

Multirésolution adaptative pour imageurs CMOS très basse consommation

A. VERDANT¹, A. DUPRET², H. MATHIAS², P. VILLARD¹, L. LACASSAGNE²

¹CEA LETI - MINATEC, 17 rue des Martyrs, 38054 Grenoble Cedex 9

arnaud.verdant@cea.fr, patrick.villard@cea.fr

²IEF, Bâtiment 220, Université de Paris 11, 91405 Orsay Cedex

antoine.dupret@ief.u-psud.fr, herve.mathias@ief.u-psud.fr, lionel.lacassagne@ief.u-psud.fr

Résumé – Destiné à un imageur CMOS intelligent, nous présentons un algorithme robuste basé sur des opérations récursives, permettant l'adaptation de l'acuité du capteur en fonction de l'activité de la scène observée. L'intérêt principal de la détection de mouvement avec seuillage adaptatif décrite ici est de n'acquérir en haute résolution que les zones de la scène comportant un mouvement, le fond restant en sous résolution. Une réduction majeure de la consommation de puissance est alors obtenue, par une réduction drastique du nombre de données à traiter.

Abstract – To be implemented on an analog CMOS image sensor, a robust algorithm based on recursive operations is presented. It allows sensor's acuity adaptation to the scene activity. The main interest of motion detection with adaptive thresholding here described is that in a context of embedded steady camera, such a system allows focusing on targets with high resolution while keeping background in low resolution. Drastic power consumption reduction is achieved by tremendously reducing the amount of processed data.

1. Introduction

Les systèmes abandonnés de surveillance sont confrontés à la fois aux exigences de performances et de gestion de leur consommation. Une manière d'aborder ce problème est le développement d'architectures spécifiques pour le traitement local des images, au niveau du capteur (en amont de la conversion analogique numérique), permettant une gestion efficace de la consommation. Avec l'essor des technologies CMOS submicroniques, ce traitement local peut être réalisé sans dégrader de manière significative la résolution ou la sensibilité de l'imageur. Cependant, de nouveaux algorithmes spécifiques doivent être développés. La totale autonomie requise par ces capteurs impose à ces algorithmes d'être à la fois robustes et adaptables à divers environnements, tout en ne nécessitant pas de ressources importantes de mémoire ou de calcul, sources de consommation de puissance.

Notre approche repose sur l'implantation physique d'un système de détection et de suivi de cible permettant l'identification des zones d'intérêt de l'image avant toute interprétation humaine. Ce système permet d'ajuster la résolution du capteur en fonction de l'activité de la scène et donc de réduire la consommation en appliquant une pleine résolution sur les cibles et en maintenant le fond de la scène en sous résolution.

Parmi l'ensemble des études ayant trait à la détection et au suivi de mouvement, quelques méthodes

particulièrement efficaces ont été proposées. Le traitement vidéo par mesure du flot optique est par exemple une technique bien connue et son implémentation dans des capteurs d'images CMOS a été abordée dans [1] et [2]. La segmentation d'objets en mouvement par des techniques de différences au fond et seuillage adaptatif à partir d'opérations récursives a également été étudiée dans [3] et [4]. Dans [5], un algorithme robuste et efficace du point de vue des ressources requises (mémoires et calculs) est présenté. Il est implanté sur rétine artificielle et basé sur des opérations non linéaires telles que la modulation Σ - Δ .

Toutes ces précédentes approches sont motivées par l'amélioration de la détection du mouvement, mais ne sont pas concernées par le traitement très basse consommation des images. Une importante économie d'énergie est cependant possible en combinant multirésolution et algorithmes de détection de mouvement

Notre approche, consistant à « réveiller » l'imageur lorsqu'un événement se produit dans la scène, a également été abordée dans [6]. La basse résolution est ici obtenue en réalisant une décimation des pixels lus, puis en acquérant toute l'image en pleine résolution en cas de mouvement détecté, sans focalisation sur les cibles. Dans [7], nous avons comparé différentes techniques de basse résolution pour la détection de mouvement. La réalisation de moyennes spatiales locales carrées (macropixels) s'est avérée être la solution permettant la meilleure détection.

Dans ce papier, nous présentons brièvement l'architecture analogique programmable, ainsi que le principe de multirésolution proposé pour réduire la consommation. Nous décrivons ensuite succinctement l'algorithme de détection de mouvement avec seuillage adaptatif. Nous exposons par la suite les résultats obtenus, en comparaison avec un algorithme de référence. Enfin, une évaluation du gain en consommation est indiquée.

2. Architecture du système

2.1 Principe de détection

Le mode basse consommation de notre système est obtenu en réduisant la résolution de l'image à l'aide de macropixels. Un macropixel est la moyenne spatiale des niveaux de gris d'un carré de pixels. En mode basse consommation, cette moyenne est générée par la connexion physique des photocapteurs, au coût d'une consommation très réduite (quelques dizaines de nW). Seule la valeur du macropixel est prise en compte pour détecter le mouvement, avec une amélioration de 85% des résultats par rapport à une sous résolution avec décimation de pixels [6].

Dans [7], nous avons analysé la fonction de « réveil », i.e. passage de la basse résolution à la pleine résolution, lorsqu'une variation notable du niveau de gris d'un macropixel est détectée.

Nous présentons ici la fonction de transition inverse, i.e. de « mise en veille », lorsque aucun mouvement n'est décelé au sein d'une région d'intérêt, i.e. un macropixel et son voisinage 5×5 . Cette fonction implique la conservation de l'information faible résolution dans les zones « réveillées » dans le but de scruter les variations des moyennes de pixels (figure 1).

La combinaison du réveil du voisinage 5×5 et de la conservation de l'information basse résolution dans les zones « réveillées » permet donc à la fois la transition du mode actif au mode faible consommation, mais aussi d'appliquer une pleine résolution sur une cible avant son entrée dans un macropixel. Un suivi de cible haute résolution et faible consommation est alors réalisé. La figure 2 montre l'amélioration de la détection et du suivi obtenue par notre méthode.

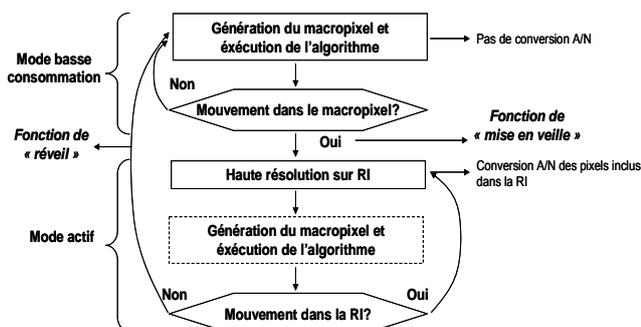


FIG. 1 : Fonctionnement du capteur (RI : Région d'Intérêt)



FIG. 2 : Amélioration de la détection de cible : a) image en pleine résolution ; b) résultats précédents de détection à partir d'image sous résolues [7] ; c) réduction de la distorsion de l'image avec notre nouvelle solution.

2.2 Architecture

Dans la mesure où les traitements appliqués à chacun des macropixels sont identiques et indépendants, de nombreuses opérations sont à effectuer en parallèle. L'architecture programmable envisagée est donc une machine SIMD basée sur un traitement analogique des informations [8] (figure 3).

Cette architecture est composée d'une matrice de $A \times B$ photorécepteurs à laquelle est associée une matrice de $A \times (mB)$ points mémoire analogiques, m représentant le nombre de points mémoire par pixel. La matrice ainsi formée est bordée d'un côté par un vecteur de A processeurs analogiques à capacités commutées. Une colonne de multiplexeurs permet la sélection des colonnes de macropixels ou mémoires à traiter par ces processeurs. Un séquenceur digital (CPU) délivre les instructions successives des processeurs.

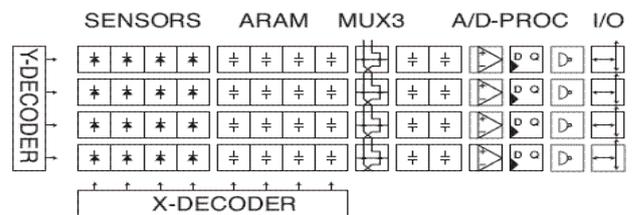


FIG. 3 : Architecture du capteur

L'efficacité d'une telle architecture d'un point de vue consommation a pu être mise en avant dans [9].

L'architecture choisie permet l'implantation à moindre consommation de puissances d'algorithmes faiblement itératifs, à noyau spatial ou temporel réduit.

3. Algorithme de détection

Nous avons pris pour référence l'algorithme présenté dans [5], réalisant une détection de mouvement à partir d'opérations non linéaires non intensives d'un point de vue itérations. L'algorithme proposé est basé sur la génération de moyennes récursives avec deux constantes de temps différentes ($RA1$ et $RA2$) : la plus rapide permet le filtrage des parasites haute fréquence du signal sans génération d'un effet de traînée trop important (constante de temps dépendante de N) (1); la plus lente permet une estimation

du fond (constante de temps dépendant de M) (2). Pour chaque pixel, les étapes de traitement sont décrites ci-dessous, où n représente l'indice de l'image courante, S_n le niveau de gris du macropixel traité et $k.\delta_n$ le seuil local (figure 4).

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

$$RA1_n = RA1_{n-1} - \frac{1}{N} RA1_{n-1} + \frac{1}{N} S_n \quad (1)$$

$$RA2_n = RA2_{n-1} - \frac{1}{M} RA2_{n-1} + \frac{1}{M} S_n \quad (2)$$

$$\text{if } \Delta_n = |RA1_n - RA2_n| > k.\delta_n \rightarrow \text{motion} \quad (3)$$

$$\delta_n = \delta_{n-1} - \frac{1}{P} \delta_{n-1} + \frac{1}{P} \Delta_n \quad (4)$$

Le seuil local adaptatif $k.\delta_n$ (3) est obtenu en amplifiant d'un facteur k la moyenne récurrente de la différence absolue Δ_n des deux moyennes récurrentes du signal δ_n (4). Ce seuil dépend alors directement de l'amplitude, de la période et de la persistance des perturbations du signal obtenues par le biais de Δ_n qui agit à la manière d'un filtre passe bande. Ce seuil doit néanmoins posséder une constante de temps relativement importante dans le but de ne prendre en compte que les perturbations persistantes. Nous considérons alors qu'une différence absolue Δ_n supérieure au seuil local $k.\delta_n$ (3) indique un mouvement. Le macropixel correspondant et son voisinage connexe passent alors en haute résolution.

Ces opérations récurrentes ne nécessitent pas de ressources de calcul ou de mémoire importantes et sont facilement implantables sur notre architecture.

Dans cet algorithme, les quatre constantes (k , M , N , P) dépendent des propriétés des objets à détecter (principalement la taille et la vitesse). Cependant, en ayant connaissance des types d'objets à détecter, un seuillage adaptatif local est assuré par cet algorithme.

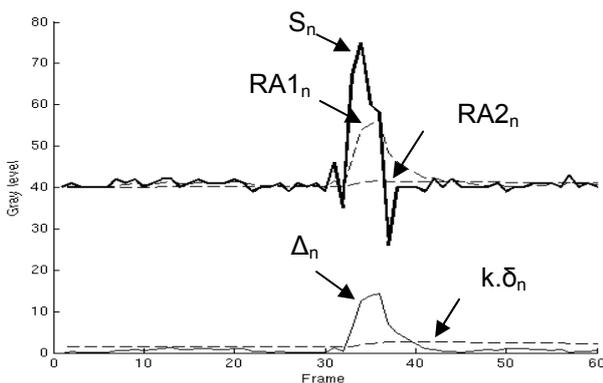


FIG. 4 : Traitement du signal temporel d'un pixel

Dans le but de réduire les faux positifs induits par les zones bruitées de l'image, un filtre supplémentaire a été ajouté. Dérivé d'une approche markovienne déterministe, ce filtre ne considère comme mouvement pertinent que les variations de macropixels inscrits dans une translation locale de mouvement. Ainsi, une variation suffisante du

signal temporel d'un pixel ne sera considérée comme un mouvement pertinent que si elle fait suite à des variations antérieures des pixels voisins dans une direction donnée (N, NW, W, SW, S, SE, E, NE) (figure 5).

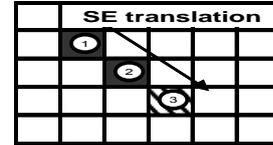


FIG. 5 : Détection de translation locales. Le mouvement est successivement détecté dans les pixels 1 et 2. Seul le pixel 3 pourra être considéré comme un mouvement pertinent dans les trames suivantes.

4. Résultats

4.1 Mesures de performances

Les simulations de détection de mouvement avec l'algorithme présenté ont été réalisées sous MATLAB©, avec une sous résolution en macropixels. Différentes séquences présentant une grande variété de conditions intérieures et extérieures ont été testées : *Hall Monitor* (figure 2), *dtneu_schnee* et *kwbB* (i21www.ira.uka.de), *Walk* (séquence IEF), *Pets 2002* (respectivement a, b, c et d sur figure 6).

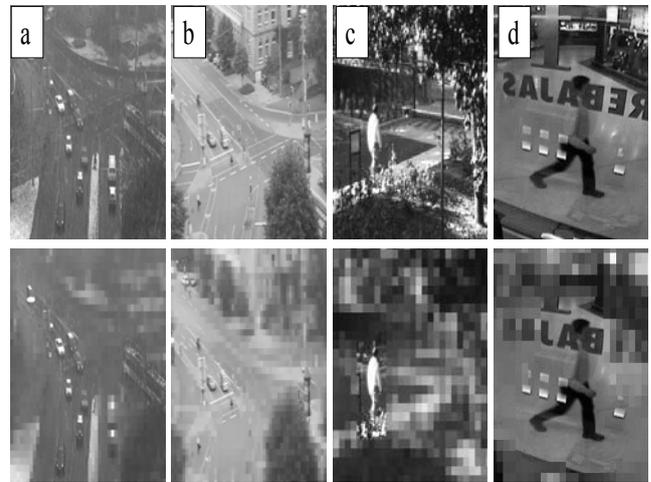


FIG. 6 : Séquences testées

Nous avons utilisé les mesures de performance basées sur [10] et présentées tableaux 1 et 2, respectivement pour l'algorithme Σ - Δ de référence et pour notre algorithme: taux de détection ((DR=TP/(TP+FN)), taux de fausses alarmes (FAR=FP/(TP+FP)), et taux de faux positifs (FPR=FP/(FP+TN)) avec TP, TN, FP, FN les vrais ou faux positifs ou négatifs. Un paramètre spécifique, le taux de mise en veille (SB=(TN+FN)/nombre de macropixels), donnant le pourcentage de l'image restant en sous résolution, a été introduit. 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.

Tab. 1 : Résultats avec algorithme Σ - Δ [5] (N=3) avec et sans fonction de détection de translation locale (F et NF)

Séquence (taux de réduction %)	Métrique (%)							
	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 (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

Tab. 2 : Résultats avec algorithme RAE

Séquence (taux de réduction %)	Métrique (%)							
	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 (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

Ces résultats indiquent des taux de détection DR équivalents pour les deux algorithmes. 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, la fonction de détection de translations locales apporte une amélioration des résultats des paramètres FAR et FPR.

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 Σ - Δ . Etant donné que faible résolution implique faible consommation, un gain de consommation supérieur est escompté avec notre algorithme, comparé au Σ - Δ , sans détériorer l'efficacité de la détection.

4.2 Consommation du système

Une modélisation SystemC de l'architecture a pu être réalisée [11] afin de valider sa cohérence et de prendre en compte l'impact des paramètres technologiques sur la robustesse de l'algorithme présenté. Ces simulations réalisées sur la séquences *Hall Monitor* (320×240), sous résolues en macropixels (32×24), ont permis de mesurer une diminution de 0.4% du taux de détection (DR) ainsi qu'une hausse de 0.5% du taux de faux positifs (FPR) et de 4.5% du taux de fausses alarmes (FAR).

Une estimation de la consommation a de plus été rendue possible par cette modélisation haut niveau de l'architecture. Le traitement par l'algorithme RAE requérant 21 instructions sur notre architecture induit une consommation de 301 μ W avec la sous résolution précédemment citée et une horloge de 40kHz. A titre de

comparaison, un traitement SIMD pleine résolution d'une matrice de 128×128 pixels [12] induit une consommation de 240mW.

5. Conclusion

Nous avons présenté un algorithme permettant une détection de mouvement faible consommation, à partir d'images sous résolues, en utilisant une architecture programmable déjà existante, pour des capteurs d'images CMOS. Une amélioration de la robustesse par rapport à un algorithme de détection de mouvement bas niveau a été présentée. Une pseudo-poursuite faible consommation est obtenue en ne traitant que 1% des données par rapport à la résolution originale des images.

Au-delà des applications de vidéo surveillance, ce système peut aussi servir de premier étage à la compression de flux vidéo.

Références

- [1] Moini A. and al., "An insect vision-based motion detection chip", Solid-State Circuits, IEEE, Volume 32, Issue 2, Feb. 1997.
- [2] Mehta S., Etienne-Cummings R., "Normal optical flow measurement on a CMOS APS imager", Circuits and Systems, ISCAS '04, IEEE.
- [3] Joo S., Zheng Q., "A temporal variance-based moving target detector", IEEE Workshop PETS, Jan. 2005.
- [4] Abdelkader M.F., Chellappa R., Qinfen Zheng, Chan A.L., "Integrated Motion Detection and Tracking for Visual Surveillance", ICVS '06, IEEE.
- [5] Manzanera A., Richefeu J., "A robust and computationally motion detection algorithm based on background estimation", ICVGIP'04.
- [6] Sohn S.M. 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.
- [7] Verdant A., Dupret A., Mathias H., Villard P., Lacassagne L., "Low power motion detection with low spatial and temporal resolution for CIS", CAMPS 2006.
- [8] Moutault S., Mathias H., Klein J.-O., Dupret A., "An improved analog computation cell for Paris II, a programmable vision chip", ISCAS '04.
- [9] Klein J.-O., Lacassagne L., Mathias H., Moutault S., Dupret A., "Low power image processing: analog versus digital comparison", CAMP 2005.
- [10] Black J., Ellis T.J., and Rosin P., "A novel method for video tracking performance evaluation", IEEE, VS-PETS, pages 125-132, October 2003.
- [11] Verdant A., Dupret A., Mathias H., Villard P., "SystemC Validation of A Low Power Analog CMOS Image Sensor Architecture", Newcas 2007, IEEE.
- [12] P. Dudek and S.J. Carey, "General-purpose 128×128 SIMD processor array with integrated image sensor", Electronic letters, 8th June 2006 Vol. 42 No. 12.