

Modélisation, évaluation et contrôle d'une chaîne de traitement d'images routières

A. DOMINGUES¹, Y.LUCAS², D.DRIOUCHI³, P.MARCHE¹

¹Laboratoire Vision et Robotique, ENSI de Bourges 10 bd Lahitolle 18000 Bourges

²Laboratoire Vision et Robotique, IUT Mesures Physiques 63 av. de Lattre 18020 Bourges cedex

³Laboratoire de Statistique Théorique et Appliquée, Université Pierre et Marie Curie, 175 av. du chevaleret 75013 Paris

Yves.Lucas@bourges.univ-orleans.fr, {Antonio.Domingues, Pierre.Marche}@ensi-bourges.fr, driouchi@ccr.jussieuu.fr

Résumé – La mise au point d'une chaîne complète de traitement d'image (CTI) est toujours très délicate. Jusqu'à présent la communauté s'est intéressée à l'évaluation d'algorithmes isolés, sur un petit nombre d'image test et avec un réglage adhoc des paramètres indépendant des données d'entrée. Nous montrons ici comment en combinant la modélisation statistique par plans d'expérience, l'optimisation numérique et l'apprentissage neuronal, on peut élaborer une CTI adaptative performante. Il faut aussi pour cela disposer d'une base d'images conséquente, être capable de caractériser en entrée les images brutes et enfin évaluer les sorties de la CTI. Nous montrons comment, dans le cas d'une CTI dédiée à la détection d'obstacles routiers, cette méthodologie expérimentale et l'architecture logicielle associée, assurent en permanence l'efficacité des traitements. L'explication en est simple : la CTI est optimisée dans sa globalité, à partir d'un grand nombre d'images test réelles et en adaptant les traitements aux données d'entrée.

Abstract – *Tuning a complete image processing chain (IPC) remains a tricky step. Until now researchers focused on the evaluation of single algorithms, based on a small number of test images and ad hoc tuning independent of input data. In this paper we explain how, by combining statistical modeling with design of experiments, numerical optimization and neural learning, it is possible to elaborate a powerful and adaptive IPC. To succeed, it is also necessary to build a large image database, to describe input images et finally to evaluate the IPC output. By testing this approach on an IPC dedicated to road obstacle detection, we demonstrate that this experimental methodology and software architecture ensures a steady efficiency. The reason is simple : the IMC is globally optimized, from a large number of real images et with adaptive processing on input data.*

1. Traitements adaptatifs en vision

La conception d'une application de traitement d'images fait intervenir une succession d'opérateurs bas et moyen niveau (filtrage, extraction et chaînage de contours, détection de coins, segmentation en régions...) dont le but est d'extraire des informations pertinentes pour l'aide à la décision (reconnaissance, classification, inspection ...). A chaque étape de traitement, les paramètres de réglage ont une influence notable sur le comportement des algorithmes et sur la qualité finale des résultats. La puissance de calcul ayant augmenté considérablement en quelques décennies, la vision se tourne vers des applications comme la vidéo surveillance ou la conduite automobile où le contenu de la scène reste incontrôlé, très versatile et rapidement variable. Se pose alors le problème du réglage automatique de la chaîne de traitement d'images (CTI), car il faut en permanence assurer la qualité de la détection bas niveau, garante de la robustesse des tâches de haut niveau.

Une première difficulté est l'évaluation des traitements dont le caractère adaptatif est indissociable. Depuis quelques années, la communauté scientifique en a pris conscience et a proposé des solutions plus ou moins empiriques [1-7]. Lorsqu'une vérité terrain est disponible, on peut directement y comparer les résultats en définissant une métrique pour mesurer les écarts avec cette référence. Parfois

cette vérité terrain n'est pas disponible ou reste controversée et on préfère se baser sur l'avis d'experts proches de l'application pour une simple appréciation visuelle ou sur des critères quantitatifs mais empiriques. Ces méthodes considèrent un seul type d'opérateur à la fois [8-11]; or le réglage isolé d'un opérateur ne conduit pas forcément à un réglage optimal de la chaîne complète. De plus, les opérateurs sont généralement testés sur un nombre limité d'images test, voire sur des images de synthèse bruitées pour tester la robustesse des algorithmes. Ceci ne remplace pas le recours à une base d'images réelles conséquente sur laquelle la CTI sera testée. Comment évaluer sur un grand nombre d'images des traitements comprenant de nombreux paramètres ?

Il subsiste une deuxième difficulté : comment trouver les bons réglages et donc comment adapter les traitements pour maintenir une qualité constante des résultats ? Les traitements temps réel étant implémentés sur des composants électroniques, il faut déjà prévoir la reconfiguration de leurs paramètres. Les rétines artificielles et les caméras intelligentes permettent déjà de régler certains paramètres d'acquisition. Concernant les traitements, le volume de calcul nécessaire pour apprécier l'effet de la modification de plusieurs paramètres sur les résultats semble d'emblée dissuasif, d'autant plus que chaque image conduit à des paramètres différents. Notons que le choix des opérateurs reste ici du ressort de l'expérimentateur mais que quelques

travaux examinent aussi la possibilité de l'automatiser [12-13]].

Nous montrons ici comment on peut surmonter ces deux difficultés grâce à une approche expérimentale combinant la modélisation statistique, l'optimisation et l'apprentissage. Nous illustrons ensuite cette démarche dans le cas d'une CTI dédiée à l'extraction de lignes pour la détection d'obstacles routiers.

2. Architecture

Notre article étant sans doute rédigé en français, il est préférable d'utiliser les règles de typographie françaises. La ponctuation haute (: ; ? !) doit être précédée d'un espace, comme sur une machine à écrire. Utiliser de préférence les guillemets à la française, « comme ceci ». Enfin, les majuscules doivent être accentuées. Elle comprend les modules suivants (Fig.1) :

- une CTI composée de plusieurs opérateurs bas et moyen niveau disposant chacun de paramètres re-configurables.
- une base d'images brutes, propres à l'application et représentatives de sa versatilité.
- un module d'évaluation des sorties mesurant la qualité des images traitées. L'évaluation peut s'appuyer sur une vérité terrain ou sur des critères empiriques.
- un module d'évaluation des entrées permettant au module de contrôle d'adapter le réglage des paramètres de traitement d'image à chacune des images brutes. L'évaluation a pour but d'extraire de l'image brute des descripteurs pertinents pour le réglage de la CTI.

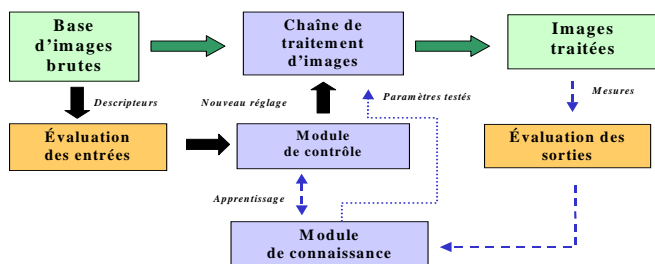


FIG.1 : architecture d'une chaîne de traitement adaptative

- un module de connaissances permettant d'une part, de modéliser l'influence des paramètres de traitement sur les images obtenues en sortie de la CTI et d'autre part de trouver un réglage optimal de ces paramètres pour une image donnée.
- un module de contrôle permettant après une phase d'apprentissage et à partir de l'évaluation des entrées de fournir pour chaque image brute des paramètres de traitement adaptés.

La modélisation de l'influence des paramètres est confiée à des plans d'expériences [14]. Cet outil couramment utilisé dans l'industrie commence à être expérimenté en vision [15-16]. Il consiste à modéliser en un nombre minimal d'essais les effets de modifications simultanées des paramètres de la chaîne. Le volume de calculs reste considérable car ces mêmes essais doivent être répétés sur un grand nombre d'images test afin de leur donner une valeur statistique. On n'obtient donc pas un réglage pour chaque image testée mais

un réglage moyen. On isole les paramètres les plus influents sur la qualité des résultats et on détecte les interactions entre ces paramètres. Pour chaque image test de la base, il faut ensuite déterminer quel est le meilleur réglage. C'est une procédure typique d'optimisation : le réglage moyen et les modalités hautes et basses des paramètres influents initialisent et délimitent le domaine d'exploration d'un l'algorithme d'optimisation du simplexe [17] qui recherche ces valeurs optimales. Une fois obtenu les réglages optimaux de la CTI pour un ensemble d'images brutes, on utilise les descripteurs d'entrée de ces images et les réglages correspondants à ces images test, comme base d'apprentissage pour un réseau de neurones. Une partie des images brutes est réservée au test du réseau. Après les étapes précédentes de modélisation statistique, d'optimisation et d'apprentissage, la CTI bascule en mode opérationnel et adapte en permanence ses paramètres de traitement aux caractéristiques des nouvelles images.

3. Application à une chaîne de traitement d'images routières

Cette application du programme PREDIT s'intègre dans le projet SPINE (Sécurité Passive INtelligentE) destiné à configurer un système d'airbags intelligent en situation de pré-crash. Le rôle du système multi capteurs (caméra haut débit EEV + télémètre laser SICK) embarqué sur un véhicule expérimental PEUGEOT 406 est de classifier les obstacles frontaux potentiels et d'estimer leur trajectoire en moins de 100 ms [18-20]. C'est pourquoi les traitement bas et moyens niveaux sont implémentés matériellement par la société MBDA, partenaire du projet, sur deux ASICs [21] intégrés avec un DSP sur une carte électronique interfacée avec le bus CAN du véhicule (Fig.2). Les premiers essais chez l'équipementier automobile FAURECIA ayant démontré qu'un réglage fixe était inopérant face à la diversité des images routières, un réglage automatique basé sur l'approche décrite plus haut a été adopté.

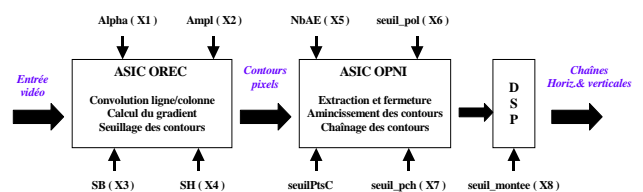


FIG.2 : Chaîne de traitement et paramètres réglables

Les huit paramètres re-configurables sont : les coefficients du filtre de Canny-Deriche (X_1) et d'amplification de l'image (X_2), les seuils bas et haut et bas des contours (X_3 , X_4), le nombre d'automates élémentaires de fermeture des contours (X_5), le seuil d'approximation polygonale (X_6), le seuil d'élimination des petites chaînes (X_7) et le seuil d'approximation des chaînes horizontales et verticales (X_8) (Fig. 2). Avec deux modalités de réglage (Tab.1), on peut appliquer 256 traitements différents sur la même image.

TAB .1 modalités pour les plans d'expérience

Facteur	Paramètre	Modalités	
X ₁	alpha	0.5	1
X ₂	ampl	33	63
X ₃	SB	5	15
X ₄	SH	15	30
X ₅	NbAE	26	30
X ₆	seuil_pol	5	6
X ₇	seuil_pch	5	10
X ₈	seuilMontee	1	3

La CTI doit extraire de l'image les familles de lignes horizontales et verticales (Fig.3) décrivant après groupement perceptuel les obstacles routiers situés à l'avant du véhicule test. L'évaluation des sorties est alors basée sur le nombre, la répartition et la longueur de ces segments dans une fenêtre d'intérêt spécifiée par un télémètre laser. Le critère utilisé, appelé taux de recouvrement, peut être mesuré pour différents réglages des paramètres (Fig.4)

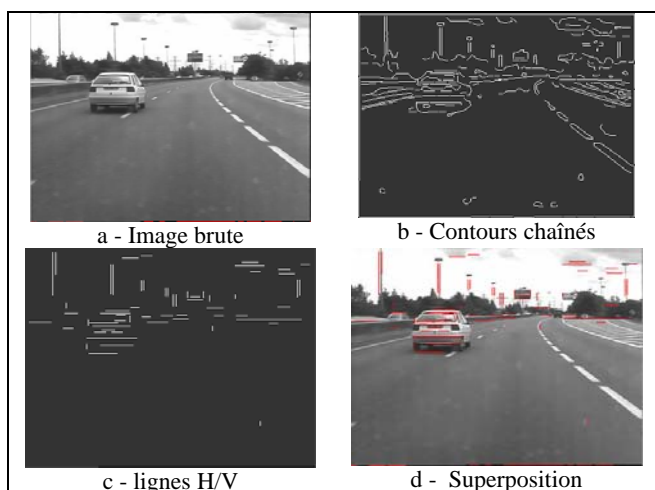


FIG. 3 : extraction des lignes H/V



FIG. 4 : évaluation des sorties

Au sein du module de connaissances, trois plans d'expérience ont été implémentés : Un plan factoriel fractionnaire 2^{k-p} à 16 essais [22] (Tab.2) pour déterminer les

paramètres vraiment influents (X₁, X₆, X₈), un plan de Rechscaffner à 37 essais [23] pour identifier les interactions des paramètres (X₁X₂ et X₆X₈) et enfin un plan quadratique ajoutant une modalité intermédiaire afin de détecter des non-linéarités. Ces plans ont été testés sur 180 images sélectionnées dans une séquence vidéo de 30 000 images urbaines et autoroutières. On en déduit un modèle statistique validé par la mesure du R-Square et du C(p) de Mallow ($Y = 40.2 + 2.06 X_1 + 0.74 X_2 - 2.47 X_6 + 5.30 X_8 - 0.92 X_1 X_2 + 0.95 X_6 X_8$). Il fournit un réglage moyen des paramètres qui initialise l'algorithme d'optimisation de Nelder & Mead utilisant le simplexe. Celui-ci détermine les réglages optimaux des paramètres pour l'ensemble des images testées.

TAB.2 plan d'expérience Factoriel 2^{k-p}

Essai	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	r (%)
1	-1	-1	-1	-1	-1	-1	-1	-1	35.535
2	-1	-1	-1	1	1	1	-1	1	40.310
3	-1	-1	1	-1	1	1	1	-1	27.859
4	-1	-1	1	1	-1	-1	1	1	42.436
5	-1	1	-1	-1	1	-1	1	1	47.328
6	-1	1	-1	1	-1	1	1	-1	30.284
7	-1	1	1	-1	-1	1	-1	1	44.034
8	-1	1	1	1	1	-1	-1	-1	37.743
9	1	-1	-1	-1	-1	1	1	1	46.517
10	1	-1	-1	1	1	-1	1	-1	40.469
11	1	-1	1	-1	1	-1	-1	1	50.680
12	1	-1	1	1	-1	1	-1	-1	33.464
13	1	1	-1	-1	1	1	-1	-1	35.169
14	1	1	-1	1	-1	-1	-1	1	49.255
15	1	1	1	-1	-1	-1	1	-1	39.715
16	1	1	1	1	1	1	1	1	44.842

Pour réaliser l'apprentissage du module de contrôle, il faut auparavant choisir des descripteurs d'entrée pour caractériser les images. L'histogramme d'homogénéité [24] de l'image brute a été utilisé pour rendre compte à la fois de zones de teinte unie (peinture des véhicules) ou de texture homogène (revêtement bitumé de la route). Le module de contrôle utilisé est un simple perceptron multicouches. Il comporte 256 neurones d'entrée associés aux niveaux de l'histogramme d'homogénéité de l'image, 48 neurones cachés correspondant à la vitesse de convergence maximale du réseau et 3 neurones de sortie pour les valeurs calculées des 3 paramètres influents de réglage. Dans (FIG.5) sont indiquées en % les valeurs du taux de recouvrement (évaluation des sorties) moyenné sur l'ensemble des images test selon la technique de réglage adoptée ainsi que le coût CPU correspondant.

En l'absence de tout réglage adaptatif (réglage fixe moyen fourni par les plans d'expériences), les résultats sont médiocres ; si on applique le réglage issu du meilleur essai d'un des plans d'expériences, les résultats sont sensiblement meilleurs mais cette méthode est inapplicable en temps réel ; les résultats optimaux sont obtenus par le simplexe (SPL) au prix du temps prohibitif d'exploration de l'espace des paramètres ; enfin les réseaux (RN) fournissent des résultats remarquables, surtout celui à 3 sorties, au prix d'un temps de calcul négligeable (\approx temps de calcul des descripteurs

d'entrée). On a mentionné volontairement dans ce tableau les résultats obtenus avec un réglage à 8 paramètres : on peut vérifier qu'il est inutile de régler les 5 paramètres jugés peu influents par les plans d'expérience.

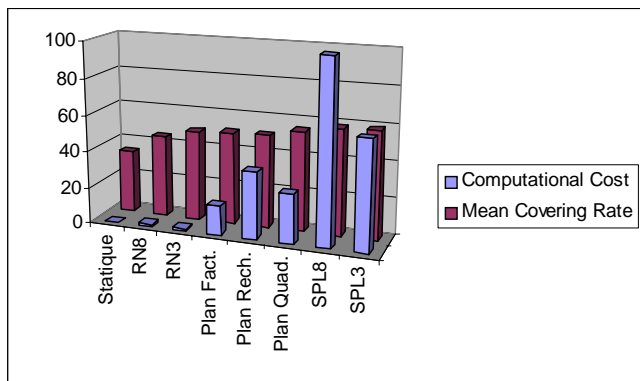


FIG.5 : comparaison de différentes techniques de réglage

Conclusion Ces résultats intéressants confirment l'intérêt de l'approche expérimentale pour le réglage adaptatif de chaînes de traitement d'images. Il s'agit maintenant d'en vérifier d'une part la portée sur d'autres applications où la typologie des images, les opérateurs de la CTI et les critères d'évaluation des données en entrée et en sortie seront à nouveau spécifiques et d'en unifier d'autre part la méthodologie.

Références

- [1] R.M. Haralick, "Performance characterization protocol in computer vision", ARPA Image Understanding Workshop, Monterey, CA, 667-673, 1994.
- [2] P.Courtney, N.Thacker, A. Clark "Algorithmic modeling for performance evaluation" Workshop on performance characteristics of vision algorithms" Cambridge, April 19 1996- 13p.
- [3] W. Forstner, "10 Pros and cons against performance characterization of vision algorithms", in Workshop on Performance Characteristics of Vision Algorithms, April 1996.
- [4] Kevin W. Bowyer, P. Jonathon Phillips "Empirical Evaluation Techniques in Computer Vision" June 1998, Wiley-IEEE Computer Society Press ISBN: 0-8186-8401-1 262 pages
- [5] P. Meer, B. Matei, K. Cho, "Input Guided Performance Evaluation", Theoretical Foundations of Computer Vision, pp. 115-124, 1998.
- [6] I.T. Phillips and A.K. Chhabra, "Empirical Performance Evaluation of Graphics Recognition Systems", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, pp. 849-870, 1999.
- [7] J.Blanc-Talon, V.Roport "Evaluation des chaînes de traitement d'images" Revue Scientifique et Technique de la Défense n°46 2000 p.29-38
- [8] S. Philipp-Foliguet, "Evaluation de la segmentation", ETIS, Cergy-Pontoise, Mars 2001.
- [9] N.Sebe,Q.Tian, E.Loupias, M.Lew, T.Huang "Evaluation of salient point techniques" CIVR 02 July 10-15 2002 London
- [10] P.Rosin, E.Ioannidis "Evaluation of global image thresholding for change detection" Pattern Recognition Letters 24 (2003) 2345-2356
- [11] Y.Yitzhaky, E.Peli "A method for objective edge detection evaluation and detection parameter selection" IEEE PAMI vol.25 n°8 Aug. 2003 p.1027-1033
- [12] V. Ropert, "Proposition d'une architecture de contrôle pour un système de vision", Thèse de l'Université René Descartes (Paris 6), Décembre 2001.
- [13] I.Levner, V.Bulitko "Machine learning for adaptive image Interpretation" The Sixteenth Innovative Applications of Artificial Intelligence Conference (IAAI-04) July 27-29, 2004 San Jose, California USA- 7p.
- [14] P. Schimmerling, J-C. Sisson, A. Zaïdi, "Pratique des Plans d'Expériences", Lavoisier Tec & Doc, ISBN 2-743-00239-5, 1998.
- [15] J.J.Rousselle, N.Vincent,"Design of experiments to set active contours" Quality control by Artificial Vision" Le Creusot France 19-22 May 2003.
- [16] S. Treuillet, "Analyse de l'influence des paramètres d'une chaîne de traitements d'images par un plan d'expériences", 19^e colloque GRETSI'03 sur le traitement du signal et des images, 8-11 sept. 2003.
- [17] Margaret H. Wright The Nelder-Mead Simplex Method: Recent Theory and Practice International Symposium on Mathematical Programming Lausanne, EPFL, August 24-29, 1997
- [18] A.Domingues, Y. Lucas , D. Baudrier, P. Marché, "Détection et suivi d'objets en temps réel par un système embarqué multi capteurs", GRETSI'01, Toulouse, Septembre 2001.
- [19] A. Domingues, "Système embarqué multicapteurs pour la détection d'obstacles routiers - Développement du prototype et réglage automatique de la chaîne de traitement d'images", Thèse de l'Université d'Orléans, 15 Juillet 2004.
- [20] Y. Lucas, A. Domingues, M. Boubal, P. Marché, "Système de vision embarqué pour la détection d'obstacles routiers " Techniques de l'Ingénieur - Recherche & Innovation. 02/2005.9 p.
- [21] P. Lamaty, "Opérateurs de niveau intermédiaire pour le traitement temps réel des images", Thèse de Doctorat, 2000.
- [22] A. Fries, J. Hunter, "Minimum aberration 2^{k-p} designs", Technometrics , vol. 22, pp. 601-608, 1980.
- [23] R. L. Rechtschaffner, "Saturated fractions of $2n$ and $3n$ factorial designs", Technometrics, 9, pp. 569-575, 1967.
- [24] H. Cheng, Y. Sun, "A hierarchical approach to color image segmentation using homogeneity", in IEEE Transactions on Image Processing, 9(12) : 2071-2082, 2000.