

Seuillage dynamique temps réel dans un système embarqué

T. Graba, D. Faura, S. Viateur, O. Romain, B. Granado et P. Garda
 LISIF, UPMC, BC 252, 4 Place Jussieu, 75252 Paris Cedex 05
 {tarik.graba}{david.faura}{bertrand.granado}@lisif.jussieu.fr

Laboratoire de Instruments et Systèmes d'Ile de France
 UPMC, BC 252, 4 Place Jussieu, 75252 Paris Cedex 05
 {tarik.graba}{david.faura}{bertrand.granado}@lisif.jussieu.fr

Résumé – Dans cet article nous présentons une méthode à faible complexité et une architecture de traitement dédiée pour la binarisation d'image. Ces travaux sont utilisés dans deux capteurs de vision embarqués: "CYCLOPE" pour la vision 3D active et "MERITE" pour la vidéo surveillance multi-capteurs. Cette méthode s'applique à des images unimodale où les objets à extraire se trouvent noyés dans un fond dominant.

Nous présentons l'architecture de traitement et sa mise en oeuvre dans le cadre de la réalisation d'un démonstrateur sur FPGA.

Abstract – In this paper we present a low complexity method and a dedicated processing architecture for image thresholding. This work is part of the design of two embedded vision sensors: "CYCLOPE" for active 3D vision and "MERITE" for multi-sensor video monitoring. This method can be used with unimodal images where the objects to extract are surrounded by a dominant background.

We present the processing architecture and its implementation in an FPGA prototype.

1 Introduction

Un réseau de capteurs sans fil est composé d'une multitude de systèmes électroniques autonomes. Ces systèmes sont capables d'acquérir, traiter et retransmettre l'information utile acquise de leur environnement vers une station de base ou vers un utilisateur final. Le type d'informations traitées peut être très varié (température, pression, son, gaz, images ...) et dépend fortement de l'application choisie pour le réseau [1, 2, 3].

Nous développons deux projets de réseaux de capteurs sans fil, un basé sur un capteur 3D de vision active CYCLOPE [4] et un autre dédiée à la vidéo surveillance multi-capteurs MERITE [5].

L'autonomie de ces balises implique que les algorithmes et les architectures de traitement de l'information répondent à des critères d'exactitude des résultats, tout en s'adaptant aux variations de leur l'environnement, et à des contraintes sur le temps de traitement et la consommation. Typiquement, pour de telles applications sans fil l'électronique de traitement fonctionne autour de 10 MHz [1, 3].

Un des besoins caractéristiques des balises de vision est l'adaptabilité du système aux changements d'intensité lumineuse de l'environnement. Ceci, nous a mené à développer une méthode de seuillage dynamique originale qui est utilisée dans les deux projets cités. Dans le cadre de CYCLOPE, elle est utilisée pour extraire un motif projeté sur la scène dans une image infrarouge alors que dans le cadre de MERITE elle est utilisée pour l'extraction des objets en mouvement entre deux images. Ce module s'insère dans la chaîne de traitement de chaque application. Les ressources qu'il utilise doivent permettre un fonctionnement en temps réel, sur des séquences vidéo, tout en

gardant la consommation aussi basse que possible.

Dans la première partie de cet article, nous verrons que les images des deux applications présentent des similitudes qui permettent le développement d'une méthode commune. Nous verrons ensuite brièvement quelles sont les méthodes de seuillage utilisées et pourquoi nous avons développé notre méthode. Nous finirons par décrire l'architecture électronique développée pour cette méthode et les résultats de sa mise en oeuvre.

2 Contexte

Dans le cadre de Cyclope, un motif IR est projeté sur une scène. La déformation de ce motif permet de retrouver la profondeur [6]. Une caméra multispéctrale (IR+Visible) permet de récupérer la texture dans le visible et l'image du motif seul dans l'IR. La figure 1(a) montre un exemple d'image et la figure 1(b) l'histogramme de cette image.

Dans le cadre de MERITE, un algorithme de compression MMJPEG2000 [7] encode la différence entre une image de référence et l'image courante. Cette différence représente les mouvements entre les deux images et sert à définir les zones d'intérêt pour la compression. L'image 1(c) représente la valeur absolue de la différence entre deux images d'une séquence vidéo ainsi que son histogramme 1(d).

Dans les deux cas présentés, l'histogramme présente un lobe dominant qui représente le fond (ou le bruit à éliminer). L'information à extraire a quand à elle une intensité supérieure à celle du fond et une population beaucoup moins importante étalée dans le reste de l'histogramme. Pour la suite des développements nous considérerons la distribution de l'intensité du fond comme gaussienne. La moyenne et la variance de cette gaussienne varient avec

les conditions de prise de vue et l'application.

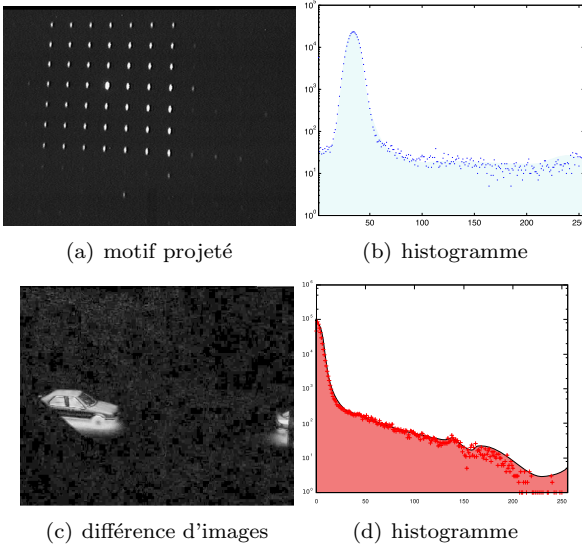


FIG. 1 – Similarité des images pour les deux applications

Nous pouvons approcher cette moyenne et cet écart type par la moyenne $\tilde{\mu}_f$ et l'écart type $\tilde{\sigma}_f$ de l'image. A partir de ces valeurs nous définissons un seuil de binarisation $k = \alpha \tilde{\mu}_f + \tilde{\sigma}_f$. Mais comme ces calculs s'effectuent sur la globalité de l'image, l'erreur commise par cette approximation peut rapidement devenir importante.

Des méthodes utilisant l'histogramme des images ont été étudiées [8]. L'analyse de la distribution de l'intensité des pixels de l'image permet de séparer les classes représentant le fond et les objets comme par la méthode d'Otsu [9] qui maximise la variance inter-classes. L'efficacité de cette dernière est discutable dans notre cas du fait de la différence de population entre les deux classes à extraire. Rosin [10] montre cette inefficacité et propose une méthode basée sur une analyse graphique de l'histogramme pour extraire le premier lobe.

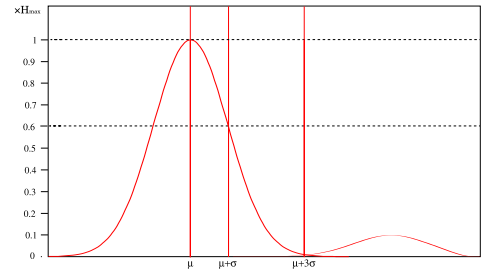
3 Détermination du seuil par analyse d'histogramme

En nous basant sur la méthode de Rosin [10] et sur les propriétés de la fonction gaussienne, nous proposons une estimation de la moyenne et de l'écart type du fond à partir de l'histogramme de l'image. Le premier lobe de cet histogramme est en fonction de l'intensité i :

$$B(i) = H_{max} \exp\left(-\frac{(i - \mu)^2}{2\sigma^2}\right)$$

Nous identifions deux points particuliers, $B(\mu) = H_{max}$ qui est le premier maximum de l'histogramme de l'image et $B(\mu) = H_{max} \exp(-\frac{1}{2}) \simeq 0,6 H_{max}$ par un balayage séquentiel de l'histogramme et obtenons ainsi la moyenne μ et la variance σ de la gaussienne (voir figure 2(a)).

A partir de ces deux points on définit le seuil comme $k = \mu + \alpha \sigma$. Le paramètre α permet d'ajuster le taux d'élimination du fond. Une valeur de 3 permet ainsi d'éliminer 99% de la gaussienne de fond (voir tableau 2(b)).



(a) Modélisation de l'histogramme

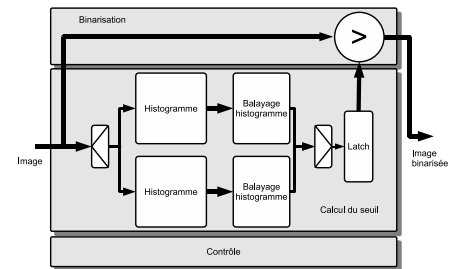
α	surface de la gaussienne
1	68,28%
2	95,44%
3	99,74%

(b) Taux d'élimination du fond

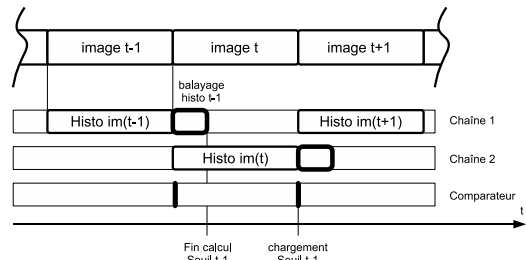
FIG. 2 – Détermination du seuil de binarisation

L'intérêt de l'approche que nous proposons est la réduction de la complexité algorithmique puisque nous effectuons une lecture unique de l'histogramme pour trouver les valeurs particulières qui permettent de définir le seuil. De plus, contrairement aux méthodes d'estimation globale, on approche plus précisément les statistiques du bruit à éliminer.

4 Architecture pour le seuillage dynamique temps réel



(a) Architecture de traitement globale



(b) Séquencement des tâches

FIG. 3 – Mise en œuvre architecturale

Pour réaliser ce module, une architecture spécifique a été conçue. Un système de ping pong (voir figure 3(a)) et l'utilisation de deux blocs de calcul permettent l'acquisition, le calcul de l'histogramme, la détermination du seuil à partir de l'histogramme de l'image précédente et la binarisation de l'image durant son acquisition. Comme la variation de la luminosité entre deux images est faible, on

peut approximer la valeur du seuil de binarisation par la valeur du seuil calculée sur l'image précédente.

Le bloc servant au calcul de l'histogramme (figure 4(a)) est composé principalement d'une mémoire pour sauvegarder le nombre d'occurrences d'un niveau de gris et d'un registre pour garder la valeur de ce niveau. La sortie du registre est utilisée pour adresser la mémoire. Pour une image sur 256 niveaux de gris, ce bus d'adresse est de 8 bits.

Le calcul de l'histogramme nécessite la lecture d'une image complète et est effectué de la manière suivante : Attente d'un pixel (front sur l'horloge pixel p_clk), Lecture de la valeur en mémoire et incrémentation, Sauvegarde de la nouvelle valeur.

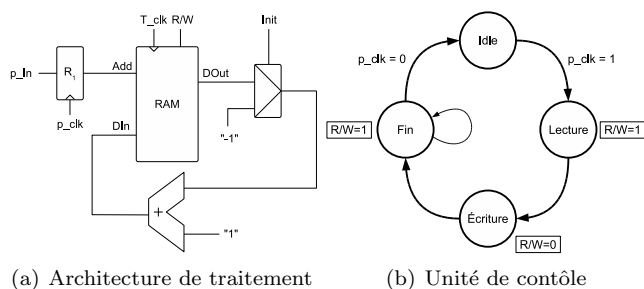


FIG. 4 – Bloc de calcul de l'histogramme

Pour permettre l'exécution de ces étapes durant une période de l'horloge pixel (p_clk), l'horloge de traitement (T_clk) doit être au moins 4 fois plus rapide que l'horloge d'acquisition de l'image.

Le contenu de cette mémoire est ensuite relue pour déterminer les points particuliers de l'histogramme (premier maximum et valeur correspondant à l'écart type) et calculer le seuil.

La figure 3(b) montre le séquencement de ces tâches. Le calcul de l'histogramme nécessite la lecture d'une image entière, les traitements qui suivent sont plus rapides car ils ne nécessitent que la relecture des 256 valeurs de l'histogramme à la fréquence de l'horloge de traitement.

5 Mise en œuvre sur le FPGA

Cette architecture a été mise en œuvre sur une carte de développement XUP2P [11] contenant un Virtex2Pro. Nous avons rajouté une caméra CIF OV6620 (352×288 pixels) munie d'un filtre optique coupe visible pour simuler le fonctionnement d'un capteur multispectrale, un projecteur de motif infrarouge à 830 nm produisant une matrice régulière de points et un écran VGA pour la visualisation des résultats.

L'utilisation des ressources du FPGA pour l'architecture de binarisation est donnée dans le tableau 5(a). Le tableau 5(b) montre quand à lui, les temps et les fréquences d'acquisition et de traitement de chaque pixel pour une cadence vidéo de 25 image/s sur une image CIF.

Ressource	utilisée	disponible	%
Clbs slices	107	13696	0,78%
Latches ou Dffs	192	29060	0,66%
LUT	214	27392	0,78%
Blocs RAM	1	136	0,74%

(a) Utilisation des ressources

Acquisition et calcul de l'histogramme	Détermination du seuil
40 ms (25 im/s)	21,5 us
2,53 MHz	12 MHz

(b) Temps d'exécution

FIG. 5 – résultat de la mise en œuvre

6 Conclusion

Les travaux présentés dans cet article sont destinés à être intégrés dans deux capteurs pour des applications de compression et de vision 3D.

Les résultats obtenus montrent que la méthode de binarisation et l'architecture électronique proposées permettent une mise en œuvre utilisant peu de ressources pour des traitements en temps réel à cadence vidéo dans un système embarqué.

La réalisation d'un démonstrateur fonctionnel utilisant un FPGA nous a permis de valider expérimentalement notre approche. La mise en œuvre dans le circuit finale se fera en adaptant la technologie aux contraintes de consommation auxquelles ne permet pas de répondre le FPGA utilisé.

Références

- [1] I.F. Akyildiz et al. "A survey on wireless multimedia sensor networks", Elsevier Comput. Networks, 2006
- [2] I.F. Akyildiz et al. "Wireless sensor networks : a survey", Elsevier Computer Networks, 2002
- [3] P. Baront et al. "Survey on Wireless Sensor Networks and 802.15.4/ZigBee standards", ACM : C.2 COMPUTER COMMUNICATION NETWORKS, 2006
- [4] T. Graba et al. "Cyclope : an integrated real-time 3D image sensor", XIX Conference on design of circuits and integrated systems 2004
- [5] O. Romain et al. "MERITE A versatile platform for wireless sensor network", IEEE International Symposium on Industrial Electronics, 2004
- [6] T. Graba et al. "Reconstruction 3D temps réel dans un VSIP", Actes du 20ème colloque GRETSI : Traitement du signal et des images
- [7] D.Faura et al. "MMJPEG2000 : A video compression scheme based on JPEG2000", IEEE International Conference on Image Processing, ICIP'06, 2006
- [8] P. K. Sahoo et al. "A survey of thresholding techniques", Computer Vision, Graphics, and Image Processing, Volume 41-2, 1988
- [9] N. Otsu, "A threshold selection method from gray level histogram", IEEE Transaction on System Man Cybernetics, 1979
- [10] P. Rosin, "Unimodal thresholding", Scandanavian Conference on Image Analysis, 1999
- [11] <http://www.xilinx.com/univ/xupv2p.html>

