2(-Nm + Neq) \rceil \quad (3)$$

Dans notre cas, avec une dynamique de scènes de l'ordre de 120 dB, il faut donc disposer de l'équivalent d'un codage sur $Neq = 20$ bits. Par ailleurs, le SNR introduit par la quantification est fixé par le nombre de bits de la mantisse. Un SNR de 50 dB nécessite un codage de la mantisse sur au moins 8 bits. L'équation [3] nous montre que la dynamique de 120 dB avec un SNR de 50 dB peut être atteinte en prenant $Nm = 8$ bits et $Ne = 4$ bits avec 13 valeurs distinctes de l'exposant ou $Nm = 10$ bits et $Ne = 4$ bits avec 11 valeurs distinctes de l'exposant.

Ce codage permet de passer de 20 bits par pixel à 12 ou 14 bits par pixel selon le nombre de bits Nm de la mantisse. Dans la plupart des applications d'acquisition d'images, les pixels sont codés sur 8 à 10 bits. Pour limiter d'avantage le flux de données imposé par l'augmentation de la dynamique, nous proposons une approche originale. Celle-ci se base sur le principe du codage flottant mais en appliquant une valeur d'exposant commune à tous les pixels d'un macropixel. L'étude de différentes méthodes de choix de l'exposant commun (Guezzi *et al.*, 2011 ; 2010), permet d'établir qu'en termes de PSNR (*Peak Signal to Noise Ratio*), le meilleur choix de l'exposant commun à tous les pixels du même macropixel est l'exposant maximum de tous les exposants des pixels du même groupe (cf. figure 3). En factorisant ainsi les exposants des pixels, au sein d'un même macropixel, les pixels ne se différencient alors que par leurs mantisses. Ce codage a été appliqué à des images HDR de 20 bits pour évaluer la qualité des images obtenues à l'aide du PSNR.

Cette technique introduit un taux de compression dont l'expression est donnée ci-dessous :

$$Tc = \frac{Nm + \frac{Ne}{Q^2}}{Nm + Ne} \quad (4)$$

où Q^2 représente le nombre de pixels dans un macropixel. Lorsque Q^2 tend vers l'infini, la limite du taux de compression est :

$$T_{c_{limite}} = \frac{Nm}{Nm + Ne} \quad (5)$$

En pratique, au-delà de $Q^2 = 16$ le taux de compression n'évolue plus de façon significative.

Le circuit de calcul du temps d'exposition doit être placé sous chaque macro-pixel. Il est donc nécessaire de déterminer la dimension maximale des macro-pixels qui permettent à la fois de réduire la quantité de données à transmettre, conserver un PSNR supérieur à 50 dB et assurer une surface suffisante pour intégrer le circuit de calcul de temps d'exposition.

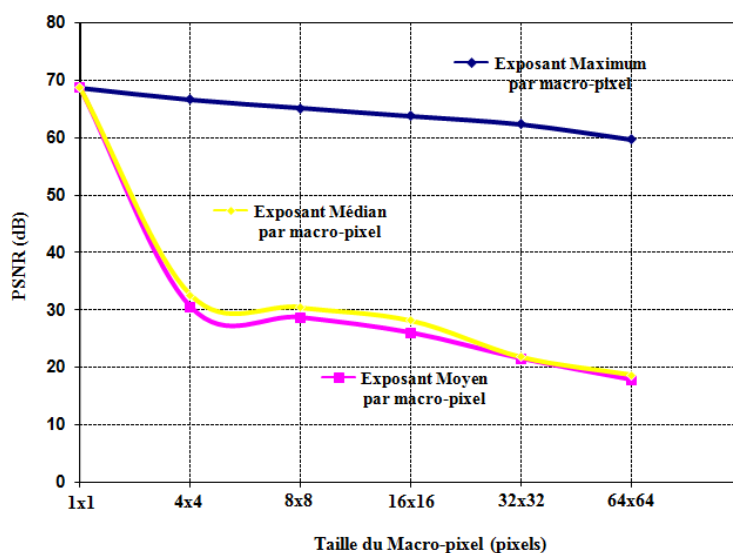


Figure 3. PSNR de l'image obtenue avec différentes méthodes d'exposant commun par macropixel

Par ailleurs, le choix du nombre de bits des mantisse/exposant doit permettre l'utilisation d'un convertisseur compact et à faible consommation, tel qu'un convertisseur à rampe, celui-ci devant également être intégré dans chaque macropixel. Notre choix s'est porté sur une mantisse codée sur 8 à 10 bits et un exposant codé sur 4 bits. Nous avons donc étudié l'évolution du PSNR en fonction de la taille des macropixels en codant la mantisse avec 8 bits et l'exposant avec 4 bits. La figure 3 montre l'évolution du PSNR en fonction de la taille des blocs. Il apparaît que les macropixels peuvent atteindre des dimensions de 32×32 pixels tout en conservant un PSNR élevé, c'est-à-dire supérieur à 60 dB.

La figure 4 permet de vérifier que la factorisation de l'exposant par macropixel n'a qu'un faible impact sur la qualité des images.

Cette technique de factorisation permet d'assurer un premier niveau de compression. De plus cette factorisation d'exposant permet d'avoir suffisamment de surface pour intégrer un bloc de calcul de temps d'exposition par macropixel.

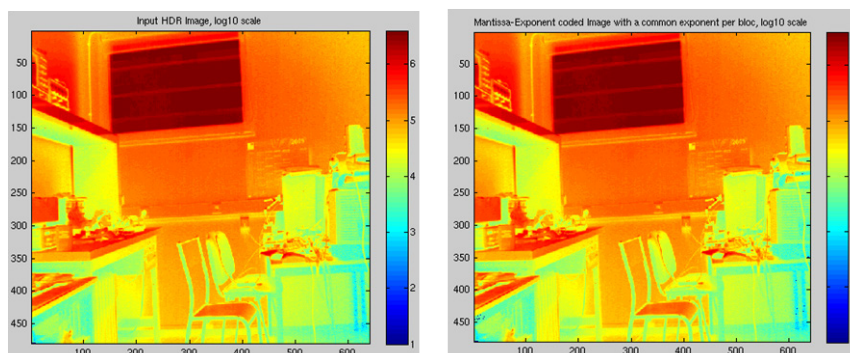


Figure 4. Image HDR originale (à gauche) et image reconstruite après codage m-e (à droite)

3.2. Architecture d'ajustement du temps d'exposition

Nous partons de l'hypothèse qu'au sein de la scène, l'éclairage varie lentement vis-à-vis du temps trame. L'objectif est de fixer un temps d'intégration qui permet de garantir qu'aucun des pixels ne sature.

Pour déterminer la plus forte intensité lumineuse reçue parmi tous les pixels d'un macropixel ainsi que la valeur de la chute de tension correspondante et en déduire le temps d'intégration optimal, nous utilisons le circuit dont l'architecture est présentée en figure 5 et dont les chronogrammes correspondants sont donnés en figure 6. Nous nous appuierons sur l'exemple d'un pixel 4T en mode Global Shutter.

Dans une première phase, la photodiode et le nœud flottant de mémorisation de chaque pixel sont réinitialisés à V_{dd} via le signal de remise à zéro (RST) qui est appliqué simultanément à tous les pixels de la matrice. Dans chaque pixel, le courant I_{ph} est intégré dans la capacité C_{int} ce qui produit une tension décroissante V_{int} proportionnelle au temps d'intégration et à l'intensité lumineuse reçue.

Dans chaque macropixel, la chute de tension ΔV maximale, correspondant au pixel qui reçoit la plus forte intensité lumineuse, est déterminée à l'aide d'un circuit Winner Take All. Classiquement utilisé dans les circuits neuromorphiques, le WTA (Winner-Take-All) est un circuit permettant d'extraire la tension maximum parmi toutes celles présentes à ses entrées (Carvajal *et al.*, 2000 ; Soleimani *et al.*, 2009) et (Lopez *et al.*, 2006).

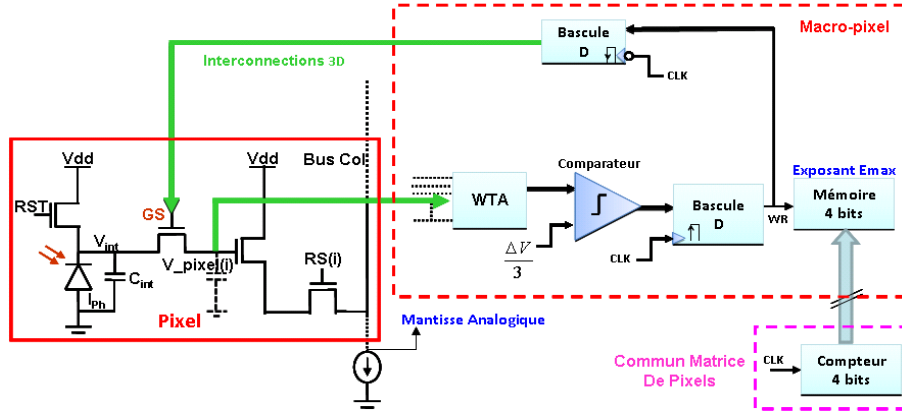


Figure 5. Architecture proposée du capteur d'images 3D HDR

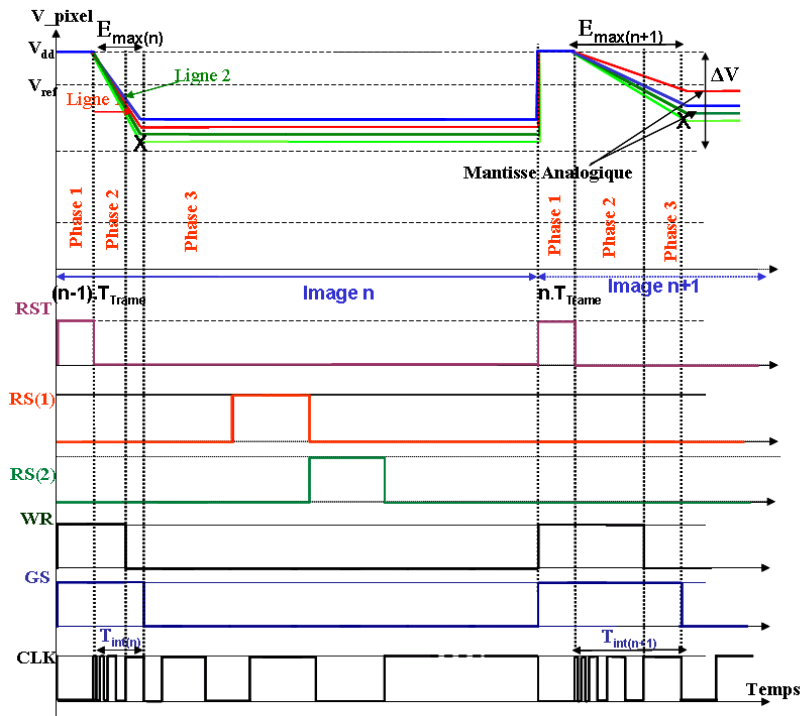


Figure 6. Différentes phases de fonctionnement pendant un temps trame

Tant que la sortie du WTA reste inférieure à $V_{ref} = \Delta V/3$, le pixel continue à intégrer. Dès que ce seuil est atteint, l'expositif correspondant au temps d'exposition

maximum est mémorisé. Au coup d'horloge suivant, le signal GS est activé. L'activation de ce signal bloque le transistor GS et, en isolant le nœud V_{pixel} , permet de mémoriser la tension du pixel dans le nœud flottant. En considérant que les fuites sont négligeables, cette tension sera figée jusqu'à la phase de réinitialisation.

Pour refléter le codage exponentiel, le temps d'intégration varie exponentiellement. Le compteur compte à la cadence d'une horloge logarithmique évoluant en $1/2^i$ avec i allant de 0 à N_e . Au cours de la durée d'une trame, la fréquence d'horloge diminue exponentiellement en fonction du temps (cf. figure 6). Chaque ligne i de pixels est lue en mode « fenêtre déroulante » (Rolling Shutter) lors de l'activation du signal $RS(i)$ correspondant. Puisque l'intégration (signal GS) se termine un demi-coup d'horloge ($1/2t_{\text{com}}$) après la commutation du signal WR, cette commutation peut intervenir après un temps double (t_{com}). Le choix de $V_{\text{ref}} = \Delta V/3$ permet d'assurer que la variation de tension en fin de commutation est au plus de ΔV . En effet, le front d'horloge suivant interviendra après un temps t_{com} . En supposant l'éclairement constant sur le temps total d'intégration, la tension du pixel n'aura pas variée de plus de $\Delta V/3 + 2\Delta V/3 = \Delta V$, ce qui permet d'éviter que le pixel recevant l'éclairement le plus fort ne sature (cf. figure 6).

La figure 6 illustre ces différentes phases de fonctionnement pendant un temps trame pour le cas d'un macropixel 2×2 (ligne 1 : pixels rouge et bleu, ligne 2 : pixels vert et vert clair).

Les différents opérateurs de cette architecture peuvent introduire des imperfections plus ou moins gênantes (offset, gain...). L'objectif principal de cette partie d'architecture de l'imageur étant d'estimer l'exposant convenable à chaque macropixel, ces imperfections peuvent alors introduire une erreur sur l'estimation de l'exposant approprié. Par conséquent, l'impact des non-idéalités sur l'évaluation de l'exposant est principalement étudié et le risque d'erreur est alors évalué. Dans le paragraphe suivant, on explicite cette étude.

4. Étude du WTA et évaluation de l'erreur induite

4.1. Principe du WTA

Comme nous l'avons vu, le temps d'exposition est déterminé en un cycle trame. Toutes les tensions des pixels d'un même macropixel sont évaluées simultanément. Le pixel ayant la plus forte variation de tension impose le temps d'exposition en fonction de sa chute de tension ΔV_{pixel} (cf. figure 7). Pour déterminer le temps d'intégration optimal, donc la valeur de l'exposant, le WTA doit pouvoir déceler une différence de tension au moins égale à la variation de tension correspondant au plus petit pas d'exposant, c'est-à-dire de distinguer des différences de tension proches de $\Delta V/2^{\text{ExpMax}}$ lorsque ΔV_{pixel} est inférieure ou égale à $\Delta V/2^{\text{ExpMax}}$.

En considérant une dynamique de tension ΔV de 1V avec 10 bits codant la mantisse et 4 bits codant l'exposant, il faut au moins 11 valeurs distinctes (de 0 à 10)

de l'exposant pour atteindre une dynamique de 120 dB équivalente 20 bits en linéaire. La précision théorique souhaitée du WTA est alors :

$$\text{précision_WTA} = \frac{\text{Dynamique_en_Tension}}{2^{\text{Exposant_Max}}} = \frac{\Delta V}{2^{10}} = 1mV \quad (6)$$

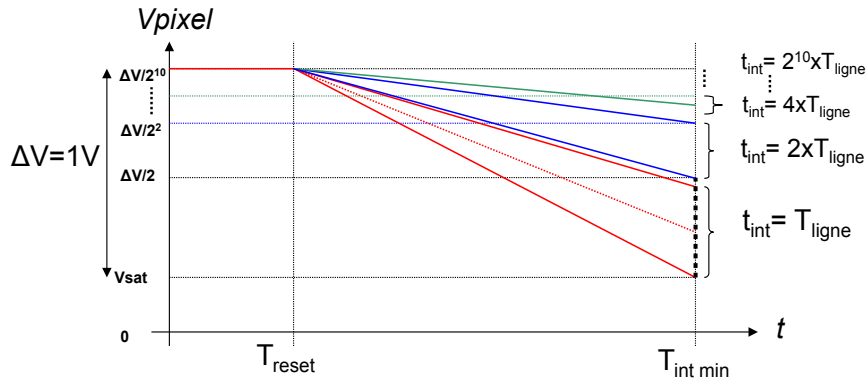


Figure 7. Chronogrammes des différents signaux du circuit proposé

Nous avons évalué les performances d'architectures basées sur des Winner Take All. Depuis (Lazzaro, 1988), de nombreux travaux ont été menés pour améliorer les performances des WTA notamment en termes de consommation et de précision (Carvajal *et al.*, 2000 ; Soleimani *et al.*, 2009 ; Lopez *et al.*, 2006).

Nous considérons le circuit WTA présenté en figure 8, dont toutes les entrées V_{in_i} sont connectées simultanément à tous les pixels du même macropixel.

Comme le montre la figure 8, les n paires différentielles se partagent la même source de courant de polarisation. Les entrées non inverseuses des paires différentielles sont connectées aux pixels tandis que les entrées inverseuses sont reliées entre elles et aux sorties des étages différentiels.

Rappelons le principe de fonctionnement du WTA en nous appuyant sur le cas de deux pixels. Ce principe se généralise sans peine à N paires différentielles. On considère tout d'abord la condition initiale $V_{in1} = V_{in2}$ ce qui induit $I1 = I2 = Ic/2$ avec V_{in1} la tension en entrée de la 1^{re} paire différentielle, V_{in2} la tension en entrée de la 2^e paire différentielle, Ic le courant de polarisation, $I1$ et $I2$ les courants circulant respectivement dans les paires différentielles (T11, T12) et (T21, T22). Le même courant $Ic/4$ circule dans chaque transistor des 2 paires différentielles. Dans cet état de repos, tous les transistors fonctionnent en régime de saturation.

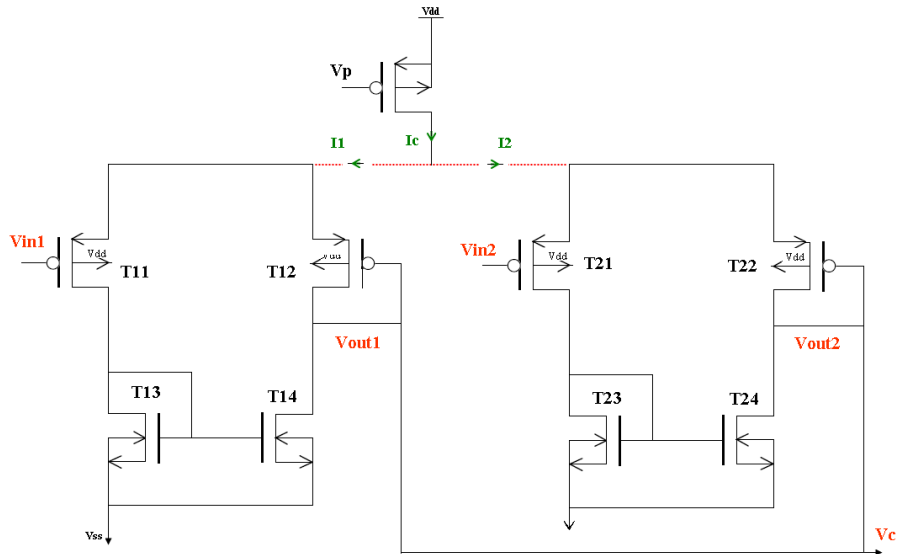


Figure 8. Schéma du WTA

Considérons maintenant que V_{in1} descende au-dessous de V_{in2} . Cela induit une augmentation du courant dans T11. Le potentiel V_c diminue alors et le courant absorbé dans les transistors T12 et T22 augmente. La tension différentielle $V_{GS21} - V_{GS22}$ augmente et tend à bloquer T21.

Par abus de langage, nous appelons « paire » différentielle, le transistor T11 et l'ensemble des transistors T12, T22... qui se partagent le courant I_c . La « paire » différentielle gagnante fonctionne donc en suiveur tandis que les transistors T21, T31,... des autres « paires » différentielles sont bloquées.

Aux offsets, dispersions et erreurs de gain près, la tension V_c en sortie est identique à la tension d'entrée de la « paire » différentielle qui a absorbé tout le courant I_c . Ce raisonnement reste valable pour N paires différentielles connectées entre elles. Afin d'évaluer l'impact des offsets, dispersions et erreurs de gain, nous avons déterminé la relation entre V_c et V_{in} , donnée par :

$$V_c = G \times V_{in1} + V_{Offset} \quad (7)$$

V_{Offset} peut être évalué de la façon suivante :

Nous nous plaçons dans le cas où toutes les branches d'entrée des paires différentielles « perdantes » sont bloquées. Afin de réduire la consommation tout en gardant un gain G proche de 1, nous choisissons un courant de polarisation I_c tel que seuls les transistors bouclés en suiveurs de ces paires différentielles fonctionnent en faible inversion et aient alors un courant i de la forme :

$$i = I_s \times e^{\frac{V_{sg}(i) - V_{TH}}{n \cdot V_t}} = \frac{Ic}{2 \cdot N} \quad (8)$$

Avec $V_t = \frac{kT}{q}$, où k est la constante de Boltzmann, T la température et q la charge de l'électron.

Cela permet de déterminer la tension $V_{sg}(i)$:

$$V_{sg}(i) = n \cdot V_t \ln\left(\frac{Ic}{2 N \cdot I_s}\right) + V_{TH} \quad (9)$$

L'équation ci-dessus peut se réécrire :

$$V_{sg}(i) = n \cdot V_t \ln\left(\frac{Ic}{2 I_s}\right) + V_{TH} - n \cdot V_t \ln(N) \quad (10)$$

La branche gagnante voit la moitié de courant de polarisation et fonctionne en limite régimes faible-forte inversion. Le courant de cette branche s'écrit alors :

$$i_w = I_s \times e^{\frac{V_{sgw} - V_{th}}{n \cdot V_t}} = \frac{Ic}{2} \quad (11)$$

d'où la tension V_{sgw} :

$$V_{sgw} = n \cdot V_t \cdot \ln\left(\frac{Ic}{2 I_s}\right) + V_{th} \quad (12)$$

En substituant l'équation (12) dans l'équation (10), l'expression de $V_{gs}(i)$ devient :

$$V_{gs}(i) = V_{sgw} + n \cdot V_t \cdot \ln(N) \quad (13)$$

La tension $n V_t \cdot \ln(N)$ représente l'offset introduit par le WTA. Cette tension d'offset peut être compensée lors du calcul de l'exposant. L'impact des dispersions est réduit par la mise en parallèle d'un grand nombre de branches. Le gain G est alors donné par :

$$G = \frac{Vc}{Vinw} = \left(\frac{g_{mw1}}{g_{mw3} + g_{dsw1} + g_{dsw3}} \right) \left(\frac{g_{mw4}}{\sum_i^N g_{mi2} + \sum_i^N g_{dsi2} + g_{dsw4}} \right) \quad (14)$$

où, d'après [8] et [11], les transconductances sont données par :

$$g_{mw1} = g_{mw3} = g_{mw4} = \frac{I_c}{2 \cdot n \cdot V_t} \quad (15)$$

et

$$g_{mi2} = \frac{I_c}{2 \cdot N \cdot n \cdot V_t} \quad (16)$$

Ce qui permet d'établir :

$$G = \frac{V_c}{V_{inw}} = \left(\frac{g_{mw1}}{g_{mw1} + g_{dsw1} + g_{dsw3}} \right) \left(\frac{g_{mw1}}{g_{mw1} + \sum_i^N g_{dsi2} + g_{dsw4}} \right) \quad (17)$$

En première approximation, les transconductances sont grandes devant les conductances et l'expression du gain devient : $G = 1$

En fait, avec la technologie choisie et en prenant des transistors de faibles dimensions (cf. tableau 1) pour limiter l'encombrement du WTA, le rapport g_m/g_{ds} est voisin de 100. Dans ces conditions, G est proche de 0,98. Le calcul de l'expression de la tension de sortie du WTA en fonction de la tension du pixel gagnant nous permet de vérifier que la tension de sortie du WTA est bien une expression linéaire de la tension du pixel gagnant. Dans la section suivante, nous présentons les limites de ces calculs et l'impact sur la capacité du WTA à discriminer.

Tableau 1. Dimensions, transconductance et conductances des transistors du WTA (branche gagnante)

Transistor	Ti1	Ti2	Ti3	Ti4
W (µm)	0,5	0,5	0,36	0,36
L (µm)	0,35	0,35	1	1
g_m (µS)	19,77	0,0197	19,77	20,19
g_{ds} (nS)	171	0,262	217	202

4.2. Implémentation du WTA : résultats et limitations et comment s'en affranchir

Le circuit de la figure 8 a été conçu en technologie CMOS 32 nm double oxyde. Cette technologie permet en effet de concilier les caractéristiques d'une technologie CMOS 0,25 µm pour les transistors fonctionnant en mode analogique, avec des densités d'interconnexions et d'intégrations élevées pour les fonctions numériques.

La caractéristique tension de sortie en fonction de la tension d'entrée du WTA, est obtenue en simulant le circuit pour différentes tensions d'entrée. On prélève ensuite la valeur de la tension d'entrée et la tension V_c en sortie du WTA. Les différentes simulations nous ont permis de tracer la courbe de la figure 9. L'équation caractéristique de la courbe obtenue montre un gain de 0,973 et une tension d'offset de 321 mV. Ces valeurs sont cohérentes avec les expressions analytiques déterminées dans la section 4.1. L'erreur due à l'offset peut être prise en compte dans l'évaluation de l'exposant car il s'agit d'une valeur constante. En revanche, comme le montre l'évolution de la dérivée de la tension de sortie du WTA par rapport à la tension d'entrée, en fonction de la tension pixel (cf. figure 10), le gain du WTA n'évolue pas de manière linéaire.

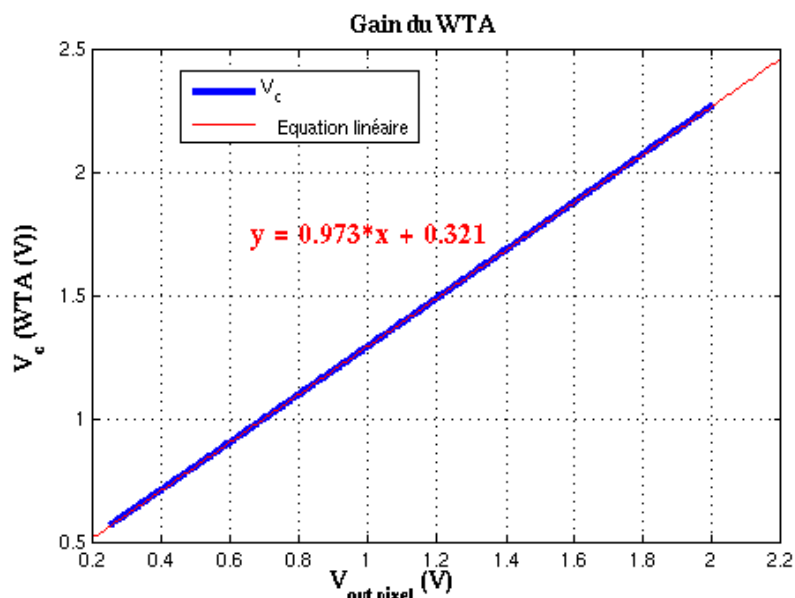


Figure 9. Tension de sortie du WTA en fonction de la tension du pixel

Les résultats de la Figure 10 confirment alors que l'erreur liée au gain dépend de la tension du pixel ayant gagné. Par conséquent, l'erreur du gain peut entraîner un décalage de la valeur analogique du signal lue vers un intervalle proche ce qui induit une erreur sur l'évaluation du temps d'intégration.

En conclusion, le WTA est une solution compacte en termes de ressources électroniques et simple à implémenter. Néanmoins, avec de faibles éclaircissements, elle peut donner lieu à des erreurs sur les valeurs d'exposant. Cette mesure peut être corrigée en allongeant le temps d'exposition pour les cellules concernées. Nous proposons donc d'effectuer le calcul de l'exposant en deux étapes. À chaque étape l'exposant est codé sur 3 bits en ne prenant que les valeurs de 0 à 5. Dans un premier

temps, le temps d'intégration est fixé au minimum afin d'établir une évaluation grossière de l'exposant. À l'étape suivante, c'est-à-dire en introduisant une pseudo-trame suivante, le temps d'intégration des pixels avec l'exposant minimum est multiplié par 2^5 . L'exposant est alors de nouveau estimé avec le WTA ce qui revient à en effectuer une évaluation fine. À l'issue des deux pseudo-trames, le temps d'intégration adapté est donc connu avec une précision compatible avec l'échelle de dynamique de 120 dB.

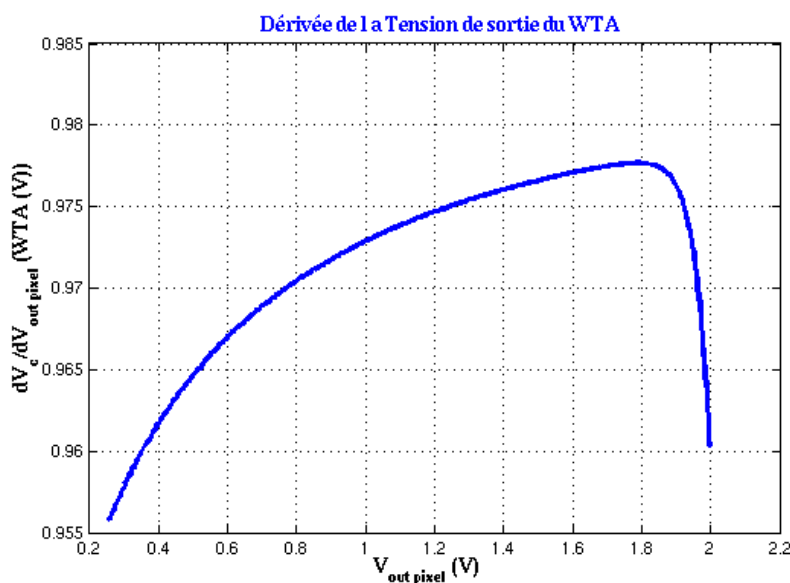


Figure 10. Gain en tension de sortie du WTA en fonction de la tension d'entrée

Enfin, la lecture de l'imageur est effectuée comme pour un imageur classique. La matrice des exposants est lue en même temps que l'imageur mais à une cadence Q fois plus lente afin que l'exposant reste constant pendant toute la durée de lecture de l'ensemble des pixels du macropixel.

5. Deuxième niveau de compression à base de DCT

La chaîne de traitement de compression d'une image est classiquement constituée d'une transformée qui va séparer les composantes hautes fréquences et basses fréquences. Cette transformée est suivie d'une quantification qui permet de réduire l'amplitude de ces coefficients. Enfin, un codeur entropique va les encoder avec un code à longueur variable. L'association de la quantification et du codeur à longueur variable permet ainsi d'obtenir la réduction du nombre de bits de l'image, mais c'est particulièrement la quantification qui permet de jouer sur le taux de

compression et la qualité de l'image résultante en annulant plus ou moins de coefficients issus de la transformée.

Dans le cas de notre imageur, cette chaîne doit être simplifiée pour tenir compte des contraintes locales et globales de l'architecture. La contrainte locale tient principalement au traitement massivement parallèle où chaque macropixel travaille indépendamment de ses voisins. La compression se fait au sein du macropixel et doit donc tenir dans la surface d'un macropixel. Au niveau de l'architecture globale, les données issues et compressées dans chaque macropixel doivent ensuite être récupérées et réunies pour être transmises à l'extérieur de l'imageur. Afin de garder un contrôle et une architecture simple la quantité de donnée transmise par chaque macropixel doit être constante et identique pour chaque macropixel de l'imageur.

Ce résultat est obtenu en ne mettant pas de codeur entropique en sortie de la transformée. La compression est obtenue par une quantification agressive en abandonnant les coefficients basses fréquences qui seront donc considérés comme étant à zéro lors de la reconstruction.

Une transformée en cosinus discrète est utilisée car elle offre d'une part une grande régularité de traitement et, d'autre part, elle peut être facilement reconfigurée pour que ses opérateurs soient utilisés dans d'autres types de filtrage. Classiquement ce type de transformée est appliqué sur des blocs de 8×8 pixels. Ici elle est appliquée sur les mantisses des macropixels comme illustré sur la figure 2, c'est donc une DCT 32×32 qui est utilisée. En plus d'éliminer les coefficients basses fréquences la quantification en sortie de la DCT ramène les coefficients conservés sur le nombre de bits des mantisses de départ (8 bits). Pour un macropixel ce sont ces coefficients qui sont transmis en plus de l'exposant commun (cf. figure 11). Hors de l'imageur, l'image est reconstituée en faisant la DCT inverse sur les coefficients reçus complétés par des 0 pour les coefficients qui ont été éliminés puis en multipliant par l'exposant commun du macropixel.

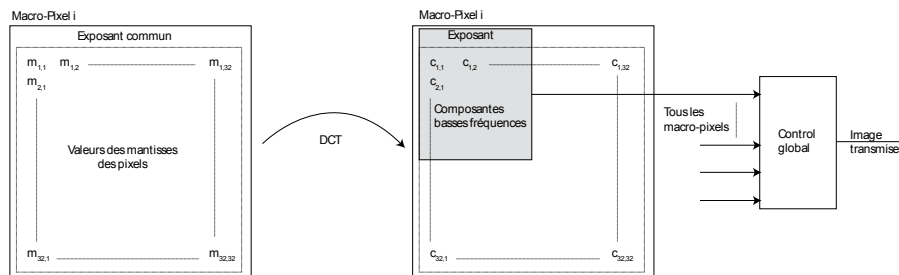


Figure 11. Compression d'un macro-pixel à l'aide de la DCT

L'impact du taux de compression et de la précision des coefficients de la DCT 32×32 ont été étudiés principalement en termes de PSNR sur l'image reconstituée après transformée inverse, réalisée hors de l'imageur (Guezzi *et al.*, 2010).

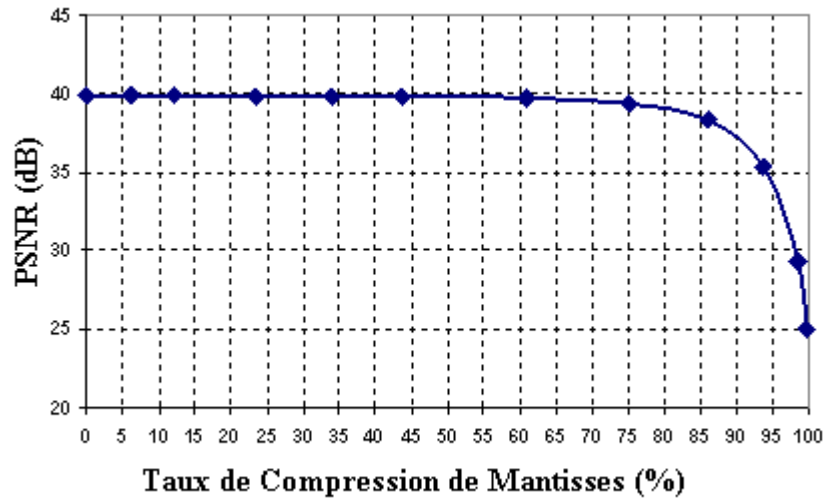


Figure 12. Courbe de PSNR d'une image reconstruite pour différents taux de compression (coefficients de la DCT sur 3 bits)

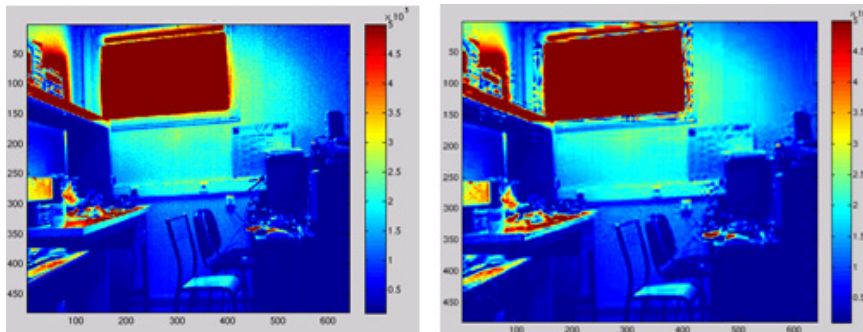


Figure 13. Image reconstruite après compression de 75 % sur les mantisses, (coefficients de la DCT sur 3 bits) : image à pleine échelle de luminance (gauche), image dans les faibles à moyennes luminance (droite)

La figure 12 représente le PSNR d'une image traitée par macropixel de 32×32 en fonction de différents taux de compression pour des coefficients de DCT codés sur 3 bits. Ces résultats ont été obtenus en simulation sous Matlab avec des calculs intermédiaires fait en pleine précision. L'impact des calculs en virgule fixe amènera sûrement un certain abaissement des valeurs du PSNR (étude en cours).

Un taux de compression sur les mantisses de 75 % peut être atteint tout en conservant une qualité d'image acceptable (PSNR d'environ 40 dB). Il est possible d'aller jusqu'à un taux de compression de 95 % pour un PSNR de 30 dB, ce qui constitue la limite minimale acceptable.

La figure 13 représente le résultat d'une compression avec un taux de 75 % sur les mantisses et des coefficients de la DCT sur 3 bits. L'effet de bloc de 32×32 devient visible et certains détails sont perdus, cependant l'image conserve sa grande dynamique et reste utilisable pour faire des post-traitements tels que de la détection de contours.

6. Conclusion

Nous avons proposé et analysé les performances d'une nouvelle architecture de capteur d'images HDR destinée à être mise en œuvre dans ces circuits à intégration 3D. Cette architecture permet d'adapter le temps d'intégration par groupe de pixels et, ainsi, d'étendre la dynamique de l'imageur à 120 dB tout en conservant les performances des pixels 3T et 4T classiques car, et ce point est essentiel, cette architecture permet de garder des pixels ayant une architecture classique (3T, 4T), en particulier sans leur rajouter de transistors supplémentaires. Le fonctionnement HDR est alors équivalent à un codage Mantisse-Exposant par groupe de pixels. Ce type de codage permet un premier niveau de compression, un deuxième niveau de compression, effectué sur la matrice de mantisses, étant obtenu par DCT suivi de la suppression d'une partie des coefficients de quantification.

L'architecture décrite nécessite peu de ressources électroniques supplémentaires. En effet, les 4 transistors requis pour chaque branche Winner Take All nécessitent une surface inférieure à celle de la photodiode plus les 3 ou 4 transistors du pixel. Un seul signal de commande supplémentaire par pixel est nécessaire. Les interconnexions entre les deux circuits n'introduisent donc que peu de contraintes sur le pas des interconnexions 3D.

Dans cette architecture, le calcul de l'exposant maximum se fait par le WTA analogique. Une telle solution permet d'obtenir une cellule compacte, cependant au détriment d'une erreur sur la détermination du temps d'intégration optimal. Néanmoins cette limitation peut aisément être contournée en rajoutant une pseudo-trame supplémentaire pour déterminer jusqu'à 10 valeurs d'exposant et ainsi atteindre les 120 dB.

L'association du codage Mantisse-Exposant et la double compression permet de réduire le flux de données de 22 bits par pixel à environ 2 bits par pixel ce qui est équivalent à un taux de compression de 93 %, tout en gardant une dynamique supérieure à 120 dB et une bonne qualité d'image (PSNR de l'ordre de 30 dB).

Bibliographic

- Akahane N., Adachi S., Mizobuchi K., Sugawa S. (2009, nov.). Optimum Design of Conversion Gain and Full Well Capacity in CMOS Image Sensor with Lateral Overflow Integration Capacitor. *IEEE Transactions On Electron Devices*, vol. 56, n°. 11.
- Bermak A., Kitchen A. (2006). A Novel Adaptive Logarithmic Digital Pixel Sensor. *IEEE Photonics Technology Letters*, vol. 18, n° 20, Oct. 15.
- Carvajal R.G., Ramirez-Angula J., Tombs J. (2000). High-speed high-precision voltage-mode MIN/MAX circuits in CMOS technology. *Circuits and Systems, Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on*.
- Guezzi Messaoud F., Dupret A., Peizerat A., Blanchard Y. (2011). A novel 3D architecture for High Dynamic Range image sensor and on-chip data compression. *Proceedings of the Sensors, Cameras, and Systems for Industrial, Scientific, and Consumer Applications XII*, San Francisco, SPIE 2011.
- Guezzi Messaoud F., Dupret A., Peizerat A., Blanchard Y. (2010). On-Chip Compression for HDR Image Sensors. *Proc.DASIP*, p. 90-96.
- Lazzaro J. (1988). *Winner Take-All Networks Of O(N) Complexity*. Computer Science Department Technical Report CALTECH-CS-TR-21-88, California Institute of Technology.
- Lopez-Morillo E., Carvajal R.G., Galan J., Ramirez-Angulo J., Lopez-Martin A., Rodriguez-Villegas E. (2006). A low-voltage low-power QFG-based Sigma-Delta modulator for electroencephalogram applications. *Biomedical Circuits and Systems Conference, 2006. BioCAS. IEEE*.
- Magnan P., Battude P., Gagnard X., Leyris C., Depoyan L., Vinet M., Cazaux Y., Giffard B., Ancy P., Coudrain P. (2009, June). Towards a Three-Dimensional Back –Illuminated Miniaturized CMOS Pixel Technology using 100 nm Inter-Layer Contacts. *International Image Sensor Workshop*, p.1-4.
- Mase M., Kawahito S., Sasaki M., Wakamori Y., Faruta M. (2005). A Wide Dynamic Range CMOS Image Sensor with Multiple Exposure-Time Signal Outputs and 12-bit Column Parallel Cyclic A/D Converters. *IEEE Journal of Solid-State Circuits*, vol. 40, n° 12.
- Rhee J., Joo Y. (2003 Feb.). Wide dynamic range CMOS image sensor with pixel level ADC. *IEEE Electron. Letters*, vol. 39, p. 360-361.
- Rhee J., Park D. (2009 May). Analysis and Design of a Robust Floating Point CMOS Image Sensor. *IEEE Sensors Journal*, vol. 9, n° 5.
- Soleimani M., Khoei A., Hadidi K., Kazemi Nia S. (2009). Design of High-speed high-precision voltage-mode MIN/MAX circuits with Low Area and Low Power Consumption. *IEEE*.
- Spivak A., Belenky A., Fish A., Yadid-Pecht O. (2009). Wide-Dynamic-Range CMOS Image Sensors Comparative Performances Analysis. *Electron Devices, IEEE Transactions on Electron Devices*, p. 2446-2461.
- Suntharalingam V., Berger R. (2005). Megapixel CMOS Image Sensor Fabricated in Three-Dimensional Integrated Circuit Technology. *IEEE International Solid-State Circuits Conference. ISSCC proceeding*, p. 356-357.

Vatteroni M., Covi D., Stoppa D., Crespi B., Sartori A. (2007). High Dynamic Range CMOS Image Sensors in Biomedical Applications. *Proceedings of the 29 th Annual International Conference of the IEEE EMBS*, Cité Internationale, Lyon, France, August 23-26.

Yang D., El Gamal A., Fowler B., Tian H. (1999, Dec). A 640×512 CMOS image sensor with ultrawide dynamic range floating-point pixel-level ADC. *IEEE J. Solid State Circuits*, vol.34, p. 1821-1834.

Fadoua Guezzi-Messaoud est doctorante au CEA-Leti. Diplômée de l'Ecole Nationale d'Ingénieurs de Caen (ENSICAEN) avec un M2R en microélectronique de l'Université de Basse Normandie, elle a débuté ses travaux de recherches en octobre 2009 au laboratoire L3I spécialisé en Imagerie.

Antoine Dupret est expert senior au CEA Léli et membre du Laboratoire d'Électronique et de Technologie de l'Information (LETI) après avoir été Maître de Conférences (IUT de Villetaneuse/Institut d'Electronique Fondamentale) puis professeur à l'ESIEE. Ses travaux portent sur les architectures de capteurs d'image et de traitements des signaux lumineux.

Arnaud Peizerat est ingénieur-chercheur au CEA de Grenoble. Il est membre du Laboratoire d'Électronique et de Technologie de l'Information (LETI). Ses travaux portent sur la conception de capteurs d'image en technologie CMOS dans les domaines de l'infrarouge, du visible ou des rayons X.

Yves Blanchard est Directeur Scientifique du French South African Institut of Technoloy (F'SATI). Ses travaux portent sur la modélisation et l'intégration d'architectures de traitement numérique de l'information.

