

# Détection robuste de véhicules par imagerie multi-dimensionnelle de polarisation

WANG FAN<sup>1</sup>, SAMIA AINOUS<sup>1</sup>, FABRICE MERIAUDEAU<sup>3</sup>, ABDELAZIZ BENSRAIR<sup>1</sup>

<sup>1</sup> Laboratoire d'Images de Traitement de l'Information et des systèmes, Normandie université  
685 Avenue de l'Université, BP 08, Saint Etienne Du Rouvray, France

<sup>3</sup> Université Technologique de Petronas, CISIR  
32610 Bandar Seri Iskandar Malaysia

<sup>1,2</sup> Fan\_wang@med.unc.edu,

<sup>1</sup> samia.ainouz@insa-rouen.fr, <sup>3</sup>fabrice.meriaudeau@utp.edu.my, <sup>1</sup>abdelaziz.bensrair@insa-rouen.fr

Résumé – L'analyse des scènes routières est une tâche primordiale pour les systèmes d'aide à la conduite (ADAS). Le véhicule est l'un des obstacles majeurs qu'il faut tenir au maximum à distance sur la route. Sa détection est nécessaire pour une conduite plus sûre. Les méthodes de détection classiques basées sur les images couleurs manquent d'informations physiques sur la scène et sont, de ce fait, instables face aux problématiques liées aux scènes routières notamment les fortes réflexions, les bruits d'acquisitions ou les distorsions radiométriques. L'imagerie polarimétrique, en revanche, est capable de caractériser un objet par sa structure et son contenu physique (structure géométrique de la surface, rugosité, réflexion,...) même dans des conditions d'illumination défavorables. Cette modalité d'images, riches en informations physiques peut parfaitement compléter les attributs colorimétriques ou géométriques extraits des images conventionnelles. Dans le but d'améliorer la robustesse de la détection des véhicules, nous proposons dans ce papier, un modèle de fusion des caractéristiques pertinentes issues des images polarimétriques et des images couleurs. La méthode proposée, basées sur les modèles à parties déformables (DPM, *Deformable Part based Models*) a été évaluée sur notre propre base de données et a montrée des résultats très prometteurs quant à l'apport de la polarimétrie pour l'analyse des scènes routières.

Abstract - Road scene understanding is a vital task for driving assistance systems. Robust vehicle detection is a precondition for obstacle avoidance and secure driving. Color images provide limited information concerning the physical properties of the object, resulting in unstable detection against the usual problems of road scenes like strong reflexions, noises or radiometric distortions. Polarization images, however, can robustly describe important properties of the object (the surface geometric structure, material and roughness, reflexion, etc), even under poor illumination or strong reflections. This modality gives reach physical information, which could perfectly complete the features, extracted from classical color images. In order to improve the robustness of the vehicle detection problem, we propose in this paper a fusion model using polarization information and color image attributes. Our method is based on a feature selection procedure to get the most informative polarization feature and color-based ones. The proposed method, based on the Deformable Part based Models (DPM), has been evaluated by our self-collected database, showing good performances and encouraging results about the use of the polarimetric modality for road scenes analysis.

## 1 Introduction

Pour les systèmes d'aide à la conduite, la voiture est souvent considérée comme l'objet le plus dangereux de la route. Quelle que soit sa position sur la route, une distance de sécurité doit toujours être maintenue pour éviter la collision. Due à sa large structure et à sa variation d'apparences (couleur, forme, point de vue) [6], la détection automatique de voitures dans une scène routière reste un défi important pour la navigation autonome. Le schéma de base, habituellement utilisé pour la détection des obstacles est le détecteur de Dallal et Trig [2] (DT), basé sur l'association d'un descripteur type HOG (*Histogramm of Oriented Gradient*) et du détecteur SVM (*Support Vector Machine*) linéaire. En revanche, pour mieux intégrer l'objet et gérer ses variations intra-classe, Felzenszwalb et al. [5] ont développé le DPM où l'objet est représenté par

plusieurs parties déformables et où des détecteurs de type HOG sont appliqués dans chaque partie. La position de ces parties est latente dans l'image et donc à optimiser. C'est pour cette raison qu'ils ont associé à ce descripteur un détecteur appelé SVM latent (LSVM).

Les méthodes de détection basées sur les images d'intensité sont complètement dépendantes des conditions d'illumination extérieures. Or, dans les scènes routières, ces conditions ne sont pas toujours commodes (fortes réflexions, météo dégradées) et font que les canaux RGB sont très limités en terme d'information et sont peu robustes face à ces conditions.

La polarisation de la lumière réfléchiée est fortement liée aux propriétés physiques de la surface. Après réflexion, l'onde lumineuse devient linéairement partiellement polarisée [4] et traduit de ce fait les caractéristiques intrinsèques de la surface considérée. Comparé aux

autres capteurs, un système imageur polarimétrique a l'avantage de caractériser l'aspect géométrique de la surface (orientation, angle de réflexion), d'identifier la nature de chaque matériau de la scène (indice de réfraction, type de matériaux) et de fournir des attributs liés aux propriétés physiques et optiques de la surface (absorption, réflexion, dépolarisation,...). La polarimétrie est capable de fournir des caractéristiques supplémentaires sur l'objet autre que son intensité.

Nous proposons dans ce papier, d'utiliser les images polarimétriques comme source d'information complémentaire pour la détection de véhicules. A notre connaissance aucun travail n'a tenté d'exploiter cette richesse informationnelle pour l'analyse des scènes routières. Un processus de sélection de caractéristiques est proposé afin d'en choisir les plus pertinentes. Deux DPM sont appris séparément, en utilisant respectivement les caractéristiques sélectionnées des images polarimétriques et celles des images couleur. Un schéma de fusion est proposé afin de combiner les résultats des deux détecteurs pour une détection plus robuste. La méthode proposée est évaluée sur notre base de données et a montré des résultats concluants.

## 2 Principe d'acquisition polarimétrique

La lumière réfléchiée par un objet est en général partiellement linéairement polarisée [4]. Elle est décrite par un vecteur mesurable ayant trois composantes  $S = [S_0, S_1, S_2]$ , appelé vecteur de Stokes. Le premier paramètre est lié à l'intensité de l'image, les autres traduisent des informations physiques sur l'interaction matière/lumière telles que la réflexion, l'absorption ou la dépolarisation. La mesure de ces trois composantes pour chaque pixel passe par l'acquisition des images polarimétriques. Un polariseur linéaire en rotation autour d'un angle  $\alpha$  est placé en face d'une caméra. Pour chaque angle de rotation  $\alpha_i$  une image d'intensité  $I(\alpha_i)$  est mesurée par la caméra. La relation entre les paramètres de Stokes et l'intensité  $I(\alpha_i)$  est donnée par [1]:

$$I(\alpha_i) = \frac{1}{2}(1 + \cos(2\alpha_i) \sin(2\alpha_i)) \cdot (S_0 \ S_1 \ S_2)^t \quad (1)$$

La matrice reliant  $I$  et  $S$  est la matrice optique caractérisant le polariseur linéaire utilisée [1]. Pour chaque pixel, la mesure des trois paramètres de Stokes, inconnus, requiert au moins trois acquisitions.

Le vecteur de Stokes donne accès à deux autres entités physiques très importantes : l'angle de polarisation (AOP) et le degré de polarisation (DOP). Ces deux quantités mesurent respectivement l'orientation et la quantité de lumière totalement polarisée dans une onde après réflexion :

$$AOP = \frac{1}{2} \tan^{-1} \frac{S_2}{S_1}, \quad DOP = \frac{\sqrt{S_1^2 + S_2^2}}{S_0} \quad (2)$$

## 3 Extraction et sélection des caractéristiques

### 3.1 Caractéristiques polarimétriques

Chaque pixel d'une image de polarisation est doté de 5 composantes  $S = [S_0, S_1, S_2, DOP, AOP]$ . Les conditions d'illumination extérieures sont non

contrôlées et font que ces images peuvent être bruitées. Certaines informations polarimétriques sont imprécises et difficiles à interpréter. Un processus de sélection des caractéristiques les plus pertinentes et significatives est par conséquent nécessaire afin d'en choisir la plus pertinente pour l'apprentissage du modèle de détection.

Similaire au HOG classique, les HOG polarimétriques (HOGP) sont basés sur des blocs et des cellules pour mieux englober la structure de l'objet considéré. L'image est divisée en blocs de taille 8x8, chaque bloc est divisé en quatre cellules 4x4.

Pour chaque cellule, un vecteur de caractéristiques  $[S_0, S_1, S_2, DOP, AOP]$  moyen (moyenne calculée sur tous les pixels de la cellule) est extrait. Chaque cellule est, de ce fait, dotée d'une caractéristique à 5 dimensions. Par concaténation de ces caractéristiques, à chaque bloc, composé de 4 cellules, est associée une caractéristique de 20 dimensions. Ces caractéristiques appelées  $f_{20-d}$  de tous les exemples sont utilisées pour apprendre le détecteur DT afin de produire un filtre  $w$  qui indique le poids de chaque caractéristique. Plus le poids est grand, mieux est l'importance de la caractéristique.

### 3.2 Caractéristique sélectionnée : l'AOP

Dans le but, de tester dans un premier temps l'apport de la polarimétrie à la détection des voitures, notre choix de caractéristique s'est porté sur celle qui présente le poids  $w$  le plus important. Le processus de sélection présenté dans [2] a montré que l'angle de polarisation est la caractéristique la plus pertinente pour le processus de détection de voitures. L'AOP dépend fortement de l'angle d'incidence, de l'orientation de la surface et du type du matériau de l'objet. Pour les surfaces rugueuses, l'AOP est distribué d'une manière irrégulière. A contrario, pour les surfaces lisses, il varie continuellement. La Figure.1 illustre ce constat. Dans cette figure, on distingue l'AOP de l'arbre qui est bruitée de celui du sol qui l'est moins et encore de celui de la voiture qui est régulier et qui varie d'une manière graduelle. La géométrie de la voiture est clairement distinguée même sur les zones clairsemées. La structure de la voiture est bien rehaussée.



Figure 1. Image originale (à gauche) et son AOP (à droite)

## 4 Modèle de détection

Le DPM est connu pour sa capacité à gérer les fortes variations intra-classe des objets [5]. Le modèle DPM est un modèle basé sur un filtre appliqué à une carte de caractéristiques dans le but d'avoir un score de détection. Une carte de caractéristiques est un tableau où chaque entrée est une caractéristique de dimension  $d$

calculée à partir de différentes positions (matérialisée par des fenêtres englobantes) dans l'image correspondante. Soit  $f$  la carte des caractéristiques avec  $f(x, y)$  l'entrée correspondante à chaque position  $(x, y)$ . Le filtre  $w$  est de même dimension que la carte des caractéristiques. Le score de détection est calculé comme le produit scalaire  $w \cdot f(x, y)$ . Le score de détection pour une position donnée est proportionnel à la possibilité de l'existence d'un objet à cette position (se référer à [5] pour d'avantage de détails sur les processus d'initialisation et d'optimisation de la méthode). Soit  $S_c = M(f(x, y))$  ce score, où  $M$  est le DPM considéré, un seuil  $t_r$  est appris de sorte à ce que si  $S_c \geq t_r$ , la boîte englobante (*Bounding Box*) centrée à la position  $(x, y)$  contient un positif et vice versa.

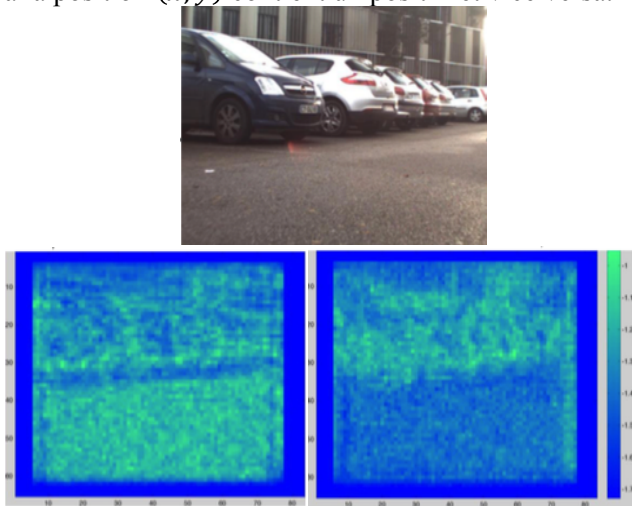


Figure 2. Cartes des scores par le DPM couleur (image du milieu) et le DPM appliqué sur l'AOP (à droite)

Du fait que la polarisation produit des informations supplémentaires et complémentaires avec celles des images d'intensité, la fusion des caractéristiques venant de ces deux modalités améliore considérablement les performances de la détection. Nous proposons un modèle de fusion qui combine les résultats du DPM appliqués séparément sur la modalité polarimétrique et sur les images couleur. Notons par  $f_p$  et  $f_c$  les cartes de caractéristiques de type HOG issues respectivement de l'image AOP et de l'image couleur RGB. De même soit  $M_p$  et  $M_c$  respectivement les DPM appliqués aux images polarimétriques et aux images couleur. Les scores obtenus de chaque DPM sont illustrés sur la Figure.2. Sur la carte produite par  $M_c$  les voitures et le sol ont un grand score ce qui complique leur distinction. En revanche, sur  $M_p$ , la zone contenant des voitures a un score nettement élevé par rapport au reste de la scène. Le modèle de fusion proposée utilise cette complémentarité entre les deux modalités et propose le score donné par  $M_p$  comme une confirmation du score obtenu par  $M_c$ . La décision finale est positive si les scores des deux modèles sont supérieurs aux seuils respectifs. Sinon, elle est considérée comme négative. Ce modèle de fusion est basé sur l'intersection (*AND fusion*) des deux détections fournies indépendamment par les deux modalités.

## 5 Résultats expérimentaux

L'état de l'art autour de l'imagerie polarimétrique ne contient aucune base de données de scènes routières. Nous avons acquis une base de données contenant 153 images polarimétriques de scènes routières<sup>1</sup>. Pour des raisons de simplicité de mise en oeuvre, trois images sont prises avec trois angles du polariseur pour chaque scène. Les angles du polariseur sont  $0^\circ$ ,  $45^\circ$  et  $90^\circ$  [1]. Les véhicules de chaque image ont été labellisés par des boîtes englobantes. La base de données est divisée de telle sorte à ce que 115 scènes (75% de la base) sont utilisées pour la phase d'apprentissage et 38 scènes pour la validation. Le processus d'évaluation est inspiré du schéma d'évaluation utilisée sur la base de données PASCAL VOC [3], [7].

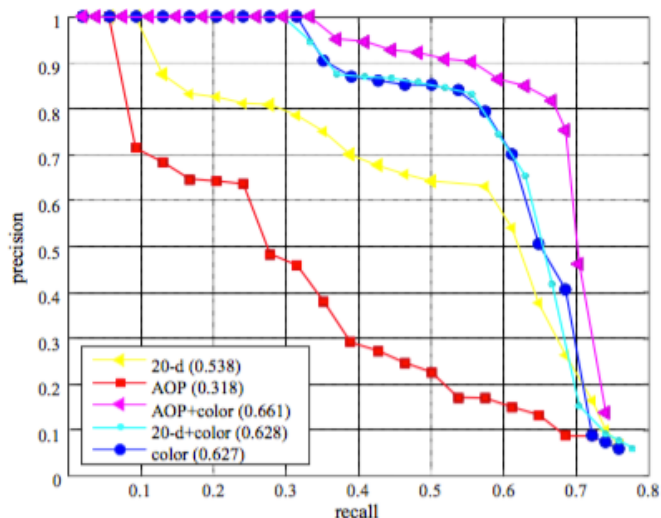


Figure 3. Courbe PRC appliquée sur la base de validation

Afin d'évaluer l'apport de l'information de polarisation pour le processus de détection, la courbe Précision-Rappel (PRC, Precision-Recall Curve) est calculée sur plusieurs configurations, sans et avec fusion de données multimodales. Afin de justifier notre choix de fusion AOP-couleur, le DPM est appris sur l'image AOP seul, sur l'image couleur seule (RGB), sur la fusion des deux modalités (AOP+couleur) et également sur le modèle 20\_d (avec les 5 canaux polarimétriques) et le modèle 20\_d fusionné avec les images couleur (20\_d+couleur). Les Courbes PRC de toutes les configurations proposées sont illustrées sur la figure.3

Tab 1 : Précision moyenne (PM) correspondante à la figure.3

	AOP	20_d	20_d+ couleur	RGB	AOP+ couleur
PM (%)	31.8	53.8	62.8	62.7	66.1

La précision moyenne (PM) issues de la figure.3 après la phase de validation est résumée dans le tableau Tab.1. La figure. 3 et le tableau.1 montrent que la fusion des caractéristiques de l'AOP et celles des images couleurs donnent les meilleurs résultats avec un taux de

<sup>1</sup> <http://pagesperso.litislab.fr/fwang/fichiers/>

détection de 66% et une amélioration de 4% par rapport à la détection classique sur les images couleurs, suivie de la fusion de la couleur et du modèle 20\_d polarimétrique (62.8%) et du modèle basé sur la couleur seule (62.7%). Ces deux derniers résultats sont très proches, alors qu'il devrait y avoir un écart. Ceci est dû au bruit des canaux polarimétriques qui empêche parfois l'exploitation de leurs informations. En revanche, les résultats ont montré que la polarisation seule (AOP seul ou modèle 20\_d.) n'est pas concluante pour la tâche de détection (31.8% et 53.8% respectivement).



Figure 4. Résultats de la détection par le DPM couleur (1<sup>ère</sup> colonne) vs le DPM couleur+Aop (2<sup>ème</sup> colonne)

Il est important de souligner que l'AOP seul occupe un taux de détection de 31% sur les 53% réalisée sur le modèle 20\_d issu de tous les canaux polarimétriques. Ces résultats montrent que l'AOP a un poids important et est complémentaire aux informations provenant des images couleurs. L'efficacité des résultats de la fusion de l'AOP et de la couleur est illustrée dans la figure.4 sur 3 exemples de scènes où les voitures sont sous des points de vues différents. Pour chaque scène, la première colonne montre les résultats du schéma basée que sur la couleur seule et la deuxième colonne, le résultat de la fusion proposée. Cette figure montre que par rapport à la méthode classique couleur, la méthode proposée permet de garder les fenêtres où il y a des vrais positifs et élimine les faux négatifs.

## 6 Discussion et conclusion

La méthode proposée montre la complémentarité apportée par l'information polarimétrique aux caractéristiques classiques basées sur l'intensité. Le modèle de fusion proposée, très sélectif, permet d'éliminer les négatifs vu qu'il est basé sur l'intersection des deux détections indépendantes mais ne permet pas de créer de nouveaux positifs. Un objet non détecté dans les deux modalités ne peut l'être après le

processus de fusion. En dépit de cette contrainte, ces premiers résultats montrent que la polarimétrie est une alternative très intéressante aux méthodes de détection classiques existantes. Ces résultats ouvrent le champ à plusieurs perspectives dans le futur. Il serait intéressant de tester des modèles de fusion conjoints aux deux modalités avant la phase d'apprentissage du détecteur afin de tirer pleinement profit des deux modalités et avoir un modèle de détection plus robuste et efficace. D'autres méthodes plus récentes comme le *transfert learning* basées sur les toutes nouvelles architectures en réseau de neurones profonds [6] méritent d'être testés sur cette nouvelle modalité. Ces méthodes sont connues pour leur efficacité dans la plupart des applications en vision par ordinateur. Il serait également intéressant d'appliquer une amélioration de la qualité des images polarimétriques, plus particulièrement celle du DOP, afin d'éliminer une partie du bruit et de maximiser le nombre de canaux polarimétriques significatifs pour le processus de détection. En conclusion, cette première étude a montré des résultats prometteurs et encourageants quant à l'apport de la polarimétrie pour l'analyse des scènes routières

## 7 Références

- [1] S. Ainouz, O. Morel, D. Fofi, et al., "Adaptive processing of catadioptric images using polarization imaging: towards a pola-catadioptric model," *Optical engineering* 52(3), 370011–370018, 2013.
- [2] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," In *IEEE CVPR*. 1, 886–893, 2005.
- [3] M. Everingham, L. Van Gool, C. K. I. Williams, et al., "The pascal visual object classes (voc) challenge," *IJCV* 88, 303–338, 2010.
- [4] Lawrence B. Wolff and Terrance E. Boult. Constraining object features using a polarization reflectance model. *IEEE Transaction on PAMI*, vol. 13, no 7, p. 635-657, 1991.
- [5] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D. "Object detection with discriminatively trained part-based models". *IEEE Transactions on PAMI*, 32(9), 1627-1645, 2010.
- [6] Kumar Santosh, Xiaoying Gao, and Ian Welch. "Learning Under Data Shift for Domain Adaptation: A Model-Based Co-clustering Transfer Learning Solution." *Pacific Rim Knowledge Acquisition Workshop*. Springer International Publishing, 2016.
- [7] T. Wu, B. Li, and S.-C. Zhu, "Learning and-or models to represent context and occlusion for car detection and viewpoint estimation," *IEEE Transactions on PAMI* 38(9), 1829–1843, 2015.