Estimation du Mouvement Global en Vue de son Intégration sur Silicium

Fabrice Gensolen, Guy Cathébras, Lionel Martin

▶ **To cite this version:**

Fabrice Gensolen, Guy Cathébras, Lionel Martin. Estimation du Mouvement Global en Vue de son Intégration sur Silicium. JNRDM: Journées Nationales du Réseau Doctoral de Microélectronique, 2004, Marseille, France. lirmm-00108646

HAL Id: lirmm-00108646

https://hal-lirmm.ccsd.cnrs.fr/lirmm-00108646

Submitted on 23 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation du mouvement global en vue de son intégration sur silicium

Fabrice Gensolen
LIRMM - UMR 5506
161, Rue Ada - 34392 Montpellier
STMICROELECTRONICS
ZI Rousset BP 2 – 13106 Rousset

Guy Cathébras
LIRMM - UMR 5506
161, Rue Ada - 34392 Montpellier

Lionel Martin
STMICROELECTRONICS
ZI Rousset BP 2 - 13106 Rousset

Email : fabrice.gensolen@st.com
fabrice.gensolen@lirmm.fr

Résumé

L'estimation du mouvement dans un système embarqué se réalise classiquement en post-traitement des images acquises et numérisées. Il est cependant possible de modifier l'architecture au niveau du pixel pour extraire des informations dès cette première étape d'acquisition de l'image. Notre cadre d'étude est l'estimation du mouvement global dans une séquence d'images, nous présentons dans cet article notre démarche de conception.

1. Introduction

Aujourd'hui, les progrès des procédés de fabrication CMOS permettent de réaliser des capteurs d'image dans des technologies ''quasi-standard'' en s'affranchissant des contraintes et des coûts liés à la réalisation des registres à décalages CCD.

Le développement de ces technologies permet d'envisager des modifications de la structure du capteur afin d'obtenir un véritable ''processeur'' d'images. De tels circuits entrent dans la catégorie des ''rétines électroniques'' que nous définirons mieux dans la section 2.

Le travail qui est exposé dans le présent article a pour objectif l'étude de l'intégration d'un système de mesure du mouvement global de l'image sur un ''imageur''.

Nous définissons plus précisément dans la section 3 ce que nous appelons le mouvement global, et les problèmes qui peuvent se poser pour le déterminer. Nous décrirons ensuite les principales méthodes d'estimation du mouvement global, et nous en dégagerons quelques techniques qui nous semblent les mieux adaptées à une intégration près du capteur. Puis nous présenterons notre méthode d'étude et de validation en vue de la réalisation d'un système de ce type.

2. Types de capteurs d'image et contraintes de conception

On peut classer les différents types de circuits photosensibles en deux catégories :

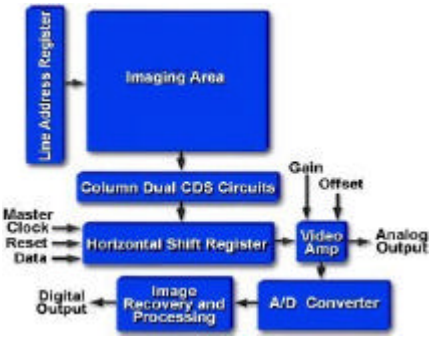

Figure 1 : Architecture d'un imageur.

- les *imageurs*, qui sont exclusivement dédiés à l'acquisition de l'image et à sa numérisation (cf Figure 1). Dans certains cas, il est possible de paramétrer la stratégie d'adressage (lecture) des pixels. Les paramètres peuvent être alors : le temps d'intégration des pixels, la vitesse de lecture, ou la zone de l'image acquise. Cette flexibilité s'avère intéressante dans le cadre de poursuite d'objets dans une scène par exemple, où la fenêtre d'intérêt est variable.

- les *rétines* ou *circuits de vision*, qui sont plus évolués et effectuent un traitement sur l'information lumineuse acquise pour en extraire une information , qui n'est pas nécessairement l'image. Ces systèmes sont :

  - soit *dédiés*, pour répondre à une application ou à un besoin spécifique. Un exemple de ce type de système est la souris optique [1].

  - soit *généralistes*, et programmables pour pouvoir ainsi réaliser des tâches diverses souvent basées sur des architectures SIMD (Single Instruction Multiple Data). Un exemple est la rétine programmable de deuxième génération [2].

Dans notre cas, nous nous intéressons à une rétine fournissant l'information de mouvement global, et ayant la fonction d'imageur. Dans l'état actuel de la technique, le nombre et la complexité des architectures réalisables sont très vite limités car tout ajout de composants

électroniques au niveau du pixel se traduit par une diminution relative de la surface photosensible et une augmentation de la taille générale du capteur. Cependant, on voit émerger des technologies séparant la fonction "transduction" du substrat (couche photosensible déposée en surface du circuit, technologies TFA [3]). Ainsi, il n'existe plus d'antagonisme entre surface photosensible et surface dédiée au traitement, ce qui autorise de nouvelles perspectives de réalisations, et de nouvelles applications.

3. Estimation du mouvement global

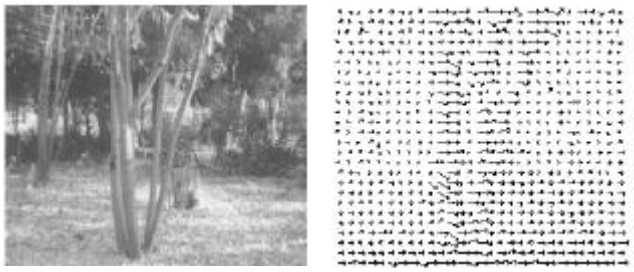

Figure 2 : Exemple de scène courante, et le champ de vecteurs vitesses apparent lors d'un mouvement de translation vers la droite.

Lors de l'acquisition vidéo d'une scène courante à l'aide d'un dispositif portable de type caméscope ou téléphone mobile, il apparaît plusieurs classes de mouvements (cf Figure 2).

Dans notre cas de l'estimation du mouvement global à des fins de stabilisation d'image, ou d'interface homme-machine comme par exemple la reconnaissance de mouvement intentionnel à partir du capteur, ce sont les mouvements de l'arrière plan fixe de la scène qui sont les plus porteurs d'information. En effet, l'information de mouvement global est alors directement donnée par la mesure de ces mouvements. Il paraît donc naturel de considérer en priorité ces régions là de l'image.

4. Méthodes envisagées

La mesure du mouvement global dans une séquence d'images doit être abordée par une analyse géométrique du mouvement. Il s'agit d'estimer les paramètres décrivant un modèle de mouvement global entre une image considérée comme référence, et l'image courante (classiquement deux images consécutives).

De plus, les mouvements intentionnels comme le zoom, ou le panoramique sont des mouvements lents et réguliers par rapport aux mouvements non voulus de types tremblements par exemple. Ainsi, les paramètres du mouvement intentionnel peuvent être estimés par une opération de type filtrage passe-bas.

Les choix critiques pour réaliser une estimation du mouvement global performante sont donc :

1- *le modèle du mouvement*, qui sera considéré dans le plan de l'image, afin qu'il ne soit pas trop complexe à estimer.

2- *la technique d'estimation des paramètres,*

3- *le filtrage du mouvement observé,* qui doit prendre en compte la nature des mouvements intentionnels.

En considérant l'aspect intégration sur silicium du capteur d'image, et particulièrement le fait que la somme des luminances de plusieurs pixels se réalise simplement en connectant à un même nœud leur photo-courant respectif, nous avons décidé de nous intéresser dans une première approche à la technique de Kim [4]. En effet, comme nous l'explicitons ci-dessous, cette méthode se base sur la somme des luminances des pixels de l'image.

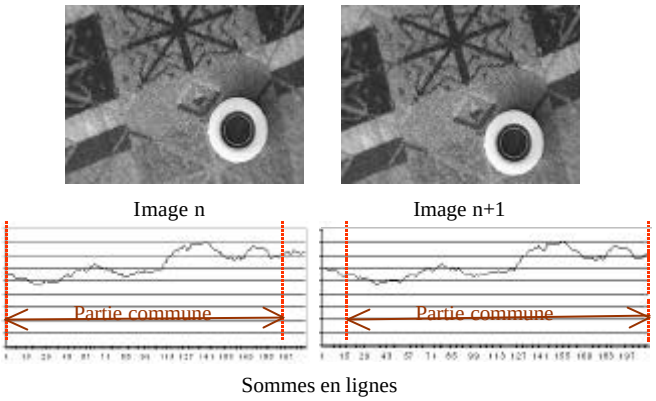

Figure 3 : Représentation de l'image en projection verticale, avec mouvement de la caméra vers le haut.

On considère ici un modèle de mouvement à deux paramètres, correspondant aux déplacements horizontaux et verticaux de l'image.

Le principe est de retrouver la partie commune entre deux images consécutives (cf Figure 3). En effet, lors d'un mouvement global d'une vidéo représentant une scène considérée comme un arrière plan fixe, d'une image à l'autre, une partie est identique, une autre est apparue, et une a disparue.

Pour estimer les paramètres du mouvement, nous raisonnons sur les projections en lignes et colonnes des luminances des pixels de chaque image.

A partir des projections verticales, le déplacement global vertical est estimé en recherchant le minimum de la différence absolue des vecteurs sommes en lignes de chaque image. Inversement, pour déterminer le mouvement horizontal. Un coefficient pondérateur est associé au résultat de cette somme des différences absolues, égal à la taille de la zone mise en correspondance. Le calcul de la correspondance recherche donc le minimum de la fonction suivante :

$$\xi = \frac{\sum_i abs\left[X_1(i) - X_2(i+D)\right]}{Pondération}$$

Où D représente le déplacement (taille de la zone apparue ou disparue des images), et X les vecteurs sommes des lignes ou des colonnes.

5. Environnement de simulation

Tout dispositif de traitement d'image implanté sous forme de rétine perçoit l'image sans aucune correction des

défauts du capteur. Il est donc important de pouvoir évaluer leur influence dès la phase de conception.

A cet effet, on peut utiliser un outil de simulation tel que SPW de Cadence, qui intègre à la fois des primitives de traitements d'images, ainsi que des modèles paramétrables de capteurs [5]. Ce sont des modèles décrits à haut-niveau (C++) pour une validation fonctionnelle du système, puis nous pourrons par la suite passer à une description structurelle (VHDL et VHDL-AMS) afin de valider cette fois les choix technologiques.

Nous avons dans un premier temps réalisé un banc de simulation au niveau fonctionnel en C++, nous permettant d'étudier les méthodes d'estimation du mouvement global.

Comme l'illustre la figure 4, le banc de simulation est constitué d'un bloc générique (ici appelé Kim), qui réalise l'estimation du mouvement entre les deux images présentes sur ses deux entrées. On automatise les simulations en appliquant la séquence vidéo désirée sur une entrée du bloc Kim, et la même séquence retardée d'une image sur la deuxième entrée. On récupère alors les mouvements obtenus dans un fichier ascii en sortie, pour pouvoir les analyser à posteriori.

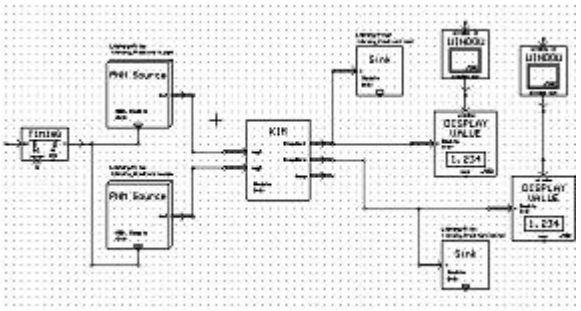

Figure 4 : Banc de simulation pour l'étude des techniques d'estimation de mouvement global.

Nous avons commencé par caractériser la méthode de Kim à partir de séquences d'images dont le mouvement était connu, et dans des situations choisies pour tester la robustesse de l'estimation du mouvement à certains paramètres. Les premiers résultats obtenus en considérant une scène comme un arrière plan fixe sont les suivants :

- nous obtenons une erreur nulle quelque-soit le mouvement de translation en entrée, pour des mouvements de la caméra pouvant représenter un tiers de l'image.
- la précision de mesure est le pixel.
- dans le cas où la luminosité globale de la scène varie d'une image à l'autre, l'estimation du mouvement est aussi correcte.

Par contre lorsque des perturbations interviennent dans la scène, comme le mouvement d'un objet, ou le changement de luminosité de certaines zones de l'image, l'estimation n'est plus parfaite. Et la caractérisation précise de ces phénomènes nécessitent de nouvelles acquisitions vidéo. On peut dore et déjà affirmer que cette méthode est utilisable dans une scène que l'on peut assimiler à un arrière plan, ce qui n'est pas suffisant dans notre champ d'applications.

6. Conclusion

Nous avons présenté dans cet article une approche de conception de rétine silicium pour l'estimation du mouvement global dans une séquence d'images. Nous avons précisé notre cadre de recherche, puis au vu des principales techniques d'estimation du mouvement existantes, nous en avons choisi une première, qui se prête bien à l'intégration. Nous avons ensuite mis en place une phase de validation de la méthode retenue à priori, en réalisant un banc de simulation modélisant à plus ou moins haut niveau les phénomènes liés à l'intégration dans le plan focal (sur silicium). Nous sommes actuellement à l'étape de validation fonctionnelle des méthodes, ce qui nous permettra ensuite de passer à la validation structurelle.

Références

[1] J. Tanner and C. Mead, "An integrated analog optical motion sensor," VLSI Signal Processing, II, S. Y. Kung, Ed. New York: IEEE Press, 1986, pp 59-76

[2] F. Paillet, D. Mercier, T. Bernard, E.Senn, "Low Power Issues in a Digital Programmable Artificial Retina," IEEE Alessandro Volta Memorial International Workshop on Low Power Design, IEEE Computer Society Press, Como, Italie, Mars 1999

[3] H. Fischer, J. Schulte, J. Giehl, M. Böhm, J. P. M. Schmitt, "Thin Film on ASIC - a Novel Concept for Intelligent Image Sensors," in *Mat. Res. Soc. Symp. Proc.*, vol. 285, pp. 1139-1145, 1992.

[4] Min-Kyu Kim, Ealgoo Kim, Daeyun Shim, Seong-Ik Jang, Gyudong Kim and Wonchan Kim, "An efficient global motion characterization method for image processing applications," IEEE Transactions on Consumer Electronics, Vol. 43, No. 4, November 1997

[5] SPW User's Guide, Cadence Design Systems, Product version 4.8.1, September 2002