



**HAL**  
open science

# SYSTEMES DE VISION INTELLIGENTS : ADEQUATION ARCHITECTURE ALGORITHME, CONCEPTION ET IMPLANTATION MATERIELLES

Julien Dubois

► **To cite this version:**

Julien Dubois. SYSTEMES DE VISION INTELLIGENTS : ADEQUATION ARCHITECTURE ALGORITHME, CONCEPTION ET IMPLANTATION MATERIELLES. Systèmes embarqués. Université de Bourgogne, 2013. tel-02121907

**HAL Id: tel-02121907**

**<https://u-bourgogne.hal.science/tel-02121907v1>**

Submitted on 6 May 2019

**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.

SPIM



école doctorale sciences pour l'ingénieur et microtechniques  
UNIVERSITÉ DE BOURGOGNE

Dossier présenté par **Julien DUBOIS**

Docteur en Electronique

Habilitation à Diriger des Recherches

Intitulé

**SYSTEMES DE VISION INTELLIGENTS :**

**ADEQUATION ARCHITECTURE ALGORITHME, CONCEPTION ET IMPLANTATION MATERIELLES**

18 octobre 2013

Composition du jury :

**Pr. Olivier Déforges**, INSA de Rennes – Rapporteur

**Pr. Bertrand Granado**, Université Pierre et Marie Curie (Paris 6)-Rapporteur

**Dr. Gilles Sicard**, Université Joseph Fourier - Rapporteur

**Dr. François Berry**, Université Blaise Pascal de Clermont Ferrand II- Examineur

**Pr. Dominique Ginhac**, Université de Bourgogne - Examineur

**Pr. Johel Mitéran**, Université de Bourgogne – Examineur

**Pr. Michel Paidavoine**, Université de Bourgogne - Examineur



# Remerciements

Que de personnes à remercier pour ces dix ans déjà de recherche au Le2i...

Tout d'abord, je commencerai par Johel Mitéran, Michel Paindavoine et Vincent Vajnovszki pour m'avoir permis de co-encadrer 5 thèses sous leurs directions : expérience au combien enrichissante.

Je souhaiterais remercier Messieurs Déforges, Granado et Sicard pour avoir accepté de rapporter ce manuscrit et Messieurs Berry, Gin hac, Mitéran et Paindavoine pour avoir accepté de l'examiner.

Mes remerciements vont aussi naturellement aux étudiants qu'il m'a été donné d'encadrer au cours de ces années et bien sûr, en particulier, aux doctorants que j'ai co-encadrés avec un très grand plaisir : Romuald Mosqueron, Khalil Khattab, Elfitrin Syahrul, Wajdi Elhamzi et Imen Charfi.

Merci aussi à mes collègues du Le2i avec une pensée particulière à tout le personnel technique Cédric Clerc, Matthieu, Mimi, et toutes les filles du secrétariat, pour leur efficacité, et qui nous permettent de travailler.

Merci bien sûr à mes collègues de l'ESIREM, à toute l'équipe InfoTronique avec qui je travaille tous les jours, merci Christine, Gérard, Patrick, Nader et la liste ... est bien longue...

Merci bien sûr à mes deux complices, Bart et Savério. Merci pour leur bonne humeur, et pour le plaisir de partager avec eux les journées de travail. Et j'ajouterais aussi *caro amico che condivide ogni momento, il buono, il brutto e il molto buono (mi dispiace che non ero in grado di scrivere in Calabrese) ...* mais, sans doute ai-je bu trop de bon café italien !

Évidemment, merci à ma toute ma famille, qui par sa joie de vivre et l'amour qu'elle m'apporte chaque jour, est la source de mon énergie. Merci.

Et puis surtout, j'espère que je n'oublie personne...

# Sommaire du dossier

Sommaire du dossier .....	2
Table des figures.....	5
Introduction générale.....	6
Partie1 : Dossier administratif .....	7
1 Introduction .....	9
2 Curriculum Vitae .....	11
2.1 Fonctions Actuelles .....	11
2.2 Formation Universitaire .....	12
2.3 Parcours Professionnel.....	12
2.4 Bilan des publications scientifiques .....	13
2.5 Renseignements complémentaires.....	13
3 Synthèse des activités pédagogiques & administratives liées à la pédagogie .....	15
3.1 Enseignement et responsabilités de module .....	15
3.2 Synthèse des activités administratives liées à la pédagogie .....	17
4 Rayonnement, animations et responsabilités scientifiques.....	19
4.1 Au niveau local .....	19
4.2 Au niveau national et international.....	19
✓ Organisation et de colloques.....	20
5 Encadrements de thèses et de DEA .....	21
5.1 Thèses soutenues (4) .....	21
5.2 Thèse en cours (1) .....	22
5.3 Encadrements de Master Recherche .....	22
6 Activités contractuelles .....	23
6.1 Contrats publics retenus : .....	23
6.2 Contrats publics soumis ou non retenus (2 soumis, 4 non retenus).....	24
6.3 Contrats privés & transfert de technologies .....	26
7 Collaborations scientifiques.....	29
7.1 Collaborations internationales : .....	29

7.2	Collaborations nationales : .....	29
8	Publications et Communications .....	31
8.1	Récapitulatif .....	31
8.2	Publications dans une revue internationale (17) .....	32
8.3	Livres (1) .....	33
8.4	Revue de vulgarisation (1) .....	33
8.5	Articles soumis dans une revue internationale (1) .....	33
8.6	Conférences internationales (41) .....	33
8.7	Conférences nationales (10).....	36
8.8	Autres présentations internationales (5) .....	36
9	Récapitulatif des travaux de recherche et responsabilités administratives .....	37
9.1	Publications .....	37
9.2	Encadrement de thèse, master recherche et tutorat stage ingénieur .....	37
9.3	Activités contractuelles .....	37
9.4	Responsabilités administratives .....	37
9.5	Diagramme temporel résumé des activités de recherche et administratives au sein du cycle d'ingénieur .....	38
	Partie 2 : Dossier de recherche.....	39
1	Introduction .....	41
2	Développement de nouvelles méthodes et mise en œuvre.....	43
2.1	Contexte scientifique .....	43
2.2	Chaîne de compression d'images incluant des transformées combinatoires.....	43
2.3	Analyse de scènes vidéo pour la détection de chutes .....	46
2.4	Estimation du mouvement pour images multi-exposées : application à l'épandage « intelligent » d'engrais.....	48
2.5	Synthèse personnelle sur la recherche de nouvelles méthodes et leur mise en œuvre .....	50
2.6	Bibliographie relative à ce chapitre.....	51
2.7	Publications jointes .....	51
3	Développement d'accélérateurs matériels dans une démarche « Adéquation Algorithme Architecture » : définition, modélisation, validation, implantation.....	97

3.1	Contexte scientifique .....	97
3.2	Plateforme de co-traitement destinée à la validation d'accélérateurs matériels .....	97
3.3	Implantation d'accélérateurs matériels dédié à la reconnaissance de visages à partir de modélisation en SystemC .....	100
3.4	Implantation d'accélérateurs matériels pour l'estimation du mouvement dédiés la compression vidéo. ....	104
3.5	Prototypage rapide d'accélérateurs matérielles : génération automatique d'interfaces de communication.....	108
3.6	Synthèse personnelle sur la conception d'accélérateurs matériels.....	111
3.7	Bibliographie relative à ce chapitre.....	111
3.8	Publications jointes .....	112
4	Conception et implantation d'architectures innovantes de caméras intelligentes .....	153
4.1	Contexte scientifique .....	153
4.2	Conception d'une caméra rapide intelligente haute résolution .....	153
4.3	Caméra intelligente à unité de traitement hétérogène SW/HW .....	156
4.4	Caméra intelligente pour la haute dynamique .....	158
4.5	Synthèse personnelle sur la conception et la réalisation de caméras intelligentes .....	161
4.6	Bibliographie relative à ce chapitre.....	162
4.7	Publications jointes .....	162
5	Conclusions & perspectives de recherche .....	191

# Table des figures

Figure 1 : Insertion de la transformée de Burrows-Wheeler au sein d'une chaîne de compression de données.....	44
Figure 2 : IRMA - base de données médicale ouverte à la communauté .....	45
Figure 3. Représentation schématique de la méthode de détection de chutes.....	46
Figure 4 : Images traitées et extraites d'une scène de chute disponible dans la base de données ouverte à la communauté .....	47
Figure 5. Estimation des vitesses des granules de fertilisant avec une illustration de cette dernière sur deux images acquises consécutivement dans le temps ( <i>F. Cointault; S. Villette</i> ).....	49
Figure 6 : Une PC-card à base de FPGA – La Wildcard .....	98
Figure 7. Phase d'apprentissage et phase de décision.....	100
Figure 8. Exploitation du parallélisme de données dans la cascade de boosting .....	102
Figure 9. Détection de visages obtenue à partir de l'architecture proposée.....	103
Figure 10. Implantation matérielle/logicielle de l'estimation de mouvement .....	108
Figure 11. Implantation matérielle de l'estimation de mouvement.....	108
Figure 12. Modélisation flot de donnée et définition des interfaces communication .....	109
Figure 13. Evolution du nœud de communication.....	110
Figure 14. Caméra rapide intelligente à haute résolution .....	154
Figure 15. Chaîne de traitement pour l'extraction de centre de gravité .....	155
Figure 16. Caméra intelligente à élément de calcul hétérogène et architecture globale développée .....	156
Figure 17. Différents champs d'applications du projet HiDRaLoN.....	160

# Introduction générale

Les travaux présentés dans le cadre de cette Habilitation à Diriger des Recherches s'inscrivent à la fois dans la problématique dite d'« **Adéquation Algorithme Architecture** » (A<sup>3</sup>) mais aussi « Systèmes embarqués », de la **section 61 du CNU** « Génie informatique, automatique et traitement du signal » . De fait, ces activités de conception de systèmes électroniques s'inscrivent également dans les champs disciplinaires de la **section 63 du CNU** « Electronique, Optronique et Systèmes » tels que « Composants et systèmes électriques », « Instrumentation » ou « Traitement du signal appliqué ».

Ils ont pour objectif commun la mise en œuvre de systèmes principalement matériels et mais aussi logiciels dédiés à la vision artificielle à fortes contraintes en termes de temps d'exécution. Ces systèmes matériels reposent principalement sur des capteurs d'images spécifiquement développés en technologie CMOS et interfacés avec des structures de calculs plus traditionnelles telles que des FPGA ou des DSP.

**Partie1 :**

**Dossier administratif**



## 8 Publications et Communications

Les paragraphes suivants sont consacrés à la présentation exhaustive de toutes les publications et communications réalisées depuis 2000. Le classement a été fait en six catégories :

- Publications dans une revue internationale :
- Brevet
- Chapitres de livre
- Communications dans un congrès international avec actes
- Communications dans un congrès national avec actes
- Autres présentations internationales

A titre d'information, les publications internationales soumises et les publications en cours de rédaction ont été également ajoutées en fin de liste.

### 8.1 Récapitulatif

Publications dans une revue internationale :	<b>17( + 1 soumise)</b>
Brevets en cours de dépôt :	<b>1</b>
Chapitres de livres :	<b>1</b>
Revue de vulgarisation	<b>1</b>
Conférences internationales :	<b>41</b>
Conférences nationales :	<b>10</b>
Workshop (contributions MPEG)	<b>5</b>
<b>Total des publications :</b>	<b>76</b>

## 8.2 Publications dans une revue internationale (17)

- [R1] Imen Charfi, Johel Mitéran, Julien Dubois, Mohamed Atri, Rached Tourki, "**Optimised spatio-temporal descriptors for real-time fall detection : comparison of SVM and Adaboost based classification**", Journal of Electronic Imaging (JEI), 2013, **In Press**
- [R2] Wajdi Elhamzi, Julien Dubois, Johel Mitéran, Mohamed Atri, Barthélémy Heyrman, Dominique Gin hac, "**Efficient smart-camera accelerator: an configurable motion estimator dedicated to video codec**", Journal of System Architecture, Ed. Elsevier, 2013, **In Press**
- [R3] Tomasz TOCZEK, Faouzi HAMDI, Barthélémy HEYRMAN, Julien DUBOIS, Johel MITERAN, Dominique GINHAC, "**Scene-based non-uniformity correction: from algorithm to implementation on a smart camera**", Journal of Systems Architecture, (2013) 1-18, 2013, **In Press**
- [R4] Wajdi Elhamzi, Julien Dubois, Johel Mitéran, Mohamed Atri, "**An Efficient Low-cost FPGA Implementation of a Configurable Motion Estimation for H.264 Video Coding**", Journal of Real-Time Image Processing (JRTIP), August 2012
- [R5] Bilal Hijazi, Jurgen Vangeyte, Frédéric Cointault, Julien Dubois, Sébastien Coudert, Michel Painsavoine, Jan Pieters, "**Two-step cross correlation based algorithm for motion estimation applied to fertilizer granules motion during centrifugal spreading**", Optical Engineering, Vol.50, Number 6, June 2011.
- [R6] Elfitrin Syahrul, Julien Dubois, Vincent Vajnovski, "**Combinatorial Transforms : Application in Lossless Image Compression**", Journal Of Discrete Mathematical Sciences & Cryptography, Editor TARU, Vol. 14, Number 2, pp 129-147, april 2011
- [R7] Richard Thavot, Romuald Mosqueron, Julien Dubois and Marco Mattavelli, "**Generation of Hardware/Software systems based on CAL dataflow description**", Algorithm-Architecture Matching for Signal and Image Processing, Best papers from Design and Architectures for Signal and Image Processing 2007 & 2008 & 2009, Series Lectures Notes in Electrical Engineering, Vol. 73, Springer Verlag, pp 275-292, jan. 2011.
- [R8] Bilal Hijazi, Frédéric Cointault, Julien Dubois, Sébastien Coudert, Jurgen. Vangeyte, Michel Painsavoine, "**Multi-phase cross-correlation method for motion estimation of fertilizer granules during centrifugal spreading**", Precision Agriculture, pp 684-702, vol.11, No.6, Nov. 2010.
- [R9] Khalil Khattab, Julien Dubois and Johel Miteran, "**Cascade Boosting Based Object Detection from High Level Description to Hardware Implementation**", EURASIP Journal of Embedded Systems, Special Issue "Design and Architectures for Signal Image Processing", Hindawi, 12 pages, 2009.
- [R10] Christophe Lucarz, Marco Mattavelli, Julien Dubois, "**A platform for the development and the validation of HW IP components starting from reference software specifications**", EURASIP Journal of Embedded Systems, Special Issue "Design and Architectures for Signal Image Processing", Hindawi, 16 pages, 2008.
- [R11] Romuald Mosqueron, Julien Dubois, Marco Mattavelli, David Mauvilet, "**Smart camera based on an Embedded HW/SW Co-Processor**", EURASIP Journal of Embedded Systems, Special Issue "Design and Architectures for Signal Image Processing", Hindawi, 13 pages, 2008.
- [R12] Fethi SMACH, Johel MITERAN, Mohamed ATRI, Julien DUBOIS, Mohamed ABID, Jean Paul GAUTHIER, "**An FPGA-based accelerator for Fourier Descriptors computing for color object recognition using SVM**", Journal of Real-Time Image Processing (JRTIP), Springer, vol.2, pp. 249-258, 2007.
- [R13] Romuald MOSQUERON, Julien DUBOIS, Michel PAINDAVOINE, "**High-speed smart camera with high resolution**", EURASIP Journal of Embedded Systems, Special Issue "Embedded Vision System", Hindawi, vol. 14, no. 1, Article ID 24163, 15 pages, 2007.
- [R14] Johel MITERAN, Jean-philippe ZIMMER, Michel PAINDAVOINE, Julien DUBOIS, "**Real-time 3D face acquisition using reconfigurable hybrid architecture**", EURASIP Journal on Image and Video Processing, Hindawi, vol. 8, no. 3, Article ID 81387, 6 pages, 2007.
- [R15] Johel MITERAN, Sébastien BOUILLANT, Michel PAINDAVOINE, Fabrice MERIAUDEAU, Julien DUBOIS, "**Real-time flaw detection on complex object: comparison of results using classification with SVM, Boosting and Hyper-rectangle based method**", Journal of Electronic Imaging, SPIE, vol.15, 2006
- [R16] Fan YANG, Michel PAINDAVOINE, Dominique GINHAC, Julien DUBOIS, "**Développement d'un système rapide pour le mosaïquage et la reconnaissance de visages panoramiques**", Traitement du Signal, Vol. 22 (No. 5), pp. 549-562, Septembre 2005.

- [R17] Johel MITERAN, Jiri MATAS, Elbey BOURENNANE, Michel PAINDAVOINE, Julien DUBOIS "**Automatic Hardware Implementation Tool for a Discrete Adaboost-Based Decision Algorithm**", EURASIP Journal on Applied Signal Processing, Hindawi, 2005 (7), pp. 1035-1046, 2005

### 8.3 Livres (1)

- [R18] Khalil Khattab, Philippe Brunet, Julien Dubois, Johel Mitéran, "**Real Time Robust Embedded Face Detection Using High Level Description**", InTech, [New Approaches to Characterization and Recognition of Faces](#), Ed. Peter Corcoran, August 2011.

### 8.4 Revues de vulgarisation (1)

- [R19] Fethi SMACH, Johel MITERAN, Julien DUBOIS, Irène Foucherot, "**Les Descripteurs de Fourier Généralisés dans le cadre de la reconnaissance de formes**", article paru dans la revue de l'Université de Bourgogne en 2008.

### 8.5 Articles soumis dans une revue internationale (1)

- [R20] Mohammad Ali Mirzaei, Julien Dubois, Darin Tsoupikova, "**Run-time Supervised Multi-stage Classifier: Theory and Implementation on FPGA**", *IEEE Transactions in Industrial Electronics*.

### 8.6 Conférences internationales (41)

- [C1] Imen CHARFI, Johel Miteran, Julien DUBOIS, Mohamed ATRI, Rached Tourki "**Spatio-temporal descriptor for SVM and Adaboost based fall detection**", Internataional conference TAIMA, special session on Shape Descriptor and application (IDMA), Hammamet, Tunisia, may 2013
- [C2] Wajdi Elhamzi, Julien Dubois, Johel Mitéran, Mohamed Atri, Rached Tourki, "**High Efficiency Architecture of Half-Pel Motion Estimation for H.264 Video Coding**", Proceedings of International Conference on Information Processing and Wireless Systems (IP-WIS), Djerba, Tunisia, March 2013.
- [C3] Julien Dubois, "**FPGA-based smart camera : industrial applications**", Keynote at the 9<sup>th</sup> Portuguese Meeting on Reconfigurable Systems, Institute of Systems and Robotics, University of Coimbra, Portugal, Feb. 2013.
- [C4] Imen CHARFI, Johel Miteran, Julien DUBOIS, Mohamed ATRI, Rached Tourki, "**Definition And Performance Evaluation Of A Robust SVM Based Fall Detection Solution**" Proceedings of IEEE International Conference on Signal-Image Technology & Internet-Based Systems (SITIS 2012), Italy, November 2012.
- [C5] Wajdi Elhamzi, Julien Dubois, Johel Mitéran, Mohamed Atri, Rached Tourki, "**Hardware implementation of a configurable motion estimator for adjusting the video coding**", Advanced Concepts for Intelligent Vision Systems (ACIVS2012), Brno, Czech Republic, Sept. 2012
- [C6] Imen Charfi, Wajdi Elhamzi, Julien Dubois, Mohamed Atri, Johel Mitéran, "**Video Scene analysis for a configurable hardware accelerator dedicated to Smart Camera**", IEEE/ACM First Workshop on Architecture of Smart Camera (sponsored by ICDSC conference), Clermont Ferrand, France, April 2012.
- [C7] Wajdi Elhamzi, Julien Dubois, Mohamed Atri, Richard Thavot, Jérôme Gorin, Johel Miteran and Rached Tourki, "**An Efficient Hardware Implementation of Diamond Search Motion Estimation Based on CAL Dataflow Language**", The 23rd International Conference on Microelectronics (ICM), Hammamet, Tunisia, December 2011.

- [C8] Elfitrin SYAHRUL, Julien DUBOIS, Asep Juarna, Vincent VAJNOVSKI, "**Lossless Compression Based on Combinatorial Transform : Application to Medical Images**", International Congress on Computer Applications and Computational Science (CACCS), Singapore, Dec, 2010
- [C9] Bilal Hijazi, Frédéric Cointault, Jurgen Vangeyte, Julien Dubois, Michel Paindavoine, Jan Pieters "**A 3-D stereovision system for fertilizer granule characterization**", International Conference of Agricultural & Biosystems Engineering for a Suitable World (AgEng 2010), Clermont-Ferrand, France, Sept. 2010.
- [C10] Elfitrin SYAHRUL, Julien DUBOIS, Vincent VAJNOVSKI, "**Burrows Wheeler Compression Algorithm (BWCA) in Lossless image Compression**", International Conference on Soft Computing, Intelligent System and Information Technology (2<sup>nd</sup> ICSIT) , Bali, Indonesia, July 2010.
- [C11] Bilal HIJAZI , Frédéric COINTAULT , Julien DUBOIS , Jurgen VANGEYTE, Michel PAINDAVOINE, "**New Power-LEDs based illumination system for fertilizer granule motion estimation**", Proceedings of 10<sup>th</sup> International Conference on Precision Agriculture (ICPA2010), Denver, USA, July 2010.
- [C12] Bilal HIJAZI, Jeroen BAERT, Frédéric COINTAULT, Julien DUBOIS, Michel PAINDAVOINE, Jan PIETERS, Jurgen VANGEYTE, "**A device for extracting 3D information of fertilizer trajectories**", Proceedings of 17<sup>th</sup> Congress of the International Commission of Agricultural Engineering (CIGR2010), Québec, Canada, July, 2010.
- [C13] Julien DUBOIS, Richard THAVOT, Romuald MOSQUERON, Johel MITERAN, Christophe LUCARZ, "**Motion Estimation Accelerator with User Search Strategy in an RVC Context**", Proceedings of IEEE International Conference On Image Processing 2009 (ICIP09), Cairo, Egypt, Nov. 2009.
- [C14] Richard THAVOT, Romuald MOSQUERON, Julien DUBOIS, Marco MATTAVELLI, "**Hardware synthesis of complex standard interfaces using cal dataflow descriptions**", Workshop on Design and Architectures for Signal and Image Processing (DASIP09), Sophia Antipolis, France, September 2009.
- [C15] Bilal HIJAZI , Frédéric COINTAULT , Julien DUBOIS , Sylvain VILLETTE , Jurgen VANGEYTE, Fan YANG , Michel PAINDAVOINE, "**New high speed image acquisition system and image processing techniques for fertilizer granule trajectory determination**", Joint International Agricultural Conference 2009 (JIAC09), Wageningen (Netherlands), July 2009.
- [C16] Richard THAVOT, Romuald MOSQUERON, Mohammad ALISAFEE, Christophe. LUCARZ, Marco MATTAVELLI, Julien DUBOIS, "**Dataflow design of a co-processor architecture for image processing**", Workshop on Design and Architectures for Signal and Image Processing (DASIP 2008), Bruxelles (Belgium), November 2008.
- [C17] Elfitrin SYAHRUL, Julien DUBOIS, Vincent VAJNOVSKI, Taoufik SAIDANI, Mohamed ATRI, "**Lossless Image Compression Using Burrows Wheeler Transform (Methods and Techniques)**", Proceedings of IEEE International Conference on Signal-Image Technology & Internet-Based Systems (SITIS'08), Bali (Indonesia), December 2008.
- [C18] Christophe Lucarz, Julien Dubois, Marco Mattavelli "**A Co-design platform for algorithm/architecture design exploration**", Proceedings of IEEE International Conference Multimedia and Expo (ICME08), Hannover (Germany), June 2008
- [C19] Frédéric Cointault , Jürgen Vangeyte, Julien Dubois, Jeroen Baert, Cedric Clerc "**High-speed Imaging Systems for Granule Fertilizer Motion Characterization**", Proceedings of Agricultural & Biosystems Engineering for a Suitable World (AgEng08), Hersonissos, Crete(Greece), 23-25 June 2008.
- [C20] Khalil Khattab, Julien Dubois, Johel Miteran, "**Hardware Implementation of Boosting Based Object Detection Using a High Level Description**", IEEE International Symposium on Industrial Electronics (ISIE08), Cambridge (UK), June 2008.
- [C21] Julien Dubois, Romuald Mosqueron, Michel Paindavoine, Cedric Clerc, Khalil Khattab, "**High-speed smart camera with embedded feature extactions and profilometry measurements**", Optical and Digital Image Processing conference (SPIE08), Strasbourg, France, April 2008.
- [C22] Romuald Mosqueron, Julien Dubois, Marco MATTAVELLI, "**Smart camera with embedded co-processor: a postal sorting application**", Optical and Digital Image Processing conference (SPIE08), Strasbourg, France, April 2008.
- [C23] Christophe LUCARZ, Marco MATTAVELLI, Julien DUBOIS, "**A HW/SW codesign platform for Algorithm-Architecture mapping**", Workshop on Design and Architectures for Signal and Image Processing, DASIP 2007, Grenoble, France, November 2007.

- [C24] Romuald MOSQUERON, Julien DUBOIS, Marco MATTAVELLI, "**High Performance Embedded Co-Processor Architecture For CMOS Imaging Systems**", Workshop on Design and Architectures for Signal and Image Processing, DASIP 2007, Grenoble, France, November 2007.
- [C25] Michel PAINDAVOINE, Julien DUBOIS, Romuald MOSQUERON, Barthélemy HEYRMAN, Jérôme DUBOIS, Dominique GINHAC, "**High speed camera with embedded image processing**", 6th International Workshop on Embedded System, September 6-7, 2007, Vaasa, Finland, September 2007.
- [C26] Romuald MOSQUERON, Julien DUBOIS, Michel PAINDAVOINE, "**Embedded Image Processing/Compression For High-Speed CMOS Sensor**", Proceedings of EUSIPCO2006, Florence, Italy, 6 September 2006.
- [C27] Michel PAINDAVOINE, Romuald MOSQUERON, Julien DUBOIS, "**High-speed embedded real-time image processing**", 27th International Congress on High-Speed Photography and Photonics, SPIE, Xian, China, September 2006.
- [C28] Khalil KHATTAB, Johel MITERAN, Julien DUBOIS, Jiri MATAS, "**Embedded System Study For Real Time Boosting Based Face Detection**", IECON 06, Paris, France, pp. 3461-3465, 2006.
- [C29] Julien DUBOIS, Marco MATTAVELLI, Lionel PIERREFEU, Johel MITERAN, "**Configurable motion-estimation hardware accelerator module for the MPEG-4 reference hardware description platform**", Proceedings of IEEE International Conference On Image Processing (ICIP05), Genova, 2005, September 2005.
- [C30] Michel PAINDAVOINE, Romuald MOSQUERON, Julien DUBOIS, Cédric CLERC, Jean Claude GRAPIN, "**High-speed camera with internal real-time image processing**", SPIE Optics and Photonics (SPIE05), San Diego, California, USA, August 2005.
- [C31] Julien DUBOIS, Marco MATTAVELLI, "**Embedded Co-Processor Architecture and hardware processing modules for CMOS based Image Acquisition**", Proceedings of Journées Francophones sur Adéquations Algorithmes Architecture (JFAAA2005), Dijon, France, January 2005.
- [C32] Johel MITERAN, Jiri MATAS, Julien DUBOIS, Elbey BOURENNANE, "**BOOSTING : from data to hardware using automatic implementation tool**", EUSIPCO2004, Vienne, Autriche, September 2004.
- [C33] Johel MITERAN, Jiri MATAS, Julien DUBOIS, Elbey BOURENNANE, "**Automatic FPGA based implementation of a classification tree**", SCS'2004 Signaux, Circuits and Systems, Monastir, Tunisia, pp. 189-192, 18 March 2004.
- [C34] Michel PAINDAVOINE, Romuald MOSQUERON, Julien DUBOIS, Cédric CLERC, Jean Claude GRAPIN, Lionel PIERREFEU, "**High-speed camera with internal real-time Image compression**", 26th International Congress on High-Speed Photography and Photonics (HSPP04), Alexandria, Virginie, USA, pp. 308-319, sept. 2004.
- [C35] Julien DUBOIS, Marco MATTAVELLI, "**Embedded Co-Processor Architecture for CMOS based Image Acquisition**", Proceedings of IEEE International Conference on Image Processing (ICIP03), Barcelona, 14 September 2003.
- [C36] Lionel LELONG, Guy MOTYL, Julien DUBOIS, Alain AUBERT, Gérard JACQUET, "**The smart sensors contribution for measurement by image processing in fluid mechanics**", 4th Pacific Symposium on Flow Visualization and Image Processing, Chamonix, France, 3 June 2003.
- [C37] Julien DUBOIS, Gérard JACQUET, Guy MOTYL, Viktor FISCHER, Alain AUBERT, "**Design of a binary correlator component and its integration in Round-About architecture for real time motion measurement**", 25th International Congress on High Speed Photography and Photonics (HSPP02), 29 September 2002.
- [C38] Julien DUBOIS, Gérard JACQUET, Viktor FISCHER, Guy MOTYL, Robert FOUQUET, "**Parallel Architecture for Physical Measurements by Image Processing**", 15th Conference on Digital Design on Circuits and Integrated Systems (DCIS'2000), Montpellier, pp. 444-449, 21 November 2000.
- [C39] Viktor FISCHER, Julien DUBOIS, "**Flexible Didactic Card with Embedded Processor**", 15th Conference on Digital Design on Circuits and Integrated Systems (DCIS'2000), pp. 392-396, 21 November 2000.
- [C40] Viktor FISCHER, Julien DUBOIS, "**Realisation of a Parameterised Processor in Reconfigurable Hardware**", Proceedings of IEEE 4th International Conference on Electronics Computers and Informatics 2000, Kosice, Slovakia, pp. 151-161, September 2000.
- [C41] Julien DUBOIS, Gérard JACQUET, Guy MOTYL, Viktor FISCHER, "**System for real time motion measurement**", Third European DSP Education & Research Conference, Paris, France, September 2000.

## 8.7 Conférences nationales (10)

- [C42] Wajdi Elhamzi, Imen Charfi, Julien Dubois, Johel Mitéran, Mohamed Atri, Rached Tourki, **"Un accélérateur matériel flexible pour l'estimation du mouvement dans le cadre de la vidéo assistance avec détection d'événement"**, 6<sup>ème</sup> colloque national du GDR SOC-SIP (System-on-Chip & System-in-Package), Paris, juin, 2012.
- [C43] Michel PAINDAVOINE, Romuald MOSQUERON, Julien DUBOIS, Jean Claude GRAPIN, Cédric CLERC, **"Réalisation et architecture d'une caméra rapide intelligente."**, 9<sup>ème</sup> Rencontres Nationales du Réseau des Electroniciens (CNRS), Dijon, 2007
- [C44] Julien DUBOIS, **"PowerPC emabrqués et accélérateurs matériels pour cibles de type FPGA"**, 9<sup>ème</sup> Rencontres Nationales du Réseau des Electroniciens, Dijon, 2007.
- [C45] Michel PAINDAVOINE, Romuald MOSQUERON, Julien DUBOIS, Cédric CLERC, Jean Claude GRAPIN, **"Analyse en temps réel de la locomotion du petit animal par caméra rapide"**, IMVIE-3, June 2006.
- [C46] Romuald MOSQUERON, Michel PAINDAVOINE, Julien DUBOIS, Cédric CLERC, Jean Claude GRAPIN, Johel MITERAN, **"Caméra rapide avec compression embarquée temps réel"**, RETINE ELECTRONIQUE, ASIC-FPGA ET DSP POUR LA VISION ET LE TRAITEMENT D'IMAGES EN TEMPS REEL (READ05), Evry, France, 2005.
- [C47] Romuald MOSQUERON, Julien DUBOIS, Michel PAINDAVOINE, **"Compression embarquée temps réel pour caméras rapides"**, Manifestation des Jeunes Chercheurs en STIC-(MAJECSTIC 2005), Rennes, France, 2005.
- [C48] Julien DUBOIS, Gérard JACQUET, Frédéric CELLE, Viktor FISCHER, **"Round-About une architecture de traitement d'images pour la mesure en temps réel de paramètres physiques"**, 18<sup>ème</sup> Colloque GRETSI on Signal and Image Processing, 10 September 2001.
- [C49] Julien DUBOIS, Tardi TJAHJADI, Gérard JACQUET, Sébastien COUDERT, **"Mesure temps réel de mouvement avec une précision sub-pixel en PIV à partir d'images binaires"**, Actes du Colloque Interdisciplinaire en Instrumentation (C2I'2001), Hermès, Paris, France, 1, pp. 441-449, 2001.
- [C50] Julien DUBOIS, Gérard JACQUET, Guy MOTYL, **"Architecture Parallèle pour la Vélocimétrie par Images de Particules"**, XVIII<sup>ème</sup> Colloque Photonique et Imagerie Rapide, Arcueil, France, May 2000.
- [C51] Julien DUBOIS, Guy MOTYL, Frédéric CELLE, Viktor FISCHER, **"A non intrusive measurement in real time"**, Journées Signal et Images pour le contrôle dans l'industrie, May 2000.

## 8.8 Autres présentations internationales (5)

- [C52] Julien DUBOIS, Barthélémy HEYRMAN, Marco MATTAVELLI, Johel MITERAN, **"Wildcard Platform Vs ML310 Platform in a motion estimation context"** ISO/IEC JTC1/SC29/WG11 M13136 Contribution to AHG on MPEG-4 Part 9, San Jose, USA, April 2007.
- [C53] Julien DUBOIS, Barthélémy HEYRMAN, Marco MATTAVELLI, Johel MITERAN, Romuald MOSQUERON, **"Motion estimation Hardware Accelerator with Variable Search Window Size"**, ISO/IEC JTC1/SC29/WG11 M13136 Contribution to AHG on MPEG-4 Part 9, Montreux, Switzerland, April 2006.
- [C54] Andrew KINANE, Marco MATTAVELLI, Julien DUBOIS, Christophe LUCARZ, **"Specification of directory structure for SW, HDL and documentation of MPEG-4 Part 9 Reference HW Description"**, ISO/IEC JTC1/SC29/WG11 N8059, Contribution to AHG on MPEG-4 Part 9: Reference Hardware,, Montreux, Switzerland, April 2006.
- [C55] Julien DUBOIS, Marco MATTAVELLI, Johel MITERAN, **"Motion estimation Hardware Accelerator for reduced-search strategy"**, ISO/IEC JTC1/SC29/WG11 M12934 Contribution to AHG on MPEG-4 Part 9: Reference Hardware, Bangkok, Thailand, January 2006.
- [C56] Julien DUBOIS, Lionel PIERREFEU, Marco MATTAVELLI, **"Motion estimation Hardware Accelerator Implementation for Virtex-II Devices"**, ISO/IEC JTC1/SC29/WG11 M10821 Contribution to AHG on MPEG-4 Part 9: Reference Hardware, Redmond, USA, July 2004.

## **9 Récapitulatif des travaux de recherche et responsabilités administratives**

### **9.1 Publications**

- Nombre d'articles publiés dans des revues scientifiques : 17
- Nombre d'articles soumis : 1
- Brevet : 1
- Nombre de chapitres de livres: 1
- Nombre d'articles publiés dans des conférences internationales : 41
- Nombre d'articles publiés dans des conférences nationales : 10

### **9.2 Encadrement de thèse, master recherche et tutorat stage ingénieur**

- Thèses soutenues : 4
- Thèses à soutenir : 1
- Stages de recherche de DEA : 1
- Tutorat de stage ingénieur : 26

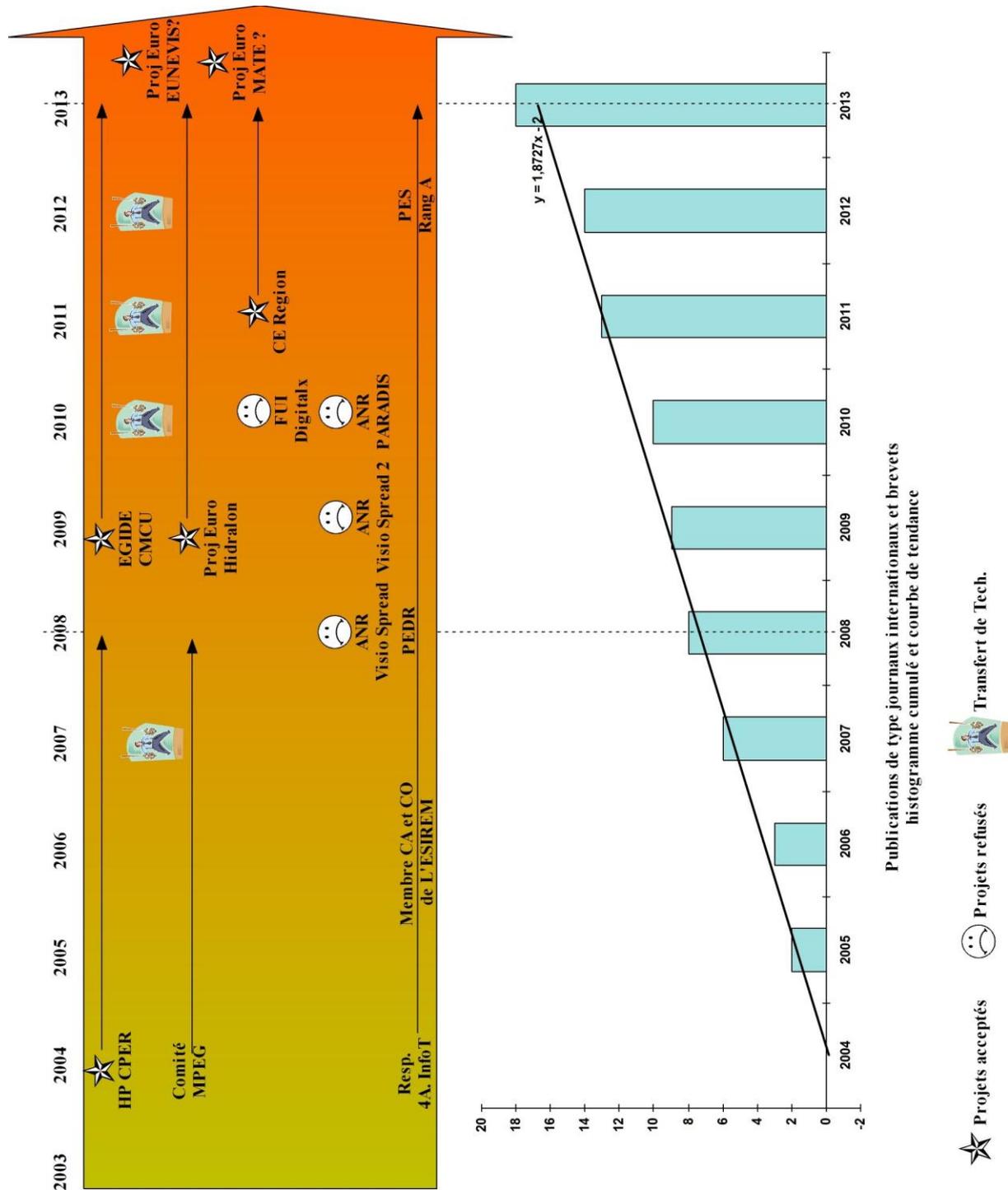
### **9.3 Activités contractuelles**

- Contrats de type public : 4
- Contrats de type privé : 4
- Montant cumulé : 383 k€

### **9.4 Responsabilités administratives**

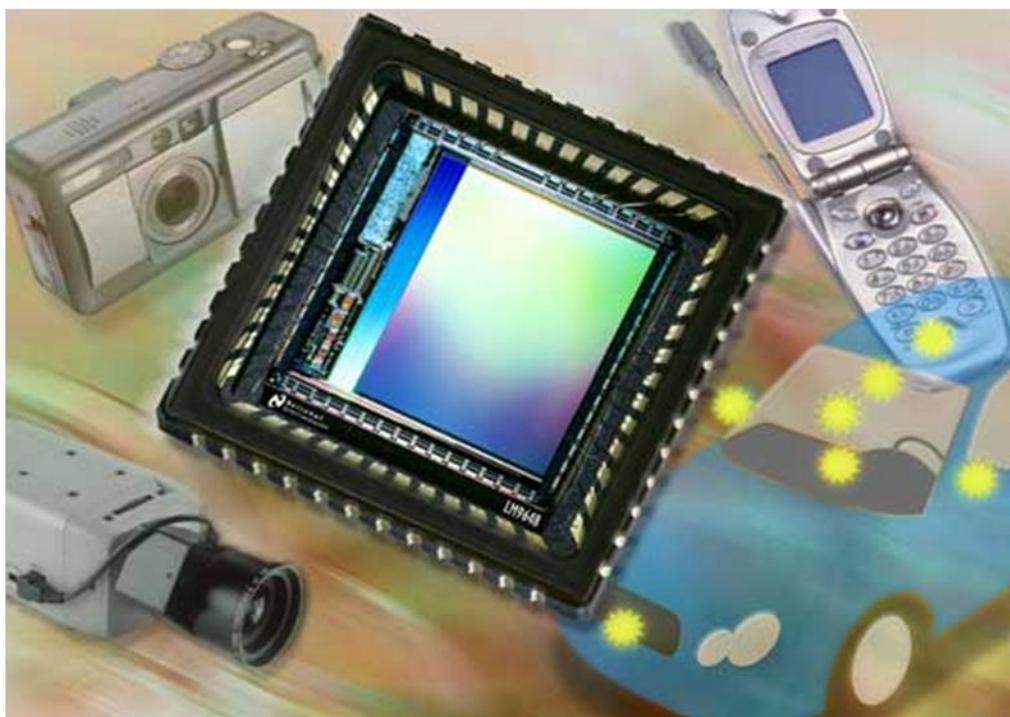
- Responsable pédagogique de la 4<sup>ème</sup> année du Département InfoTronique de l'ESIREM
- Membre élu du conseil d'administration de l'ESIREM
- Membre nommé du conseil d'administration de l'ESIREM
- Membre externe de la Commission de spécialistes de l'Université Blaise Pascale (Clermont Ferrand)

### 9.5 Diagramme temporel résumé des activités de recherche et administratives au sein du cycle d'ingénieur



## Partie 2 :

### Dossier de recherche





# 1 Introduction

Outre le besoin pédagogique lié à l'ouverture du département InfoTronique au sein de l'école d'ingénieur ESIREM, d'un point de vue recherche, mon recrutement s'inscrivait pleinement dans la problématique de recherche menée par l'équipe « Architecture » de ce laboratoire. Il s'agissait ainsi à la fois de renforcer la thématique « **Adéquation Algorithme Architecture** » ( $A^3$ ) du Le2i et en raison de mes expériences antérieures j'ai été rapidement amené à contribuer à des travaux de recherche dans le domaine de « **vision intelligente à haute cadence** ».

Mon parcours professionnel au sein du Le2i est demeuré fortement lié à la thématique  $A^3$  et à celle de vision intelligente au sens large. Au cours de ces dix dernières années (enfin presque...), j'ai pu, aux contacts de mes collègues et grâce à leurs apports, développer une expérience autour des systèmes de vision intelligents et porter cet axe de recherche labellisé « smart camera » au sein du laboratoire le2i dès 2010. Je vais ainsi présenter dans cette partie recherche trois activités qui ont, à mon sens, tout leur place dans la conception et l'implantation de ces systèmes de vision intelligents et présentent des interactions fortes entre elles telle que le schématise la Figure 1 :

- ✓ Développement de nouvelles méthodes pour/ou en vue d'une mise en œuvre dans un système de vision intelligent,
- ✓ Développement d'accélérateurs matériels dans une démarche  $A^3$ ,
- ✓ Conception et l'implantation d'architectures innovantes de caméras intelligentes.

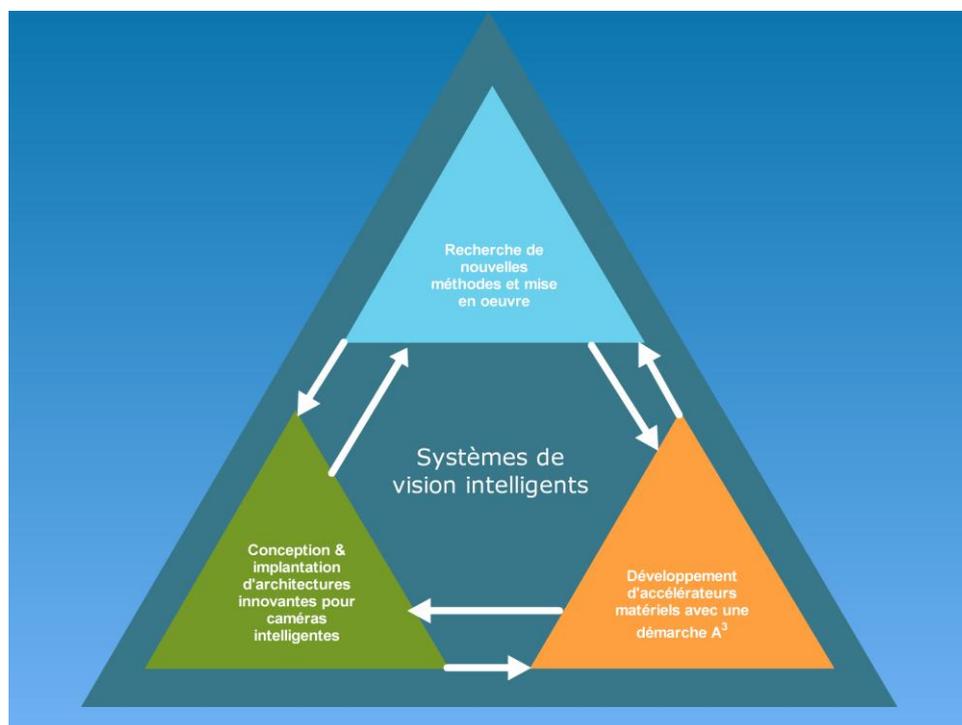


Figure 1. Activités réalisées ayant pour objectif la conception de systèmes de vision intelligents

Pour chacune de ces activités, les contributions pour l'axe de recherche sont ainsi exposées, le bilan en termes de valorisation scientifique est proposé, suivi d'une sélection des publications, jugées significatives.

## 2 Développement de nouvelles méthodes et mise en oeuvre

### 2.1 Contexte scientifique

La conception de systèmes embarqués temps réel, nous amène, en suivant une démarche Adéquation Algorithme Architecture ( $A^3$ ), à la définition de l'architecture des systèmes de visions ou des accélérateurs matériels conjointement à l'étude et à la modification des techniques d'imagerie (de traitement du signal) sélectionnées. Nous pouvons être aussi amenés à travailler de manière plus amont en développant les algorithmes eux-mêmes et en évaluant leurs performances et robustesse indépendamment de la plateforme matérielle support.

Les travaux qui sont ainsi présentés dans ce chapitre s'inscrivent complètement dans ce contexte. Ces études se focalisent sur les thèmes algorithmiques principaux de ma recherche, à savoir la compression, l'analyse d'images et l'estimation du mouvement. Ces traitements sont fondamentaux puisque présents dans un grand nombre de systèmes de vision actuels. Ces traitements seront la base des implantations matérielles proposées dans les chapitres suivants. Trois études sont présentées dans ce chapitre. Dans le cas des deux dernières études, nous avons mis en oeuvre une chaîne complète d'acquisition et de traitement, en levant les verrous technologiques liés à l'application considérée. Les trois études sont les suivantes :

- ✓ La première étude se focalise sur une approche originale dédiée à la compression de données. Cette recherche algorithmique pure s'accompagne d'une validation sur plateforme PC standard.
- ✓ La seconde étude propose le développement d'une méthode d'analyse de scènes vidéo permettant la détection de chute. Cette validation a été réalisée dans un premier temps sur une plateforme PC standard puis nous avons réalisé le portage de cette application sur un processeur embarqué dual corps ARM cortex -A9 MPCore (avec OS embarqué – distribution Ubuntu).
- ✓ La troisième étude consiste à développer une technique d'estimation du mouvement d'images multi—exposées dans le domaine applicatif de l'agro—équipement. Le challenge réside aussi dans la mise en oeuvre d'un système complet de mesure pour l'obtention d'un épandeur d'engrais « intelligent ».

### 2.2 Chaîne de compression d'images incluant des transformées combinatoires

Les méthodes classiques de compression d'image sont communément basées sur des transformées telles que la transformée en Cosinus Discret (DCT) ou encore la transformée discrète en ondelettes. Nous présentons dans ce document une méthode originale basée sur une transformée combinatoire,

celle de Burrows-Wheeler (BWT) [1]. Cette transformée est à la base d'un réagencement des données du fichier servant d'entrée au codeur à proprement parler. Ainsi après utilisation de cette méthode sur l'image originale, les probabilités pour que des caractères identiques initialement éloignés les uns des autres se retrouvent côte à côte sont augmentées. Cette technique est utilisée pour la compression de texte, comme le format BZIP2 qui est actuellement l'un des formats offrant un des meilleurs taux de compression.

La chaîne originale de compression basée sur la transformée de Burrows-Wheeler est composée de 3 étapes, présentées Figure 2. La première étape est la transformée de Burrows-Wheeler elle-même qui réorganise les données de façon à regrouper certains échantillons de valeurs identiques. Burrows et Wheeler conseillent d'utiliser un codage Move-To-Front (MTF) qui va maximiser le nombre de caractères identiques et donc permettre un codage entropique (EC) (principalement Huffman ou un codeur arithmétique). Ces deux codages représentent les deux dernières étapes de la chaîne de compression.

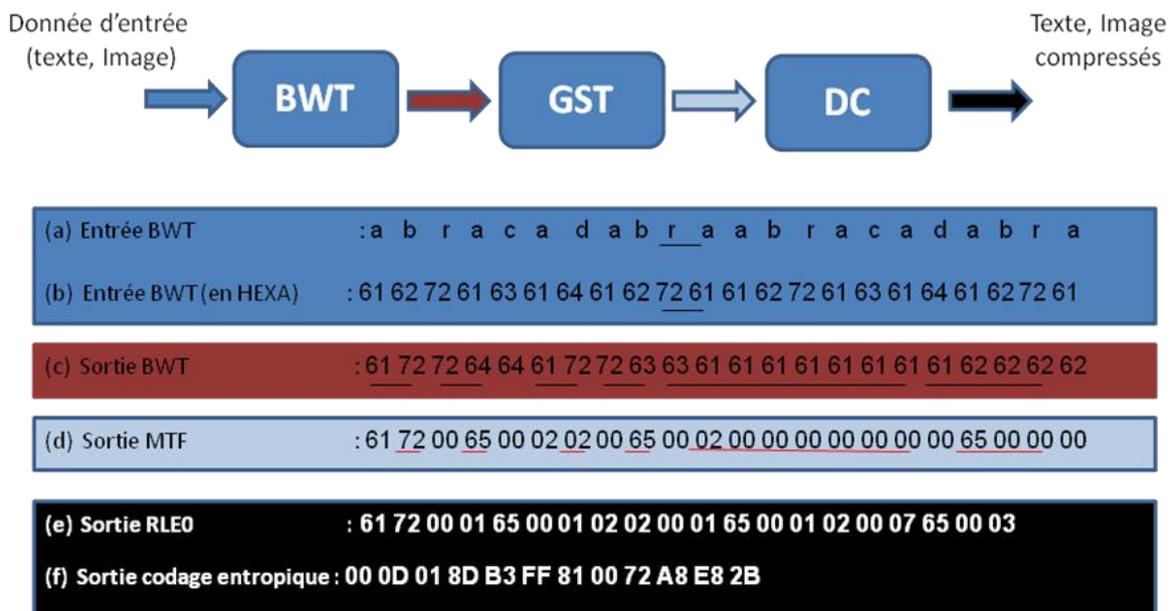


Figure 2 : Insertion de la transformée de Burrows-Wheeler au sein d'une chaîne de compression de données

La méthode présentée est une approche originale intégrant l'utilisation de transformées combinatoires dans une chaîne de compression de données. La méthode est issue des travaux de thèse de Mademoiselle Syahrul Elfitrin. J'ai assuré le co-encadrement de cette thèse dirigée par le Professeur Vincent Vajnovszki. Ces travaux ont bénéficié du support de financier du ministère de la recherche indonésien et de l'ambassade de France en Indonésie.

Nous avons étudié l'état de l'art et réalisé des études empiriques de chaînes de compression basées sur la transformée BWT pour la compression d'images sans perte. Les données empiriques et les analyses approfondies se rapportant à des variantes de MTF et EC. De plus, contrairement à son utilisation pour la compression de texte, et en raison de la nature 2D de l'image, la lecture des données apparaît importante. Ainsi un prétraitement est utilisé lors de la lecture des données et

améliore le taux de compression. Nous avons comparé nos résultats avec les méthodes de compression standard et en particulier JPEG 2000 et JPEG-LS. En moyenne le taux de compression obtenu avec la méthode proposée est supérieur à celui obtenu avec la norme JPEG 2000 ou JPEG-LS. Ainsi sur la base IRMA (base images médicale) [2] présentée Figure 3, les gains moyens sont respectivement de 4,9 et 6,3 % de gain en moyenne sur JPEG-LS et JPEG2K.

Il est important de noter que la valeur moyenne masque l'écart de performances obtenu. Ainsi le taux de compression obtenu avec cette méthode est supérieur à celui obtenu avec JPEG-LS pour seulement 40 % des images de la base. Sur ces images, le gain est alors en moyenne de 20% par rapport à JPEG-LS. Par conséquent, le taux de compression est inférieur pour 60 % des images de la base. Dans ce cas la perte enregistrée au niveau du taux de compression est égale à 3,33 %. Des résultats quasi identiques ont été observés avec le standard JPEG2000.

Aucun critère discriminant entre ces deux classes d'images n'ayant à ce jour été trouvé, nous avons proposé la mise en oeuvre d'un double codeur JPEG-LS et la méthode basée la BWT. Le flot de donnée doit alors être compressé en parallèle par les deux codeurs. Le flot compressé de plus petite taille est conservé et marqué en fonction de la méthode de compression utilisée. Dans ce cas, le gain serait alors en moyenne de 8 % par rapport à JPEG-LS.



Figure 3 : IRMA - base de données médicales ouverte à la communauté scientifique

D'un point de vue pratique, une version presque sans perte d'images médicales a aussi été proposée. Une segmentation de la zone critique est alors réalisée. Cette zone est compressée sans perte tandis que le reste de l'image est très fortement compressée avec perte (voir éliminée pour certaines configurations).

Des perspectives existent pour ces travaux, notamment sur deux axes :

- ✓ Le développement d'un accélérateur matériel réalisant de traitement de la transformée de Burrows-Wheeler en temps réel (25 images/s), pour proposer in-fine la réalisation d'un système d'acquisition permettant l'obtention d'un flux compressé sans perte. La conception d'un tel accélérateur demeure un challenge en raison du très nombre de permutations à réaliser séquentiellement.
- ✓ L'utilisation de la méthode par plan de bits et non pas sur la dynamique total de l'échantillon afin de réduire la statistique des échantillons à regrouper.

## 2.3 Analyse de scènes vidéo pour la détection de chutes

Les applications liées à la vidéo assistance nécessitent une bande passante élevée notamment lorsque la sortie est composée d'informations à la fois bas niveau (pixel, image) et haut niveau (informations extraites automatiquement, comme le nombre de personnes présentes dans la scène, la détection d'événements inhabituels de type chute, etc.). Dans ce contexte, nous étudions un système de détection de chutes adapté à une implantation matérielle permettant de réduire la bande passante finale nécessaire (compression vidéo adaptative). En effet, l'analyse de l'évolution de la scène permet d'adapter la compression dynamiquement en fonction des besoins pour lesquels différents taux de compression peuvent être affectés.

La méthode originale proposée de détection automatique de chutes, nous a permis de définir des descripteurs de chute et d'étudier leurs performances. Ces descripteurs sont calculés à partir de la géométrie de la forme en mouvement dans la scène, et de la transformée de Fourier ou des ondelettes combinés avec des classifieurs de type SVM et Boosting. La représentation schématique de la méthode de détection est représentée Figure 4. L'approche proposée est issue des travaux de thèse de Mademoiselle Charfi Imen. J'ai assuré le co-encadrement de cette thèse réalisé en cotutelle avec l'Université de Monastir. Cette thèse a été dirigée par le Professeur Johel Mitéran. Elle s'appuie sur des études préalables menées conjointement avec le Professeur Mitéran [R12][R15][R17]. Le domaine applicatif est la vidéo assistance. Le projet a été supporté financièrement par le projet EGIDE UTIQUE dont j'ai assuré la responsabilité.

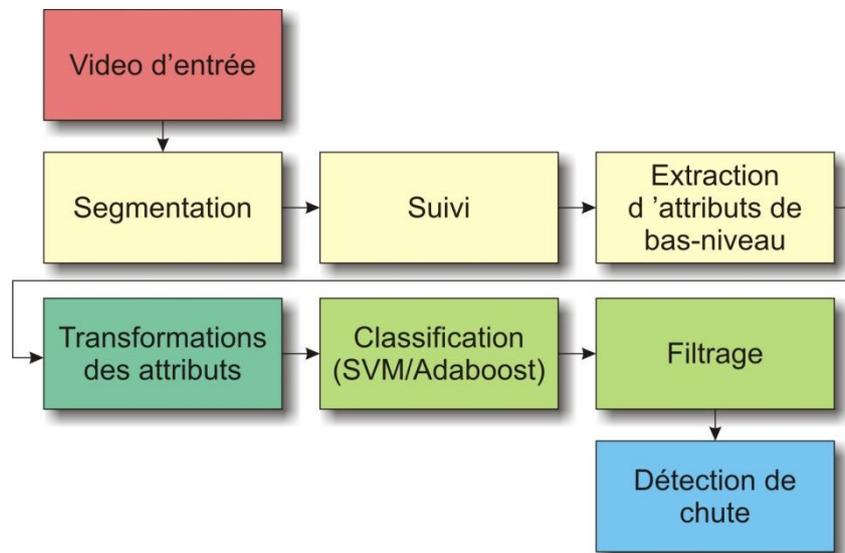


Figure 4. Représentation schématique de la méthode de détection de chutes

De manière plus précise, la détection est basée sur une élimination de l'arrière-plan, un suivi de la personne en mouvement et l'extraction d'attributs de premier niveau (moments, caractéristiques de la boîte englobante, position de l'ellipse englobante de la zone en mouvement, etc.) tels que représentés sur la Figure 5. Au total 14 attributs sont définis puis réduits à 7 par sélection d'attributs de type « Sequential Backward Floating Selection » (SBFS). Une chute étant caractérisée par une variation brutale de certains de ces attributs, nous avons appliqué différentes transformées (dérivées première et seconde, FFT et transformée en ondelette). Les données brutes et ces transformées,

calculées sur une fenêtre temporelle glissante (dont nous avons étudié la taille optimum) forment des attributs de second niveau constituant le vecteur d'entrée du classifieur, qui donne une décision pour cette plage d'images. Nous appliquons de plus à la sortie du classifieur une règle temporelle prenant en compte 18 décisions successives, ce qui correspond à la durée moyenne d'une chute.

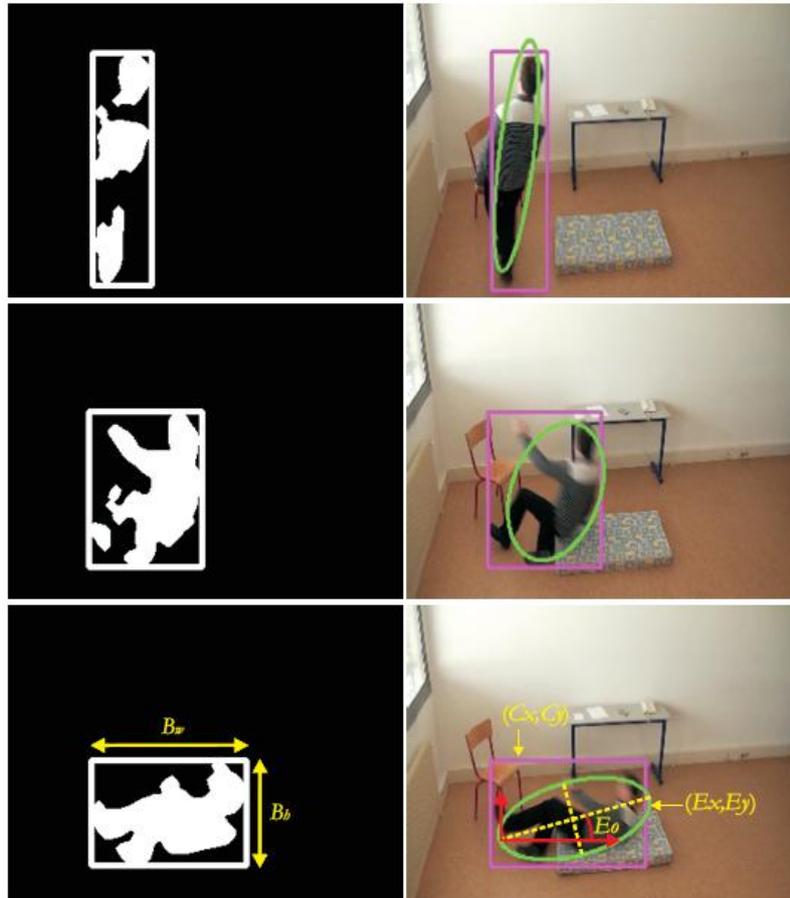


Figure 5 : Images traitées et extraites d'une scène de chute disponible dans la base de données ouverte à la communauté

En outre, nous avons constitué une base de 200 vidéos annotées comprenant des chutes et des situations diverses afin de réaliser l'apprentissage puis l'évaluation de la méthode, ainsi que la recherche de la meilleure combinaison possible d'attributs. Cette base de données est ouverte à la communauté scientifique et accessible sur le site <http://le2i.cnrs.fr/Fall-detection-Dataset>. Afin d'évaluer la robustesse de la méthode au changement de point de vue et d'environnement, les vidéos ont été enregistrées dans 4 locaux différents. Différents protocoles d'évaluation ont été proposés et testés afin d'obtenir une solution réaliste en vue du déploiement au sein d'une habitation personnelle ou encore au sein de structure médicalisée (hôpital, maison de retraite).

Les résultats expérimentaux préliminaires nous permettent de conclure que dans le cas des SVM, la meilleure combinaison d'attributs est constituée des données brutes et de leur dérivée première. Dans ce cas, le meilleur résultat obtenu est une seule fausse alarme constatée sur l'ensemble des vidéos, alors que toutes les chutes sont détectées. Lorsque le test est réalisé sur des vidéos d'un

environnement qui n'a jamais été appris, le nombre de fausses alarmes passe à 4 (rappel : 0.89, précision : 0.94).

Dans le cas du Boosting, les performances finales sont très proches des SVM, (rappel : 0.86, précision : 0.96) et beaucoup mieux adaptées à une implantation matérielle de type FPGA, puisque la fonction de décision est constituée de simples comparateurs qui peuvent être largement parallélisés [R17].

Les contributions de ce travail sont : la constitution d'une base de 200 vidéos annotées, enregistrées en plusieurs endroits ; une analyse de la trajectoire et de la déformation de la forme en mouvement réalisées directement par le classifieur, laissant ce dernier prendre en compte l'aspect temporel de la chute, contrairement à [3], [4] ; une comparaison en faveur du Boosting qui permet, à performances de classification comparables, d'envisager une implantation matérielle complète sur FPGA à faible coût, compatibles avec nos objectifs initiaux d'intégration de l'ensemble dans une caméra intelligente. D'autre part, nous avons réalisé l'intégration de la détection de chute sur une cible matérielle de Zynq. L'idée essentielle de ce développement réside dans le prototypage rapide de la méthode sur une plateforme embarquée autonome. Le composant Zynq possède d'une part une architecture hétérogène qui allie un processeur double-cœur et de la logique configurable et d'autre part propose un grand panel d'interfaces de communication. Ainsi, le déploiement d'un OS de type Ubuntu sur cette cible a permis la réutilisation du code original en C/C++ et des bibliothèques associées (principalement OpenCV) ainsi qu'une gestion simplifiée des interfaces de communications. Le système proposé, assimilable à une caméra intelligente, est composé d'une webcam et d'une carte d'évaluation intégrant un composant Zynq. Les communications sont assurées par des liaisons USB et Ethernet respectivement pour les liaisons :

- ✓ entre capteur vidéo et carte d'évaluation
- ✓ entre capteur d'évaluation et l'environnement extérieur.

L'adaptation algorithmique mais aussi la transformation du code dans une version « multi-thread » a permis d'obtenir une intégration performante (environ 10 images/s) de la méthode proposée sur ce système embarqué et autonome. Des perspectives intéressantes demeurent naturellement dans le passage de ces parties de l'algorithme en implantation matérielle dans l'objectif de traiter le flux vidéo en temps réel à 25 images/s.

## **2.4 Estimation du mouvement pour images multi-exposées : application à l'épandage « intelligent » d'engrais**

La fertilisation excessive des terres agricoles est considérée comme une contribution majeure au déséquilibre environnemental croissant observé dans les 20 dernières années. Un meilleur contrôle de la précision des méthodes d'épandage limiterait les excès de fertilisant inutilement déversés dans l'environnement. L'objectif de ces travaux est de se concentrer sur le contrôle de l'épandage centrifuge qui représente le mode de fertilisation le plus communément utilisé en Europe (Figure 6).

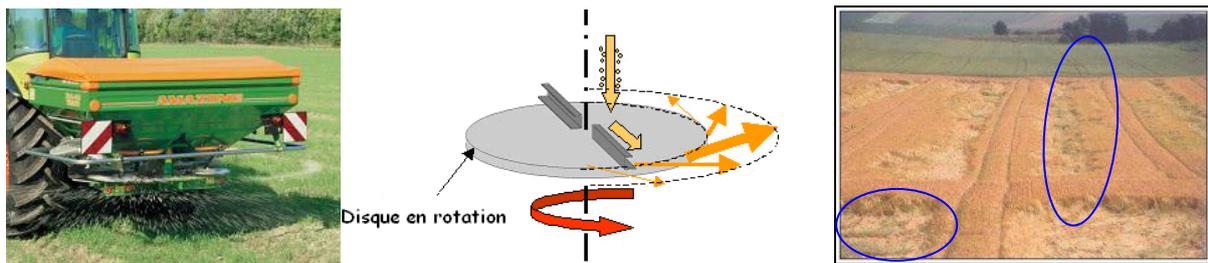


Figure 6.. a- Un système d'épandage centrifuge commercial (Amazone S.A.); b-principe de l'épandeur centrifuge (S. Villette) ; c-images illustrant des problèmes occasionnés par des mauvaises fertilisations (cercles).

L'utilisation de matériel permettant l'acquisition et le traitement d'images a permis de la définition de nouvelles technologies permettant d'observer et de décrire le comportement des granules de fertilisant durant la phase d'éjection au niveau de l'épandeur centrifuge. La répartition au sol du fertilisant peut ainsi être prédite en utilisant un modèle de vol balistique à condition de pouvoir déterminer, lors de la phase d'éjection, les vitesses et les directions des granules. Nos travaux ont ainsi consisté à proposer l'amélioration d'un système de mesure de ces informations, basé sur de l'imagerie rapide mais aussi sur le système d'illumination stroboscopique spécifique à cette application. Ces travaux ont été portés et initiés par F. Cointault d'AgroSup Dijon avec le soutien du Professeur M. Paindavoine. Ces travaux ont majoritairement été développés lors de la thèse de Monsieur Hijazi Bilal (soutenue en 2012). Cette thèse a été co-encadrée par Monsieur Frédéric Cointault et dirigé par le Professeur Michel Paindavoine. J'ai collaboré sur ces études d'une part sur la définition de la méthode du mouvement en 2D et d'autre part sur la mise en place du système expérimental. Ces méthodes ont été empruntées aux techniques de vélocimétrie par images de particules utilisées en mécanique des fluides. Ces techniques m'étant particulièrement familières puisqu'aux cœurs de mes travaux de thèse. Ces méthodes ont dû être adaptées aux contraintes des images multi-exposées avec large déplacement inter-expositions.

Les images multi-exposées sont ainsi constituées d'une série de jet d'engrais ayant une forme se rapprochant d'arcs tel qu'illustré Figure 7. Chacun d'entre eux doit être identifié et extrait de manière automatique. Une technique basée principalement sur de l'inter-corrélation et réalisée sur de petites zones (typiquement 16x16 ou 32x32) permet de réaliser l'identification et le suivi des particules entre deux jets consécutifs. Cette technique d'estimation du mouvement directement inspirée de la méthode de vélocimétrie par images de particules (P.I.V) a été comparée avec d'autres méthodes en particulier une méthode basée sur les champs de Markov (MRF), mise en œuvre dans les travaux précédemment menés durant la thèse de F. Cointault [5].

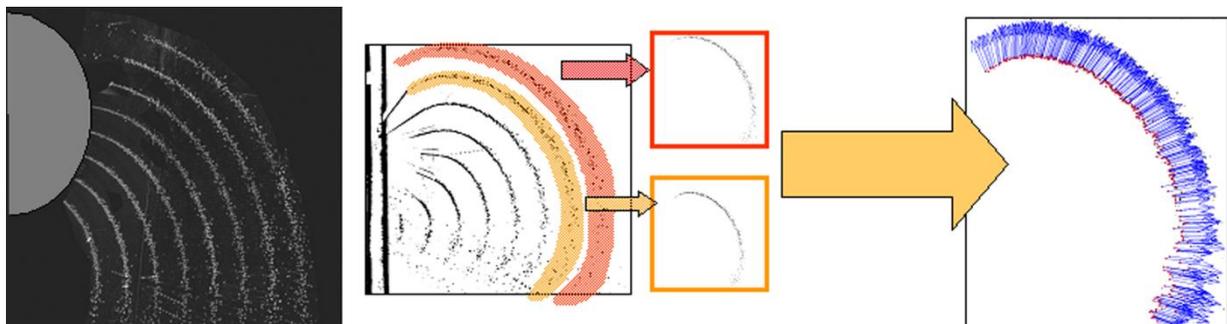


Figure 7. Estimation des vitesses des granules de fertilisant avec une illustration de cette dernière sur deux images acquises consécutivement dans le temps (F. Cointault; S. Villette).

Les tests ont montré que la méthode basée sur l'inter-corrélation est nettement plus précise que les champs de Markov et garantissait l'obtention d'une précision sub-pixel. Sachant qu'à partir de la résolution spatiale choisie (640x480) et de la distance caméra – épandeur (figée pour des raisons de sécurité du matériel) une erreur d'un pixel entraîne une erreur de prédiction au sol comprise entre 200 et 500 mm, la méthode proposée permet d'obtenir une précision sur les composantes horizontales et verticale de la vitesse comprise entre 0.1 et 0.4 pixels contre une erreur située entre 3 et 9 pixels avec la méthode MRF. La précision sub-pixel a été démontrée par l'utilisation d'images de synthèse modélisant des déplacements contrôlés au niveau des grains de fertilisant. Cette précision sub-pixel permet ainsi d'envisager des résolutions spatiales réduites, avec des caméras standard. La réduction de prix de revient du système résultant est alors nettement inférieure à l'utilisation de caméras rapides possédant des résolutions spatiales supérieures au méga-pixel. Ces améliorations rendent alors envisageable le déploiement de ce type de système, probablement sous la forme de caméras intelligentes, sur les épandeurs classiques. Pour autant, la source lumineuse doit permettre un éclairage uniforme sur une large surface (typiquement 1 m<sup>2</sup>) de type stroboscopique synchronisé sur le système d'acquisition. Nous avons ainsi développé ce type d'éclairage à base de LED de puissance en vue du déploiement industriel du système de mesure. Ce système d'éclairage a fait l'objet d'un dépôt de brevet.

La transposition de la plateforme 2D à un système de mesure 3D a été investiguée par l'équipe de recherche d'AgroSup en collaboration avec l'institut d'ILVO (Belgique). La mise en place d'un système stéréoscopique a permis une mesure précise des trois composantes des vecteurs vitesses et améliorer ainsi l'estimation de la distribution de fertilisant au sol.

L'intégration du système 2D sur composant de type FPGA est actuellement à l'étude en collaboration avec l'institut d'ILVO. La caméra intelligente qui sera alors obtenue pourra ainsi être déployée sur les épandeurs d'engrais centrifuges

## **2.5 Synthèse personnelle sur la recherche de nouvelles méthodes et leur mise en oeuvre**

Les méthodes définies ont été développées dans le but de répondre aux contraintes fortes des applications visées. Leur originalité doit être un apport fondamental pour améliorer les performances des systèmes de vision intelligent (bien souvent pour une application donnée). Nous avons toujours développé ces techniques dans l'objectif d'une mise en oeuvre au sein d'un démonstrateur. Ainsi une démarche A<sup>3</sup> a accompagné ces développements.

D'autre part, nous réalisons le constat que les modes d'acquisition et généralement fortement lié aux traitements envisagés. La liaison entre acquisition et traitement justifie l'idée même de systèmes dits « intelligents ». Où se situe la limite entre acquisition et traitement ? Cette question est tout à fait légitime dans le cas d'un capteur intelligent ! Ainsi la détection d'évènements dans la scène vidéo, comme la détection de chutes, peut permettre de contrôler temporellement le débit d'un codeur vidéo (ou encore la qualité de l'encodage).

Le contrôle en dynamique de l'acquisition comme par exemple l'extraction de zones d'intérêt dans les images s'avère fondamental pour les performances du système global. Ainsi, le contrôle de l'acquisition peut permettre de :

- ✓ simplifier le traitement à mettre en jeu,
- ✓ sélectionner les zones utiles dans l'image à un instant donnée et réduisant par conséquent les bandes passantes nécessaires pour les transmissions des données ou encore le temps de calcul.

La répartition du flot de données entrant vers les différentes unités de traitement pour un fonctionnement en parallèle de ces dernières, s'avère tout aussi fondamental pour obtenir de hautes performances de traitement. Nous essayerons de l'illustrer dans la section 4 qui traite de la définition d'architectures innovantes pour les systèmes intelligents de visions.

Finalement, l'utilisation de cibles matérielles telles que le Zynq permet (avec déploiement d'un OS et une distribution adaptée) la migration rapide vers un prototype opérationnel. Bien sûr, une perte de performance peut être alors observable si aucune optimisation de type logicielle (par exemple une version « multi-thread ») ou matérielle (intégration de certaines fonctionnalités en matériel) n'est réalisée. Pour autant, ces solutions ou plus généralement doivent être considérées pour réduire les temps de développement (souvent très important) du système global. Nous présenterons cette approche en section 3.

## 2.6 Bibliographie relative à ce chapitre

[1] M. Burrows and D. J. Wheeler. "A block-sorting lossless data compression algorithm". Technical report, System Research Center (SRC) California, May 10, 1994.

[2] T. M. Lehmann, M. O. Güld, C. Thies, B. Fischer, K. Spitzer, D. Keysers, H. Ney, M. Kohnen, H. Schubert, and B. B. Wein. Content-based image retrieval in medical applications. 2004.

[3] C. Rougier, J. Meunier, A. St-Arnaud, J. Rousseau, "Robust Video Surveillance for Fall Detection Based on Human Shape Deformation", IEEE Transactions on Circuits and Systems for Video Technology, 2011, pp 611-622.

[4] A. Hervieu, P. Bouthemy, J.P. Le Cadre, "Reconnaissance d'évènements vidéos par l'analyse de trajectoires à l'aide de modèles de Markov", Traitement du signal, 2009, pp 187-197.

[5] F. Cointault, Automatisation de la mesure du débit d'engrais et trajectographie des granules d'engrais : application à la fertilisation raisonnée dans le contexte de l'agriculture de précision, thèse, 2001

## 2.7 Publications jointes

Les travaux relatifs à la thématique présentée dans cette section ont permis une valorisation scientifique significative avec notamment la publication de 5 articles dans des journaux scientifiques internationaux, un dépôt d'un brevet et treize présentations dans des conférences internationales.

---

Une sélection de deux articles récents et publiés dans des revues internationales est proposée :

- [R1] Imen Charfi, Johel mitéran, Julien Dubois, Mohamed Atri, Rached Tourki, **"Optimised spatio-temporal descriptors for real-time fall detection : comparison of SVM and Adaboost based classification"**, Journal of Electronic Imaging (JEI), pp.31, 2013, **In Press**
- [R5] Bilal Hijazi, Jurgen Vangeyte, Frédéric Cointault, Julien Dubois, Sébastien Coudert, Michel Paindavoine, Jan Pieters, **"Two-step cross correlation based algorithm for motion estimation applied to fertilizer granules motion during centrifugal spreading"**, Optical Engineering, Vol.50, Number 6, pp.13, June 2011.

### **3 Développement d'accélérateurs matériels dans une démarche « Adéquation Algorithme Architecture » : définition, modélisation, validation, implantation**

#### **3.1 Contexte scientifique**

La définition et l'implantation d'accélérateurs matériels est réalisée dans une démarche d'Adéquation entre Algorithme et Architecture (AAA). Le développement d'architectures innovantes représente le cœur de mes travaux de recherche. J'ai ainsi contribué au développement d'un certain nombre d'architectures innovantes depuis mon intégration au sein du Le2i, parfois en tant qu'initiateur, parfois en tant que participant. De plus, en partant du constat que la conception des systèmes de vision embarqués permet généralement la réutilisation d'un certain nombre de fonctionnalités (aussi bien en termes de traitement que de communication) d'une application à une autre, la démarche « AAA » a été alors accompagnée par le souci d'exploiter cette caractéristique ou plus généralement de réduire les temps de conception et d'implantation de ces développements. Ainsi aux cours des travaux de recherche décrits dans la suite de ce paragraphe, nous avons proposé :

- ✓ la réutilisation de plateformes logicielles/matérielles (SW/HW) et de bibliothèques spécifiques pour diminuer le temps nécessaires à la validation des accélérateurs,
- ✓ l'évaluation de l'utilisation de la modélisation haut niveau et des outils associés de synthèse de haut niveau (HLS) pour le développement d'accélérateurs logiciels et ou matériels.

#### **3.2 Plateforme de co-traitement destinée à la validation d'accélérateurs matériels**

Le projet consistait en la conception d'une plateforme « Recherche et Développement », plateforme unifiée de co-processing pour implantation Software/Hardware de traitements vidéo. Ce projet a été initié au sein du comité MPEG, auquel j'ai participé activement jusqu'en 2009, pour un besoin de validation d'accélérateurs matériels sophistiqués et dédiés à la compression, et dont nous avons pu montrer par ailleurs qu'il était transposable plus généralement aux accélérateurs matériels dédiés au traitement d'images.

De nos jours, la compression et la décompression d'images numériques sont devenus des outils technologiques matures qui commencent à apparaître dans les applications de tous les jours tels que la télévision numérique, le stockage d'images numériques statiques, les web-cams. Cependant, la vidéo numérique est loin d'être utilisée à son potentiel maximum sur terminaux mobiles ou encore sur plate-formes PC standard. La raison principale réside dans le fait que les tâches de compression et de décompression nécessitent l'utilisation de plate-formes spécifiques dans le but de parvenir à une performance temps-réel. En général, les terminaux sont équipés de composants hardware spécifiques dans le but de réaliser les tâches de compression/décompression. C'est le cas de la télévision numérique et des webcams où des composants MPEG-2 sont en charge de l'intégralité du

chargement vidéo. Dans d'autres cas, la résolution vidéo est réduite dans le but de correspondre aux capacités de traitement des terminaux, même en utilisant des taux de transmission vidéo bien en dessous des capacités offertes par la bande passante du réseau ou des capacités de stockage. C'est le cas des terminaux mobiles de troisième génération (UMTS) ou d'un PC portable sans capacités hardware qui lisent des fichiers vidéo au format DVX.

La difficulté d'intégrer les capacités nécessaires en termes de traitement dans les terminaux multimédia est accrue par le fait qu'un terminal ne doit pas seulement répondre à un simple algorithme spécifique de compression/décompression vidéo (vidéo standard) mais aussi à différents standards dépendant du type de matériel dont le terminal dispose ou de la connexion réseau imposée par le codeur spécifique. Si d'un côté les standards vidéo correspondent à des algorithmes très différents (H.261, MPEG-1, MPEG-2, H.263, MPEG-4 vidéo, MPEG/ITU advanced video coding AVC, Movie JPEG2000, etc etc ...) car ils représentent l'état de l'art du codage vidéo au moment de leur définition, pour une gamme de taux de transmission spécifiques ou pour un ensemble spécifique de fonctionnalités (efficacité du codage, modularité, erreur ..., code orienté objet, etc, ...), d'un autre côté, ces standards partagent différents outils de codage utilisés dans des contextes différents ou avec des paramètres différents.

Ainsi l'idée de ce projet était de définir une plate-forme ouverte destinée au co-processing vidéo qui devait supporter des implémentations mixtes hardware/software, flexibles et reconfigurables, de codeur vidéo. **Cette plate-forme avait pour ambition de correspondre dans le domaine de compression vidéo à l'équivalent de la librairie OpenGL dans le domaine graphique sur ordinateur.** En d'autres termes, il s'agissait d'un ensemble de traitements codés à haut niveau tels que les blocs DCT, DWT, compensation de mouvement, estimation du mouvement, VLC, VLD, codage de forme, codage arithmétique, etc., et qui, grâce à la définition de leur API, seront exécutables par les « fonctions » de co-processing de la plate-forme, et ceci d'une manière complètement transparente pour l'application multimédia y faisant appel. Les fonctions ont été codées en VHDL ou Verilog et compatible avec la « PC-card WildCard » (Figure 8) comportant un FPGA de type Xilinx Virtex2 XCV300E.



Figure 8 : Une PC-card à base de FPGA – La Wildcard

L'intervention du Le2i s'est effectuée en partenariat avec ses autres collaborateurs universitaires (Universités de Taïwan, de Calgary, de Dublin et l'École Polytechnique Fédérale de Lausanne) et industriel (Xilinx). Le laboratoire Le2i a eu en charge le développement de la fonction d'estimation du mouvement présente dans la plupart des standards de compression vidéo récents. Le Le2i a été supporté financièrement par le contrat d'étude « Hors Contrats de Plan Etat-Région » (Référence: 04516CPO10S184) et a permis un certain nombre de contributions MPEG sur la période 2004-2008. Ce projet a été achevé avec succès en 2008. D'une part, il a permis d'atteindre les objectifs annoncés pour faciliter la mise en œuvre de codec vidéo. D'autre part, il a permis d'initier un projet plus ambitieux nommé « Reconfigurable Video Coding » (RVC). Ainsi le développement d'un formalisme de type flot de données nommé CAL-RVC et des outils de prototypages rapides associés a été proposé comme plateforme de développement des futurs standards MPEG. A partir d'une description de ce type, il est possible de générer une partition logicielle et/ou matérielle. Les langages de description des décodeurs vidéo, le langage flot de données CAL-RVC permettant de décrire notamment l'ensemble des composants décodeurs ont été respectivement normalisés sous les noms **MPEG-B Decoder Description Language (ISO 23001-4)** (normative languages for the specification of a decoder configuration) et **MPEG-C MPEG Toolbox (ISO 23002-4)** (normative language for the specification of the toolbox).

En outre, nous avons cherché à montrer que cette plateforme pouvait être utilisée en dehors du domaine de la compression d'images. Ainsi la plateforme Wildcard a été utilisée dans le cadre de la thèse de Fethi Smach. Cette thèse portant sur la reconnaissance de formes a été dirigée par le Professeur Johel Mitéran. Les applications de reconnaissance de formes représentent un nombre croissant d'application de la vision artificielle au domaine industriel. La littérature abonde de techniques en reconnaissance de formes. Bon nombre de ces études ont pour objectifs de résumer une image à un nombre réduit de descripteurs qui ont comme propriété d'être invariants notamment aux déplacements (translation, rotation). Considérant les déplacements dans le plan, Gauthier et al [6] ont proposé une famille d'invariants, appelés descripteurs de mouvement, qui sont invariants en translation, en rotation, insensible aux changements d'échelle et à l'effet miroir. H. Fonga [7] a étendu l'utilisation des descripteurs de mouvement, définis de manière identique à ceux de Gauthier mais appliqués aux images en niveau de gris. Notre but, ici, était de montrer que tels descripteurs pouvaient être utilisés de manière souple et robuste pour la reconnaissance de formes basées sur des images couleur. Les descripteurs sont utilisés pour alimenter un classifieur de type « Support Vector Machine » (SVM), qui est une méthode supervisée. Les SVM cherchent un hyperplan permettant de séparer l'ensemble d'apprentissage de sorte que tous les points d'une même classe soient d'un même côté de l'hyperplan. Parmi l'ensemble des hyperplans satisfaisant ces conditions, les SVM cherchent celui qui maximise la distance entre l'hyperplan et les points les plus proches de chaque classe. Nous appliquons les phases classiques d'apprentissage et de décision ( Figure 9) de la manière suivante :

- ✓ L'image est redimensionnée au format 128x128 pixels.
- ✓ La FFT2D est calculée pour les 3 canaux rouge, vert et bleu.
- ✓ Les descripteurs de Fourier généralisés sont calculés pour chaque canal.

- ✓ L'ensemble des descripteurs (soit  $3 \times 64 = 192$  valeurs) est utilisé pour alimenter le classifieur, qui génère un modèle en ce qui concerne la phase d'apprentissage, ou classe l'élément inconnu en ce qui concerne la phase de décision.

Au cours de la phase de décision, les descripteurs sont calculés de manière identique à la phase d'apprentissage. La prise de décision s'effectue directement grâce à la fonction de décision des SVM.

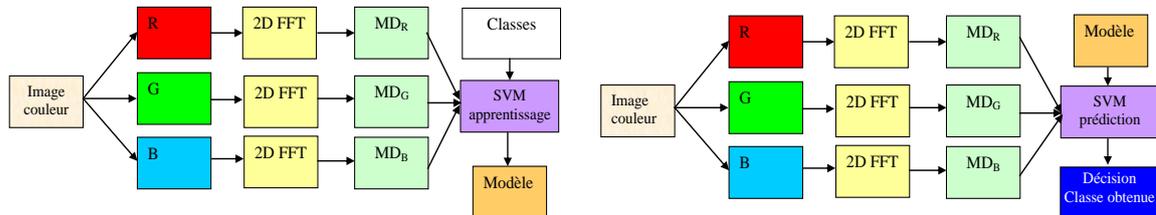


Figure 9. Phase d'apprentissage et phase de décision

Nous avons évalué les invariants des descripteurs généralisés de Fourier appliqués à la reconnaissance d'objets couleur sur les bases COIL et AR-faces. Pour montrer la robustesse et les performances de cette approche, nous avons réalisé des tests avec variation d'éclairage et de bruit, ainsi qu'une comparaison avec d'autres méthodes appliquées aux mêmes bases. Nous avons implanté la chaîne de calcul des descripteurs de Fourier généralisées sur la carte WildCard PCMCIA, (Xilinx Virtex2 XCV300E). La fonction FFT2D a été décrite en VHDL et intégrée au reste de l'algorithme (décrit) en C/C++. La plateforme de co-processing a ainsi permis d'intégrer rapidement cette fonctionnalité afin d'assurer un traitement temps réel de la phase de décision. Le débit bus PCI a permis de supporter le flot important des données.

Ces différents projets ont démontré l'intérêt de ce type de plateforme. Pour autant, l'utilisateur est limité à l'utilisation de l'architecture proposée. Par conséquent, nous nous sommes intéressés à une autre approche : l'utilisation d'une modélisation haut niveau associée à des outils de HLS pour permettre d'obtenir un prototypage rapide d'un accélérateur matérielle de manière « indépendante » de la plateforme visée.

### 3.3 Implantation d'accélérateurs matériels dédié à la reconnaissance de visages à partir de modélisation en SystemC

Ces travaux se situent dans la continuité des études menées sur l'analyse d'images pour la reconnaissance d'objets en temps réel menées durant la thèse de Monsieur F. Smach. Cette approche est issue des travaux de thèse de Monsieur Khalil Khattab, portant sur la définition d'un accélérateur matériel permettant la détection et la localisation de visages dans une scène vidéo. J'ai assuré le co-encadrement de cette thèse dirigée par le Professeur Johel Mitéran. Cette thèse a été soutenue en 2009. Lors de ces travaux, nous avons par ailleurs estimé l'impact de l'utilisation d'un outil de HLS à partir d'une modélisation en langage SystemC sur la définition et l'implantation d'architecture matérielle. L'outil de HLS utilisé était l'outil commercial System Crafter.

Plusieurs solutions algorithmiques robustes ont été proposées avant le début de ces travaux. Toutefois, les recherches dans le domaine de vision par ordinateur et notamment la reconnaissance d'objets et des formes se focalisaient généralement sur les parties algorithmiques et fonctionnelles des techniques étudiées. Ceci mène généralement à des implantations sur des stations de travail peu contraintes en puissance de calcul et en taille mémoire. La plupart de ces techniques, même si elles permettent d'obtenir de bonnes performances en termes de détection, ne sont pas adaptées à des applications ou systèmes temps réels, c'est-à-dire en ce qui nous concerne de l'ordre de 25 images/s, pour une résolution standard de type 320x240. Toutefois, la publication par Viola et Jones en 2001 d'une nouvelle méthode est venue modifier significativement l'état de l'art [8]. Leur algorithme, basé sur une cascade de classifieurs de type boosting a permis des performances en termes de vitesses de détection plusieurs fois plus rapides que les autres méthodes de la littérature, tout en conservant un taux de détection élevé et un nombre de fausses alarmes acceptable. De plus la méthode de Viola et Jones peut également être utilisée pour détecter d'autres types de formes au sens large, telles que les véhicules, les vélos et les motos, les piétons, etc. Le prototype initial de cette méthode réalisé par Viola et Jones est une implantation logiciel sur un PC, en utilisant des classifieurs de type Adaboost Discret. L'approche utilisée est un algorithme optimisé pour une implantation séquentielle qui consiste à construire une cascade d'étages de classification dont la complexité est progressive. Cette première mise en œuvre a montré de bons potentiels en réalisant de bonnes performances en termes de qualité de classification et de vitesse. En effet Viola et Jones ont obtenu une vitesse de traitement de 15 images par seconde sur un poste de travail, pour des images de résolution de 320x240. Une telle implantation sur des postes de travail ordinaires offre une flexibilité importante, et peut être développée et optimisée en peu de temps et à faible coût, grâce à une large variété d'outils de conception bien établie pour le développement logiciel. Cependant, cette mise en œuvre peut occuper toute la puissance de calcul du processeur pour cette seule tâche. Or la détection de visages est considérée comme étape préalable pour de nombreuses applications. D'autre part, des capteurs haute résolution et rapides ont été développés récemment et conduisent à de nouvelles exigences. Dans ce cas, les solutions logicielles standard ne permettent plus d'obtenir des performances temps réel.

L'objectif de ces travaux était donc de contribuer à accélérer et embarquer, sur des systèmes disposant de ressources limitées, des opérateurs de détection de visages multiples dans une scène non contrainte. L'accent a été mis sur la généricité du détecteur choisi, afin que ces travaux puissent être étendus à d'autres domaines. Dans ce contexte, nous proposons d'utiliser une méthodologie de conception comprenant :

- ✓ une étude des flots de données d'un algorithme,
- ✓ des modélisations à plusieurs niveaux d'abstraction,
- ✓ la synthèse haut-niveau des modélisations proposées et les raffinements correspondants.

La cascade proposée par Viola-Jones comprend 25 étages de classifieurs. Les trois premiers étages sont respectivement constitués de 3, 9 et 16 classifieurs faibles. Le nombre de classifieurs faibles pour les étages supérieurs peut être supérieur à 200. A partir des analyses faites sur le nombre de classifieurs faibles évalués et le nombre d'accès mémoire nécessaire par étage, nous avons remarqué que, malgré la différence de complexité entre les premiers étages et les derniers, le nombre de

classifieurs faibles évalués et le nombre d'accès mémoire nécessaires dans les deux premiers étages sont supérieurs à l'ensemble des nombre de classifieurs faibles évalués et nombre d'accès mémoire des étages restants. En effet 34,5% et 33,5% des accès mémoire du système entier sont réalisés respectivement dans le premier et le deuxième étage de la cascade, tandis que l'ensemble de tous les étages restants nécessite seulement 32% des accès mémoire. Ces analyses nous ont amenés à proposer une nouvelle solution en pipeline, formée de 3 blocs qui fonctionnent en parallèle sur 3 images différentes. Dans les deux premiers blocs (Etage1 et Etage2), nous avons implanté respectivement le premier et le deuxième étage du classificateur, et dans le dernier bloc (EtageA), nous positionnons une cascade séquentielle formée de tous les étages restants tels que représentés en Figure 10.

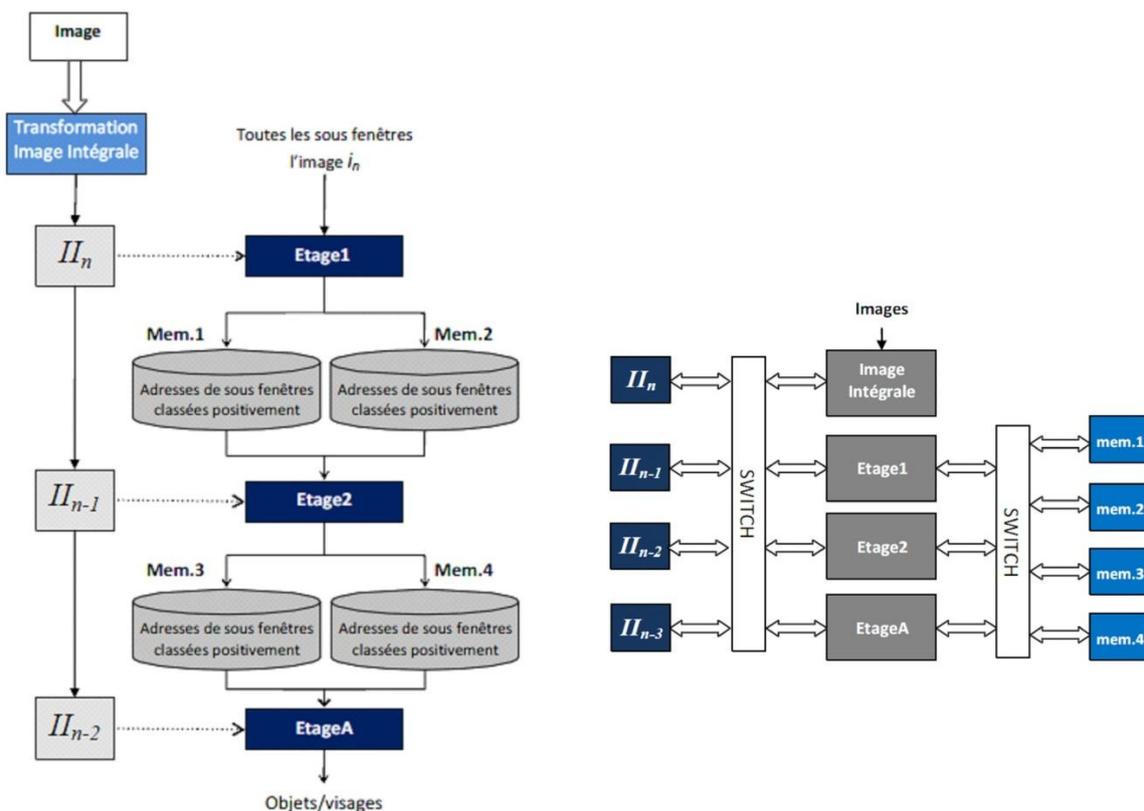


Figure 10. Exploitation du parallélisme de données dans la cascade de boosting

L'ensemble des traitements mais aussi des interfaces de communication a été modélisé en langage SystemC. L'architecture pipeline permet de réaliser les calculs des trois étages ainsi que celui de l'image intégrale en parallèle. Le temps de calcul est alors intrinsèquement réduit d'un facteur légèrement supérieur à 3. La difficulté majeure réside alors dans l'accès aux résultats de calcul de l'image intégrale sur plusieurs images décalées dans le temps mais aussi dans la transmission des résultats intermédiaires entre les étages. L'utilisation de différents bancs mémoires dans un mode multi-ports (double voir triple accès en ce qui concerne la gestion de l'image intégrale) est une solution envisageable. De manière réaliste, il apparaît nécessaire de regrouper ces bancs mémoires dans le minimum de composants physiques. Deux boîtiers de commutation, nommés SWITCH, ont été ainsi conçus afin de réaliser ces fonctionnalités de transfert de données en temps partagé sur une unique liaison physique. Une description SystemC à haut niveau d'abstraction des processus de communication permet d'élaborer efficacement l'architecture en charge des transferts de données.

Ces boîtiers intègrent des files d'attente « First In First Out ». Leur utilisation en simulation avec une taille infinie et l'analyse de leur taux de remplissage en cours d'utilisation permet de définir leur taille définitive, les largeurs de bus et les fréquences des différents éléments. Ces informations permettent in-fine la fusion de la plupart des bancs mémoires. La modélisation en SystemC à haut niveau d'abstraction des transferts de données a permis de définir un compromis intéressant utilisant respectivement une mémoire de type SDRAM pour les bancs relatifs au stockage des résultats intermédiaires et une mémoire de type SRAM pour ceux relatifs à l'image intégrale. Pour les deux boîtiers les simulations ont montrées que des fréquences inférieures à 25 MHz et une largeur de bus égale à 32 bits étaient suffisantes pour satisfaire les contraintes de notre application en termes de bande passante. Ces modules SWITCH représentent des éléments critiques pour l'architecture globale, ils ainsi ont été décrit en VHDL afin d'optimiser l'implantation matérielle. La possibilité de réaliser des simulations mixtes (HDL/SystemC) sur les environnements de simulation classique tels que ModelSim facilite le test des modules et réduit le temps de développement.

L'efficacité de modélisation SystemC est clairement apparue pour la mise en place de l'architecture globale. Cependant, l'outil de synthèse de haut niveau utilisé, à savoir SystemCrafter, a produit des résultats modestes à partir d'une description en SystemC à haut niveau d'abstraction. L'architecture globale possède alors une cadence inférieure de 4 images/s. Des raffinements du code proposé ont été alors nécessaires pour obtenir une architecture plus performante. Plusieurs procédures et mécanismes ont été utilisés afin de raffiner le niveau de détail de la modélisation et de décrire le parallélisme de l'algorithme, notamment le dépliage des boucles, le pipeline des instructions, le traitement en parallèle des différents flots de données et de contrôle, et enfin l'utilisation de primitives issues de bibliothèques optimisées. Il faut noter ainsi que selon les cas nous pouvons utiliser ces techniques séparément ou de les associer. Les raffinements réalisés ont mené à des modélisations possédant le même degré de détail qu'une description de type HDL, notamment pour la plupart des modules de traitement. L'amélioration de performances est alors notable. Le système implanté est capable de traiter jusqu'à 44 images par seconde pour des résolutions d'images de 320x240. L'objectif de traitement en temps réel a été atteint pour cette résolution à partir d'une modélisation SystemC. Le résultat de la détection est illustré Figure 11.

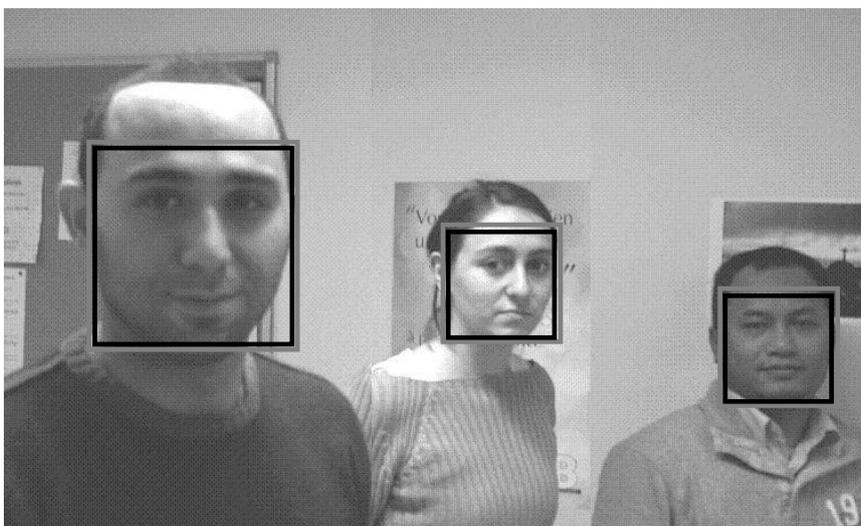


Figure 11. Détection de visages obtenue à partir de l'architecture proposée

D'autres modélisations à partir de langages à haut niveau d'abstraction et d'autres outils de HLS peuvent être bien entendu considérées. Nous avons aussi utilisé le flot de conception associé au langage flot de donnée CAL-RVC proposé par le comité MPEG. Cette étude sera décrite dans la section suivante.

### **3.4 Implantation d'accélérateurs matériels pour l'estimation du mouvement dédiés la compression vidéo.**

Au vue de la diversité des standards de compression, de l'hétérogénéité des média de communication et des terminaux utilisés, il paraît primordial de pouvoir adapter les performances de codage en fonction des évolutions des contraintes liées à l'environnement, voire du contenu de l'image ou des événements détectés dans la scène (comme notamment décrit en partie 2.2).

D'autre part, les outils existants combinant rapidité, faible encombrement et faible consommation et qui réalisent la compression sont la plupart du temps basés sur des architectures spécifiques. En effet, devant la demande croissante en puissance de calcul, des applications de traitement du signal et particulièrement, des applications de codage d'images fixes et vidéos, il est de plus en plus courant de faire appel à la puissance des architectures matérielles afin d'accélérer le temps de traitement et respecter les contraintes temps réel. Ce type d'approche présente une puissance potentielle importante, mais la complexité de la description et le coût d'implantation d'une application peuvent s'avérer très élevés. Pour cette raison, plusieurs langages à haut niveau ont été mis au point afin de faciliter la conception hardware tout en profitant de la souplesse d'une description à un niveau d'abstraction le plus élevé possible.

Le premier objectif de ce travail était donc de proposer une architecture capable de concurrencer l'état de l'art, en termes de performances brutes d'encodage, et d'envisager plusieurs outils pour la description de cette architecture.

D'autre part, le deuxième objectif était que cette architecture soit flexible, permettant d'une part de modifier la stratégie de recherche et d'autre part de sélectionner les raffinements préconisés par les standards les plus récents. En effet, l'opération de mise en correspondance de blocs dépend de la stratégie de recherche choisie. L'étude exhaustive de toutes les positions du macro-bloc dans la zone de recherche représente la technique de référence. Il s'agit aussi de la méthode la plus coûteuse en temps de calcul. D'autres stratégies, sub-optimales en termes de performances de codage, utilisent un parcours particulier de la fenêtre de recherche pour limiter le nombre de tentatives de mises en correspondances, et sont de ce fait beaucoup plus rapides.

La flexibilité vise à pouvoir adapter et optimiser l'estimation du mouvement réalisée pour la prédiction de l'image courante par rapport à une image précédemment encodée/décodée selon le choix d'utilisateur ainsi que le type d'application. Il nous a donc été nécessaire d'identifier les éléments pouvant être utilisés au sein d'un modèle commun d'architecture afin d'arriver à notre objectif final, l'implantation de l'ensemble de l'architecture et la comparaison avec des architectures existantes.

Les travaux réalisés s'articulent autour de deux phases de recherche ayant pour but de proposer de nouvelles architectures matérielles d'estimateurs de mouvement flexibles et améliorant les performances en compression des codeurs vidéo actuels. Cette approche est issue des travaux de thèse de Monsieur Elhamzi Wajdi, soutenue en février 2013. J'ai assuré le co-encadrement de cette thèse réalisé en cotutelle avec l'Université de Monastir. Cette thèse a été dirigée par le Professeur Johel Mitéran. Le projet a été supporté financièrement par le projet EGIDE UTIQUE dont j'ai assuré la responsabilité. Ce travail a été effectué dans le cadre d'une coopération franco-tunisienne, projet CMCU intitulé « Plateforme de compression vidéo pour des applications basse résolution incorporant des mesures temps réel de reconnaissance de formes » au sein de deux laboratoires : Laboratoire Electronique, Informatique et Image (Le2i) de l'Université de Bourgogne et le Laboratoire d'Electronique et Microélectronique (LEME) de l'Université de Monastir.

La première phase de nos travaux a été davantage orientée vers l'étude générale algorithmique et architecturale des différentes stratégies de recherche afin de proposer une solution permettant de supporter les différents paramètres de configuration visant à améliorer l'encodeur standard H.264/AVC de façon incrémentale. Dans cette phase, nous avons identifiés principalement quatre contributions qui apportent des gains significatifs par rapport aux travaux présents dans la littérature :

- ✓ Une analyse des différentes stratégies et architectures existantes, qui nous a amenés à définir un modèle modulaire et flexible.
- ✓ Une nouvelle architecture matérielle d'un estimateur de mouvement pixélique à taille de bloc variable, capable de supporter différentes stratégies de recherche.
- ✓ Deux nouvelles architectures des unités d'interpolation demi et quart de pixel, profitant de parallélisme des données, afin de réduire le temps de traitement de raffinement sub-pixélique.
- ✓ Deux nouvelles architectures concernant l'estimation de mouvement sub-pixélique (séquentielle et pipeline).

L'architecture réalisant l'estimation pixélique est comparée d'un point performance avec les architectures les plus populaires utilisant 16 unités de calcul [9][10][11]. Nous obtenons des performances identiques en termes de nombre de cycles d'horloge pour le traitement d'un macro-bloc 16x16 avec une fenêtre de recherche de 16x16 (4096 cycles). De plus, suivant les besoins de l'utilisateur, cette architecture permet de modifier la stratégie de recherche, par

exemple en choisissant l'algorithme « Diamond Search (DS) » à la place de la recherche exhaustive « Full Search (FS) ». Il est ainsi possible de modifier la qualité d'encodage, les débits de transmission, ou encore le temps d'encodage en fonction des contraintes liées à l'application.

L'utilisation stratégies de recherche non exhaustive réduit de manière particulièrement significative le temps de calcul (le temps de calcul peut être divisé par facteur 100). Ainsi, il apparaît nécessaire de porter une attention particulière à l'implantation de la phase d'estimation sub-pixélique, puisque en effet cette phase peut se révéler alors comme l'étape la plus consommatrice en temps de la phase d'estimation du mouvement globale. Par conséquent, nous avons proposé des architectures exploitant le parallélisme de données de cette phase de raffinement sub-pixélique. La définition d'une unité d'interpolation originale permettant les traitements simultanés de plusieurs sous-blocs de largeur 4 ou 8 a permis d'améliorer de manière significative les performances de l'état de l'art. D'un point de vue performances de codage, les deux solutions architecturales proposées (séquentielle et pipeline), basées sur cette unité d'interpolation permettent d'atteindre les plus forts débits en termes de nombre de MacroBlocs traités par seconde à la fois pour le raffinement demi pixel mais pour celui quart de pixel. Ainsi l'architecture séquentielle permet d'atteindre 649 K MacroBlocs/s pour le raffinement demi-pixel, soit un facteur d'accélération respectivement de 6.49, 1.30 et 1.25 obtenu par rapport aux architectures de Chen [12], Yang [13] et Thang Ta [14]. En ce qui concerne l'architecture pipeline, elle permet d'atteindre 538 K MacroBlocs/s pour le raffinement quart de pixel, soit un facteur d'accélération respectivement de 10.98, 2.15 et 2.07 obtenu par rapport aux mêmes architectures.

La seconde phase de nos travaux est davantage consacrée l'étude de l'optimisation de temps de mise sur le marché d'un tel système, par réduction du temps de développement. Une modélisation et description de type flot de donnée en langage CAL [15] de l'estimateur de mouvement pixélique est réalisée. L'utilisation de la plateforme ORCC [16] a permis de traduire cette description en langage VHDL. La comparaison d'une implantation basée sur le code généré par cette plateforme a pu alors être réalisée avec celle basée sur une description manuelle de l'architecture proposée. Une discussion est aussi menée dans le mémoire de thèse correspondant pour démontrer l'apport de telles descriptions « haut niveau » au niveau de la conception d'une architecture hétérogène d'un estimateur du mouvement supportant la stratégie sélectionnée.

L'étude des résultats de l'implantation pour une cible de type FPGA Virtex6 a fait apparaître le surcoût matériel apparaît relativement important puisque le nombre de blocs logiques (ou Slices) est supérieur d'un facteur 3 aux résultats obtenus lors de l'implantation de la description manuelle. Les débits sont quant à eux réduits d'un facteur 2,25. Finalement le temps de conception a été réduit d'un facteur 4. Des travaux menés antérieurement [C16] avaient montrés qu'il était possible d'égaliser les performances d'une implantation réalisées à partir d'une description manuelle. Cette perte de

performance au niveau de l'unité de traitement, s'explique par les opérateurs mis en jeu (notamment extraction de bits) et leur implantation par l'outil de HLS comme décrit dans [C7][C13].

Bien que l'implantation matérielle obtenue à partir du code automatiquement généré à partir de la description flot de données ne permette d'obtenir que des performances relativement modestes par rapport à l'implantation réalisée à partir du code VHDL décrit manuellement, cette solution nous apparaît tout de même très intéressante en raison de sa flexibilité, de la compacité de la description, de la réduction du temps de développement mais aussi et surtout, car elle offre la possibilité de proposer une architecture hétérogène à partir d'un modèle unifié. En effet la possibilité de générer du code en C, à partir de la même description en CAL-RVC, permet de proposer une flexibilité supplémentaire vis-à-vis de la problématique traitée. En fonction de la complexité et de la régularité de l'algorithme de recherche choisi, il est très simple de modifier la partition matérielle/logicielle et de traduire cet acteur en code C ou VHDL.

Ainsi, en ce qui concerne notre problématique particulière de compression vidéo, plusieurs solutions d'implantation peuvent être proposées très rapidement grâce aux outils de génération automatique de codes. L'approche de conception CAL facilite les tâches de validation de cette nouvelle architecture, ce qui nous permet de choisir, pour l'acteur génération des adresses (définissant la stratégie de recherche au pixélique), une implantation logicielle/matérielle (PC/FPGA) ou exclusivement matérielle comme respectivement illustré Figure 12 et Figure 13. Les outils de génération automatique associés au formalisme CAL peuvent fournir l'architecture complète. En effet, outre les partitions matérielles et logicielles, les interfaces de communication (par exemple une interface PCI), peuvent être générées automatiquement ce qui facilite encore l'implantation d'une architecture hétérogène tel que nous le décrivons dans la section 3.5

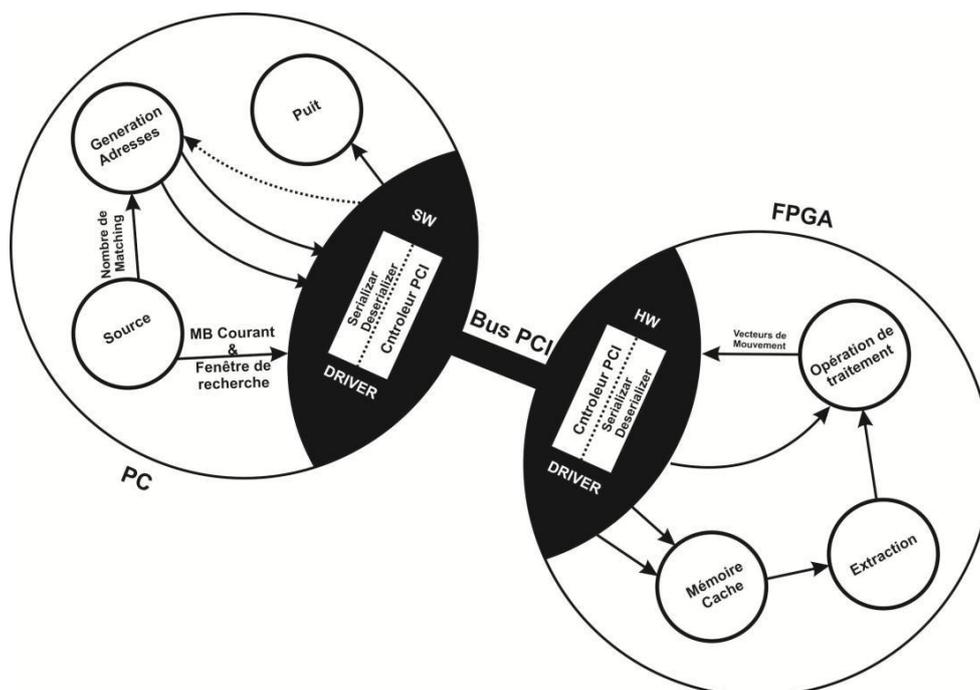


Figure 12. Implantation matérielle/logicielle de l'estimation de mouvement

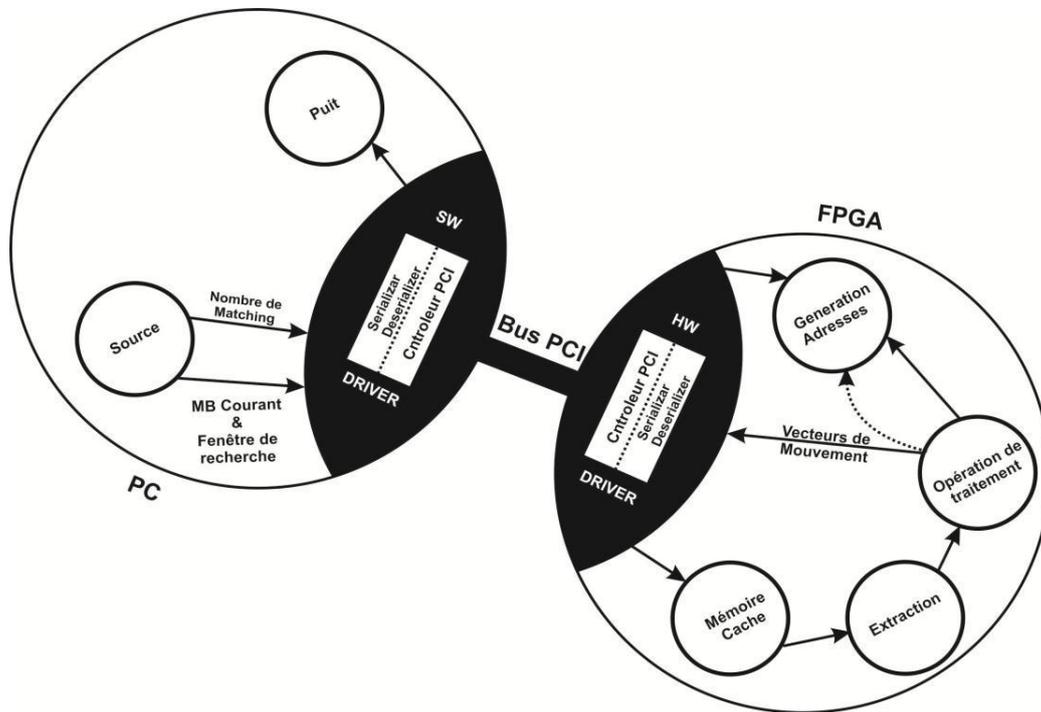


Figure 13. Implantation matérielle de l'estimation de mouvement

### 3.5 Prototypage rapide d'accélérateurs matérielles : génération automatique d'interfaces de communication

De nos jours, un grand nombre de méthodologies existent pour concevoir et déployer les systèmes. Cependant, l'intégration des communications entre les unités de traitement en fonction des spécifications fonctionnelles nécessite un temps important. Dans ce contexte, le défi est de fournir un moyen automatique pour les intégrer tout en étant conscient de l'environnement du système et des contraintes du concepteur. Par conséquent, une méthode de modélisation de système a été étudiée pour définir et générer automatiquement des communications en utilisant une description de type de flot de données CAL-RVC. La méthodologie proposée permet de gérer les interfaces, les contrôleurs et négociateurs réseaux, tout en tenant compte de la future génération de réseau avec les aspects de gestion de flux multiple pour tous les périphériques traditionnels. Ces travaux ont été menés lors de la thèse de Monsieur Richard Thavot au sein de l'École Polytechnique Fédérale de Lausanne (EPFL). Cette thèse, soutenue en décembre 2012, a été dirigée par le Professeur Marco Mattavelli. Notre contribution a été essentiellement la mise en place d'un nœud de communication générique.

La modélisation d'un traitement d'images en langage CAL permet, comme traditionnellement dans les langages flot de données, de décrire des actions (ici nommé acteurs) et de les associer sous la

forme de réseaux. Les acteurs sont donc reliés par des arcs de communication unidirectionnels. Comme précisé précédemment, le flot de conception associé au langage CAL permet de spécifier la partition logicielle et la partie matérielle (Figure 14). L'idée originale de ce travail est de simplifier la mise en place des interfaces de communication entre ces deux partitions mais aussi avec l'environnement extérieur.

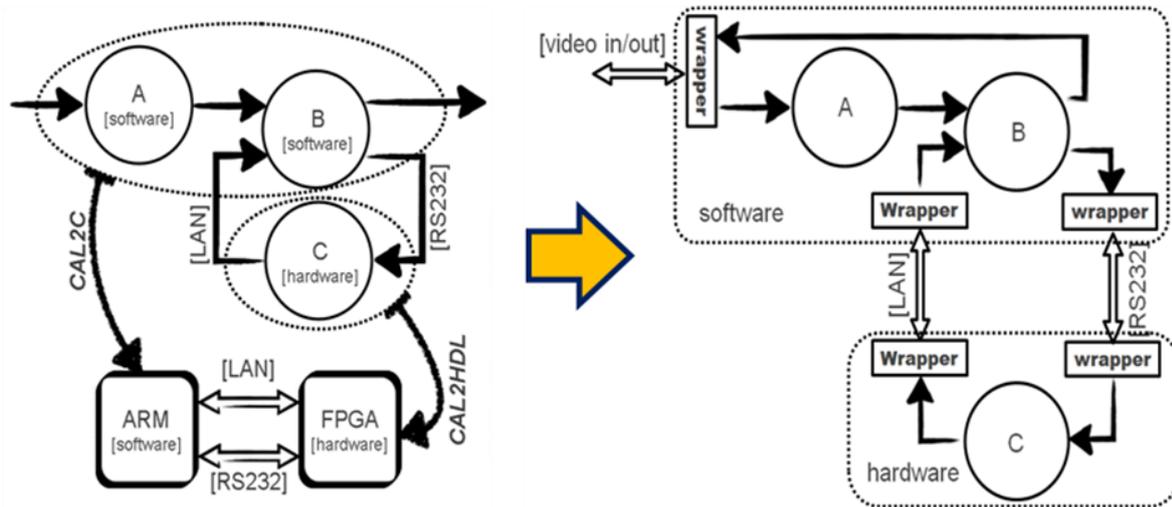


Figure 14. Modélisation flot de donnée et définition des interfaces communication

Ainsi la méthode proposée permet de sélectionner une interface de communication standard pour un arc de communication donné, puis de regrouper plusieurs arcs sur la même interface physique, et enfin de générer automatiquement les contrôleurs matériels et les pilotes permettant d'insérer l'interface de communication au sein de la cible matérielle. Pour cela, un nœud de communication générique a été défini. Il comporte :

- ✓ des sérialiseurs et désérialiseurs permettant de connecter plusieurs arcs de communications sur une même interface physique,
- ✓ un contrôleur matériel (IP) issu des bibliothèques proposé par le fabricant de FPGA,
- ✓ un processeur embarqué permettant le séquençage global du système.

Les flux de données issus des sérialiseurs et des désérialisateurs sont contrôlés par une unité nommée « flow translator » qui permet de gérer la multi-connectivité. Différentes stratégies ont été implantées afin de gérer efficacement les échanges de données dans cette configuration, tel que le simple « round-robin » pour une gestion identique des différents flots mais aussi l'algorithme des banquiers pour affecter dynamiquement une priorité à chaque lien de communication relativement à la quantité de données transitant dans ce lien [R7].

Les pilotes logiciels sont générés automatiquement par l'outil pour chaque contrôleur matériel. La structure générique permet d'insérer les contrôleurs matériels sur le bus local lié au processeur. Le processeur est de type « soft-core » directement décrit en VHDL. Le concept a été testé à la fois pour la technologie Altera et Xilinx. Différentes implantations ont été réalisées [R7] [17] mettant en œuvre des liaisons Ethernet ou encore diverses liaisons séries pour valider avec succès le concept.

L'avantage de cette solution est de pouvoir utiliser la performance des contrôleurs développés par les fabricants de FPGA ou de leurs partenaires. Pour autant, des pertes de performances peuvent apparaître en fonction de la puissance du processeur mis en jeu et de la complexité du pilote à mettre en œuvre. Une étude devra être réalisée afin de chiffrer précisément ces pertes pour chaque liaison disponible.

Une autre configuration a fait l'objet d'une étude préliminaire. Le processeur peut éventuellement permettre simultanément de gérer plusieurs contrôleurs matériels tel que présenté Figure 15.

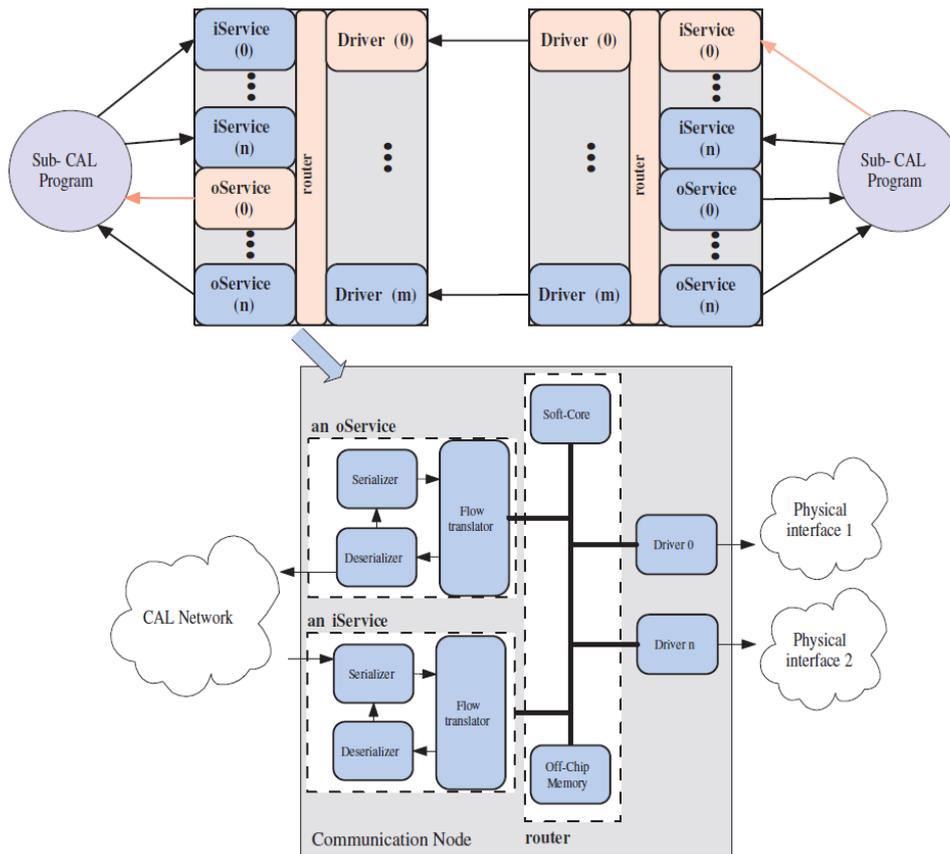


Figure 15. Evolution du nœud de communication

Le nœud de communication subit alors de légères modifications. Chaque arc de communication est associé à un service d'entrée ou de sortie (respectivement iService et OService) conservant la structure précédente (sérialisateur, désérialisateur et flow-translator). L'ensemble des interfaces de communication est connecté au processeur unique. Le processeur joue le rôle de routeur, permettant d'atteindre l'interface physique désirée. Grâce à l'association des services et le routeur, le re-routage dynamique des paquets de communication est alors rendu possible. Ce processus s'inspire des spécificités décrites au sein du protocole « Stream Control Transmission Protocol » (SCTP). De la même façon, il est ainsi possible d'envisager l'équilibrage dynamique de charge sur les différentes interfaces physiques (dans le cas d'un fonctionnement en mode « multi-homing ») ou encore la modification dynamique du choix d'interface de communication en cas de défaillance matérielle d'une liaison ou encore d'une attaque visant à pirater des informations. Ces perspectives

représentent, selon moi, des axes de recherche de travail très intéressants pour la mise en place de réseaux de capteurs intelligents.

### 3.6 Synthèse personnelle sur la conception d'accélérateurs matériels

L'exploitation du parallélisme potentiel de l'application représente certainement une clé pour l'obtention d'architectures matérielles efficaces. Le compromis entre le degré d'exploitation de ce parallélisme et les ressources matérielles requises permet d'ajuster les performances de l'architecture définie.

Une représentation permettant de décrire le parallélisme de flot de données tels que le langage CAL ou bien encore dans une moindre mesure le langage SystemC, peut apparaître très appropriée pour représenter ce parallélisme. D'autre part les outils de HLS associés permettent d'accélérer le prototypage des systèmes embarqués. Selon notre expérience avec les outils que nous avons évalués, les performances des implantations matérielles ainsi obtenues peuvent être tout à fait satisfaisantes moyennant des raffinements de la description. Ces méthodes et les outils associés demeurent perfectibles mais représentent certainement, selon nous, des pistes à considérer pour diminuer de manière importante le temps de développement des systèmes de vision embarqués. L'outil Vivado HLS apparaît lui aussi intéressant. Son évaluation est en cours au sein du laboratoire.

A l'heure actuelle, le degré de raffinement (éventuellement nécessaire) pour obtenir de hautes performances pour certaines parties de l'implantation matérielle, ou encore améliorer les performances de transmission obtenues pour certaines interfaces de communication générées et insérées automatiquement, nous amène à penser que des solutions hybrides permettant de remplacer (mais d'une manière « simple » et « rapide » !) lors de la synthèse le code généré automatiquement par un code décrit manuellement, seraient particulièrement intéressantes.

Les architectures figées comme la Wildcard associée à un PC standard ou encore de type System on Chip (SoC), comme la plateforme Zynq, offrant une architecture hétérogène de traitement logicielle/matérielle, apparaissent, elles aussi, comme une solution alternative intéressante non seulement en raison de leurs performances de traitement mais aussi pour la flexibilité proposée lors de l'utilisation d'un OS. Dans le cas la plateforme Zynq, cette solution permet en effet notamment une intégration rapide d'interfaces standard de communication. Pour autant, les performances de traitement obtenues pour ce type de solutions sont limitées par la plateforme utilisée. L'utilisation de langages à haut niveau d'abstraction associés aux outils HLS ne présente pas cette limitation puisque le choix des cibles n'est pas figé à l'avance.

### 3.7 Bibliographie relative à ce chapitre

- [6] J. P.Gauthier, G. Bornard, M. Zilbermann, "Harmonic analysis on motions groups and their homogenous spaces", IEEE Trans on Systems, Man and Cyb, vol 21, 1991.
- [7] H. Fonga, "Analyse harmonique sur les groupes et reconnaissance de formes", thèse, université de Grenoble, 1992.

- [8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features." CVPR, *Proceedings of the 2001 IEEE Computer Society Conference*. vol. 1, pp. I-511-I-518, 2001
- [9] Swee, Y.Ya., McCanny, J.V. : A VLSI Architecture for Variable Block Size Video Motion Estimation. IEEE Trans. on Circuits and Systems for Video Technology, vol.51, no. 7, 384–389 (2004)
- [10] Y. Song, Z.Liu, T. Ikenaga, S. Goto, : A VLSI Architecture for Variable Block Size Motion Estimation in H.264/AVC with Low Cost Memory Organization. IEICE Trans. Fundamentals, Vol E89, no. 12, pp. 3594–3601, (2006)
- [11] M.R.H. Fatemi, H.F. Ates, R. Salleh : A bit-serial sum of absolute difference accelerator for variable block size motion estimation of H.264. In Proceedings of of the Conference on Innovative in Intelligent Systems and Industrial Applications, pp. 1–4, (2009)
- [12] T.C. Chen , Y.W. Huang, L.G. Chen: Fully utilized and reusable architecture for fractional motion estimation of H.264/AVC. In Proc. IEEE ICASSP, pp. 9-12 (2004)
- [13] C. Yang, S. Goto, T. Ikenaga: High Performance VLSI Architecture of Fractional Motion Estimation in H.264 for HDTV. In Proc. IEEE ISCAS, pp. 2605-2608, Greece (2006)
- [14] N.T Ta,. J.R. Choi,: High performance fractional motion estimation in H.264/AVC based on one-step algorithm and 8x4 element block processing. Signal Processing: Image Communication, vol. 26, pp. 85-92, (2011).
- [15] J. Eker, J. Janneck, CAL language report. Tech. Rep. ERL Technical Memo UCB/ERL M03/48. University of California at Berkeley, December 2003.
- [16] G. Roquier, M. Wipliez, M. Raulet, J.W. Janneck, I.D. Miller, D.B. Parlour, "Automatic software synthesis of dataflow program : An MPEG-4 simple profile decoder case study", In IEEE workshop on Signal Processing Systems (SiPS), pp. 281-286, 2008.
- [17] R. Thavot, "High-level dataflow programming for complex digital systems: Methodology and tools for design, validation, and implementation", these, 2012

### 3.8 Publications jointes

Les travaux relatifs à la thématique présentée dans cette section ont permis une valorisation scientifique significative avec notamment la publication de 7 articles dans des journaux scientifiques internationaux, et 11 présentations dans des conférences internationales.

Une sélection de 3 articles publiés dans des revues internationales est proposée :

- [R2] Wajdi Elhamzi, Julien Dubois, Johel Mitéran, Mohamed Atri, Barthélémy Heyrman, Dominique Ginhac, "**Efficient smart-camera accelerator: an configurable motion estimator dedicated to video codec**", Journal of System Architecture, Ed. Elsevier, pp.8, 2013, **In Press**
- [R7] Richard Thavot, Romuald Mosqueron, Julien Dubois and Marco Mattavelli, "**Generation of Hardware/Software systems based on CAL dataflow description**", Algorithm-Architecture Matching for Signal and Image Processing, Best papers from Design and Architectures for Signal and Image Processing 2007 & 2008 & 2009, Series Lectures Notes in Electrical Engineering, Vol. 73, Springer Verlag, pp 275-292, jan. 2011.
- [R9] Khalil Khattab, Julien Dubois and Johel Miteran, "**Cascade Boosting Based Object Detection from High Level Description to Hardware Implementation**",

EURASIP Journal of Embedded Systems, Special Issue "Design and Architectures for Signal Image Processing", Hindawi, pp.12, 2009.

## 4 Conception et implantation d'architectures innovantes de caméras intelligentes

### 4.1 Contexte scientifique

L'idée de base de la caméra intelligente réside dans le fait d'associer au sein d'un même système le capteur en charge de l'acquisition des images et les unités en charge de leur traitement. Bien sûr, la communication avec le monde extérieur nécessite l'ajout d'interfaces de communication adaptées aux caractéristiques du capteur d'images et des traitements implantés au sein de la caméra. Les unités de traitement peuvent être basés sur trois catégories composants (ou leur combinaisons) :

- ✓ les composants programmables (processeurs génériques ou spécialisés comme par exemples les processeurs dédiés au traitement du signal (DSP),
- ✓ les composants de type configurables tel que les FPGA,
- ✓ ou finalement les composants figés tel que les ASICs.

Les accélérateurs matériels tels que présentés dans la section 3 permettent l'intégration d'algorithmes innovants (comme par exemple ceux présentés en section 2) au sein de caméras destinées à fonctionner en temps réel (c'est-à-dire à la cadence du capteur ou celle requise par l'application ciblée). Un des aspects les plus fondamentaux dans ce type de systèmes est certainement la gestion des différents flots de données au sein même du système comme à l'extérieur de ce dernier. Les trois architectures proposées dans les sections suivantes ont été élaborées en respectant ce critère qui est selon moi fondamental. De plus, en s'appuyant sur l'expérience forte du Le2i [18][19][20][21] dans la conception de démonstrations pour le domaine du contrôle par la vision et en particulier dans conception de caméras intelligentes chacun de ces travaux s'est accompagné de la réalisation d'un démonstrateur opérationnel.

### 4.2 Conception d'une caméra rapide intelligente haute résolution

Les caméras rapides sont de puissants outils pour étudier, par exemple, la dynamique des fluides ou le mouvement des pièces mécaniques lors d'un processus de fabrication. Durant la dernière décennie, l'utilisation des capteurs CMOS à la place des CCDs a facilité le développement des caméras rapides en offrant des interfaces numériques, une flexibilité de lecture et un faible coût de fabrication.

Dès 2003, l'objectif de nos recherches fut de produire à bas coût, des caméras intelligentes basées sur l'utilisation de capteurs CMOS mais non standard puisque fonctionnant à des hautes cadences (500 images par seconde) et possédant une résolution spatiale importante (> 1Mpixels). Le contrôle du capteur a été réalisé au sein d'un composant de type FPGA ce qui permis l'intégration des pré-traitements d'images et l'insertion de module de communication pour permet la connexion avec une

plateforme PC standard. L'intégration d'une liaison de type USB 2.0 a permis notamment la suppression de carte d'acquisition au niveau du PC. Cette approche est issue des travaux de thèse de Monsieur Mosqueron Romuald. Cette thèse a été soutenue en décembre 2006. J'ai assuré le co-encadrement de cette thèse dirigée par le Professeur Michel Paindavoine. Le système final est présenté Figure 16.



Figure 16. Caméra rapide intelligente à haute résolution

D'une manière générale, lors de l'utilisation de caméras rapides, l'important flot de données provenant du capteur ne pouvant être facilement transféré directement au PC, il doit donc être stocké temporairement dans une RAM local rapide. La taille de la RAM étant limitée, le temps total d'enregistrement ne dépassera pas quelques secondes pour des résolutions importantes. Nous avons ainsi proposé de développer une solution alternative permettant un enregistrement continu. Nous avons développé une compression d'images en temps réel pour réduire ce flot de données. Nous avons implanté et comparé 3 algorithmes: le codage par longueur de plage (run-length encoding), le codage par blocs et la compression en ondelettes. Ces algorithmes de compression ont été implantés dans un FPGA Virtex II-1000. Les images de résolution 1280x1024 pixels peuvent alors être codées à la cadence maximale de 500 images par seconde et ceci avec un taux de compression de 30:1 et un PSNR supérieur à 30dB. Le flot d'images peut alors être transmis en temps réel par liaison USB2. Avec de telles performances, il est possible de mémoriser de longues séquences d'images directement sur le PC sans utiliser de mémoire spécifique, ce qui est un avantage car nous pouvons profiter des améliorations constantes sur les PCs, spécialement sur les mémoires.

Avec cette approche, il nous sera possible de créer de nouvelles caméras avec des capteurs plus rapides (comme 1Mpixels@1000images par seconde ou plus) ou de plus grandes résolutions. En utilisant la technologie FPGA, l'intégration de nouveaux traitements d'images temps réel à l'intérieur de la caméra sont alors envisageables, par exemple le suivi d'objet (tracking), l'analyse d'image, ou reconnaissance de modèles. La Figure 17 illustre un algorithme de détection de centre de gravité mis en œuvre pour l'analyse du mouvement du petit animal.

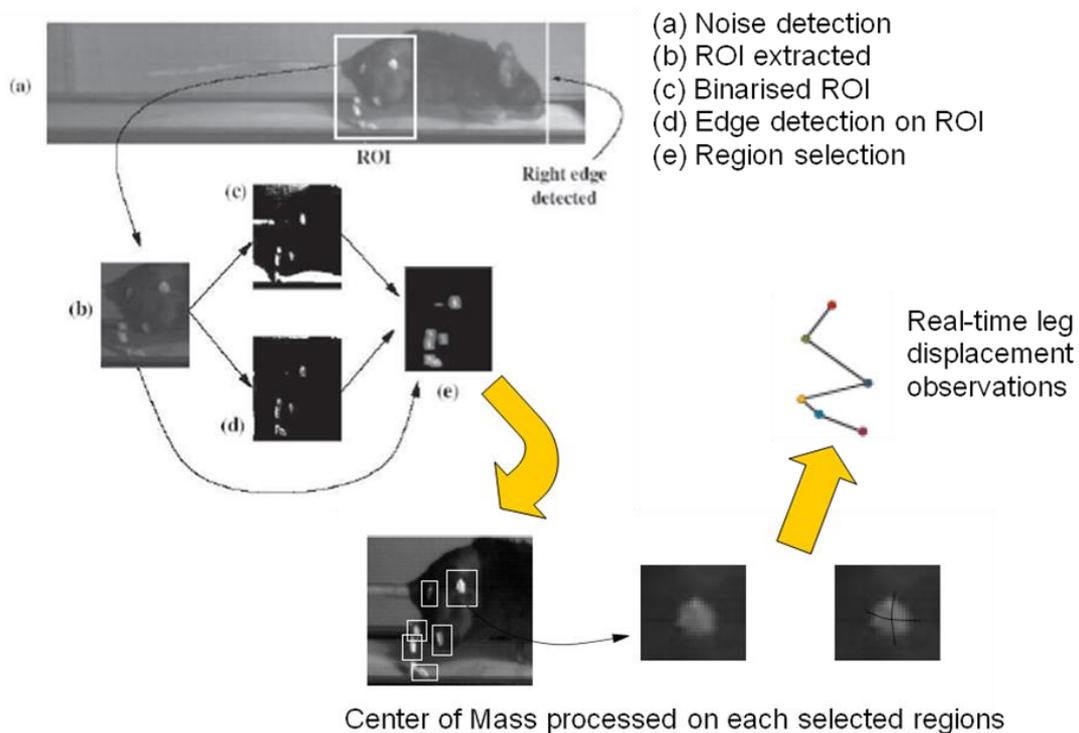


Figure 17. Chaîne de traitement pour l'extraction de centre de gravité

Cette figure illustre la mise en œuvre de pré-traitements simples de type morphologique, ou binarisation de zone d'intérêt et du calcul de centre d'intérêt. La mise en place de ces traitements nécessitant l'utilisation de voisinage de points est compliquée par la nature même du flux de pixels provenant de la caméra. En effet, 10 pixels sont transmis simultanément à chaque cycle d'horloge (fréquence de 50 MHz) en sortie de la caméra. Pour autant, une amélioration significative des performances de l'état de l'art a été obtenue avec ce système puisque d'une manière générale le traitement de 16 000 fenêtres de résolution 1280x32 par seconde est alors rendu possible.

Le centre de gravité peut être aussi alors mesuré pour chaque objet sur chaque ligne. Une application de profilométrie est ainsi présentée Figure 18. Il est alors possible de traiter 500 profils/s en pleine résolution ou encore jusqu'à 50 000 profils/s mais pour une image de 10 lignes de large.

Outre une valeur scientifique significative, ces travaux ont abouti à l'élaboration d'un partenariat fort avec la société Eynetics (<http://www.eynetics.fr/home.php>) mais aussi la création de la société FastVision (<http://www.fastvisiontechnologies.com/fvtsocietefr.html>).

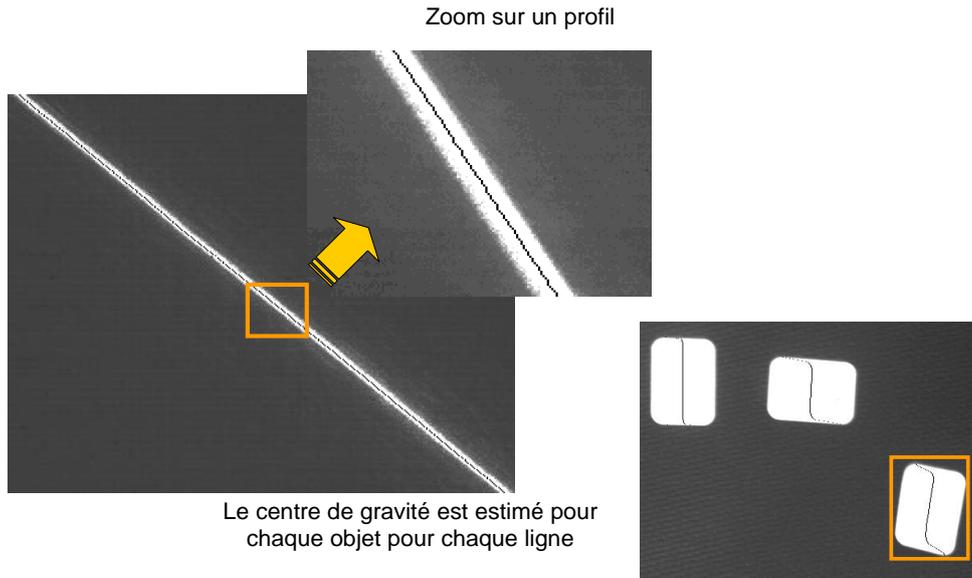


Figure 18. Mesure de centre de gravité par ligne et par objet

### 4.3 Caméra intelligente à unité de traitement hétérogène SW/HW

Ces travaux présentent une caméra intelligente basée une architecture hétérogène de traitement. Cette architecture regroupe un processeur multi-média de type Nexperia et un co-processeur vidéo développé spécifiquement pour ce projet. Le système permet l'acquisition d'images de résolution spatiale supérieure à 1 Mpixels et leur traitement à fréquence maximale de 30 images par seconde. Le système global est représenté Figure 19.

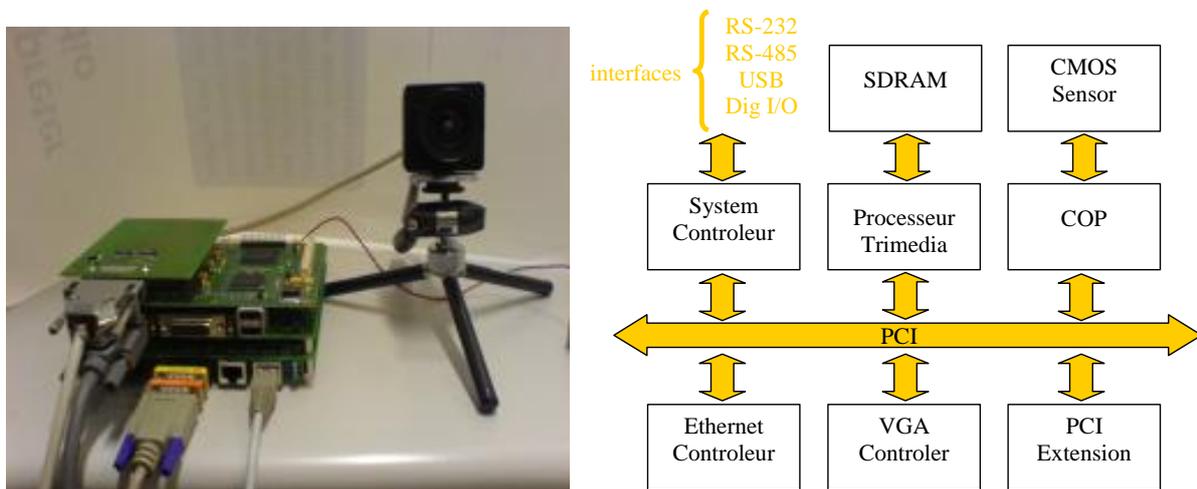


Figure 19. Caméra intelligente à élément de calcul hétérogène et architecture globale développée

Un bus fond de panier de type PCI permet la communication entre les deux composants principaux et les différentes interfaces de communication disponibles. La communication entre capteur CMOS et le co-processeur vidéo est assurée par une liaison de type CameraLink. Le co-processeur (COP) est basé sur deux composants de type FPGA (Figure 20). Outre ses larges capacités de calcul, il permet notamment la gestion en parallèle du flot de données issu de la caméra CMOS mais aussi des différents flots de données lors des traitements. Le flux vidéo provenant du capteur est

temporairement stocké au sein de la mémoire SDRAM. La relecture de la mémoire s'effectue en temps partagé et permet le transfert des images ou de zone d'intérêt à l'unité de traitement ou même au processeur multi-média. La mémoire SRAM est utilisée pour les phases de traitement. Son architecture originale permet l'enchaînement des traitements disponibles [R13]. La liste des traitements est évolutive en fonction des applications ciblées et l'utilisation d'un composant de type FPGA.

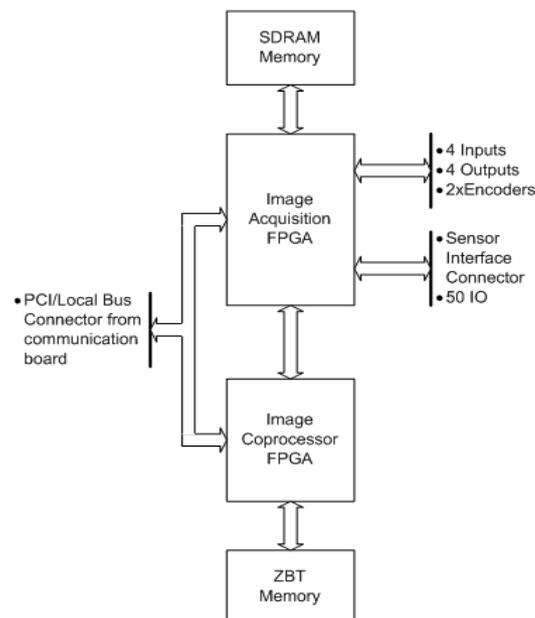


Figure 20. Architecture du co-processeur vidéo développé

Ces travaux ont été menés en collaboration avec l'Ecole Polytechnique Fédérale de Lausanne. J'ai initié ces travaux lors de mon parcours dans cet institut de recherche entre 2001 et 2003, et ceux-ci ont été poursuivis par le Dr Mosqueron Romuald dès 2006. Ce produit a été élaboré en partenariat avec la société Akatech (<http://www.akatech.ch/>) basée à proximité de Lausanne. Cette société réalise le tri postal pour l'ensemble de la poste suisse, danoise et une partie de celle française.

Les capacités de calcul ainsi que l'intérêt du partitionnement logiciel/matériel des tâches sont illustrées à l'application de tri postal développée pour la société Akatech. Ainsi l'objectif initial était de traiter le flot continu d'enveloppes circulant à 4 m/s devant un système en charge de réaliser le tri de ces dernières en fonction de leur destination, celle étant indiquée par l'intermédiaire d'un code barre présent sur ces enveloppes. Le co-processeur est en charge du traitement des tâches les plus régulières et le reste des tâches étant confiées au processeur multi-média. L'architecture de ce dernier a été conçue par la société Philips conjointement au compilateur en charge de génération de l'exécutable. L'objectif étant de conserver un code en langage C et ne pas décrire les tâches dans un langage assembleur. La chaîne de traitements mise en œuvre est représentée sur la Figure 21. La partie contrôle permet la sélection de zone utile dans l'image. Une série de filtrage définie par la société Akatech est tout d'abord réalisée afin de rehausser l'information utile dans l'image, puis l'image subit des opérations morphologiques. Un sous-échantillonnage de l'image permet de réduire la zone à transmettre au processeur multi-média. Ce dernier est en charge d'une opération de blobbing permettant de déterminer et donc de situer la zone utile. Cette zone est alors transmise

dans sa taille d'origine au processeur multimédia. Elle est lue et interprétée au niveau de cette unité. Le système permet alors une lecture à la vitesse de 8 m/s soit le double du cahier des charges original.

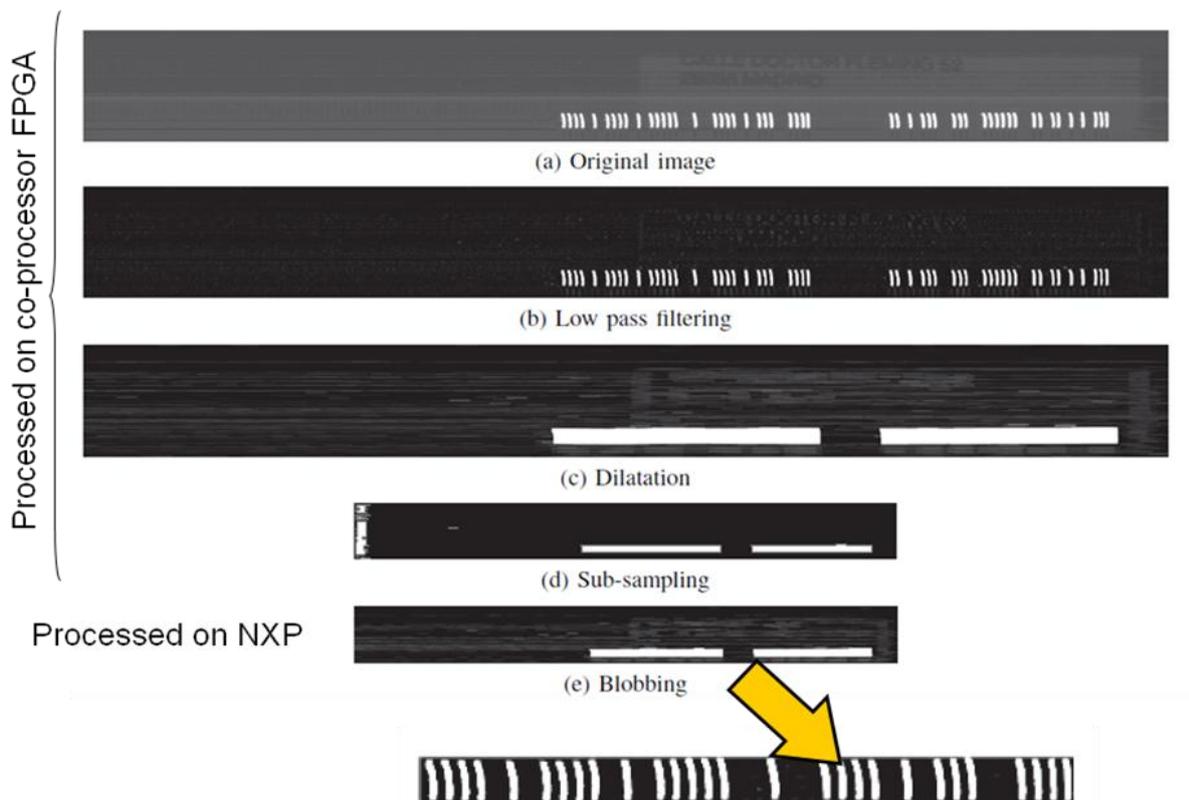


Figure 21. Séquençage des traitements pour l'extraction en temps réel de codes barres

#### 4.4 Caméra intelligente pour la haute dynamique

Le projet HiDRaLoN (High Dynamic Range Low Noise CMOS imagers) est un projet européen déposé en avril 2008 dans le cadre du programme EUREKA CATRENE (Cluster for Application and Technology Research in Europe on NanoElectronics). Ce projet a été évalué positivement par le comité scientifique de CATRENE réuni en assemblée en juillet 2008. Ce projet est géré au sein du Le2i par le Professeur Dominique Ginhac. Le projet HiDRaLoN s'intéresse au développement de nouvelles approches de conception de capteurs d'images CMOS. L'objectif était de concevoir, de développer et de mettre au point de nouvelles méthodes permettant d'accroître de manière drastique la dynamique des capteurs d'images CMOS tout en conservant, voire améliorant les autres caractéristiques importantes que sont la résolution, la qualité d'images, la sensibilité.

La participation du LE2I dans ce projet se situe à deux niveaux. Le premier est le développement d'algorithmes de correction d'images permettant d'éliminer les valeurs aberrantes des pixels. Il s'agit ici de veiller à fournir en sortie de la matrice de pixels une image de la plus grande qualité possible après avoir supprimé les pixels défectueux et les non-uniformités de type bruit fixe de colonne (FPN). Deuxièmement, le projet HiDRaLoN vise à augmenter de manière significative la dynamique des capteurs CMOS. Pour cela, les travaux menés par le LE2I avaient pour objectif de développer un système innovant de lecture de l'image réalisant plusieurs lectures partielles de l'image tout au long

de l'acquisition. Au final, l'image à la sortie du capteur est constituée de l'assemblage des différentes zones de l'image acquises à différents moments. Un tel système devait permettre d'une part d'éviter les problèmes de saturation pour les zones fortement éclairées et d'autre part de disposer d'informations suffisantes pour les zones faiblement éclairées.

Les retombées potentielles de ce projet et les champs d'applications de ces capteurs CMOS innovants sont très nombreuses. On peut citer les caméras professionnelles pour la télévision haute définition ou le cinéma numérique dans le domaine artistique et culturel. Mais également, on peut envisager diverses applications pour le contrôle et la surveillance des routes, pour la surveillance des lieux publics, pour les systèmes automatisés de contrôle de qualité par vision artificielle, ... Selon les applications visées, les résultats issus de ce projet HiDRaLon, et plus particulièrement les travaux sur le pixel, pourront permettre de concevoir des systèmes de vision parfaitement adaptés au domaine souhaité selon plusieurs directions différentes (Figure 22) :

- ✓ La diminution de la taille des pixels associée pour une diagonale du capteur constante permet de libérer de l'espace en périphérie de chaque pixel, espace qui peut être utilisé pour intégrer au niveau du plan focal des traitements d'images spécifiques. Le champ d'application visé ici est principalement le domaine des machines de vision pour applications industrielles.
- ✓ La réduction de la taille des pixels pour une même diagonale de capteur permet de créer des sous-groupes de pixels sur lesquels sont disposées des microlentilles dans l'objectif de pouvoir créer des images 3D des scènes acquises.
- ✓ Dans le cas d'une diagonale du capteur constante, l'utilisation de pixels de taille plus petite permet également d'augmenter la résolution horizontale et/ou verticale du capteur pour toutes les applications nécessitant une grande finesse de détails (surveillance, télévision, cinéma, photographie, ...).
- ✓ Pour une résolution constante, des pixels de taille plus petite permettent de concevoir des capteurs plus compacts et donc de réduire le coût de fabrication. Ceci est très bénéfique pour les systèmes grand public comme les webcams, la téléphonie mobile, etc.

J'ai pu collaborer à la fois sur la correction des défauts mais aussi la conception d'une caméra intelligente permettant de générer un flux vidéo traité à 60 images/s. Ces travaux font actuellement l'objet de la thèse de Pierre Jean Lapray dirigée par le Professeur Dominique Ginhac et co-encadrée par le Dr Barthélémy Heyrman.

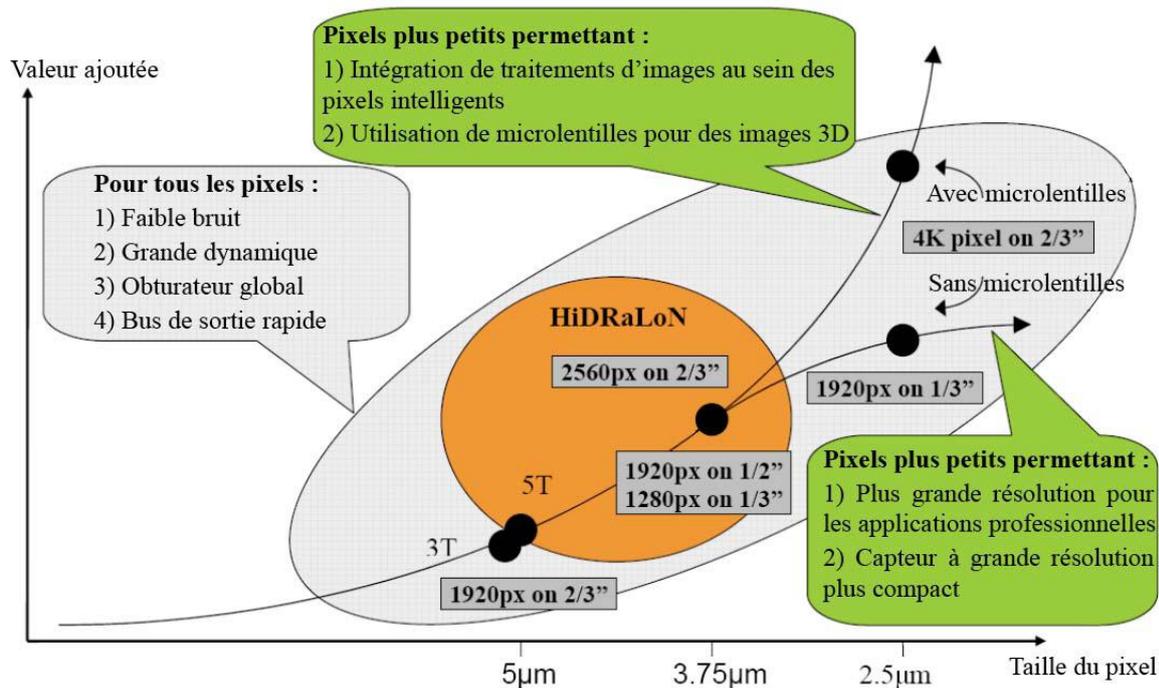


Figure 22. Différents champs d'applications du projet HiDRaLoN

Le système et l'architecture associée sont représentés Figure 23. L'architecture proposée a été implantée sur une carte évaluation de type ML605. Elle permet l'ajout d'une carte capteur fabriquée au sein du laboratoire le2i. Le capteur CMOS utilisé a été fourni par la société E2V partenaire du projet HiDRaLoN. Il permet l'acquisition des images de 1.3 Mpixels et ceci à la fréquence de 60 images/s.

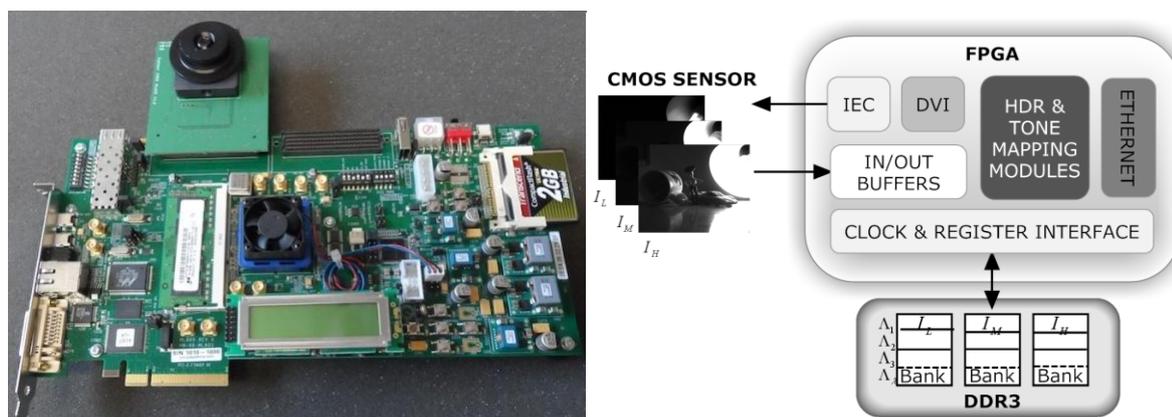


Figure 23. Caméra HDR et architecture associée

L'acquisition de trois images décalées dans le temps, chacune d'elle possédant un temps d'intégration différent (court, moyen et long) permet de générer l'image de type HDR (Figure 24).

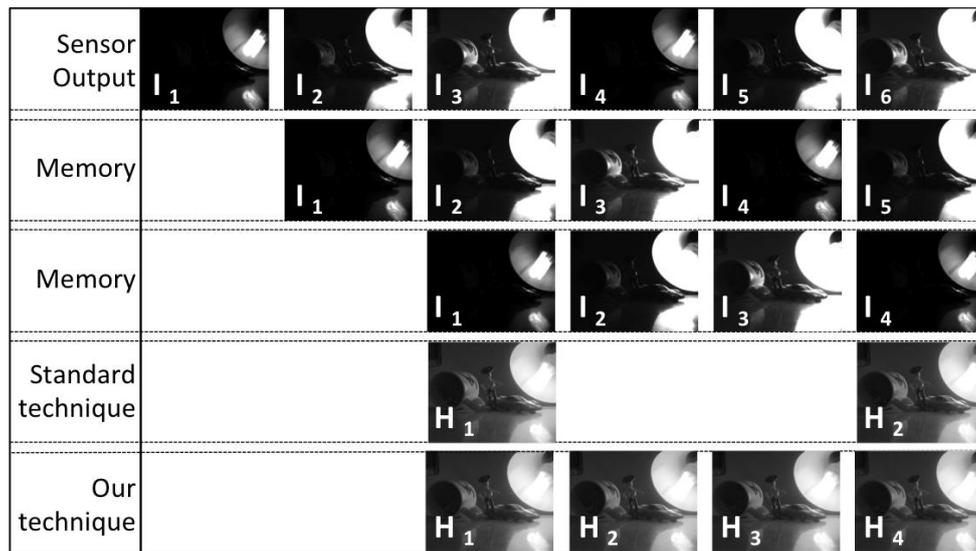


Figure 24. Séquencement des acquisitions et génération d'un flot continu d'images HDR.

Chaque pixel est codé sur un nombre flottant (norme IEEE 754) de 32 bits. Un algorithme de type « tune mapping » a aussi été intégré au sein du FPGA. Il permet la génération d'une image visualisable au niveau des écrans standards, chaque pixel possédant une dynamique de 8 bits. La transmission du flux vidéo peut alors être réalisée via une liaison Ethernet. L'implantation de l'ensemble du système sur le FPGA Virtex-6 XC6VLX240T présent sur la carte d'évaluation nécessite moins de 12 % des ressources matérielles présentes au sein de ce composant. Il est par conséquent envisageable d'intégrer in-situ des algorithmes de traitements utilisant directement les images HDR codées sur 32 bits.

Ces travaux font actuellement l'objet d'une phase de valorisation scientifique. Les perspectives sont par conséquent nombreuses et très prometteuses. L'intégration d'une correction de fantômes permettant de s'affranchir du phénomène de bouger obtenu en raison du décalage temporel présent entre les trois images nécessaires à la génération d'une image HDR est actuellement envisagée.

## 4.5 Synthèse personnelle sur la conception et la réalisation de caméras intelligentes

La caméra intelligente peut être considérée comme un capteur, certes relativement sophistiqué, permettant d'extraire des informations d'une scène vidéo. Le traitement mis en jeu, responsable de l'extraction de ces informations, peut être considéré comme le cœur du système global. En effet, il réalise le lien entre l'acquisition et la communication vers l'extérieur du système. Ainsi des commandes utilisateur provenant de l'extérieur du système mais aussi les résultats obtenus au niveau du traitement peuvent permettre de configurer dynamiquement à la fois l'acquisition (modification de temps d'exposition, extraction de zones d'intérêts...) et les caractéristiques de la communication choisie (la liaison physique choisie, débit de la liaison, ...). Selon moi, la flexibilité offerte par la caméra intelligente pour s'adapter dynamiquement aux évènements détectés ou à une

## 5 Conclusions & perspectives de recherche

L'ensemble des activités réalisées au cours de ces dix dernières années s'inscrivent dans la thématique de conception et d'implantation de systèmes de vision « intelligents ». Le savoir-faire acquis au cours de travaux se situe au niveau :

- ✓ du développement de nouvelles méthodes de mesure ou traitement d'information de type image ou vidéo,
- ✓ du développement et l'implantation d'architectures nouvelles permettant une optimisation des traitements et de la gestion des flux de données,
- ✓ de la réduction des temps de développement de ce type de systèmes,
- ✓ de l'intégration des systèmes complets et la réalisation de prototypes opérationnels.

Cette expertise permettra, je l'espère, de proposer les systèmes de vision de demain car apparaissent de nombreuses perspectives à ces travaux. Ainsi, comme nous l'avons décrit précédemment (section 4.5), au sein d'un système de vision intelligent l'acquisition, le traitement et la communication avec l'environnement extérieur sont étroitement liés. L'intelligence des systèmes se mesure certainement, selon moi, aux interactions reliant les trois aspects. La flexibilité obtenue peut impacter non seulement les performances du système global mais aussi éventuellement faciliter au sein d'un réseau de capteurs (comme par exemple l'ajustement dynamique au cours du temps des traitements ou encore du débit de sortie du système intelligent en fonction des besoins ou contraintes du réseau).

L'intégration de systèmes « intelligents » au sein des réseaux de capteurs représente une des plus intéressantes perspectives sur laquelle nous avons commencé à travailler depuis plusieurs années. Une collaboration avec l'Université Technologique de Petronas (UTP) est en train de voir le jour sur cette activité. Les aspects de communications, de répartition des charges de calcul entre différentes unités disponibles et enfin le regroupement des informations ainsi que les heuristiques pour réaliser la prise de décision seront alors fondamentaux. L'étude de la consommation des systèmes intelligents a été très peu considérée dans les travaux exposés précédemment. Elle est pourtant essentielle dans le cas où les capteurs intelligents sont autonomes, comme ils peuvent l'être dans certain cas dans les réseaux de capteurs. La minimisation de la consommation pourra ainsi faire, elle aussi donc l'objet d'études.

Devant l'ensemble de ces challenges futurs (et la charge des projets actuels, d'ailleurs...) il est apparu indispensable de fédérer un nombre significatif d'enseignant chercheurs sur cette thématique de recherche. Au sein du le2i, ce projet regroupe ainsi quatre personnes Dominique Gin hac, Johel Mitéran, Barthélémy Heyrman et moi-même. Dans cette esprit fédérateur, nous organisons depuis cette année ainsi des réunions scientifiques bimensuelles où les projets les informations sur les activités liés à ce projet sont diffusés aux participants et où l'on réalise des réflexions collégiales sur des sujets techniques portant sur cette thématique (étude d'un article issu de l'état de l'art, tutoriel sur une action précise). D'un point de vue national ou international, nous participons activement aux

actions possibles au sein des GDR, du workshop international WASC et contribuons à l'établissement d'un réseau autour de cette activité, ce qui s'est d'ailleurs concrétisé récemment par le dépôt du projet COST **EUNEVIS (European Network on Embedded Vision Systems)** regroupant plus de 15 partenaires européens. Finalement, la communication autour de cette thématique auprès de partenaires éventuels (industriels ou académiques) s'est aussi intensifiée grâce à notre activité.