

Multirésolution adaptative pour un système de vision ultra faible consommation avec détection de mouvement et suivi intégrés

Integrated Motion Detection and Tracking for Ultra Low Power Vision Chip with Adaptive Acuity

Arnaud Verdant¹, Antoine Dupret², Hervé Mathias², Patrick Villard¹

¹CEA LETI-MINATEC, 17 rue des Martyrs, 38054 Grenoble Cedex 9, France
arnaud.verdant@cea.fr

²IEF, Bâtiment 220, Université Paris Sud 11, 91405 Orsay Cedex, France,

Manuscrit reçu le 14 mars 2008

Résumé et mots clés

Cet article décrit un système de vidéo surveillance ultra faible consommation. Basé sur un principe d'acquisition multi résolution, ce Capteur d'Image CMOS (CIC) réalise une détection de mouvement avec suivi intégré en appliquant une haute résolution sur les cibles, tout en maintenant le fond de la scène en sous résolution. Dans un contexte de camera fixe embarquée, cette adaptation de l'acuité de capteur suivant l'activité de la scène permet à l'imageur de se réveiller sur des régions d'intérêt et de réduire ainsi la quantité de données à mémoriser et traiter. Une réduction drastique de la consommation est ainsi obtenue. Les différentes étapes de traitement sont exposées, de même que les résultats de simulations obtenus à partir d'une description haut niveau intégrant les paramètres technologiques du capteur d'image.

Traitement analogique, capteur d'image CMOS, détection de mouvement, consommation d'énergie, multirésolution, architecture parallèle, système sur puce, suivi, vidéosurveillance, seuillage adaptatif.

Abstract and key words

Abstract – An ultra low power vision chip dedicated to video surveillance is presented. Based on a multiresolution acquisition scheme, this CMOS image sensor (CIS) performs integrated adaptive motion detection and tracking by focusing on targets with high resolution while keeping background in low resolution. In a context of embedded steady camera, this sensor's acuity adaptation to the scene activity by waking up on region of interest leads to a reduced amount of processed and memorized data. Drastic power consumption reduction is so achieved. Processing steps are exposed as well as simulations results obtained from a high level modeling of our image sensor taking into account process parameters.

Adaptive thresholding, analog processing, CMOS image sensor, motion detection, power consumption, multiresolution, parallel architecture, SoC, tracking, video surveillance.



Introduction

VIDEO systems dedicated to surveillance have to process and store a large amount of data, generally containing little relevant information. This situation leads to reduced power autonomy for embedded systems and to uselessly large memory size for captured videos.

In surveillance systems based on quasi-steady camera, adaptive environment modeling constitutes a key point to reduce the amount of information. In this paper we focus on left-behind vision sensors, for which a compromise has to be found between computational efficiency and robustness. Our way of addressing vision issues for left behind sensor is to design a smart CMOS image sensor (CIS) and the corresponding algorithms that may be implemented on it.

The approaches [2] to [5] focus on optimizing motion detection in CIS but are not concerned with very low power image processing. Other methods based on background modeling have been developed to perform low power motion detection in CIS ([5], [6]).

Further power saving may be achieved by combining multiresolution with motion detection algorithms. The idea is to wake up the areas of the pixel array where relevant events occur. This allows A/D converting and processing only within the selected Regions Of Interest (ROI). This simultaneously saves computation time and power because only a reduced set of pixels is processed. Using submicron CMOS technology, specific architectures performing some low level analog image processing at sensor level (before A/D conversion) can be designed without significantly impairing the device resolution and sensitivity. With such a physical implementation of ROI detection, the sensor performance and power consumption can be adjusted according to the scene activity.

Multiresolution architectures for CIS have been developed ([7] to [12]), leading to variable spatial acuity imaging, i.e. ROI. Image sub-sampling, including the «waking up» scheme stated above in a context of embedded video surveillance, has been explored in [13]. However, this sub-sampling technique significantly reduces fill factor [14].

In this paper we present an integrated ultra low power motion detection system based on the codesign of a specific SIMD CIS and a set of robust associated algorithms. The latter are based on a multi-resolution scheme, as to optimize at the same time the performance and the power consumption.

Algorithm Constraints Due To Architecture

Our goal being to perform motion detection for embedded video surveillance systems with high power efficiency requirements, we have to develop robust low power motion detection algorithms. The system on which these algorithms are implemented should also offer high compactness and low power consumption. Analog-based computational systems are an efficient response to these constraints. To face different situations found in video monitoring, an implementation based on a versatile architecture is thus proposed. The counterpart of this much-reduced power cost is that this architecture only allows the implementation of algorithms based on temporal or spatial kernel filters of reduced dimension.

Metrics definition and Performance Estimation Based on Benchmark Algorithms

Benchmark Algorithms

In our surveillance scheme, we aim at performing motion detection in unknown environments. In such configuration, algorithms meet hard constraints of robustness and adaptability. In order to perform adaptive motion detection, background modeling has been chosen because of its computationally efficient implementation ([5] and [6]).

1) Recursive Average [5]

This first algorithm relies on recursive operations.

The static global threshold implies limited adaptability, which constitutes the main limitation of this technique.

2) Sigma-delta modulations [6]

The algorithm presented in [6] features non-intensive computation operations and high robustness.

Instead of the global threshold used in RA, the $\Sigma\Delta$ algorithm computes a local adaptive threshold for each pixel signal to get more robustness on noisy elements, while keeping enough sensitivity on static background.

Performance evaluation metrics

The evaluation of the algorithm performance combines motion detection results together with the corresponding induced power consumption.

During the simulation, the motion detection result is compared to ground truth information for each frame to count the true and false positives and negatives.

In order to get a more accurate evaluation of the future performance of our architecture, a SystemC model of the architecture has been developed to co-simulate the proposed algorithms and the processing architecture [23].

Low Power Functioning Strategy

Image sub-sampling

Combining specific motion detection algorithms with images of low resolution allows only A/D converting in high-resolution areas of interest, i.e. where some motion occurs, while static areas remain in low resolution.

Figure 8 shows the considered low resolution schemes for a given image. The processed images are in each case split into a number of blocks corresponding to the desired low resolution.

Impact of subsampling on motion detection

Results show that Sigma-delta is better than Recursive average. However, we chose to implement a versatile architecture on which different algorithms could be implemented, to get the better performance regardless of the kind of scene processed.

Acquisition scheme

In low power mode, all the pixels remain active but only to compute the macropixels value. Only this spatial average value is then taken into account to detect motion.

The wake up function consists in switching from low resolution to high resolution when a sufficient temporal variation of a macropixel gray level is detected.

All the constitutive pixels of this ROI are then A/D converted and read out from the imager. Detecting macropixels variation is then equivalent to motion detection.



Improved Motion Detection on Low Resolution Image

Proposed motion detection algorithm

Our idea is to develop an algorithm based on RA computations, as to take benefit of the compactness of these operations in the analog domain, but at the same time to improve its robustness by introducing adaptive thresholding like in $\Sigma\Delta$ algorithm.

The principle of the proposed RAE algorithm (Recursive Average and Estimator) is to perform local adaptive thresholding with 'band pass filtering' to get rid of the high frequency background variations inducing false positives.

In order to further reduce false positives induced by noisy elements of the scene, an activation function is used in addition to the algorithm steps [25].

Motion detection results

Simulations have been performed for the considered RAE algorithm and the reference $\Sigma\Delta$ algorithm with a macropixel-based low-resolution configuration (figure 14), showing better results offers with RAE algorithm.

Effective Power Consumption And Performance Estimation



In order to check our power saving concept, we compared the estimated power consumption induced by our motion detection scheme, implemented on our analog architecture, and its counterpart for a digital implementation.

With our analog SIMD architecture, 21 micro-instructions per pixel are necessary to implement the presented RAE algorithm (Table 4). A power consumption of $301\mu\text{W}$ is thus obtained as long as the system remains in power saving mode. As a comparison, the $\Sigma\Delta$ algorithm requires 22 micro-instructions inducing a power consumption of $297\mu\text{W}$ for the same simulation configuration (Table 4).

These values are to be compared with the power consumption of about 10mW of a standard 320×240 APS CMOS sensor, with no processing performed.

Conclusion

A power efficient integrated sensor for video surveillance has been presented. It allows low power tracking by applying high resolution only on moving objects while keeping static parts in low resolution. The co-design approach we used allowed us to develop an efficient power saving strategy using a programmable analog architecture performing SIMD processing. Indeed, from the accurate SystemC model we have developed, we expect a power consumption far below 1mW, which represents 1/30 of the power of a stand alone imager. An efficient low-level motion detection algorithm has been developed that exploits the available compact operators. This algorithm shows an improved robustness compared to state of the art algorithms of the same class, featuring detection rate over 92% in most cases. This gain in robustness also leads to lower power consumption considering our motion detection scheme. A low power «pseudo-tracking» is thus achievable from only 1% of data with respect to full resolution. An associated power saving gain greater than 100 can thus be expected with respect to systems performing motion detection in full resolution. The on-chip processing thus proves to be a robust, low power way to address the issues of left behind video surveillance sensors.

1. Introduction

Les systèmes vidéo dédiés à la surveillance doivent être en mesure de traiter et de mémoriser une quantité importante de données, contenant généralement peu d'informations pertinentes, induisant ainsi une autonomie en énergie réduite pour les systèmes embarqués et une augmentation des ressources mémoires nécessaires pour le stockage vidéo.

Nous considérons ici le contexte de caméras fixes, pour lequel la modélisation adaptative du fond est largement utilisée. Parmi les nombreux travaux concernant la vision artificielle, la problématique de vidéo surveillance est étudiée dans [1], où des approches conventionnelles de détection de mouvement sont exposées. Les méthodes présentées sont efficaces mais requièrent de nombreuses opérations et, de ce fait, ne sont pas adaptées à notre principe de détection basse consommation. Nous avons également retenu une classe de capteurs spécifiques, dits « abandonnés », pour lesquels un compromis doit être établi entre complexité et robustesse des algorithmes. Pour répondre aux contraintes de consommation de tels capteurs, notre approche consiste à concevoir des imageurs CMOS intégrant des architectures versatiles sur lesquels peuvent être intégrés différents opérateurs de calculs.

Parmi les nombreuses études existantes sur la détection et le suivi de mouvement dans les CIC, des approches particulièrement intéressantes ont été proposées, dans le cadre de caméra fixe. L'intégration sur silicium de mesures du flot optique est une technique bien connue, dont l'implémentation est proposée dans [2] et [3]. La différence inter trame est aussi utilisée dans [4], où chaque pixel est codé sous forme de mantisse et exposant, permettant ainsi une grande dynamique de scène. Cependant, la comparaison étant effectuée sur les bits de poids fort, la détection de mouvements de faible amplitude ou de faible contraste s'en trouve limitée.

Les approches précitées s'intéressent à l'optimisation de la détection de mouvement dans les CIC mais n'intègrent pas la notion de détection très faible consommation. D'autres méthodes basées sur une modélisation du fond ont été développées en vue d'une détection induisant une consommation réduite. Ainsi, dans [5], une segmentation bas niveau du mouvement est décrite et dans [6], un algorithme performant basé sur des modulations Σ - Δ pour rétines artificielles est présenté. Dans ce travail, de manière à obtenir une robustesse plus importante dans les zones bruitées de la scène (e.g. bruissement de feuillage), un seuillage local adaptatif est réalisé. Pour chaque pixel, l'estimation du fond est calculée à l'aide d'opérations non-linéaires.

Une économie d'énergie plus importante peut être obtenue en combinant multirésolution et algorithmes de détection de mouvement. L'idée est ici de réveiller les zones de la matrice de pixels dans lesquelles un mouvement pertinent intervient. Cela permet de ne convertir numériquement et de ne traiter que les pixels inclus dans les régions d'intérêt (ROI) sélectionnées dans

l'image. Cette méthode permettra de réduire simultanément la consommation et les temps d'exécution des traitements du fait du nombre réduit de pixels à analyser. Par le biais des technologies CMOS submicroniques, des architectures spécifiques permettant l'exécution analogique de tâches de traitement d'image au niveau du capteur (en amont de la Conversion Analogique Numérique-CAN) peuvent être conçues sans réduire de manière importante la résolution et la sensibilité du système.

Avec une telle implémentation de la détection de ROI, les performances de l'imageur et sa consommation peuvent être ajustées suivant l'activité de la scène.

Des architectures pour CIC permettant une variation de l'acuité du capteur dans l'image ont également été développées. Dans [7] et [8], les auteurs présentent des architectures réalisant des images multi résolues en dehors de la matrice de pixels, en effectuant une moyenne en charges des informations de la matrice sur des bancs de capacités, augmentant néanmoins la surface du circuit. Dans [9], la multirésolution est obtenue en sommant les courants des pixels. Dans [10], une résolution variable est acquise par regroupement des pixels suivant leur luminosité, en utilisant un algorithme type quadtree, cette technique nécessitant néanmoins de lire et traiter l'ensemble des pixels de la matrice. Dans [11] et [12], des architectures permettant un partage de charges entre pixels au sein de la matrice sont décrites. L'ensemble de ces techniques nécessite globalement une augmentation du nombre de transistors dans le plan focal et de la surface des pixels, impactant de ce fait le fill factor, et donc la sensibilité et la résolution de l'imageur.

Dans [13], une technique de réveil avec lecture sous échantillonnée de l'image dans un contexte de vidéo surveillance embarquée est explorée. L'architecture permet deux modes de fonctionnement. En absence d'activité dans la scène, le capteur reste en mode de faible consommation, en scrutant la luminosité de 1% des pixels de la matrice. Si la variation de cette luminosité moyenne évolue de façon significative, le système bascule en mode actif et vient lire l'ensemble des données de la matrice. Cependant, cette technique réduit sensiblement le fill factor [14]. L'approche alternative utilisant dans [11] et [12] des macropixels est plus adaptée pour préserver ce paramètre à des valeurs raisonnables.

Bien que la détection de mouvement dans des scènes sous résolues puisse nécessiter des opérations relativement difficiles à réaliser dans un CIC [15], cette approche a été implémentée avec succès dans [16]. La multirésolution permet ici d'augmenter la cadence trame sur les parties en mouvement de la scène en réduisant la résolution de celles-ci, tout en maintenant une résolution spatiale plus importante sur le fond de la scène avec une cadence trame plus faible, dans le but de réduire l'effet de flou engendré par les mouvements rapides. Cependant, l'objectif de ce travail n'est pas de réduire la consommation du capteur.

Afin de réaliser une détection de mouvement intégrée ultra faible consommation, nous avons développé conjointement une architecture SIMD spécifique et les algorithmes de détection robustes associés. Ces derniers sont basés sur un mode de sous

S

résolution, de façon à optimiser à la fois les performances des algorithmes et la consommation induite. Dans ce papier, après une présentation succincte de nos choix architecturaux et des opérateurs disponibles (partie 2), nous décrivons les algorithmes de détection de la littérature pris en référence (partie 3). Nous comparons ensuite l'impact de différentes multirésolution réalistes sur les performances des algorithmes sélectionnés (partie 4). En partie 5, quelques améliorations apportées à cette classe d'algorithmes, afin d'améliorer leur robustesse, sont illustrées. Enfin, en partie 6, la consommation d'énergie estimée pour les différents cas traités est donnée et la solution la plus efficace d'un point de vue compromis robustesse/consommation est déterminée.

2. Contraintes algorithmiques imposées par l'architecture

Notre but étant de réaliser une détection de mouvement pour un système embarqué de vidéo surveillance avec de fortes contraintes en consommation, nous avons développé des algorithmes robustes requérant de faibles ressources de calcul. Le système sur lequel ces algorithmes sont intégrés doit également disposer d'opérateurs compacts et n'induire qu'une faible consommation. Un circuit analogique a donc été choisi en vue de répondre à ces contraintes. Dans un souci de compatibilité avec les différents environnements dans lequel évolue un tel dispositif, une architecture versatile est proposée. Dans la mesure où les traitements appliqués à chacun des pixels sont identiques et indépendants, de nombreuses opérations sont à effectuer simultanément. Une implémentation parallèle des algorithmes est donc intéressante pour réduire la consommation globale.

Dans de précédents travaux, une architecture de traitement SIMD semblable, dédiée à la réalisation de filtrages spatiaux [17-18] a été développée. Son efficacité d'un point de vue consommation est plus importante qu'un PowerPC dans le cas par exemple de l'implémentation d'un filtre de Deriche [19]. La consommation est en effet vingt fois plus faible (7mJ au lieu de 144mJ pour un PowerPC), avec une image de 1024 × 1024. La contrepartie de ce gain en consommation réside dans la complexité moindre des algorithmes intégrables, au sens de la taille réduite des noyaux temporels ou spatiaux des filtres intégrables. L'architecture programmable envisagée est donc une machine SIMD basée sur un traitement analogique des informations (figure 1). Cette architecture est composée, pour notre démonstrateur, d'une matrice de 320 × 240 photorécepteurs. Cette matrice de pixels permet l'acquisition des images en pleine résolution, mais également de générer les images sous résolues à partir desquelles s'effectue la détection de mouvement. Les images sous résolues sont réalisées sous forme de macropixels,

tel que nous le verrons en partie 4. L'image sous résolue à traiter est ainsi constituée de 32 × 24 macropixels. À chaque colonne de macropixels est associé un processeur possédant 120 points mémoire analogiques (5 points mémoire par macropixel). Le processeur analogique est un montage à base de capacités commutées permettant la réalisation d'opérations linéaires mais également conditionnelles. Un séquenceur numérique délivre les instructions successives. Les traitements étant effectués en bordure de matrice, une telle architecture autorise une taille réduite de pixels.

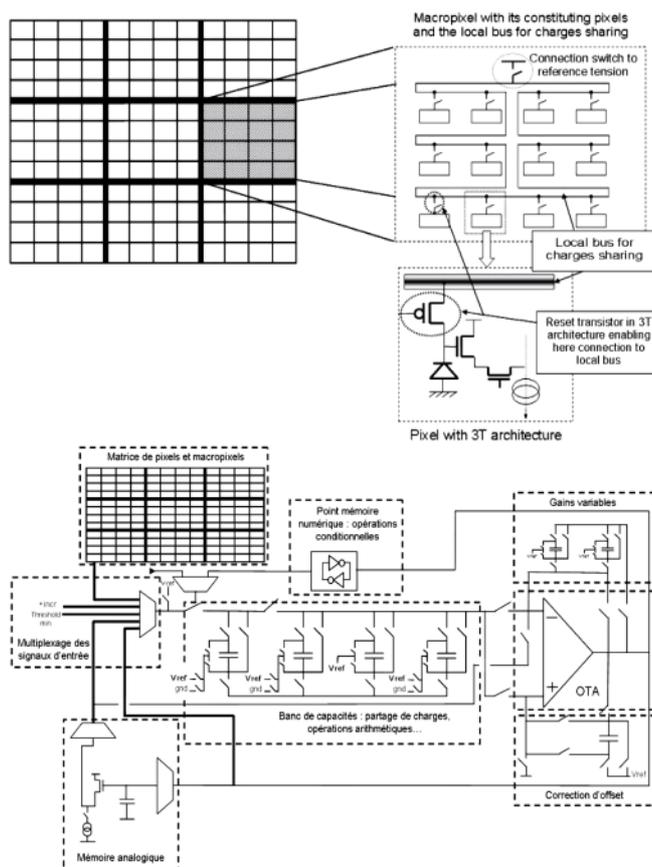


Figure 1. Architecture du capteur

Les diverses opérations réalisables grâce à cette architecture parallèle sont principalement les suivantes :

- Moyenne de valeurs de pixels
- Amplification
- Moyennes récursives (*i.e.* sommes pondérées)
- Σ - Δ (incrémention/décrémention de pas fixes)
- Mémorisation (état, résultats intermédiaires)

Les opérateurs les plus utilisés sont la somme, la multiplication d'une variable par un coefficient fixe, l'incrémention, la valeur absolue et la comparaison. Des opérations conditionnelles sont également possibles.

Avec une telle architecture, une moyenne spatiale de pixels peut être réalisée en effectuant un mélange de charges au niveau du pixel à l'image de [11] ou [12]. Une implémentation numérique

nécessiterait une première CAN des données à traiter et de nombreux transferts de données, induisant ainsi une consommation importante avant même le début de tout traitement algorithmique des valeurs de pixels.

3. Définition des métriques et estimation des performances avec les algorithmes de la littérature

A. Algorithmes de références

Le système de surveillance est conçu pour la détection de mouvement dans des environnements inconnus. Dans une telle configuration, les algorithmes sont soumis à de fortes contraintes de robustesse et d'adaptabilité. Dans le but de réaliser une détection de mouvement adaptative, le principe de modélisation du fond a été retenu en raison du faible coût en consommation et de son implémentation aisée.

Dans [5] et [6], deux techniques de référence permettant une modélisation adaptative du fond pour rétines artificielles ont été présentées. Ces algorithmes de l'état de l'art s'exécutent à partir d'opérations locales (*i.e.* pour chaque valeur de pixel) de manière à réaliser un filtrage passe bas de la scène observée.

Dans les paragraphes suivants, nous rappelons le principe de ces algorithmes et nous présentons par la suite les métriques utilisées pour évaluer leurs performances.

1) Moyenne réursive RA [5]

Une première technique est basée sur des opérations récursives. Soit S_n une valeur de pixel dont le fond RA_n est estimé avec une constante de temps importante fixée par N (1).

$$RA_n = RA_{n-1} - \frac{1}{N}RA_{n-1} + \frac{1}{N}S_n \quad (1)$$

La figure 2 illustre le filtrage passe bas d'un signal de pixel avec cette technique de moyenne réursive.

La figure 3 montre le fond estimé avec deux constantes de temps différentes. Le mouvement est considéré quand la différence absolue entre le fond estimé et le niveau du pixel traité est supérieur à un seuil global fixé *a priori* (2). Le seuil global fixe induit cependant une adaptabilité limitée, ce qui constitue le principal inconvénient de cette technique.

$$|RA_n - S_n| \geq \text{seuil} \implies \text{mouvement} \quad (2)$$

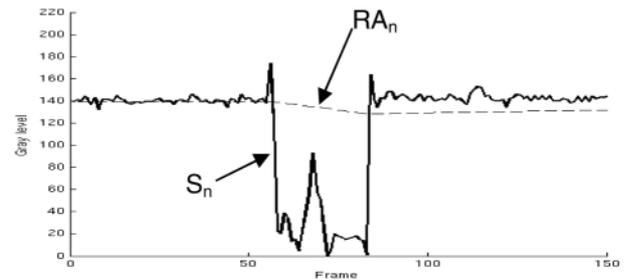


Figure 2. Estimation du fond par filtrage réursive [5].

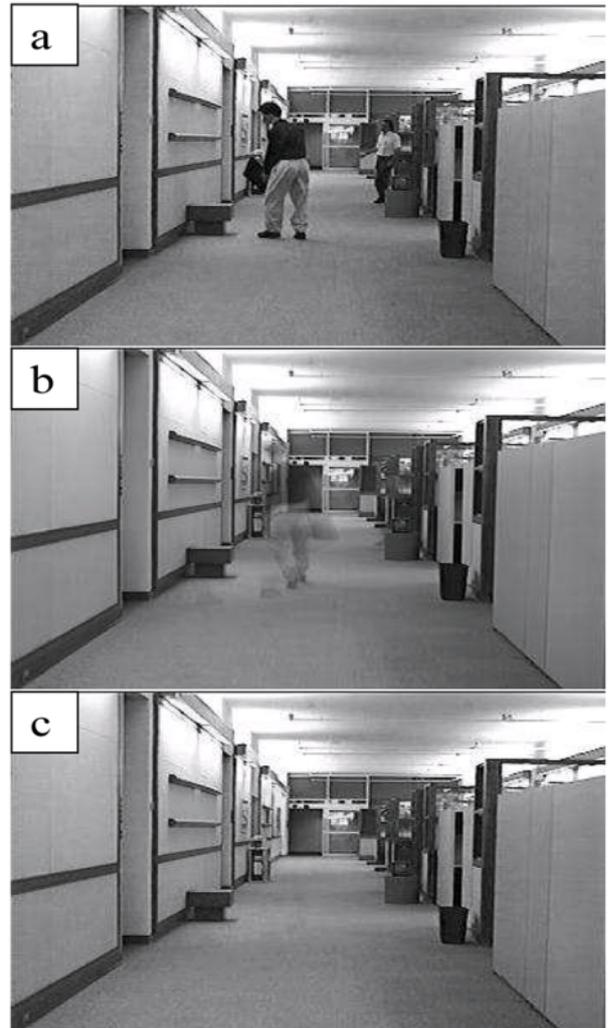


Figure 3. Fond estimé de l'image originale (a) (séquence Hall Monitor), avec $N = 2^5$ (b) et $N = 2^8$ (c).

2) Modulations Σ - Δ [6]

Les algorithmes markoviens sont connus pour donner des résultats de haute qualité et sont généralement utilisés pour la détection de mouvement en terrains inconnus. Cependant, ils doivent dans notre cas être simplifiés, en utilisant par exemple des algorithmes markoviens déterministes, dans le but de répondre aux contraintes de consommation et de ressources de calculs, tout en préservant leur robustesse.

La seconde méthode [6] introduit des opérations algorithmique-ment peu coûteuses avec une robustesse importante. Utilisant des opérations non linéaires basées sur des modulations Σ - Δ , cet algorithme suit le modèle de Markov utilisé pour des implémentations temps réel [20] et [21]. Les principales améliorations apportées sont une détection par estimation du fond plus robuste que des méthodes utilisant des différences inter-frames, et un seuillage local adaptatif. Suivant les comparaisons successives avec la valeur du signal (3), une variable M_n est incrémentée (5) ou décrémentée (4) par une valeur constante de manière à tendre vers la valeur du signal S_n .

$$\Delta_n = M_{n-1} - S_n \tag{3}$$

$$\text{si } \Delta_n > 0 \longrightarrow M_n = M_{n-1} - 1 \tag{4}$$

$$\text{si } \Delta_n < 0 \longrightarrow M_n = M_{n-1} + 1 \tag{5}$$

La figure 4 illustre le filtrage passe bas d'un signal pixel avec la méthode de modulation Σ - Δ .

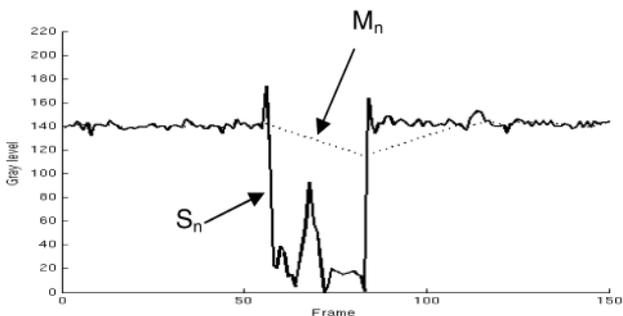


Figure 4. Estimation du fond par modulation Σ - Δ [6].

Comparé à la technique de moyenne réursive *RA* précédente et en considérant une implémentation analogique, le principal avantage de cette méthode réside dans la plus grande flexibilité obtenue. En effet, les variations d'estimation du fond peuvent être ajustées par les niveaux d'incrément/décrément alors que les constantes de temps des filtres passe bas type *RA* sont limitées par leur implémentation physique (dans notre architecture: les rapports de valeurs de condensateurs).



Figure 5. Résultat de l'estimation du fond sur la séquence Hall Monitor par modulations Σ - Δ . [6].

La figure 5 illustre l'estimation du fond avec la méthode de modulations Σ - Δ .

Une variable V_n basée sur les mêmes opérations conditionnelles que (4) ou (5) est générée pour la détection du mouvement. Cette variable est à l'image de la variance du signal et permet le seuillage de la différence absolue Δ_n entre la valeur S_n du pixel et le fond estimé (figure 6). Un mouvement est alors considéré lorsque la valeur de Δ_n devient supérieure à V_n .

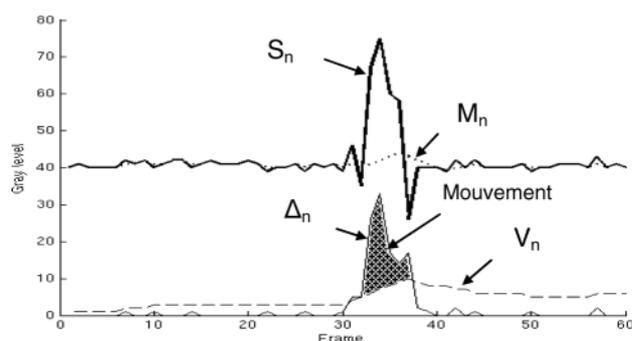


Figure 6. Évolution des variables de l'algorithme Σ Δ [6].

À l'inverse du seuillage global utilisé dans *RA*, l'algorithme $\Sigma\Delta$ génère un seuil adaptatif propre à chaque pixel afin d'obtenir une robustesse accrue sur les éléments parasites de la scène, tout en conservant une sensibilité raisonnable sur le fond statique.

B. Stratégie d'évaluation des performances

Nous devons évaluer les performances des algorithmes en fonction à la fois de la qualité de la détection et de la consommation induite par ces algorithmes.

Les résultats présentés en parties 4 et 5, ont été obtenus à l'aide de simulations MATLAB® sur différentes séquences en niveaux de gris avec caméra fixe représentant une grande diversité de conditions intérieures et extérieures: *Hall Monitor* (figure 3), *Walk* (séquence IEF), *Pets2002*, *dtneu_schnee* et *kwbB* (i21www.ira.uka.de) (respectivement a, b, c, d figure 7).

Durant la simulation, les résultats de détection du mouvement sont comparés à leur «vérité terrain» correspondante pour chaque trame afin de compter le nombre de faux positifs et négatifs générés (TP, TN, FP, FN). Nous avons utilisé les mesures de performance basées sur [22]: taux de détection (Detection Rate *DR* donné par: $DR = TP/(TP + FN)$) indiquant la capacité de l'algorithme à détecter les informations pertinentes, taux de fausses alarmes (False Alarm Rate, *FAR* donné par: $FAR = FP/(TP + FP)$) estimant la qualité de la détection, et taux de faux positifs (False Positive Rate, *FPR* donné par: $FPR = FP/(FP + TN)$). Un paramètre spécifique, le taux de mise en veille ($SB = (TN + FN)/\text{nombre de macro-pixels}$), donnant le pourcentage de l'image restant en sous résolution, a été introduit. Il est important de noter que la robustesse

de la détection impacte directement sur la consommation lorsque le nombre de fausses alarmes est élevé.

Dans le but d'obtenir une évaluation plus précise des performances de l'architecture, un modèle systemC a été développé afin de simuler conjointement les algorithmes et l'architecture de traitement [23]. Une description bas niveau des blocs analogiques constituant cette architecture a été réalisée, basée sur des blocs dont le niveau de granularité va jusqu'à la description de switches, capacités ou OTA. Pour l'ensemble de ces blocs élémentaires, les principales non idéalités ont été modélisées en fonction de la technologie CMOS employée. Le modèle ainsi obtenu permet une évaluation de l'impact des paramètres technologiques sur les performances des algorithmes en prenant en compte par exemple les dispersions, les non linéarités, les fuites ou le bruit temporel et spatial. De plus, la consommation de chaque bloc élémentaire a été modélisée afin d'évaluer fidèlement la consommation globale de l'architecture induite par les algorithmes implémentés.

Ainsi, l'ensemble de l'architecture a pu être validé fonctionnellement et ses performances effectives en termes de robustesse et de consommation déterminées. Les résultats correspondants seront présentés en partie 6.

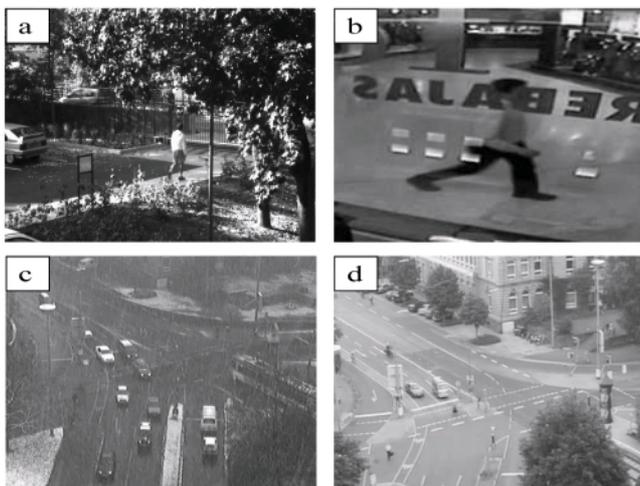


Figure 7. Séquences testées pour la détection du mouvement.

4. Détermination du mode de basse consommation

A. Techniques de sous échantillonnage spatial

La combinaison d'algorithmes spécifiques de détection avec des images de faible résolution permet de n'effectuer la CAN (*i.e.* appliquer une forte résolution) que sur les ROI de l'image, où le mouvement apparaît, tout en maintenant le fond en faible réso-

lution (*i.e.* absence de CAN ou conversion de la valeur moyenne uniquement).

Différentes méthodes de réduction de résolution spatiale ont été envisagées pour le mode de basse consommation de notre système. Celles-ci découlent de travaux existants ou de nouvelles propositions. L'objectif est à ce niveau de déterminer la méthode de sous résolution la plus pertinente en terme de compromis entre efficacité de la détection de mouvement et consommation des algorithmes.

La figure 8 illustre les différents modes de sous résolution envisagés. Les images traitées dans chaque cas sont divisées en un nombre de blocs correspondant à la sous résolution souhaitée. Quatre types différents de sous résolution ont été considérés :

- Macropixels (figure 8 a) : blocs contenant la moyenne spatiale de leurs pixels constituant ;
- Décimation de macropixels (figure 8 b) : une partie seulement des macropixels est traitée ;
- Bandes uniformes (figure 8 c), contenant la moyenne de chaque colonne de pixels ;
- Décimation de pixels (figure 8 d) : blocs dans lesquels seul le pixel central est traité.

Dans tous les cas, les paramètres tels que le taux de décimation, la taille des macropixels ou autres, sont liés aux propriétés des objets à détecter, par rapport à une profondeur de champ donnée. Pour toutes les configurations considérées, excepté la décimation de pixels, tous les pixels de chaque bloc sont actifs durant la phase d'acquisition de l'image. La moyenne spatiale de chacun est obtenue sans consommation supplémentaire (hormis l'activation de switches) en effectuant un partage de charges entre les photodiodes de pixels par connexions parallèles. Seule la moyenne spatiale (une valeur par bloc) est analysée pour la détection du mouvement. Ces trois configurations induisent ainsi une consommation très légèrement supérieure à la décimation de pixels, mais permettent en contre partie une meilleure détection du mouvement (voir paragraphe 5.B.).

Pour un algorithme donné, la consommation du capteur est directement liée à la quantité de données traitées. Par exemple,



Figure 8. Image originale et images sous résolues correspondantes obtenues avec les méthodes présentées.

dans le cas d'une sous résolution 64×48 obtenue à l'aide de macropixels (à partir d'une image originale 640×480), (64×48) données sont traitées, ce qui représente 1% de l'ensemble de données par rapport à la résolution initiale (figure 8). Un taux de réduction de données de 99% est ainsi atteint, induisant un taux de réduction de la consommation de même proportion.

B. Impact du sous échantillonnage sur la détection du mouvement

En se basant sur les algorithmes de références, nous avons sélectionné la technique de sous résolution la plus pertinentes vis-à-vis des métriques sélectionnées. Parmi les résultats exposés dans [24], le Tableau 1 illustre ceux obtenus pour les configurations en macropixels et décimation de pixels avec les algorithmes *RA* et $\Sigma\Delta$, sur les séquences présentées figure 7.

Les résultats obtenus avec les configurations de bandes uniformes et la décimation de macropixels conduisent à une réduction moyenne de *DR* respectivement de 2.9% et 2.5%, ainsi qu'à une diminution moyenne du *FPR* de 5.2% et 2.4%, par rapport aux résultats obtenus avec la configuration de macropixels. Le principal désavantage de la configuration en bandes uniformes est que lors du passage en haute résolution, une partie importante de l'image est convertie inutilement, induisant une consommation rédhibitoire. De plus, la détection du mouvement n'est efficace que pour des mouvements orthogonaux aux colonnes de la matrice. De plus, cette dernière configuration est également plus délicate à manier en terme de détermination du seuil car les colonnes couvrent l'image dans la totalité de sa hauteur et peuvent contenir à la fois des éléments statiques ou de bruit du fond. Dans ce cas, le seuil doit être suffisamment élevé pour ne pas engendrer de fausses alarmes, conduisant ainsi à diminuer la sensibilité de détection dans les zones statiques des bandes considérées. Concernant la configuration de décimation de macropixels, la réduction du fill factor engendre une diminution de *DR*.

Les tailles de blocs retenues pour les séquences testées sont les suivantes : *Monitor* (44×24 de 352×240), *kwbB* (74×56 de

740×560) et *Walk* (40×40 de 640×480). La dimension de ces blocs est déterminée de façon à être sensiblement de mêmes dimensions que les objets à détecter.

Le tableau 1 montre que les taux de détection sont élevés dans toutes les configurations. Compte tenu de l'application visée, le taux de fausses alarmes et le taux de faux positifs deviennent alors les critères importants. Dans les simulations réalisées, une robustesse plus importante est acquise avec l'algorithme $\Sigma\Delta$ qu'avec *RA*, grâce au seuil local adaptatif.

La détection de mouvement la plus pertinente est obtenue en utilisant la configuration en macropixels. De plus, une robustesse plus importante est obtenue dans le cas du seuil local adaptatif de l'algorithme $\Sigma\Delta$ en comparaison du seuil global fixe de l'algorithme de moyenne récursive *RA*. En effet, le paramètre *DR* de l'algorithme *RA* est légèrement plus haut dans la plupart des cas tandis que ses paramètres *FAR* et *FPR* sont significativement plus importants. Pour deux séquences (*Hall Monitor* et *kwbB*), l'algorithme $\Sigma\Delta$ présente des résultats significativement plus importants que la moyenne récursive *RA* (i.e. : *FPR* 3 à 4 fois plus faible, *FAR* 2 à 3 fois plus faible, à *DR* équivalent). Pour la séquence *Walk*, la moyenne récursive est significativement avantageuse par rapport à l'algorithme $\Sigma\Delta$ seulement pour le paramètre *FPR*. Ces éléments nous incitent donc à considérer l'algorithme $\Sigma\Delta$ comme étant plus performant que celui de moyenne récursive. Néanmoins, nous avons retenu une architecture versatile sur laquelle différents algorithmes peuvent être implémentés, pour obtenir les meilleures performances, indépendamment du type de scène.

C. Principe d'acquisition des images

En mode de faible consommation, tous les pixels restent actifs, mais seulement afin de générer les macropixels. Seule cette valeur de moyenne spatiale est prise en compte pour la détection du mouvement.

La fonction de réveil de l'imageur consiste à basculer de basse à haute résolution quand une variation temporelle critique du niveau de gris du macropixel est détectée. Si aucun mouvement n'est détecté au sein d'un macropixel, les valeurs des pixels constitutants ne sont ni converties ni transmises à l'extérieur de l'imageur. À l'inverse, si un mouvement est détecté, le macropixel bascule en mode haute résolution, simultanément à son voisinage. L'ensemble des pixels constituant de cette ROI est alors converti numériquement et transmis. La détection d'une variation de macropixel est assimilable à la détection du mouvement d'un objet dans la scène.

Afin d'effectuer un suivi de cible en appliquant une haute résolution sur cible, nous avons également considéré la transition inverse, lorsque le mouvement disparaît de la ROI: un bloc de pixel (i.e. un macropixel) repasse en mode de basse résolution si aucun mouvement n'est détecté dans l'ensemble du voisinage. Ceci implique également de maintenir l'information de basse résolution dans les zones de haute résolution afin de scrupuleusement

Tableau 1. Résultats obtenus avec les algorithmes *RA* et $\Sigma\Delta$ avec les configurations de décimation de pixels (*D*) et macropixels (*M*).

Séquences	Performances (%)					
	<i>DR</i>		<i>FAR</i>		<i>FPR</i>	
	<i>D</i>	<i>M</i>	<i>D</i>	<i>M</i>	<i>D</i>	<i>M</i>
<i>Hall (RA)</i>	96,2	98,7	43,5	16,8	16,0	4,0
<i>kwbB (RA)</i>	97,4	96,6	78,4	27,7	39,9	4,2
<i>Walk (RA)</i>	98,1	99,8	87,3	88,7	25,2	29,2
<i>Hall ($\Sigma\Delta$)</i>	98,0	94,3	56,7	8,3	27,7	1,4
<i>kwbB ($\Sigma\Delta$)</i>	94,6	99,1	56,3	10,3	13,8	1,0
<i>Walk ($\Sigma\Delta$)</i>	99,0	100	93,0	87,3	49,3	34,8

ter en continu les variations des moyennes. Le traitement de la moyenne spatiale de chaque bloc est donc à la fois réalisé en mode « réveil » ou « mise en veille » (figure 9). Compte tenu de la dimension des objets et de leur vitesse de déplacement nous avons retenu un voisinage 5×5 .

Le coût en consommation de cet algorithme est proportionnel à la quantité de données traitées. Cette réduction de la consommation est à relativiser par rapport au coût de la génération analogique de cette moyenne (charge et décharge des capacités des photodiodes). Cette consommation reste cependant négligeable par rapport au gain induit par la réduction de données traitées.

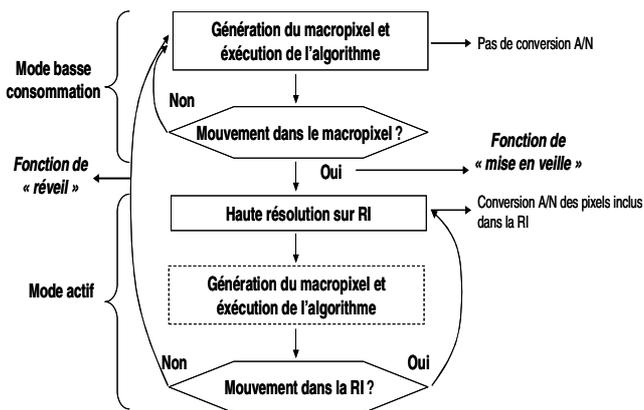


Figure 9. Comportement du capteur (ROI: région d'intérêt).

Un tel système permet donc à la fois un mode de suivi de cible par application d'une haute résolution mais également de passer d'un mode de faible acuité à un mode actif en amont du passage de la cible dans un macropixel.

En effet, dans le principe de détection de mouvement évoqué, la tâche de fond consistant à scruter les variations de la moyenne spatiale de chaque bloc à la fois sur les zones de basse et de haute résolution permet « d'anticiper » le passage en mode actif. Quand un bloc bascule en haute résolution, son voisinage passe également dans ce mode actif même si aucun mouvement n'y a été décelé. Quand l'objet atteint ce voisinage (déjà en haute résolution), le mouvement peut être détecté dans ces blocs, introduisant de nouveau le basculement en haute résolution de leurs voisins correspondants. De proche en proche, la région d'intérêt haute résolution est redéfinie naturellement en amont du déplacement de la cible. Un suivi haute résolution est ainsi obtenu.

En l'absence de cette tâche de fond, le réveil de l'imageur s'effectuerait juste à la suite de l'entrée de la cible dans les zones faiblement résolues, induisant ainsi une distorsion des objets détectés. La figure 10, illustre ce principe.

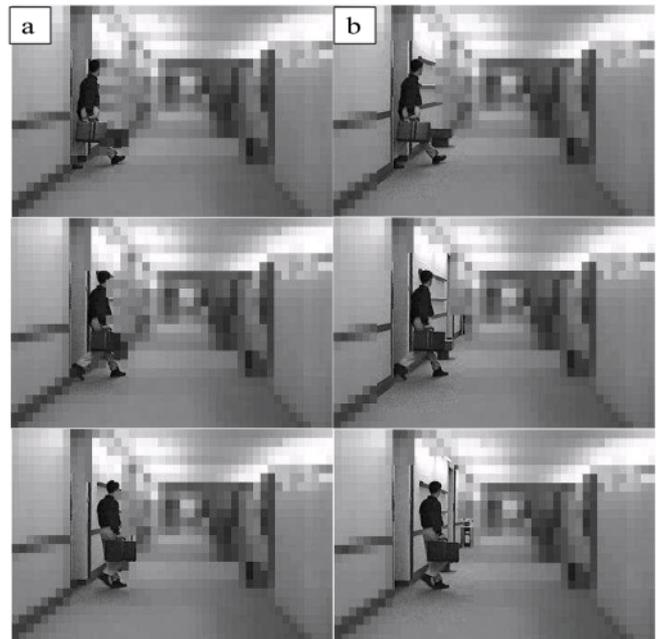


Figure 10. Amélioration de la détection de cible: colonne (a) sans tâche de fond: perte d'information sur la cible [24]; colonne (b) réduction de la distorsion de l'objet détecté.

5. Amélioration des algorithmes de détection de mouvement de référence

A. Algorithme proposé

Nous investiguons dans cette partie un algorithme permettant de détecter pertinemment les variations critiques des niveaux de gris des macropixels. Au tableau 1, nous avons pu montrer que l'algorithme $\Sigma\Delta$ était plus efficace que le RA dans cette configuration. Cependant, avec notre architecture analogique, une estimation du fond type RA (1) nécessite un nombre inférieur d'instructions (3 instructions) qu'avec les modulations $\Sigma-\Delta$ (3)-(5) (6 instructions).

Nous avons donc investigué un algorithme basé sur ce type de filtrage récursif RA, afin de bénéficier de la compacité de cet opérateur en analogique, tout en intégrant une méthode de seuillage adaptatif à l'image de celle exposée pour l'algorithme $\Sigma\Delta$. L'introduction du seuillage adaptatif permettant une amélioration de la robustesse de la détection est en effet cruciale pour notre stratégie de réduction de la consommation, car la génération de faux positifs se traduit par une perte d'énergie due à la CAN des pixels faussement détectés.

L'algorithme RAE envisagé (Recursive Average and Estimator) réalise un seuillage local adaptatif avec un filtrage de type

« passe bande », permettant la réjection des variations hautes fréquences du fond modélisé, à l'origine des faux positifs. L'algorithme proposé est basé sur la génération de deux moyennes récursives ($RA1_n$ (6) et $RA2_n$ (7)), ayant leurs propres constantes de temps (fixées par les paramètres N et M): la moyenne ayant la fréquence de coupure la plus basse ($RA1_n$) permet l'estimation du fond tandis que la plus haute ($RA2_n$) permet le filtrage des hautes fréquences du signal correspondant au « bruit » de la scène observée (figure 11).

$$RA1_0 = S_0 \text{ et } RA2_0 = S_0$$

$$RA1_n = RA1_{n-1} - \frac{1}{N} RA1_{n-1} + \frac{1}{N} S_n \tag{6}$$

$$RA2_n = RA2_{n-1} - \frac{1}{M} RA2_{n-1} + \frac{1}{M} S_n \tag{7}$$

$$\text{si } \Delta_n = |RA1_n - RA2_n| > k \cdot \delta_n \longrightarrow \text{motion} \tag{8}$$

n représente l'index de la trame en cours, S_n le niveau de gris courant du bloc traité et $k \cdot \delta_n$ le seuil local (8).

Ces opérations récursives, nécessitant peu de ressources mémoires, sont idéalement adaptées à une implémentation sur l'architecture présentée. La constante de temps de la moyenne récursive « rapide » ($RA2_n$) est déterminée de manière à permettre un filtrage efficace des hautes fréquences du signal, tout en induisant un déphasage réduit avec le signal, à l'origine d'effets de traînée. À partir de la transformée en z de la moyenne récursive, cette constante est exprimée en (9).

$$\frac{RA2(z)}{S(z)} = \frac{z}{M \left(z - \left(1 - \frac{1}{M} \right) \right)} = \frac{z}{M \left(z - e^{-\frac{\tau}{T}} \right)},$$

$$\text{avec } \tau = \frac{-T e}{\ln \left(1 - \frac{1}{M} \right)} \tag{9}$$

Le seuil adaptatif est obtenu par une amplification de la moyenne récursive δ_n de Δ_n (10). L'évolution du niveau de seuil $k \cdot \delta_n$ suivant les variations du signal est illustré figure 11. Par cette méthode, le seuil $k \cdot \delta_n$ dépend de l'activité temporelle de Δ_n . Si l'estimateur Δ_n devient supérieur au seuil local $k \cdot \delta_n$, le bloc correspondant et son voisinage basculent haute résolution. Δ_n agit à la manière d'un filtre passe bande ne conservant que les objets d'intérêt mobiles de la scène.

$$\delta_n = \delta_{n-1} - \frac{1}{P} \delta_{n-1} + \frac{1}{P} \Delta_n \tag{10}$$

La constante de temps de ce seuil doit également être suffisamment grande de manière à ne s'adapter qu'aux variations lentes du fond.

Dans cet algorithme, les trois constantes (M, N, P) sont fonction des propriétés des objets à détecter (principalement taille et vitesse) et de la cadence trame. Cependant, connaissant le type

d'objet à détecter, un seuillage local adaptatif est assuré par notre algorithme. Dans les paragraphes suivants présentant les résultats de simulation, ces constantes (M, N, P) ont été respectivement fixées à ($2^2, 2^4, 2^6$), avec une fréquence de 25 images par secondes.

Typiquement, la constante k est fixée à 2 et peut être augmentée en vue de réduire le nombre de faux positifs.

La figure 11 illustre le traitement d'un signal de macropixel avec l'algorithme RAE proposé.

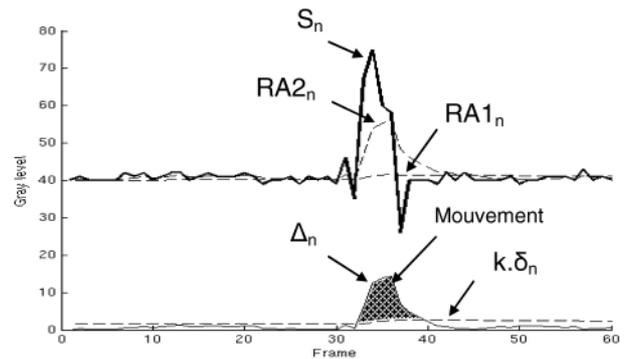


Figure 11. Traitement d'un signal avec l'algorithme RAE .

En vue de réduire le nombre de faux positifs induits par les éléments parasites de la scène, une fonction d'activation a été ajoutée [25]. Basée sur un scénario *a priori*, cette fonction favorise la détection de translations locales. Le passage en haute résolution d'un macropixel n'est ainsi autorisé que si le mouvement local observé est précédé dans les trames antérieures d'un mouvement ou d'une activation des macropixels voisins dans une direction donnée (figure 12).

Pour ce filtre, une mémoire numérique additionnelle est nécessaire dans notre architecture, afin de mémoriser l'état des macropixels des images précédentes. Nous avons retenu une profondeur typique de 5 images. Un circuit logique dédié réalise cette fonction d'activation.

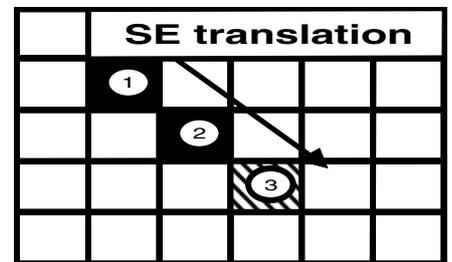


Figure 12. Verrouillage de la détection sur translations locales. Le mouvement est successivement détecté en 1 puis 2. Seul le macropixel 3 est alors autorisé à basculer en haute résolution.

Cette fonction d'activation est un moyen efficace d'améliorer l'adaptabilité des algorithmes de détection et permet de plus

d'améliorer leur robustesse dans le cas où la constante k (8) ne serait pas adaptée aux niveaux des perturbations de la scène. La figure 13 illustre la détection de mouvement réalisée sur la séquence *Walk* présentant un piéton marchant sous un arbre avec un bruissement important des feuilles. À la figure 13-b, k a été volontairement réglé à une valeur induisant de nombreuses fausses alarmes. En utilisant cette même valeur de k avec la fonction d'activation présentée, on peut alors voir figure 13-c la réduction du nombre de fausses alarmes générées.

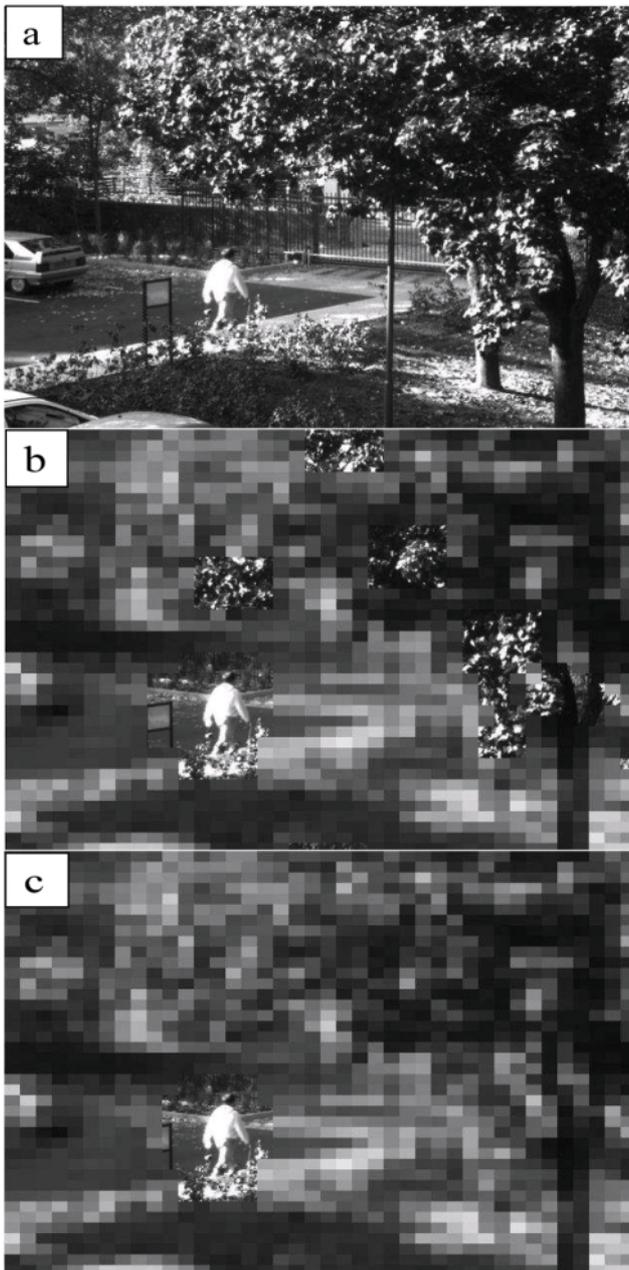


Figure 13. Illustration de la fonction d'activation : (a) image originale, (b) sans fonction d'activation, (c) avec fonction d'activation réalisant une restriction de la détection de mouvement aux translations locales seulement.

B. Résultats des algorithmes de détection

Les simulations ont été réalisées pour les algorithmes *RAE* et $\Sigma\Delta$ présentés avec l'application d'une sous résolution des images en macropixels sur les images testées.

Une taille spécifique de macropixels dépendant de la résolution initiale de l'image ainsi que de la taille des objets à détecter a été choisie pour chacune des séquences testées : *Hall* (44×24 de 352×240 , figure 3), *Walk* (40×40 de 640×480), *Pets2002* (64×24 de 640×240), *dtneu_schnee* (64×48 de 768×576) and *kwbB* (74×56 de 740×560) (respectivement a, b, c, d in figure 7).

Le taux de réduction indiqué pour chaque séquence représente la proportion de données non traitées par la technique de sous résolution en macropixels, par rapport un traitement de tous les pixels.

Tableau 2. Mesures de performances sans fonction d'activation (*NF*) et avec fonction d'activation (*F*) avec l'algorithme $\Sigma\Delta$ ($N = 3$).

Séquences (taux réduction %)	Performances (%)							
	DR		FAR		FPR		SB	
	NF	F	NF	F	NF	F	NF	F
<i>Pets 2002 (99)</i>	96,4	94,8	57,0	32,0	15,7	5,0	76,2	86,5
<i>Hall (98, 75)</i>	94,3	92,0	8,3	5,8	1,4	0,8	86,6	88,3
<i>kwbB (99)</i>	99,1	98,3	10,3	8,2	1,0	0,7	91,6	92,4
<i>dtneu_schnee (99,3)</i>	99,8	99,7	70,2	60,8	44,0	28,5	47,7	60,9
<i>Walk (99,5)</i>	100	99,8	87,3	76,9	34,8	15,0	57,8	75,5

Tableau 3. Performances avec l'algorithme *RAE*.

Séquences (taux réduction %)	Performances (%)							
	DR		FAR		FPR		SB	
	NF	F	NF	F	NF	F	NF	F
<i>Pets 2002 (99)</i>	94,6	92,7	15,9	16,5	2,1	1,9	88,4	89,7
<i>Hall (98, 75)</i>	96,9	96,6	12,9	12,4	2,6	2,4	83,4	84,0
<i>kwbB (99)</i>	99,0	98,0	25,4	23,2	2,8	2,2	90,1	91,2
<i>dtneu_schnee (99,3)</i>	99,5	98,6	23,9	22,6	5,3	4,6	81,5	83,1
<i>Walk (99,5)</i>	98,7	94,5	64,4	20,1	6,8	0,7	83,3	89,4

Les tableaux 2 et 3 donnent un taux de détection *DR* des deux algorithmes équivalent pour l'ensemble des séquences. Cependant, de meilleurs résultats sont obtenus avec notre algorithme concernant les paramètres *FPR* et *FAR* sur les 3 séquences ayant les perturbations les plus fortes (*Pets2002*, *Walk* et *dtneu_schnee*). Pour la séquence *kwbB*, les oscillations de caméra induisent des faux positifs plus nombreux du fait de l'effet de traînée dû aux filtrages récursifs. Cependant, malgré une légère détérioration du paramètre *DR*, la fonction de détection de translations locales apporte une amélioration des résultats des

paramètres *FAR* et *FPR*, notamment pour la séquence *Walk* avec des parasites importants dus au bruissement des feuilles.

Le paramètre de mesure *SB* montre globalement une portion plus importante de la scène restant en sous résolution avec notre algorithme, pour des taux de détection équivalents à l'algorithme $\Sigma\Delta$. Étant donné qu'une faible résolution implique une consommation réduite, un gain de consommation supérieur est escompté avec notre algorithme, comparé au $\Sigma-\Delta$, sans détériorer l'efficacité de la détection.

L'algorithme *RAE* combiné à la fonction d'activation permet donc un fort taux de détection *DR*, tout en conservant un taux de mise en veille *SB* important, même dans le cas de scène avec de fortes perturbations.

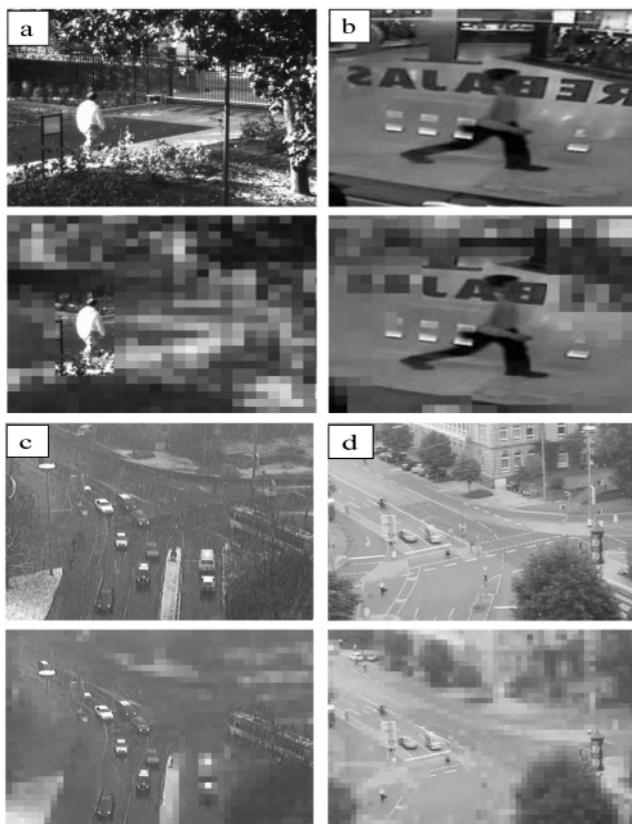


Figure 14. Séquences de test en haute et basse résolution.

Avec notre architecture analogique, 21 micro instructions sont nécessaires par macropixel pour l'algorithme *RAE* (Tableau 4). Une consommation de 301 μW est ainsi engendrée durant les phases de mises en veille de l'imager. Cette consommation a été simulée avec la séquence *Hall Monitor* 320×240 25 images/seconde et une sous résolution de 32×24 macropixels. La fréquence d'horloge des instructions est fixée à 100kHz.

L'algorithme $\Sigma\Delta$ requiert 22 micro instructions avec une consommation de 297 μW dans la même configuration de simulation (Tableau 4). Bien que l'algorithme $\Sigma\Delta$ nécessite 1 instruction supplémentaire, la consommation est légèrement inférieure du fait de la mémorisation d'une variable supplémentaire dans l'algorithme *RAE* et de calculs indépendants de ratios de capacités.

Pour ces deux algorithmes, la consommation supplémentaire induite par le séquenceur d'instructions doit également être prise en compte. En effet, les instructions de chaque algorithme sont stockées au préalable dans un registre numérique. Ainsi, la consommation de cette entité est directement liée à sa taille mémoire. Pour notre CIC, chaque instruction paramètre l'état d'environ 84 switchs. Avec une fréquence trame de 25Hz et 24 lignes de macropixels à traiter, cette mémoire numérique doit être lue aux fréquences de 1.1 MHz pour les algorithmes *RAE* et $\Sigma\Delta$. Pour la fonction d'activation, $(24 \times 32 \times 5)$ données sont lues, correspondant à une fréquence de lecture de 96kHz. En conséquence, la fréquence de lecture de cette mémoire analogique est dans ces deux cas proche de 2MHz. Dans [26], une consommation de 3.27 mW est mesurée lors de la lecture d'une mémoire utilisant la même technologie $0.35 \mu\text{m}$ que celle de notre capteur. Nous estimons donc cette consommation à 65 $\mu\text{W}@2\text{MHz}$ dans notre cas. Pour les algorithmes *RAE* et $\Sigma\Delta$, des consommations globales de 365 μW sont ainsi obtenues.

Ces valeurs doivent également être comparées aux 10mW consommés par un capteur CMOS APS 320×240 , sans prétraitement de l'image. Un gain en consommation de 30 est obtenu avec notre système en comparaison de ce capteur effectuant la seule acquisition de l'image. Un gain supérieur peut également être obtenu si le traitement lié à la détection du mouvement est également pris en compte. En effet, nous avons pu calculer que la consommation induite par une implémentation duale numé-

6. Consommation d'énergie effective des algorithmes

Afin de valider la pertinence du système développé, nous avons comparé la consommation estimée de notre principe de basse consommation du capteur, intégré sur l'architecture analogique présentée, à son implémentation duale numérique.

Tableau 4. Simulation systemC avec prise en compte des paramètres technologiques.

Mesures	Algorithmes	
	<i>Sigma Delta</i>	<i>RAE</i>
<i>Instructions</i>	22	21
<i>Consommation en mode faible résol.</i>	297 μW	301 μW
<i>DR</i>	- 7,3 %	- 12,8 %
<i>FPR</i>	+ 5,1 %	+ 2,8 %

rique de notre système (PowerPC G4 (7447)) serait d'environ 200mW, en supplément de la consommation du capteur seul. Le gain en consommation est alors de 550.

Le tableau 4 illustre l'impact des paramètres technologiques extraits de simulations post layout sur les performances de détection pour chaque algorithme.

Les simulations effectuées à partir de la modélisation SystemC de notre architecture illustrent l'influence des paramètres technologiques sur les performances des algorithmes. Des réductions respectives du taux de détection DR de 7 à 13 % sont ainsi mesurées pour les algorithmes $\Sigma\Delta$ et RAE , avec des hausses du FPR de 5.1 % et 2.8 %.

En comparaison d'architectures de type rétine programmable, l'implémentation de l'algorithme $\Sigma\Delta$ exposée dans [5] requiert 312 instructions. La compacité de nos opérateurs analogiques permet donc de réduire les cadences de traitement des images. Les figures 15 et 16 illustrent respectivement l'impact des paramètres technologiques sur DR avec l'algorithme RAE et sur FPR avec l'algorithme $\Sigma\Delta$. Néanmoins, étant donné le nombre légèrement inférieur de faux positifs induits avec notre algorithme RAE , un gain en consommation plus important peut être envisagé du fait du nombre réduit de pixels convertis numériquement par erreur.

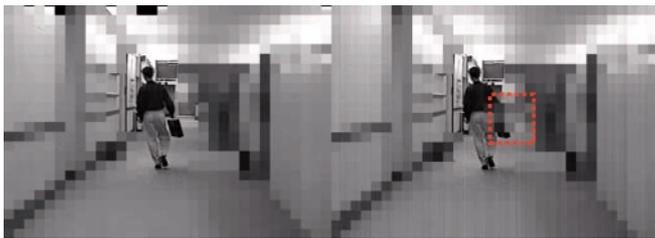


Figure 15. Image de gauche – simulation Matlab de l'algorithme RAE ; Image de droite – simulation SystemC prenant en compte les paramètres technologiques induisant une réduction de DR .



Figure 16. Image de gauche – simulation Matlab de l'algorithme RAE ; Image de droite – simulation SystemC prenant en compte les paramètres technologiques induisant une hausse de FPR .

7. Conclusion

Nous avons présenté l'architecture et l'évaluation d'un capteur faible consommation dédié à des applications de vidéosur-

veillance. Nous avons développé un algorithme de suivi de cible basse consommation obtenu à partir d'images en haute résolution focalisées sur les objets mobiles tout en maintenant le fond de la scène en basse résolution. L'approche co-design employée nous a permis de développer une stratégie d'économie d'énergie par le biais d'une architecture analogique programmable de type SIMD. Un système de «pseudo suivi» très faible consommation fonctionnant à partir de l'analyse de 1 % des données de l'image a été présenté. Les simulations SystemC effectuées permettent ainsi d'envisager une consommation de notre capteur de l'ordre de 1/30 de la consommation d'un capteur seul.

Un algorithme de détection de mouvement bas niveau exploitant les opérateurs compacts disponibles a été réalisé. Cet algorithme présente une robustesse accrue en comparaison des algorithmes de l'état de l'art de la même classe. Ce gain en robustesse permet également d'assurer la pertinence de notre principe de détection basse consommation, en réduisant le nombre de fausses alarmes et le nombre de pixels à convertir numériquement.

Les tests à venir de l'ASIC réalisés permettront de vérifier ces estimations obtenues à l'aide de nos modèles.

Références

- [1] W. HU, T. TAN, L. WANG, and S. MAYBANK, « A Survey on Visual Surveillance of Object Motion and Behaviors », IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, Vol. 34, N°3, 10/2004.
- [2] MOINI A. and *al.*, «An insect vision based motion detection chip», Solid-State Circuits, IEEE, Volume 32, Issue 2, Feb.1997 Page(s): 279-284.
- [3] MEHTA S., ETIENNE-CUMMINGS R., «Normal Optical Flow measurement on a CMOS APS imager», Circuits and Systems, ISCAS'04, IEEE, Volume 4, 23-26 May 2004 Page(s): IV-848-51 Vol.4.
- [4] MILIRUD V. and *al.*, «A wide dynamic range CMOS active pixel sensor with frame difference», Circuits and Systems, ISCAS 2005, IEEE, 23-26 May 2005 Page(s): 588-591 Vol. 1.
- [5] RICHEFEU J., MANZANERA A., «Motion Detection with smart sensor», 9th Congress Young Searchers in Computer Vision-ORASIS 05, May 05.
- [6] MANZANERA A., RICHEFEU J., «A robust and computationally motion detection algorithm based on background estimation», ICVGIP'04, December 2004.
- [7] KEMENY SE. and *al.*, «Multiresolution image sensor», Circuits and Systems for Video Technology, IEEE, Volume 7, Issue 4, Aug.1997 Page(s): 575-583.
- [8] PAIN B. and *al.*, «Dynamically reconfigurable vision with high performance CMOS active pixel sensors (APS) Sensors », Proceedings of IEEE, Volume 1, 12-14 June 2002 Page(s): 21-26 vol. 1.
- [9] COULOMBE J. and *al.*, «Variable resolution CMOS current mode active pixel sensor», Circuits and Systems, Proceedings. ISCAS 2000 Geneva, IEEE, Volume 2, 28-31 May 2000 Page(s): 293-296 vol.2.
- [10] ARTYOMOV E., YADID-PECHT O., «Adaptive multiple resolution CMOS active pixel sensor», Circuits and Systems, 2004. ISCAS'04, Volume 4, 23-26 May 2004 Page(s): IV-836-9 Vol.4.
- [11] SAFFIH F. and *al.*, «Multiresolution CMOS Image Sensor», Technical Digest of SPIE Opto-Canada 2002, Ottawa, Ontario, Canada 9-10, May 2002, p.425.

- [12] MASSIE M. and *al.*, «Vision chip for navigating and controlling Micro Unmanned Aerial Vehicles», Circuits and Systems. ISCAS'03. Volume 3, 25-28 May 2003 Page(s): III-786-III-789 vol.3.
- [13] SUNG-MIN SOHN and *al.*, «A CMOS image sensor (CIS) architecture with low power motion detection for portable security camera applications», Consumer Electronics, IEEE, Volume 49, Issue 4, Nov. 2003.
- [14] MCCARLEY P.L. and *al.*, «Large format variable spatial acuity superpixel imaging: visible and infrared systems applications», SPIE 2004, Orlando, FL.
- [15] MOELICH M., «Autonomous Motion Segmentation of Multiple Objects in Low Resolution Video Using Variational Level Sets», 2003, Lawrence Livermore National Laboratory.
- [16] JAEHYUK C., SANG-WOOK H., SEONG-JIN K., SUN-IL C., EUISIK Y., «A Spatial-Temporal Multi-Resolution CMOS Image Sensor with Adaptive Frame Rates for Moving Objects in the Region-of-Interest», ISSCC 2007.
- [17] MOUTAULT S., MATHIAS H., KLEIN J.-O., DUPRET A., «An improved analog computation cell for Paris II, a programmable vision chip», ISCAS'04.
- [18] VERDANT A., DUPRET A., MATHIAS H., VILLARD P., «Architecture for a low power image sensor with motion detection based ROI», ICECS 2007, IEEE.
- [19] KLEIN J.-O., LACASSAGNE L., MATHIAS H., MOUTAULT S., DUPRET A., «Low power image processing: analog versus digital comparison», CAMP 2005.
- [20] LACASSAGNE L., MILGRAM M. and GARDA P., «Motion detection, labeling, data association and tracking in real-time on risc computer», In Proc. IEEE ICIAP, pages 520.525, 1999.
- [21] DENOULET J., MOSTAFAOUI G., LACASSAGNE L., MÉRIGOT A., «Implementing Motion Markov Detection on General Purpose Processor and Associative Mesh», IEEE, CAMP 2005, 4-6 July 2005, Palermo, Italy.
- [22] BLACK J., ELLIS T.J. and ROSIN P., «A novel method for video tracking performance evaluation», IEEE, VS-PETS, pages 125-132, October 2003.
- [23] VERDANT A., DUPRET A., MATHIAS H., VILLARD P., «SystemC Validation of A Low Power Analog CMOS Image Sensor Architecture», Newcas 2007, IEEE.
- [24] VERDANT A., DUPRET A., MATHIAS H., VILLARD P., LACASSAGNE L., «Low power motion detection with low spatial and temporal resolution for CIS», CAMPS'06, IEEE.
- [25] VERDANT A., DUPRET A., MATHIAS H., VILLARD P., LACASSAGNE L., «Multirésolution adaptative pour imageurs CMOS très basse consommation», Colloque GRETSI 2007, IEEE.
- [26] JINN-SHYAN W., HONG-YU L., «A new current-mode sense amplifier for low-voltage low-power SRAM», Eleventh Annual IEEE International ASIC Conference, 13-16 Sept. 1998 Page(s): 163-167.



Arnaud Verdant

Arnaud Verdant obtained his Ph.D. at the University Paris XI, Orsay, France, in October 2008. He actually works at the CEA LETI and his work focuses on adaptive architectures dedicated to the in-sensor image processing, as well as low power read-out circuits of imagers.



Antoine Dupret

Antoine Dupret received the Ph.D. degree in electronics from Paris XI University, Orsay, France, in 1995. He is currently an Assistant Professor at the Institut Universitaire de Technologie de Villetaneuse, Villetaneuse, France, and is a Researcher in the Fundamental Electronics Institute, Paris XI University. He has been working on electronics retina, and his current interests include the modeling and the design of low-power vision circuits.

Hervé Mathias

Hervé Mathias received a Ph.D. degree in micro-electronics from the Ecole Centrale, Lyon, in 1996, for his work on analog integrated circuits layout automation. He then worked on the design of parallel integrated architectures for image processing, at Henri Poincare University, Nancy. He is currently associate professor at Institut d'Electronique Fondamentale (IEF), Paris XI University, where he's mainly working on modeling, design and testing of MEMS and their associated electronics.

Patrick Villard

Patrick Villard received the engineer degree from the Ecole Supérieure d'Electricité, France, in 1992, and the Ph.D. degree in physics, microelectronics option, from the University Paris XI, Orsay, France, in 1995. After its Ph.D. dedicated to the effects of ionizing radiation on CMOS integrated circuits, he joined the CEA-LETI where he worked on ASIC design for imaging and RFID applications. Since 2004 he has been head of the imagers & biochips IC design lab at CEA-LETI.