

eISP, une architecture de calcul programmable pour l'amélioration d'images sur téléphone portable.

Thevenin Mathieu, Paindavoine Michel, Laurent Letellier, Heyrman Barthelemy

▶ **To cite this version:**

Thevenin Mathieu, Paindavoine Michel, Laurent Letellier, Heyrman Barthelemy. eISP, une architecture de calcul programmable pour l'amélioration d'images sur téléphone portable.. 2009. cea-00445727

HAL Id: cea-00445727
https://cea.hal.science/cea-00445727

Preprint submitted on 11 Jan 2010

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.

eISP, une architecture de calcul programmable pour l'amélioration d'images sur téléphone portable.

Mathieu THEVENIN[1], Michel PAINDAVOINE[2], Barthelemy HEYRMAN[1], Laurent LETELLIER[2]

[1]CEA LIST – Laboratoire Calculs Embarqués (LCE), Bat 528 PC 94 – F-91191 Gif sur Yvette – France

[2]Université de Bourgogne – LE2i UMR CNRS 5158 – Rue Alain Savary – 21000 Dijon

Mathieu.Thevenin@cea.fr,Laurent.Letellier@cea.fr,
Michel.Paindavoine@u-bourgogne.fr,Barthelemy.Heyrman@u-bourgogne.fr

Résumé – Les téléphones portables actuels intégrent des capteurs de résolutions telles qu'il est difficile d'allier surface silicium, consommation électrique (1 à 3 mm^2 pour 0,5 watt) et flexibilité pour les composants traitant les images qui en sont issues. Pour pouvoir répondre à ces contraintes, les intégrateurs utilisent des composants dédiés induisant des systèmes peu flexibles, pour lesquels il est impossible de faire évoluer les traitements d'images embarqués alors que la demande en flexibilité est croissant. C'est de ce contexte que ce papier présente la nouvelle architecture eISP (Embedded Image Signal Processor) qui est entièrement programmable. Cette architecture, validée en technologie TSMC 65 nm assurer la capacité de calcul nécessaire aux traitements de flux vidéo haute définition HD 1080p (16,8 GOPs à 233 MHz) pour une surface silicium de 1,5mm^2 et une consommation de 250 mW. Le support des formats vidéo HD, qui forment les futurs standards en téléphonie mobile, place l'architecture eISP parmi les plus performantes dans le domaine de l'embarqué avec sa capacité de calcul de 67 MOPs/mW.

Abstract – Today's smart phones, with their embedded high-resolution video sensors, require computing capacities that are too high to easily meet stringent silicon area and power consumption requirements (some one and a half square millimeters and half a watt) especially when programmable components are used. To develop such capacities, integrators still rely on dedicated low resolution video processing components, whose drawback is low flexibility. With this in mind, our paper presents eISP – a new, fully programmable Embedded Image Signal Processor architecture, now validated in TSMC 65nm technology to achieve a capacity of 16.8 GOPs at 233 MHz, for 1.5 mm^2 of silicon area and a power consumption of 250 mW. Its resulting efficiency (67 MOPs/mW), has made eISP the leading programmable architecture for signal processing, especially for HD 1080p video processing on embedded devices such as smart phone.

1 Introduction

Les capteurs vidéo font maintenant partie de notre quotidien et notamment dans les téléphones portables. Les contraintes de consommation électrique et de surfaces silicium sont drastiques. Il nécessaire d'associer au capteur un processeur de traitement du signal permettant non seulement la reconstruction des images couleur, mais aussi d'en améliorer la qualité intrinsèque. Les consommations électriques consenties dans le domaine de l'embarqué sont de l'ordre de quelques centaines de milliwatts pour une puissance de calcul de plusieurs milliards d'opérations par seconde. Aujourd'hui les intégrateurs utilisent des composants dédiés malgré leur manque de flexibilité. Par ailleurs la vidéo Haute Définition (HD), qui n'est pour l'heure pas supportée par les téléphones portables, nécessite une capacité de calcul telle (plusieurs dizaines de GOPS) qu'il est impossible de l'embarquer sur des dispositifs mobiles en utilisant les architectures programmables usuelles. Dans la mesure où les clients intégrateurs veulent souvent utiliser leurs propres fonctions d'amélioration d'image, il est crucial de rendre flexible et programmable l'ensemble de la structure de calcul en aval du capteur. Dans une première partie, différentes méthodes usuelles de reconstruction et d'amélioration d'image sont présentées, puis dans une seconde partie, une estimation des ressources nécessaires à leur exécution est donnée. Ensuite, l'architecture Embedded Image Signal Processor (eISP) est présentée. Cette architecture entièrement programmable est destinée à exécuter des traitements en HD 1080p – 1920×1080 pixels à 25 images par seconde – ce qui permet d'anticiper les capacités à venir des capteurs vidéo des futurs téléphones portables. Cette architecture, synthétisée en technologie TSMC 65 nm est caractérisée en surface, consommation et puissance de calcul.

2 Chaîne de reconstruction vidéo

Un ensemble de traitements sont nécessaires pour l'acquisition et l'amélioration des images issues des capteurs capteurs vidéo. De nombreux types de traitements et algorithmes peuvent être utilisés dans la chaîne de reconstruction de l'image.

Les images capturées ne couvrent généralement pas l'ensemble de la dynamique du capteur en raison de la difficulté à évaluer les paramètres d'exposition. C'est pourquoi une normalisation

d'histogramme[1] ou encore des méthodes locales adaptatives[2] permettent d'optimiser l'utilisation de la dynamique disponible pour le codage des valeurs des pixels.

La réduction du bruit est une étape essentielle. Le rapport signal sur bruit des images diminue avec la réduction de la taille des pixels ce qui est la tendance actuelle dans le domaine des dispositifs mobiles de grande consommation, mais d'autres types de bruits viennent dégrader l'image. Différentes approches permettent la réduction du bruit[3]. Le bruit spatial fixe est essentiellement lié à la non uniformité des propriétés du substrat et peut être caractérisé et supprimé. Le bruit électronique causé par l'agitation thermique des électrons, prend la forme d'un bruit blanc gaussien. Son effet peut être limité en appliquant un flou gaussien, ou encore un filtre local adaptatif. Enfin, le bruit dit « poivre et sel », est causé par des valeurs de pixels aberrantes par rapport à leur voisinage. Ce type de bruit est particulièrement perceptible à l'œil. Des filtres médians ou des techniques permettant de supprimer les valeurs de pixels extrêmes au sein d'un groupe sont adaptées pour la suppression de ce type de bruit[1].

Les systèmes mono-capteurs capturent une image de luminance. Pour reconstituer l'information couleur, le capteur est recouvert d'un filtre alternant les couleurs primaires. Le plus connu de ces filtres est le filtre de Bayer. La balance des blancs permet d'ajuster le gain entre ces différentes couleurs primaires pour assurer à l'image finale une teinte naturelle. La littérature propose de nombreuses méthodes allant du simple « grey world » au retinex. Le démosaïquage permet de générer trois plans pour chacune des couleurs (ou luminance/chrominances) à partir de l'image brute[4]. Enfin, il est possible d'améliorer la qualité visuelle des images par le rehaussement des contrastes et des contours.

Après avoir réalisé ces traitements, il est généralement nécessaire d'appliquer un rehaussement des contours, en utilisant un filtre passe-haut ou encore des filtres locaux adaptatifs. De même une amélioration des contrastes en vue d'un affichage de l'image est souvent nécessaire et peut aller de la simple correction gamma à des techniques plus élaborées de « tone-mapping » [5].

3 Ressources de calcul

L'étude d'algorithmes représentatifs des traitements décrits dans la section précédente permet de déterminer la puissance de calcul nécessaire à l'exécution d'une chaîne de reconstruction vidéo à des résolutions HD 720p et 1080p. Les traitements sélectionnés pour réaliser cette étude sont présentés en table 1. C'est ainsi qu'il apparaît que plus de la moitié des opérations sont dédiées au calcul d'adresses et au contrôle [6]. La capacité de calcul nécessaire à l'exécution d'une chaîne de reconstruction utilisant les algorithmes cités ci-dessus sur un flux HD 1080p – 1920×1080 à 25 images par seconde – est de 44,6 GOPS tandis que la part dédiée au calcul est de moins 15 GOPs. L'essentiel des ressources est utilisé pour le rehaussement des contours (51%), le filtrage médian (22%), et le démosaïquage de l'image brute (21%), comme le montre la table 1. Cette étude permet aussi de déterminer les opérateurs nécessaires au déroulement des algorithmes.

TAB. 1 – Exemple de ressources de calcul nécessaires à la reconstruction d'un flux vidéo HD 1080p.

Traitement	HD 720p 19,5 Mpx/sec (GOPs)	HD 1080p 51,8 MPx/sec (GOPs)	Représentativité (%)
Suppr FPN	0,39	1,06	2%
Bal. blancs	0,71	1,90	4%
Démosaïquage	3,46	9,23	21%
Médian	3,68	9,82	22%
Accentuation	8,49	22,63	51%
Total	16,74	44,65	100%

4 L'architecture eISP

Cette section décrit l'architecture programmable eISP, qui est en mesure de traiter des flux vidéo de différentes résolution et notamment HD 720p et HD 1080p. Sa programmabilité lui permet de supporter un large spectre d'algorithmes et notamment ceux nécessitant l'accès aux voisinages d'un pixel.

4.1 Etat de l'art

Actuellement la plupart de produits commercialisés intègrent des Systems On Chip (SOCs) construit autour de processeurs embarqués généralistes couplés à des unités de calcul dédiées. Ces solutions manquent toutefois de flexibilité puisque les traitements intégrés sont figés durant toute la durée du vie du SOC le plus souvent intégré au sein de différentes générations de produits. La forte demande du marché en termes de flexibilité au niveau des composants de traitement de vidéo en temps réel implique une importante activité des établissements de recherche, qu'ils soient académiques ou industriels. Par exemple, l'université de Standford et le Massachusetts Institute of Technology (MIT) proposent, au travers de la société Stream Processors Inc., une famille de processeurs[7] capable de traiter des flux vidéo en temps réel. Le processeur CRISP[8] pour « Coarse-Grained Reconfigurable Image Stream Processor » permet de supporter le traitement de flux vidéo 1080p. Bien qu'il ne s'agisse pas d'une solution entièrement programmable, sa flexibilité est améliorée au regard des composants dédiés. Il s'avère toutefois nécessaire de lui associer une mémoire d'image, particulièrement coûteuse en surface. La société SiliconHive licencie l'architecture HiveFlex[9] comportant 4 à 128 processeurs programmables, tandis que Xetal propose un processeur[10] programmable massivement parallèle intégrant 320 unités de calcul élémentaires. L'architecture SIMPil[11] est conçue pour

que chacun de ses 4096 processeurs travaille sur un bloc de pixels. Ces solutions peuvent difficilement être intégrées dans des systèmes fortement contraints en raison de leur importante surface silicium ou consommation électrique. De plus, nombreuses sont celles qui doivent être associées à une mémoire d'image ce qui nécessite plusieurs mm^2 en technologie TSMC 65 nm.

4.2 Parallélisme

En considérant un flux HD 1080p devant être traité en temps réel par un processeur fonctionnant à 233 MHz, seuls trois cycles d'horloge processeur sont disponibles pour réaliser les opérations de traitement, ce qui est insuffisant au regard des dizaines de cycles nécessaires pour la réalisation d'une simple convolution. Afin de maximiser le temps processeur disponible par pixel, il est crucial d'exploiter toutes les formes de parallélisme. Le parallélisme au niveau des instructions est supporté par l'utilisation de processeurs Very Long Instruction Word (VLIW). Le parallélisme spatial l'est par l'utilisation de plusieurs processeurs en parallèle en leur associant à chacun un pixel différent à traiter. Le parallélisme temporel est exploité en enchaînant les différents traitements ainsi qu'en réalisant les opérations de contrôle, d'adressage et d'accès aux données en temps masqué.

4.3 Tuile de calcul

L'architecture eISP est composée de tuiles de calcul programmables telles que celle présentée en figure 1. Elles intègrent des processeurs VLIW deux voies spécifiquement conçus pour limiter la surface silicium et la consommation électrique. Contrairement aux processeurs VLIW traditionnels, l'ensemble des opérateurs est mutualisé entre les deux voies d'où leur nom « SplitWay ». Un tel processeur, de largeur de chemin de 24 bits, présente une complexité de 5200 portes logiques, et peut être associé à une mémoire de travail supplémentaire de 2^{24} mots de 24 bits. Avec deux instructions par cycle d'horloge processeurs, sa capacité est de 466 MOPs à 233 MHz. Ces processeurs, fonctionnant en mode Multi Single Instruction Multiple Data (Multi-SIMD). Cette association constitue la partie calculatoire des tuiles, à laquelle est ajouté un gestionnaire de voisinages, dont la fonction est de transformer le flux de pixels entrant afin de le rendre directement accessible aux processeurs, tant au niveau du pixel que du voisinage considéré. Ceci, permet de masquer les temps d'accès aux données, et de supprimer les calculs d'adresses au niveau du programme. Ce gestionnaire de voisinage est conçu pour limiter la consommation et la surface silicium tout en permettant un accès direct à l'ensemble des données par les processeurs. La tuile comporte aussi un module d'entrées et sortie qui permet d'intégrer les valeurs des pixels du flux à traiter, mais aussi de reconstruire un flux avec les données calculées par la tuile. Une telle tuile peut intégrer typiquement quatre à seize processeurs. Leur fréquence de fonctionnement est déterminée lors de la compilation du programme, ce qui permet de gérer au mieux la

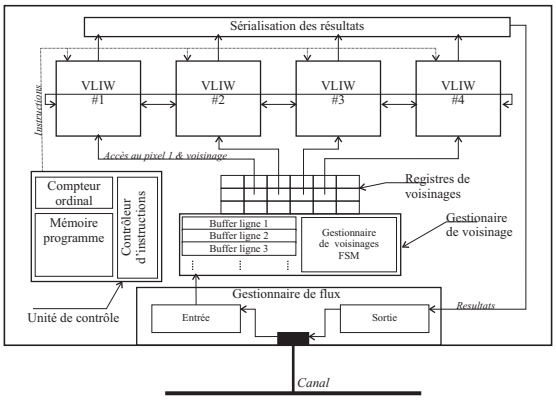


FIG. 1 – Exemple de tuile de calcul contenant 4 processeurs SplitWay, ce nombre peut être adapté à la surface disponible.

consommation du système. Afin d'exécuter des traitements complexes, plusieurs tuiles sont associées afin de former l'architecture eISP dans son ensemble. Le transfert des pixels aux différentes tuiles est réalisé par un bus Time Delay Multiplexed Access (TDMA) permettant de rendre configurable l'ordre dans lequel les tuiles se transmettent leurs données. Par ailleurs ce bus permet de maintenir l'ensemble des tuiles synchronisées.

5 Implémentation et résultats

Cette section présente le portage d'une chaîne de reconstruction d'image sur l'architecture eISP. Les algorithmes sont d'abord présentés puis leur arrangement sur les différentes tuiles de calcul est décrit. Les surfaces silicium et consommations électriques sont enfin présentées.

5.1 Algorithmes

Afin de valider le fonctionnement de l'architecture, une chaîne de reconstruction d'image est portée sur les processeurs SplitWay de l'architecture. Cette chaîne comporte les opérations suivantes :

- La suppression du bruit spatial fixe nécessite 5 cycles par pixel et ceci sans accès au voisinage du pixel à traiter.
- La normalisation d'histogramme nécessite 8 cycles par pixel, la mémoire de travail locale à chaque processeur contenant les histogrammes partiels.
- La réduction du bruit est réalisée par remplacement de la valeur du pixel central par la moyenne de ses voisins, si celui-ci n'en est pas proche (tolérance de 12,5% soit 1/8). Cette opération, moins coûteuse qu'un filtre médian, nécessite un accès au voisinage 5×5 du pixel et 14 cycles pour être réalisée.
- La correction gamma est réalisée en 4 cycles grâce à l'utilisation d'une LUT stockée dans une mémoire de travail locale aux processeurs.
- La correction de la balance des blancs est réalisée par la méthode du « monde gris » et nécessite 4 cycles.

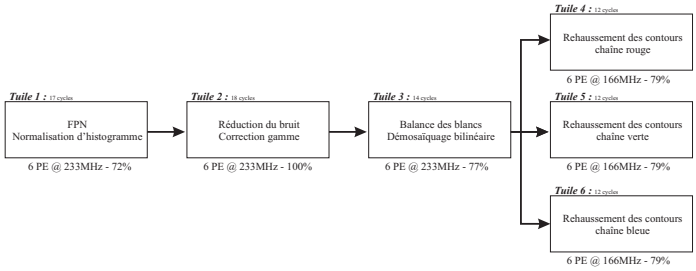


FIG. 2 – Proposition d'arrangement des algorithmes sur les tuiles de calcul.

- Le démosaïquage bilinéaire, porté sur un processeur de l'architecture eISP nécessite 10 cycles et l'accès à un voisinage 3×3.
- Le rehaussement des contours des trois chaînes de couleurs issues du démosaïquage est réalisé par une convolution 3×3 et coûte 12 cycles processeurs.

5.2 Adéquation algorithme-architecture

L'ensemble des algorithmes précédents ne peuvent être exécutés sur la même tuile de calcul. Notamment parce que des dépendances de données rendent complexe une telle implémentation au niveau de l'accès aux voisinages. C'est pourquoi l'organisation de ces algorithmes sur différentes tuiles permet de rendre possible leur intégration. Il est donc préférable d'organiser ces algorithmes tel que présenté en Figure 2, soit 6 tuiles de 6 processeurs chacune. Une telle architecture développe 16,8 GOPS – 2,8 GOPS par tuile – à 233 MHz. Avec une consommation de 40 mW par tuile, l'efficacité de cette architecture est de 70 MOPs/mW. La surface totale est de 1,5mm^2 pour une consommation de 240 mW.

6 Conclusion

Dans un contexte embarqué « dur » où les calculs embarqués au sein de dispositifs mobiles sont essentiellement constitués de composants dédiés, cet article a décrit l'architecture de calcul eISP entièrement programmable. Non seulement, la consommation électrique et la surface silicium sont maîtrisées (1,5mm^2, 250 mW), mais la puissance de calcul disponible (16,8 GOPS effectifs) permet de gérer des flux vidéo haute définition et aux standards 720p ou 1080p, bientôt incontournables en téléphonie mobile. Les résultats présentés s'appuient sur une synthèse en technologie TSMC 65 nm pour une fréquence pouvant aller jusqu'à 400 MHz.

De par sa conception, il est possible d'adapter la capacité de calcul de l'architecture en fonction des besoins, en augmentant le nombre de processeurs par tuile, leurs fréquences de fonctionnement ou encore le nombre de tuiles de calcul. Une telle architecture permet d'intégrer les dernières avancées algorithmiques. D'autres améliorations sont à l'étude comme l'intégration de coprocesseurs dédiés. Le spectre des applications couvert par l'architecture eISP peut ainsi s'élargir à d'autres domaines tels que les télécommunications.

Références

[1] J. C. Russ, *The Image Processing Handbook - Fifth Edition*. 270 Madison Avenue - New York, NY 100 : CRC Press, 1997.

[2] K. Goh, Y. Huang, and L. Hui, "Automatic video contrast enhancement," *Consumer Electronics, 2004 IEEE International Symposium on*, pp. 359–364, 1-3, 2004.

[3] M. Motwani, M. Gadiya, R. Motwani, and F. C. Harris, "Survey of image denoising techniques," in *Proceedings of GSP 2004*, Santa Clara, CA, September 2004, pp. 27–30.

[4] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicing : color filter array interpolation," in *IEEE Signal Processing Magazine*, 2005, pp. 44–54.

[5] J. Zhang and S. Kamata, "Adaptive local contrast enhancement for the visualization of high dynamic range images," in *ICPR 2008 : :International Conference on Pattern Recognition*, 2008, pp. 1–4.

[6] M. Thevenin, M. Paindavoine, L. Letellier, and B. Heyrman, "Processor extensions for CMOS sensor imaging in camera phone," in *Photonics Europe 2008 - Photonics in Multimedia*, vol. 7001, Apr 2008.

[7] B. Khailany, J. Williams, T. adn Lin, E. Long, M. Rygh, D. Tovey, and W. Dally, "A Programmable 512 GOPS Stream Processor for Signal, Image, and Video Processing," *IEEE J. Solid-State Circuits*, vol. 43, no. 1, pp. 202–213, Jan 2008.

[8] J. Chen and S.-Y. Chien, "Crisp : Coarse-grained reconfigurable image stream processor for digital still cameras and camcorders," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1223–1236, Sept 2008.

[9] C. A. Pinto, A. Beric, S. P. Singh, and S. Farfade, "Hiveflex-video vsp1 : Video signal processing architecture for video coding and post-processing," *Multimedia, 2006. ISM'06. Eighth IEEE International Symposium on*, pp. 493–500, Dec. 2006.

[10] R. Kleihorst, A. Abbo, A. van der Avoird, M. Op de Beeck, L. Sevat, P. Wielage, R. van Veen, and H. van Herten, "Xetal : a low-power high-performance smart camera processor," *Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, vol. 5, pp. 215–218 vol. 5, 2001.

[11] A. Gentile, S. Vitabile, L. Verdoscia, and F. Sorbello, "Image processing chain for digital still cameras based on the SIMPil architecture," *Parallel Processing, 2005. ICPP 2005 Workshops. International Conference Workshops on*, pp. 215–222, June 2005.