

Détection d'Obstacles dans une Séquence d'Images de Rétrovision par Analyse du Mouvement

Rong YANG, Sébastien AMBELLOUIS et François CABESTAING

Centre d'Automatique de Lille,
Université des Sciences et Technologies de Lille,
59655 Villeneuve d'Ascq CEDEX, France. Tel +33 3.20.43.41.69
email : ry@cal.univ-lille1.fr

RÉSUMÉ

Dans cet article, nous abordons le problème de l'amélioration de la sécurité de conduite sur autoroute. Le champ de vision du conducteur dans le rétroviseur, qui se trouve limité par la présence d'un angle mort, ne lui permet pas toujours de détecter à temps un obstacle potentiel. Nous utilisons les images issues d'une caméra placée au niveau du rétroviseur latéral gauche, et effectuons une détection d'obstacles en utilisant un rééchantillonnage suivi d'une analyse de mouvement. L'estimation du mouvement est réalisée par un réseau de neurones à temps de propagation (TDNN).

ABSTRACT

In this paper, we present an obstacle detection framework which aims at avoiding collisions during overtaking on a highway. Potential obstacles are detected by a vision system which processes the images given by a video camera installed near the left lateral rearview mirror of the car. Images of the sequence are resampled, following an adaptative model of the road geometry. Then a motion analysis step allows an efficient discrimination between dangerous and non dangerous potential obstacles. Motion estimation is carried out by a Time Delayed Neural Network (TDNN).

1 Introduction

L'amélioration de la sécurité de conduite, principalement sur autoroute est une des priorités des pouvoirs publics. Des études statistiques ont montré que de nombreux accidents graves sont causés par une collision latérale ou arrière avec un autre véhicule lors d'un changement de voie ou un dépassement. Le champ de vision perçu par le conducteur dans le rétroviseur latéral gauche est relativement restreint et ne lui permet pas toujours de détecter à temps les véhicules constituant des obstacles potentiels. Ceci est dû en partie à l'existence d'un angle mort dans la zone de visibilité, qui occulte totalement toute une partie du champ de vision.

Pour remédier à ce problème, plusieurs moyens ont été proposés. Une solution consiste à placer des capteurs sur le côté et à l'arrière du véhicule [5, 10, 2]. Plusieurs technologies de capteurs ont été testées : capteur ultrasonore, radar ou lidar. On peut également utiliser comme capteur une caméra CCD couplée à un système de vision artificielle assurant la détection des obstacles potentiels dans les images fournies par la caméra [2, 8, 3].

Dans ce travail, nous utilisons les images issues d'une caméra placée au niveau du rétroviseur latéral gauche. Les conditions de prise de vue ont été choisies de telle sorte que l'angle mort soit réduit au minimum. Les obstacles potentiels sont détectés grâce à une analyse de mouvement effectuée sur la séquence d'images fournie par la caméra. Les images sont tout d'abord rééchantillonnées pour diminuer le nombre de pixels traités et pour compenser les déformations dues à la perspective. Ensuite, l'analyse du mouvement est réalisée par

un réseau de neurones à temps de propagation (TDNN) [1]. Dans la dernière section de cet article, nous présentons les résultats de détection obtenus sur une séquence réelle acquise sur autoroute.

2 Pré-traitement des images

Les images fournies par la caméra de rétrovision présentent une vue en perspective de la scène située à l'arrière gauche du véhicule, dans laquelle les dimensions et l'aspect des autres véhicules se modifient en fonction de leur éloignement. Pour compenser ces déformations dues à l'effet de perspective, nous avons utilisé une procédure de rééchantillonnage des images [7, 4, 11]. Cette procédure de rééchantillonnage permet également de diminuer le nombre total de pixels pris en compte durant la phase d'analyse du mouvement, ce qui autorise une diminution considérable du temps de calcul.

Lorsqu'on connaît précisément les caractéristiques géométriques et optiques du système de prise de vue (distance focale, hauteur de caméra, inclinaison, orientation . . .), on peut aisément définir la projection dans le plan image d'une grille régulière positionnée sur la route (cf. figure (1)). Inversement, on peut obtenir une image pour laquelle la perspective a été corrigée en sélectionnant dans l'image initiale les pixels situés sur la projection de la grille. Cette correction de l'effet de perspective n'est qu'approximative, puisqu'elle ne peut pas prendre en compte le caractère tridimensionnel de la scène initiale.

Les paramètres définissant la procédure de rééchantillonnage sont issus à la fois de la connaissance du système optique

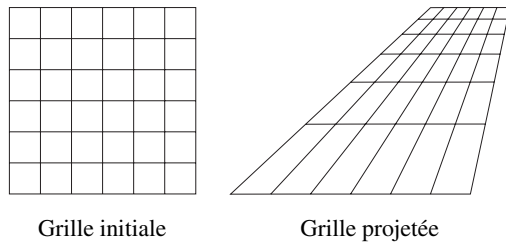


FIG. 1 — Projection d'une grille

et d'une estimation de la géométrie de la route. Dans le modèle simplifié que nous avons implanté, nous supposons que la route est plane et que la trajectoire des véhicules est soit linéaire soit circulaire. Il suffit alors de déterminer le rayon de courbure de la route à l'arrière du véhicule, ainsi que la position du véhicule sur les voies.

Ces deux paramètres sont fournis par une procédure de recherche des lignes blanches présentes dans les images, qui se décompose en deux phases successives. Sur quelques lignes extraites de l'image initiale, on détecte tout d'abord les variations brutales du niveau de gris par une dérivée première estimée par un filtre de Deriche. Une ligne blanche fait apparaître simultanément un maximum et un minimum dans la dérivée. La distance qui sépare ces deux extréma dépend naturellement de la largeur réelle de la ligne blanche, mais également de sa position dans l'image. Dans une deuxième phase, la position de la voiture sur les voies et le rayon de courbure de la route sont estimés à partir de la position des lignes blanches dans l'image par une méthode de moindres carrés.

3 Estimation du mouvement par réseaux de neurones

Après rééchantillonnage, nous supposons que les déformations dues à la perspective ont été corrigées. Ainsi, si on suppose qu'un véhicule présent dans l'image de rétrovision se déplace à une vitesse relative constante dans la scène réelle, le champ des déplacements créé par ce véhicule dans les images rééchantillonnées sera également uniforme et constant [9].

Dans [1], Ambellouis a proposé une méthode d'analyse du mouvement local basée sur l'utilisation de réseaux de neurones à temps de propagation sur les connexions (TDNN : Time Delayed Neural Network). Cette méthode s'avère très efficace pour estimer de façon précise les paramètres du mouvement lorsque la vitesse de l'objet varie très peu dans la séquence d'image. Elle peut s'assimiler à d'autres méthodes basées sur un filtrage spatio-temporel de l'image [6].

Lorsqu'on connaît la position d'un objet dans l'une des images d'une séquence, une simple procédure d'extrapolation permet de définir sa position dans toutes les images, si on suppose que le mouvement est uniforme. Inversement, s'il existe une corrélation entre une séquence d'images prédites et la séquence étudiée, on peut affirmer que le mouvement est uniforme et en déduire ses paramètres. Toutes les méthodes d'analyse du mouvement basées sur une corrélation exploitent

ce principe. L'estimation du mouvement s'effectue en deux étapes : les pixels de l'image affectés par le mouvement sont marqués grâce à un algorithme de détection du mouvement. Ensuite le vecteur vitesse est déterminé en recherchant un maximum de corrélation.

Comme toutes les méthodes de corrélation appliquées en traitement d'images, l'analyse de mouvement requiert une capacité mémoire importante et un temps de calcul prohibitif. En effet, il faut calculer la corrélation entre les images de la séquence réelle et les images prédites pour un nombre important de valeurs possibles de la vitesse de déplacement. Afin de palier à ce dernier inconvénient, Ambellouis a proposé une implantation sur une architecture parallèle, en utilisant un formalisme neuronal [1]. Une autre simplification consiste à modifier légèrement l'expression de la corrélation pour permettre le calcul de la réponse de façon récursive.

Le traitement récursif repose sur le principe suivant : lorsque qu'un pixel de coordonnées (x, y) dans l'image courante (à l'instant t) est situé dans une zone de l'image affectée par un mouvement, la vraisemblance d'une vitesse $v = (v_x, v_y)^T$ en ce point est déterminée en utilisant la vraisemblance de cette même vitesse v déterminée à l'instant $t - \Delta t$ au point $(x - v_x \cdot \Delta t, y - v_y \cdot \Delta t)$. Ainsi l'estimation de la vraisemblance d'une vitesse en un point ne nécessite que très peu de calculs élémentaires.

Dans notre application, nous ne recherchons qu'une seule composante du vecteur vitesse, celle qui correspond au déplacement des véhicules dans la direction de la route. De plus, un véhicule ne constitue un obstacle potentiel que s'il se rapproche de la caméra, donc si sa vitesse relative est positive. Dans ces conditions, le temps de calcul nécessaire à l'estimation du mouvement par le réseau TDNN est très largement diminué, ce qui permet de l'utiliser dans un contexte temps-réel.

4 Résultats

La figure (2) présente une série d'images de rétrovision extraites d'une séquence filmée sur autoroute. Cette séquence a été acquise dans des conditions de prise de vue non idéales, la forte luminosité ambiante ayant entraîné une saturation des niveaux de gris dans le haut de l'image. Sur cette séquence, on peut voir une voiture qui dépasse notre véhicule avec une vitesse relative assez élevée.

La figure (3) présente la grille de rééchantillonnage utilisée si on suppose que la route est rectiligne. Sur la figure (4), la grille est définie pour corriger la perspective en faisant intervenir la courbure de la route.

La figure (5) montre les images résultant du rééchantillonnage de la séquence de rétrovision présentée sur la figure (2). Sur chaque petite image, l'axe de la chaussée est ramené sur l'horizontale, la zone proche de notre véhicule correspond à la partie gauche de l'image. La première série d'images a été obtenue en utilisant une grille rectiligne, alors que la seconde est issue d'une grille suivant la courbure de la route. On constate que le deuxième modèle est légèrement plus précis, les lignes blanches discontinues ayant été transformées en segments horizontaux réguliers.



FIG. 2 — Séquence d'images de rétrovision

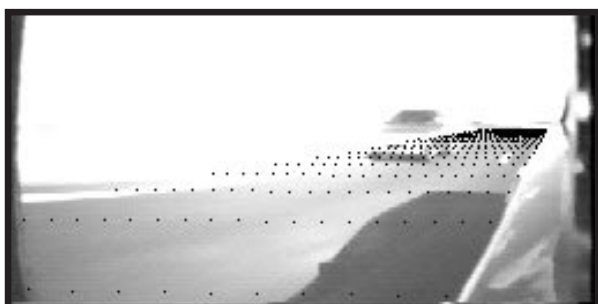


FIG. 3 — Grille d'échantillonnage rectiligne

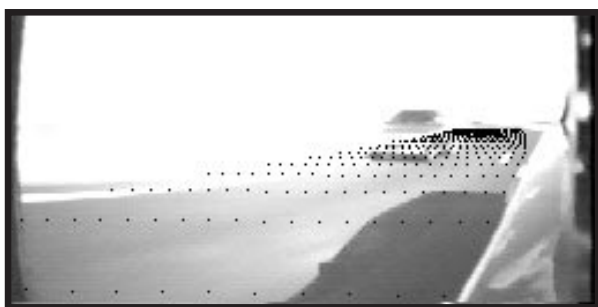


FIG. 4 — Grille d'échantillonnage courbe

L'analyse de mouvement réalisée sur les images rééchantillonnées fournit un maximum de corrélation pour une vitesse de 0.8 pixel par image. Cela correspond à une vitesse relative réelle d'environ 26 kilomètres à l'heure.

5 Conclusion

Les algorithmes décrits précédemment ont été testés en laboratoire sur des séquences d'images de rétrovision pour vérifier leur efficacité. Ces séquences, créées grâce à un logiciel de synthèse, présentaient différentes configurations : route droite ou virage, dépassement par un autre véhicule ou changement de voie.

Cela nous a permis de vérifier la nécessité du rééchantillonnage pour obtenir une détermination précise de la vitesse relative des véhicules détectés. La procédure d'analyse du mouvement par réseaux de neurones est parfaitement adaptée au problème de la surveillance de la route dans lequel la caméra et les véhicules sont mobiles.

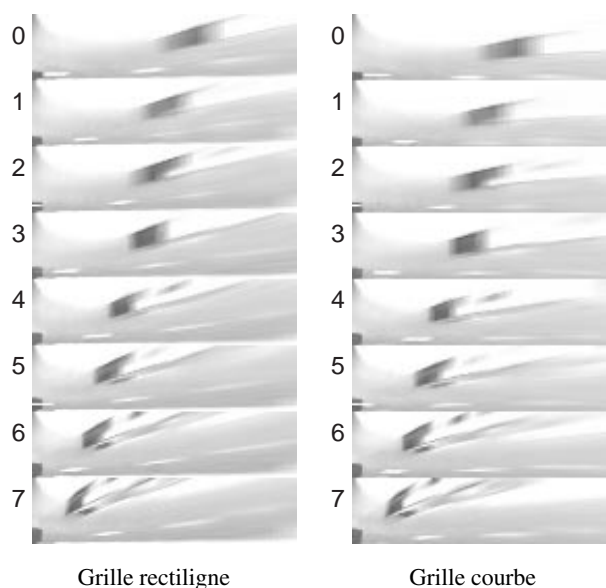


FIG. 5 — Séquences rééchantillonnées

Nous travaillons actuellement sur une implantation logicielle des procédures de traitement autorisant une analyse du mouvement à la cadence vidéo de 25 images par seconde. Enfin, nous envisageons une série de tests dans des conditions réelles d'utilisation.

Références

- [1] S. Ambellouis and F. Cabestaing. Motion analysis with a time delayed neural network. In *CESA'96 IMACS Multiconference, Symposium on Robotics and Cybernetics*, pages 328–332, Lille, France, 1996.
- [2] H. Araki, K. Yamada, Y. Hiroshima, and T. Ito. Development of rear-end collision avoidance system. In *IEEE Intelligent Vehicles Symposium*, pages 224–229, Seiki University, Tokyo, Japan, 1996.
- [3] M. Brattoli, R. Tasca, A. Tomasini, E. Chioffi, D. Gerna, and M. Pasotti. A vision-based off-road alert system. In *IEEE Intelligent Vehicles Symposium*, pages 195–220, Seiki University, Tokyo, Japan, 1996.
- [4] A. Broggi. Robust real-time lane and road detection in critical shadow conditions. In *ISCV'95 International*

Symposium on Computer Vision, pages 353–358, Coral Gables, Florida, 1995.

- [5] T. Butsuen, T. Yoshioka, and K. Okuda. Introduction of the mazda advanced safety vehicle ASV. In *IEEE Intelligent Vehicles Symposium*, pages 242–247, Seiki University, Tokyo, Japan, 1996.
- [6] D.J. Fleet and A.D. Jepson. Hierarchical construction of orientation and velocity selective filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11(3) :315–324, 1989.
- [7] V. Garefe and W. Efenberger. A novel approach for the detection of vehicles on freeways by real-time vision. In *IEEE Intelligent Vehicles Symposium*, pages 363–368, Seiki University, Tokyo, Japan, 1996.
- [8] T. Mimuro, Y. Miichi, T. Maemura, and K. Hayafune. Functions and devices of mitsubishi active safety ASV. In *IEEE Intelligent Vehicles Symposium*, pages 248–253, Seiki University, Tokyo, Japan, 1996.
- [9] S. Ohnaka, T. Ikeda, and M. Mizoguchi. Vehicle detection and velocity estimation using image sensor. In *ACCV'95 Second Asian Conference on Computer Vision*, pages 690–694, Singapore, 1995.
- [10] T. Tokoro. Automotive application systems of a millimeter-wave radar. In *IEEE Intelligent Vehicles Symposium*, pages 260–265, Seiki University, Tokyo, Japan, 1996.
- [11] J. Yang, S. Guo, T. Hasegawa, and S. Ozawa. Roadway matching technique from perspective image. In *ACCV'95 Second Asian Conference on Computer Vision*, pages 747–750, Singapore, 1995.