

Détection de signalisation routière par ajustement de formes prototypes

Guillaume DUTILLEUX, Pierre CHARBONNIER
Laboratoire des Ponts et Chaussées de Strasbourg
11 rue Jean Mentelin, BP 9, 67035 STRASBOURG Cedex, FRANCE
guillaume.dutilleux@equipement.gouv.fr
pierre.charbonnier@equipement.gouv.fr

Résumé – Nous présentons une approche originale en matière de détection automatique robuste de signalisations verticales de danger en lumière naturelle. Le détecteur s'appuie sur la combinaison de formes prototypes et d'un algorithme d'optimisation globale de type stratégie d'évolution hybridé avec une méthode locale de type gradient conjugué. L'étape de pré-détection fait appel à un calcul de dominante chromatique robuste face aux éclairagements incontrôlés. L'emploi d'une forme prototype adaptée au problème permet de s'affranchir du bruit. La fonction minimisée est la somme de termes d'*a priori* et de deux termes d'attache aux données. Le réglage du seuil de détection et l'évaluation globale du détecteur s'appuient sur le tracé de courbes COR.

Abstract – We present an original approach regarding the robust automatic detection of vertical danger road signs under natural light. The detector is based on the combination of template shapes and an optimization algorithm called evolution strategy hybridized with a local search method of the conjugate gradient kind. The pre-detection step uses a computation of chromatic dominance robust to uncontrolled illumination conditions. Using a template shape helps reducing the sensitivity to noise. The cost function is the sum of *a priori* terms and of two likelihood terms. The adjustment of the detection threshold and the global performance evaluation of the detector are founded on ROC curves.

1 Introduction

La détection et la reconnaissance automatiques de panneaux de signalisation font l'objet d'un intérêt soutenu depuis plusieurs années (voir par exemple [4] pour une bibliographie récente), en relation avec l'alimentation de bases de données d'images routières que les gestionnaires d'infrastructures utilisent en particulier pour l'inventaire de signalisation et les études de sécurité. Les systèmes d'aide à la conduite sont un autre domaine d'application.

Cet article traite de la conception d'un système de détection hors-ligne qui fonctionne sur des séquences d'images enregistrées en lumière naturelle visible. On traite ici le cas des panneaux de danger. Le système de détection s'accommode bien d'une illumination incontrôlée ou d'images où la balance des blancs est incorrecte. Le système est aussi conçu pour être robuste vis-à-vis des occultations et des signaux endommagés. L'approche choisie est basée sur la combinaison de formes prototypes [1][2] et d'une stratégie d'évolution (ES), qui représente une alternative intéressante aux algorithmes génétiques [5][6][7]. Cet algorithme d'optimisation globale est *hybridé* avec une méthode de recherche locale de type gradient, dans le but d'obtenir une convergence fine. Dans ce qui suit, les caractéristiques principales du détecteur sont d'abord présentées. Les aspects de traitement d'image sont d'abord décrits. Ils concernent la pré-détection robuste de la composante rouge, la définition d'une forme prototype adaptée au problème et du critère optimisé. Ensuite, les caractéristiques de l'algorithme d'optimisation choisi sont explicitées. La dernière partie fournit des résultats pour des images réelles et illustre la fiabilité du détecteur global.

2 Le système de détection

2.1 Traitement d'images

2.1.1 Pré-détection

Pour identifier les pixels présentant une dominante rouge, on travaille dans l'espace RGB. Le pré-détecteur considère qu'un pixel est rouge si ses composantes vérifient

$$\begin{aligned} R &> \alpha(G + B) \\ R - \max(G, B) &> 2\alpha[\max(G, B) - \min(G, B)] \end{aligned} \quad (1)$$

Le premier critère revient à poser une contrainte sur la composante rouge normalisée, $R/(R+G+B)$, qui doit être dominante. La normalisation donne sa robustesse au pré-détecteur face aux variations d'illumination. Le deuxième critère vérifie que le pixel ne tende ni vers le jaune, ni vers le magenta. Le paramètre α peut être ajusté globalement pour une séquence d'images. Ses valeurs habituelles varient de 0.5 à 0.75. On privilégie des valeurs de α relativement faibles pour ne pas risquer de voir un panneau disparaître lors de la pré-détection. En contrepartie, le bruit dans l'image de pré-détection obtenue est plus important, mais l'utilisation d'une forme prototype adaptée (Cf. 2.1.2) permet de s'affranchir de ce problème. Enfin, une balance des blancs incorrecte est malheureusement assez fréquente dans les séquences à traiter. Comme les images analysées contiennent toujours une portion de route, une égalisation de type « monde gris » [3] peut être effectuée automatiquement dans la plupart des situations.

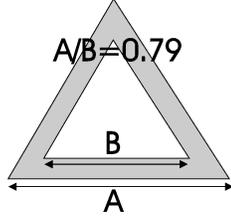


FIG. 1 – Forme prototype

2.1.2 Forme prototype

Pour la détection de panneaux de danger, une forme prototype particulière est utilisée (Cf. figure 1), dans le but de minimiser les fausses détections. Les deux triangles emboîtés définissent une couronne où l'on peut rechercher des pixels rouges.

La forme prototype est ajustée sur l'image à l'aide de trois transformations :

- mise à l'échelle,
- rotation,
- translations horizontale et verticale.

Ceci définit un espace de recherche à 4 dimensions. Un jeu de paramètres de transformation est représenté par la variable Θ dans la suite.

2.1.3 Fonction d'énergie

La fonction d'énergie est définie classiquement comme une somme de termes d'*a priori* et de termes de vraisemblance. Le calcul de l'*a priori* sur la forme obtenue par transformation concerne les points suivants :

- la position usuelle des panneaux dans l'image (dans la moitié droite),
- la gamme de taille possible pour un panneau,
- l'orientation usuelle d'un panneau.

Chaque critère sur un paramètre x supposé se trouver dans un intervalle de pertinence $[x_{min}, x_{max}]$ est représenté par une fonction de pénalité en forme de vallée :

$$\begin{aligned} x < x_{min}, g(x) &= a(x_{min} - x) \\ x_{min} \leq x \leq x_{max}, g(x) &= 0 \\ x > x_{max}, g(x) &= b(x - x_{max}) \end{aligned} \quad (2)$$

Les bornes des différents intervalles de pertinence peuvent être considérées comme des données établies par la « vérité terrain ». Les différentes pentes a et b ont été réglées empiriquement.

Le terme global d'*a priori* est une somme pondérée des différents termes de pénalité.

La vraisemblance se compose de deux termes : un terme orienté contour et un terme orienté région. Le terme orienté contour est largement inspiré de [1]. Il dépend de l'angle entre le gradient local de l'image et le vecteur directeur local de la forme prototype. Sa valeur approche 0 quand le modèle est aligné sur les contours d'un objet d'intérêt et 1 quand le modèle est sur un domaine homogène. Ce terme est calculé le long du triangle extérieur. Pour une forme Θ , le terme orienté région est calculé sur la carte binaire de pré-détection de dominante rouge. Il se compose de deux termes reliés aux deux zones du modèle choisi : la

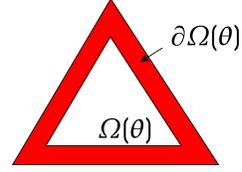


FIG. 2 – Domaines d'intégration

couronne $\partial\Omega$ et l'intérieur du petit triangle Ω (voir figure 2) :

$$U_c(\Theta) = \frac{\int \int_{\partial\Omega(\Theta)} I(x) dx}{\int \int_{\partial\Omega(\Theta)} dx} \quad \text{et} \quad U_i(\Theta) = \frac{\int \int_{\Omega(\Theta)} I(x) dx}{\int \int_{\Omega(\Theta)} dx} \quad (3)$$

Le terme global orienté région s'écrit :

$$U_t(\Theta) = \min(1, 1 + U_c(\Theta) - U_i(\Theta)) \quad (4)$$

Quand Θ est parfaitement ajusté, U_t vaut zéro. Au contraire, $U_t(\Theta)$ est égal à 1 lorsque Θ est sur une zone uniforme « rouge » ou « non rouge » de l'image.

2.2 Optimisation

2.2.1 Optimisation globale par stratégie d'évolution

La fonction d'énergie globale obtenue en sommant les termes d'*a priori* et de vraisemblance est hautement multimodale. De ce fait, un algorithme d'optimisation globale est nécessaire. Ici, c'est une stratégie d'évolution (ES) qui a été choisie [5]. Cet algorithme appartient à la classe des algorithmes évolutionnaires, au même titre que les algorithmes génétiques (GA) [7]. Comme les GA, les ES manipulent une population de solutions candidates, et utilisent des opérateurs de croisement, mutation et sélection. A la différence des GA, les ES manipulent des vecteurs aléatoires gaussiens plutôt que des chaînes binaires. De ce fait, elles ne nécessitent pas de codage/décodage des paramètres à optimiser. Ce fondement probabiliste permet à la théorie de la convergence des ES d'être mieux établie que celle des GA [6]. Un autre point intéressant par rapport aux GA est que les ES incorporent naturellement un mécanisme d'auto-adaptation des paramètres internes, que l'on peut rapprocher du schéma de refroidissement du recuit simulé [8]. Ainsi la probabilité de mutations de grande amplitude décroît au fil des itérations.

Pour le problème d'optimisation discuté ici, une configuration à 4 parents et 7 fois plus de descendants à été retenue. Comme dans [4], la population initiale est basée sur une analyse des composantes connexes de l'image de dominante rouge. Lorsque le nombre de composantes connexes est insuffisant, les individus sont obtenus par tirage de variables aléatoires uniformes centrées et mises à l'échelle relativement à la définition des termes d'*a priori*. En remplacement de la sélection déterministe, un opérateur mixte roulette/déterministe est préféré ici, de façon à réduire la probabilité de convergence prématurée.

2.2.2 Recherche locale par gradient conjugué

En pratique, la convergence est habituellement très rapide au début de l'optimisation. Plus tard, les solutions candidates ont souvent tendance à tourner autour de la solution sans véritablement la trouver. De façon à résoudre ce problème, l'ES a été hybridée ici avec une méthode locale, de type gradient conjugué. Ceci permet d'exploiter la topologie de la fonction optimisée au voisinage de la solution, et donc à l'optimisation de converger plus vite et plus finement.

Du fait de la complexité des expressions des termes d'attache aux données, le gradient est calculé numériquement, par différences finies. En outre, comme les termes d'*a priori* ne sont pas différentiables sur la totalité de leur domaine de définition, le gradient est calculé sur le terme de vraisemblance uniquement. Ceci n'est pas préjudiciable, étant donné que la recherche locale présente surtout un intérêt au voisinage de la solution, donc dans une zone où l'*a priori* est nul.

3 Résultats

La figure 3 montre une image où le panneau est à peine visible du fait d'une situation de fort contre-jour, ainsi que de la présence de végétation et d'une levée de terrain sur le bas-côté. La figure 4 montre l'image de pré-détection de rouge, où le panneau est bien visible. Avec une telle image, la détection automatique ne pose pas de difficulté.

En guise d'illustration du fonctionnement global de l'algorithme, la figure 5 montre une image originale, la figure 6 l'image de pré-détection de rouge associée, la figure 7 le résultat de l'optimisation. Ici, l'image de pré-détection contient de nombreux leurres potentiels pour le détecteur. L'utilisation d'une forme prototype spécifique permet de les éviter.

Enfin, nous montrons figure 8 une comparaison des courbes COR (Caractéristique Opérationnelle du Récepteur) obtenues pour différents scénarios d'utilisation de l'algorithme, à partir des résultats de 2 analyses d'une séquence de 52 images, dont 32 contenant des panneaux. Pour le calcul de ces courbes, une vérité-terrain concernant la position des objets d'intérêt a été obtenue par analyse manuelle de la séquence d'images. A la fin de l'optimisation, soit 500 itérations de l'ES, on considère la valeur résiduelle de la fonction d'énergie correspondant au prototype optimal : on considère qu'il y a eu détection d'un panneau lorsque celle-ci est inférieure à un certain seuil. Si la position de l'objet est dans la zone de tolérance définie autour de la vérité-terrain (qui peut être vide s'il n'y a effectivement pas de panneau dans l'image), il s'agit d'une bonne détection, sinon, il s'agit d'une fausse alarme. Pour cette évaluation, les coordonnées des centres des formes sont arrondies au pixel près et la zone de tolérance a un rayon de 2 pixels. La probabilité de bonne détection est calculée comme le rapport du nombre de détections valides sur le nombre de panneaux présents dans la séquence. La probabilité de fausse alarme est le rapport du nombre de détections erronées sur le nombre total de fausses détections possibles. Ces probabilités sont représentées sous formes

de courbes, paramétrées par la valeur du seuil de détection. On utilise ici une échelle semi-logarithmique pour améliorer la lisibilité des courbes COR. La courbe idéale correspondrait au cas de 100% de bonnes détections pour 0% de fausses alarmes. Les courbes montrent la supériorité d'une initialisation tenant compte des composantes connexes issues de l'étape de pré-détection du rouge, par rapport à une initialisation purement aléatoire. D'autre part, elles illustrent l'intérêt de l'approche hybride ES-gradient conjugué par rapport à l'algorithme ES classique.

4 Perspectives

Un détecteur de panneaux de signalisation à partir de prises de vues en lumière naturelle visible a été présenté. Il se montre peu sensible aux conditions d'illumination grâce à un pré-détecteur de dominante rouge robuste. L'emploi d'une stratégie d'évolution hybridée avec une méthode locale permet d'affiner la convergence. L'application du détecteur sur des séries de plusieurs dizaines d'images montre des résultats prometteurs en termes de robustesse.

Le développement du détecteur se porte actuellement sur le traitement efficace du cas général de la détection de signalisation : identification précoce de l'absence de panneau, traitement des images contenant plusieurs panneaux, en vue de l'utilisation du détecteur en phase opérationnelle.

Références

- [1] M.P. Dubuisson Jolly, S. Lakshmanan et A.K. Jain, "Vehicle segmentation and classification using deformable templates," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 18, N°3, pp. 293-308, mars 1996.
- [2] M. Mignotte, C. Collet, P. Perez et P. Bouthemy, "Hybrid genetic optimization and statistical model-based approach for the classification of shadow shapes in sonar imagery," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, Vol. 22, N°2, pp. 129-141, février 2000.
- [3] G. Buchsbaum, "A Spatial Processor Model for Object Colour Perception", *Journal of the Franklin Institute*, 310 :1-26, 1980.
- [4] A. de la Escalera, J.M. Armingol, J.M. Pastor, et F.J. Rodrigo, "Visual sign information extraction and identification by deformable models for intelligent vehicles," *IEEE Trans. on Intelligent Transportation Systems*, Vol. 5, N°2, pp. 57-68, 2 juin 2004.
- [5] I. Rechenberg, *Evolutionsstrategie '94* Fromann-Holzboog, 1994 (en allemand).
- [6] T. Bäck, *Evolutionary algorithms in theory and practice*, Oxford university Press, 1996.
- [7] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer, 1999.
- [8] J. Dréo, A. Pérowski, P. Siarry, E. Taillard, *Métaheuristiques pour l'optimisation difficile*, Eyrolles, 2003.



FIG. 3 – Image originale

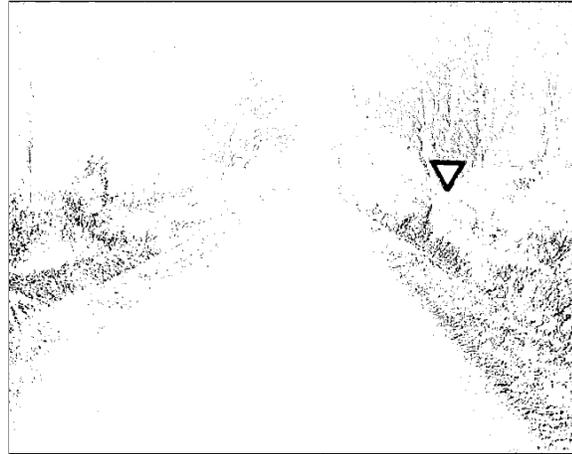


FIG. 4 – Pré-détection de rouge



FIG. 5 – Image originale

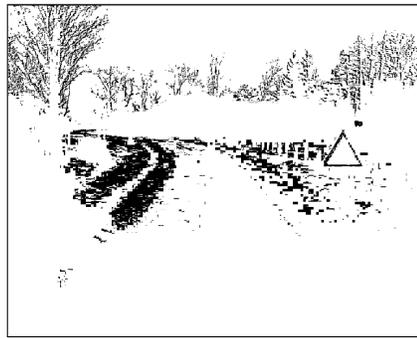


FIG. 6 – Pré-détection de rouge

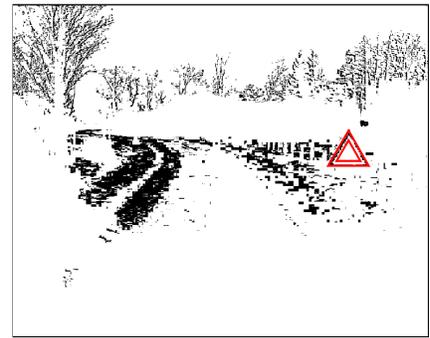


FIG. 7 – Résultat de l'optimisation

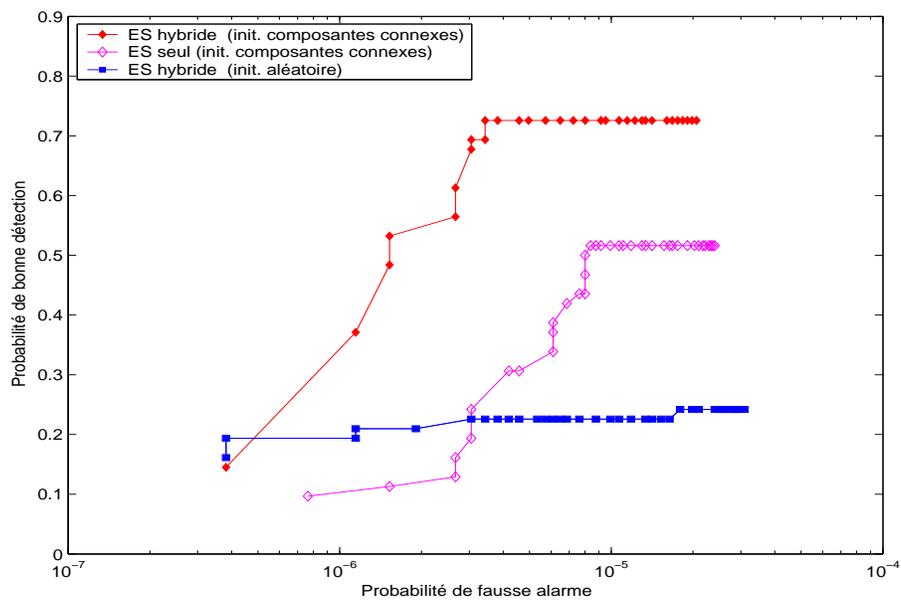


FIG. 8 – Évaluation de l'algorithme : courbes COR