

Système de stéréovision pour la détection d'obstacles et de véhicule en temps réel

Gwenaëlle TOULMINET¹, Abdelaziz BENSRAHAI¹, Stéphane MOUSSET¹, Alberto BROGGI², Pierre MICHÉ¹

¹Laboratoire Perception Systèmes Information, UPRES EA 2120, Université/INSA de Rouen
Place Emile Blondel, 76131 Mont-Saint-Aignan Cedex, France

²Dipartimento di Informatica e Sistemistica, Università di Pavia
Via Ferrata, 1, I-27100 Pavia, Italie

gwenaelle.toulminet@insa-rouen.fr, abdelaziz.bensrahai@insa-rouen.fr
stephane.mousset@insa-rouen.fr, alberto.broggi@unipv.it
pierre.miche@insa-rouen.fr

Résumé – Dans le cadre de l'aide à la conduite automobile, nous présentons deux méthodes de détection d'obstacles et de détection de véhicule à partir de notre système embarquable de stéréovision. Ces deux tâches sont effectuées en temps réel en segmentant des cartes éparses de profondeur par sélection de segments 3D. Pour la détection d'obstacles, la sélection des segments 3D s'effectue à partir du calcul de leur angle d'inclinaison. La détection de véhicule s'effectue à partir des données fournies par ARGO, le véhicule expérimental autonome développé à l'Université de Parme.

Abstract – Within the framework of driving assistance, an obstacle detection method and a vehicle detection method are presented. Based on the configuration of our embedded stereovision system, these two tasks are computed in real time by segmenting sparse depth maps. Obstacles depth map is constructed by thresholding 3D segments inclination angles. Vehicle depth map is constructed using the data of ARGO, the experimental autonomous vehicle developed at the university of Parma.

1 Introduction

Dans le cadre de l'aide à la conduite automobile, deux techniques existent pour la perception de l'environnement d'un mobile, suivant qu'elles font appel à des capteurs actifs ou passifs. Lorsque plusieurs véhicules sont présents dans un même environnement, les systèmes actifs peuvent interférer entre eux et compromettre la fiabilité et l'utilité des informations qu'ils apportent [1]. Ainsi, nous présentons des méthodes originales de détection d'obstacles et de détection de véhicule¹ à partir de notre système embarquable de stéréovision passif. Le choix d'une géométrie épipolaire de notre système de stéréovision nous permet d'effectuer chacune de ces tâches en temps réel.

Pour détecter des obstacles et des véhicules dans un environnement routier, nous avons choisi de segmenter des cartes éparses de profondeur par sélection de segments 3D. Les cartes éparses de profondeur représentent les scènes routières réelles 3D et sont obtenues par mise en correspondance de points de contour en utilisant une méthode de programmation dynamique. Les points de contour sont extraits de façon auto-adaptative en utilisant un opérateur développé au laboratoire et appelé déclivité. Les segments 3D [2] sont construits à partir de points 3D (résultats de la mise en correspondance) en appliquant des critères de connexité, de profondeur et d'unicité. Les méthodes de sélection des segments 3D sont différentes suivant qu'il s'agit de la détection d'obstacles ou la détection de véhicule.

Dans le premier cas, nous les sélectionnons à partir d'une estimation de leur angle d'inclinaison. Dans le deuxième cas, notre approche consiste à construire une carte de profondeur d'un véhicule détecté par ARGO, le véhicule expérimental autonome développé au *dipartimento di Informatica dell'Informazione università di Parma*. Nous sélectionnons alors les segments 3D à partir des données fournies par ARGO [3].

Dans la deuxième section de ce papier, nous présentons notre algorithme de construction de segments 3D. Puis, les méthodes de sélection de segments 3D sont décrites dans la section 3 pour la détection d'obstacles, et dans la section 4 pour la détection de véhicule. Des résultats expérimentaux sont présentés à la fin de ce papier.

2 Construction de segments 3D

Nous construisons des segments 3D à partir de courbes 3D qui sont obtenues à partir de cartes éparses de profondeur. En se basant sur la configuration de notre système de stéréovision, l'algorithme de construction des courbes 3D est appliqué sur chaque ligne image.

2.1 Configuration du système de stéréovision

Notre système de stéréovision est constitué d'un bâti rigide, de deux objectifs identiques, de deux caméras cartes Philips VMC3405 et de la carte d'acquisition Imaging

1. Le travail de détection de véhicule a été financé par une action intégrée franco-italienne GALILEE 2000

Technology PC-RGB (figure 1).

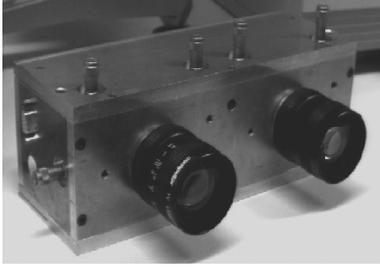


FIG. 1: *Le capteur de vision stéréoscopique.*

Les deux caméras cartes sont placées dans une configuration spécifique: les blocs caméras-objectifs sont positionnés de façon à avoir leurs axes optiques parallèles. Afin de respecter la contrainte épipolaire, la droite joignant les deux centres optiques des objectifs est parallèle aux lignes de balayage des deux caméras. Ainsi, la profondeur Z d'un point 3D noté P d'un objet est donnée par la relation suivante :

$$Z = \frac{f e}{p \delta} \quad (1)$$

e étant la distance entre les deux centres optiques, p la largeur du pixel du capteur CCD, f la distance focale des deux objectifs et δ la disparité horizontale des deux points stéréo correspondants du point 3D P . Soit P_g et P_d ces deux points de coordonnées images (x_g, y_g) et (x_d, y_d) avec $(y_d = y_g)$. La disparité δ , exprimée en pixels, est égale à $(x_d - x_g)$.

2.2 Construction de cartes éparses de profondeur

L'algorithme de construction des cartes éparses de profondeur est composé de deux étapes et est appliqué pour chaque ligne image. Dans un premier temps, les points de contour de deux lignes images droite et gauche sont extraits de façon auto-adaptative en utilisant un opérateur développé au laboratoire et appelé déclivité. Dans un deuxième temps, les points de contour de la ligne image droite sont mis en correspondance avec ceux de la ligne image gauche par une méthode de programmation dynamique en tenant compte de contraintes géométriques, photométriques, d'ordre et d'unicité. Les résultats de mise en correspondance sont des points de contour 3D dont la profondeur est donnée par l'équation 1. En utilisant cet algorithme de mise en correspondance, nous obtenons des résultats très satisfaisants : pour le traitement d'une paire d'images correspondant à une scène d'intérieur, 92.6% des points de contour sont appariés, et parmi ces appariements 98% sont corrects.

2.3 Construction de courbes 3D

A partir des points 3D de contour, nous construisons les courbes 3D. Pour chaque point 3D de la ligne image deux cas sont possibles. Si ce point 3D, noté N , respecte des

critères de connexité, de profondeur et d'unicité, il prolonge une courbe 3D. Sinon, il génère une nouvelle courbe 3D. Pour cela, une contrainte est définie entre un point N et un point 3D, noté M , appartenant à une des deux lignes images précédentes. Le critère d'unicité permet la construction de courbes 3D « sans ramification ». Si les critères de connexité et de profondeur sont validés alors le point N prolonge la courbe 3D à laquelle appartient le point M .

La dernière étape de notre calcul concerne la construction des segments 3D à partir des courbes 3D. Puisque l'environnement routier est structuré, alors les courbes 3D peuvent être approximées par un ou plusieurs segments 3D. Ainsi, les courbes de petite longueur sont éliminées, les autres sont approximées par des segments 3D en utilisant une méthode itérative de partition [4].

3 Détection d'obstacles

Pour détecter des obstacles en temps réel dans un environnement routier, nous avons choisi de segmenter les cartes éparses de profondeur par sélection de segments 3D, à partir des estimations de leur angle d'inclinaison (c.f. figure 2).

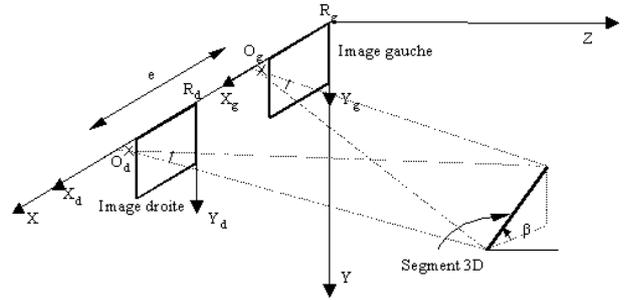


FIG. 2: *Angle d'inclinaison β d'un segment 3D.*

En effet, les primitives d'une scène routière peuvent être classées en fonction de leur appartenance soit à un obstacle soit à la route. Le choix de l'angle d'inclinaison comme moyen de sélection est donc approprié, et permet de faire de la détection d'obstacles [5]. Afin de sélectionner correctement les primitives d'une scène routière réelle appartenant à des obstacles, nous avons choisi de calculer l'angle d'inclinaison sur des segments 3D. Pour une bonne robustesse des résultats, nous avons choisi d'utiliser les équations des projections des segments 3D dans l'image droite et des projections estimées dans l'image gauche.

La construction des courbes 3D utilise un critère de connexité (paragraphe 2.3). Elle dépend donc de la construction de sa projection dans l'image droite. Ainsi, un segment 3D est associé à sa projection dans l'image droite. En utilisant cette projection et à partir des résultats de mise en correspondance, nous obtenons l'estimée de la projection du segment 3D dans l'image gauche. Les équations de ces deux segments sont calculées dans le repère $(R_d X_d Y_d)$ (c.f. figure 2), et dans le repère $(R_g X_g Y_g)$ pour l'estimée de la projection dans l'image gauche (équation 3).

$$x_d = m_d y + b_d \quad (2)$$

$$x_g = m_g y + b_g \quad (3)$$

(x_d, y) et (x_g, y) sont les coordonnées images des points stéréo correspondants aux points 3D. m_g , m_d , b_g et b_d sont calculés par la méthode des moindres carrés. Par géométrie, la tangente de l'angle d'inclinaison du segment 3D est donnée par la relation suivante :

$$\tan \beta = \frac{p_y (b + m \frac{h}{2})}{\sqrt{p_x^2 ((m_d b_g - m_g b_d) - \frac{w}{2} m)^2 + f^2 m^2}} \quad (4)$$

avec $m = m_d - m_g$, $b = b_d - b_g$, p_x et p_y la largeur et la hauteur du pixel du capteur CCD, f la distance focale des deux objectifs et $w \times h$ la résolution en pixels des caméras cartes.

En supposant que la route est plate, nous sélectionnons par seuillage de l'angle d'inclinaison les segments qui appartiennent à un obstacle et nous éliminons ceux qui appartiennent à la route. Dans une dernière étape, pour mettre en évidence les obstacles, nous utilisons une méthode d'interpolation associée à un critère d'éloignement entre les différents segments.

4 Détection de véhicule

Dans le cadre d'une action intégrée franco-italienne GALILEE 2000, notre laboratoire s'est associé au laboratoire *Dipartimento di Informatica dell'Informazione università di Parma* pour évaluer les performances et envisager des améliorations des systèmes passifs embarqués temps réel des deux équipes. Dans une première approche du programme GALILEE, nous avons conçu un système temps réel de détection de véhicule [6]. Pour cela, nous utilisons les données associées à un véhicule détecté par ARGO comme moyen de sélection des segments 3D. Le véhicule détecté est modélisé par un rectangle, dont ARGO nous fournit la profondeur D , la hauteur H , la largeur W , et les coordonnées du point milieu de l'arrête du bas. Ces données nous permettent ainsi de sélectionner les segments 3D contenus dans un cube de hauteur $H + 2 \times \Delta_H$, de largeur $W + 2 \times \Delta_W$ et de profondeur $2 \times \Delta_D$. Le centre de ce cube est à une distance D . Δ_H , Δ_W et Δ_D sont utilisés pour compenser le bruit sur les données. Puisque un véhicule est un obstacle, nous avons récemment intégré la méthode de détection d'obstacles. Enfin, pour chaque segment 3D nous calculons la moyenne de la profondeur de tous ses points. Le segment 3D le plus proche nous donne la distance du véhicule par rapport au système de stéréovision.

5 Résultats expérimentaux

Dans cette section, nous présentons des résultats expérimentaux des méthodes de détection d'obstacles et de véhicule sur une paire d'images réelles (figure 3). Les photos (a) et (b) représentent l'image gauche et l'image droite

d'une scène d'extérieur. Ces images ont été acquises simultanément par notre système de stéréovision embarqué dans le véhicule ARGO. Ces images sont au format 720 x 284 x 8 bits. Les résultats d'extraction des points de contour de l'image gauche et de l'image droite sont présentés sur les cartes (c) et (d). La carte (e) représente les estimées des projections des segments 3D dans l'image gauche et (f) représente leur projection dans l'image droite. Les cartes (g) et (h) sont les projections des segments 3D considérés comme appartenant à des obstacles. (i) est la carte de profondeur des obstacles, et est codée en utilisant des valeurs de niveaux de gris : les niveaux de gris élevés représentent des objets proches du système de vision stéréoscopique. Enfin, (j) est la carte de profondeur du véhicule. La distance estimée du véhicule par rapport au système de stéréovision est de 18,8m. La distance réelle est de 18,4m. Le temps de calcul pour chacune des deux méthodes est de environ 450 ms avec un PC Pentium III 800 Mhz.

6 Conclusion

Dans cette communication, nous avons présenté des méthodes originales de détection d'obstacles et de véhicule en temps réel, utilisant des données issues de capteurs passifs. La détection d'obstacles est effectuée par sélection de segments 3D à partir du calcul de leur angle d'inclinaison. La détection de véhicule s'effectue par coopération à partir des données fournies par ARGO. Des résultats expérimentaux ont également été présentés.

Basé sur la configuration de notre système de stéréovision, nous avons développé des algorithmes de construction de cartes de vitesses axiales en temps réel (0.8 s sur PC Pentium III 800 Mhz) à partir de séquences de cartes de profondeur [7]. Afin de compléter nos méthodes de détection, nous souhaitons construire des cartes denses de la vitesse axiale des obstacles et des véhicules.

Références

- [1] A. Broggi et E.D. Dickmanns, *Applications of computer vision to intelligent vehicles*, Image and Vision Computing, p.365-366, vol. 18, 2000
- [2] O. Faugeras, *Tree-Dimensional Computer Vision*, The M.I.T. Press
- [3] A. Broggi, M. Bertozzi et A. Fascioli, *Stereo Vision-based Vehicle Detection*, In Proceedings IEEE IV-2000, Intelligent Vehicles Symposium, p. 39-44, Octobre 2000
- [4] T. Pavlidis et S.L. Horowitz, *Segmentation of Plane Curves*, IEEE trans. on Computers, 23(8), p. 860-870, 1974
- [5] K. Sobottka et H. Bunke, *Obstacle detection in range image sequences using radial slope*, In 3rd IFAC Symposium on Intelligent Autonomous Vehicles, Madrid Espagne, Mars 1998
- [6] A. Bensrhair, M. Bertozzi, A. Broggi, P. Miché, S. Mousset et G. Toulminet, *A Cooperative Approach*

to Vision-based Vehicle Detection, In the 4th International IEEE Conference on Intelligent Transportation Systems, Oakland USA, Août 2001, à paraître

- [7] S. Mousset, A. Bensrhair et P. Miché, *A fast and automatic stereovision system for axial motion detection*, International Conference on Computer Vision, Pattern Recognition and Image Processing CV-PRIP'2000, USA, février 2000

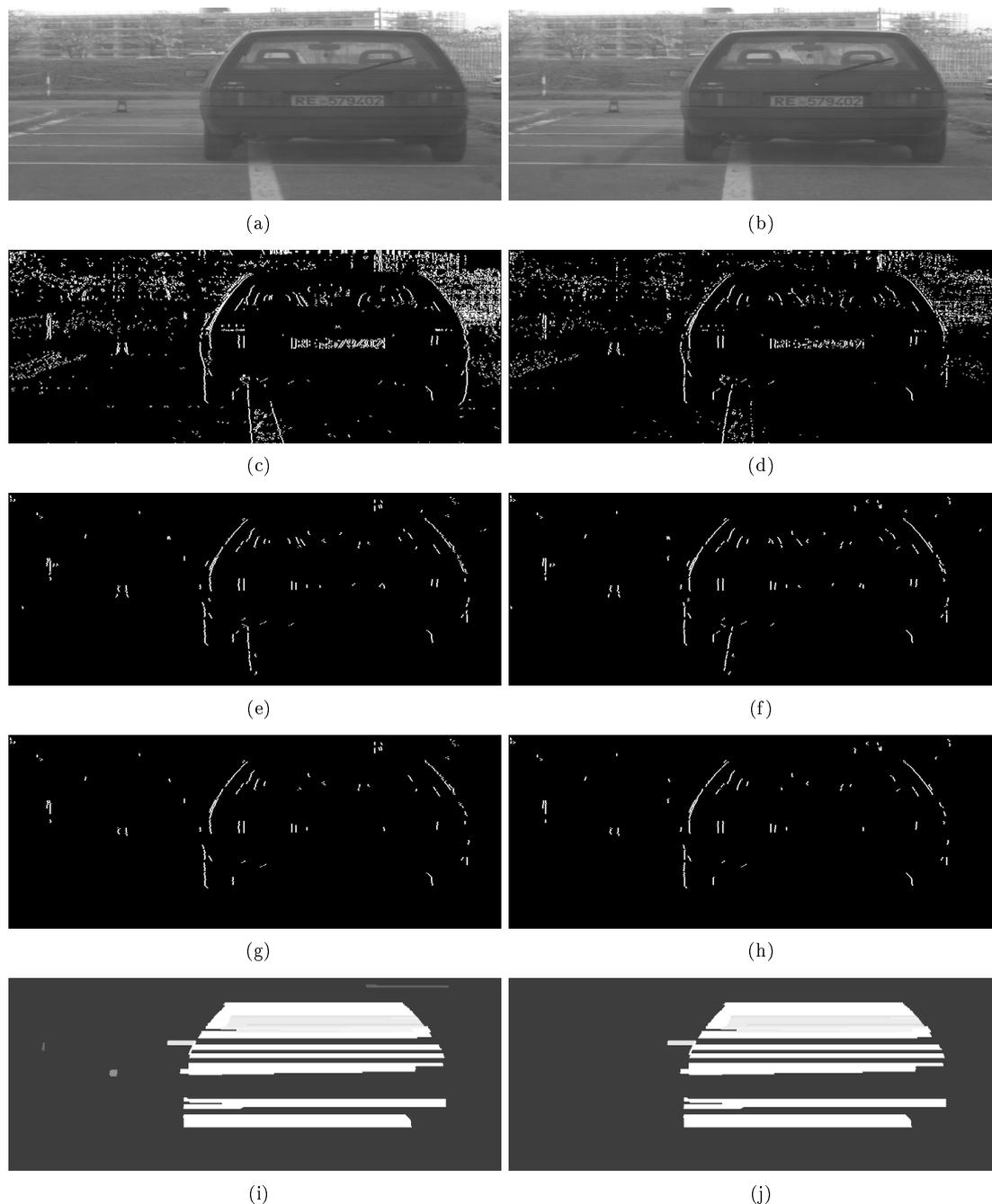


FIG. 3: Résultats expérimentaux: (a): image gauche, (b): image droite, (c): points de contour de l'image gauche, (d): points de contour de l'image droite, (e): estimées des projections des segments 3D dans l'image gauche, (f): projections des segments 3D dans l'image droite, (g): estimées des projections des segments 3D appartenant à des obstacles dans l'image gauche, (h): projections des segments 3D appartenant à des obstacles dans l'image droite, (i): carte de profondeur des obstacles, (j): carte de profondeur du véhicule détecté par ARGO.