NOTE TO USERS

This reproduction is the best copy available.

UMI®

UNIVERSITÉ DE MONTRÉAL

Détection et localisation des panneaux routiers

DENIS FOURT

DÉPARTEMENT DE GÉNIE ÉLECTRIQUE ET DE GÉNIE INFORMATIQUE

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION

DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES (M.Sc.A.)

(GÉNIE ÉLECTRIQUE)

JUILLET 2000

© Denis FOURT, 2000.

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services
Acquisitions et services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-57404-0

Canada

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

Détection et localisation des panneaux routiers

présenté par: FOURT Denis,

en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées

a été dûment accepté par le jury d'examen constitué de:

BRAULT Jean-Jules, Ph.D., président

COHEN Paul, Ph.D., membre et directeur de recherche

HERVÉ Jean-Yves, Ph.D., membre et codirecteur de recherche

BAASS Karsten, Ph.D., membre

Remerciements

Je tiens, en premier, à remercier mon directeur de recherche Paul Cohen pour m'avoir accueilli dans son laboratoire, et dirigé au cours de ma maîtrise. Je tiens aussi à remercier mon codirecteur Jean-Yves Hervé pour son aide et ses conseils.

Je remercie aussi le CRSNG, le CNRC et la compagnie GIE technologies pour avoir financé mon projet de maîtrise.

Je tiens également à remercier les autres membres, passés ou présents du Groupe de Recherche en Perception et Robotique.

Je suis reconnaissant envers toutes celles et ceux qui m'ont soutenu durant ces années d'études.

Résumé

Les panneaux routiers contribuent à des routes plus sûres et à une circulation plus fluide par les informations qu'ils fournissent aux conducteurs. Pour qu'ils puissent toutefois jouer pleinement leur rôle, il est nécessaire qu'ils restent bien visibles et que les informations fournies restent cohérentes entre elles. Pour maintenir la qualité de ce service, il faut procéder régulièrement à des inspections. Si ces inspections sont réalisées en parcourant le réseau routier à l'aide de véhicules équipés de caméras et de systèmes de positionnement (pour estimer la position des panneaux), le jugement visuel d'un opérateur reste indispensable pour trouver puis reconnaître les panneaux dans les images. L'objectif de cette étude est de permettre l'automatisation de cette tâche en détectant les apparitions successives des panneaux (pour contrôler leur bonne visibilité) puis en les localisant (pour pouvoir les identifier sans ambiguïtés) à l'aide d'un système de vision artificielle. Nous avons considéré que ce système disposerait des informations recueillies à bord des véhicules d'inspection les plus avancés, qui embarquent un système de vision stéréoscopique ainsi qu'un système donnant très précisément la position et l'orientation de ce système de vision.

Les problèmes de la détection et de la reconnaissance automatique des panneaux ont déjà été étudiés dans le cadre des recherches sur les véhicules intelligents, dans le but d'assister les conducteurs. Ces problèmes sont complexes en raison de la très grande variété des scènes et des conditions d'éclairage rencontrées. Cependant les panneaux sont normalisés à des fins d'uniformité, ainsi la plage des variations possibles reste peu étendue.

Notre approche pour la détection des panneaux consiste à rechercher des régions de l'image ayant une couleur et une forme particulières. Nous procédons en trois étapes. L'emploi de l'espace de représentation des couleurs (Teinte, Saturation, Intensité) permet d'isoler les pixels pouvant appartenir à un panneau et donc de segmenter l'image. L'étude des contours des régions obtenues permet de sélectionner ceux qui sont suffisamment réguliers pour correspondre à des éléments de panneaux. Les panneaux sont finalement détectés grâce à une corrélation avec des masques présentant les formes recherchées sur l'image des contours. Bien que nécessitant une charge importante de calculs, la corrélation permet de détecter des panneaux qui ont été segmentés en plusieurs éléments ou fusionnés à une autre région. Une telle approche exclut les panneaux achromatiques. Nous avons en outre supposé que les éventuelles déformations des formes dans l'image (liées à la perspective, aux distorsions induites par la caméra et au bruit de mouvement) étaient négligeables.

Chacune de ces trois étapes dépend d'un certain nombre de seuils qui doivent être déterminés à l'aide d'un apprentissage sur un ensemble d'images présentant tous les

types de panneaux et de conditions qu'il est possible de rencontrer. Notre algorithme de détection a été testé avec succès pour différentes formes et couleurs. La recherche des petits panneaux produit toutefois des fausses alarmes.

L'estimation de la visibilité et de la position des panneaux requiert l'appariement des apparitions successives de ceux-ci. Nous proposons de le réaliser au moyen d'un système de poursuite couplé à la détection selon un schéma récursif de prédiction de la position du panneau puis détection de celui-ci. Ce qui permet de restreindre la recherche à quelques sous-régions de l'image. De plus, comme la détection des grands panneaux produit nettement moins de fausses alarmes, cette poursuite est initialisée à partir de l'une des dernières apparitions du panneau, puis celui-ci est recherché dans les images précédentes. La prédiction de la position du panneau dans une des images est réalisée en estimant la localisation par rapport à la caméra du panneau au moyen d'un filtre de Kalman étendu dont le vecteur d'état est formé en ajoutant à la localisation du panneau, une de ses dimensions. Il permet en outre d'estimer l'incertitude sur la prédiction et donc de dimensionner la zone de recherche. L'algorithme développé requiert la connaissance de la position et de l'orientation des caméras. Comme il nous est impossible de savoir si le filtre a bien convergé, la localisation des panneaux est envisagée a posteriori. L'association des diverses détections d'un panneau aux positions des caméras aux instants de prise de vue correspondants, permet d'obtenir un système d'équations linéaires sur-déterminé. La détermination des valeurs singulières permet de le résoudre.

Nous avons évalué ces deux algorithmes, en simulant le déplacement d'un véhicule d'inspection le long d'une route ainsi que l'acquisition des images, la détection des panneaux et l'estimation de la position et de l'orientation du système de vision. Cette évaluation a été réalisée pour différents modèles d'erreur sur la position des caméras. Nous avons obtenu de bons résultats pour le suivi des panneaux : l'algorithme permet de retrouver rapidement les panneaux qui sortent momentanément du champ de vision et conduit à des réductions importantes de la zone de recherche dans l'image. La localisation s'est avérée très sensible à l'erreur sur la position des caméras. Néanmoins avec un très bon système de positionnement, la position absolue des panneaux peut être estimée avec une erreur inférieure à 3 m dans 98% des cas. De telles performances sont inférieures à celles obtenues en réalisant manuellement la détection et l'appariement, mais notre système devrait conduire à des inspections plus rapides.

Cette étude a donc conduit au développement de trois outils fondamentaux pour automatiser l'inspection des panneaux. Nous avons d'abord conçu un système de détection qui soit capable de trouver toutes les apparitions des panneaux dans les images. Nous avons ensuite développé un système de suivi pour apparier ces apparitions successives et ainsi estimer la portion de route depuis laquelle un panneau est visible. Finalement nous avons proposé un algorithme de localisation afin de pouvoir identifier les panneaux. Un système d'inspection devrait en outre pouvoir reconnaître les panneaux. Les recherches futures devraient porter sur l'intégration de la détection au suivi.

Abstract

Information displayed on traffic signs play an important role in road safety and traffic flow. This requires that traffic signs remain perfectly visible and that there is no incoherence between the various indications they provide. Inspections must be scheduled in order to keep the quality of service to a high level. Despite the use of vehicles embedding motion cameras and positioning devices (in order to be able to locate signs), there is still a need for a visual analysis from an operator to locate and identify signs in the images. We aim to automate these inspections by developing an artificial vision system for detecting traffic signs in image sequences (to control their visibility) and for locating them (as the position allows to identify a sign). We have taken for granted that this artificial vision system would have access to all the measurements that can provide one of the most advanced inspection vehicles which embed a stereoscopic vision system and a measurement system to precisely estimate the position and the orientation of the stereoscopic system.

Traffic signs detection and recognition have already been addressed by some studies in the intelligent vehicles field in order to assist drivers. These are difficult problems

to solve because of the variety of backgrounds and lighting conditions encountered. Nevertheless, traffic signs are normalized, so the change in their appearance from one image to another is limited.

Our approach for traffic signs detection is based on the search of areas of the image that exhibit particular colors and shapes. We proceed in three steps. The use of the (Hue, Saturation, Intensity) color space allows to select the pixels which can belong to a sign and therefore leads to a segmentation of the image. Then, regions whose border is not regular enough to be signs are discarded. Traffic signs are finally detected by the mean of a correlation between masks exhibiting searched shapes and the image of the borders of the segmented regions. Despite the fact of being computationally expensive, correlation allows to detect traffic signs that have been segmented in more than one region or merged with an other region. Such an approach can not be applied to achromatic signs. We have also assumed that shape warping (induced by perspective, by the camera and by the motion of the vehicle) can be neglected.

Each of these steps depends on thresholds which are estimated during a learning on set of images. This set must be representative of all signs and conditions encountered. Our detection algorithm has been successfully tested for different shapes and colors. Detection of small signs leads, however, to false alarms.

The estimation of the visibility and of the position of signs implies to establish correspondance between their various appearances in the images. We propose to use

a tracking algorithm to predict the area of the next image where the sign should be searched by our detection algorithm, following a recursive process. This allows to reduce the area of the image where signs have to be searched. As detection of large signs is less prone to false alarms, the tracking starts from one of the last appearances of the sign, which is then searched in the previous images. An extended Kalman filter is used to estimate the location and the size of the signs which then gives us a prediction in the image. The filter allows also to estimate the uncertainty on the prediction and therefore to evaluate the dimensions of the searched area. Our tracking algorithm requires the knowledge of the positions and the orientations of the cameras. As we have no clue to know wether the extended Kalman filter has converged, signs location is performed in batch mode. In associating the various positions of a sign in the images to the positions of the cameras, we obtain a overdetermined linear system which is solved by the mean of singular value decomposition.

We evaluated these algorithms using simulations of a moving inspection vehicle along a road. Traffic signs detection is simulated by adding an error on the position of the signs in the image. The effect of various error models for camera positions was studied. We obtained good results for signs tracking : the tracking algorithm is able to find again signs that went previously out of a camera view field. On the other hand, sign location is very sensitive to the camera position error. However, using a high quality positioning system, it is possible to locate 98% of signs with an accuracy within three meters. Such results are inferior to those that can provide a

GPS receptor at the sign but we expect a fully automated moving system to be much faster.

During this study we developed three key tools for the automation of traffic signs inspection. First, a traffic sign detection algorithm was designed to detect all the appearances of the signs in the images. Then a tracking algorithm was developed to establish correspondance between these appearances and therefore to estimate the visibility of the sign. Finally a simple locating algorithm was proposed in order to be able to identify without ambiguities the sign. An inspection system should also perform traffic sign recognition to get the exact model of the sign. Further researches should aim to a complete integration of signs tracking and detection.

Table des matières

Remerciements i

Résumé ii

Abstract vi

Table des matières x

Liste des tableaux xvii

Liste des figures xix

Liste des annexes xxiii

Liste des notations xxiv

1 Introduction 1

1.1 Contexte 1

1.1.1 Surveillance des panneaux routiers 1

1.1.2 Détection et reconnaissance des panneaux routiers pour des véhicules intelligents 3

1.2 Les panneaux routiers 4

1.2.1 Généralités 4

1.2.2 Normes et emploi des panneaux au Québec (d'après Baass (1993)) 7

1.2.3 Conclusion 9

1.3 Obtention et gestion de la position des panneaux 10

1.4 Objectifs 11

1.5 Organisation du mémoire 13

2 Description du système d'inspection et présentation des données 15

2.1 Le système d'inspection 15

2.1.1 Module d'acquisition d'images 16

2.1.2 Module de positionnement du système de vision 18

2.1.3 Module de détection des panneaux 18

2.1.4 Module de poursuite des panneaux 20

2.1.5 Module de localisation des panneaux 20

2.1.6 Modules et sous-modules étudiés 21

2.2 Données 22

2.2.1 Images de l'université de Koblenz-Landau 22

2.2.2 Images du Groupe de Recherche en Perception et Robotique 27

2.2.3 Utilisation des images et des séquences 33

3 Segmentation basée sur la couleur 35

3.1 Introduction 35

3.2 Revue bibliographique 36

3.2.1 Avantages et inconvénients de la couleur 36

3.2.2 Représentation des couleurs 36

3.2.3 Segmentation des panneaux routiers 40

3.2.4 Implantation 44

3.2.5 Conclusion 45

3.3 Solution proposée pour la segmentation 46

3.3.1 Principe 46

3.3.2 Choix des seuils 47

3.4 Expériences 48

3.5 Résultats 49

3.6 Discussion 50

4 Détection des panneaux grâce à leur forme 61

4.1 Introduction 61

4.2 Revue bibliographique 62

4.2.1 Avantages et inconvénients de la forme 62

4.2.2 Signatures non-basées sur la forme 62

4.2.3 Le problème de la détection 63

4.2.4 Conclusion 66

4.3 Solution proposée pour la détection des panneaux 67

4.4 Pré-reconnaissance de forme 68

4.4.1 Introduction 68

4.4.2 Les opérations morphologiques de type "hit or miss" 68

4.4.3 Elimination des pixels parasites 69

4.4.4 Extraction des contours 72

4.4.5 Rejet des contours n'étant pas des panneaux 72

4.4.6 Interaction avec la segmentation 73

4.4.7 Rejet des contours des panneaux de direction 75

4.5 Localisation des panneaux dans l'image 80

4.5.1 Introduction 80

4.5.2 Principe et définitions 80

4.5.3 Description des masques utilisés 81

4.5.4 Sélection des pics de corrélation 82

4.5.5 Choix des seuils 84

4.6 Expériences 86

4.6.1 Détermination des performances de la détection 88

4.6.2 Robustesse vis-à-vis de l'ensemble d'apprentissage 88

4.6.3 Détection des grands panneaux 89

4.7 Résultats 89

4.7.1 Résultats de la détection 89

4.7.2 Apprentissage des seuils 89

4.7.3 Détection des panneaux réels avec les plus grands masques . . 91

4.8 Discussion 92

4.8.1 Analyse des résultats 92

4.8.2 Analyse de l'algorithme 98

5 Suivi et localisation des panneaux 102

5.1 Introduction 102

5.2 Revue bibliographique 103

5.2.1 Introduction 103

5.2.2 Suivi des panneaux 104

5.2.3 Estimation du mouvement et de la structure de la scène . . . 104

5.2.4 Conclusion 108

5.3 Utilisation de la redondance temporelle 109

5.3.1 Objectifs 109

5.3.2 Principe 110

5.3.3 Poursuite et localisation des panneaux 110

5.3.4 Intégration 112

5.4 Suivi des panneaux 118

5.4.1 Filtrage de Kalman 118

5.4.2 Repères utilisés et définitions 120

5.4.3 Premier modèle 123

5.4.4 Second modèle 125

5.4.5 Prédiction de la zone de recherche 127

5.5 Localisation des panneaux a posteriori 128

5.5.1 Introduction 128

5.5.2 Formulation 128

5.6 Simulation de l'inspection d'une portion de route 130

5.6.1 Simulation d'une trajectoire 131

5.6.2 Simulation du système de vision 133

5.6.3 Simulation de la détection des panneaux 133

5.6.4 Simulation du système de positionnement 135

5.7 Expériences 136

5.7.1 Objectifs 136

5.7.2 Initialisation du filtre 137

5.7.3 Dimensions de la zone de recherche 138

5.7.4 Étude de la prédiction dans l'image 139

5.7.5 Étude du suivi 140

5.7.6 Étude de la localisation 140

5.8 Résultats 141

5.8.1 Étude de la prédiction dans l'image 141

5.8.2 Suivi des panneaux 146

5.8.3 Localisation des panneaux 146

5.9 Discussion 147

5.9.1 Efficacité de la prédiction 147

5.9.2 Suivi des panneaux 150

5.9.3 Localisation des panneaux 150

5.9.4 Vers des expérimentations réelles 151

5.9.5 Emploi d'autres techniques d'analyse d'images 152

6 Conclusion 154

Bibliographie 166

Liste des tableaux

2.1 Conditions d'éclairage pour les images de l'université de Koblenz . . . 23

2.2 Conditions d'éclairage pour les images du GRPR 28

2.3 Conditions de prise de vue pour les images du GRPR 28

4.1 Masque pour détecter les losanges 83

4.2 Masque mince pour détecter les cercles 84

4.3 Nombre de panneaux recherchés dans les images de l'université de Koblenz 87

4.4 Résultats pour tous les types de panneaux 90

4.5 Résultats de l'apprentissage 91

4.6 Résultats du module détection 91

5.1 Facteur de réduction par image moyen et maximal pour l'image de droite 141

5.2 Facteur de réduction par image moyen et maximal pour l'image de gauche 141

5.3 Pourcentages d'apparitions de panneaux manqués pour l'image de gauche. 146

5.4 Erreurs sur la position du panneau (en m), sans biais 146

Liste des figures

1.1 Exemples de panneaux employés au Québec. 6

2.1 Organisation du système d'inspection 17

2.2 Exemples d'images de scènes routières avec panneaux européens. . . . 24

2.3 Exemples d'images de scènes routières avec panneaux européens. . . . 25

2.4 Exemples d'images de scènes routières avec panneaux européens. . . . 26

2.5 Exemples d'images de scènes routières avec panneaux canadiens. . . . 29

2.6 Exemples d'images de scènes routières avec panneaux canadiens. . . . 30

2.7 Exemples d'images de scènes routières avec panneaux canadiens. . . . 31

2.8 Exemples d'images de scènes routières avec panneaux canadiens. . . . 32

2.9 Image utilisée pour illustrer l'algorithme de détection. 34

2.10 Agrandissement de la portion de l'image 2.9 contenant le panneau stop de gauche. 34

3.1 Le plan TS ou disque chromatique. 40

3.2 Cartes dans le plan TS pour les pixels rouges de panneaux au soleil, les cercles correspondent aux valeurs de saturation 0.1 et 1 (le maximum). 50

3.3 Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge. 51

3.4 Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge. 51

3.5 Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge. 52

3.6 Segmentation pour le rouge et le bleu. 54

3.7 Segmentation pour le rouge et le bleu. 55

3.8 Segmentation pour le rouge et le bleu. 56

3.9 Segmentation pour le rouge et le bleu. 57

3.10 Segmentation pour le jaune. 58

3.11 Segmentation pour le jaune. 59

3.12 Segmentation pour le jaune. 60

4.1 Implantation des opérations de morphologie mathématiques. 69

4.2 Exemple de Liaison entre un panneau et une région rouge produites à la suite de la segmentation. 70

4.3 Elimination des pixels parasites. 71

4.4 Image des contours des régions rouges obtenues. 72

4.5 Contours des régions rouges sélectionnés pour la reconnaissance de forme. 74

4.6 Problèmes dans les zones très saturées. 76

4.7 Suppression des régions inutilisables. 77

4.8 Seconde segmentation et pré-reconnaissance pour les zones très saturées. 78

4.9 Exemple de contour correspondant à un panneau de direction. 79

4.10 Problème de détections multiples. 85

4.11 Influence de l'éclairage sur la segmentation. 94

4.12 Fausses alarmes pour la détection des rectangles bleus en milieu urbain. 95

4.13 Erreur lors de la détection des cercles rouges 96

4.14 Détection de deux panneaux, celui du bas étant segmenté en deux régions, celui du haut étant fusionné à celui du haut. 99

5.1 Organigramme pour l'emploi du mouvement et de la stéréo. 115

5.2 Organigramme pour la recherche d'un panneau dans les images précédentes. 116

5.3 Organigramme pour la mise à jour de la liste de panneaux. 117

5.4 Les deux principaux repères utilisés 121

5.5 Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le premier trajet simulé. 142

5.6 Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le second trajet simulé. 143

5.7 Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le troisième trajet simulé. 144

5.8 Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le quatrième trajet simulé. 145

5.9 Distribution de l'erreur sur la position et l'erreur latérale sur la position des panneaux en fonction de la distance du panneau à la position de la caméra de droite au début de la poursuite. 148

B.1 Principe de la stéréoscopie. 165

Liste des annexes

A Les systèmes de positionnement par satellites (d'après (Hofmann-Wellenhof, Lichtenegger et Collins, 1997)) **161**

A.1 Principe 161

A.2 Défauts 162

B La stéréoscopie (d'après (Horn, 1986)) **164**

Liste des notations

- TSI : espace de représentation des couleurs teinte (T), saturation (S), Intensité (I) (en anglais HSI);
- RGB : espace de représentation des couleurs rouge (R), vert (G), bleu (B);
- YIQ, espace de représentation des couleurs du standard NTSC;
- $Nrgb$: espace de représentation des couleurs rouge, vert, bleu normalisé;
- Irg : espace de représentation des couleurs intensité (I), rouge (r), vert (g);
- $.^T$: exposant indiquant une transposée de matrice ou de vecteur;
- $\mathbf{t}$: centre de la terre et origine du repère terrestre;
- $._T$: indice indiquant qu'une variable est exprimée dans le repère terrestre;
- $\mathbf{o}$: centre optique de la caméra et origine du repère de la caméra;
- $._C$: indice indiquant qu'une variable est exprimée dans le repère de la caméra;
- $\mathbf{m}$: un point de l'espace;

- $_k$: indice faisant référence à l'instant k;
- $\mathbf{R}$: matrice de rotation;
- κ : roulis;
- τ : tangage;
- ω : azimut;
- θ : latitude;
- ϕ : longitude;
- $c\alpha$: cosinus de l'angle α;
- $s\alpha$: sinus de l'angle α;
- $\mathbf{0}_N$: vecteur nul de dimension N;
- $\mathbf{I}_N$: matrice identité de dimension N;
- $\mathbf{x} = (X, Y, Z, D)^T$: vecteur d'état formé de la position d'un panneau $(X, Y, Z)^T$ et d'une de ces dimensions D;
- $\mathbf{v} = (r, s, d)^T$: vecteur de mesure formé de la position $(r, s)^T$ et d'une des dimensions d d'un panneau dans l'image;
- $\mathbf{A}_k$: matrice de changement d'état pour l'instant k;
- g_k : fonction de mesure à l'instant k;

- $\mathbf{C}_k$: matrice jacobienne de la fonction de mesure à l'instant k;
- $\hat{\mathbf{x}}_{k|k}$: estimé du vecteur d'état à l'instant k;
- $\hat{\mathbf{x}}_{k|k-1}$: prédiction du vecteur d'état effectuée pour l'instant k;
- $\hat{\mathbf{v}}_{k|k-1}$: prédiction de la mesure effectuée pour l'instant k;
- $\mathbf{X}_{k|k-1}$: matrice de covariance de la prédiction du vecteur d'état pour l'instant k;
- $\mathbf{V}_{k|k-1}$: matrice de covariance de la prédiction de la mesure pour l'instant k;
- $E(m)$: espérance de la variable aléatoire m;
- σ : écart type.

Chapitre 1

Introduction

1.1 Contexte

1.1.1 Surveillance des panneaux routiers

Les panneaux routiers sont un des vecteurs d'information employés pour indiquer aux usagers de nombreux renseignements dont les consignes à respecter et les dangers potentiels présents sur la route où ils circulent. Comme tous les autres éléments de la signalisation routière (feux tricolores, marquages au sol ...) les panneaux contribuent à la sécurité et à la fluidité du transport sur les routes. Les coûts, humains et financiers, des accidents et des retards dans les transports sont élevés pour la société. L'efficacité de la signalisation repose sur l'intégrité et la cohérence des informations fournies. Cependant les panneaux se dégradent sous l'effet de diverses contraintes telles que les intempéries et la lumière, et peuvent être endommagés ou détruits suite

à un accident ou à un acte de vandalisme. Pour maintenir des conditions de circulation les meilleures possibles il est donc nécessaire de connaître l'état des panneaux routiers.

Pour qu'un panneau soit opérationnel, il faut qu'il soit visible suffisamment à l'avance quelles que soient les conditions d'éclairage, de jour comme de nuit. Afin de déterminer les éventuelles opérations de maintenance à réaliser, les services en charge de l'entretien de la signalisation contrôlent les caractéristiques suivantes (Baass et Dion, 1993) :

- la visibilité,
- la rétroréflectivité, soit la capacité à réfléchir la lumière dans la direction d'incidence, et donc à être bien visible sous l'éclairage des phares d'un véhicule,
- l'état du panneau,
- l'état du support,
- la conformité aux normes.

Ces contrôles peuvent être effectués sur place par un expert, visuellement ou à l'aide de matériel de mesure spécialisé, par exemple pour la rétroréflectivité. Il est aussi possible d'opérer depuis un véhicule circulant sur la chaussée, et muni d'un dispositif de prise de vue (caméra cinématographique ou vidéo) à des fins d'analyses ultérieures (Baass et Dion, 1993). Travailler depuis un véhicule en mouvement offre

en outre des conditions plus sécuritaires à l'expert. Le résultat de ces inspections est ensuite utilisé pour déterminer les opérations de maintenance à effectuer.

Les panneaux délivrent des messages visuels aux conducteurs. Le contrôle des conditions de délivrance de ces messages fait notamment appel à l'observation directe d'un expert. Un système de vision artificielle détectant et reconnaissant les panneaux est donc fort intéressant pour deux types d'applications. D'abord un tel système est utile pour automatiser certaines des opérations de surveillance de la signalisation que nous venons de présenter : cet aspect constitue l'objet du travail présenté ici. La détection et la reconnaissance automatique des panneaux permettent de réaliser des applications d'aide à la conduite: de telles applications ont été l'objet d'un certain nombre d'études, dans le cadre des recherches sur les véhicules intelligents, dont nous allons donner maintenant un bref aperçu.

1.1.2 Détection et reconnaissance des panneaux routiers pour des véhicules intelligents

La détection et la reconnaissance des panneaux routiers ont d'abord été étudiées pour les systèmes de transport intelligents, principalement dans le cadre de deux projets : PROMETHEUS (PROgraM for a European Traffic with Highest Efficiency and Unprecedented Safety) en Europe et IVHS (Intelligent Vehicle and Highway Systems) aux Etats-Unis. Le but de ces projets est d'améliorer l'efficacité et la sécurité de la circulation automobile. Il s'agit notamment d'offrir aux conducteurs une meilleure

perception de leur situation au sein du traffic ainsi que des informations disponibles et de compenser éventuellement leurs erreurs. On peut évoquer, à titre d'exemple, les problèmes de détection d'obstacles, de pilotage automatique et de vision nocturne. La reconnaissance automatique des panneaux permet de s'assurer que le pilote dispose de toute l'information fournie par la signalisation, ce qui implique un fonctionnement en temps réel, pour que chaque panneau soit signalé à temps. Les applications de détection et de reconnaissance des panneaux sont réalisables car les choix des couleurs, des formes et des emplacements des panneaux sont normalisés, de manière à ce que la signalisation soit uniforme sur l'ensemble du réseau routier : ainsi les conditions de perception des panneaux varient peu ce qui facilite leur détection. La section suivante porte donc sur les principes de la signalisation routière implantée à l'aide de panneaux.

1.2 Les panneaux routiers

1.2.1 Généralités

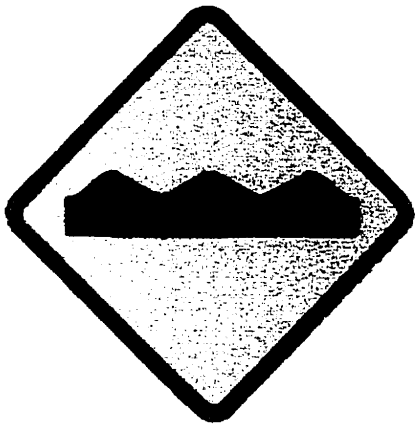
Il existe différents systèmes de signalisation routière utilisés dans le monde. Ces systèmes sont établis sur la base de principes généraux et ont les mêmes objectifs. Cela se traduit par des normes internationales. A l'intérieur d'un système, il existe des variations propres à chaque pays, mais elles ne sont pas fondamentales dans la mesure où les formes et les couleurs de base employées sont les mêmes. Notre intérêt

se porte principalement sur la signalisation employée au Québec et c'est de celle-ci dont il s'agit dans le reste de cette section; néanmoins la signalisation européenne a aussi été considérée dans cette étude.

Le répertoire des signaux rencontrés au Québec (Transport Québec, 1993) distingue quatre classes de panneaux :

- la signalisation de prescription et d'information,
- les signaux de danger,
- les signaux de travaux,
- les signaux d'indication.

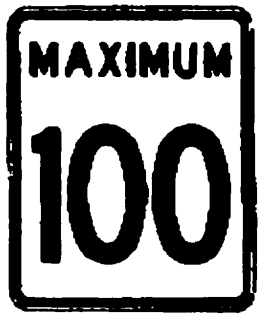
La figure 1.1 montre des exemples de panneaux de prescription et de danger utilisés au Québec. Sur ces cinq exemples, il est possible de voir comment le message est codé en associant un fond ayant une couleur et une forme particulière, et un pictogramme. Ces fonds sont l'octogone rouge, le losange jaune, et les rectangles noirs et blancs et sont caractéristiques des panneaux. Les pictogrammes peuvent être alphanumériques comme pour la limite de vitesse ou symboliques comme pour la chaussée cahoteuse. Excepté le cas de l'arrêt, les pictogrammes sont nécessaires à la compréhension du message et ne peuvent être interprétés seuls, comme le montre le signal avancé d'arrêt (on verrait alors un arrêt). La seconde partie de cette section va nous permettre de voir plus en détail l'utilisation des panneaux ainsi que les normes associées.



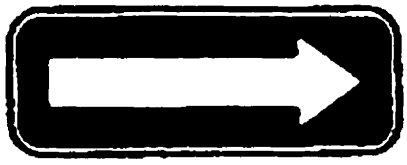
(a) Chaussée cahoteuse (danger).

(b) Signal avancé d'arrêt (danger)

(c) Arrêt (prescription)



(d) limite de vitesse (prescription).



(e) Sens unique (prescription).

Figure 1.1 : Exemples de panneaux employés au Québec.

1.2.2 Normes et emploi des panneaux au Québec (d'après Baass (1993))

La signalisation routière, et plus particulièrement, les panneaux obéissent à des normes précises afin que les conducteurs trouvent dans des situations semblables la même signalisation et puissent donc l'interpréter rapidement sans ambiguïtés. Le message est transmis par l'association de formes, de couleurs et d'un pictogramme (éventuellement de type alphanumérique). Les couleurs doivent normalement établir un contraste entre le panneau et le reste de l'environnement pour faciliter sa détection. Les formes et les couleurs employées doivent permettre de déterminer que l'on est en présence d'un panneau, ainsi que sa classe ou sous classe (c'est-à-dire le type de message) sans avoir à comprendre le message complet. Cela doit permettre de diriger au mieux l'attention des conducteurs sur le panneau. Les normes portent donc sur les couleurs, les formes, les symboles et les caractères employés, ainsi que sur les dimensions et l'emplacement au voisinage de la chaussée.

Les couleurs sont définies dans un système de représentation tri-chromatique associé à un blanc de référence. Chaque couleur correspond à une région dans l'espace des couleurs qui donne la plage de variation acceptable. Un nombre réduit de formes est employé. Les normes internationales spécifient un nombre idéal de cinq et un maximum de quinze. Les angles doivent être arrondis, sauf ceux des panneaux de signalisation installés en hauteur. Certaines formes comme l'octogone et le pentagone ne sont utilisées que pour un panneau. Il existe de même une liste de symboles que l'on

peut employer. Dans le cas de pictogrammes alphanumériques la taille des caractères est fonction de la plage de distances sur laquelle le panneau peut être lu.

Les dimensions et la position d'un panneau sont déterminées par le type de route, les contraintes de visibilité et le type du panneau. L'objectif est que le panneau soit visible suffisamment longtemps à l'avance pour que les conducteurs aient le temps de comprendre le message et de réagir en conséquence. Les cas réels pouvant être fort complexes, nous nous limitons à présenter les considérations suivantes :

- Il est souhaitable que le panneau puisse être lu en restant à l'intérieur d'un champ de vision de 10° horizontalement et de 6° verticalement, sans pour autant constituer une gène et un danger pour les usagers (piétons inclus).

- La plupart des panneaux sont situés à droite de la chaussée voire au dessus de celle-ci. Dans certaines situations (chaussée à sens unique par exemple) ils peuvent être répétés sur le bord gauche, mais un panneau ne peut être dédoublé au même endroit.

- L'orientation du panneau par rapport à l'axe de vision normal des conducteurs varie de quelques degrés par rapport à l'orthogonale en fonction de sa position par rapport à la chaussée.

- L'éloignement latéral doit être suffisant pour minimiser les risques de collision avec un véhicule sans toutefois compromettre la lisibilité. Des valeurs minimales et maximales sont définies suivant le type de route.

- Le type de panneau, le type de route et l'éloignement latéral du panneau se traduisent par des dimensions minimales pour ce panneau.

- La hauteur des panneaux par rapport à la chaussée, ainsi que par rapport au sol, est normalisée selon le type de route et ne peut sortir d'un intervalle de valeurs. Il s'agit là aussi d'un compromis entre la sécurité et la bonne lisibilité des panneaux.

1.2.3 Conclusion

La détection des panneaux repose donc sur la détection d'une couleur et d'une forme. L'identification finale d'un panneau repose sur l'association de ces deux informations au contenu du pictogramme. La réglementation sur la signalisation fixe des contraintes sur l'apparence des panneaux et les conditions sous lesquelles ils sont visibles depuis la chaussée, afin de faciliter la tâche des conducteurs : la détection automatique doit donc bénéficier de cette uniformité des situations rencontrées. Une fois un panneau détecté et reconnu, celui-ci ne peut être identifié sans ambiguïté qu'au moyen de sa position géographique. La section suivante porte donc sur les outils pour obtenir et manipuler cette information.

1.3 Obtention et gestion de la position des panneaux

Dans cette étude, pour des raisons de clarté, nous emploierons le terme positionnement dans un sens absolu: ainsi, nous parlerons de positionnement sur le globe terrestre. Nous emploierons le terme localisation dans un sens relatif et parlerons donc de localisation par rapport à un véhicule.

Chaque panneau est situé à un point précis dans l'espace et est en relation avec une route dont l'une des caractéristiques est une position géographique. De ce fait, les systèmes d'information géographique sont des outils intéressants pour gérer les informations relatives aux panneaux. Il s'agit d'ensembles de cartes géographiques numériques, chacune représentant un type information: cette information peut être à support continu, et prend alors une valeur en tout point, ou bien à support discret, et n'est définie que pour un point, une courbe ou une surface. Par exemple les panneaux et les routes sont associés respectivement à des points et des courbes particuliers sur la carte, tandis que l'altitude est une information à support continu prenant une valeur en chaque point de la carte. Ces systèmes d'information géographique permettent d'associer diverses informations dont la caractéristique principale est qu'elles sont reliées à une position géographique, et d'analyser ces données. Ces systèmes sont ainsi des outils très puissants pour la gestion de l'infrastructure routière. Il est possible d'associer des informations supplémentaires de n'importe quelle nature, des images par exemple, à des panneaux, ou à des routes dans une même base de données.

La position géographique d'un panneau est intéressante dans la mesure où l'on

est capable de l'estimer facilement. Les systèmes de positionnement par satellite permettent d'accéder aisément à cette information. Il existe actuellement deux systèmes de positionnement : NAVSTAR, le système américain, et son concurrent russe GLONASS. Ces systèmes ont été développés à des fins militaires. Ils sont couramment désignés sous l'abréviation GPS pour Global Positioning System (Hofmann-Wellenhof et al., 1997). Pour connaître la position d'un point, il est cependant nécessaire d'amener un récepteur GPS en ce point (l'annexe A explique le principe du positionnement par GPS).

La photogrammétrie élimine ce problème en permettant de localiser à distance un objet par triangulation (se reporter à l'annexe B pour plus de détails). Un véhicule embarquant un système de vision stéréoscopique et un dispositif donnant la position et l'orientation de ce système de vision permet, outre la surveillance classique, de déterminer la position des panneaux. La position et l'orientation du système stéréoscopique sont obtenues à l'aide d'une centrale de navigation inertielle et d'un récepteur GPS embarqués sur le véhicule.

1.4 Objectifs

Un certain nombre des opérations de contrôle des panneaux sont effectuées à partir d'images prises depuis un véhicule se déplaçant sur les routes. La vérification de la présence et de la bonne visibilité du panneau repose sur la détection de celui-ci dans ces images. Le but de ce projet est d'employer des techniques de vision artificielle

pour automatiser cette détection. Une telle inspection permet de faire rapidement un bilan de la situation et d'orienter les opérations de mesure sur le terrain (qui sont les plus onéreuses et les plus dangereuses pour ceux qui les effectuent) vers les panneaux qui en ont le plus besoin.

La détection du panneau indique que celui-ci est évidemment présent mais aussi en suffisamment bon état pour être reconnu comme tel. Plus un panneau ou son support se dégradent (déformations, modifications de l'orientation, écaillage de la peinture), plus le panneau est différent de ce que l'on cherche, et moins il est probable de le détecter. Un échec ne signifie pas obligatoirement un problème mais invite à se consacrer à un cas particulier : comme l'on travaille avec des images, celles-ci peuvent être utilisées par un opérateur pour établir un premier diagnostic.

En travaillant avec un système de vision stéréoscopique dont la position est connue pour chaque image prise, la détection d'un même panneau dans les images gauche et droite prises au même instant, nous permet de le localiser par photogrammétrie puis d'estimer sa position géographique. La position caractérise chaque panneau car en un point donné il ne peut y avoir deux panneaux (ou plus) identiques. La connaissance des résultats de l'inspection avec la liste des panneaux existants (sur la base du modèle et de la position) permet alors de déterminer les panneaux qui n'ont pas été détectés.

Un panneau est normalement visible le long d'une certaine portion de route. Il doit donc en théorie apparaître dans toutes les images prises durant le parcours de ce tronçon. En s'assurant que les images soient prises à des positions suffisamment

rapprochées pour pouvoir détecter un problème d'occlusion majeur, la ou les zones depuis lesquelles le panneau est visible peuvent être estimées. Ces informations peuvent alors être comparées à ce qui devrait être la norme. Cette zone de visibilité normale est déterminée par le type et les caractéristiques géométriques de la route où se trouve le panneau. En ayant détecté un panneau dans différentes images prises depuis des points de vue différents, on dispose d'informations supplémentaires pour améliorer la localisation de celui-ci.

1.5 Organisation du mémoire

Le second chapitre présente les données utilisées et décrit l'organisation du système développé. Le premier problème à résoudre est celui de la détection des panneaux présents dans une image. Parmi les critères qui permettent de distinguer les panneaux les deux plus importants (et les plus accessibles) sont la couleur et la forme. Il s'agit d'extraire des régions dont la forme et la couleur font qu'elles sont très probablement des panneaux; le problème de la segmentation basée sur la couleur est abordé au troisième chapitre préalablement à la détection de forme, qui est traitée au quatrième chapitre, en accord avec le mode opératoire de l'algorithme. Le cinquième chapitre traite du suivi et de la localisation des panneaux. L'intérêt du suivi des panneaux est d'abord démontré. Le problème fondamental du suivi est présenté ensuite, car il est requis pour la localisation.

Une revue des travaux dans le domaine ainsi que des techniques applicables au

problème est présentée en début de chacun des chapitres. Chaque chapitre se conclut par une présentation et une discussion des résultats expérimentaux obtenus pour le problème traité. Finalement un récapitulatif des résultats obtenus ainsi qu'une discussion sur les développements futurs constituent le dernier chapitre.

Chapitre 2

Description du système d'inspection et présentation des données

2.1 Le système d'inspection

Le système que nous présentons repose sur le principe d'une inspection automatisée depuis un véhicule circulant sur le réseau routier à une vitesse suffisamment élevée pour ne pas perturber le trafic. Cette solution permet de contrôler, aussi rapidement que le permet la vitesse de parcours, une portion du réseau sans que l'inspection ne crée de risques d'accidents supplémentaires pour les opérateurs et les autres usagers. Il est de plus possible de réaliser en même temps d'autres mesures : par exemple certaines caractéristiques de la chaussée telle que l'ornièrage et la qualité de roulement peuvent actuellement être estimées depuis un tel véhicule (Tessier, 1990).

L'architecture du système embarqué à bord du véhicule est décrite par le schéma bloc de la figure 2.1. Le système se compose des cinq modules suivants :

1. acquisition des séquences d'images,
2. positionnement du système de vision,
3. détection des panneaux,
4. poursuite des panneaux,
5. localisation des panneaux.

Pour chacun de ces modules, nous présentons maintenant les fonctions réalisées, leurs interactions avec les autres modules et les sous-modules supportant ces fonctions et interactions. Nous détaillerons ensuite les problèmes étudiés dans le cadre de ce projet.

2.1.1 Module d'acquisition d'images

Le rôle de ce module est de prendre des paires images stéréoscopiques de la route et de ses abords, tels qu'ils apparaissent à un conducteur. Il comporte donc un système de vision stéréoscopique formé de deux caméras vidéo, couplé à un système d'acquisition d'images. Comme les modules de poursuite et de localisation des panneaux utiliseront les images fournies par ce module à des fins de triangulation, il est nécessaire de déterminer la géométrie du système. En pratique cela implique de calibrer séparément

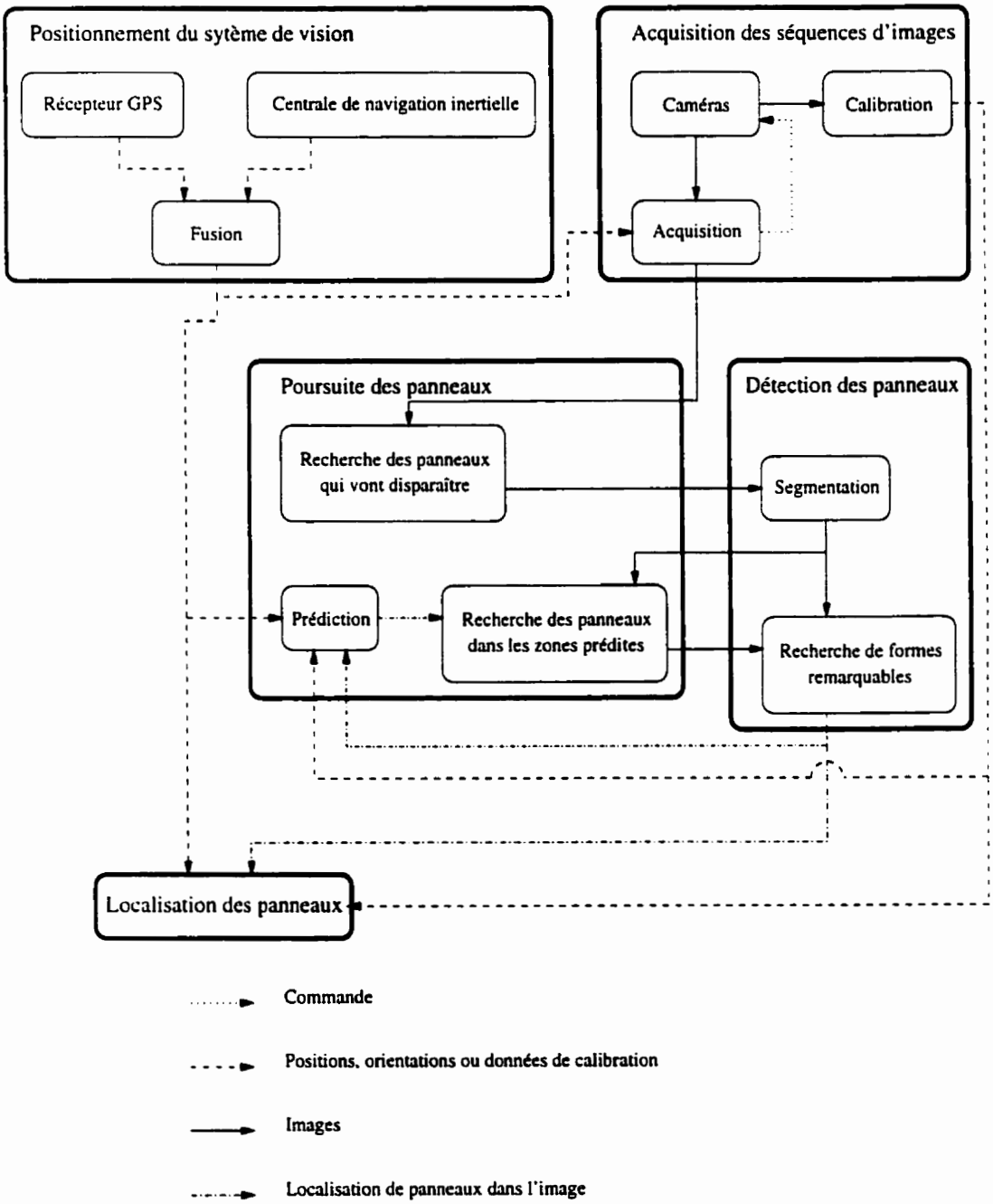


Figure 2.1 : Organisation du système d'inspection

les caméras afin de pallier les distorsions géométriques induites par l'objectif et le senseur ainsi que l'ensemble stéréoscopique pour connaître la position relative des deux caméras, et ce avant toute inspection. Pour éviter de fournir des séquences d'images statiques, le système d'acquisition utilise l'information de position fournie par le module de positionnement pour espacer les prises de vues.

2.1.2 Module de positionnement du système de vision

Le rôle de ce module est de fournir la position absolue et l'orientation des caméras du système de vision. Il comporte pour ces fins un récepteur GPS, une centrale de navigation inertielle et un dispositif pour fusionner les mesures produites par ces deux capteurs. Cette association permet de palier les défauts du récepteur GPS (inopérant dans les tunnels, imprécis entre des bâtiments élevés) et de la centrale inertielle (erreur augmentant avec le temps) et d'obtenir une bonne précision de positionnement (Sukkarieh, Nebot et Durrant-Whyte, 1998). Les positions et orientations estimées sont utilisées par les modules de poursuite et de localisation.

2.1.3 Module de détection des panneaux

Ce module est en charge de déterminer la position des panneaux dans les images. Ce problème est envisagé comme celui de la localisation de régions ayant une couleur et une forme particulières. Cette localisation est effectuée en deux étapes, chacune correspondant à un sous-module : une segmentation de l'image en régions basée sur la

couleur et une recherche des formes associées parmi les régions obtenues. Ce module est employé par celui de poursuite des panneaux dans deux circonstances différentes: la détection des panneaux avant qu'ils ne disparaissent du champ de vision, ce qui implique une recherche sur toute l'image, et la détection d'un panneau lors de sa poursuite qui implique une recherche sur une petite région de l'image. Dans ce dernier cas, la segmentation ayant déjà été effectuée, la recherche de forme est directement utilisée.

La segmentation opère sur des images de scènes extérieures prises dans des conditions d'éclairage variées. Son but est d'isoler les régions dont la teinte correspond à celle d'un panneau vu sous l'un des éclairages possibles. Elle requiert un apprentissage préalable dont l'objectif est de modéliser ces teintes. Ces modèles sont ensuite utilisés lors de la segmentation proprement dite pour classer les pixels, ce qui permet ensuite de former des régions connexes.

La recherche de forme a lieu dans chacune des images binaires produites par la segmentation et montrant les régions correspondant à une teinte. Elle se déroule en deux temps. Une analyse morphologique locale rejette les régions dont les frontières sont trop irrégulières pour correspondre à un panneau; ensuite, par corrélation, les formes recherchées sont localisées. Ces deux étapes nécessitent un apprentissage préalable. Ces apprentissages peuvent être couplés avec celui de la segmentation.

2.1.4 Module de poursuite des panneaux

Ce module a pour objectif de détecter toutes les apparitions d'un panneau dans les différentes images. Il comprend pour cela un sous-module chargé de détecter les panneaux avant qu'ils ne sortent du champ de vision, ce qui correspond à un emploi particulier du module de détection; ce sous-module permet d'initialiser la poursuite. Celle-ci est effectuée par l'enchaînement en boucle d'un sous-module prédisant la position du panneau et d'un sous-module le recherchant autour de cette position prédite; cette recherche correspond aussi à un cas particulier de la détection des panneaux. Ce sous-module de prédiction se base sur les détections préalablement effectuées du panneau et sur les positions et orientations des caméras lors de la prise de vue des images correspondantes. Le résultat de la détection des différentes apparitions d'un panneau permet alors de le localiser.

2.1.5 Module de localisation des panneaux

Le rôle de ce module est de localiser les panneaux par rapport à la position du véhicule à un instant donné (prise comme référence) puis de déduire leur position géographique absolue. Il intervient une fois que toutes les apparitions d'un panneau ont été repérées. Cela peut avoir lieu durant la phase d'inspection ou après celle-ci. En ayant localisé un panneau dans une image et en connaissant l'orientation et la localisation de la caméra par rapport à la position de référence du véhicule, à l'instant où cette image a été prise, on obtient une droite passant par le centre optique de la caméra et le panneau

réel. Cette droite est parfaitement connue. En utilisant l'ensemble des images où un panneau est visible, on obtient un ensemble de droites s'intersectant au panneau; nous obtenons ainsi un système d'équations dont les inconnues sont les coordonnées du panneau. Ce système est linéaire et très souvent surdéterminé, et est donc résolu en employant la méthode des moindres carrés. Alors par changement de repère, la position géographique du panneau est estimée.

2.1.6 Modules et sous-modules étudiés

Les deux premiers modules (système de vision et positionnement) ont déjà fait l'objet d'études ayant conduit à des réalisations commerciales : ce sont les systèmes d'arpentage depuis un véhicule tel que ceux de (*LambdaTech International*, n.d.) et (*VISAT Technologies Inc.* n.d.). Nous nous concentrons donc sur les autres aspects du problème à savoir la détection des panneaux, leur poursuite et leur localisation. Le sous-module de prédiction sera étudié avec le module de localisation car ils utilisent les mêmes données (positions et orientations des caméras, positions des panneaux dans les images). Les sous-modules formant la détection seront étudiés séparément. La fin de ce chapitre est consacrée à la présentation des images utilisées pour l'étude de la détection.

2.2 Données

Pour développer et évaluer notre algorithme de détection nous avons employé des images de scènes routières présentant des panneaux, prises en Europe et au Québec. Si, à quelques exceptions près, la signalisation européenne est différente, le problème est du même type, à savoir localiser une région de l'image ayant une forme et une couleur précises.

2.2.1 Images de l'université de Koblenz-Landau

Ces images sont disponibles sur le site ftp du laboratoire d'analyse d'images de l'université de Koblenz (`ftp://ftphost.uni-koblenz.de/outgoing/vision`). Elles sont au nombre de 248, réparties en 16 séquences présentant diverses conditions d'éclairage (voir le tableau 2.1). Elles ont été réalisées en Europe depuis un véhicule en mouvement. Les caméras vidéo employées acquièrent les champs formés respectivement des lignes paires et impaires avec un décalage de 0.02 s. Cela induit un bruit de mouvement d'une ligne sur l'autre. Pour éliminer ce problème nous employons des images formées d'un seul champ et donc avec une résolution verticale deux fois moins élevée et de ce fait paraissent aplaties. La définition est donc de 512 pixels pour la largeur et de 256 pixels pour la hauteur.

Ces images ont été prises avec une caméra ayant trois capteurs CCD, ce qui donne une très bonne définition des couleurs, sauf pour trois des séquences qui ont été prises sous la pluie; elles ne sont pas employées en raison de leur faible qualité.

Tableau 2.1 : Conditions d'éclairage pour les images de l'université de Koblenz

éclairage	Nombre de séquences
ciel couvert	5
ensoleillé	3
contre-jour	2
ombre	3
pluie	4
début ou fin du jour	1

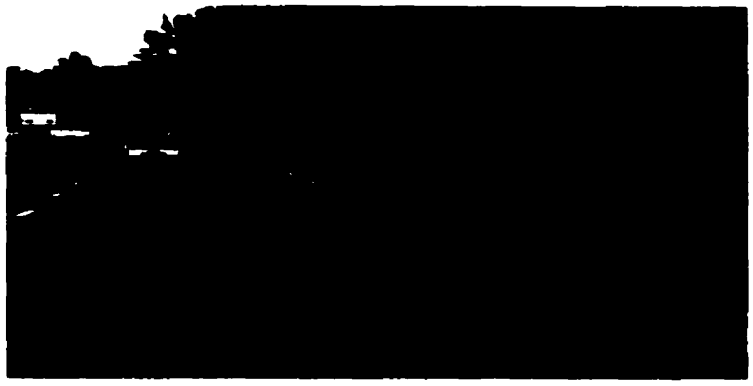
En pratique nous travaillons donc avec 186 images. Les figures 2.2 à 2.4 montrent des exemples de ces scènes routières avec des panneaux. Les images 2.2a et 2.2b permettent de juger de la différence entre une image prise avec une caméra à un seul capteur CCD et une image prise avec une caméra à trois capteurs CCD, dans des conditions de pluie. Ces images présentent également un éventail de problèmes rencontrés. Les images 2.2c 2.3a et plus particulièrement l'image 2.3b présentent des panneaux faiblement éclairés et donc contrastant faiblement avec l'arrière plan, ce qui rend les panneaux plus difficiles à distinguer que pour l'image 2.3c. Inversement le panneau de l'image 2.4a est éclairé directement par le soleil et présente un reflet où la teinte n'est pas discernable. L'image 2.4b montre le cas particulier des milieux urbains où l'on retrouve dans l'architecture de nombreuses lignes droites qui peuvent former des rectangles. Finalement l'image 2.4c montre un des plus petits panneaux reconnaissables.

(a) Image prise par temps de pluie, capteur à 1 CCD

(b) Image prise par temps de pluie, capteur à 3 CCD

(c) Image prise par temps couvert, capteur à 3 CCD

Figure 2.2 : Exemples d'images de scènes routières avec panneaux européens.



(a) Image présentant deux panneaux dans une zone d'ombre, capteur à 3 CCD

(b) Image prise à l'aube ou au crépuscule, capteur à 3 CCD

(c) Image prise par beau temps, capteur à 3 CCD

Figure 2.3 : Exemples d'images de scènes routières avec panneaux européens.

(a) Image présentant un panneau avec un reflet, capteur à 3 CCD

(b) Image prise en milieu urbain, capteur à 3 CCD

(c) Image présentant un petit panneau, capteur à 3 CCD

Figure 2.4 : Exemples d'images de scènes routières avec panneaux européens.

2.2.2 Images du Groupe de Recherche en Perception et Robotique

Ces images ont été prises au Québec. Elles sont au nombre de 249 et se divisent en 49 séquences prises dans des conditions d'éclairage variées (voir le tableau 2.2). Elles ont été obtenues avec un appareil photo 35mm puis numérisées. La numérisation a été effectuée à partir d'épreuves de 10 cm par 15 cm obtenues depuis un négatif avec un scanner plat. L'équilibrage des blancs n'est donc pas réalisé. Les photographies ont été prises soit à main levée à l'arrêt (pour la plupart) soit depuis l'intérieur d'un véhicule en mouvement (voir le tableau 2.3). La définition est de 586 pixels pour la largeur et de 392 pour la hauteur. Les figures 2.5 à 2.8 montrent des exemples de ces scènes routières avec des panneaux. L'image 2.7a a été prise depuis un véhicule en mouvement au travers du pare-brise. Ces images donnent un aperçu des conditions d'éclairage rencontrées et de leur effet sur la teinte des panneaux jaunes, ainsi que d'autres difficultés à surmonter. Ainsi certains panneaux disparaissent dans les feuillages (image 2.5b) et la chaussée présente des marquages au sol de la même couleur que les panneaux (image 2.8b).

Tableau 2.2 : Conditions d'éclairage pour les images du GRPR

Eclairage	Nombre de séquences
ciel couvert	18
ensoleillé	13
contre-jour	8
ombre	5
pluie	0
début ou fin du jour	5

Tableau 2.3 : Conditions de prise de vue pour les images du GRPR

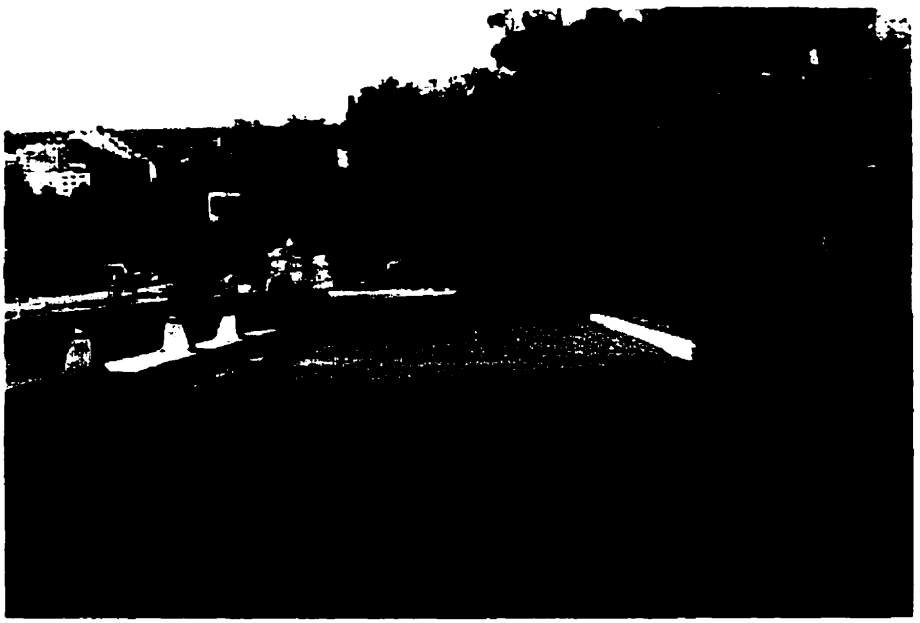
Type de prise de vues	nombre de séquences
en mouvement	8
à l'arrêt	41

(a) Panneaux directement éclairés par le soleil.

(b) Panneaux dans les feuillages

Figure 2.5 : Exemples d'images de scènes routières avec panneaux canadiens.

(a) Panneaux en contre-jour.



(b) Panneaux dans l'ombre.

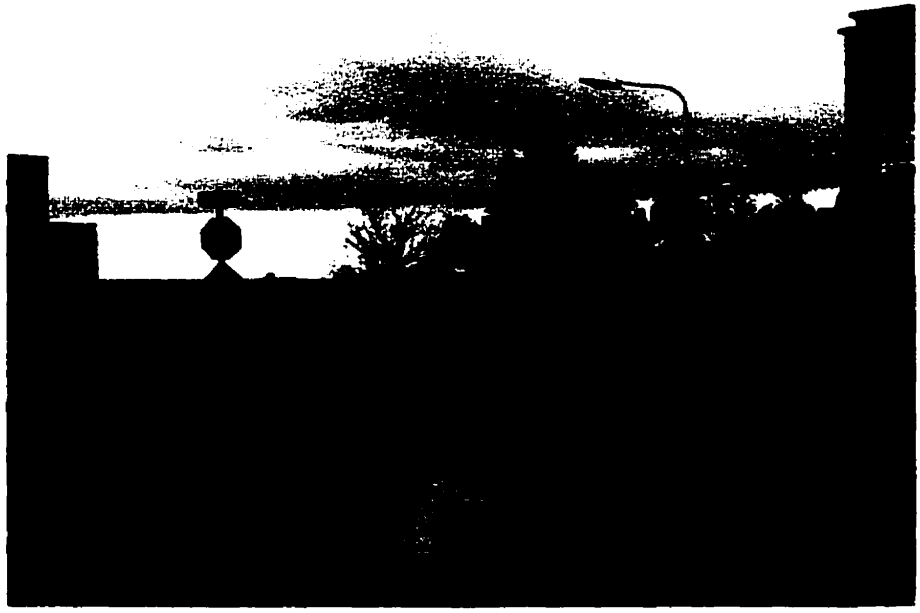
Figure 2.6 : Exemples d'images de scènes routières avec panneaux canadiens.

(a) Image prise au travers du pare-brise d'un véhicule en mouvement.

(b) Panneau sous l'éclairage rasant de fin de journée

Figure 2.7 : Exemples d'images de scènes routières avec panneaux canadiens.

(a) Image prise par temps couvert.



(b) Image prise par temps couvert avec des marquages au sol jaunes importants.

Figure 2.8 : Exemples d'images de scènes routières avec panneaux canadiens.

2.2.3 Utilisation des images et des séquences

Ces images font partie de séquences. Un même panneau est donc visible dans plusieurs images. Le suivi et la localisation de panneaux nécessitent des séquences d'images. Cet aspect est traité au cinquième chapitre. La détection des panneaux est en revanche considérée comme un problème portant sur une image isolée. Les notions de séquences et de panneaux réels sont laissées de côté et les images sont employés sans liens particuliers entre elles. Ces dernières considérations concernent les deux phases de la détection que sont la segmentation, exposée dans le chapitre qui suit et la recherche de forme traitée au quatrième chapitre. Nous allons illustrer les différentes étapes de l'algorithme de détection à l'aide de l'image de la figure 2.9, en nous intéressant plus particulièrement à la région autour du panneau arrêt de gauche présentée à la figure 2.10

Figure 2.9 : Image utilisée pour illustrer l'algorithme de détection.

Figure 2.10 : Agrandissement de la portion de l'image 2.9 contenant le panneau stop de gauche.

Chapitre 3

Segmentation basée sur la couleur

3.1 Introduction

La localisation d'un panneau dans l'image implique de rechercher les régions ayant des propriétés identiques, ou du moins similaires, à celles des panneaux. La couleur est le premier facteur qui permet de distinguer un panneau de l'arrière plan. Nous proposons donc de segmenter l'image en régions ayant ou n'ayant pas de couleur remarquable. Notre approche s'appuie sur le système TSI (Teinte, Saturation, Intensité, HSI en anglais pour Hue, Saturation, Intensity) de représentation des images en couleur. Cette représentation permet d'établir, dans le cadre de cette application, une classification des pixels par seuillage qui soit relativement robuste vis-à-vis de l'éclairage. Une méthode pour juger de l'efficacité de ce seuillage est également présentée. Cette technique de segmentation peut être implantée à l'aide d'une table

de mise à jour et donc être très rapide.

3.2 Revue bibliographique

3.2.1 Avantages et inconvénients de la couleur

Travailler avec la couleur plutôt qu'avec seulement l'intensité implique en général de traiter un plus grand volume de données. Les variations des conditions d'éclairage de la scène influent sur l'enregistrement des teintes comme sur celui de l'intensité. La seule connaissance des couleurs d'un objet inconnu ne permet pas non plus de fournir une description géométrique de celui-ci, mais dans le cas d'objets connus et plus particulièrement d'objets dont la ou les couleurs traduisent des propriétés de ces objets, elle constitue un très bon indice, indépendant de la résolution, pour la localisation de ces objets (Swain et Ballard, 1991). De plus la couleur fournit des indices qui sont indépendants de la géométrie de la scène (Healey, 1992). Dans le cas qui nous intéresse, la couleur est vecteur d'une signification (par exemple le rouge est souvent associé à une interdiction), et est aussi une caractéristique distinguant le panneau du reste de l'environnement (Janssen, Ritter, Stein et Ott, 1993).

3.2.2 Représentation des couleurs

La notion de couleur correspond d'abord à une sensation qui varie selon les individus, (certains, daltoniens, n'étant pas capables de différencier certaines ou toutes

les couleurs) et est décrite par les notions de teinte, de brillance et de saturation; en outre la perception des couleurs est relative car elle dépend de l'adaptation à l'illumination et du contenu de la scène (Pratt, 1991). L'œil possède deux types de photo-senseurs : les bâtonnets et les cônes. Il existe trois types de cônes sensibles à différentes longueurs d'onde et qui permettent la perception des couleurs (Jähne, 1997).

La colorimétrie repose sur le modèle tri-chromatique, c'est-à-dire la décomposition d'une couleur sur trois couleurs primaires. Ce modèle se justifie, outre le nombre de sortes de cônes, par la possibilité de décrire une sensation lumineuse à l'aide de trois sources lumineuses différentes. Cette décomposition donne une représentation vectorielle des couleurs. La procédure à suivre, pour obtenir cette décomposition, est de créer deux couleurs donnant des sensations lumineuses identiques en superposant (synthèse additive) des sources de lumières sur un écran blanc. la première couleur correspond soit à la couleur inconnue, soit à un mélange de celle-ci avec une ou deux des trois sources de références. La seconde couleur correspond à un mélange des sources de références pas encore employées. Dans le cas particulier où l'on additionne les trois sources de référence, les trois coordonnées obtenues sont toutes positives et cette décomposition peut aussi être obtenue avec trois capteurs dont la sensibilité spectrale correspond au spectre de la couleur primaire associée (Pratt, 1991). Les capteurs employés donnent une décomposition sur trois couleurs qui sont le rouge, le vert et le bleu : il s'agit de l'espace RGB. L'espace ainsi défini a été normalisé par la

Commission Internationale de l'Éclairage (CIE) (Pratt, 1991). Les longueurs d'onde des trois couleurs primaires sont fixées aux valeurs suivantes :

- pour le rouge : 700 nm,
- pour le vert : 546,1 nm,
- pour le bleu : 435,8 nm,

et le blanc de référence obtenu doit être équiénergétique. Si la décomposition des couleurs dans ce système est directement accessible, il présente néanmoins deux inconvénients : les valeurs de chaque canal sont difficiles à interpréter et sont fortement corrélées entre elles, les fluctuations de l'intensité lumineuse ont ainsi un effet sur chacune (Baghdassarian, 1995). Par une transformation linéaire, le système de transmission employé en télévision sépare la luminance de la chrominance. Pour le standard nord-américain (NTSC) la transformation est la suivante (Pratt, 1991) :

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,177 & 0,813 & 0,011 \\ 0,540 & -0,263 & 0,174 \\ 0,246 & -0,675 & 0,404 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}.$$

où Y est la luminance normalisée pour un œil humain moyen et I et Q donnent la chrominance. Baghdassarian (1995) considère toutefois que cette transformation n'élimine pas les problèmes de corrélation entre les canaux, ni ne facilite l'interprétation. Une autre approche pour gagner en robustesse vis-à-vis des variations

de l'intensité lumineuse réside dans l'emploi de transformations non-linéaires. La normalisation des valeurs R, G et B par la norme L^1 donne l'espace $Nrgb$ qui est indépendant des variations uniformes de l'intensité (Perez et Koch, 1994) :

$$
\begin{aligned}
r &= \frac{R}{R+G+B}, \\
g &= \frac{G}{R+G+B}, \\
b &= \frac{B}{R+G+B}.
\end{aligned}
$$

En pratique seules deux des composantes sont indépendantes. On leur adjoint donc l'intensité pour obtenir l'espace Irg (Janssen et al., 1993). Comparé aux autres systèmes de représentation, l'espace Teinte, Saturation, Intensité (TSI) reflète mieux la perception humaine des couleurs (Baghdassarian, 1995) :

$$
\begin{aligned}
T &= \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right) \\
S &= 1 - \frac{\min(R,G,B)}{I} \\
I &= \frac{R+G+B}{3}
\end{aligned}
$$

Malgré la singularité de T pour $S = 0$ et son instabilité au voisinage de cette valeur, Perez et Koch (1994) ont montré que cette variable était indépendante des ombres, de l'ombrage et des reflets quand l'équilibrage des blancs était correctement effectué. Priese et al. (1994) indiquent que les segmentations de scènes extérieures basées sur la teinte et la saturation sont robustes aux conditions d'éclairage.

Le sous espace TS est utile dès que l'on ne se préoccupe pas de l'intensité (voir figure 3.1).

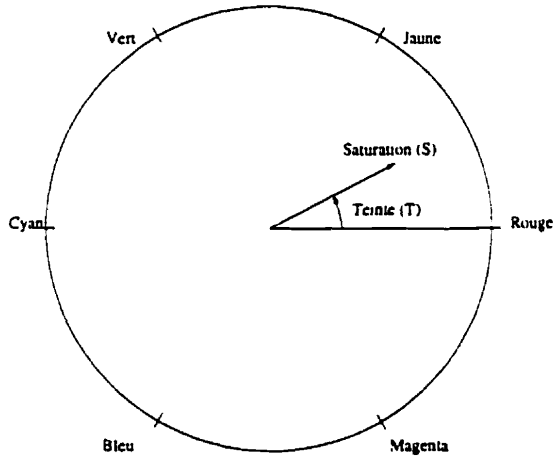


Figure 3.1 : Le plan TS ou disque chromatique.

3.2.3 Segmentation des panneaux routiers

Cette opération doit permettre d'isoler les panneaux du reste de l'image. Deux types d'approches sont possibles. L'approche par contours cherche à déterminer les fron-

tières entre des régions sur la base de différences locales. L'approche par régions cherche à obtenir des régions connexes sur lesquelles un critère d'uniformité est vérifié.

3.2.3.1 Approche par contours

Piccioli, de Micheli, Parodi et Campani (1994) ont développé un algorithme de détection s'appuyant sur les arêtes d'intensité. Des variations importantes de l'intensité ne sont cependant pas nécessairement révélatrices d'un changement d'objet et peuvent correspondre à un reflet ou à une ombre. Le problème inverse peut se produire, à savoir que le contraste d'intensité entre un panneau et l'environnement ne soit pas suffisamment élevé pour pouvoir être discriminé. L'emploi d'un filtre rouge (de Saint Blancard, 1992) peut augmenter le contraste mais seulement pour cette couleur. L'emploi d'un détecteur d'arêtes fonctionnant sur des images multi-spectrales tel que celui de Cumani (1991) permet d'exploiter les contrastes de couleurs en plus de ceux d'intensités. Une solution pour n'obtenir que les arêtes correspondant à un contraste de couleur est d'utiliser la teinte (Perez et Koch, 1994) : cette méthode ne peut cependant détecter que les arêtes associées à des variations simultanées de l'intensité et de la teinte dans les zones faiblement saturées: en pratique cela correspond à des saturations inférieures à 0,2 ce qui limite le domaine d'application.

3.2.3.2 Approche par régions

Il existe de nombreuses techniques pour partitioner une image en régions uniformes suivant le type d'image et les objectifs poursuivis (Haralick et Shapiro, 1993); nous ne considérons dans cette brève revue de la littérature que les techniques ayant été employées pour la détection des panneaux ainsi que certaines plus spécifiques basées sur la couleur.

La technique de l'étiquetage qui consiste à classer les pixels en fonction de leur seule valeur, a été proposée par plusieurs auteurs (Baghdassarian, 1995; Estevez et Kehtarnavaz, 1996; Janssen et al., 1993; Li et Lalonde, 1995) : la segmentation est alors immédiate. Baghdassarian (1995) définit une classe par les deux teintes extrêmes et la saturation minimale dans le plan (teinte, saturation). Plutôt que de chercher à englober toutes les valeurs possibles dans une seule classe, Estevez et Kehtarnavaz (1996) l'adaptent en fonction de l'éclairement en estimant l'intensité moyenne du ciel; l'illumination des panneaux n'est cependant pas corrélée à celle du ciel. Le principal intérêt de cette technique de segmentation est sa simplicité. Janssen et al. (1993) considèrent cependant qu'elle présente des faiblesses pour des images prises par temps de pluie ou lors de conditions d'éclairage particulières telles que des contre-jours et dans le cas des panneaux achromatiques et, re-segmentent par conséquent les régions dignes d'intérêt par division puis fusion. La détermination des agrégats correspondants aux teintes recherchées, dans l'espace YIQ (celui-ci étant un des plus adaptés pour différencier les deux teintes cherchées), permet ensuite d'isoler dans l'image les régions

caractérisées par ces teintes (Kehtarnavaz et Ahmad, 1995). La détection des petits panneaux (donc fournissant peu de points dans l'espace des couleurs) n'est alors pas forcément aisée. Ces techniques d'étiquetage présentées ici définissent une segmentation depuis un point de vue global et sont à même d'échouer dans des situations du type "rouge sur fond rouge".

Pour éviter ce genre de problèmes, Janssen et al. (1993) proposent de segmenter par division puis fusion en estimant l'uniformité à l'aide de sa matrice de covariance dans l'espace RGB. Cette mesure indique de plus comment partitioner les régions inhomogènes. Les auteurs jugent cependant cette méthode trop complexe pour pouvoir l'appliquer en temps réel à toute l'image. Priese, Rehrmann, Schian et Lakmann (1993) ont développé un algorithme de croissance hiérarchique de régions autorisant seulement la fusion entre régions dont les dimensions sont du même ordre de grandeur afin de limiter l'apparition de chaînages. Cette approche requiert cependant un maillage hexagonal.

A l'opposé de l'approche de Perez et Koch (1994) qui consiste à employer la variable de haut niveau qu'est la teinte, un modèle physique de la couleur peut être utilisé pour juger de l'uniformité d'une région (Klinker, Shafer et Kanade, 1990). Les limites de la méthode sont celles du modèle : les objets de la scène doivent être en plastique et les interréflections ne sont pas considérées, ce qui empêche l'application de la méthode à des scènes extérieures. Healey (1992) fait aussi appel à un modèle physique, nettement plus simple, et utilise un modèle de décision bayésien. Il con-

vient également de signaler sa technique de segmentation combinant la technique de division-fusion et la segmentation hybride. Baghdassarian (1995) emploie aussi une méthode bayésienne pour la segmentation de la chaussée dans des scènes extérieures. Il faut cependant disposer d'une bonne estimation des probabilités conditionnelles a priori pour chaque classe.

3.2.4 Implantation

La rapidité du système développé est un facteur important; ainsi Estevez et Kehtarnavaz (1996) ont mis l'accent sur la simplicité afin de pouvoir traiter plusieurs images par seconde. Le calcul de la teinte est un opération complexe : Perez et Koch (1994) ont développé un circuit VLSI analogique pour réduire ce temps de calcul. Ce circuit ne donne cependant pas la teinte de manière exacte et est destiné à être utilisé avec des circuits analogiques de segmentation spécifiques. Du point de vue logiciel, la région définie dans le plan (teinte, saturation) correspond dans l'espace RGB à une région comprise entre deux plans (Baghdassarian, 1995). De plus, toute segmentation par étiquetage peut être implantée efficacement à l'aide d'une table de recherche dans l'espace RGB, au prix toutefois du coût en mémoire requis par la table (Janssen et al., 1993). Des techniques de segmentation plus complexes sont nécessairement plus coûteuses en temps de calcul. Le développement d'algorithmes facilement parallélisables associés à des implantations spécifiques permet de limiter le problème (Priese et al., 1993; Priese, Klieber, Lakmann, Rehrmann et Schian, 1994).

3.2.5 Conclusion

La couleur est une information fort utile pour segmenter puis détecter les panneaux routiers. A moins de ne s'intéresser qu'à seulement un ou deux types de panneaux, il semble plus simple de choisir un système de représentation ne privilégiant aucune couleur en particulier. Le système teinte, saturation, intensité semble bien approprié à la recherche des panneaux pour les raisons suivantes : ceux-ci ayant des teintes saturées, la teinte délivre une information fiable, elle offre ensuite une bonne robustesse vis-à-vis des conditions variées d'éclairage et, finalement, cet espace offre une description proche de la description humaine des couleurs pour lesquels les panneaux ont été conçus. Il faut aussi noter qu'un éventuel problème de constance de couleur ne peut être résolu avec seulement un système de représentation.

La saturation élevée des panneaux incite à penser qu'une approche par contours basée sur la teinte pourrait donner de bons résultats sans régularisation; c'est cependant aux frontières que la saturation est la plus faible, en raison de la quantification spatiale. En contrepartie une approche région permet d'accéder aisément à la teinte des régions obtenues, ce qui est utile pour les analyses ultérieures. La dernière opération à réaliser est l'attribution d'une teinte à chaque région pour laquelle cela a du sens. Cette opération est similaire à l'étiquetage d'un pixel : il s'agit d'une décision d'ordre absolue. Ainsi on aurait pu classer correctement les pixels de ces régions par étiquetage. C'est pourquoi la segmentation par étiquetage, comparée à une autre méthode ne risque pas plus de manquer les régions correspondant aux panneaux mais

de les fusionner avec l'arrière plan (avec lequel elles doivent normalement contraster): cette méthode est très rapide et donc très intéressante.

La théorie de la décision bayésienne peut être employée pour classer des pixels et segmenter des images sur la base de la couleur. Cela implique cependant de pouvoir estimer les probabilités de chaque classe ainsi que les probabilités conditionnelles a priori associées.

Il convient de signaler que les teintes achromatiques (noir, blanc, gris) sont difficiles à distinguer les unes des autres dans l'image sans connaissance a priori sur la scène.

3.3 Solution proposée pour la segmentation

3.3.1 Principe

Le but est d'extraire dans des images couleur les régions correspondant aux panneaux. Nous ne traitons pas ici du cas des panneaux achromatiques. L'illumination de la scène est complexe et varie avec la géométrie de celle-ci (interréflections, ombres, reflets), les conditions météorologiques (beau temps, ciel couvert ...) et le moment de la journée (épaisseur de l'atmosphère filtrant le rayonnement solaire). Les réglages de la caméra influent sur la quantité de lumière qui atteint le senseur. Afin de s'affranchir autant que possible de ces variations, nous passons de l'espace RGB vers l'espace TSI, et nous nous restreignons au plan TS.

Les images sont segmentées par étiquetage. Nous formons des classes de pixels en

nous appuyant sur les constatations suivantes. La teinte pour un type de panneau va prendre ces valeurs dans un intervalle dont les bornes dépendent des normes sur le panneau et des spectres possibles de l'illumination. Un équilibrage des blancs réduit l'ampleur de cet intervalle en normalisant l'éclairement. Le même raisonnement est possible pour la saturation: en pratique la limite supérieure pour la saturation est le maximum possible. Pour chaque couleur remarquable on dispose donc d'une plage de valeurs possibles pour la teinte et la saturation. Ce triple seuillage définit des régions du type secteur angulaire de couronne dans le plan TS. Cela permet de classer les pixels comme faisant soit partie de l'arrière plan, soit des régions ayant une teinte remarquable.

3.3.2 Choix des seuils

Les panneaux occupent une faible portion de l'image: il est donc difficile d'estimer ces seuils pour chaque image. Il faut par conséquent procéder à l'aide d'un ensemble d'apprentissage suffisamment représentatif des situations rencontrées. Un choix de seuils définissant une segmentation peut être évalué seul ou avec l'ensemble du processus de détection. Dans ce cas, une première évaluation visuelle qualitative permet d'estimer des seuils conduisant à des performances correctes.

Afin d'évaluer quantitativement une segmentation, il est impératif de disposer d'une classification de référence. Idéalement il est souhaitable d'en avoir plusieurs, produites par des opérateurs différents et prenant en compte tous les pixels de toutes

les images de l'ensemble; en pratique, en raison du nombre de pixels à classer que cela représente, on peut se limiter au voisinage de chaque panneau.

Une comparaison entre le produit d'une segmentation et un classement de référence peut ne pas révéler de mauvaises performances : ainsi il est impossible de dire si les pixels mal classés appartiennent tous à la même image ou à quelques panneaux. Il convient donc de juger des performances d'abord à l'échelle de chaque panneau, puis pour l'ensemble de ceux-ci. Pour une couleur, pour chaque voisinage associé à un panneau, on obtient deux taux de pixels mal classés. Une segmentation peut donc se voir caractériser par deux histogrammes, montrant pour différents intervalles de valeurs de taux, le nombre de panneaux dans ce cas.

3.4 Expériences

Nous disposons de deux groupes d'images. Dans les images prises en Europe, les zones rouges des panneaux de danger et d'interdiction et les zones bleues des panneaux d'obligation et d'indication ont été recherchées. Dans les images prises au Québec, le fond jaune des panneaux de danger a retenu notre attention. Pour ces trois teintes, les seuils ont été estimés visuellement : pour le rouge et le bleu, cela a été fait avec l'ensemble des images, pour le jaune avec un sous-ensemble des images disponibles afin de permettre une évaluation complète du processus de détection dont la segmentation constitue la première étape.

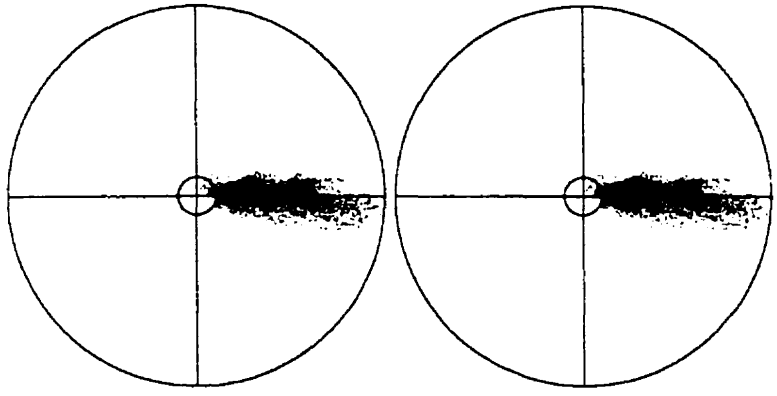
Une classification manuelle a été réalisée pour le rouge. Cela a permis de comparer

des segmentations reposant sur un choix de seuils voisins de ceux définis visuellement.

3.5 Résultats

La figure 3.2 présente la distribution des pixels considérés comme rouges par un opérateur dans le voisinage des panneaux rouges au soleil. La présence est une carte binaire montrant les valeurs dans le plan TS prises par au moins un pixel rouge; La densité est une carte dont les niveaux de rouge montrent le nombre de pixels ayant pris une valeur particulière dans le plan TS. Cette opération n'a pas été effectuée pour tous les panneaux car dans les zones de faible intensité et de faible saturation, il est difficile de classer les pixels un par un, même si l'opérateur est parfaitement capable de distinguer et de reconnaître le panneau. 6% des panneaux sont dans ce cas.

La figure 3.3 montre les histogrammes pour les pixels manqués et ceux détectés en trop pour la segmentation définie visuellement par un humain par rapport à cette classification des pixels. Les figures 3.4 et 3.5 montrent les mêmes histogrammes obtenus pour deux autres segmentations définies par des seuils dont les valeurs sont proches de celles employées pour la première segmentation. Les niveaux de performances atteints sont très semblables. L'agrandissement de la zone définissant le rouge conduit à une légère amélioration pour les pixels manqués ainsi qu'à une légère dégradation pour les pixels en trop. Sur les histogrammes une amélioration se traduit par un déplacement de la courbe vers la gauche, une dégradation par un déplacement de



(a) Carte témoignant de la présence.

(b) Carte témoignant de la densité.

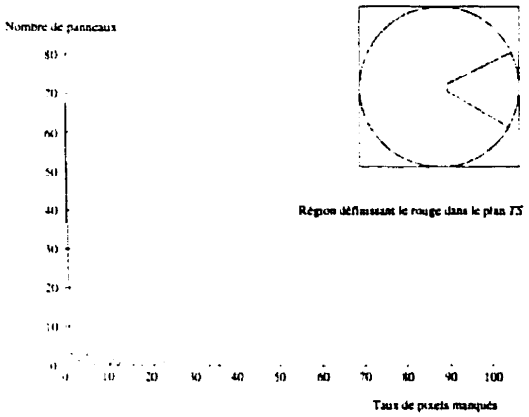
Figure 3.2 : Cartes dans le plan TS pour les pixels rouges de panneaux au soleil, les cercles correspondent aux valeurs de saturation 0.1 et 1 (le maximum).

la courbe vers la droite.

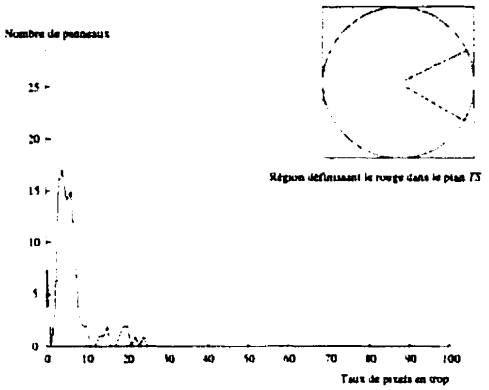
La figure 3.6 montre les régions rouges et bleues obtenues pour l'image de la figure 2.9. Outre les zones rouges et bleues des cinq panneaux, un certain nombre d'autres régions ont été trouvées dans l'arrière plan. Les figures 3.7 à 3.12 présentent d'autres exemples de résultats de la segmentation. Les couleurs recherchées sont le rouge et le bleu pour les images des figures 3.7 à 3.9 et le jaune pour les images des figures 3.10 à 3.12. La section suivante est consacrée à l'analyse de ces résultats.

3.6 Discussion

La comparaison entre les performances de la segmentation définie visuellement pour les panneaux rouges, et celles des segmentations définies par des seuils dont les valeurs

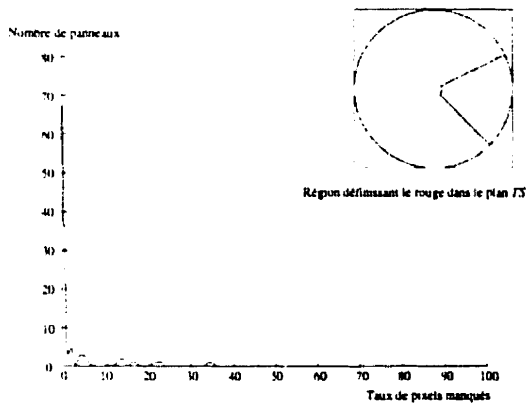


(a) Histogramme des panneaux en fonction du taux de pixels manqués.

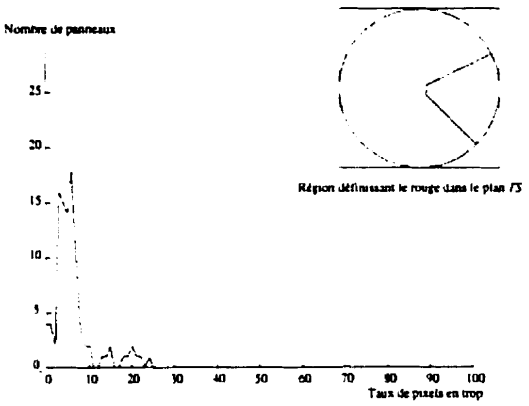


(b) Histogramme des panneaux en fonction du taux de pixels en trop.

Figure 3.3 : Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge.

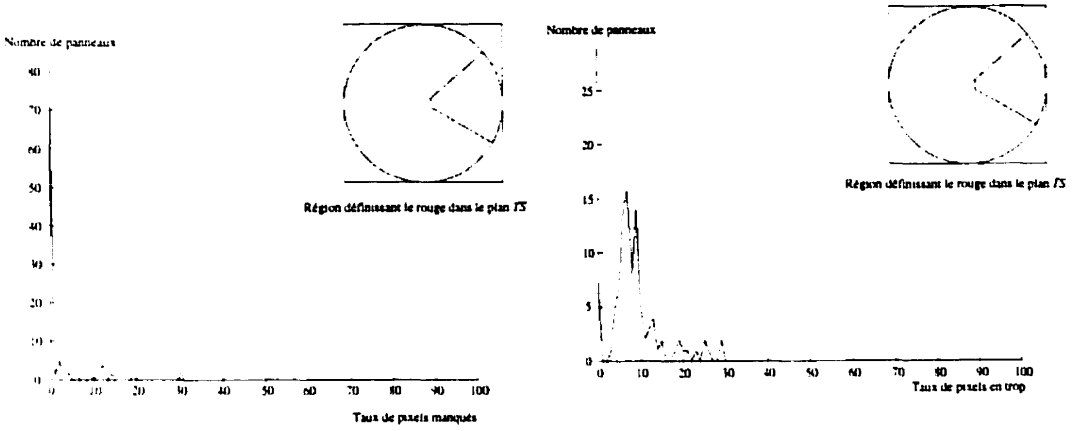


(a) Histogramme des panneaux en fonction du taux de pixels manqués.



(b) Histogramme des panneaux en fonction du taux de pixels en trop.

Figure 3.4 : Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge.



(a) Histogramme des panneaux en fonction du taux de pixels manqués.

(b) Histogramme des panneaux en fonction du taux de pixels en trop.

Figure 3.5 : Segmentation des régions rouges : la zone inscrite dans le plan TS définit le rouge.

sont proches de celles employées pour la première segmentation montre que l'on arrive à des résultats similaires par les deux méthodes. Il convient de signaler que la classification des pixels et le choix des seuils de manière visuelle ont été réalisés par la même personne. Classer des pixels d'une image selon leur teinte est une opération difficile quand la saturation ou l'intensité diminuent. Cela est donc problématique pour prendre en compte les panneaux situés dans l'ombre, ainsi que pour fixer un seuil pour la saturation.

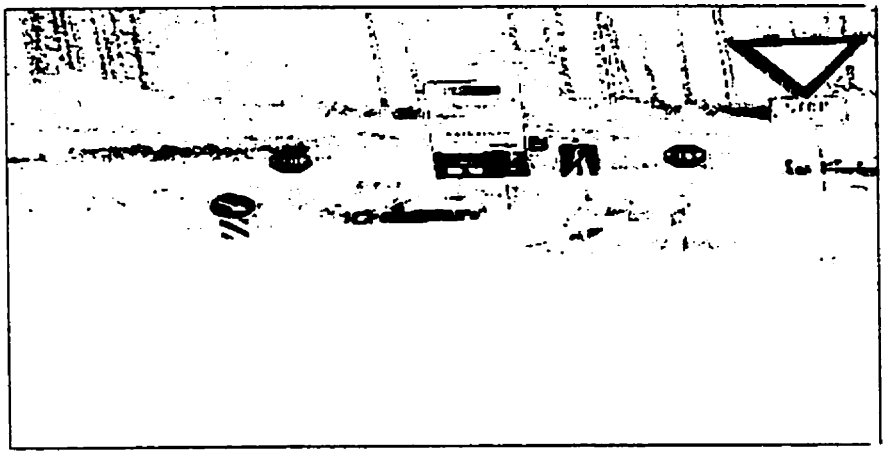
L'étiquetage des pix

La technique développée pour juger de l'efficacité d'une méthode de segmentation a pour principal défaut d'accorder la même importance à tous les pixels composant

un panneau. En fait, les pixels se trouvant aux frontières de la région définissent la forme, information que l'on souhaite exploiter par la suite. En outre, étant situés dans une zone de transition leur saturation est plus faible que celle des autres pixels et leur teinte s'écarte de la valeur moyenne. Leur classification est donc plus sensible aux variations des seuils.

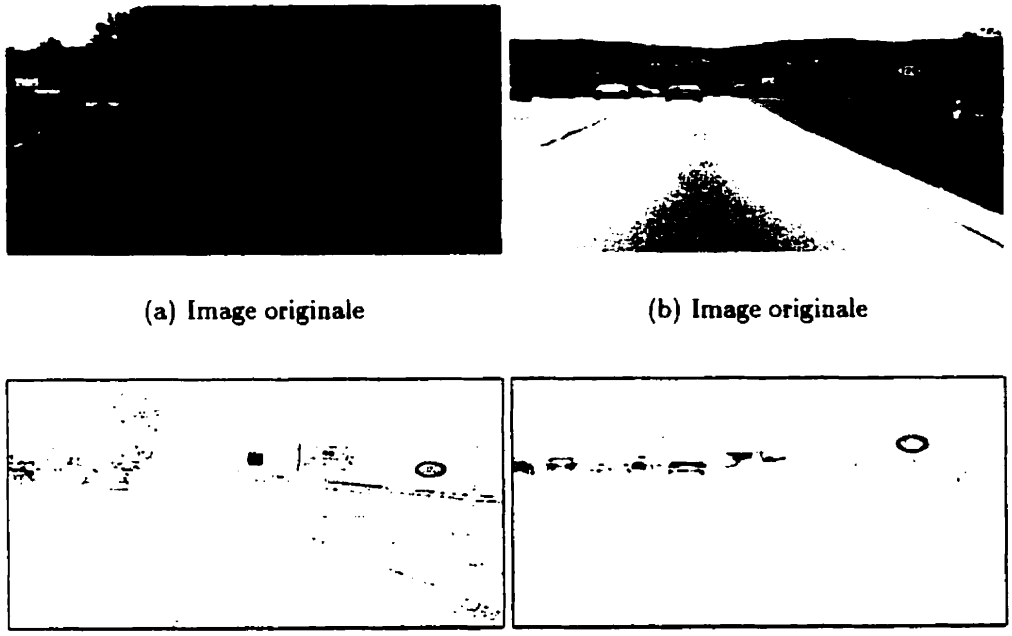
Dans la mesure où l'on doit être capable de trouver les zones de l'image correspondant aux panneaux, dans différentes conditions d'éclairage, des régions de l'image ne correspondant pas à des panneaux ont été segmentées; des fusions entre les panneaux et l'arrière plan peuvent aussi se produire. La segmentation seule ne suffit donc pas. C'est pourquoi elle est complétée par une recherche basée sur la forme exposée au chapitre suivant.

(a) Image originale



(b) Régions rouges et bleues produites par la segmentation

Figure 3.6 : Segmentation pour le rouge et le bleu.



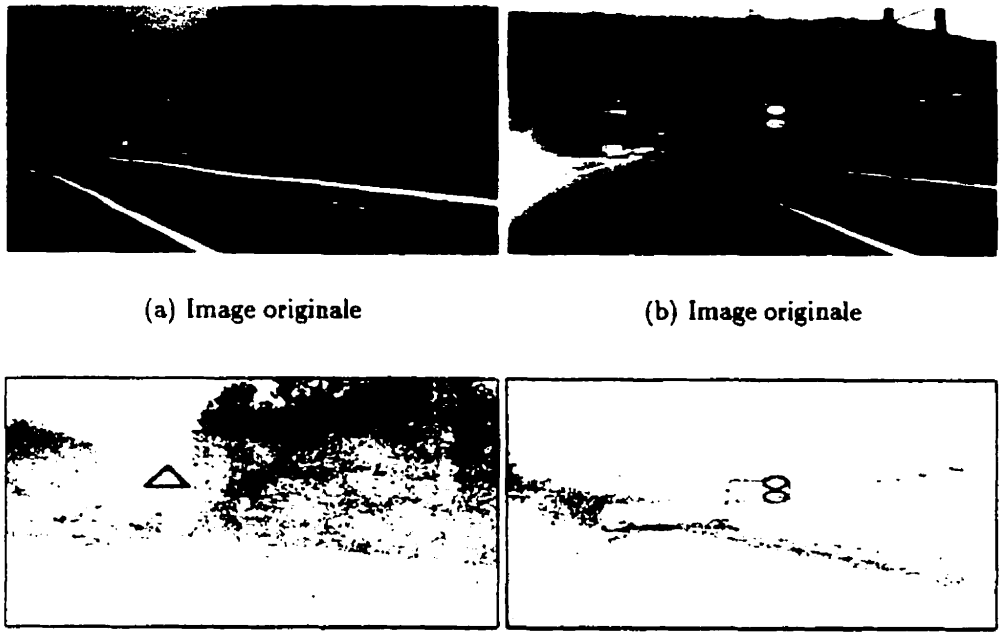
(a) Image originale

(b) Image originale

(c) Régions rouges et bleues produites par la segmentation

(d) Régions rouges et bleues produites par la segmentation

Figure 3.7 : Segmentation pour le rouge et le bleu.



(a) Image originale

(b) Image originale

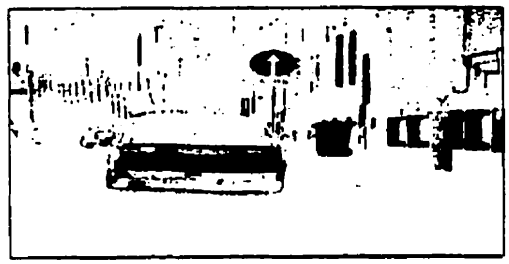
(c) Régions rouges et bleues produites par la segmentation

(d) Régions rouges et bleues produites par la segmentation

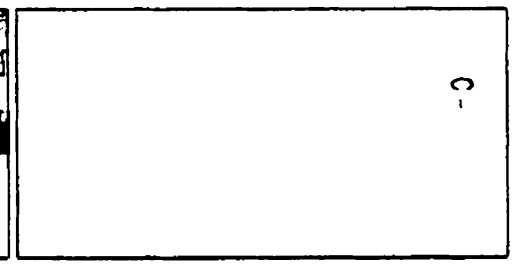
Figure 3.8 : Segmentation pour le rouge et le bleu.

(a) Image originale

(b) Image originale

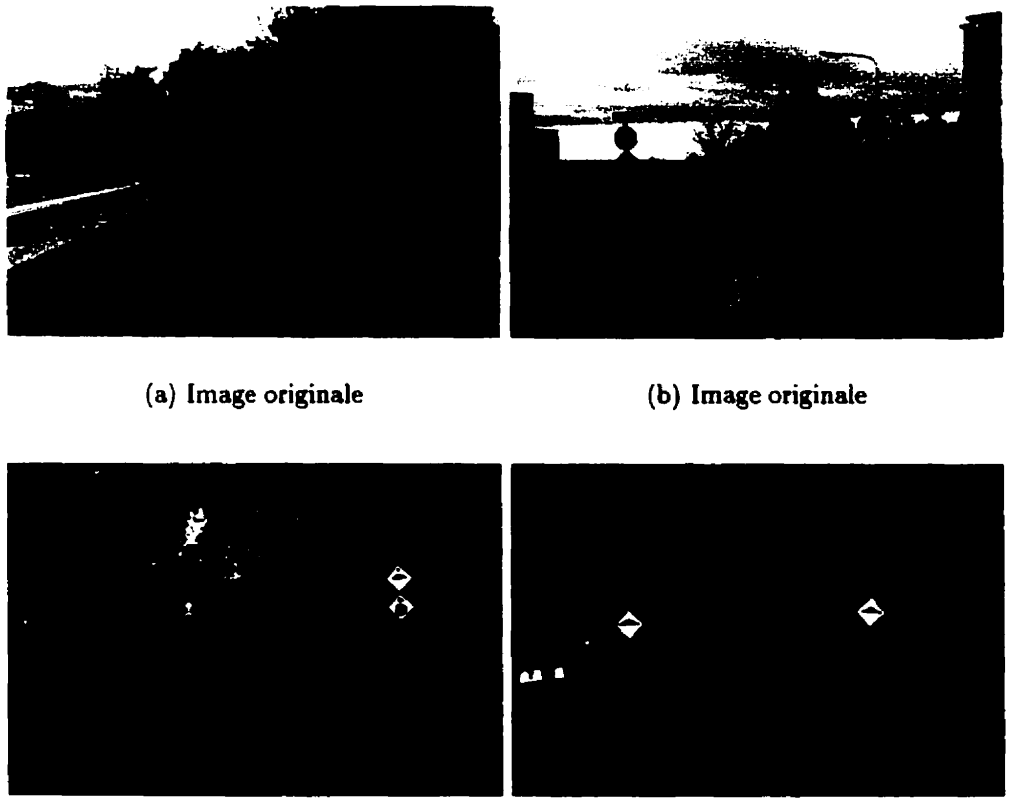


(c) Régions rouges et bleues produites par la segmentation



(d) Régions rouges et bleues produites par la segmentation

Figure 3.9 : Segmentation pour le rouge et le bleu.



(a) Image originale

(b) Image originale

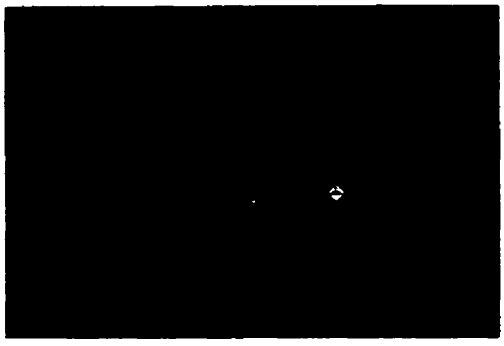
(c) Régions jaunes produites par la segmentation

(d) Régions jaunes produites par la segmentation

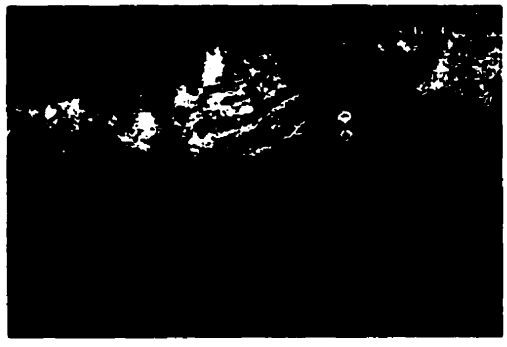
Figure 3.10 : Segmentation pour le jaune.

(a) Image originale

(b) Image originale

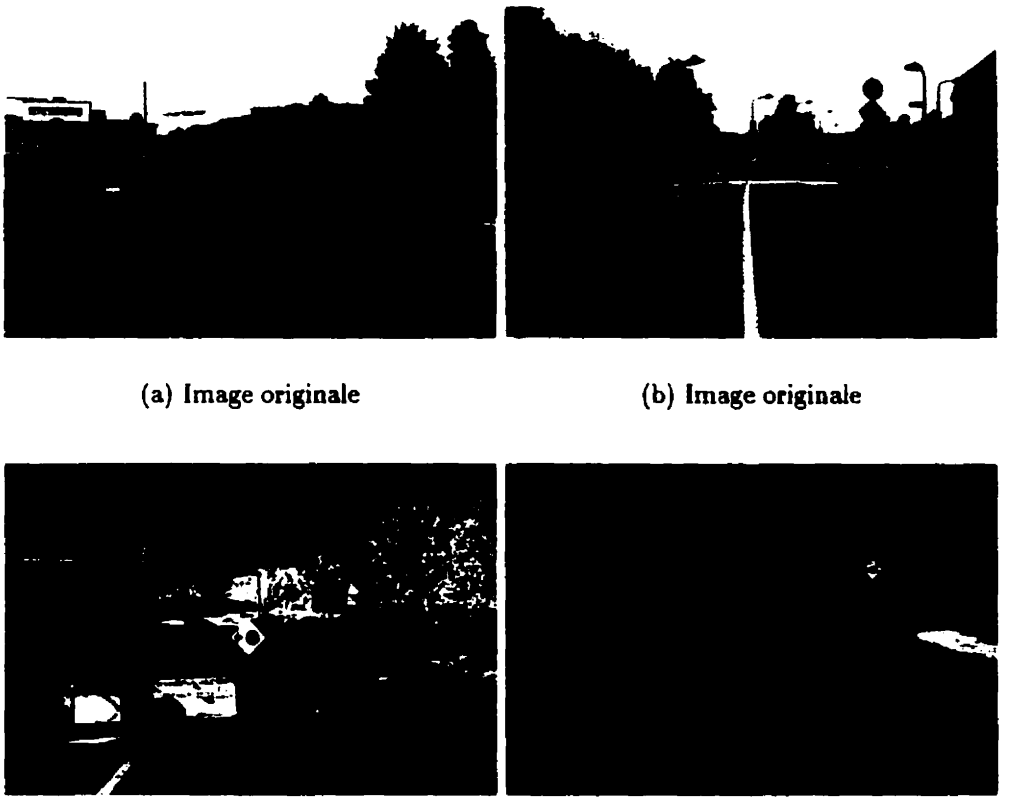


(c) Régions jaunes produites par la segmentation



(d) Régions jaunes produites par la segmentation

Figure 3.11 : Segmentation pour le jaune.



(a) Image originale

(b) Image originale

(c) Régions jaunes produites par la segmentation

(d) Régions jaunes produites par la segmentation

Figure 3.12 : Segmentation pour le jaune.

Chapitre 4

Détection des panneaux grâce à leur forme

4.1 Introduction

Si la couleur permet de distinguer un panneau de son environnement, c'est l'association d'une forme et d'une couleur qui permet de le reconnaître comme étant un panneau. L'approche proposée vise à tirer parti de cette association en recherchant la présence de ces formes dans l'image formée par les régions ayant la teinte correspondante. Comme la forme dans l'image dépend de l'état du panneau, de sa position par rapport à la caméra, des occlusions et des erreurs de la segmentation, la corrélation est employée afin de disposer d'un outil de détection robuste. Cette recherche est précédée d'une élimination des régions dont le contour est trop irrégulier pour

correspondre à l'une des formes simples cherchées. Ce processus nécessite de déterminer des seuils de décision : nous présentons une technique d'optimisation pour un ensemble d'images donné et avons testé ces seuils sur un ensemble distinct.

4.2 Revue bibliographique

4.2.1 Avantages et inconvénients de la forme

La forme est un élément fondamental pour reconnaître un objet. Cependant les informations sur la forme dépendent de la résolution et les seules qui soient indépendantes du point de vue sont difficiles à extraire (Swain et Ballard, 1991). La perception de la forme ne dépend toutefois pas des conditions d'éclairage et, dans le cas particulier des panneaux routiers, ceux-ci apparaissent dans les mêmes conditions d'observations (Piccioli et al., 1994).

4.2.2 Signatures non-basées sur la forme

Outre la couleur, Janssen et al. (1993) utilisent les associations spatiales de couleurs car celles-ci sont moins sensibles à la quantification spatiale que les bords de régions. C'est aussi aux frontières que les couleurs sont le moins bien définies, et les régions les plus petites ne sont pas prises en compte. Cette approche semble appropriée pour les panneaux associant au moins deux couleurs, dans la mesure où la segmentation en région préserve la topologie de celles-ci. Les histogrammes de couleurs ne sont en

revanche pas reliés à la géométrie des objets, sont robustes vis-à-vis des changements de point de vue (Swain et Ballard, 1991) et constituent donc de ce fait des outils efficaces pour la reconnaissance d'objets. Cependant ils sont sensibles aux variations de l'illumination (Funt et Finlayson, 1995). Li et Lalonde (1995) ont adapté cet outil au problème de la détection des panneaux en prenant en compte les teintes identifiées après un étiquetage des pixels. Cette solution a le mérite d'être simple et rapide, mais le nombre maximal de teintes employées sur un panneau au Québec est de quatre (Transport Québec, 1993) et ne justifie pas nécessairement l'emploi d'un tel outil.

4.2.3 Le problème de la détection

Ce problème est envisagé après une segmentation de l'image. Trois types d'approches peuvent être distingués. Il est tout d'abord possible de caractériser progressivement les éléments segmentés, chaque étape du processus pouvant éventuellement être l'occasion d'une sélection. Dans la mesure où il arrive que les panneaux ne soient pas segmentés en un seul élément, des regroupements peuvent s'avérer nécessaires. Pour prendre du recul vis-à-vis de la segmentation, on peut déterminer les régions de l'image présentant certaines conjonctions typiques des panneaux, puis étudier ces régions. Il est aussi possible d'estimer une statistique sur l'ensemble de l'image et de décider de la présence de panneaux dans celle-ci. Cependant en se limitant à une petite région autour d'un candidat potentiel, on réduit l'influence sur l'analyse d'autres

objets segmentés.

4.2.3.1 Identification des éléments segmentés

Dans le cas d'une approche par contours, Piccioli et al. (1994) recherchent dans les arêtes les segments de droite dont l'orientation et la dimension en font de bons candidats pour un bord de panneau. Dans le cas d'une approche par régions, les différents auteurs (Baghdassarian, 1995; Kehtarnavaz et Ahmad, 1995; Priese et al., 1994) déterminent d'abord la couleur des régions segmentées et l'utilisent comme premier critère de sélection puis s'intéressent à la forme. Kehtarnavaz et Ahmad (1995) emploient une transformation logarithmique-polaire-exponentielle définie par rapport au barycentre, suivie d'une transformée de Fourier pour obtenir une description des régions retenues qui soit invariante par translation, rotation et changement d'échelle; cette description n'est cependant pas robuste vis-à-vis des occlusions. Baghdassarian (1995) retient des régions en fonction de leur finesse (rapport de l'aire sur le carré du périmètre), puis utilise une approximation polygonale du contour pour reconnaître leur forme. Priese et al. (1994) se limitent à l'étude d'une approximation de l'enveloppe convexe et estiment les probabilités d'appartenance aux différentes formes possibles. L'histogramme des couleurs d'une région et de ces inclusions fournit une seconde probabilité : la combinaison des deux permet de décider de la présence où de l'absence d'un panneau. Si considérer l'enveloppe convexe permet de gagner en robustesse vis-à-vis des occlusions, il est aussi parfois nécessaire de rassembler des régions qui correspondent à un

même panneau qui a été ainsi scindé lors de la segmentation.

4.2.3.2 Recherche et étude d'autres régions d'intérêt

Il s'agit de déterminer une sous-région de l'image (rectangulaire pour des raisons de simplicité) contenant un panneau potentiel; suivant le produit de la segmentation divers critères peuvent être employés et combinés. Dans le cas d'une approche par contours, la corrélation entre un masque présentant la forme cherchée et l'image d'arêtes donne des pics là où cette forme est présente (Piccioli et al., 1994). La recherche de segments de droite dont l'association peut définir une forme géométrique simple donne un indice sérieux quant à la présence d'une telle forme (Piccioli et al., 1994). Dans le cas d'une approche par régions, Li et Lalonde (1995) considèrent la plus petite sous-image rectangulaire contenant une région segmentée ou deux de la même couleur et suffisamment proches. Janssen et al. (1993) emploient une méthode similaire mais pour des associations topologiques particulières de couleurs: les proportions de la sous-image sont par ailleurs prises en compte.

Piccioli et al. (1994) analysent le contenu des régions ainsi obtenues par corrélation normalisée avec différentes images de référence en niveau de gris; ils font toutefois l'hypothèse que les panneaux présentent toujours la même orientation dans l'image. Li et Lalonde (1995) emploient les histogrammes de couleurs pour effectuer une comparaison avec une liste d'objets de référence, ce qui offre une bonne robustesse vis-à-vis du point de vue. Janssen et al. (1993) recherchent la forme associée à la teinte jugée

la plus importante : un classifieur du type plus proche voisin opère sur la sous-image après normalisation de sa taille et filtrage de la couleur relativement à cette teinte. Cette classification nécessite un apprentissage.

4.2.3.3 Détection de la seule présence du panneau

Estevez et Kehtarnavaz (1996) déterminent un histogramme de la distribution locale des pixels des arêtes correspondant à la frontière d'une zone rouge : cet histogramme permet de détecter la présence de l'un des trois panneaux nord-américains contenant du rouge. Ce système ne permet pas de localiser le panneau dans l'image et requiert que celui-ci soit d'une taille importante pour le détecter.

4.2.4 Conclusion

Les histogrammes de couleur offrent une description simple et facile à manipuler: il existe cependant des panneaux bien différents comme l'entrée interdite et le céder le passage dont les histogrammes doivent être fort semblables. L'information apportée par la forme est donc indispensable. L'étude des seules régions segmentées est limitée par la performances de la segmentation, tâche délicate notamment pour les petits panneaux et lorsque les limites des plages de variation des senseurs sont atteintes (les couleurs sont alors modifiées). De ce point de vue les techniques développées pour les images d'arêtes sont intéressantes car elles sont prévues pour exploiter une information partielle sur les frontières : adaptées à une approche région (les contours

sont facilement obtenables), elles doivent permettre de s'affranchir des certains problèmes de fusion de régions, de segmentation en plusieurs régions et d'occlusions. La corrélation présente l'avantage de traiter toutes les formes de la même manière et ne requiert pas d'analyse préalable de l'image.

4.3 Solution proposée pour la détection des panneaux

A la suite de la segmentation, nous disposons, pour chaque couleur recherchée, d'une image binaire contenant les régions obtenues. Notre algorithme repose sur le principe d'une corrélation entre l'image des contours de ces régions et des masques présentant les formes recherchées. La segmentation produit toutefois des régions aux frontières complexes qui n'apportent aucune information pertinente. Afin de ne pas avoir à traiter les contours correspondants lors de la corrélation, celle-ci est précédée d'une pré-reconnaissance de forme dont l'objectif est d'éliminer ces contours et que nous présentons maintenant.

4.4 Pré-reconnaissance de forme

4.4.1 Introduction

La pré-reconnaissance de forme a pour objectif de rejeter les régions issues de la segmentation qui ne peuvent correspondre à des panneaux afin de faciliter la reconnaissance finale. Cette opération se déroule en trois étapes. D'abord les pixels porteurs d'information parasite sont éliminés. Ce nettoyage est suivi de l'obtention des contours des différentes régions. Le tri des régions est alors effectué en se basant sur des caractéristiques locales de ces contours.

Cette pré-reconnaissance agit sur les images binaires présentant les régions segmentées correspondant à une teinte. Nous considérons qu'une région est un ensemble connexe de pixels qui sont 8-connectés. Ce choix permet de traiter l'ensemble des pixels d'une mince couronne comme une seule et même région. Les trois phases procèdent par classement de pixels et font pour cela appel à un type d'opération de morphologie mathématique que nous allons d'abord présenter.

4.4.2 Les opérations morphologiques de type "hit or miss"

Les opérations morphologiques de type "hit or miss" font partie des techniques classiques en traitement d'images (Pratt, 1991). Ces techniques sont appliquées à des images binaires. Il s'agit de conserver ou de changer la valeur d'un pixel en fonction de son voisinage (généralement de taille trois par trois).

Chaque type de voisinage est répertorié et se voit attribuer un code unique. Ce code permet de choisir la valeur finale du pixel au centre du voisinage à l'aide d'une table de recherche (figure 4.1).

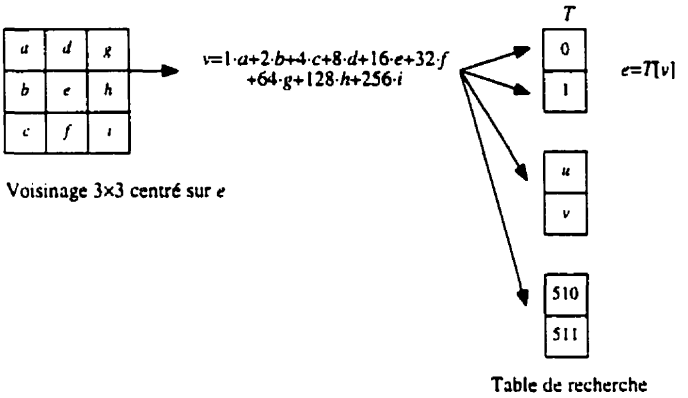


Figure 4.1 : Implantation des opérations de morphologie mathématiques.

Ces techniques peuvent aussi être employées pour localiser des agencements remarquables de pixels dans une image (Canning, Kim, Netanyahu et Rosenfeld, 1987). Lors de l'emploi de ces techniques nous nous limitons à des voisinages de taille trois par trois. Bien que parfaitement possible, l'utilisation de voisinages de taille cinq par cinq, est en pratique irréalisable en raison de l'explosion combinatoire menant à 33554432 entrées différentes dans la table.

4.4.3 Elimination des pixels parasites

L'objectif final est d'être en mesure de rejeter certaines régions, sans pour autant se livrer à une analyse complète de celles-ci. Il est donc souhaitable de casser toute

liaison entre une zone qui pourrait être un panneau et une autre zone qui pourrait ne pas l'être. En effet si la seconde, prise individuellement, devait être rejetée, alors la première subirait le même sort, car faisant partie de la même région. La figure 4.2 présente un agrandissement (de la figure 3.6) où une telle situation se produit. Ces cassures sont réalisées en supprimant les pixels qui peuvent être interprétés comme des liaisons entre deux zones distinctes sans l'être comme étant des éléments de panneaux. Pour une plus grande efficacité, les points et les trous isolés sont aussi effacés ou comblés : ils n'apportent en effet aucune information pertinente pour la recherche des régions qui peuvent correspondre aux panneaux. La figure 4.3 montre l'effet de ces traitements sur l'agrandissement déjà présenté. On peut constater qu'à l'issue de ce traitement l'octogone du panneau d'arrêt a été séparé de la région adjacente. Les deux nouvelles régions obtenues vont donc pouvoir être traitées séparément. Nous allons donc nous intéresser à leur forme en déterminant leur contour.

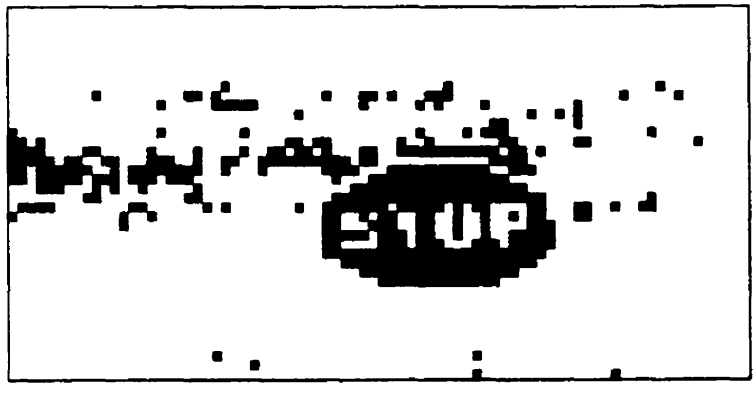
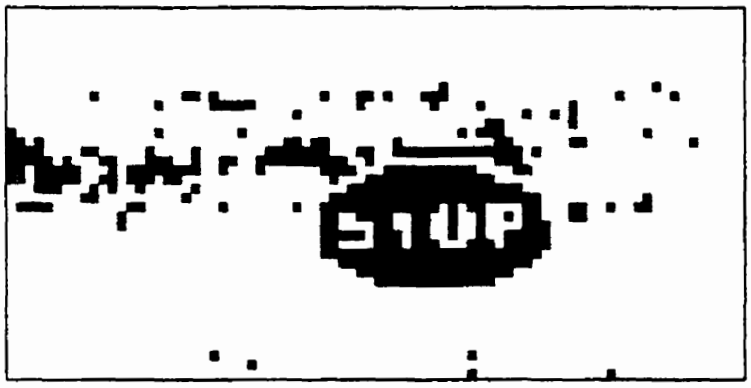
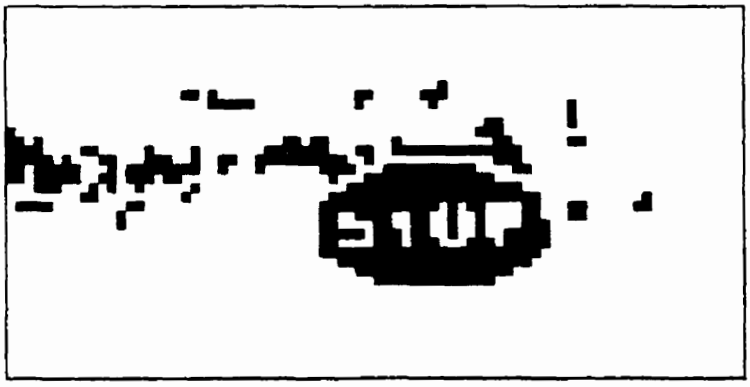


Figure 4.2 : Exemple de Liaison entre un panneau et une région rouge produites à la suite de la segmentation.

(a) Remplissage des trous isolés



(b) Elimination des liaisons entre régions



(c) Suppression des points isolés

Figure 4.3 : Elimination des pixels parasites.

4.4.4 Extraction des contours

Nous considérons des régions connexes au sens 8-connectées. Les contours permettent de conserver l'information sur la forme pour un minimum de points dans l'image. Ils sont obtenus en effaçant les pixels ayant quatre voisins similaires (au sens 4-connectés) (Pratt, 1991). Les contours obtenus sont donc 8-connectés et à une région correspond un seul contour. La figure 4.4 montre le résultat obtenu à ce stade à partir de l'image segmentée (figure 3.6).

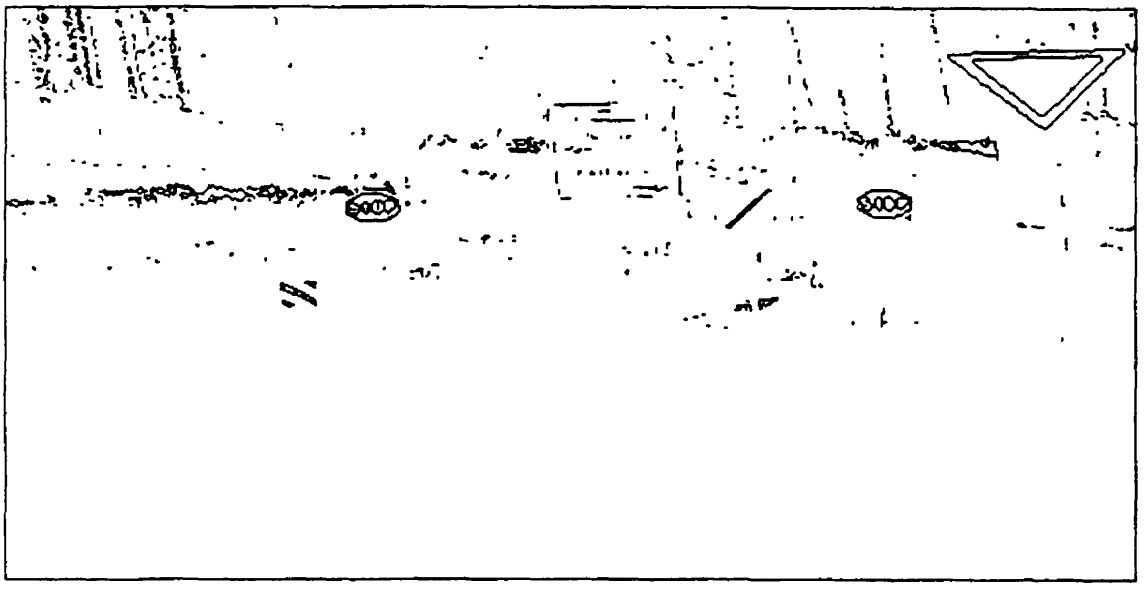


Figure 4.4 : Image des contours des régions rouges obtenues.

4.4.5 Rejet des contours n'étant pas des panneaux

Afin de pouvoir procéder à ce tri des contours, ceux-ci sont étiquetés en deux parcours de l'image. Les panneaux présentent des formes géométriques simples, cercles ou

polygones dont les angles ont été arrondis, bien régulières, et de ce fait les contours des régions qui leur correspondent ne devraient pas présenter d'extrémités libres, c'est-à-dire de pixels qui lors du parcours du contour constituent un cul-de-sac. Des extrémités libres peuvent toujours apparaître si le panneau a été segmenté sous la forme de plusieurs régions ou suite à des faiblesses de la segmentation, mais leur nombre reste limité. La détection de ces extrémités libres s'effectue à l'échelle d'un voisinage de taille trois par trois avec toujours la même technique.

La connaissance de la taille minimale dans l'image et du type de panneaux que l'on recherche, donne un critère sur la longueur (en nombre de pixels) minimale que doit avoir un candidat sérieux à la reconnaissance de forme. La figure 4.5 montre les contours finalement retenus dans le cas de l'image de la figure 2.9. On peut constater que la plupart des régions correspondant à l'arrière plan ont été éliminées. Avant d'aborder le problème de la localisation dans l'image des panneaux parmi les régions sélectionnées, nous étudions d'abord la dépendance de cette préreconnaissance avec la segmentation et nous présentons le cas particulier des régions correspondant aux panneaux de direction.

4.4.6 Interaction avec la segmentation

Le traitement par la pré-reconnaissance d'une région dépend des caractéristiques locales de son contour et, de ce fait dépend de la segmentation. En faisant varier les seuils définissant une couleur, on peut restreindre ou agrandir une région, voire

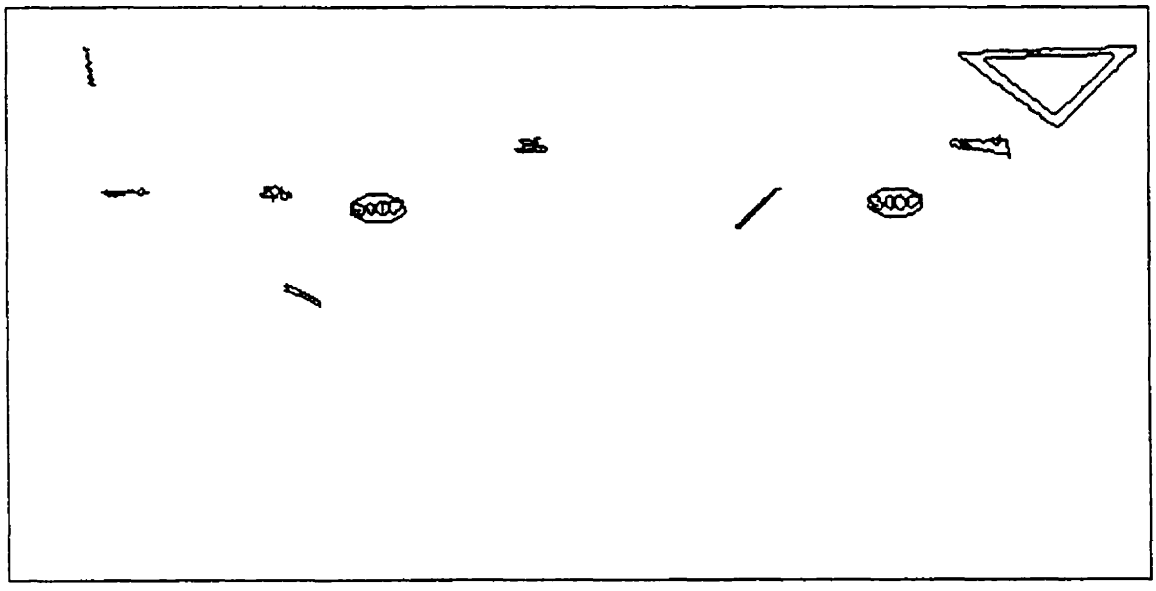


Figure 4.5 : Contours des régions rouges sélectionnés pour la reconnaissance de forme.

fusionner plusieurs régions.

Les panneaux doivent normalement contraster avec l'arrière plan. Sous certaines conditions (en particulier les éclairages rasants de début et fin de journée), il arrive que des panneaux soient difficiles à discriminer d'arrière-plans qui présentent alors les mêmes teintes avec des saturations élevées. Cela conduit à des régions difficilement reconnaissables comme on peut le constater sur les images de la figure 4.6. Plutôt que de chercher à exploiter ces régions, nous adaptons la pré-reconnaissance pour les éliminer autant que possible. Le résultat de cet opération dans le cas de notre exemple peut être vu sur la figure 4.7. La segmentation de ces zones spécifiques est réalisée avec un choix de seuils plus sévère; cette segmentation ne peut, de ce fait, créer de nouvelles régions sauf là ou la première segmentation a échoué. Cette

segmentation est suivie d'une pré-reconnaissance de forme qui lui est propre. Les images de la figure 4.8 montrent les résultats de cette seconde segmentation et de la pré-reconnaissance associée. Il est alors possible de fusionner par simple superposition les deux images segmentées, pour la reconnaissance de forme.

4.4.7 Rejet des contours des panneaux de direction

Les panneaux indiquant des directions constituent une classe à part : leur proportions, au contraire de la plupart des autres panneaux, ne sont pas fixées et leur graphisme constitué de flèches, de caractères et de lignes est complexe (Transport Québec, 1993). La figure 4.9 montre un exemple du type de contour que l'on peut obtenir. De ce fait, les contours résultants de ces panneaux peuvent être fort complexes et peuvent ainsi être source de fausses alarmes, d'autant plus qu'ils sont de grande taille. Dans la mesure où nous n'envisageons pas ici de les détecter, il convient de pouvoir les éliminer. Comme ils correspondent à des panneaux, il n'est pas possible de se baser sur une quelconque irrégularité. En revanche, ces contours présentent de nombreux trous et ont une grande longueur. Le nombre de trous peut être estimé à partir du nombre d'Euler (Pratt, 1991) de l'image contenant ce seul contour; la taille à partir du nombre de pixels. Cela permet de les rejeter. Nous présentons maintenant la recherche des panneaux parmi les régions sélectionnées.

(a) Image originale

(b) Régions jaunes segmentées

Figure 4.6 : Problèmes dans les zones très saturées.

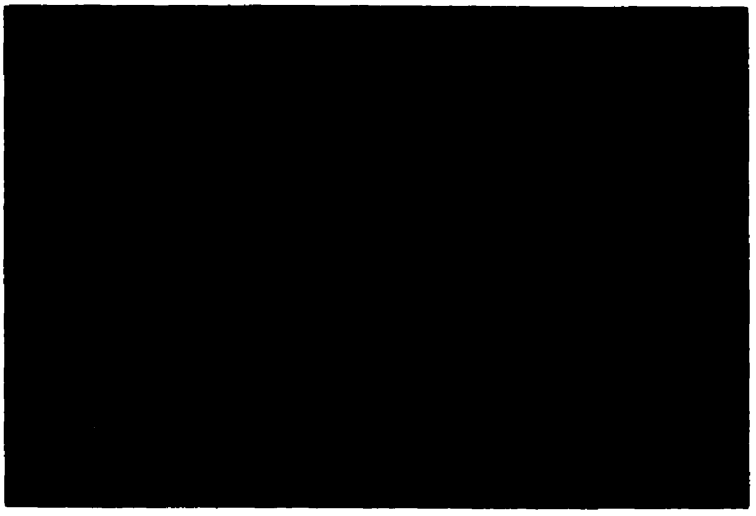
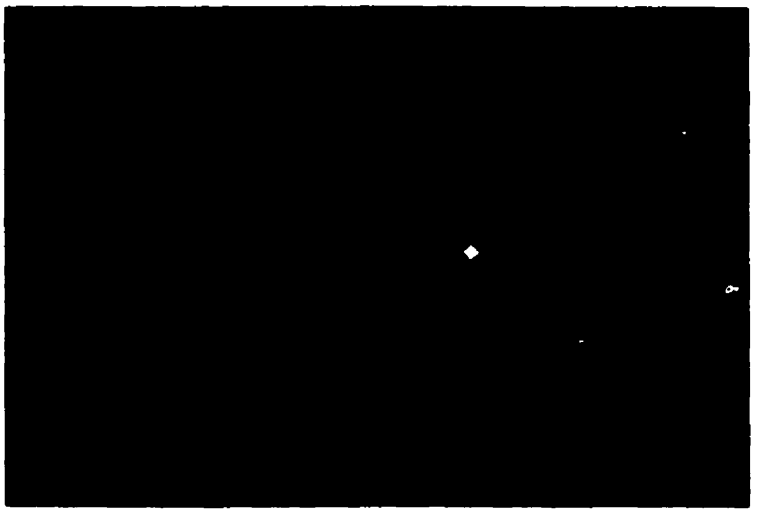
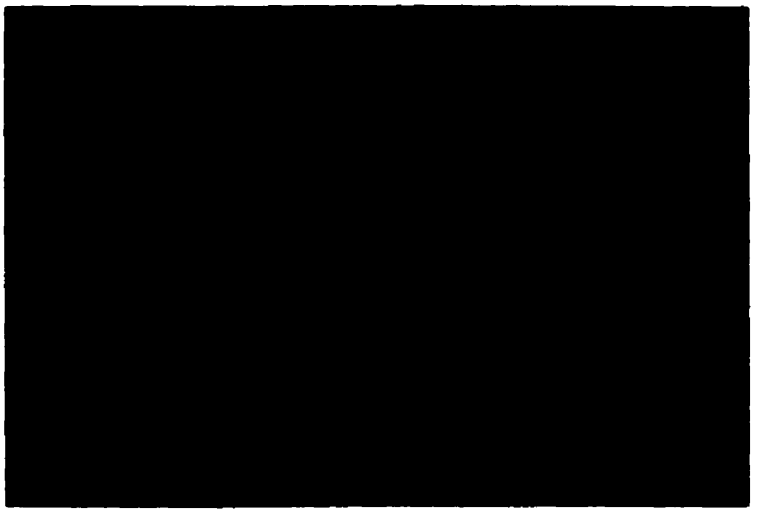


Figure 4.7 : Suppression des régions inutilisables.



(a) Régions jaunes produites par la seconde segmentation



(b) Contours retenus suite à la seconde segmentation

Figure 4.8 : Seconde segmentation et pré-reconnaissance pour les zones très saturées.

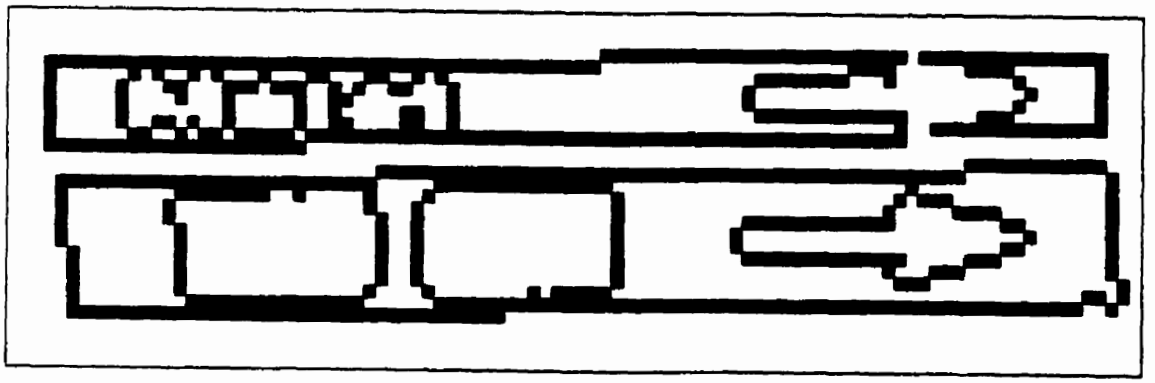


Figure 4.9 : Exemple de contour correspondant à un panneau de direction.

4.5 Localisation des panneaux dans l'image

4.5.1 Introduction

Chaque teinte remarquable conduit à une image de contours correspondante. Pour chacune de ces teintes il existe différents panneaux et un nombre limité de formes possibles. Comme l'orientation des panneaux varie très peu et qu'ils sont quasi-orthogonaux à l'axe optique de la caméra, ces formes se retrouvent à un facteur d'échelle près dans l'image. Les dimensions dans l'image restent faibles et nous faisons donc l'hypothèse que ni les aberrations de l'objectif ni les écarts acceptables par rapport aux normes de positionnement des panneaux n'induisent de déformations sur la forme dans l'image qu'il faille prendre en compte. Ces formes sont recherchées dans l'image de contours correspondante par corrélation avec des masques couvrant la plage des tailles possibles.

4.5.2 Principe et définitions

Nous réalisons une corrélation entre l'image des contours et une image présentant le contour théorique recherché. Cette seconde image est le masque de corrélation. Nous employons autant de masques que nécessaire pour couvrir l'ensemble des formes et des tailles de panneaux présentes dans les images. La corrélation du masque avec lui-même se traduit par un maximum de l'image résultant de la corrélation. Nous recherchons donc les maximums locaux (les pics de corrélation) dont la valeur dépasse

un certain seuil. Pour pouvoir évaluer notre système nous considérons qu'un panneau est localisé si un pic de corrélation satisfaisant est trouvé suffisamment proche de ce panneau. Sur la base de ce critère de détection, les termes fausse alarme et erreur sont alors ainsi définis :

- Une **fausse alarme** correspond à un pic de corrélation qui ne correspond pas à un panneau réel dans l'image: les statistiques sur les fausses alarmes se font en terme de nombre par image, puisqu'elles ne sont pas reliées à la présence de panneaux.

- Une **erreur** correspond à un pic de corrélation qui détecte un panneau qui n'a pas la forme cherchée. Le taux d'erreur peut être exprimé comme un pourcentage du nombre de panneaux associés aux couleurs recherchées et réellement présents dans l'image.

4.5.3 Description des masques utilisés

Il est possible d'utiliser deux types de masques : l'arrière plan de l'image constituant le masque peut être neutre ou avoir une contribution négative au résultat de la corrélation. Si par exemple, on cherche à détecter des cercles, le premier type de masque conduit au même résultat pour un cercle et un disque de rayon supérieur ou égal au rayon du cercle cherché. En revanche le second type permet de discriminer le cercle de ces disques. Comme nous travaillons sur l'image des contours des régions segmentées, le second type peut s'avérer plus robuste sur des portions de l'image

présentant des contours complexes et proches les uns des autres. Cependant le contour du pictogramme du panneau, de même qu'une partie du contour du panneau (dans l'éventualité où la segmentation n'est pas parfaite) peuvent être traités comme du bruit et ainsi conduire à un échec de la détection. De plus la corrélation avec ce type de masque nécessite bien plus d'opérations car tous les pixels du masque ont une contribution, alors que dans l'autre cas seuls les pixels appartenant au contour théorique doivent être pris en compte.

Les masques sont définis pour correspondre aux formes théoriques des panneaux, en supposant que les pixels sont carrés et se présentent sous la forme d'une couronne de 1 (le tableau 4.1 en donne un exemple). L'épaisseur standard est de deux pixels pour pouvoir s'adapter à toutes les tailles possibles en sortie de la segmentation; cependant les masques d'un pixel d'épaisseur pour les petits cercles ont montré un pouvoir de détection comparable, pour un meilleur taux de fausses alarmes (le tableau 4.2 en donne un exemple).

4.5.4 Sélection des pics de corrélation

Il peut se produire qu'un panneau conduise à des pics de corrélation pour différents masques, en divers points très proches. Cela peut être causé par les imperfections de la segmentation ou le fait que, pour les panneaux polygonaux, il est possible de détecter un panneau plus petit imaginaire dans un angle sur la base de deux côtés; le schéma de la figure 4.10 illustre ce problème pour un panneau triangulaire européen.

Tableau 4.1 : Masque pour détecter les losanges

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0
0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	0	0	0	0
0	0	0	0	0	1	1	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Comme on ne peut trouver qu'un panneau à un endroit, nous sélectionnons les pics de corrélation en appliquant les deux règles suivantes :

- tout pic de corrélation prime sur tout autre pic correspondant à un masque plus petit et situé dans la zone couverte par le masque ayant généré le premier pic; cette règle permet de traiter le problème illustré à la figure 4.10;

- si la situation précédente se produit pour plusieurs pics issus d'un même masque, le pic le plus élevé prime sur les autres.

Tableau 4.2 : Masque mince pour détecter les cercles

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0
0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
0	0	0	1	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0
0	0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

4.5.5 Choix des seuils

Chaque panneau apparaît dans un certain nombre d'images. L'objectif est de détecter un nombre maximum de ces apparitions, pour un nombre de fausses alarmes le plus faible possible. Cela différencie le système d'une application dans le domaine des véhicules intelligent où le but est de détecter les panneaux réels le plus tôt possible, en minimisant d'abord le nombre de fausses alarmes afin de ne pas perturber le conducteur.

Le choix des seuils se fait par apprentissage sur un ensemble d'images pris comme référence. Cet ensemble doit refléter fidèlement les conditions réelles rencontrées et donc présenter des panneaux de tous les types et les tailles cherchés. Nous décrivons maintenant pour une forme et une couleur la procédure d'apprentissage.

1. La position des panneaux de cette couleur est estimée par un opérateur dans chaque image. Cela donne deux ensembles de panneaux : celui des panneaux

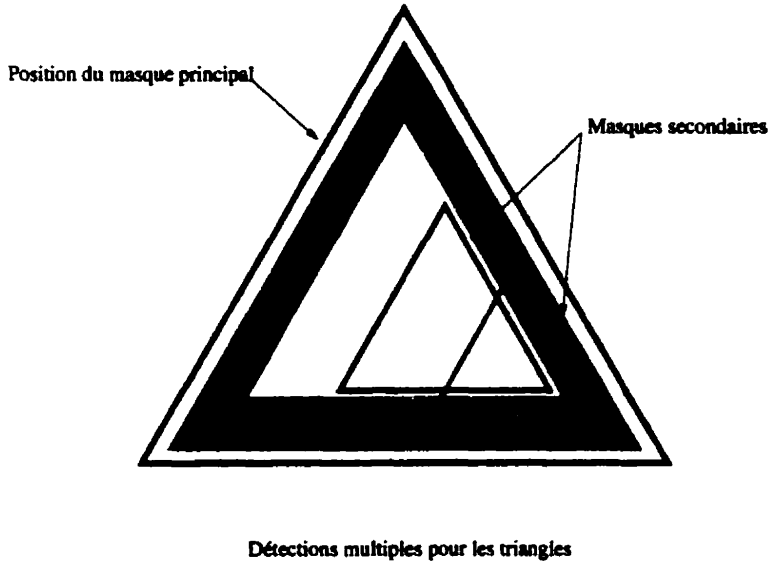


Figure 4.10 : Problème de détections multiples.

associant la forme et la couleur cherchées et celui des autres panneaux présentant la même couleur mais une autre forme.

2. Chaque image est segmentée, puis tous le processus de détection jusqu'à obtention des pics de corrélation est effectué.

3. Une fois les deux premières étapes complétées, il est possible de déterminer pour chaque combinaison de seuils les nombres de panneaux détectés, de fausses alarmes et d'erreurs. Il s'agit alors de trouver la combinaison de seuils optimale. Dans la mesure où, plus le masque est grand, plus il est improbable d'avoir une adéquation parfaite entre un contour et un masque, la contrainte de décroissance des seuils avec l'augmentation de la taille des masques est imposée. En respectant cette contrainte, on fait varier les seuils entre deux valeurs avec un

pas de variation donné; les performances de chaque combinaison sont évaluées.

4. Pour un apprentissage plus rapide, les combinaisons de seuils conduisant à trop de fausses alarmes sont d'abord éliminées si nécessaire.

5. On recherche ensuite la meilleure détection possible.

6. Parmi les combinaisons de seuils atteignant ce taux de détection on retient celles ayant le plus bas taux de fausses alarmes. On retient en pratique la combinaison pour laquelle les seuils pour les plus petits masques sont les plus élevés, ces masques-ci ayant tendance à générer plus de fausses alarmes.

Comme les performances atteignables dépendent de la qualité des images, des tailles de panneaux présentes et recherchées, il peut être nécessaire de répéter les trois dernières étapes.

4.6 Expériences

Dans les images provenant de l'université de Koblenz, les panneaux suivants sont recherchés : interdictions (cercle rouge), danger (triangle rouge), obligation (cercle bleu) et indication (rectangle bleu). Le tableau 4.3 donne le nombre d'apparition de ces panneaux. Dans les images du GRPR, on recherche les panneaux de danger (losange jaune); comme des panneaux de très petites tailles sont présents et difficilement reconnaissable pour un observateur, nous avons décidé de ne pas rechercher les panneaux dont la taille est inférieure à 14 pixels par 14 pixels. Cela correspond

approximativement aux plus petites tailles de panneaux présents dans l'autre groupe d'images. L'appréciation de la taille des panneaux étant délicate, la détection d'un panneau jugé trop petit est considéré comme une fausse alarme. En fin de compte, 330 panneaux de danger nord-américains sont recherchés.

Tableau 4.3 : Nombre de panneaux recherchés dans les images de l'université de Koblenz

Type de panneau	Nombre
rouges circulaires	87
rouges triangulaires	47
bleus circulaires	27
bleus rectangulaires	44

Nous travaillons avec la segmentation définie au chapitre précédent. Nous avons réalisé trois groupes d'expériences. Le but du premier est d'estimer les performances atteignables. Avec la seconde série nous avons voulu voir la robustesse du système vis-à-vis des images employées pour l'apprentissage. Finalement, nous avons testé les performances du système pour les panneaux de grande taille car un tel système constitue le fondement d'un système de suivi. Les critères de détection utilisés sont les suivants. Un panneau est considéré comme détecté si l'on obtient un pic de corrélation dans une petite région centrée sur ce panneau. Pour les cercles rouges, nous avons employé un cercle ayant un rayon de 5 pixels; ce critère ayant été jugé trop strict a *posteriori* et ne reflétant pas le fait que l'on disposait que d'une ligne sur deux, pour les cercles et rectangles bleus ainsi que les triangles rouges, un rectangle de 7 pixels par 13 pixels a été utilisé; pour les losanges jaunes un carré de 13 pixels de côté à été

employé. Les deux dernières régions sont équivalentes car pour les images avec les losanges jaunes on dispose de toutes les lignes.

4.6.1 Détermination des performances de la détection

L'apprentissage des seuils est réalisé sur l'ensemble des images disponibles. Cela a été effectué pour les panneaux rouges (cercles et triangles) et bleus (cercles et rectangles). Les paramètres pour la suppression de régions ne pouvant correspondre à des panneaux ont été déterminés par contrôle visuel sur l'ensemble des images. Pour les panneaux jaunes nous avons seulement retenu les performances obtenues sur l'ensemble des images à partir de seuils définis sur un sous-ensemble.

4.6.2 Robustesse vis-à-vis de l'ensemble d'apprentissage

Deux expériences ont été réalisées. Avec les cercles rouges, les seuils pour la corrélation ont été déterminés sur un ensemble d'apprentissage formé en sélectionnant la moitié des séquences avec pour seul critère d'être représentatif du point de vue des tailles rencontrées. Les seuils obtenus ont été appliqués sur le complémentaire. Avec les losanges jaunes, l'opération est similaire sauf que tous les seuils (segmentation, élimination de région, corrélation) ont été d'abord estimés sur un ensemble d'apprentissage. Cet ensemble a été formé en sélectionnant 40% des séquences avec pour seul critère que d'être représentatif du point de vue des tailles de panneaux et des conditions éclairages rencontrées. Une double segmentation (avec chacune un

seuil différent pour la saturation) est effectuée.

4.6.3 Détection des grands panneaux

Pour initialiser la poursuite d'un panneau, il est nécessaire de détecter une de ses dernières apparitions. Cela revient à détecter les plus grands panneaux dans l'image. La détection d'une seule de ces apparitions suffit. Il s'agit donc de détecter les panneaux réels sans utiliser les petits masques. Une liste des panneaux réels qui doivent être détectés avec ces masques est établie; toute détection d'un autre panneau réel est considéré comme une fausse alarme. Les panneaux retenus sont ceux ayant au moins une apparition dont la largeur est strictement supérieure à 25 pixels.

4.7 Résultats

4.7.1 Résultats de la détection

Le tableau 4.4 donne les performances de la détection par type de panneau, de forme, de couleur et pour l'ensemble des types.

4.7.2 Apprentissage des seuils

Le tableau 4.5 donne les résultats des apprentissages pour deux types de panneaux, sur les ensembles d'apprentissage ainsi que sur les complémentaires.

Tableau 4.4 : Résultats pour tous les types de panneaux

Couleur \ Forme		rouge	bleu	jaune	toutes
	détectés	91%	78%		89%
cercle	fausses alarmes	0.02/im	0.01/im		0.02/im
	erreurs	4%	0%		3%
	détectés	91%			91%
triangle	fausses alarmes	0.02/im			0.02/im
	erreurs	0%			0%
	détectés		84%		84%
rectangle	fausses alarmes		0.38/im		0.38/im
	erreurs		2%		0.6%
	détectés			88%	88%
losange	fausses alarmes			0.64/im	0.64/im
	erreurs			0%	0%
	détectés	93%	82%	88%	88%
toutes	fausses alarmes	0.04/im	0.39/im	0.64/im	0.44/im
	erreurs	4%	2%	0%	2%

Tableau 4.5 : Résultats de l'apprentissage

type de panneaux / Ensembles		cercle rouge	losange jaune
	détectés	94%	90%
Apprentissage	fausses alarmes	0.01/im	0.59/im
	erreurs	6%	0%
	détectés	91%	86%
Complémentaire	fausses alarmes	0.07/im	0.68/im
	erreurs	1%	0%

4.7.3 Détection des panneaux réels avec les plus grands masques

Le tableau 4.6 montre les performances de la détection des panneaux physiques en employant uniquement les plus grands masques, dans le but de limiter le nombre de fausses alarmes.

Tableau 4.6 : Résultats du module détection

panneaux physiques détectés avec les plus grands masques détectes	97%
fausses alarmes	0.09/im
erreurs	2%

Les résultats de ces trois groupes d'expériences sont analysés à la section suivante.

4.8 Discussion

4.8.1 Analyse des résultats

La procédure de choix des seuils par apprentissage (tableau 4.5) dans le cas des losanges jaunes se caractérise par une légère dégradation des performances lorsque l'on passe de l'ensemble d'apprentissage à l'ensemble de test. Le taux de détection chute en effet de 4 points. Pour une couleur et une forme données il est donc possible de régler l'ensemble des opérations réalisant la détection indépendamment des images.

L'expérience similaire d'apprentissage de la corrélation sur les panneaux rouges circulaires (tableau 4.5) montre qu'à condition de disposer d'une bonne segmentation, la reconnaissance des cercles rouges est bonne et indépendante des images d'apprentissage.

Nous ne disposons pas de résultats directs et formels prouvant que les segmentations pour les autres couleurs soient applicables à d'autres images avec succès. Cependant le tableau 4.4 montre que de bons résultats sont obtenus avec les autres couleurs, ce qui nécessite à la base une bonne segmentation. Dans le cas du bleu, les fausses alarmes sont produites lors de la recherche de rectangles sur une séquence d'images tandis que la détection des cercles bleus présents dans ces images conduit à de bons résultats. Comme la segmentation traite les pixels indépendamment les uns des autres, si ces fausses alarmes étaient causées par une mauvaise segmentation du bleu pour ces images cela devrait se ressentir aussi sur la détection des cercles; si

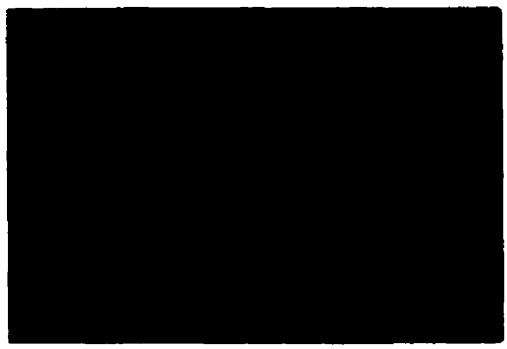
cela était du à une compensation par la détection de mauvaises performances de la segmentation sur d'autres images, alors cela se traduirait très probablement par des fausses alarmes pour la recherche des cercles bleus dans ces images. Comme nous ne constatons aucune de ces deux situations, il y a de très forte chance que la segmentation pour le bleu donne effectivement de bons résultats. Savoir si cette segmentation donnerait de bons résultats sur d'autres images obtenues avec le même matériel revient à savoir si l'ensemble actuel d'images contient un échantillonnage de conditions suffisamment variées : cela semble être le cas pour le rouge.

L'étude des performances par couleur et par forme (tableau 4.4) montre que les losanges jaunes conduisent à un taux élevé de fausses alarmes : cela est dû essentiellement à la présence de feuillages et de marquages au sol jaunes dans les images ainsi qu'à l'éclairage rasant de fin de journée pour certaines prises de vues qui hausse la saturation moyenne et crée des zones semblables du point de vue de la couleur aux panneaux. Ce phénomène est d'autant plus sensible que l'équilibrage des blancs n'est pas fait pour les images contenant des losanges jaunes. La figure 4.11 permet de comparer les contours obtenus pour une même scène en cours puis en fin de journée : l'image prise en fin de soirée comporte beaucoup plus de régions perçues comme jaunes, ce qui conduit à des fausses alarmes.

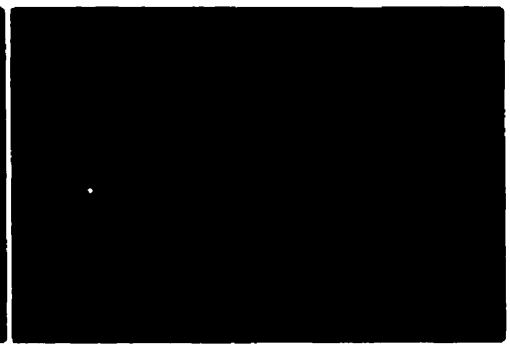
Les rectangles génèrent aussi beaucoup de fausses alarmes, ce en raison de la présence de cette forme en milieu urbain : les constructions présentent de nombreuses lignes verticales et horizontales qui s'intersectent. De plus dans les rues à l'ombre la

(a) Image originale prise dans la journée.

(b) Image originale prise en fin de journée.



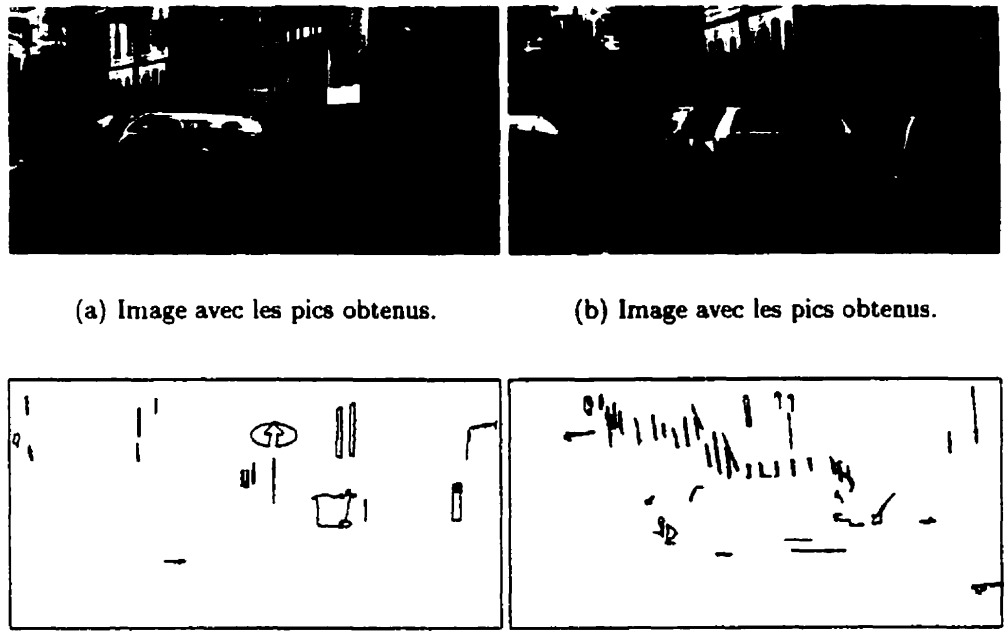
(c) Image du jour après segmentation et traitement.



(d) Image du soir après segmentation et traitement.

Figure 4.11 : Influence de l'éclairage sur la segmentation.

principale source d'éclairage est le ciel, ce qui fait virer vers le bleu les teintes grises (achromatiques donc facilement modifiables), couleur pour laquelle on recherche de telles formes. La figure 4.12 montre les images originales, avec les pics de corrélations indiqués par le cadre du masque correspondant ainsi que les images de contours associées. Pour bien illustrer le problème, tous les pics de corrélation obtenus sont montrés.



(a) Image avec les pics obtenus.

(b) Image avec les pics obtenus.

(c) Image après segmentation et traitement.

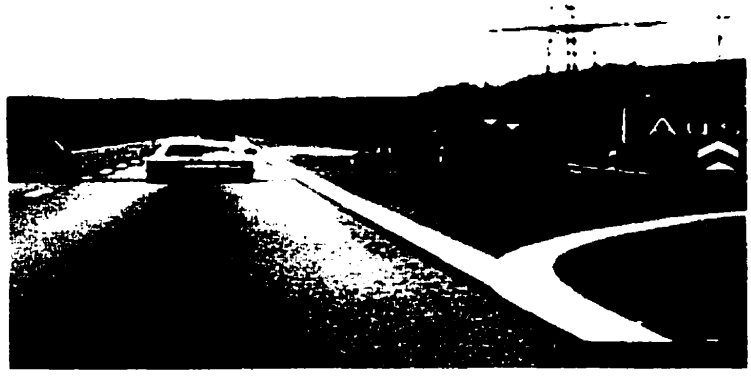
(d) Image après segmentation et traitement.

Figure 4.12 : Fausses alarmes pour la détection des rectangles bleus en milieu urbain.

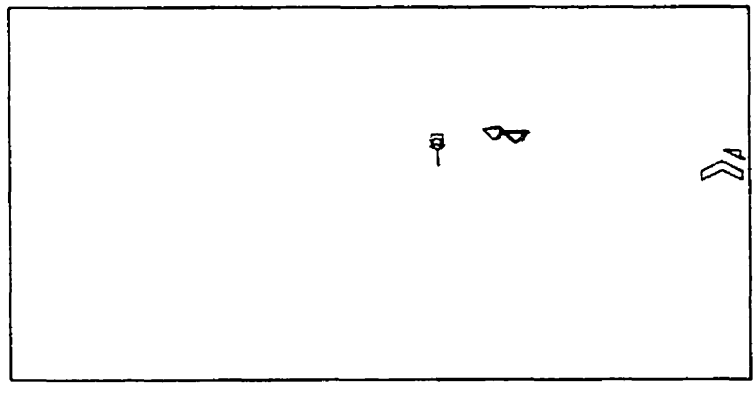
Il faut noter que les faibles performances pour les cercles bleus s'expliquent par la présence de panneaux coupés par le bord de l'image, qui sont éliminés avant la corrélation car leur contour devient alors trop irrégulier. Sinon 100% des panneaux

restants sont correctement détectés.

Pour les panneaux rouges il n'y a pas de problèmes, exceptés des triangles qui sont vus comme des cercles. Pour les petites tailles, le cercle inscrit est détecté. La figure 4.13 donne un exemple de ce problème.



(a) Image avec les pics obtenus.



(b) Image après segmentation et traitement.

Figure 4.13 : Erreur lors de la détection des cercles rouges

Les scènes routières sont complexes et fort variées; nous devons de plus déterminer par apprentissage un nombre important de paramètres. Nous ne disposions que d'un

nombre relativement limité d'images (quelques centaines) pour estimer la valeur de ces paramètres. Il n'est donc pas sur que l'on puisse obtenir des résultats similaires à ceux que nous venons de présenter avec d'autres ensembles d'images (obtenus avec le même matériel de prise de vue) sans réaliser un nouvel apprentissage à l'aide d'un ensemble d'images plus important.

La première phase d'une recherche avec suivi des panneaux dans les images précédentes consiste à chercher les grands panneaux avec seulement les grands masques. Le tableau 4.6 montre que le taux de détection est alors très bon (97%) et que le taux de fausses alarmes qui était un des principaux problèmes a été ramené à 0.09 par image soit une réduction d'un facteur 5.

Il est difficile de comparer ces résultats à ceux obtenus avec d'autres algorithmes dans la mesure où les ensembles de test ne sont pas les mêmes. Il convient aussi de prendre en compte le but poursuivi : nous cherchons à détecter les panneaux sur la plus grande distance possible pour pouvoir estimer leur visibilité et les localiser, tandis qu'une application d'aide à la conduite cherche à évidemment reconnaître les panneaux le plus tôt possible, mais doit d'abord minimiser son nombre de fausses alarmes afin de ne pas perturber le conducteur. Par rapport aux travaux de Priese et al. (Priese et al., 1994), qui correspondent à une détection sans reconnaissance du pictogramme, nous obtenons des résultats similaires pour les grands panneaux (plus de 30 pixels); pour le panneaux plus petits notre taux de détection est meilleur, mais nous obtenons en contrepartie plus de fausses alarmes.

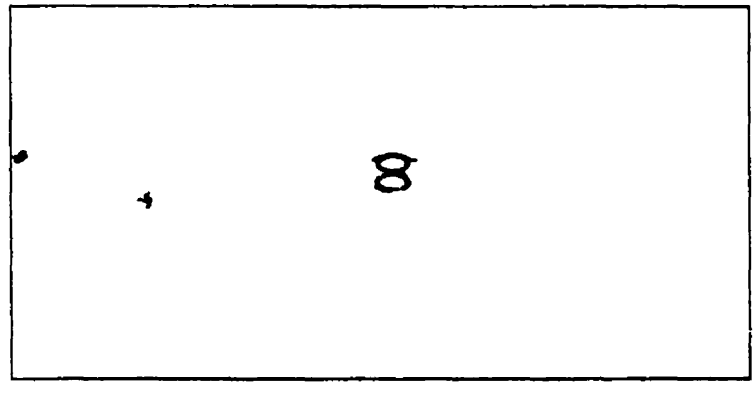
4.8.2 Analyse de l'algorithme

La corrélation, en dépit de la charge de calcul qu'elle entraîne, présente le net avantage d'être robuste vis-à-vis d'un certain nombre de problèmes qu'aucun algorithme de segmentation ne pourra éviter. Comme on peut le voir sur la figure 4.14, le panneau du bas a beau avoir été segmenté en deux régions, dont l'une correspond aussi au panneau du haut, nous sommes capables de les détecter correctement.

Le point faible est inévitablement la segmentation puisque ses performances conditionnent celles de l'ensemble. Il faut cependant remarquer que la qualité des images est primordiale : plus la définition des couleurs est bonne, plus les transitions entre les panneaux et l'arrière plan sont rapides. Ainsi une légère variation des seuils (en particulier celui sur la saturation qui caractérise bien les transitions entre des objets très et peu saturés) n'influe alors pas sur la région segmentée associée au panneau. De plus cela permet de tirer profit de la bande blanche ou noire situé parfois au bord du panneau pour résoudre des conflits du type rouge sur fond rouge. Une caméra tri-CCD à haute résolution montre alors tout son intérêt.

La segmentation par étiquetage ne peut donner de bonnes performances dans ce genre de contexte. Elle reste néanmoins rapide et simple. La technique d'évaluation des segmentations présente le mérite de permettre de travailler de façon systématique en l'absence du système de détection complet. Cependant elle accorde la même importance à tous les pixels, alors qu'en pratique ceux aux frontières des panneaux sont fondamentaux.

(a) Image avec les pics obtenus.



(b) Image après segmentation et traitement.

Figure 4.14 : Détection de deux panneaux, celui du bas étant segmenté en deux régions, celui du haut étant fusionné à celui du haut.

La pré-reconnaissance de forme vient pallier le principal défaut de la corrélation sans prise en compte de l'arrière plan : dans les zones où des régions présentent des contours irréguliers il est possible de reconnaître n'importe quoi. L'efficacité de ces traitements repose sur la qualité des images et de la segmentation.

La procédure d'apprentissage nous permet de déterminer des seuils de décision sans avoir à développer un modèle de décision probabiliste complexe tel qu'une classification bayésienne. Une sélection bayésienne des pics de corrélation nécessite l'estimation des probabilités conditionnelles a priori ce qui n'est pas immédiat. Le même problème se présenterait pour une classification des régions segmentées. La simplicité de notre solution se paie toutefois en durée d'apprentissage.

L'algorithme repose au départ sur une approche par régions, mais n'utilise plus cette connaissance en phase finale, car l'on souhaite pouvoir disposer d'un certain recul par rapport aux régions obtenues et donc traiter les contours comme de simples arêtes. Cela se traduit en outre par une solution algorithmiquement simple. Cependant l'information découlant de cette approche "régions" peut au moins être employée pour limiter la recherche de formes remarquables aux seules zones potentiellement intéressantes. Il semble aussi possible d'employer cette connaissance pour discriminer parmi les panneaux potentiels obtenus et donc réduire le nombre de fausses alarmes.

Notre algorithme de détection nous permet de trouver les panneaux dans l'image. Il ne nous permet pas cependant de faire le lien entre les apparitions d'un même panneau physique : un algorithme de suivi va nous permettre de réaliser cet appariement.

ce qui rendra possible l'exploitation des positions dans l'image pour localiser le panneau. Ces deux algorithmes vont être présentés dans le prochain chapitre.

Chapitre 5

Suivi et localisation des panneaux

5.1 Introduction

Un panneau est normalement visible sur une certaine distance lors du parcours de la route. Il apparait ainsi dans plusieurs images dans lesquelles il est possible de le détecter. Si l'on ne cherche pas à exploiter spécifiquement cette redondance, cela se traduit simplement par une augmentation des chances de détection d'un panneau réel. Pour exploiter cette redondance temporelle, il faut pouvoir relier les apparitions successives d'un panneau; notre premier objectif est donc de suivre les panneaux dans des séquences d'images. La connaissance de ces apparitions permet alors d'estimer la portion de la route depuis laquelle ce panneau est visible et offre de l'information supplémentaire pour estimer la position de ce panneau dans l'espace. Nous proposons de plus de coupler le suivi avec la segmentation et la détection afin de pouvoir limiter

la recherche des panneaux à des sous-régions de l'image. Notre approche s'appuie sur un système de poursuite dans des séquences monoculaires à l'aide d'un filtre de Kalman étendu permettant d'estimer la position des panneaux par rapport à la caméra. La localisation spatiale des panneaux est envisagée a posteriori comme un problème d'estimation en minimisant la somme des carrés des erreurs.

5.2 Revue bibliographique

5.2.1 Introduction

Peu d'attention a été portée sur le problème de la localisation et du suivi des panneaux routiers une fois ceux-ci détectés, la majorité des études étant orientée vers le domaine des véhicules intelligents, pour lesquels les objectifs fondamentaux sont la détection et la reconnaissance. Le problème de la localisation des panneaux routiers constitue cependant un cas particulier du problème de l'estimation de la structure et des mouvements de la scène relatifs au système de prise de vue, à partir d'une ou des séquences d'images produites. L'analyse de ces séquences d'images peut être conduite selon deux approches : soit en se basant sur le flux optique obtenu à partir des variations de l'intensité dans l'image, soit en suivant des éléments d'une image sur l'autre (Zhang et Faugeras, 1992). Nous ne nous intéressons pas ici aux techniques basées sur le flux optique car notre problème porte sur des éléments discrets (les panneaux) et car l'estimation du flux optique (notamment dans les régions de faibles variations

de l'intensité) est délicate et son exploitation difficile (Zhang et Faugeras, 1992).

5.2.2 Suivi des panneaux

Pour améliorer la robustesse de la détection, Piccioli et al. (1994) combinent le résultat des différentes détections d'un panneau en fonction de sa taille dans l'image correspondante : afin de réduire du même coup la région où le panneau est recherché (et donc le temps qui y est consacré) un système de poursuite est employé à partir de la première détection du panneau (lorsque la taille de celui-ci dans l'image dépasse dix pixels). Le filtre donne une estimation de la position du panneau par rapport à la caméra ce qui permet de prédire sa position dans l'image suivante: le modèle employé suppose que le mouvement se fait selon l'axe de la caméra à vitesse constante. Ce système visant à assister un conducteur, la poursuite est initialisée à partir de la détection de la plus petite apparition du panneau et donc sur laquelle on a le moins de certitudes.

5.2.3 Estimation du mouvement et de la structure de la scène

Il s'agit de déterminer la position des éléments suivis ainsi que leur mouvement par rapport au système de prise de vue. La plupart des auteurs supposent que la scène est rigide donc que tous les éléments suivis sont régis par le même mouvement; toutefois Zhang et Faugeras (1992) prennent en compte la possibilité de solides rigides ayant des mouvements indépendants, ce qui constitue une simple extension du cas le plus

simple. A partir des positions des éléments suivis dans les images, l'objectif est de construire un système d'équations dont les inconnues sont la position de ces éléments et les paramètres décrivant leur mouvement, puis de le résoudre. Ce formalisme dépend des choix sur le système de prise de vue, le repère, le type d'objet suivi, la modélisation des incertitudes et les techniques de résolution qui conduisent à des approches différentes.

5.2.3.1 Système de prise de vue

Il est possible de travailler avec une seule caméra comme avec un système de vision stéréoscopique. Le modèle du sténopé est généralement adopté, la formation des images correspond alors à une projection de type perspective; cependant dans le cas d'un système monoculaire, dont la distance focale est très supérieure aux dimensions du senseur, lorsque l'amplitude des déplacements de ce système le long de l'axe optique est très inférieure à la distance à la scène, une projection orthographique peut être employée comme modèle de formation des images (Tomasi et Kanade, 1992).

L'utilisation de séquences monoculaires conduit néanmoins à une incertitude sur le couple translation-position de la caméra. Dans le cas d'une projection de type perspective, ces deux vecteurs ne peuvent être connus qu'à un facteur près (Broida et Chellappa, 1991). Dans le cas d'une projection orthographique, leur composante selon l'axe optique ne peut être connue, mais dans sa zone de validité ce modèle conduit à de meilleurs résultats (Tomasi et Kanade, 1992). Un système stéréoscopique permet

de lever cette ambiguïté ainsi que d'éliminer la non-linéarité induite par la projection de type perspective (Broida et Chellappa, 1989). Un appariement est alors nécessaire entre les deux images.

5.2.3.2 Repère

Ce problème peut donc être décrit dans un repère lié soit au système de prise de vue soit à la scène. Ce choix n'influe pas sur la complexité de l'ensemble des équations (Matthies et Shafer, 1987). Cependant, Tomasi et Kanade (1992) considèrent que le choix d'un repère à l'origine du dispositif de prise de vue conduit à une plus grande sensibilité au bruit pour l'estimation de la géométrie de la scène, lorsque les mouvements sont faibles et lorsque les dimensions dans la scène deviennent nettement inférieures à la distance entre la scène et le système de prise de vue.

5.2.3.3 Objets poursuivis

Le but étant de retrouver un même objet dans différentes images, deux aspects doivent être considérés : la nature géométrique de cet objet, c'est-à-dire l'information qu'il apporte, et ce qui permet de le localiser dans une image. Le point est l'élément le plus simple. Broida et Chellappa (1991) localisent manuellement des points à l'aide de balises disposées dans la scène. Matthies et Shafer (1987) se limitent aux points dont le voisinage présente une configuration particulière de l'intensité dans l'image mais peuvent donc travailler avec des scènes inconnues. L'appariement est effectué sur la base d'une corrélation (Zheng et Chellappa, 1995). Un segment de droite apporte

plus d'information qu'un point, car son mouvement est celui d'un solide rigide, en raison des occlusions il est difficile de connaître la position exacte des deux extrémités (Zhang et Faugeras, 1992). Le choix d'un objet connu possédant un certain nombre de points facilement identifiables permet de disposer d'un solide indéformable fournissant suffisamment de mesures : Wu, Rink, Caelli et Gourishanker (1989) utilisent ainsi une grille de calibration.

5.2.3.4 Modélisation des incertitudes

L'estimation des densités de probabilité réelles est difficile; cependant, en pratique, les deux premiers moments se révèlent suffisants (Zhang et Faugeras, 1992). Cela revient à considérer des variables aléatoires gaussiennes. L'erreur sur les mesures obtenues par triangulation à l'aide d'un système de vision stéréoscopique peut ainsi être modélisée mais il convient d'estimer les moments à partir d'un modèle d'erreur gaussien uniforme sur les positions dans l'image (Matthies et Shafer, 1987).

5.2.3.5 Techniques de résolution

Le principe général est de minimiser la somme des carrés des erreurs. Deux types de méthodes sont employés : soit un traitement en une seule étape a posteriori, soit un traitement récursif image (ou paire d'images) après image à l'aide d'un filtre de Kalman étendu (Wu, Chellappa et Zheng, 1995).

L'intérêt d'une approche récursive est de pouvoir traiter des séquences de longueur indéterminée à l'avance et de produire des résultats à chaque image ce qui est intéres-

sant pour des problèmes de navigation (Matthies et Shafer, 1987) et pour pouvoir prédire les positions futures des objets suivis (Chandrashekhar et Chellappa, 1991). Le filtre de Kalman étendu nécessite cependant une bonne initialisation pour pouvoir converger (Chandrashekhar et Chellappa, 1991).

L'approche a posteriori permet de respecter la non-linéarité du problème, et ne nécessite pas d'initialisation des paramètres que l'on souhaite estimer (Taylor, Kriegman et Anandan, 1991); certaines méthodes numériques peuvent cependant nécessiter une initialisation. En considérant toutes les images en même temps il est de plus peu probable de rencontrer des difficultés si pendant un certain temps la séquence d'images n'apporte aucune information (Tomasi et Kanade, 1992).

La diversité des expériences rend difficiles les comparaisons, mais il faut toutefois noter les deux résultats suivants. Leung, Liu et Huang (1991) ont estimé qu'il était possible d'atteindre les limites imposées par la quantification spatiale des images que l'on utilise des séquences mono ou binoculaires. Wu et al. (1995) obtiennent les mêmes niveaux de performances avec les approches récursives et a posteriori.

5.2.4 Conclusion

L'approche récursive offre l'avantage de pouvoir être employée pour faciliter l'appariement. Le problème de la lenteur de la convergence disparaît dès que l'on dispose d'une bonne estimation du mouvement du système de vision, au moins pour le suivi dans l'image. L'approche a posteriori a pour principal intérêt de ne pas nécessiter

d'initialisation et de ne pas souffrir de problème de convergence comme le filtre de Kalman étendu. Les systèmes stéréoscopiques permettent d'estimer dès le départ la position des objets dans la scène à condition que ceux-ci soient visibles pour les deux caméras; cela n'est pas nécessairement le cas pour des objets excentrés. L'emploi de la stéréo implique de plus un appariement qui peut être difficilement réalisable dans un environnement complexe où les possibilités d'occlusions sont relativement importantes.

5.3 Utilisation de la redondance temporelle

5.3.1 Objectifs

Il est intéressant, pour localiser un panneau, de connaître ses apparitions successives afin de disposer d'un maximum d'informations pour pouvoir déterminer sa position; le contrôle de sa visibilité présente les mêmes exigences. Plutôt que de mettre simplement un module de suivi en aval des opérations de segmentation et de détection, nous proposons, pour réaliser l'appariement des apparitions des panneaux, d'intégrer un système de poursuite à ces deux opérations. Cette intégration s'appuie sur une analyse des forces et des faiblesses du processus de localisation dans l'image ainsi que sur les senseurs que l'on doit trouver à bord du véhicule. L'idée directrice est qu'en suivant les panneaux, nous serons en mesure de restreindre la zone de recherche dans l'image et donc du fait de cette spécialisation d'améliorer l'ensemble. Parmi les

progrès possibles, un des plus évidents est le gain en temps d'exécution.

5.3.2 Principe

Comme on vient de le voir, les meilleurs résultats sont obtenus pour les panneaux qui sont de grandes dimensions dans l'image, c'est-à-dire juste avant qu'ils ne sortent du champ de vision. Le système en développement ne vise ni l'aide à la conduite ni l'automatisation de la conduite sur routes; il est donc possible de détecter un panneau au dernier moment, puis de revenir sur les images précédentes et d'améliorer la détection des premières apparitions de ce panneau.

Dans le cas d'un retour en arrière le panneau est juste repéré au dernier moment. En retraçant un panneau dans les images précédentes, il est alors possible de restreindre la zone de recherche à une fenêtre de taille bien inférieure à l'image : le gain le plus immédiat est en temps de calcul. Les probabilités d'obtenir des fausses alarmes et des erreurs diminuent : toutes celles qui se seraient produites en dehors des zones de recherches n'ont plus lieu d'être. En réduisant le domaine spatial on réduit du même coup l'énergie du bruit.

5.3.3 Poursuite et localisation des panneaux

La pièce maîtresse d'un système de détection et de localisation de la signalisation tirant parti de la redondance temporelle est la poursuite des panneaux. Le véhicule d'inspection comporte un système de vision stéréoscopique et donc cela permet d'accé-

der par triangulation à la position dans l'espace du panneau relativement au système de prise de vues. Une centrale de navigation inertielle (INS) et un système de positionnement par satellite (GPS) sont aussi embarqués à bord du véhicule. Ces senseurs permettent d'estimer la position de celui-ci et donc des deux caméras. Cet équipement permet d'implanter divers algorithmes pour réaliser le suivi des panneaux.

Ayant détecté un panneau à un instant, nous souhaitons être capable de le localiser dans l'espace avec une précision suffisante pour le situer sur une carte, mais aussi de déterminer sa position et sa taille dans les images précédentes avec le maximum de précision, (ce pour réduire le temps de recherche) et de s'assurer de sa visibilité. Pour cette seconde opération une solution récursive est nécessaire.

Le suivi des éléments discrets que sont les panneaux peut s'implanter en employant des séquences mono ou binoculaires. La vision stéréo semble en principe moins sujette au problème de convergence lente (Zhang et Faugeras, 1992; Faugeras, Ayache et Zhang, 1988; Chandrashekhar et Chellappa, 1991). En milieu routier les panneaux seront bien souvent plus visibles par la caméra de droite, et ce d'autant plus si la ligne de base est grande (1 à 2 m). Le problème de l'ambiguïté sur la profondeur ne se pose pas en raison des informations fournies par le système de positionnement embarqué. La caméra de gauche permet alors de compléter l'information.

L'estimation de la position du panneau dans l'espace peut se faire soit de manière récursive soit en une fois, à la fin de la poursuite. La solution récursive avec une seule séquence souffre du problème de l'initialisation. Comme nous l'avons déjà dit l'emploi

de la stéréoscopie réduirait cet inconvénient, mais à condition que le panneau soit visible dans les deux images. Une estimation a posteriori permet d'utiliser simplement les informations, complètes ou partielles, provenant des deux séquences. Nous allons maintenant voir comment ces deux solutions s'intègrent dans le système complet.

5.3.4 Intégration

Le taux d'acquisition d'images est réglé sur l'avancement du véhicule : il est en effet sans intérêt de poursuivre des panneaux dans une image statique, de plus cela complique inutilement les structures de données en obligeant le programme à être capable de conserver un nombre quelconque d'images où un panneau peut être présent. L'odomètre du véhicule permet de connaître le déplacement de celui-ci et donc d'espacer les prises de vues. Si l'on connaît la vitesse maximale d'opération et le taux maximal auquel on peut traiter les images, on obtient alors le déplacement maximum : à titre d'exemple il est de 0,65 m à 70 km/h pour un taux d'acquisition de 30 images/s et de 5,56 m à 100 km/h pour 5 images/s.

L'algorithme suit le schéma bloc de la figure 5.1. Les trois premiers modules correspondent aux phases de segmentation et de détection. La phase suivante consiste à poursuivre les panneaux détectés dans les images précédemment produites par la caméra de droite puis celles produites par la caméra de gauche. Comme les panneaux sont situés le plus souvent à droite de la chaussée, la caméra située sur la droite du véhicule a toujours plus de chances de voir ces panneaux que celle de gauche, surtout

lorsqu'ils sont près. La poursuite dans les images prises par la caméra de gauche est donc initiée à partir des informations obtenues dans les images de droite afin de pouvoir apparier les différentes apparitions d'un panneau même s'il n'est pas visible à gauche dès le début. Cette poursuite qui associe suivi et détection est décrite par le schéma bloc de la figure 5.2. Les deux dernières opérations sont la localisation et la mise à jour de la liste des panneaux détectés.

La phase de recherche des panneaux dans les images précédentes (schéma bloc de la figure 5.2), commence par une sélection des panneaux à poursuivre. Une fois la prédiction effectuée, la recherche consiste à appliquer l'algorithme de détection à la sous-image correspondant à la zone de recherche prédite. Si pour une raison quelconque un panneau s'avère indétectable, il convient de modifier la zone de recherche pour l'image suivante, la position estimée deux images plus tôt étant moins fiable; après plusieurs échecs de ce type, les chances de fusionner deux panneaux devenant non négligeables, il vaut mieux arrêter la poursuite. Dans la mesure où les panneaux se trouvent souvent groupés, dès que la taille de la zone de recherche est double de celle du panneau les risques de confusion sont élevés. En effet, si à chaque échec on double les dimensions de la zone de recherche, avec une erreur de mesure en position d'environ 10 % de la taille, après le second échec les dimensions de la zone de recherche vont avoisiner les 80 pixels, pour des panneaux de 40 à 50 pixels de large : il vaut mieux alors arrêter la poursuite. Quand il s'avère qu'un des panneaux a déjà été suivi ou qu'il n'est plus détectable on termine sa poursuite et l'on passe à sa

localisation puis à la mise à jour de la liste des panneaux. Celle-ci est décrite par le schéma bloc de la figure 5.3. Nous allons maintenant présenter en détail l'algorithme employé pour le suivi des panneaux.

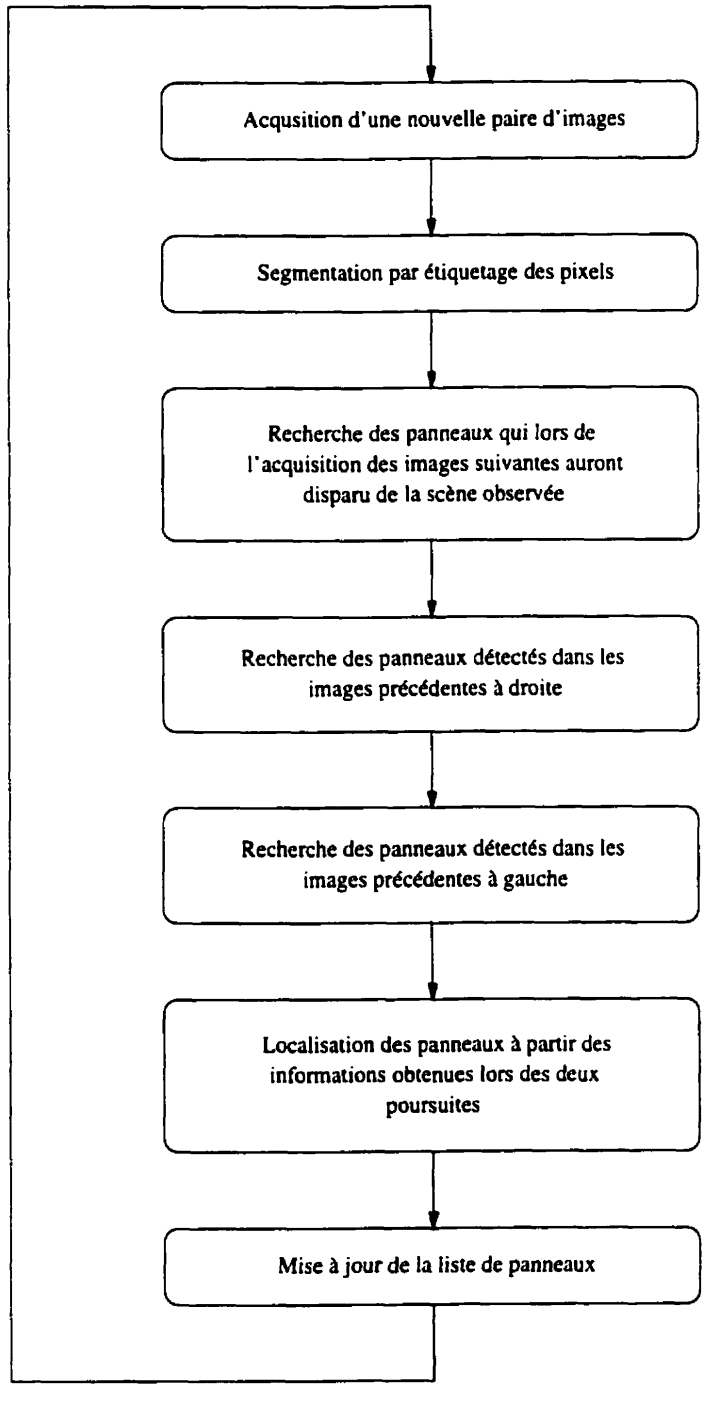


Figure 5.1 : Organigramme pour l'emploi du mouvement et de la stéréo.

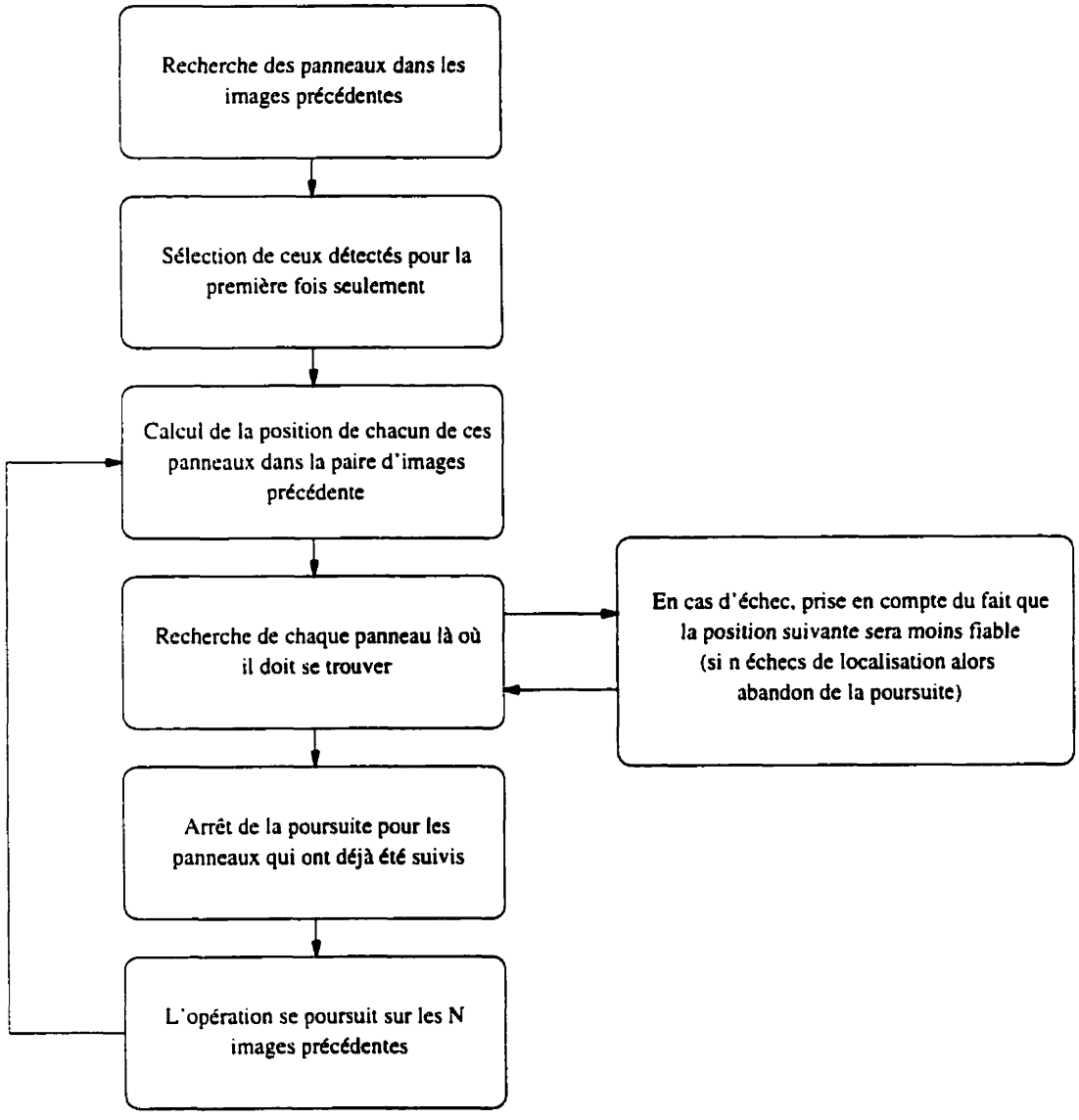


Figure 5.2 : Organigramme pour la recherche d'un panneau dans les images précédentes.

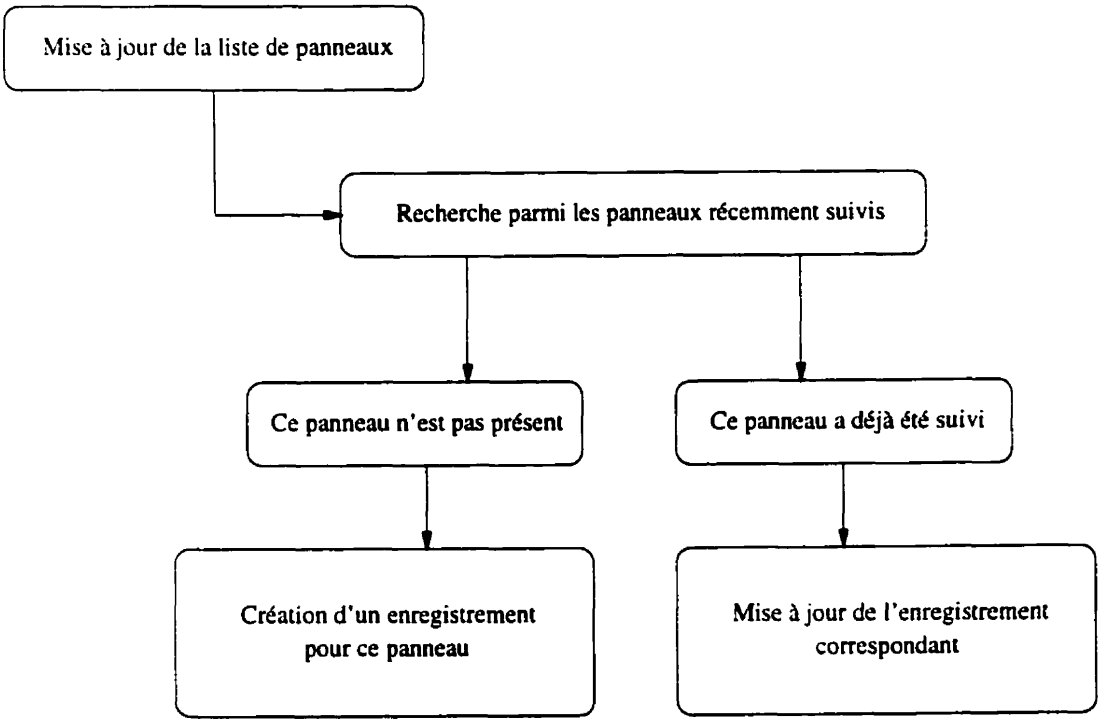


Figure 5.3 : Organigramme pour la mise à jour de la liste de panneaux.

5.4 Suivi des panneaux

5.4.1 Filtrage de Kalman

La solution retenue est celle d'un filtre de Kalman étendu (Chui et Chen, 1991) appliqué à une séquence monoculaire. Les paramètres de position et d'orientation de la caméra ne sont pas pris en compte dans le vecteur d'état. Les dispositifs de positionnement associant une centrale inertielle et un système de positionnement par satellite incorporent un filtrage de Kalman afin de fusionner les données provenant des deux senseurs. Il est donc inutile de filtrer à nouveau ces données. La phase de prédiction du filtrage nous permet d'estimer à la fois la position dans l'image suivante du panneau et l'incertitude sur cette position : cela nous permet de définir une zone de recherche.

Deux formulations du problème ont été développées. Leur différence réside dans le choix du repère de travail : la première emploie un repère géocentré et est citée ici pour mémoire, car conduisant à des problèmes de calculs numériques; la seconde utilise le repère de la caméra. En termes de complexité algorithmique elles sont équivalentes (Matthies et Shafer, 1987).

Les deux modèles d'état développés se présentent sous la forme suivante :

$$\mathbf{x}_{k+1} = \mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{u}_k + \underline{\xi}_k,$$

$$\mathbf{v}_k = g_k(\mathbf{x}_k) + \underline{\eta}_k,$$

où $\mathbf{x}_k$, $\mathbf{u}_k$ et $\mathbf{v}_k$ sont respectivement le vecteur d'état, la commande et la mesure à l'instant k, et où $\underline{\xi}_k$ et $\underline{\eta}_k$ sont les erreurs respectivement sur le modèle et la mesure, et qui sont considérées être des bruits blancs gaussiens centrés et décorrélés. L'objectif est de déterminer à chaque instant l'estimateur $\hat{\mathbf{x}}_{k|k}$ du vecteur d'état minimisant la variance de l'erreur, en partant d'une estimation à l'instant 0 $\hat{\mathbf{x}}_{0|0}$. Connaissant l'estimateur à l'instant k-1 La partie déterministe de l'équation d'état donne la prédiction du vecteur d'état pour l'instant k :

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1} \cdot \hat{\mathbf{x}}_{k-1|k-1} + \mathbf{u}_{k-1}.$$

Pour pouvoir réaliser l'estimation à l'instant k, nous devons linéariser l'équation de mesure au voisinage de la prédiction $\hat{\mathbf{x}}_{k|k-1}$ en utilisant l'approximation de Taylor suivante (Chui et Chen, 1991) :

$$g_k(\mathbf{x}_k) \simeq g_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{C}_k \cdot (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}).$$

où $\mathbf{C}_k$ est la matrice jacobienne en $\hat{\mathbf{x}}_{k|k-1}$). Nous allons maintenant voir les repères employés dans les deux modèles développés.

5.4.2 Repères utilisés et définitions

Le système de positionnement par satellite étant une des premières sources d'information, on se place à l'échelle de la terre. Nous considérons les trois repères suivants (figure 5.4) :

- repère terrestre : centre de la terre ($\mathbf{t}$), axe z vers le pôle nord, axe x correspondant à la longitude 0;
- repère terrestre local : centre de la caméra ($\mathbf{o}$), axe z radial par rapport à la terre, axe x vers l'est;
- repère associé à la caméra : centre de la caméra ($\mathbf{o}$), axe y selon l'axe optique.

Nous utilisons l'indice T pour désigner l'expression d'une variable dans le repère terrestre et l'indice C pour son expression dans le repère de la caméra. Nous réservons les indices en minuscules pour référer à un instant donné. Un point $\mathbf{m}$ représenté dans le repère terrestre par $\mathbf{m}_T$ est représenté dans le repère de la caméra par $\mathbf{m}_C$ et nous avons la relation suivante :

$$\mathbf{m}_C = \mathbf{R}_k \cdot \mathbf{m}_T + \mathbf{t}_C,$$

où $\mathbf{R}_k$ est la matrice définissant la rotation entre les deux repères à l'instant k. Elle s'exprime sous la forme suivante :

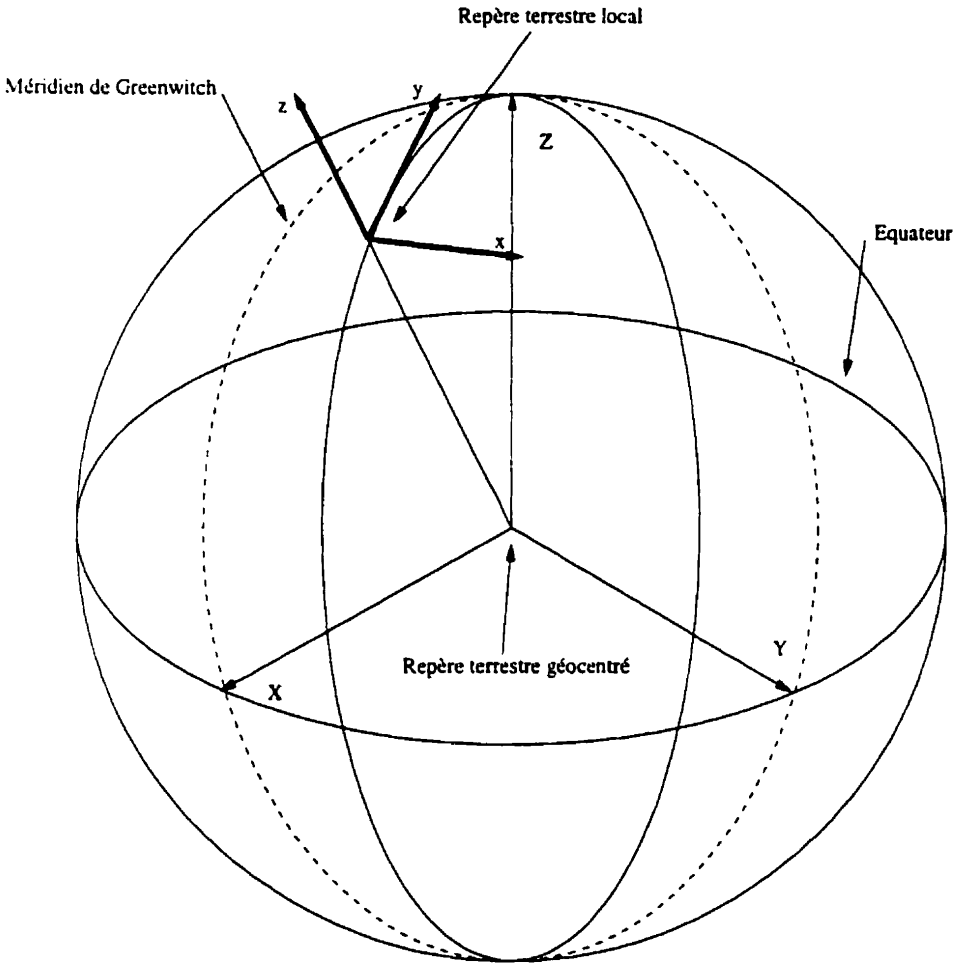


Figure 5.4 : Les deux principaux repères utilisés

$$\mathbf{R}_k = \begin{pmatrix} c\kappa & 0 & -s\kappa \\ 0 & 1 & 0 \\ s\kappa & 0 & c\kappa \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & c\tau & -s\tau \\ 0 & s\tau & c\tau \end{pmatrix} \cdot \begin{pmatrix} c\omega & s\omega & 0 \\ -s\omega & c\omega & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -s\phi & c\phi & 0 \\ -s\theta c\phi & -s\theta s\phi & c\theta \\ c\theta c\phi & c\theta s\phi & s\theta \end{pmatrix}.$$

où $c\alpha$ et $s\alpha$ désignent respectivement le cosinus et le sinus de l'angle α. κ désigne le roulis, τ le tangage, ω l'azimut, θ la latitude et ϕ la longitude. Ainsi, si le GPS fournit la position de la caméra $\mathbf{o}_T$, alors

$$\mathbf{t}_C = -\mathbf{R}_k^T \cdot \mathbf{o}_T,$$

où l'exposant T désigne la transposition. Nous désignons par $\mathbf{P}$ et $\mathbf{Q}$ les deux matrices suivantes :

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}.$$

Nous notons $\mathbf{0}_N$ un vecteur nul de dimension N et $\mathbf{I}_N$ une matrice identité de dimension N. Nous présentons maintenant deux modèles d'état pour le filtre de Kalman

5.4.3 Premier modèle

On se place dans le référentiel terrestre : le panneau est fixe. Le vecteur d'état est :

$$\mathbf{x} = \begin{bmatrix} X_T \\ Y_T \\ Z_T \\ D_T \end{bmatrix},$$

où X_T, Y_T, Z_T désignent les coordonnées du panneau dans le repère terrestre et D_T une de ses dimensions. Il est évidemment possible de prendre en compte les deux dimensions du panneau. Cela n'a pas été fait car, en pratique, pour un type de panneau donné, une dimension suffit à définir le masque. L'équation d'état est :

$$\mathbf{x}_{k+1} = \mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{u}_k + \underline{\xi}_k,$$

où

$$\mathbf{A}_k = \mathbf{I}_4$$

et

$$\mathbf{u}_k = \mathbf{0}_4.$$

La mesure est :

$$\mathbf{v} = \begin{bmatrix} r \\ s \\ d \end{bmatrix},$$

où r, s désignent la position d'un panneau dans l'image et d la taille du masque. En supposant que les panneaux soient perpendiculaires à la route et donc à l'axe optique (axe des y pour la caméra) et en utilisant un modèle de caméra sténopé, l'équation de mesure avant linéarisation est :

$$\mathbf{v}_k = g_k(\mathbf{x}_k) + \underline{\eta}_k$$

avec

$$g_k(\mathbf{x}_k) = \frac{f}{Y_C} \cdot \mathbf{P} \cdot \left(\begin{bmatrix} \mathbf{R}_k & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \mathbf{x}_k + \begin{bmatrix} \mathbf{t}_C \\ 0 \end{bmatrix} \right),$$

où f est la distance focale de la caméra et où

$$Y_C = \mathbf{Q} \cdot \left(\begin{bmatrix} \mathbf{R}_k & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \mathbf{x}_k + \begin{bmatrix} \mathbf{t}_C \\ 0 \end{bmatrix} \right)$$

est l'ordonnée du vecteur d'état exprimée dans le repère de la caméra. Nous linéarisons l'équation de mesure au voisinage de $\hat{\mathbf{x}}_{k|k-1}$, prédiction pour le vecteur d'état à l'instant k, au moyen de l'approximation suivante :

$$g_k(\mathbf{x}_k) \simeq g_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{C}_k \cdot (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}).$$

avec

$$\mathbf{C}_k = \frac{f}{\hat{Y}_{C_{k|k-1}}}\mathbf{P} \cdot \left(\mathbf{I}_4 - \frac{1}{\hat{Y}_{C_{k|k-1}}} \left(\begin{bmatrix} \mathbf{R}_k & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \hat{\mathbf{x}}_{k|k-1} + \begin{bmatrix} \mathbf{t}_C \\ 0 \end{bmatrix} \right) \cdot \mathbf{Q} \right) \cdot \begin{bmatrix} \mathbf{R}_k & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix}.$$

où $\hat{Y}_{C_{k|k-1}}$ est l'ordonnée du point de linéarisation exprimé dans le repère de la caméra.

5.4.4 Second modèle

La solution développée précédemment présente l'inconvénient pour les calculs numériques de manipuler en même temps des variables avec des ordres de grandeurs très différents : la taille des panneaux est de l'ordre du mètre alors que les coordonnées se comptent en millier de kilomètres; la mesure implique en outre une projection dans le repère de la caméra et donc on trouve des valeurs de l'ordre de la dizaine de mètres et de micromètres; le problème se retrouve dans les matrices utilisées par le filtre, ce qui ne facilite pas les calculs numériques. Il convient d'adopter un repère local. Un repère fixe présenterait l'inconvénient de n'être bien adapté que pour quelques images, et devrait en outre être modifié régulièrement. En se plaçant à chaque instant dans le repère de la caméra on évite ce problème. Cela a néanmoins pour conséquence une équation d'état un peu plus complexe. Le vecteur d'état est maintenant

$$\mathbf{x} = \begin{bmatrix} X_C \\ Y_C \\ Z_C \\ D_C \end{bmatrix},$$

où X_C, Y_C, Z_C sont les coordonnées du panneau dans le repère de la caméra, et D_C une de ses dimensions. L'équation d'état s'écrit sous la forme suivante :

$$\mathbf{x}_{k+1} = \mathbf{A}_k \cdot \mathbf{x}_k + \mathbf{u}_k + \underline{\xi}_k,$$

où

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{R}_{k+1} \cdot \mathbf{R}_k^T & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix}$$

et

$$\mathbf{u}_k = \begin{bmatrix} \mathbf{R}_{k+1} & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \left(\begin{bmatrix} \mathbf{o}_{T_k} \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbf{o}_{T_{k+1}} \\ 0 \end{bmatrix} \right).$$

Le vecteur de mesure est inchangé. L'équation de mesure est la suivante :

$$\mathbf{v}_k = g_k(\mathbf{x}_k) = \frac{f}{Y_C}\mathbf{P} \cdot \mathbf{x}_k + \underline{\eta}_k$$

Nous la linéarisons au voisinage de $\hat{\mathbf{x}}_{k|k-1}$, prédiction pour le vecteur d'état à l'instant k, au moyen de l'approximation suivante :

$$g_k(\mathbf{x}_k) \simeq g_k(\hat{\mathbf{x}}_{k|k-1}) + \mathbf{C}_k \cdot (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}),$$

avec

$$\mathbf{C}_k = \frac{f}{\hat{Y}_{C_{k|k-1}}} \left(\mathbf{I}_4 - \frac{1}{\hat{Y}_{C_{k|k-1}}} \mathbf{P} \cdot \hat{\mathbf{x}}_{k|k-1} \cdot \mathbf{Q} \right).$$

Le filtre de Kalman est un outil d'estimation: dans le cadre de cette application c'est cependant sa phase de prédiction qui nous intéresse et nous allons maintenant voir comment nous l'exploitons.

5.4.5 Prédiction de la zone de recherche

Le filtrage nous donne pour l'instant k une prédiction $\hat{\mathbf{x}}_{k|k-1}$ du vecteur d'état, ainsi que la matrice de covariance de ce vecteur $\mathbf{X}_{k|k-1}$. Grâce à l'équation de mesure nous obtenons une prédiction de la mesure :

$$\hat{\mathbf{v}}_{k|k-1} = g_k(\hat{\mathbf{x}}_{k|k-1})$$

et sa matrice de covariance

$$\mathbf{V}_{k|k-1} = \mathbf{C}_k \cdot \mathbf{X}_{k|k-1} \cdot \mathbf{C}_k^T.$$

Nous pouvons alors déterminer une zone de recherche autour de la position prédite dans l'image en fonction de la variance de ces deux coordonnées. Le suivi permet alors l'appariement des différentes apparitions du panneau. L'étape suivante est donc

la localisation de celui-ci. Nous conservons les notations employées ici pour la présentation de cet algorithme.

5.5 Localisation des panneaux a posteriori

5.5.1 Introduction

Pour chaque panneau que nous avons pu poursuivre, nous disposons d'un certain nombre de vues où la position du panneau est connue. Ces images ont été prises par les deux caméras dont les positions absolues à ces différents instants sont connues.

5.5.2 Formulation

La position du panneau dans l'espace n'est pas une variable aléatoire, et, en supposant que la moyenne du bruit sur la mesure est nulle, nous connaissons le premier moment de la mesure. Cela est suffisant pour développer un estimateur des moindres carrés. Nous obtenons :

$$E(\mathbf{v}_k) = \frac{f}{Y_{C_k}}\mathbf{P} \cdot \left(\begin{bmatrix} \mathbf{R}_k \cdot \mathbf{R}_0^T & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \mathbf{x}_0 + \begin{bmatrix} \mathbf{R}_k & \mathbf{0}_3 \\ \mathbf{0}_3^T & 1 \end{bmatrix} \cdot \left(\begin{bmatrix} \mathbf{o}_{T_0} \\ 0 \end{bmatrix} - \begin{bmatrix} \mathbf{o}_{T_k} \\ 0 \end{bmatrix} \right) \right).$$

Nous cherchons à estimer $\mathbf{x} = \mathbf{x}_0$, localisation du panneau par rapport à la dernière position de la caméra pour laquelle il a été visible. Pour cela nous cherchons $\mathbf{x}$ qui minimise

$$\sum_{k=0}^{N}(\mathbf{v}_k - E(\mathbf{v}_k))^T \cdot \mathbf{M} \cdot (\mathbf{v}_k - E(\mathbf{v}_k)),$$

où $\mathbf{M}$ est une matrice définie positive. Il est possible de prendre

$$\mathbf{M} = \mathbf{R}_{\mathbf{v}_k}^{-1},$$

où $\mathbf{R}_{\mathbf{v}_k}$ est la matrice de covariance du bruit qui s'ajoute à la mesure, si celle-ci est connue. Cette formulation conduit à une solution non-linéaire. Nous envisageons donc le problème sous un autre angle. Il peut se formuler de la façon suivante :

$$\mathbf{v}_k = f_k(\mathbf{x}) + \underline{\zeta}_k \tag{5.1}$$

où $\underline{\zeta}$ est le bruit additif sur la mesure. Si l'on multiplie l'équation 5.1 par Y_{C_k} le problème devient linéaire en $\mathbf{x}$ et s'écrit sous la forme suivante :

$$\mathbf{L}\mathbf{x} = \mathbf{j} + \underline{\varepsilon},$$

où

$$\underline{\varepsilon} = \begin{bmatrix} Y_{C_0}\underline{\zeta}_0 \\ \vdots \\ Y_{C_N}\underline{\zeta}_N \end{bmatrix}.$$

Il s'agit donc de résoudre un système linéaire surdéterminé. Cela est réalisé en réalisant une décomposition en valeurs singulières de **L** (Press, Flannery, Teukolshy et Vetterling, 1988). La fin de ce chapitre est consacrée à l'évaluation des algorithmes de localisation et de suivi.

5.6 Simulation de l'inspection d'une portion de route

Comme la position et l'orientation de la caméra aux instants de prise de vue des images auxquelles nous avons accès, ne sont pas connue, afin de pouvoir tester nos algorithmes de suivi et de localisation, nous avons simulé le déplacement d'un véhicule le long d'une route ainsi que les autres fonctionnalités embarquées. Cela nous permet de tester le suivi et la localisation sous diverses conditions. Nous n'avons toutefois pas créé d'images de synthèse sur lesquelles la détection des panneaux eut été complètement artificielle : il a été jugé préférable de simuler en même temps la projection du panneau dans l'image et sa détection, autant pour des questions de réalisme que de temps de calcul.

La simulation se décompose en plusieurs étapes. Une trajectoire est d'abord sélectionnée, puis on place des panneaux le long de cette trajectoire. Le déplacement

d'une paire de caméras le long de cette trajectoire est alors simulé et à chaque instant on peut obtenir la projection des panneaux dans l'image et déterminer ceux qui sont présentement visibles.

Pour simplifier l'interaction avec le programme de suivi, nous produisons un fichier contenant pour chaque instant la projection des panneaux dont c'est la dernière apparition ainsi qu'un second fichier, similaire au premier mais correspondant à toutes les autres apparitions. Cela s'obtient aisément en parcourant la trajectoire en sens inverse, la dernière apparition se produisant alors avant toutes les autres. L'algorithme de suivi des panneaux se voit donc proposer des positions et des dimensions dans l'image mais l'appariement des panneaux détectés dans chaque image reste à la charge du programme de poursuite.

La position et l'orientation des caméras sont aussi fournies, sous la forme de données issues d'un système formé d'un récepteur GPS et d'une centrale inertielle. Notre intérêt se porte donc sur la simulation d'une trajectoire le long de laquelle des panneaux sont placés, du système de vision, de la détection des panneaux et du système de positionnement.

5.6.1 Simulation d'une trajectoire

Les trajectoires sont simulées en respectant le modèle et les contraintes suivants. La terre est sphérique ce qui permet des calculs de positions simples à partir de la latitude, de la longitude et de l'altitude. Les problèmes d'asphéricité de la terre ne

devrait en pratique pas se faire sentir à l'échelle du suivi d'un panneau qui a lieu au maximum sur quelques centaines de mètres. Néanmoins, cette asphéricité peut constituer un problème lors du parcours de grandes distances sur le globe. La route à suivre est un arc selon un grand cercle de 0.1 degrés selon la latitude et la longitude. 50 panneaux sont régulièrement espacés le long de cette trajectoire de base et disposés selon une distribution uniforme dans une section carré de 20 m de côté. Les positions théoriques possibles pour les panneaux le long d'une route rectiligne sont en réalité beaucoup plus limitées, mais elles dépendent des types de routes et de panneaux. Une telle distribution nous permet donc de traiter un plus grand nombre des situations possibles. La largeur de ces panneaux est de 0.5 m.

La vitesse de parcours du véhicule est de 70 km/h ce qui correspond à la vitesse d'opération du camion de GIE Technologies. Nous modélisons le mouvement du véhicule de façon très simple, en ajoutant à la trajectoire de base des écarts sinusoïdaux :

- latéralement : 0.3 m d'amplitude, période de 1 s;
- verticalement : 0.2 m d'amplitude, période de 3 s.

Les périodes sont différentes afin d'avoir plus d'écarts différents. L'axe du système de vision reste tangent à la trajectoire suivie, et il n'y a pas de roulis. La simulation de ce système fait l'objet de la prochaine sous-section.

5.6.2 Simulation du système de vision

Nous simulons un système de vision stéréoscopique idéal : les deux caméras sont identiques, leurs axes optiques sont parallèles entre eux et sont orthogonaux à la ligne de base, laquelle mesure 1 m. Les caméras sont de type sténopé, avec un capteur parfaitement plan et orthogonal à l'axe optique. Ce capteur a les dimensions d'un senseur CCD pour la télévision dont la diagonale mesure 2 3". Sa résolution est de 768 pixels horizontalement et de 480 pixels verticalement. Les dimensions de chaque pixel sont 11,6 μm par 13,6 μm. La distance focale de l'objectif est de 16 mm. Nous avons choisi un taux d'acquisition de 5 images par seconde : à titre d'exemple, en 1993 le système de reconnaissance automatique des panneaux développé à l'initiative de Mercedes ne fonctionnait qu'à trois images par seconde avec des composants spécialisés.

La formation des images se limite à déterminer la position et la largeur des panneaux dans l'image. La sous-section suivante va nous permettre de montrer comment simuler la détection à partir de ces informations.

5.6.3 Simulation de la détection des panneaux

Nous considérons la détection comme une mesure introduisant une erreur, que nous simulons à l'aide de générateurs de nombres pseudo aléatoires approximant leurs distributions de probabilités. Ces erreurs de mesure dans l'image sont modélisées par des bruits blancs gaussiens centrés. Matthies et Shafer (Matthies et Shafer, 1987) ont

montré que ce modèle était suffisamment réaliste pour les erreurs de position dans les images.

Pour une distribution normale, 99,7% des valeurs se trouvent dans un intervalle de plus ou moins trois fois l'écart type autour de la moyenne. Nous considérons que l'erreur sur la mesure doit se trouver dans une telle plage : en pratique cette erreur ne peut qu'être bornée, car le masque de corrélation doit avoir un minimum de contact avec la ou les régions segmentées correspondant au panneau. Dans le cas contraire nous serions en présence d'une fausse alarme, ce qui est un problème bien différent.

Pour ce qui est de la taille, notre programme de détection emploie des masques capables en théorie de détecter des formes avec deux pixels d'écart. De plus, les panneaux, détectés grâce à leur bordure, peuvent l'être par deux masques distincts (un sur le contour externe, l'autre sur le contour interne). Sur les images dont nous disposons cela correspond à un écart de 20% entre la taille mesurée et la taille réelle.

Il reste à prendre en compte les imperfections de la segmentation. Deux phénomènes peuvent se produire : les erreurs de classification de pixels liées au bruit, et la dégradation de la teinte aux frontières des panneaux. Si l'on suppose que les erreurs de classification dues au bruit suivent une loi de répartition uniforme, alors au stade de la corrélation, elles n'ont pas d'influence sur la position et la taille, mais se traduisent par une diminution de la qualité de la détection. En revanche, hormis le cas où les bords du panneau correspondent exactement à la jonction entre les pixels (sans tenir compte des propriétés de l'objectif), il existe une bande de transition qui peut

être incluse ou non dans la région segmentée. La segmentation peut donc agrandir ou réduire, dans les deux directions, le panneau d'un demi pixel. Cela se traduit par une erreur de plus ou moins un pixel dans le pire cas sur la taille.

Ces considérations peuvent être pareillement appliquées à la position. Nous arrivons donc aux écarts types suivants :

- pour la position : $\sigma = \frac{d \times 0.1 + 1}{3}$;
- pour la taille : $\sigma = \frac{d \times 0.2 + 1}{3}$;

où d désigne la dimension du panneau dans l'image. Nous allons appliquer ces règles de variations de l'écart type pour simuler les erreurs sur la position et l'orientation du système de vision.

5.6.4 Simulation du système de positionnement

Les erreurs des mesures de positionnement par GPS et de l'orientation sont modélisées de la même manière. L'erreur maximale pour un GPS différentiel est de 2 m dans le plan horizontal et de 5 m pour l'altitude. En pratique sur les véhicules d'arpentage le GPS est couplé avec une centrale de navigation inertielle ayant des performances du niveau de celles employées encore récemment dans l'avionique militaire. Les erreurs maximales atteintes sont alors de 0.3 m (pour chaque direction) et inférieures à 0.1 degrés pour les différents angles. Dans la mesure où les sources pour ces erreurs sont commerciales et comme nous ignorons comment le couplage est utilisé pour améliorer

l'estimation de la position, afin d'avoir une certaine marge de sécurité, nous rajoutons un biais sur la position. Ce biais ne saurait excéder l'erreur maximale d'un GPS différentiel sinon le couplage avec la centrale de navigation inertielle n'aurait aucun intérêt. De plus comme nous ignorons comment se fait ce couplage nous avons réalisé des simulations avec des variations instantanées du biais. Nous abordons maintenant les expériences réalisées sur la base de ces simulations.

5.7 Expériences

5.7.1 Objectifs

Notre premier objectif est d'estimer les performances de la prédiction de la position des panneaux dans l'image, car de mauvaises prédictions rendent inutile le couplage du suivi et de la détection et peuvent conduire à une perte du panneau. Nous étudions ensuite l'efficacité du suivi, c'est-à-dire la capacité à trouver le panneau dans un maximum d'images même après l'avoir éventuellement manqué : plus le panneau a été détecté plus on dispose en effet d'informations pour le localiser. Et finalement nous nous intéressons à la précision de la localisation.

Les trois expériences correspondantes sont réalisés à l'aide de quatre des simulations d'inspection précédemment décrites en introduisant sur la position les biais suivants :

1. aucun biais.

2. (2 m, 2 m, 5 m) dans le repère terrestre local, période 6 s (soit 30 images), que nous désignerons sous l'expression biais de période 6 s.

3. (2 m, 2 m ,5 m) dans le repère terrestre local, période 12 s (soit 60 images), que nous désignerons sous l'expression biais de période 12 s.

4. ($\pm$1m, $\pm$1m, $\pm$1m) dans le repère terrestre local, les valeurs changeant aléatoirement toutes les 6 s (soit 30 images), que nous désignerons sous l'expression biais aléatoire.

Nous n'introduisons pas de biais constant car, comme la distance sur laquelle un panneau est poursuivi est trop faible pour que la sphéricité de la terre se fasse sentir, cela revient à translater le problème et conduit de ce fait à localiser les panneaux avec une erreur supplémentaire, égale à ce biais. Les autres erreurs simulées (y compris la partie gaussienne de l'erreur sur la position du système de vision) sont les mêmes quel que soit le biais choisi. Avant de décrire ces trois expériences, nous présentons les choix effectués pour l'initialisation du filtre et le dimensionnement de la zone de recherche.

5.7.2 Initialisation du filtre

Même si dans le cadre de ces simulations, en raison de la répartition uniforme des panneaux, les caméras ont a priori la même chance de détecter un panneau, la caméra de droite suit d'abord le panneau, puis transmet son estimation de la position de celui-

ci à la caméra de gauche qui retrace alors à nouveau le panneau.

Nous initialisons le filtre sur la base de la première détection en estimant que la largeur du panneau est de 0.7 m. En pratique les normes réelles ne sont pas assez strictes pour que l'on puisse connaître les dimensions des panneaux sur la base du type de route et de panneau. Nous surestimons donc cette dimension car avec une telle initialisation le filtre de Kalman étendu diverge dans l'espace (il converge toutefois dans l'image). Ainsi nous ne nous plaçons pas dans le cas le plus favorable.

5.7.3 Dimensions de la zone de recherche

La zone de recherche est estimée à partir de la variance de la prédiction de la position dans l'image. Si celle-ci était une variable gaussienne, il suffirait de prendre six fois l'écart type correspondant pour obtenir 99.7% des panneaux. En pratique cela n'est certainement pas le cas et comme le biais sur la position des caméras n'est pas pris en compte il convient de prendre une valeur plus élevée. En restreignant cette zone, nous pouvons manquer artificiellement des panneaux. Après un échec les dimensions la zone de recherche dans l'image suivante sont doublées. Dans le cas de l'étude de la prédiction dans l'image nous nous assurons de ne pas avoir de panneaux manqués en prenant neuf fois l'écart type.

5.7.4 Étude de la prédiction dans l'image

Afin de pouvoir évaluer indépendamment l'algorithme de prédiction de la position des panneaux dans l'image, le panneau doit toujours pouvoir être trouvé (s'il est visible). Cela est obtenu en surdimensionnant la zone de recherche maximale. A l'intérieur de cette zone de recherche maximale, le choix se porte sur le pic de corrélation le plus proche de la position prédite : c'est le moyen le plus sûr d'éviter de confondre deux panneaux en l'absence d'un couplage des suivis pouvant avoir lieu en parallèle (il est évidemment toujours possible de trouver une fausse alarme plus proche). A partir de la prédiction, pour trouver le panneau il a donc fallu explorer au moins une région carrée centrée sur la position prédite et contenant le pic de corrélation. Le temps de calcul requis pour cela est proportionnel au nombre de pixels de cette région, c'est-à-dire l'aire de celle-ci. Nous jugeons de l'efficacité de la prédiction de la position d'un panneau dans l'image d'après l'aire de cette zone de recherche. Comme la recherche d'un panneau est limitée à l'image, nous rapportons cette aire à celle de l'image et obtenons alors le facteur de réduction de la zone de recherche. Ce facteur est aussi le facteur de réduction du temps de calcul réalisé grâce au suivi. Nous nous intéressons donc à la valeur maximale de ce facteur, ce qui correspond au pire cas, à sa valeur moyenne et à son évolution au cours de la poursuite, et ce pour les divers biais envisagés pour la position. Pour caractériser l'évolution de ce facteur au cours de la poursuite, nous considérons la moyenne ainsi que la médiane, car la moyenne est très influencée par les valeurs les plus élevées du facteur de réduction, celui-ci

correspondant à une aire (c'est en fait une sorte de moyenne des carrés de l'erreur maximale selon l'un des deux axes).

5.7.5 Étude du suivi

Nous réduisons les dimensions de la zone de recherche à deux fois l'écart type correspondant (ces dimensions peuvent toujours être doublées si pour l'image précédente, le panneau a été manqué) et nous comptabilisons le nombre d'images où le panneau a été manqué. Cela est évidemment fait pour les différents biais possibles. Les différentes apparitions détectés sont alors utilisées pour localiser les panneaux.

5.7.6 Étude de la localisation

Partant des informations obtenues lors du suivi, nous estimons la localisation puis la position des panneaux sur le globe. Connaissant leur position exacte nous déterminons les erreurs de positionnement. Nous considérons alors les erreurs moyennes et maximales, la médiane,l'écart type de l'erreur ainsi que la distribution de l'erreur en fonction de la distance entre le panneau et la position de la caméra de droite au début de la poursuite.

5.8 Résultats

5.8.1 Étude de la prédiction dans l'image

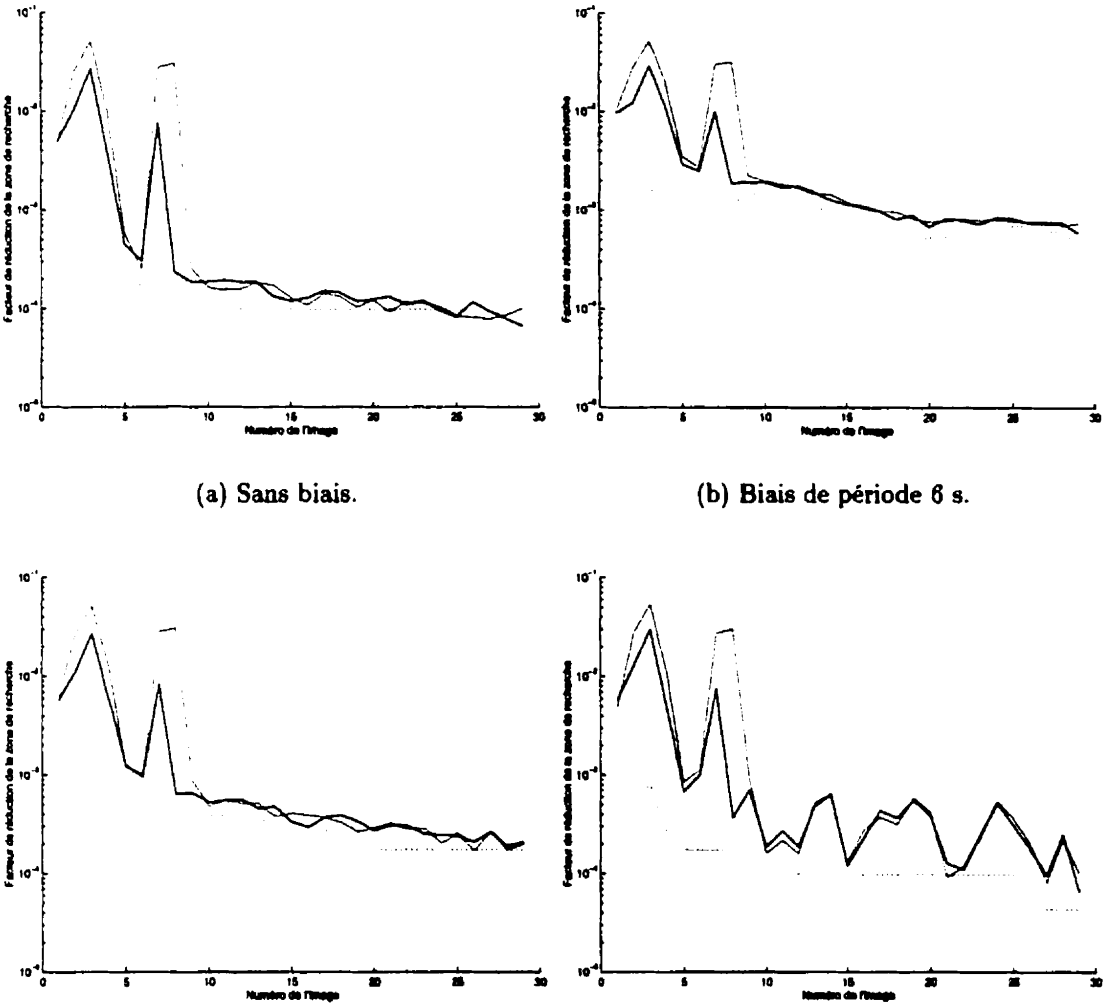
Les figures 5.5 à 5.8 montrent les évolutions de la moyenne et de la médiane du facteur de réduction au cours du suivi pour les quatre inspections qui ont été simulées. Le tableau 5.1 fournit les facteurs de réduction moyens et maximums pour la caméra de droite; le tableau 5.2 donne les informations correspondantes pour la caméra de gauche. Ce facteur peut dépasser la valeur de 0,1, cependant cette situation ne se produit que pour 1% des images.

Tableau 5.1 : Facteur de réduction par image moyen et maximal pour l'image de droite

Biais	sans	période 6 s	période 12 s	aléatoire
Facteur moyen	$5,29 \cdot 10^{-3}$	$6,94 \cdot 10^{-3}$	$5,69 \cdot 10^{-3}$	$5,66 \cdot 10^{-3}$
Facteur maximal	$1,35 \cdot 10^{-1}$	$1,38 \cdot 10^{-1}$	$1,36 \cdot 10^{-1}$	$1,35 \cdot 10^{-1}$

Tableau 5.2 : Facteur de réduction par image moyen et maximal pour l'image de gauche

Biais	sans	période 6 s	période 12 s	aléatoire
Facteur moyen	$2,74 \cdot 10^{-3}$	$4,15 \cdot 10^{-3}$	$3,07 \cdot 10^{-3}$	$3,14 \cdot 10^{-3}$
Facteur maximal	$8,10 \cdot 10^{-2}$	$8,10 \cdot 10^{-2}$	$8,07 \cdot 10^{-2}$	$8,04 \cdot 10^{-2}$



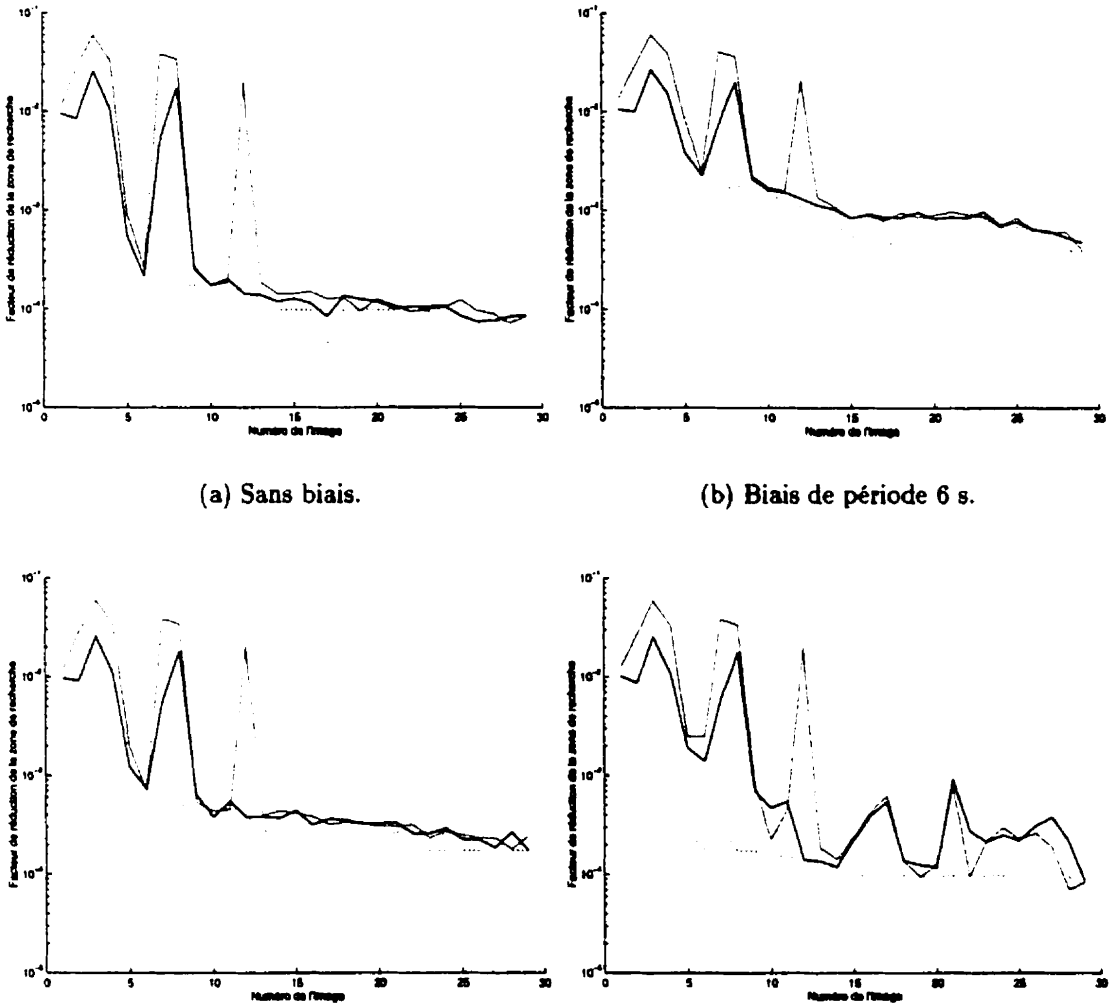
(a) Sans biais.

(b) Biais de période 6 s.

(c) Biais de période 12 s.

(d) Biais aléatoire.

Figure 5.5 : Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le premier trajet simulé.



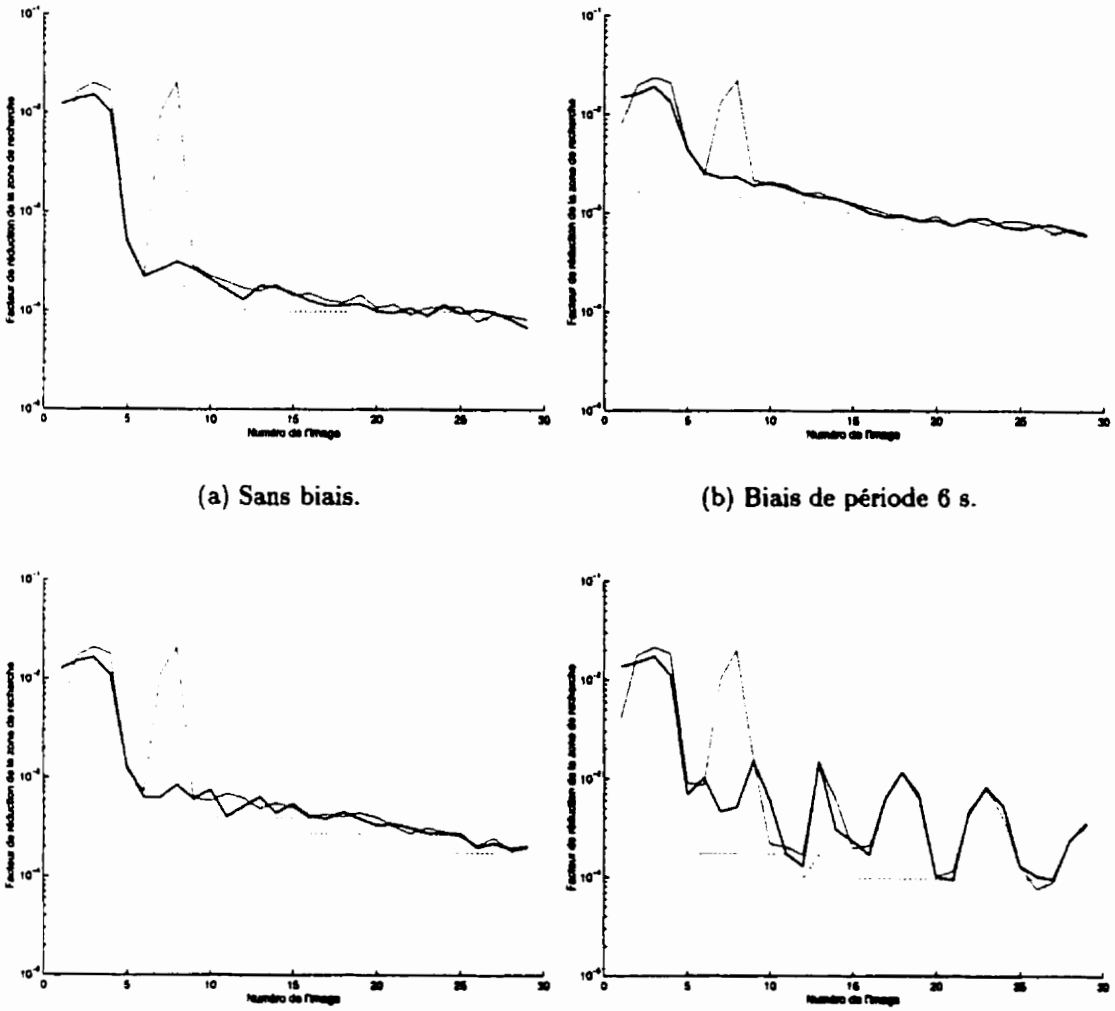
(a) Sans biais.

(b) Biais de période 6 s.

(c) Biais de période 12 s.

(d) Biais aléatoire.

Figure 5.6 : Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le second trajet simulé.



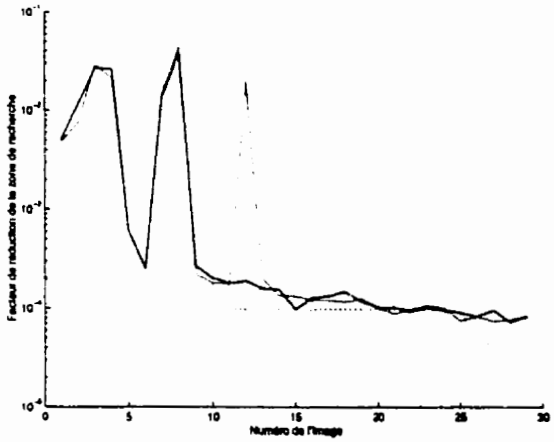
(a) Sans biais.

(b) Biais de période 6 s.

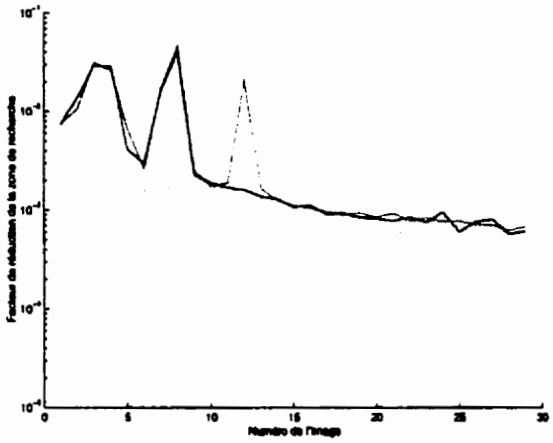
(c) Biais de période 12 s.

(d) Biais aléatoire.

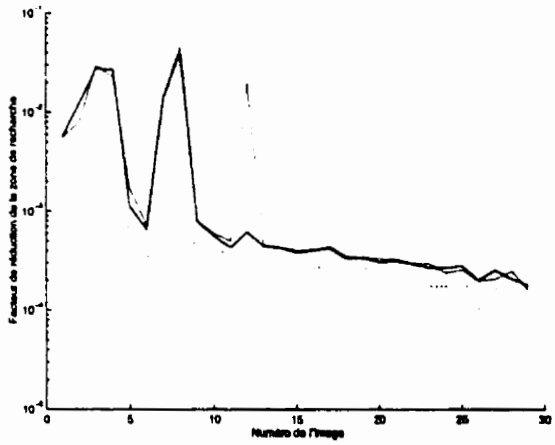
Figure 5.7 : Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le troisième trajet simulé.



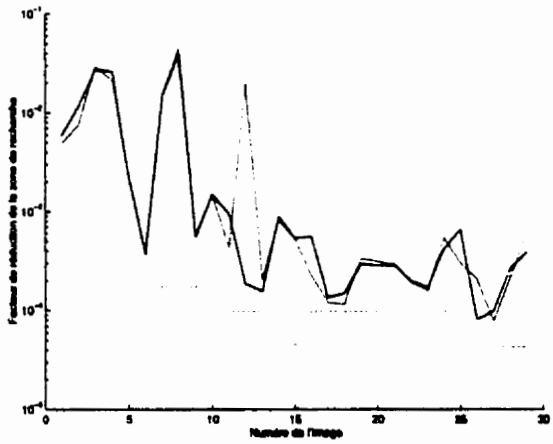
(a) Sans biais.



(b) Biais de période 6 s.



(c) Biais de période 12 s.



(d) Biais aléatoire.

Figure 5.8 : Facteur de réduction de la zone de recherche moyen (trait continu) et médian (trait pointillés) pour l'image de droite (rouge) et l'image de gauche (en bleu) pour le quatrième trajet simulé.

5.8.2 Suivi des panneaux

La caméra de droite ne manque aucun panneau. Ce n'est pas le cas pour celle de gauche : le tableau 5.3 indique les pourcentages maximums et moyens (par rapport à la durée de la poursuite) d'apparitions dans l'image de panneaux manquées. Cela prend en compte aussi bien les mauvaises prédictions que les cas où le panneau est en dehors du champ de vision.

Tableau 5.3 : Pourcentages d'apparitions de panneaux manqués pour l'image de gauche.

Biais	sans	période 6 s	période 12 s	aléatoire
Moyen	0,20	0,20	0,20	0,20
Maximum	0.036	0.036	0.036	0.047

5.8.3 Localisation des panneaux

Tableau 5.4 : Erreurs sur la position du panneau (en m), sans biais

Séquence	1	2	3	4
Erreur maximale	33,2	2,46	12,3	19,2
Erreur moyenne	1,77	1,11	1,18	1,75
Médiane	1,07	1,03	0,883	1,10
Écart type	4,57	0,637	1,68	3,10

Le tableau 5.4 montre les caractéristiques de l'erreur sur la position des panneaux sans introduction de biais parasite sur la position des caméras. Il faut aussi noter que pour 98% des panneaux l'erreur est inférieure à 3 m. La figure 5.9 montre la

distribution de l'erreur sur la position et celle de l'erreur latérale dans le repère de la caméra correspondant à la première détection du panneau. Nous désignons sous le terme d'erreur latérale la composante de l'erreur située dans le plan image de la caméra pour la première détection.

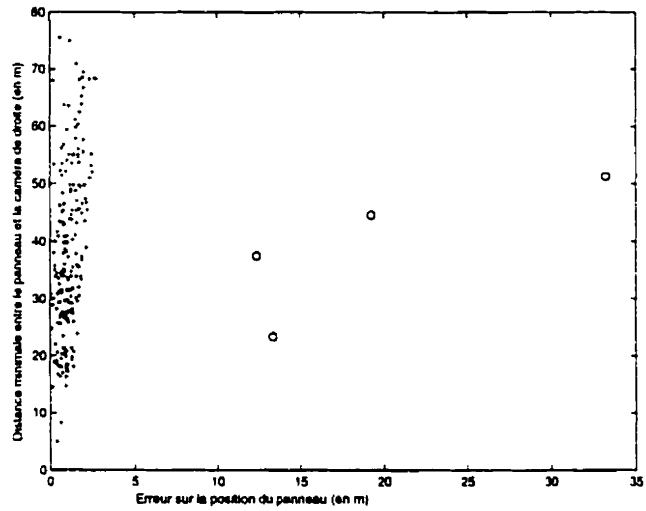
Lorsque l'on introduit un biais parasite variable sur la position des caméra, on obtient des résultats fort médiocres que nous ne détaillerons pas. Nous allons maintenant analyser ces résultats à la section suivante et évaluer nos algorithmes.

5.9 Discussion

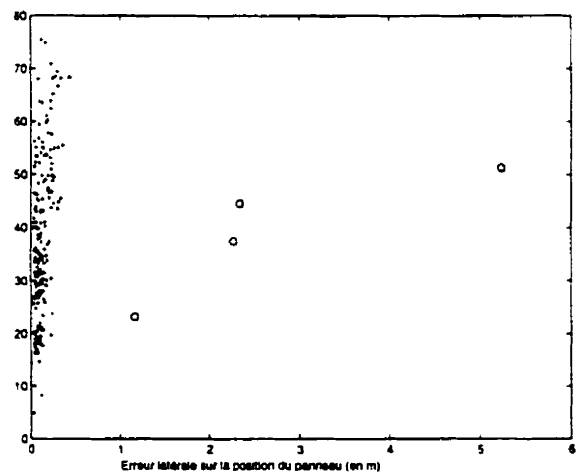
Nous allons d'abord analyser les résultats obtenus pour les trois expériences réalisées, puis nous porterons notre intérêt sur les limites des simulations et les améliorations possibles lors d'une intégration complète avec la détection.

5.9.1 Efficacité de la prédiction

Les courbes des figures 5.5 à 5.8 montrent que l'évolution de la moyenne du facteur de réduction de la zone de recherche est beaucoup plus irrégulière que celle de la médiane : quand un panneau est présent dans l'image, c'est le carré de la plus importante des deux erreurs de prédiction qui est prise en compte dans le facteur de réduction, ainsi une mauvaise prédiction a plus d'incidence sur la moyenne qu'une bonne prédiction. La médiane du facteur de réduction diminue (avec des irrégularités locales) au cours de la poursuite : cela indique une amélioration de la prédiction. Un biais variant



(a) Distribution de l'erreur sur la position.



(b) Distribution de l'erreur latérale.

Figure 5.9 : Distribution de l'erreur sur la position et l'erreur latérale sur la position des panneaux en fonction de la distance du panneau à la position de la caméra de droite au début de la poursuite.

brutalement se traduit par des prévisions médiocres à ces instants. En revanche un biais en constante évolution semble limiter les capacités de prédiction d'autant plus que ces variations sont rapides. Le modèle d'état utilisé ne suppose pas d'erreur sur la position et il est donc normal que l'introduction d'un biais conduise à de moins bons résultats. Il faut donc considérer ces cas critiques.

Le facteur de réduction maximum caractérise le risque d'échanger le panneau pour une fausse alarme. Dans les pires situations, il peut avoisiner la valeur de 0,15, soit une recherche sur un sixième de l'image complète. Signalons que les valeurs supérieures à 0,1 se produisent pour 1% des images; en outre les courbes des figures 5.5 à 5.8 indique qu'en moyenne ce phénomène se produit durant les 15 premières images de la poursuite, c'est-à-dire quand on recherche de grands panneaux et que l'on a donc le moins de chances d'obtenir des fausses alarmes. En dépit de ces problèmes on peut estimer l'apport de l'intégration à la détection.

La réduction du temps de calcul est proportionnelle à celle de la zone de recherche. C'est le facteur moyen qu'il faut aussi considérer, car à chaque nouvelle image où un panneau va être détecté, une poursuite va être lancée, et la réduction globale va être proportionnelle à cette valeur. Cette réduction est au moins d'un facteur 100 ce qui est loin d'être négligeable. Nous disposons donc d'un outil pour effectuer un suivi des panneaux.

5.9.2 Suivi des panneaux

La caméra de droite ne manque aucune apparition d'un panneau : la zone de recherche est suffisamment grande pour que cela soit possible. Celle de gauche peut manquer jusqu'à une apparition sur cinq mais le système de suivi se révèle capable de le retrouver. Du point de vue du problème du contrôle de la visibilité dans la mesure où la caméra de droite a parfaitement réussi à suivre le panneau cela est sans conséquences. Ces pertes du panneau reste momentanées. Elles s'expliquent de la manière suivante : quand un panneau a été suivi avec la caméra de droite cette opération doit être répétée avec les images de gauche; le panneau n'est pas forcément visible dès le début et la poursuite démarre à partir d'une prédiction et donc pas forcément dans de bonnes conditions. Nous obtenons en fin de compte au moins 90% des apparitions pour essayer de localiser les panneaux.

5.9.3 Localisation des panneaux

Sans biais sur la position des caméras, l'erreur constatée sur la position est inférieure à 3 m dans 98% des cas, et en moyenne, d'après le tableau 5.4 les erreurs sont proches du mètre. La figure 5.9a montre que nous avons quatre panneaux pour lesquels la localisation est très mauvaise : ils sont repérés par des cercles. Pour les autres panneaux (repérés par des points), la figure 5.9a montre que plus la distance minimale à la caméra de droite augmente, plus la localisation peut devenir imprécise. La figure 5.9b montre que l'erreur latérale est beaucoup plus faible; l'erreur latérale

des quatre panneaux pour lesquels l'erreur de position est supérieure à 3 m, (repérés par des cercles) est inférieure à 6 m. les erreurs importantes trouvent donc leur origine dans une mauvaise estimation de la coordonnée de profondeur du panneau. L'erreur sur la position des panneaux correspond à un bruit sur la mesure plus important que l'incertitude résultant de la détection des panneaux. L'influence des incertitudes sur la position et l'orientation des caméras est donc loin d'être négligeable. Ce sont des résultats inférieurs à ceux des systèmes manuels qui permettent d'être en dessous du mètre : ces systèmes nécessitent cependant qu'un opérateur amène un récepteur GPS à la base du panneau pour effectuer la mesure de la position ce qui prend beaucoup plus de temps. Dès qu'un biais non constant est introduit, notre algorithme n'est pas capable de produire de bons résultats. Avec un biais constant, le problème est translaté et l'erreur sur la position est augmentée (vectoriellement) de ce biais. Il s'agit là d'une des limites de notre système; nous allons maintenant aborder plus particulièrement celles reliées à nos simulations.

5.9.4 Vers des expérimentations réelles

Nous avons seulement adressé le problème du suivi des panneaux dans le cadre de simulations d'un déplacement en ligne droite. La mise en œuvre dans des conditions réelles implique d'adresser un certain nombre de problèmes qui n'ont pas été abordés dans le cadre de ces simulations. Notre modèle suppose d'abord que les caméras sont de type sténopé, donc parfaites. Il convient donc d'évaluer l'influence des distorsions

induites par les caméras et, si nécessaire, de les compenser. Le problème du choix de la focale des objectifs n'a pas non plus été étudié car une route rectiligne n'induit pas de contraintes sur l'angle de champ des caméras : la distance focale influe seulement sur la taille d'un panneau dans l'image ainsi que la portion de route depuis laquelle il est visible. Nous indiquons maintenant deux contraintes sur la distance focale à prendre en considération. Premièrement, pour pouvoir détecter et suivre les panneaux dans les courbes, il faut que ceux-ci restent dans le champ de vision, ce qui implique d'utiliser des objectifs dont la distance focale est suffisamment petite. Deuxièmement, tous les autres paramètres restant inchangés, une focale plus longue permet d'obtenir plus d'images pour lesquelles les dimensions d'un panneau dans l'image restent comprises dans un intervalle donné : une longue focale permet donc soit d'obtenir plus d'images, soit d'employer moins de masques. Nous allons finalement voir les possibilités d'améliorations de la détection qu'offre son intégration au suivi.

5.9.5 Emploi d'autres techniques d'analyse d'images

Estimer seulement la position d'un panneau dans les images précédentes ne permet cependant que de réduire la zone de recherche et de diminuer le nombre de fausses alarmes. En cas d'échec de la détection dans la zone de recherche, il est possible de resegmenter cette région au moyen de techniques plus performantes que l'étiquetage mais dont l'emploi sur toute l'image représenterait une charge de calcul trop impor-

tante. Cela pourrait permettre une amélioration du taux de détection dans la mesure où il s'agit de son point faible. Le dernier chapitre va nous permettre maintenant de faire le bilan de cette étude et d'envisager les autres développements futurs possibles.

Chapitre 6

Conclusion

Nous avons traité le problème de l'application de la reconnaissance par analyse d'images des panneaux routiers à l'inventaire de la signalisation routière, afin d'automatiser celui-ci. Le choix a été fait d'aborder ce problème en développant un algorithme de détection des panneaux, en lui rajoutant les fonctionnalités nécessaires pour pouvoir établir un inventaire tout en cherchant à utiliser ces fonctionnalités pour améliorer les performances de la détection. Un module de reconnaissance finale déterminant si l'on est bien en présence d'un panneau et dans cette éventualité le type de ce panneau devrait normalement compléter l'ensemble.

Nous avons développé un algorithme de détection utilisant deux techniques relativement simples : la segmentation par étiquetage et la reconnaissance de forme par corrélation. Nous avons adjoint un traitement intermédiaire afin d'améliorer les performances de l'ensemble. Notre algorithme nécessite un apprentissage dont

les résultats dépendent des types de panneaux recherchés et du matériel acquisition d'images. Sur les images dont nous disposons, de bons résultats ont été obtenus.

Les simulations du déplacement d'un véhicule le long d'une route ont montré que le filtre de Kalman constituait un outil efficace pour le suivi des panneaux d'une image sur l'autre dans la mesure où l'on était capable de détecter correctement les panneaux. Ces résultats nous permettent d'espérer une réduction notable du temps consacré à la détection d'un panneau. La détection des panneaux uniquement lorsqu'ils sont grands dans l'image conduit à un taux de détection très proche de un et à un nombre de fausses alarmes très faibles : c'est l'étape de départ d'un système unifiant détection et suivi des panneaux. L'implantation d'un tel système devrait donc permettre de maintenir un taux de fausses alarmes très bas tout en conservant un taux de détection élevé.

Notre algorithme de localisation a posteriori des panneaux sur la base des résultats du suivi nécessite de connaître la position des caméras avec une meilleure précision que l'algorithme de suivi. Quand ce degré de précision est disponible, notre algorithme permet, dans la très grande majorité des cas, de positionner les panneaux avec une précision aussi bonne que celle obtenue avec un simple récepteur GPS différentiel placé à la base du panneau. Les systèmes stéréoscopiques similaires avec appariement manuel ainsi que les récepteurs GPS différentiels intégrant plusieurs mesures sont plus précis mais notre approche devrait conduire à des systèmes entièrement automatisés, plus rapides et moins coûteux.

Notre étude apporte plusieurs contributions originales au problème de la détection des panneaux routiers. Notre algorithme de détection, s'il suit le schéma couramment employé de la recherche de certaines couleurs puis de certaines formes, ne lie pas strictement ces deux opérations : ainsi nous ne cherchons pas à interpréter les formes des régions segmentées, mais recherchons des formes dans l'image des régions segmentées ce qui permet de traiter aisément les problèmes de fusions et de divisions de régions. Une telle approche a déjà été proposée par Estevez et Kehtarnavaz (1996), mais leur système ne permet que de détecter la présence d'un seul panneau dans l'image.

Pour chacune de ces deux étapes de la détection, nous avons développé une technique pour déterminer les valeurs des paramètres conduisant à de bonnes performances. Nous avons proposé ainsi une méthode pour juger rapidement des performances d'une segmentation par étiquetage en fonction d'une segmentation de référence établie par un opérateur et supposée idéale. Ce qui permet de comparer facilement différentes étiquetages. Pour la localisation de forme, nous avons développé une procédure simple permettant d'obtenir des seuils pour les différents masques donnant de bons résultats sur un ensemble d'images.

Notre système de détection se distingue, aussi, par le tri des régions segmentées en fonction de la régularité locale de leur frontière. Cette opération permet d'identifier rapidement les régions pouvant être des panneaux ou des éléments de panneaux sans qu'une analyse complète soit nécessaire. En présence d'un élément de panneau, une

telle analyse serait de plus difficile.

Nous proposons de plus une architecture originale pour un système d'inspection automatique de la signalisation routière. Cette architecture repose sur l'intégration d'un système de poursuite à la détection : le système de poursuite permet de relier les différentes apparitions des panneaux et de restreindre l'étendue des zones de recherche. Un tel couplage a déjà été proposé par Piccioli et al. (1994) pour améliorer les performances de la détection. La spécificité de notre approche repose sur l'exploitation de la possibilité de détecter un panneau avec un certain retard lors d'une inspection de la signalisation : nous proposons donc d'initialiser la poursuite à partir d'une des dernières apparitions puis de rechercher le panneau dans les images précédentes. Une telle initialisation, basée sur la détection d'un panneau de grandes dimensions dans l'image, est, de ce fait, moins sensible aux fausses alarmes.

Nous avons étudié trois aspects d'un système de localisation automatique des panneaux routiers. Nous n'avons cependant pas abordé tous les problèmes liés à ces trois modules et à leur intégration. Nous allons donc présenter maintenant les recherches pouvant faire suite à ce travail. Nous avons d'abord développé un algorithme de détection qui soit capable de trouver les différentes apparitions d'un panneau pour un minimum de fausses alarmes. Cet algorithme repose sur une segmentation figée qui n'est pas capable de traiter toutes les conditions d'éclairage. Nous avons résolu ce problème en combinant deux segmentations dont seul le seuil sur la saturation est différent. Il convient donc d'étudier d'autres techniques de segmentation sans oublier

que la rapidité est l'intérêt de l'étiquetage. De plus notre algorithme de détection ne permet toutefois pas de détecter les panneaux achromatiques car il se fonde sur la détection dans l'image d'une teinte; la distinction du blanc ou du noir repose sur l'intensité dans l'image et est, ainsi, difficile si l'on ne dispose d'aucune information sur l'éclairage de la scène. Nous ne pensons pas que cela soit une limitation dans la mesure où des algorithmes de détection basés uniquement sur l'intensité dans l'image (tel que celui de Piccioli et al. (1994)) ont déjà été développés. Il convient néanmoins de les intégrer en les employant dans les zones de faible saturation.

Notre algorithme de suivi a été testé à l'aide de simulations sur des routes droites ce qui ne correspond pas à la réalité. Il convient donc d'étudier ce type de mouvements. De plus nous n'avons pas considéré les problèmes des distorsions de l'image et de la calibration du système de vision. La simulation du déplacement d'un véhicule est un problème fort complexe que nous avons que partiellement adressé, il semble préférable d'étudier le suivi à partir d'images prises depuis un véhicule, pour lesquelles les informations de positions et d'orientations des caméras sont connues. Cependant, dans la mesure où le suivi s'est révélé bien robuste vis-à-vis des diverses erreurs sur la position et dans la mesure où aucune hypothèse n'est faite sur le déplacement du véhicule entre deux prises de vues, nous considérons qu'il devrait être possible d'obtenir des résultats similaires sur n'importe quelle route à condition de disposer d'un objectif ayant un angle de champ suffisamment large pour que les panneaux restent visibles.

L'algorithme de localisation s'est avéré beaucoup plus sensible aux erreurs de positionnement des caméras et dépend donc d'un système de positionnement précis. Nous n'avons cependant pas cherché à développer d'algorithme spécifique pour surmonter ce problème. Deux voies peuvent être explorées. Le développement d'un meilleur modèle d'erreur pour pouvoir adapter la méthode d'estimation d'une part et l'emploi de statistiques robustes pour limiter l'influence des erreurs d'autre part.

L'intégration de la détection et du suivi implique de gérer les poursuites en cours. Pour nos simulations nous nous sommes contentés d'associer les panneaux détectés sur la base d'un critère de distance minimale à la position prédite afin de nous concentrer sur l'étude du suivi. Un système gérant la poursuite de plusieurs panneaux doit d'abord être capable d'éviter de suivre deux fois un même panneau en partant de deux images différentes. Ensuite, pour chaque poursuite en cours, il faut essayer de trouver le panneau correspondant. Divers problèmes peuvent se poser. Le panneau peut ne pas être visible dans l'image ou bien plus d'un candidat satisfaisants peuvent être trouvés, et il faut alors éviter de poursuivre une fausse alarme. Inversement un panneau peut se trouver dans plus d'une zone de recherche, mais ne peut être attribué qu'à une seule. Différentes approches peuvent être envisagées; nous suggérons d'explorer les trois voies suivantes. Premièrement il est possible d'essayer de développer des heuristiques pour résoudre les conflits en fonctions des zones de recherches, de l'historique des poursuites et des résultats de la détection. Deuxièmement comme un même modèle de panneau ne peut se trouver plus d'une fois à un endroit, un

algorithme de reconnaissance devrait permettre un bon appariement au prix toutefois de l'étude de tous les candidats détectés. Finalement l'approche développée par (Zhang et Faugeras, 1992) pour le suivi de segments de droite pourrait être adaptée à notre problème : à chaque image, chaque poursuite se développe en arborescence afin d'envisager tous les cas possibles, puis la branche la plus satisfaisante de cet arbre est retenue.

La réalisation d'un système automatique de localisation des panneaux pour l'inspection de la signalisation routière nécessite, donc, encore des recherches dans les domaines de la détection des panneaux, de localisation d'objets à partir de plusieurs prises de vues et de la poursuite simultanée de différentes cibles. Nous avons proposé une architecture pour un tel système qui permette de l'intégrer à des systèmes déjà employés. Nous avons de plus développé les éléments fondamentaux d'un tel système. Il convient, en outre de noter que ces recherches futures doivent être conduites en parallèle afin de pouvoir exploiter au mieux les forces et les faiblesses des différents modules et finalement pour que l'ensemble puisse fonctionner en temps réel.

Annexe A

Les systèmes de positionnement par satellites (d'après (Hofmann-Wellenhof et al., 1997))

A.1 Principe

Ces systèmes reposent sur le même principe. Une constellation de satellites se trouve en orbite de façon à ce qu'à chaque instant, on puisse voir au moins quatre de ces satellites depuis n'importe quel point de la surface de la terre. Les trajectoires de ces satellites sont parfaitement connues. Chaque satellite émet en permanence la valeur de son horloge. Le décalage entre la date locale et celle émise par le satellite permet, en connaissant le temps de propagation des ondes à travers l'atmosphère, de

déterminer sa distance au satellite, dont la position à cet instant est connue. En étant visibles depuis trois satellites, on détermine trois sphères à l'intersection desquelles on se trouve, ce qui permet de connaître sa position sans ambiguïtés sur le globe terrestre. En pratique on utilise un quatrième satellite pour pouvoir estimer le biais de sa propre horloge.

A.2 Défauts

Diverses erreurs influent sur l'estimation des distances aux satellites. Pour des raisons de sécurité, les performances du système NAVSTAR sont dégradées pour tout usager autre que l'armée américaine. Ensuite des erreurs sur la vitesse de propagation, des problèmes d'interréflections (en ville notamment, entre les immeubles) viennent se rajouter. En pratique l'incertitude est de l'ordre de 100 m. Il convient de remarquer que dans le monde occidental le système NAVSTAR est actuellement le système le plus employé.

Deux techniques sont principalement employées pour améliorer les performances du système. Premièrement il est possible d'estimer l'erreur en utilisant un second récepteur placé en un point fixe de position connue, ce qui permet d'affiner la mesure dans un rayon d'environ 100 kilomètres car l'erreur induite varie lentement avec la position : c'est le principe du GPS différentiel. L'incertitude dans le plan horizontal est alors ramenée à ± 2m et celle sur l'altitude à ± 5m. Deuxièmement, cette incertitude peut être réduite en mesurant la différence de phase entre le signaux reçus par

les deux récepteurs. L'incertitude alors obtenue est de l'ordre de quelques dizaines de centimètres.

Annexe B

La stéréoscopie (d'après (Horn, 1986))

Le principe repose sur celui de la triangulation. Si un point est visible d'une caméra, il est possible de déterminer une droite passant par ce point. Avec deux caméras l'on obtient deux droites et donc la position de ce point. La figure B.1 montre le cas simplifié où les deux caméras sont identiques, de type sténopé, avec les axes optiques parallèles et la ligne de base (joignant les centres optiques) orthogonales à ces axes; pour des raisons de clarté le problème n'est montré que dans le plan épipolaire, défini par le point d'intérêt M et les deux centres optiques. Cette configuration géométrique bien particulière est celle employé dans notre système. Signalons rapidement qu'en pratique la réalisation d'un tel système est imparfaite et qu'une calibration est donc nécessaire : elle consiste à déterminer la géométrie réelle du système, à partir de la

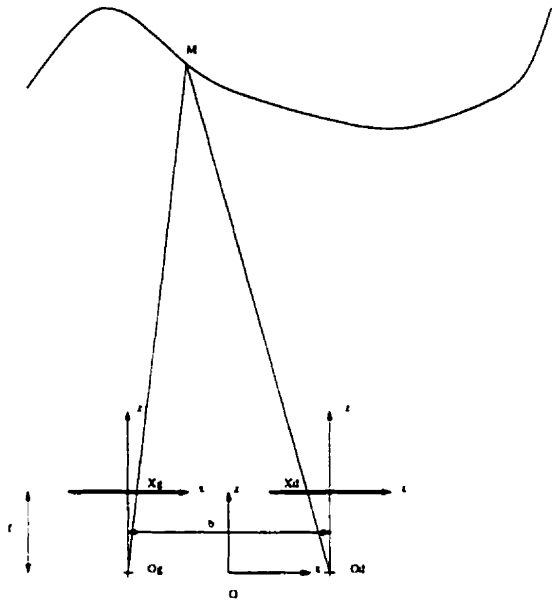


Figure B.1 : Principe de la stéréoscopie.

prise de vue d'une mire de référence connue.

Les équations sont les suivantes :

$$x_M = \frac{bX_d}{X_g - X_d} - \frac{b}{2} \tag{B.1}$$

$$z_M = \frac{bf}{X_g - X_d} \tag{B.2}$$

$$\tag{B.3}$$

Bibliographie

BAASS, K. (1993), *Précis sur la signalisation routière au Québec*, Vol. 1, Association québécoise du transport et des routes, Montréal.

BAASS, K. et DION, F. (1993), *Précis sur la signalisation routière au Québec*, Vol. 1, Association Québecoise du Transport et des Routes, Montréal, chapter 4, pp. 133–206.

BAGHDASSARIAN, C. (1995), Classification chromatique pour la reconnaissance d'objet : application aux scènes routières, Thèse de doctorat, Ecole des Mines de Paris.

BROIDA, T. et CHELLAPPA, R. (1989), Experiments and uniqueness results on object structure and kinematics from a sequence of monocular images, dans *Proceedings of the IEEE Workshop on Visual Motion*, Irvine, Californie, pp. 21–30.

BROIDA, T. et CHELLAPPA, R. (1991), Estimating the kinematics and structure of a rigid object from a sequence of monocular images, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **13**(6), 497–513.

CANNING, J., KIM, J., NETANYAHU, N. et ROSENFELD, A. (1987), Symbolic pixel labeling for curvilinear feature detection, dans *Proceedings of the Image Understanding Workshop*, Vol. 1, Los Angeles, Californie, pp. 242–256.

CHANDRASHEKHAR, S. et CHELLAPPA, R. (1991), Passive navigation in a partially known environment, dans *Proceedings of the IEEE Workshop on Visual Motion*, Princeton, New Jersey, pp. 2–7.

CHUI, C. K. et CHEN, G. (1991), *Kalman filtering : with real time applications*, 2 edn, Springer-Verlag, New-York.

CUMANI, A. (1991), Edge detection in multispectral images, *Computer Vision, Graphics and Image Processing : Graphical Model and Image Processing* **53**(1), 40–51.

DE SAINT BLANCARD, M. (1992), Road sign recognition : a study of vision-based decision making for road environment recognition, dans *Vision-based vehicle guidance*, Springer Series in Perception Engineering, Springer-Verlag, New-York.

ESTEVEZ, L. et KEHTARNAVAZ, N. (1996), A real-time histographic approach to road sign recognition, dans *Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation*, San Antonio, Texas, pp. 95–100.

FAUGERAS, O., AYACHE, N. et ZHANG, Z. (1988), A preliminary investigation

of the problem of determining ego and object motions from stereo, dans *International Conference on Pattern Recognition*, Rome, Italie, pp. 242–246.

FUNT, B. V. et FINLAYSON, G. D. (1995), Color constant color indexing, *IEEE Transaction on Pattern Analysis and Machine Intelligence* **17**(5), 522–529.

HARALICK, R. M. et SHAPIRO, L. G. (1993), *Computer and robot vision*, Addison-Wesley publishing compagny, Reading, Massachusetts.

HEALEY, G. E. (1992), Segmenting images using normalised color, *IEEE transaction on systems, man, and cybernetics* **22**(1), 64–73.

HOFMANN-WELLENHOF, B., LICHTENEGGER, H. et COLLINS, J. (1997), *Global positionning system, theory and practice*, 4 edn, Springer-Verlag, Wienne, Autriche.

HORN, B. K. P. (1986), *Robot vision*, The MIT Press, Cambridge, Massachusetts, chapter 13, pp. 299–333.

JÄHNE, B. (1997), *Pratical handbook on image processing for scientific applications*, CRC Press LLC, Boca Raton, Floride, chapter 15, pp. 480–483.

JANSSEN, R., RITTER, W., STEIN, F. et OTT, S. (1993), Hybrid approach for traffic sign recognition, dans *Proceedings of the IEEE Symposium on Intelligent Vehicles*, Tokyo, Japon, pp. 390–395.

KEHTARNAVAZ, N. et AHMAD, A. (1995), Traffic sign recognition in noisy outdoor

scenes, dans *Proceedings of the IEEE Symposium on Intelligent Vehicles*, Detroit, Michigan, pp. 460–465.

KLINKER, G. J., SHAFER, S. A. et KANADE, T. (1990), A physical approach to color image understanding, *International Journal of Computer Vision* **4**(1), 7–38.

LambdaTech International (n.d.), site web. http://www.lambdatech.com/.

LEUNG, M., LIU, Y. et HUANG, T. (1991), Estimating three-dimensional vehicle motion in an outdoor scene from monocular and stereo image sequences, dans *Proceedings of the IEEE Workshop on Visual Motion*, Princeton, New Jersey, pp. 62–68.

LI, Y. et LALONDE, M. (1995), Detection of road signs using color indexing, rapport technique, CRIM.

MATTHIES, L. et SHAFER, S. (1987), Error modelling in stereo navigation, *IEEE Journal of Robotics and Automation* **3**(3), 239–248.

PEREZ, F. et KOCH, C. (1994), Toward color image segmentation in analog VLSI : algorithm and hardware, *International Journal of Computer Vision* **12**(1), 17–42.

PICCIOLI, G., DE MICHELI, E., PARODI, P. et CAMPANI, M. (1994), Robust road sign detection and recognition from image sequences, dans *Proceedings of the IEEE Symposium on Intelligent Vehicles*, Paris, France, pp. 278–283.

PRATT, W. K. (1991), *Digital image processing*, 2 edn, John Wiley & Sons, Inc, New York.

PRESS, W. H., FLANNERY, B. P., TEUKOLSHY, S. A. et VETTERLING, W. T. (1988), *Numerical recipes in C : the art of scientific computing*, Cambridge University Press, Cambridge, Royaume-Uni.

PRIESE, L., KLIEBER, J., LAKMANN, R., REHRMANN, V. et SCHIAN, R. (1994), New results on traffic sign recognition, dans *Proceedings of the IEEE Symposium on Intelligent Vehicles*, Paris, France, pp. 249–254.

PRIESE, L., REHRMANN, V., SCHIAN, R. et LAKMANN, R. (1993), Traffic sign recognition based on color image evaluation, dans *Proceedings of the IEEE Symposium on Intelligent Vehicles*, Tokyo, Japon, pp. 95–100.

SUKKARIEH, S., NEBOT, E. M. et DURRANT-WHYTE, H. F. (1998), Achieving integrity in an INS/GPS navigation loop for autonomous land vehicle, dans *Proceedings of the IEEE International Conference on Robotics and Automation*, Leuven, Belgique, pp. 3437–3442.

SWAIN, M. J. et BALLARD, D. H. (1991), Color indexing, *International Journal of Computer Vision* **7**(1), 11–32.

TAYLOR, C., KRIEGMAN, D. et ANANDAN, P. (1991), Structure and motion in two dimensions from multiple images : a least squares approach, dans

Proceedings of the IEEE Workshop on Visual Motion, Princeton, New Jersey, pp. 242–248.

TESSIER, G.-R. (1990), *Guide de construction et d'entretien des chaussées*, Association Québécoise du transport et des routes, Montréal, Québec, chapter 17, pp. 240–270.

TOMASI, C. et KANADE, T. (1992), Shape and motion from image streams under orthography: a factorization method, *International Journal of Computer Vision* **9**(2), 137–154.

Transport Québec (1993), *La Signalisation routière au Québec repertoire des signaux*, Transport Québec, Québec.

VISAT Technologies Inc (n.d.), site web. http://www.visat.com/.

WU, J., RINK, R., CAELLI, T. et GOURISHANKER, V. (1989), Recovery of the 3-d location and motion of a rigid object through camera image (an extended kalman filter approach), *International Journal of Computer Vision* **2**(4), 373–394.

WU, T., CHELLAPPA, R. et ZHENG, Q. (1995), Experiments on estimating egomotion and structure parameters using long monocular image sequences, *International Journal of Computer Vision* **15**(1-2), 77–103.

ZHANG, Z. et FAUGERAS, O. (1992), Three-dimensional motion computation and

object segmentation in a long sequence of stereo frames. *International Journal of Computer Vision* **7**(3), 211–241.

ZHENG, Q. et CHELLAPPA, R. (1995). Automatic feature point extraction and tracking in images sequences for arbitrary camera motion. *International Journal of Computer Vision* **15**(1–2), 31–76.