N° d'ordre : 4319

Université des Sciences et Technologies de Lille

École Doctorale Sciences Pour l'Ingénieur

Thèse

pour obtenir le grade de

Docteur de l'Université des Sciences et Technologies de Lille

Automatique, Génie Informatique, Traitement du Signal et des Images

Présentée et soutenue publiquement par

Cindy Cappelle

le 5 Décembre 2008

Localisation de véhicules et détection d'obstacles

Apport d'un modèle virtuel 3D urbain

Membres du jury :

Roland Chapuis	Professeur, LASMEA, Polytech'Clermont-Ferrand	Président
Philippe Bonnifait	Professeur, HeuDiaSyC, UTC, Compiègne	Rapporteur
Dominique Meizel	Professeur, XLIM, ENSIL, Limoges	Rapporteur
David Bétaille	Chargé de Recherche, LCPC, Nantes	Examinateur
François Charpillet	Directeur de Recherche, INRIA-LORIA, Nancy	Examinateur
Juliette Marais	Chargée de Recherche, INRETS-LEOST, Lille	Examinatrice
Denis Pomorski	Professeur, LAGIS, USTL, Lille	Directeur
Maan El Badaoui El Najjar	Maître de Conférences, LAGIS, USTL, Lille	Co-encadrant

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Table des matières

Table des matières i
Liste des figures vii

Introduction générale **1**
Problématique et contexte du travail 2
Organisation du manuscrit 4

I Localisation de véhicules : état de l'art **9**
I.1 Introduction 12
I.2 Localisation relative 14
I.2.1 Odométrie 14
I.2.2 Navigation inertielle 15
I.2.3 Autocalibration *via* l'image 17
I.2.4 Recalages successifs de profils télémétriques 2D 18
I.3 Localisation absolue 18
I.3.1 Localisation par rapport à des amers 18
I.3.2 Localisation sur carte de l'environnement 23
I.3.3 Localisation sur carte routière 24
I.3.4 Positionnement par satellites 25
I.4 Localisation hybride 37
I.5 Conclusion : positionnement de notre approche 39

II Modèle virtuel 3D urbain et SIG-3D **43**
II.1 Modèle virtuel 3D urbain 46
II.1.1 Introduction 46
II.1.2 Le modèle Tecnomade de Nancy 50
II.2 Système d'information géographique 3D 55
II.2.1 Introduction 55
II.2.2 Le SIG-3D développé 56
II.3 Conclusion 65

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III Méthode de localisation proposée **67**
III.1 Observation cartographique 3D . . . 72
III.1.1 Introduction . . . 72
III.1.2 Modélisation et calibration de la caméra réelle . . . 75
III.1.3 Paramétrage de la caméra virtuelle . . . 85
III.1.4 Détection des primitives de l'image . . . 87
III.1.5 Mise en correspondance 2D/2D . . . 91
III.1.6 Calcul des points 3D correspondants aux points de Harris de l'image virtuelle . . . 94
III.1.7 Calcul de la pose de la caméra réelle à partir des appariements 2D/3D de points . . . 97
III.1.8 Incertitude associée à l'observation cartographique 3D . . . 101
III.2 Fusion de données multi-capteurs . . . 102
III.2.1 Introduction . . . 103
III.2.2 Méthode de fusion multi-capteurs choisie pour la localisation de véhicules . . . 111
III.3 Conclusion . . . 116

IV Mise en œuvre expérimentale et résultats **119**
IV.1 Moyens expérimentaux . . . 122
IV.2 Localisation par fusion multi-capteurs . . . 126
IV.2.1 Résultat sur une séquence . . . 127
IV.2.2 Test de cohérence . . . 129
IV.3 Observation cartographique 3D . . . 131
IV.3.1 Transformation rigide entre le GPS et la caméra . . . 131
IV.3.2 Résultat sur une séquence . . . 133
IV.4 Conclusion . . . 149

V Détection d'obstacles **151**
V.1 Méthode développée . . . 154
V.1.1 Présentation . . . 154
V.1.2 Détection des obstacles . . . 157
V.1.3 Extraction de la distance entre le véhicule et les obstacles . . . 159
V.1.4 Calcul de la géo-position des obstacles détectés . . . 160
V.2 Résultats expérimentaux . . . 161
V.2.1 Moyens expérimentaux . . . 161
V.2.2 Interface graphique . . . 163
V.2.3 Validation des résultats à l'aide d'un télémètre laser à balayage 166

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

V.3 Conclusion 168

Conclusion générale et perspectives **171**

Conclusions 172

Perspectives 173

Bibliographie **185**

Résumé **187**

Abstract **189**

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Table des figures

I.1 Définition du roulis, du tangage et du lacet . . . 12
I.2 Paramètres de localisation d'un véhicule . . . 13
I.3 Plateforme inertielle à cardan . . . 16
I.4 Principe de calcul de la pose avec une structure arrimée . . . 17
I.5 Principe de la localisation sur amers . . . 19
I.6 Principe de la triangulation . . . 21
I.7 Principe de la trilatération . . . 22
I.8 Constellation des satellites du système GPS . . . 26
I.9 Stations de contrôle au sol . . . 26
I.10 Positionnement absolu par satellites . . . 27
I.11 Positionnement tridimensionnel à partir de trois satellites . . . 31
I.12 Principe du DGPS . . . 34

II.1 Génération de modèles virtuels 3D urbains . . . 48
II.2 Quelques Vues du modèle 3D de la place Stanislas à Nancy . . . 51
II.3 Quelques textures du modèle 3D de la place Stanislas à Nancy . . . 52
II.4 Extrait du fichier *Geometrie.xml* du modèle 3D de la place Stanislas . 53
II.5 Extrait du fichier *SceneGraph.xml* du modèle 3D de la place Stanislas 54
II.6 Evolution des systèmes d'information géographique . . . 55
II.7 Système d'information géographique 3D . . . 57
II.8 FOV de la caméra virtuelle . . . 58
II.9 Repère attaché au modèle 3D . . . 59
II.10 Image virtuelle . . . 59
II.11 Fichier de profondeur . . . 60
II.12 Image de profondeur . . . 61
II.13 Fichier XML des segments 3D . . . 62
II.14 Images des segments 3D visibles . . . 63

III.1 Principe général de la méthode de localisation . . . 71

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.2 Synoptique de la construction de l'observation cartographique 3D . . 73
III.3 Détection et mise en correspondance des points de Harris puis calcul de la pose 74
III.4 Construction de l'observation cartographique 3D 75
III.5 Modèle sténopé d'une caméra 77
III.6 Mire utilisée pour la calibration de la caméra 84
III.7 Illustration de l'influence du paramètre FOV de la caméra virtuelle . 86
III.8 Calcul du paramètre intrinsèque FOV de la caméra virtuelle 86
III.9 Un point de Harris détecté dans l'image réelle (Gauche) et la zone de recherche de son correspondant dans l'image virtuelle (Droite) 92
III.10 Les coordonnées de P dans $\mathcal{R}_M$ se déduisent de son correspondant p, de la pose de la caméra C dans $\mathcal{R}_M$ ainsi que de la distance d lue dans le Z-Buffer 95
III.11 Appariements 2D/3D 97
III.12 Calcul de la pose de la caméra réelle à partir des appariements 2D/3D 98
III.13 Modélisation 2D du train arrière du CyCab 112
III.14 Déplacement circulaire entre deux instants d'échantillonnage 113

IV.1 CyCab 122
IV.2 Architecture distribuée du CyCab 123
IV.3 GPS Thales Sagitta02 124
IV.4 Caméra AVT MARLIN F-046C 125
IV.5 Interface du SIG-3D 125
IV.6 Trajectoire de référence du parcours (données GPS RTK) 126
IV.7 Vitesse du CyCab pendant le test 127
IV.8 Pose estimée par le filtre de Kalman sans parfum 128
IV.9 Erreur de la pose estimée 129
IV.10 Pose estimée par le filtre de Kalman sans parfum avec test de cohérence 130
IV.11 Test de cohérence des mesures GPS 131
IV.12 Positionnement des capteurs sur le CyCab 132
IV.13 Les différents repères des capteurs 133
IV.14 Observation cartographique 3D - Résultat 1 135
IV.15 Observation cartographique 3D - Résultat 2 136
IV.16 Observation cartographique 3D - Résultat 3 137
IV.17 Observation cartographique 3D - Résultat 4 138
IV.18 Observation cartographique 3D - Résultat 5 139
IV.19 Observation cartographique 3D - Résultat 6 140
IV.20 Observation cartographique 3D - Résultat 7 141

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

IV.21Observation cartographique 3D - Résultat 8 142
IV.22Observation cartographique 3D - Résultat 9 143
IV.23Observation cartographique 3D - Résultat 10 144
IV.24Observation cartographique 3D - Résultat 11 145
IV.25Problème d'appariements 2D/2D si la pose *a priori* est trop éloignée 146
IV.26Estimation de la pose donnée par l'observation cartographique 3D . . 147
IV.27Erreur sur l'axe X de l'observation cartographique 3D par rapport à un GPS RTK centimétrique 148
IV.28Erreur sur l'axe Y de l'observation cartographique 3D par rapport à un GPS RTK centimétrique 148

V.1 Principe de la méthode de détection et localisation d'obstacles 155
V.2 Calcul du paramètre intrinsèque FOV de la caméra virtuelle 156
V.3 Image réelle (haut) et image virtuelle (bas) à un instant k 158
V.4 Image résultat pour la détection des obstacles de la scène 159
V.5 Calcul de la géo-position de l'obstacle détecté 161
V.6 Télémètre laser utilisé : LMS SICK 291 162
V.7 Fixation des capteurs sur le CyCab 162
V.8 Interface graphique du détecteur d'obstacles 163
V.9 Résultats du détecteur d'obstacle 165
V.10 Distance entre le véhicule et l'obstacle donnée par notre méthode et par le télémètre laser 167
V.11 Ecart entre les distances données par notre méthode et par le télémètre laser 167

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Introduction générale

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Problématique et contexte du travail

Le travail présenté dans cette thèse s'inscrit dans le contexte des véhicules intelligents pour l'assistance à la conduite et/ou la route automatisée. Plus particulièrement, on s'intéresse à la localisation de véhicules routiers, ainsi qu'à la détection et la localisation d'obstacles. De nombreux programmes de recherche sont dédiés aux véhicules intelligents compte tenu du poids économique des transports ainsi que des préoccupations de sécurité et d'économie d'énergie.

Les applications concernant les systèmes avancés d'aide à la conduite (*Advanced Driver Assistance Systems*, ADAS) relèvent de la conduite elle-même ou de la perception de l'environnement routier par le conducteur. Les thématiques scientifiques concernées sont la localisation dynamique, précise et sûre, la fusion des données et l'observation des états, le contrôle et le diagnostic de la dynamique du véhicule, les architectures informatiques embarquées, les systèmes de communication et d'aide à la décision. Parmi les systèmes ADAS, citons par exemple, l'alerte de collision avant (*Forward Collision Warning*), le régulateur de vitesse (*Adaptative Cruise Control*, ACC), le contrôle adaptatif des phares (*Adaptative Light Control*, ALC), la fonction Stop&Go, le système anti angle mort (*Blind Spot Information Warning*), l'avertissement de sortie de route (*Lane Departure Warning*), l'assistance au suivi de voie (*Lane Keeping Assistance*) ...

Par ailleurs, l'automatisation complète ou partielle pour certaines fonctions se retrouve également dans de nombreux projets de recherche. La conduite en convoi (« platooning ») est par exemple une piste envisagée pour l'automatisation complète sur voies dédiées. Un nouveau mode de transport est également à l'étude : les Véhicules Individuels Publics (VIP). Les VIPs sont des véhicules urbains, autonomes, électriques et publics. Ce sont des alternatives aux véhicules personnels envisagées pour le futur. En effet, plusieurs villes envisagent d'interdire les véhicules personnels en centre-ville afin de faire face aux problèmes d'embouteillage, de pollutions atmosphérique et sonore. La Commission Européenne s'est engagée à améliorer les modes de transports dans le *Le livre blanc sur la politique européenne des transports à l'horizon 2010 : l'heure des choix*. Ce « livre blanc » préconise le développement massif des « systèmes de transports intelligents » . L'idée de ce système de transport alternatif au véhicule personnel est que les grandes agglomérations s'équipent de flottes de véhicules automatiques en libre service. Ces véhicules seront mis à disposition des usagers qui passeront commande de véhicule pour des trajets en fonction de leurs besoins. Ces véhicules exploiteront les nouvelles possibilités fournies par les

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sciences et Technologies de l'Information et de la Communication (STIC) et seront dotés de plus ou moins grande autonomie. Un exemple de VIP est le *Cybercar*. La société RoboSoft commercialise un tel véhicule électrique intelligent et modulaire : le *CyCab*.

Dans cette thèse, on se focalise sur deux tâches indispensables au développement des véhicules autonomes et des systèmes ADAS : la localisation du véhicule et la détection et la localisation des obstacles. Le GPS (*Global Positioning System*) est souvent considéré comme le capteur de base pour connaître la position d'un véhicule. Cependant, dans le cadre spécifique des milieux urbains, les informations délivrées par un récepteur GPS ne sont pas très précises et peuvent même être indisponibles. En effet, la localisation d'un récepteur exige la réception d'émissions d'au moins quatre satellites. Or, le nombre de satellites visibles peut varier de huit en pleine mer à aucun dans les configurations de type vallées étroites. Typiquement, les signaux en provenance des satellites peuvent être masqués par les immeubles, notamment en milieu urbain, et ainsi limiter la précision de la localisation à plusieurs dizaines de mètres. Le service GPS est donc intermittent et non fiable. Pour cette raison, il est utile de considérer l'utilisation d'autres capteurs.

Dès lors, la tâche de localisation est, dans la plupart des cas, réalisée par un couplage d'un système de géo-référencement satellitaire avec des capteurs de localisation à l'estime tels que des odomètres, des centrales inertielles, ... Ces capteurs proprioceptifs sont précis à court terme mais dérivent au cours du temps. Or, comme mentionné précédemment, en milieu urbain, les masquages des signaux des satellites peuvent parfois être long, en raison de la mauvaise visibilité satellitaire et des multi-trajets. Ainsi, sans recalage précis, la navigation à l'estime seule diverge excessivement. Un GPS hybridé avec un capteur de localisation à l'estime ne peut donc pas être utilisé pour un positionnement précis à long terme.

Une solution est alors d'utiliser d'autres capteurs extéroceptifs tels que des télémètres lasers à balayage ou des caméras. En effet, les progrès réalisés ces vingt dernières années dans le domaine de la vision par ordinateur font que la robotique mobile emploie désormais largement ce type de capteur. Il est désormais envisageable d'utiliser des informations visuelles pour réaliser une tâche de localisation en temps réel. Dans ce cadre, la localisation se fait généralement par la recherche d'amers qui structurent l'environnement. Ces amers sont ensuite mis en correspondance avec une carte de ces amers (ou un modèle de l'environnement) connue au préalable ou construite simultanément. Parallèlement au développement de méthodes de localisation basée sur la vision, les progrès dans le domaine de la réalité virtuelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

et des bases de données ont permis l'essor de modèles virtuels 3D de l'environnement. Citons, par exemple le succès récent de plateformes telles que Google Earth (http://maps.google.com), Virtual Earth 3D (http://maps.live.com) ou V3D (http://v3d.pagesjaunes.fr). Même si dans le cas de ces trois plateformes libres, la représentation en 3D des bâtiments est simple et partielle et les images plaquées sur les façades manquent souvent de réalisme ou de précision.

Dans cette thèse, nous nous proposons d'étudier l'apport de modèles 3D virtuels urbains précis pour la localisation des véhicules ainsi que pour la détection et la localisation des obstacles. Le principe de la méthode de localisation développée utilisant le modèle virtuel urbain 3D consiste à recaler le modèle 3D avec une vue réelle de la scène acquise par une caméra embarquée. L'objectif est donc, lors de longues absences des informations GPS, de maintenir une estimation de la pose avec une précision équivalente à celle d'un récepteur DGPS en mode EGNOS, soit quelques mètres.

Par ailleurs, dès lors que le véhicule est localisé dans le modèle 3D, les obstacles appartenant à l'infrastructure tels que les bâtiments sont connus et localisés. Pour détecter les obstacles n'appartenant pas à l'infrastructure (véhicules, piétons ...), on compare l'image réelle acquise par la caméra embarquée et l'image virtuelle extraite du modèle 3D en considérant que ce type d'obstacles est présent dans l'image réelle mais absent de l'image virtuelle.

Les méthodes développées ont été testées sur des données acquises par un véhicule CyCab équipé d'un récepteur GPS RTK, d'une caméra, d'odomètres et d'un gyromètre. Par ailleurs, nous avons à notre disposition un modèle virtuel 3D urbain de la société Tecnomade, modélisant le centre-ville de Nancy.

Les travaux réalisés au cours de cette thèse ont donné lieu à plusieurs publications [25, 19, 24, 23, 22, 18, 20, 21, 26].

Organisation du manuscrit

Ce manuscrit est organisé de la façon suivante.

Le **chapitre 1** dresse un état de l'art de capteurs et méthodes permettant de traiter le problème de la localisation de véhicules routiers. Les techniques de localisation sont généralement réparties en deux classes : la localisation relative et la localisation

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

absolue. La complémentarité de ces deux approches donne également lieu à une fusion de celles-ci, connue sous le terme de localisation hybride. La localisation hybride permet d'améliorer la robustesse et la précision de la localisation.

La localisation relative, également appelée localisation à l'estime, se rapporte à l'utilisation de capteurs proprioceptifs, qui prennent des mesures sur les parties mobiles du robot et donnent des informations sur le mouvement du véhicule (ex : odomètres, capteurs inertiels ...). La localisation absolue utilise des capteurs extéroceptifs qui permettent de percevoir le monde qui entoure le véhicule (ex : caméras, télémètres lasers, radars, GPS, ...).

Le **chapitre 2** présente la source d'information *a priori* utilisée dans le cadre de cette thèse, à savoir un modèle virtuel 3D urbain. Un modèle virtuel 3D urbain est une représentation en trois dimensions de l'environnement d'évolution du véhicule. Cette représentation est de plus photo-réaliste, texturée.

L'outil informatique, qualifié de système d'information géographique 3D (SIG-3D), permettant de manipuler la base de données du modèle 3D est également détaillé. Les entrées du SIG-3D (la base de données du modèle 3D et les paramètres intrinsèques et extrinsèques de la caméra virtuelle) ainsi que les sorties du SIG-3D (l'image virtuelle, le fichier profondeur, l'image de profondeur, le fichier des segments 3D, l'image des segments visibles) sont décrites. Par ailleurs, le géo-référencement du modèle 3D, qui permet de projeter dans le modèle 3D une position donnée dans le référentiel du GPS, est également expliqué.

Le **chapitre 3** détaille notre méthode d'égo-localisation [1], c'est-à-dire :

- d'une part la construction de l'observation de la pose du véhicule à partir du modèle virtuel 3D urbain et d'une caméra, appelée observation cartographique 3D,
- d'autre part l'algorithme de fusion multi-capteurs basé sur un filtre de Kalman sans parfum.

Le principe de l'observation cartographique 3D est de recaler le modèle virtuel 3D sur l'image réelle courante acquise par une caméra embarquée. Pour cela, on détecte les points de Harris de l'image réelle et de l'image virtuelle. L'image virtuelle est une projection 2D du modèle virtuel 3D. Les points de Harris détectés dans les deux images sont ensuite mis en correspondance. Les faux appariements sont éliminés en

1. Le problème d'égo-localisation consiste à localiser un véhicule en se mettant à la place de celui-ci, par opposition à la localisation externe qui consiste à localiser d'autres véhicules environnants.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

plongeant le calcul de l'homographie 2D des deux images dans un estimateur robuste RANSAC. Grâce à l'information de profondeur du modèle 3D, les correspondants 3D des pixels de l'image virtuelle peuvent être calculés. On a ainsi des appariements 2D/3D à partir desquels on peut calculer la pose de la caméra réelle et par suite celle du véhicule.

Concernant la partie fusion de données, dans un premier temps les équations du filtre de Kalman, du filtre de Kalman étendu et du filtre de Kalman sans parfum sont rappelées dans le cadre général. Puis, l'application du filtre de Kalman sans parfum pour traiter nos données sera présentée.

Le **chapitre 4** est dédié à la présentation des expérimentations et des résultats obtenus sur des données réelles enregistrées avec le véhicule instrumenté CyCab, développé par la société Robosoft.

Le **chapitre 5** illustre l'utilisation du modèle virtuel 3D urbain avec une caméra pour une autre application relative au domaine des véhicules intelligents : la détection et la localisation d'obstacles. Les obstacles sont classés en deux catégories : les obstacles appartenant à l'infrastructures (bâtiments, trottoirs, arbres ...) et les obstacles n'appartenant pas à l'infrastructure (piétons, véhicules, animaux ...). Les obstacles appartenant à l'infrastructure appartiennent également au modèle 3D. Une fois le véhicule localisé dans le modèle 3D, les obstacles de ce type sont donc détectés et localisés simplement. *A contrario*, les obstacles n'appartenant pas à l'infrastructure n'appartiennent pas non plus au modèle 3D. Par suite, ils n'apparaissent pas dans l'image virtuelle ; par contre ils apparaissent dans l'image réelle acquise par la caméra embarquée. Cette propriété est utilisée pour les détecter. L'information de profondeur disponible grâce au modèle 3D est ensuite utilisée pour les localiser.

La méthode développée est testée sur des données réelles. Les distances calculées entre le véhicule et les obstacles détectés sont comparées avec les distances données par un télémètre laser afin de valider notre approche.

Enfin, ce mémoire se termine par une **conclusion générale** et des **perspectives** des travaux présentés.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Chapitre I

Localisation de véhicules : état de l'art

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Résumé du chapitre

La localisation est un des problèmes fondamentaux de la robotique mobile et de la navigation de véhicules routiers. Nous présentons dans ce chapitre un ensemble de capteurs et de méthodes de traitement des mesures permettant de le résoudre.

La première partie du chapitre traite des capteurs proprioceptifs et des techniques de localisation relative, parmi lesquelles l'odométrie, la navigation inertielle ou encore le recalage relatif d'images (odométrie visuelle) ou de profils télémétriques 2D.

La deuxième partie concerne les capteurs extéroceptifs et la localisation absolue. Quelques méthodes de localisation absolue sont présentées : les méthodes basées sur des balises artificielles actives ou passives, les méthodes basées sur des amers naturels, les méthodes basées sur l'appariement de cartes 2D de l'environnement, la localisation sur carte routière numérique 2D, les systèmes de positionnement par satellites ...

La troisième partie présente la fusion multicapteurs, qui permet d'exploiter les avantages des deux méthodes de localisation, relative et absolue. La complémentarité des sources d'informations permet de recaler l'estime par des mesures absolues. La redondance des données augmente la fiabilité et la robustesse du système de localisation.

En conclusion, nous positionnerons nos travaux vis-à-vis de cet état de l'art.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sommaire

I.1 Introduction 12
I.2 Localisation relative 14
I.2.1 Odométrie 14
I.2.2 Navigation inertielle 15
I.2.3 Autocalibration *via* l'image 17
I.2.4 Recalages successifs de profils télémétriques 2D 18
I.3 Localisation absolue 18
I.3.1 Localisation par rapport à des amers 18
I.3.1.1 Les différentes classes d'amers 19
I.3.1.1.a Les amers naturels 19
I.3.1.1.b Les amers artificiels 20
Les balises passives 20
Les balises actives 20
I.3.1.2 Méthodes de calcul de la pose 20
Triangulation 21
Trilatération 22
Minimisation d'un critère 23
I.3.2 Localisation sur carte de l'environnement 23
I.3.3 Localisation sur carte routière 24
I.3.4 Positionnement par satellites 25
I.3.4.1 Constitution du système GPS 25
I.3.4.2 Les observations GPS 27
I.3.4.3 Calcul de la position 30
I.3.4.4 Les sources d'erreurs 32
I.3.4.5 GPS différentiel 33
I.3.4.6 Les autres systèmes de positionnement par satellites 35
I.4 Localisation hybride 37
I.5 Conclusion : positionnement de notre approche 39

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.1 Introduction

Le problème de la localisation de véhicules routiers consiste à répondre à la question « *Où suis-je ?* » du point de vue du véhicule. Le terme « localisation » signifie donc la connaissance de la pose du véhicule dans un référentiel connu, c'est-à-dire sa position et son orientation par rapport à son environnement.

Soient $R_W(O_W, X_W, Y_W, Z_W)$ un repère attaché à l'environnement et $R_V(O_V, X_V, Y_V, Z_V)$ le repère attaché au véhicule (cf Figure I.2). La position et l'orientation du véhicule dans le repère R_W sont respectivement déterminées par la connaissance des six paramètres suivants :

- les coordonnées cartésiennes (x, y, z) de l'origine O_V du repère R_V dans le repère R_W,
- les angles de roulis ϕ (rotation autour de l'axe longitudinal du véhicule X_V), de tangage ψ (rotation autour de l'axe transversal du véhicule Y_V) et de lacet θ (rotation autour de l'axe vertical Z_V) qui définissent l'orientation du repère R_V dans le repère R_W (cf Figure I.1).

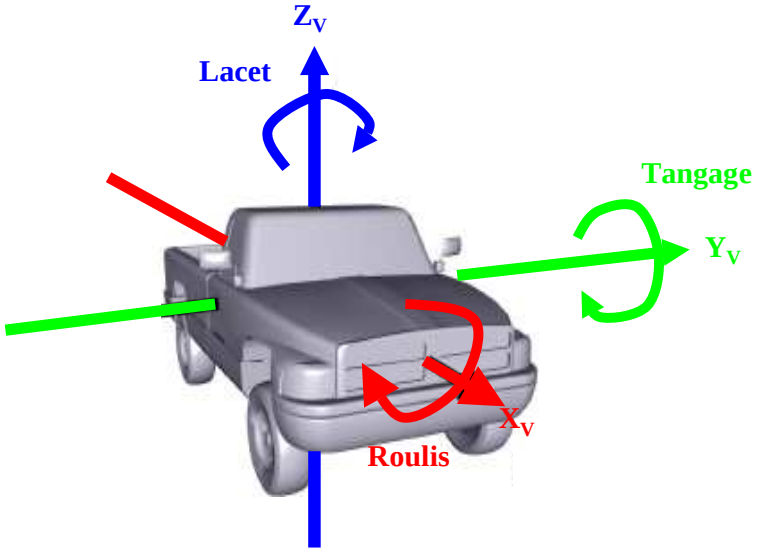


FIGURE I.1 – Définition du roulis, du tangage et du lacet

La position peut également être donnée par un ensemble de coordonnées géographiques. D'autres représentations de l'attitude du véhicule sont les quaternions ou les cosinus directeurs.

Lorsque le véhicule se déplace dans le plan (X_W, Y_W), c'est-à-dire que le roulis et le tangage sont nuls et la hauteur z est constante, la configuration du véhicule dans l'espace 2D est alors réduite au triplet (x, y, θ) avec θ le cap.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

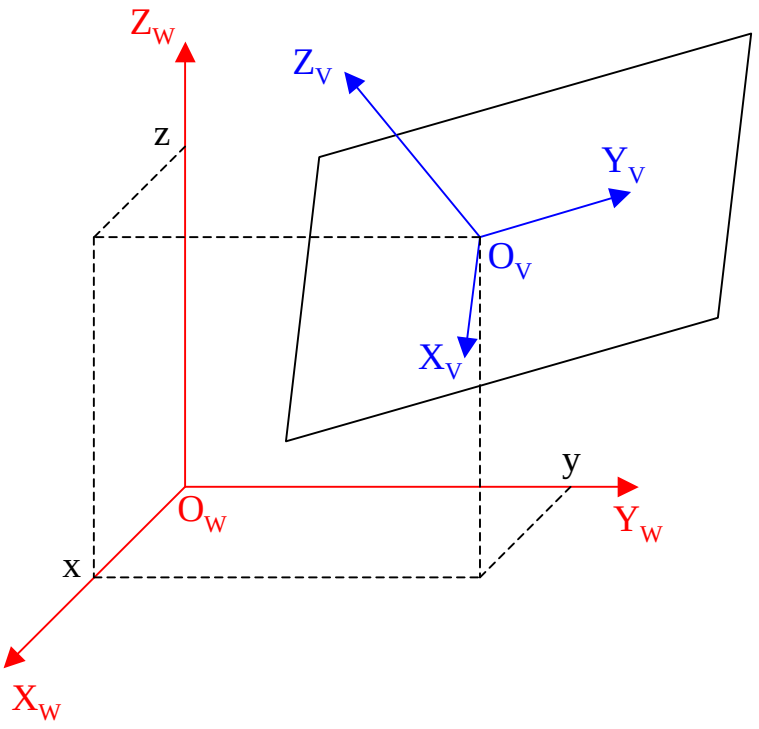


FIGURE I.2 – Paramètres de localisation d'un véhicule

La localisation est un problème important. C'est en effet souvent un élément clé dans de nombreux systèmes de navigation autonomes [112, 12, 111]. Selon certains auteurs, le problème de la localisation d'un véhicule est le problème le plus fondamental pour donner aux véhicules de vraies capacités d'autonomie [31].

Selon le type de capteurs utilisés et les méthodes mises en œuvres, on qualifie la méthode de positionnement de localisation relative (cf § I.2), localisation absolue (cf § I.3) ou localisation hybride (cf § I.4). La localisation relative se rapporte à l'utilisation de capteurs proprioceptifs, alors que la localisation absolue se rapporte à l'utilisation de capteurs extéroceptifs. Dans le cas de la localisation relative le repère de référence R_W correspond au repère du véhicule dans sa situation initiale, alors que dans le cas de la localisation absolue le repère de référence est lié à l'environnement. La localisation hybride correspond à l'utilisation conjointe des deux types de capteurs, extéroceptifs et proprioceptifs.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.2 Localisation relative

La localisation relative permet au véhicule de naviguer en utilisant les mesures de ses mouvements propres, délivrées par des capteurs dits proprioceptifs. La localisation relative, également appelée localisation à l'estime (terme anglais *: dead reckoning*), consiste alors à évaluer la position, l'orientation, et éventuellement la vitesse du véhicule par intégration (sommation) des informations fournies par les capteurs proprioceptifs depuis un point de départ. Les capteurs proprioceptifs mesurent en effet le déplacement ou le changement de cap ou d'attitude du véhicule entre deux instants mais ne donnent aucune information sur sa localisation absolue. L'intégration de leurs mesures permet alors d'estimer la situation courante du véhicule relativement à sa situation initiale. Nous verrons que la précision de la localisation relative se dégrade avec la distance parcourue notamment à cause de l'intégration. C'est pourquoi, un système permettant de recaler périodiquement le positionnement absolu du véhicule est souvent adjoint (cf § I.4).

En robotique mobile, les méthodes de localisation relative couramment utilisées sont basées sur l'odométrie (cf. § I.2.1), sur des mesures inertielles (cf. § I.2.2) ou encore sur le recalage successif d'images (cf. § I.2.3) ou de profils télémétriques (cf. § I.2.4) à partir de la position précédente.

I.2.1 Odométrie

Le principe de l'odométrie consiste à utiliser les codeurs incrémentaux montés sur les roues qui mesurent le nombre tours de roue effectués pendant une durée donnée. Cette information est utilisée pour en déduire une mesure de la vitesse de rotation des roues, et donc $(\dot{x}, \dot{y}, \dot{\theta})$. L'intégration de ces données incrémentales au cours du temps permet d'obtenir la pose (x, y, θ) du véhicule.

L'odométrie est largement utilisée parce qu'elle présente de nombreux avantages [13] :

- *Bonne précision à court terme*
- *Fréquence d'acquisition élevée*
- *Dispositif peu coûteux* : d'autant plus qu'aujourd'hui, elle est intégrée aux véhicules à travers les capteurs ABS
- *Disponibilité* : l'odométrie est un système de localisation disponible en tout temps, les odomètres sont autonomes et fiables

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- *Rapidité* : l'odométrie est basée sur des équations simples et les calculs sont réalisables en temps réel

Cependant, l'odométrie présente également des inconvénients :

- *Mauvaise précision à long terme* : à cause des dérives, du patinage des roues, des glissements, les mesures d'angle et de distance sont erronées. Par intégration, ces erreurs s'accumulent avec le temps. En particulier, les erreurs d'orientation causent d'importantes erreurs de positionnement.
- *Sensibilité au terrain* : si la surface sur laquelle le véhicule évolue n'est pas lisse, les erreurs de position peuvent être considérables, puisque l'odométrie ne détecte pas les irrégularités du terrain.
- *Localisation relative* : l'odométrie ne fournit qu'une localisation relative, elle a donc besoin d'être initialisée par un capteur externe (par exemple un GPS) pour obtenir un positionnement absolu.

Les sources d'erreurs de l'odométrie sont divisées en deux catégories : les erreurs systématiques d'une part et les erreurs non-systématiques d'autre part. Parmi les erreurs systématiques, on peut citer les diamètres inégaux des roues ; la moyenne des diamètres réels des roues qui diffère du diamètre nominal ; la distance inter-essieux du véhicule réel qui diffère de la distance nominale ; le non-alignement des roues ; la résolution finie des codeurs, et de la fréquence d'échantillonnage. Les erreurs non-systématiques sont le patinage, la navigation sur terrains accidentés, ...

I.2.2 Navigation inertielle

La navigation inertielle [4, 76, 124] utilise des accéléromètres et des gyromètres pour mesurer l'accélération et la vitesse angulaire du véhicule. Pour avoir une position en trois dimensions, le capteur inertiel doit comporter trois accéléromètres et trois gyromètres. Les accéléromètres mesurent les accélérations linéaires selon chaque axe, c'est-à-dire la pesanteur, l'accélération du véhicule, l'accélération centrifuge, l'accélération de Coriolis et du bruit. Les gyroscopes mesurent les variations d'orientation du véhicule.

On distingue deux types de capteurs inertiels :

- *Unité de mesure inertielle (Inertial Measurement Unit - IMU)* : les données brutes des accéléromètres et des gyromètres sont les seules données de sortie.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- *Système de navigation inertielle (Inertial Navigation System - INS)* : une IMU à laquelle ont été ajoutés des algorithmes de navigation afin de déduire la position et l'orientation à partir des accélérations linéaires et des vitesses angulaires.

Deux structures de systèmes inertiels existent :

- *Structure à cardan (Gimbaled)* : Les accéléromètres sont montés sur une plateforme mobile pour rester alignés avec le repère de navigation, ainsi les accélérations sont mesurées dans le repère de navigation et peuvent être directement intégrées afin d'obtenir la vitesse et la position (Figure I.3).

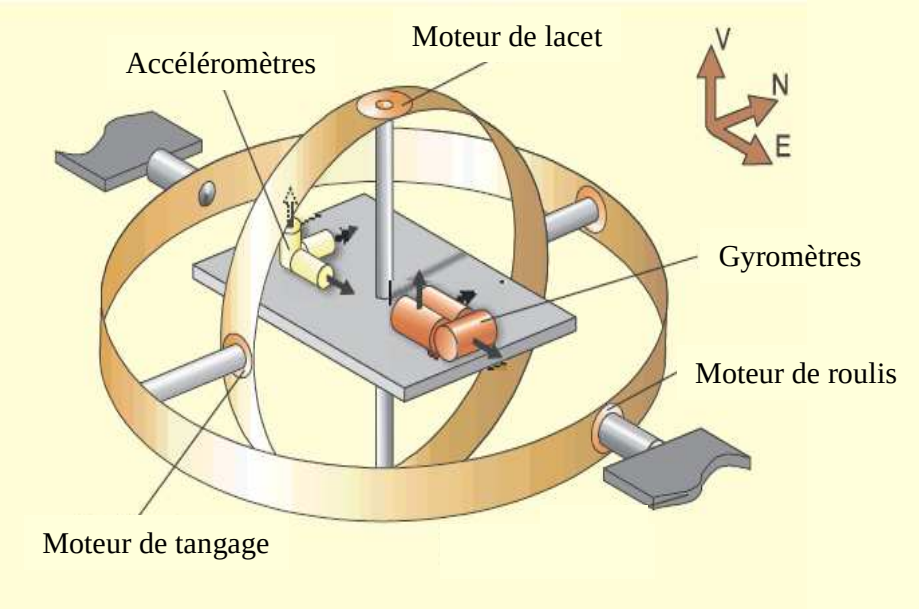


FIGURE I.3 – Plateforme inertielle à cardan

- *Structure arrimée (Strapdown)* : Les accéléromètres et les gyroscopes étant fixés rigidement, les accélérations sont mesurées dans le repère du mobile. On utilise alors les vitesses de rotation mesurées par les gyroscopes pour actualiser la transformation entre le repère du mobile et le repère de navigation. Les cardans sont en quelque sorte remplacés par un algorithme (cf Figure I.4). Ces systèmes « strapdown »sont des structures plus simples et sont plus résistants aux chocs, ils sont fortement avantageux au niveau des coûts et de la taille de l'appareil et sont particulièrement appropriés pour l'application dans le domaine automobile. En contrepartie, ils n'atteignent pas encore la précision des systèmes à cardans, et l'alignement est difficile.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

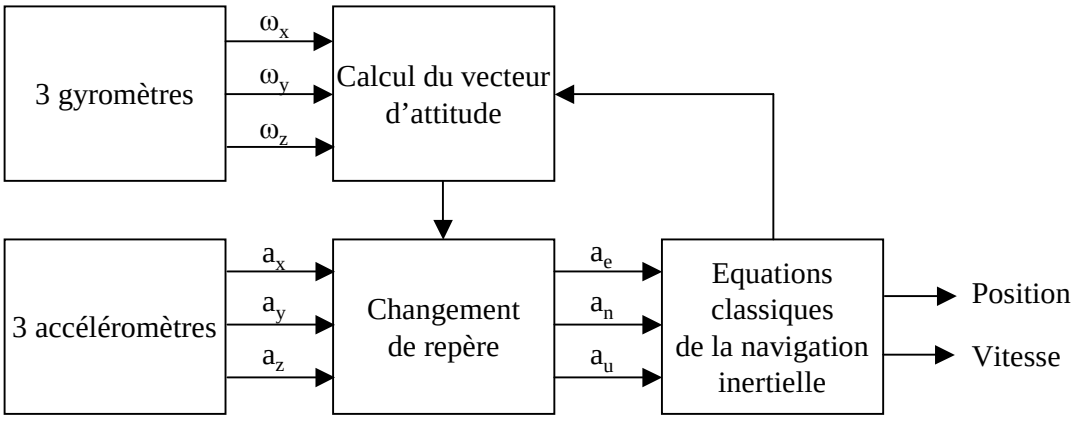


FIGURE I.4 – Principe de calcul de la pose avec une structure arrimée

Les avantages d'un système inertiel sont les suivants :

- *Bonne précision à court terme*
- *Fréquence d'échantillonnage élevée* : en général 100 à 150 Hz
- *Bonne disponibilité* : système autonome, qui ne dépend pas de dispositifs extérieurs
- *Solution complète* : position, vitesse, accélération et orientation

Ses inconvénients sont :

- *Mauvaise précision à long terme* : les effets du biais, de la marche aléatoire des accéléromètres et des gyroscopes entraînent une forte dérive après la double intégration
- *Localisation relative* : comme pour l'odométrie, un système inertiel a besoin d'être initialisé, la position et la vitesse doivent être calculées à partir de conditions initiales fournies par un capteur extéroceptif
- *Sensibilité à la gravité*

I.2.3 Autocalibration *via* l'image

Le principe de l'autocalibration *via* l'image est de s'appuyer sur les images acquises par une (ou plusieurs) caméra(s) embarquée(s) sur le véhicule. La localisation, relative ici, consiste à considérer deux images successives prises par la caméra aux instants t et $t+1$ et à déterminer quel a pu être le déplacement de la caméra entre t et $t+1$ ayant pu produire ces deux images. Cornelis et Van Gool [30] ont récemment utilisé cette approche en s'appuyant sur la détection de points caractéristiques dans la séquence vidéo et sur leurs suivis temporels [6, 33, 59, 86, 89].

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.2.4 Recalages successifs de profils télémétriques 2D

Le laser 2D, placé sur le véhicule de sorte à balayer horizontalement l'espace, permet d'obtenir un grand nombre de profils télémétriques horizontaux de l'environnement, qui, recalés entre eux, permettent d'en déduire la position du véhicule. Cette approche utilisée dans [48, 126] permet d'obtenir une grande qualité de localisation, mais là encore, à court terme. En effet, les erreurs de recalage s'accumulent (notamment lors des forts changements de direction) et la position estimée du véhicule dérive alors inévitablement.

I.3 Localisation absolue

La localisation absolue est une technique de localisation qui permet de se repérer directement dans un repère lié à l'environnement. Elle fait appel à l'utilisation de capteurs dits extéroceptifs. Les capteurs extéroceptifs mesurent la situation absolue du véhicule au moyen d'observations de points de repère naturels (amers visuels) ou artificiels (balises, satellites ...) dont la position est connue dans un référentiel attaché à l'environnement. Les capteurs extéroceptifs dédiés à la localisation sont alors des détecteurs de balises artificielles actives ou passives et des systèmes de perception qui captent une image, vidéo ou télémétrique, de l'environnement, dont on extrait les amers naturels pour le recalage du véhicule. Notons que les capteurs extéroceptifs peuvent également servir à assurer la sécurité du véhicule (perception de l'environnement proche, contrôle de l'attitude de la plateforme) et à construire en ligne un modèle de l'environnement exploré [1].

Les méthodes de positionnement absolu peuvent être divisées en plusieurs catégories, basées sur l'utilisation d'amers (cf § I.3.1), de cartes de l'environnement (cf § I.3.2) ou de cartes routières numériques (cf § I.3.3). Les systèmes de positionnement par satellites sont décrits dans le paragraphe I.3.4.

I.3.1 Localisation par rapport à des amers

Les amers sont des éléments caractéristiques de l'environnement. Ce sont des points de repère fixes et connus, que le véhicule peut repérer et identifier afin de se localiser. Avant que le véhicule puisse utiliser les amers pour la navigation, les caractéristiques des amers doivent être connus et sauvegardés. La tâche principale

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

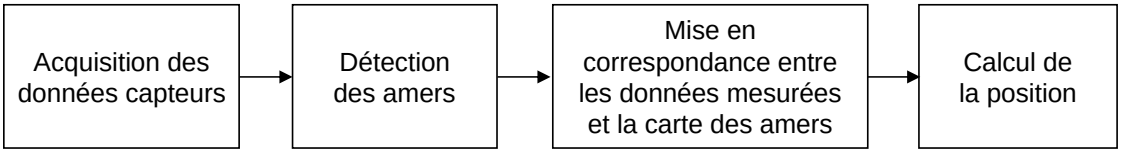


FIGURE I.5 – Principe de la localisation sur amers

de la localisation est alors de reconnaître les amers à partir des données acquises par les capteurs embarqués à bord du véhicule et de calculer la position du véhicule. Le principe de la localisation sur amers est résumé dans la figure I.5. Afin de simplifier le problème de localisation, il est souvent admis que la position courante du véhicule est connue approximativement. La qualité de la détection et de la mise en correspondance des amers dépendra alors de la précision de cette estimation initiale de la position du véhicule.

Différents types d'amers sont utilisés pour la localisation. Ces amers peuvent être classés en deux catégories : les amers naturels d'une part et les amers artificiels d'autre part. Les amers naturels sont des objets ou des éléments déjà présents dans l'environnement et ont une fonction autre que celle de la navigation. Par opposition, les amers artificiels sont des objets ou des marqueurs destinés spécifiquement à la navigation, ils sont placés dans l'environnement dans le seul but de permettre la navigation.

I.3.1.1 Les différentes classes d'amers

I.3.1.1.a Les amers naturels

Le principal problème de la localisation à partir d'amers naturels est de détecter les amers à partir des données capteurs et de les mettre en correspondance avec les amers de la base de données enregistrée dans la mémoire du véhicule. Des exemples d'amers naturels en environnement extérieur sont les portes, les fenêtres, les panneaux de circulation, les bords des routes ...

- amers télémétriques comme les poteaux
- amers visuels tels que les points de Harris, les points SIFT [79, 80]
- bords de fenêtres, marquages au sol

La tâche de perception est, en général, assurée par des capteurs télémétriques et/ou de vision.

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Un intérêt de la localisation basée sur des éléments naturels de l'environnement est qu'elle ne nécessite pas d'aménager l'environnement.

I.3.1.1.b Les amers artificiels

La localisation sur balises artificielles conduit souvent à des mesures plus fiables et plus robustes que les méthodes basées sur l'extraction et la mise en correspondance d'amers naturels. Cependant, elle requiert un aménagement préalable de l'environnement, difficilement envisageable à grande échelle, si on exclut le cas particulier du positionnement par satellites. Parmi les amers artificiels, on distingue les balises passives et les balises actives.

Les balises passives

Dans le cas des balises passives, qui ne transmettent pas de signaux, le système de perception doit dans un premier temps les détecter avant de pouvoir localiser le véhicule. Le positionnement du véhicule repose alors sur le système de perception constitué de capteurs extéroceptifs tels que des radars [36], des télémètres laser [56, 17] ou ultrason [77, 107], ou encore un système monovision ou stéréovision.

Les balises actives

Les balises actives émettent des signaux de nature luminineuse ou sonore ou encore électromagnétique, que le véhicule, équipé du dispositif de réception adéquat peut recevoir. Parmi les systèmes de localisation utilisant les balises actives, on peut citer les satellites GPS qui émettent dans le domaine des hyperfréquences, les balises acoustiques [118], les balises lumineuses du système « SIREM » [11] ou encore le positionnement par téléphonie mobile.

I.3.1.2 Méthodes de calcul de la pose

Plusieurs méthodes de calcul de la pose existent, parmi lesquelles la triangulation, la trilatération et la minimisation d'un critère. Le temps de calcul nécessaire à ses algorithmes afin de calculer la pose du véhicule est assez faible, ce qui constitue un avantage de la localisation sur amers. Cependant, ce type de méthodes de localisation suppose que la position des balises dans un repère de référence soit connue le plus parfaitement possible.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Triangulation

La triangulation permet de calculer la configuration du véhicule (deux inconnues de position et une de cap dans le cas 2D) à partir des trois mesures angulaires orientées (soit trois balises).

Le principe de la triangulation est donnée dans la figure I.6. Soient α_{21} et α_{32} les angles de gisement relatifs de trois balises B_1, B_2 et B_3 alignées. Soient x_i et y_i les coordonnées de la balise B_i dans le repère (I, J) $(y_i = 0)$. Alors la position (x, y) du véhicule est telle que :

$$\begin{aligned} y &= -x_2 \frac{W + \cot \alpha_{21}}{W^2 + 1} \\ x &= Wy + x_2 \end{aligned} \tag{I.1}$$

avec :

$$W = \frac{(x_3 - x_2)\cot \alpha_{32} - x_2 \cot \alpha_{21}}{x_3} \tag{I.2}$$

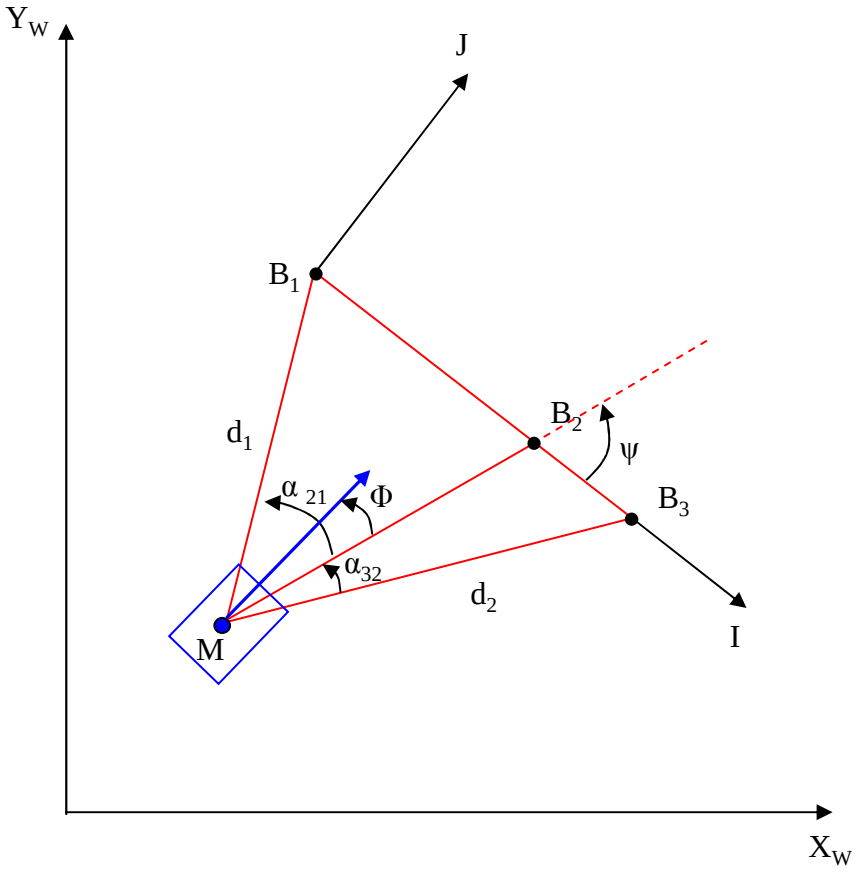


FIGURE I.6 – Principe de la triangulation

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Le cap θ est tel que :

$$\theta = \psi + \phi \tag{I.3}$$

avec

$$\psi = \arctan \frac{1}{W} \tag{I.4}$$

Trilatération

La trilatération est une méthode mathématique permettant de déterminer la position relative d'un point en utilisant la géométrie des triangles tout comme la triangulation. Mais contrairement à cette dernière, qui utilise les angles et les distances pour positionner un point, la trilatération utilise uniquement des distances.

Le principe de la trilatération est donnée dans la figure I.7. En connaissant la distance d_1 par rapport au point de référence B_1, on déduit que l'objet cherché se trouve sur le cercle de rayon d_1. En ajoutant d_2, la distance par rapport au point B_2, la position cherchée se réduit à deux points M et M'. Finalement, en ajoutant la distance d_3 par rapport au point B_3, la seule localisation possible de l'objet reste le point M.

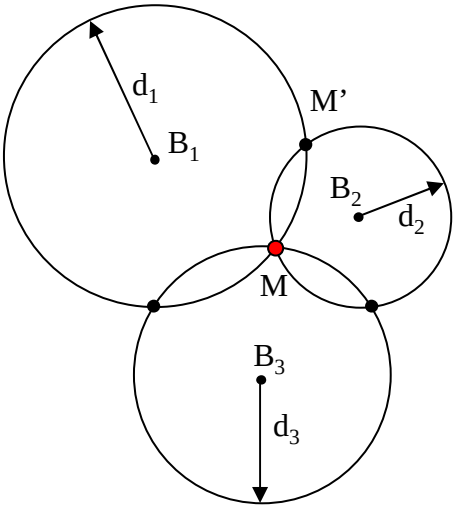


FIGURE I.7 – Principe de la trilatération

La trilatération est notamment utilisée par le système de positionnement par satellites américain Global Positioning System (GPS) qui est présenté plus en détail dans la section I.3.4.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Minimisation d'un critère

Les techniques de triangulation et de trilatération vues précédemment sont des solutions analytiques qui nécessitent peu de balises. Cependant, dans le cas de balises masquées ou de signaux fortement bruités, la qualité de la localisation décroît. Ainsi, on est amené à utiliser un nombre plus important de balises. La position du véhicule est alors estimée de manière numérique. La position optimale est obtenue en minimisant un critère basé sur l'écart entre la position exacte et la position estimée à partir de l'observation des balises. Ce critère de type moindres carrés est obtenu à l'aide de méthodes d'optimisation non linéaires tels que l'algorithme du gradient, la méthode de Newton-Gauss ...

I.3.2 Localisation sur carte de l'environnement

Afin de déterminer sa localisation, un véhicule automatisé a accès à deux types d'information. D'une part, le véhicule dispose d'information grâce aux observations et aux actions réalisées durant la navigation. D'autre part, le véhicule peut avoir à sa disposition une information *a priori*, qui a été collectée par le véhicule lui-même ou fournie par une source extérieure dans une phase d'initialisation.

En général, l'information *a priori* fournie au véhicule décrit l'environnement dans lequel il évolue. Les éléments invariants dans le temps sont spécifiés et peuvent alors être utilisés pour déterminer la position. L'information *a priori* prend souvent la forme de cartes numériques.

Une carte numérique peut être disponible sous trois formes :

- la carte peut être apprise au préalable par le véhicule dans une phase d'exploration
- la carte peut également être fournie par une source extérieure dans une phase d'initialisation
- une troisième possibilité est que le véhicule apprenne la carte de l'environnement pendant la navigation. La localisation et cartographie simultanée, connue en anglais sous le nom de SLAM (Simultaneous Localization And Mapping) consiste, pour un robot ou véhicule autonome, placé à un endroit inconnu dans un environnement inconnu, à simultanément construire d'une manière incrémentale une carte de cet environnement et de s'y localiser [110].

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.3.3 Localisation sur carte routière

La localisation d'un véhicule sur carte routière, connu sous le terme anglo-saxon de « map-matching » ou « roadmap-matching » , cherche à mettre en correspondance la position et la trajectoire estimées du véhicule avec une représentation cartographique existante du réseau routier. Une carte routière numérique est la représentation vectorielle numérique d'une carte routière. La numérisation est effectuée grâce à des systèmes d'information géographique ou des logiciels de C.A.O.

Dans les bases de données routières, les routes sont le plus souvent modélisées par leurs lignes centrales et les points d'intersection ou les points extrémités (jonctions) sont modélisés par des noeuds. Des points intermédiaires ou points de passages sont également utilisés afin de mieux modéliser les courbes. Ainsi, plus il y a de points intermédiaires, alors mieux une route sera modélisée. Outre la topologie et la géométrie du réseau routier, les cartes numériques contiennent également d'autres informations nécessaires à la navigation telles que le type et la classe des routes, les restrictions de routes, le sens de circulation, le nom des rues, les ponts et tunnels, des points d'intérêts (hôtels, restaurants ...).

La localisation sur carte routière nécessite deux étapes :

- la sélection du segment sur lequel le véhicule se situe,
- l'estimation de la position du véhicule sur ce segment

Différentes approches ont été proposées pour répondre au problème de la localisation sur carte routière. Ces approches sont très variées et vont de l'utilisation de techniques simples de recherche [73] à des techniques plus avancées telles que l'utilisation d'un filtre de Kalman étendu, de la logique floue et de la théorie des croyances [40, 91]. Les méthodes de map-matching peuvent être classées selon 4 catégories :

1. les approches géométriques [10]
2. les approches topologiques [54, 93, 123]
3. les approches basées sur les probabilités [87]
4. des approches plus avancées telles que le filtre de Kalman étendu, la théorie de l'évidence de Dempster-Shafer, le filtrage particulaire, la logique floue, l'inférence bayésienne ... [40, 57, 74, 91, 105].

Dans sa thèse, El Badaoui [39] dresse un état de l'art des méthodes de localisation sur carte routière. Le lecteur intéressé par les algorithmes de map-matching pourra également se référer à l'article de Quddus *et al.* [92].

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.3.4 Positionnement par satellites

Développé et contrôlé par les départements de la défense et des transports des Etats-Unis d'Amérique (http://pnt.gov/), le GPS a principalement été créé pour l'armée américaine. Il est actuellement utilisé par un nombre sans cesse croissant d'utilisateurs, notamment civils. Depuis le lancement du premier satellite en 1978 et son fonctionnement actuel, le système a subit de nombreuses évolutions.

Le système GPS offre deux services se différenciant principalement par la précision, l'accès et une garantie de fonctionnement.

- Le service SPS (*Standard Position Service*). Ce service est mis à disposition des utilisateurs civils. Une dégradation volontaire a été introduite jusqu'en mai 2000 et limitait la précision de positionnement à 100 m. La précision actuelle est de l'ordre de quelques mètres avec une très bonne visibilité satellitaire.
- Le service PPS (*Precise Position Service*). Ce service est réservé aux militaires américains et à certains utilisateurs autorisés par le département de la défense américain. Les récepteurs sont équipés d'algorithmes de décryptage spécifiques.

I.3.4.1 Constitution du système GPS

Le système GPS est constitué de trois parties : la partie spatiale, la partie de contrôle et la partie utilisateur.

Partie spatiale

La partie spatiale est constituée d'un ensemble de 24 satellites répartis sur 6 plans orbitaux (cf Figure I.8). Ces satellites évoluent à une altitude d'environ 20000 km et mettent 12 heures pour effectuer une rotation autour de la terre. Chaque satellite possède un oscillateur qui fournit une fréquence fondamentale de 10,23 MHz calibrée sur des horloges atomiques. L'émetteur génère deux ondes (L1 et L2) de fréquences respectives 1575,42 MHz et 1227,60 MHz. Il transmet régulièrement des signaux horaires, la description de l'orbite suivie (éphéméride) et diverses autres informations appelées *GPS data* (almanachs, état des satellites, décalage d'horloge des satellites ...).

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

FIGURE I.8 – Constellation des satellites du système GPS

Partie de contrôle

La partie de contrôle qui permet de piloter et surveiller le système, est composée de cinq stations au sol situées à travers le monde (cf Figure I.9). Elles écoutent les signaux émis par les satellites, les diagnostiquent, calculent leurs éphémérides (orbites, paramètres d'horloge, codes, etc...), transmettent les données aux satellites et assurent leur pilotage. La station principale se trouve dans le Colorado.

FIGURE I.9 – Stations de contrôle au sol

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Partie utilisateur

La partie utilisateur comprend l'ensemble des récepteurs civils et militaires qui reçoivent les informations des satellites. Pour déterminer la position, la vitesse et le temps, au moins quatre signaux satellites de la constellation GPS sont nécessaires.

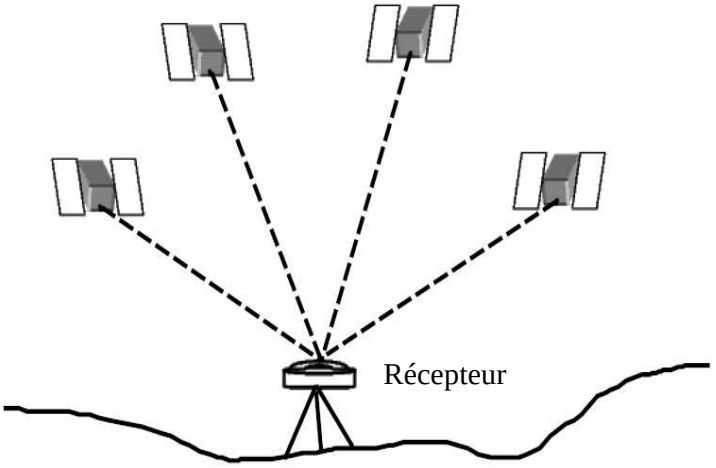


FIGURE I.10 – Positionnement absolu par satellites

Remarque

Le système GPS est en cours de modernisation et prendra le nom de GPS III à l'horizon 2017. D'ici là, il se sera doté de nouvelles fréquences (L2C et L5), d'une plus grande puissance d'émission et de davantage de stations de poursuite (contrôle).

I.3.4.2 Les observations GPS

Les satellites NAVSTAR transmettent émettent simultanément sur deux fréquences porteuses appelées L1 à 1575.42 MHz et L2 à 1227.60 MHz, dont les longueurs d'onde sont de 19 et 24 cm respectivement. Deux autres fréquences servent pour la génération de codes pseudo-aléatoires : 1,023 MHz pour le code C/A (*Coarse Acquisition*) et 10,23 MHz pour le code P (*Precise*).

Chaque satellite possède des codes qui lui sont propres et orthogonaux aux autres, ce qui permet de l'identifier. Un code pseudo-aléatoire est une séquence numérique qui semble aléatoire mais qui est en fait parfaitement connue en termes de constitution et d'instants d'émission. Les codes C/A et P sont binaires et modulent les phases des porteuses L1 et/ou L2.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

L'onde porteuse L1 est modulée par les deux codes C/A et P. Pour sa part, l'onde porteuse L2 est modulée par le code P. Le code C/A a une longueur de 1023 bits et se répète toutes les millisecondes, ce qui correspond à peu près à 300 m de parcours. Le code P a une longueur bien supérieure (de l'ordre de 10^{14} bits). La fréquence des ondes porteuses et la séquence des codes sont régies par des horloges atomiques à bord des satellites.

En parallèle avec l'émission des codes, des données de navigation sont transmises (sur L1 et L2) à un faible débit de 50 bits/sec. Elles fournissent des éphémérides, c'est-à-dire des informations telles que l'état de santé du satellite, ses paramètres d'orbite, des paramètres de correction d'horloge, ...Le message complet, long de 1500 bits, est donc transmis en 30 secondes, ce qui peut-être long dans certaines applications où le récepteur GPS démarre à froid. A la fin de chaque message, une partie des almanachs des autres satellites est transmise. La transmission de tous les almanachs dure 12 min 30 s.

Il y a 3 types d'observations GPS possibles : les mesures de pseudo-distance, les mesures de phase de l'onde porteuse et les mesures de fréquence Doppler.

Le terme observable est utilisé dans la communauté GPS pour parler des mesures brutes d'un récepteur. Sur chaque satellite en vue, un récepteur mono-fréquence civil mesure toujours le code C/A et parfois la phase et le Doppler sur L1. Un récepteur bi-fréquence peut quant à lui, mesurer le code C/A sur L1, la phase et le Doppler sur L1 et L2. Un récepteur militaire mesure en plus le code P sur L1 et L2. Il est à noter que certains récepteurs civils performants, grâce à des méthodes de traitement du signal d'extraction de phase sur L2, délivrent également des observables bruités de code P sur L1 et L2. Les erreurs d'observable d'une même porteuse sont souvent considérées comme peu corrélées du fait que la traversée dans l'atmosphère n'induit pas les mêmes effets sur le code et la phase. Le Doppler, indispensable pour démodulation des signaux des satellites au niveau des boucles à verrouillage de phase, est mis à profit notamment pour la mesure de vitesse.

La **mesure de pseudo-distance (ou la mesure de code)** est, en termes simples, une mesure du temps de propagation requis pour qu'une marque horaire transmise par un satellite atteigne le récepteur sur Terre. Ces marques horaires sont codées sur les ondes porteuses par la technique de modulation de phase. Le temps est fourni avec une très grande précision par les oscillateurs des satellites : de l'ordre de 10^{-14} secondes. Afin qu'un récepteur puisse reconnaître le satellite observé, chaque satellite transmet un code qui lui est propre. Une réplique de la séquence du code est

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

générée par le récepteur en même temps que le satellite. Le décalage que doit subir la réplique afin de coïncider avec le code reçu correspond au temps de propagation qu'a pris le signal pour parcourir la distance satellite-récepteur. Cette différence de temps multipliée par la vitesse de la lumière dans le vide (environ 3.10^8 m/s) donne une mesure de distance. Cette mesure est faussée entre autres par la propagation de l'onde dans l'atmosphère ainsi que par les erreurs de synchronisation entre les horloges du satellite et du récepteur. Une erreur de 1/1000 de seconde représente une erreur de distance de 300 km.

L'ordre de grandeur de la résolution de la mesure de pseudo-distance effectuée avec le code C/A est d'environ ±3 m, celle avec le code P est d'environ ± 0.3 m (le code P est dix fois plus précis que le code C/A). Cependant, notons que certains récepteurs, récemment introduits sur le marché, permettent des mesures de pseudo-distance sur le code C/A aussi précises que celles effectuées sur le code P. À ces erreurs s'ajoutent les erreurs d'horloges, d'orbites, et de réfractions ionosphérique et troposphérique. L'avantage des mesures avec le code P, en plus d'être plus précises, est qu'elles peuvent être corrigées pour le délai ionosphérique puisque le code P est transmis sur les 2 ondes porteuses de fréquences différentes.

Pour des fins de sécurité, l'armée américaine a remplacé, en 1994, le code P (dont la séquence est connue) par un code secret Y. Puisque la séquence complète du code Y n'est pas connue, ce signal ne peut être brouillé par un ennemi potentiel. Ce dispositif s'appelle l'antibrouillage A-S (*Anti-Spoofing*). Cependant, par différentes techniques de corrélation de code, les manufacturiers ont réussi à fabriquer des récepteurs qui peuvent tout de même effectuer des mesures de pseudo-distance avec le code Y.

La **mesure de phase** est utilisée pour des applications plus précises. Elle exploite directement la phase sur les ondes porteuses L1 et L2 en déterminant le déphasage entre le satellite et le récepteur. Le principe de la mesure de phase est la comparaison de la phase de l'onde reçue au récepteur avec la phase d'une onde générée à l'intérieur du récepteur, qui est une réplique du signal du satellite. Ce mode de mesure est plus compliqué à mettre en oeuvre que celui basé sur le code. Pour cela, il existe des récepteurs spécifiques permettant d'exploiter le code et la phase. Le traitement des données se fait en mode postopératoire « en temps différé » ou PPK (*Post Processing Kinematic*, mode de traitement cinématique en temps différé) à l'aide de logiciels spécialisés. On peut aussi traiter ces signaux en temps réel, on parle alors de RTK (*Real Time Kinematic*, cinématique temps réel). Théoriquement, la différence de phase oscille entre 0 et 2π. La mesure de phase peut être convertie en mètres puisqu'on connaît la longueur de l'onde porteuse. Malheureusement, le nombre en-

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

tier de longueurs d'onde initial contenu dans la distance récepteur-satellite n'est pas mesurable par le récepteur. Cette inconnue est appelée l'ambiguïté de phase initiale. Par contre, le récepteur est à même de compter le nombre entier de cycles (ainsi que la partie fractionnaire) cumulés depuis « l'époque » (ou le temps) d'observation initiale, s'il n'y a pas d'interruption dans la réception du signal. Les interruptions provoquent des sauts de cycles et sont principalement causées par les obstructions (bâtiments, montagnes, arbres, ...) entre les satellites et le récepteur. La mesure de phase peut être interprétée comme une mesure précise de la variation de la distance récepteur-satellite depuis l'époque initiale.

La résolution de l'ambiguïté peut se faire soit en initialisant le GPS en un point dont la position est connue avec une très grande précision (clous IGN), soit par des algorithmes d'initialisation. Dans les modèles récents de GPS centimétriques, cette phase d'initialisation dure environ une minute et peut être effectuée pendant le déplacement du véhicule, contrairement aux premiers modèles dont la phase d'initialisation était bien plus longue et exigeait un arrêt du véhicule, rendant leur utilisation dans des applications automobiles peu pratique. Si l'ambiguïté de phase initiale peut être résolue, la mesure de phase ainsi corrigée représente une mesure précise de la distance récepteur-satellite. La résolution d'une mesure de phase est de quelques millimètres.

La **mesure de fréquence Doppler** est la différence entre la fréquence reçue et la fréquence nominale de transmission causée par le mouvement relatif entre le satellite et le récepteur (Delta-range). Cette mesure est surtout utilisée pour déterminer la vitesse instantanée des récepteurs mobiles et pour détecter et corriger les sauts de cycle potentiellement présents dans les mesures de phase.

I.3.4.3 Calcul de la position

La position donnée par le GPS est celle de l'antenne du récepteur de l'utilisateur. Le référentiel GPS est le *WGS84* (*World Geodetic System 1984*). Le principe de la détermination de cette position est la trilatération : on mesure la distance entre l'utilisateur et un certain nombre de satellites dont on connaît les positions puis, de façon imagée, on obtient des sphères centrées sur des satellites dont l'intersection permet de reconstituer les trois informations de position de l'utilisateur (latitude, longitude et hauteur ellipsoïdale). La distance par rapport à chaque satellite est obtenue grâce à sa relation de proportionnalité avec le temps de parcours du signal émis par chaque satellite respectif. En effet, connaissant t_{S_i} , l'heure exacte d'émission d'un message

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

par un satellite S_i et l'heure exacte, t_R, d'arrivée de ce message sur le récepteur R, on peut en déduire le temps de propagation du signal. En multipliant ce temps par la célérité de la lumière, on obtient la distance $\Delta_{S_i,R}$ du satellite S_i au récepteur R.

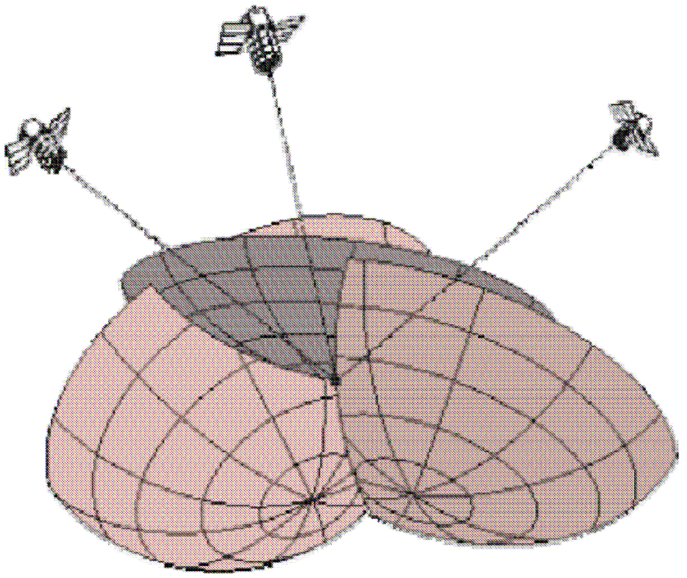


FIGURE I.11 – Positionnement tridimensionnel à partir de trois satellites

Dans le cas idéal, trois satellites suffisent pour déterminer la position du récepteur (cf Figure I.11). En effet, l'intersection des trois sphères distinctes donne 2 points. L'un de ces deux points se situe loin de la surface du globe terrestre et n'est donc pas crédible. En réalité, quatre satellites au moins sont nécessaires, si on prend en compte les erreurs d'horloge. En effet, on peut remarquer que la distance $\Delta_{S_i,R}$ mesurée (la pseudo-distance) est faussée par le décalage temporel ΔT_i entre le temps GPS et celui du récepteur. Ce décalage qui peut prendre de grandes valeurs est considéré comme une quatrième variable à estimer.

En outre, pour une utilisation plus efficace du GPS, les satellites NAVSTAR reçoivent quotidiennement des paramètres de corrections diffusés par les stations. Il est donc possible à l'utilisateur de régler son horloge lors du calcul de positionnement qui fournit l'imprécision ΔT de l'horloge utilisateur (en négligeant l'erreur de l'horloge GPS).

La prise en compte de toutes les erreurs qui perturbent la propagation du signal GPS permet d'atteindre des précisions sub-métriques. Les sources d'erreurs sont vues en détail dans le prochain paragraphe.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.3.4.4 Les sources d'erreurs

Il existe de nombreuses sources d'erreurs pouvant dégrader le positionnement : la synchronisation de l'horloge des satellites, la géométrie des satellites, la traversée des différentes couches de l'atmosphère, les effets de multi-trajet du signal GPS et les erreurs de mesures du récepteur.

Horloge des satellites

La synchronisation entre les horloges des satellites et celles des récepteurs est une source d'erreur importante sur la précision. Pour remédier à cet obstacle technique, le positionnement s'effectue alors avec au moins quatre satellites. Cependant la nécessité d'avoir quatre satellites pour effectuer une mesure pose un problème, celui de la couverture GPS. En effet, la configuration actuelle des satellites ne couvre que 95% de la planète et donc dans 5% des cas la mesure de positionnement n'est possible qu'avec trois satellites, ce qui nuit grandement à sa précision. Notons néanmoins que les 5% les moins couverts sont les pôles.

Géométrie des satellites

La disposition de la constellation induit une forte dégradation géométrique de la précision deux fois par jour dans certaines zones limitées pour des durées de 5 à 30 minutes. Par ailleurs, l'élévation des satellites et leurs positions relatives par rapport au récepteur va affecter la qualité des positions calculées. Cette qualité peut être caractérisée par le calcul des facteurs de dégradation géométrique GDOP (*Geometric Dilution Of Precision)* donnant ainsi à l'utilisateur une idée de la confiance à apporter à une mesure.

La réfraction dans l'ionosphère

Les ondes radios issues des satellites du système GPS orbitant à 20000 km d'altitude doivent traverser en premier l'ionosphère. La propagation du signal GPS est alors perturbée comme toute onde électromagnétique traversant un milieu conducteur (l'ionosphère est composée d'ions et d'électrons). Cette perturbation se manifeste par un retard et conduit à une surestimation du temps de propagation du signal et donc la distance satellite-récepteur.

La réfraction dans la troposphère

Le temps de propagation de l'onde GPS est également affecté par la teneur en

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

vapeur d'eau de la couche basse de l'atmosphère : la troposphère. Il serait donc nécessaire de connaître cette quantité avec précision tout le long du trajet suivi par l'onde pour établir les imprécisions sur les mesures de pseudo-distances qui en découlent.

Les trajets multiples

Le phénomène de multi-trajet est dû à la réflexion des ondes radios émises par les satellites, par les infrastructures en milieu urbain. Ces réflexions vont souvent augmenter la distance du trajet satellite-récepteur et donc aussi la durée de propagation du signal. Le récepteur va alors calculer une distance satellite-récepteur fausse et donc commettre une erreur sur la position.

Le bruit des mesures

Le bruit de mesure est estimé à moins de 1% de la longueur du signal sur lequel se fait la mesure, dans le cas d'une mesure de phase comme dans celui d'un calcul de pseudo-distance.

I.3.4.5 GPS différentiel

Le DGPS (*Differential Global Positioning System*) est une technique conçue pour améliorer la précision du GPS.

Le principe fondamental du DGPS est que deux récepteurs (ou plus) observant les mêmes satellites feront des erreurs de mesure d'autant plus semblables que ces récepteurs seront proches les uns des autres. En plaçant un récepteur dit de référence sur une position parfaitement connue, il est possible de calculer sa position théorique et de la comparer aux mesures fournies par les signaux des satellites de positionnement. La différence de valeurs donne l'erreur de mesure. Cette erreur mesurée sur le récepteur de référence peut ensuite servir à corriger les erreurs de mesure des récepteurs placés sur des positions inconnues dans la zone couverte par le DGPS (cf Figure I.12). La précision du positionnement par satellites passe alors de 15 mètres à quelques mètres. Cette technique permet également de compenser les erreurs propres au système de positionnement par satellites mais peut ne pas corriger systématiquement les erreurs dues à l'environnement, lorsque les signaux sont déviés par de hautes constructions ou des montagnes voisines, créant ainsi un phénomène de trajets multiples. La précision du DGPS diminue lorsque les récepteurs du système sont placés à de plus grandes distances les uns des autres.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

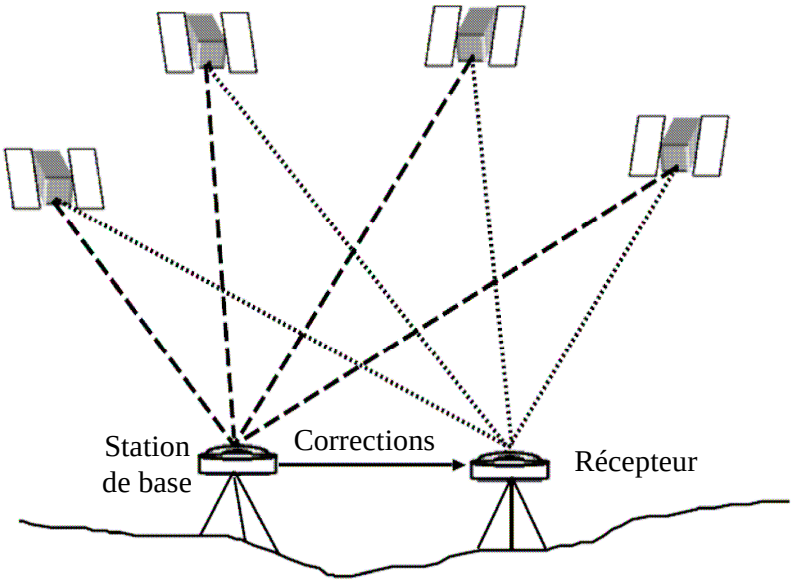


FIGURE I.12 – Principe du DGPS

Des techniques DGPS plus sophistiquées peuvent porter la précision du positionnement à quelques millimètres. Les corrections peuvent s'effectuer en temps différé ou en temps réel : dans le premier cas, les mesures « brutes » du récepteur de référence et des récepteurs « mobiles » sont enregistrées puis traitées par un logiciel ad hoc (post-traitement) ; dans le deuxième cas, les corrections sont calculées « en ligne » et diffusées immédiatement vers les récepteurs mobiles, qui les appliquent avant de calculer leur position avec une précision pouvant varier de quelques millimètres à quelques mètres, selon la méthodologie utilisée et la qualité de la transmission de données en temps réel.

Les systèmes d'augmentation de performances

Les modes différentiels de plus en plus répandus sont ceux de type SBAS (*Satellite-Based Augmentation System*). Un système d'augmentation des signaux satellitaires est un système qui utilise un réseau de satellites géo-stationnaires et de stations au sol afin d'améliorer la performance du système de navigation par satellites en apportant des corrections de signal et des satellites supplémentaires dont les positions sont connues et constantes (géostationnaires). Trois SBAS sont actuellement exploités dans le monde, couvrant des zones géographiques spécifiques : le Wide Area Augmentation System (WAAS) en Amérique du Nord, le European Global Navigation Overlay System (EGNOS) en Europe et le Multi-transportation Satellite Augmentation System (MSAS) en Asie.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Il existe également des sociétés privées qui proposent des corrections par SBAS parmi lesquelles OmniSTAR. OmniSTAR est un service de diffusion *via* satellite géo-stationnaire payant. C'est un système de diffusion de corrections GPS différentiel en temps réel couvrant l'Europe et l'Amérique du nord. Les corrections sont issues d'un ensemble de stations de base situées à travers le monde. Il utilise ce réseau de stations de référence (ou stations de base) pour mesurer les erreurs induites dans le signal GPS. Ces données de référence sont ensuite transmises au centre de contrôle où elles sont vérifiées quant à leur intégrité et à leur fiabilité, puis envoyées par liaison ascendante à un satellite géo-stationnaire qui diffuse les données sur son empreinte. Le centre de contrôle utilise une solution VRS (*Virtual Reference Station*) pour fusionner les informations des stations de base. Le principe est de modéliser finement les différentes sources d'erreurs en s'appuyant sur les stations de base, puis d'interpoler l'influence de ces erreurs à proximité de l'utilisateur en simulant les données d'une station virtuelle.

Citons également les trois opérateurs suivants : Orphéon (initialement Acthyd), Sat-Info (`www.sat-info.fr`) et Teria (`www.reseau-teria.fr`). Ils proposent un service privé payant basé sur un réseau dense de stations de référence, concentrant leurs corrections sur un serveur internet.

I.3.4.6 Les autres systèmes de positionnement par satellites

Dans cette partie, cinq autres systèmes de positionnement par satellites sont brièvement présentés, à savoir les trois systèmes globaux : GLONASS, Galileo et Compass et les deux systèmes régionaux : IRNSS et QZSS.

GLONASS

Le système GLONASS (*Global'naya Navigatsionnaya Sputnikovaya Sistema* en russe) de l'ex-union soviétique, aujourd'hui Russie, était également une constellation entièrement opérationnelle, mais avec l'écroulement de l'union soviétique, il n'était plus entretenu, provoquant des trous de couverture et une disponibilité partielle (6 à 8 satellites en 2006). En 2007, le président Poutine a déclaré vouloir relancer son programme. Aujourd'hui 17 satellites sont en service et peuvent être complétés par le service GPS pour les récepteurs compatibles GNSS. 24 satellites sont prévus pour 2011, répartis sur 3 plans orbitaux à 19100 kms de la Terre, avec une inclinaison de 64,8°.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Galileo

L'Union européenne a signé avec l'agence spatiale européenne en Mars 2002 l'accord sur le développement du système global Galileo. Le coût est estimé à environ 3 milliards d'euros. Le système est prévu opérationnel en 2013. Le premier satellite expérimental a été lancé le 28 décembre 2005. Un second satellite de validation a été lancé en 2008. Les signaux de navigation de Galileo seront compatibles avec ceux du futur GPS, permettant aux récepteurs de les combiner pour augmenter la précision ainsi que la véracité du point. 30 satellites sont prévus pour 2013, répartis sur 3 plans orbitaux à 23616 kms de la Terre, avec une inclinaison de 56,8° et une révolution en 14h21mn.

Compass

La Chine a indiqué son intention d'étendre son système régional Beidou en système global. Ce programme est appelé *Compass* par l'agence d'information chinoise officielle « Xinhua News Agency ». Le système *Compass* doit comporter 30 satellites en orbite et 5 géo-stationnaires. Cette annonce est accompagnée d'une invitation à d'autres pays désirant y collaborer, alors que la Chine est également engagée dans le programme Galileo.

IRNSS

Le système IRNSS (*Indian Regional Navigational Satellite System*) est un projet de système autonome de navigation régionale construit et contrôlé par le gouvernement indien. Il doit permettre une précision absolue de 20 mètres sur l'Inde et jusqu'à 1500 à 2000 km au voisinage. Le but est un système entièrement sous contrôle indien, le segment spatial, terrestre et les récepteurs étant développés par l'Inde. Le projet a été approuvé par le gouvernement indien en Mai 2006, avec un objectif de développement en 6 à 7 ans.

QZSS

Le système QZSS (*Quasi-Zenith Satellite System*), est développé par le Japon pour un premier lancement en 2008. Il sera constitué de trois satellites géostationnaires permettant le transfert de temps et une augmentation du GPS. Il couvrira le Japon et sa région.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

I.4 Localisation hybride

Les deux techniques de localisation précédemment présentées possèdent chacune des avantages et des inconvénients, qui sont résumés dans le tableau I.1. Le principal avantage de la localisation à l'estime est sa fréquence d'acquisition élevée, ce qui permet de calculer précisément le déplacement du véhicule à court terme. La disponibilité de la localisation à l'estime est également un facteur très intéressant. Cependant, cette méthode dérive à long terme à cause de l'accumulation des erreurs lors des intégrations successives. De plus, la localisation à l'estime nécessite une initialisation de son point de départ. En revanche, la positionnement absolu ne présente pas de problèmes de divergence, ni d'initialisation, mais est sujet aux problèmes de latence, de fréquence souvent insuffisante, et de disponibilité. En effet, le résultat de la localisation dépend de la visibilité des sources d'informations, c'est-à-dire la visibilité des satellites pour la localisation par GPS, la visibilité des stations émettrices pour la localisation par téléphone cellulaire ou encore la visibilité des amers visuels pour la localisation par vision.

On constate qu'une méthode de localisation utilisée seule ne peut pas permettre une localisation fiable du véhicule. En effet, les informations sensorielles fournissent des mesures imparfaites, c'est-à-dire incertaines, imprécises, incomplètes, contradictoires. La localisation hybride consiste alors à utiliser conjointement la localisation relative et la localisation absolue. L'objectif est d'exploiter au mieux les avantages de chacune des sources d'information, tout en essayant de pallier leurs limitations individuelles. En les fusionnant, on peut alors obtenir une information globale plus fiable et plus complète. La combinaison de plusieurs sources d'informations se base sur deux critères : la complémentarité et la redondance.

- la fusion d'informations redondantes permet de diminuer l'incertitude globale du système de positionnement, et d'augmenter la robustesse aux défaillances capteurs
- la combinaison d'informations complémentaires permet de fournir une information plus complète qui ne pourrait pas être obtenue en utilisant les sources d'information individuellement.

Le recalage dynamique consiste à fusionner des informations proprioceptives et des informations extéroceptives. En général, le positionnement absolu a pour rôle de corriger régulièrement l'estimation relative de la pose afin de limiter la dérive de celle-ci.

L'hybridation des sources d'information peut se faire par le biais de nombreuses

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

TABLE I.1 – Avantages et inconvénients de la localisation absolue et de la localisation relative

Localisation absolue	Localisation relative
(+) Aménagement parfois nécessaire de l'environnement (Balises), Sensibilité aux occultations	(-) Autonomie vis-à-vis de l'environnement
(-) Fréquence faible	(+) Fréquence élevée (fonctionnement « quasi continu »)
(+) Erreurs indépendantes de la distance et/ou du temps	(-) Accumulation d'erreurs avec le temps (intégration) ⇒ dérive à long terme
(+) Localisation par rapport à un repère attaché à l'environnement	(-) Localisation par rapport à un repère attaché à la configuration initiale du véhicule ⇒ initialisation nécessaire

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

techniques de fusion de données. On distingue les approches analytiques telles que le filtre de Kalman [71] et ses extensions, les approches numériques telles que le filtre particulaire [37, 45, 57] et ses extensions et les approches ensemblistes (à erreur bornée ou par analyse d'intervalles) [58, 72, 53].

Couplage serré et couplage lâche

Dans le cadre de l'hybridation GPS avec des capteurs proprioceptifs et/ou une cartographie de l'environnement, on distingue deux types d'approches : le couplage serré et le couplage lâche.

Dans un couplage serré, les données capteurs à utiliser sont les mesures brutes des différents capteurs. Dans sa thèse, Bonnifait [11] a utilisé cette approche pour localiser en 3D un robot mobile à évolutions lentes avec le capteur SIREM, un odomètre et des inclinomètres. Le couplage serré a également été utilisé dans [56] pour hybrider un ladar avec des capteurs proprioceptifs et dans [16] afin de fusionner un capteur goniométrique 3D avec des techniques à l'estime. Dans sa thèse, Gaudin [49] a conçu un système hybride GPS-Estime pour la localisation en condition de visibilité satellitaire réduite. Les mesures brutes qu'il utilise pour le GPS sont les pseudo-distances des satellites. Dans [44], un algorithme de positionnement par couplage serré GPS/carte 2D est présenté. Les auteurs montrent que la pose peut être estimée avec seulement 3 satellites et une carte 2D dans les canyons urbains.

Dans un couplage lâche, les données capteurs à utiliser sont les solutions de navigation, c'est-à-dire les données de position (x, y) délivrées par le récepteur GPS. Cette méthode épargne au concepteur la gestion de la position des satellites, l'implantation de modèles de propagation d'onde dans l'atmosphère, l'utilisation de corrections différentielles, elle est donc plus simple. Cependant, cette approche ne permet pas de tirer bénéfice de la visibilité de quelques satellites uniquement, contrairement au couplage serré.

I.5 Conclusion : positionnement de notre approche

La cartographie a évolué. Les premières méthodes de roadmap-matching étaient basées sur des cartes routières numériques en deux dimensions manipulées par un SIG-2D. Puis, Jabbour [66], dans sa thèse, a ajouté une couche supplémentaire au SIG-2D : des images clefs. Il développe également une autre approche consistant à utiliser un lidar embarqué. Lors d'une première étape d'apprentissage, une carte de

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

l'environnement est construite à l'aide du lidar cartographiant des amers naturels. La carte des amers initiale, élaborée lors de la phase de construction, est ensuite mise à jour de façon simultanée à la localisation du véhicule en employant une méthode d'observation d'état probabiliste inspirée des méthodes robotique de SLAM. Parallèlement, des méthodes de localisation basées sur la vision ont été développées. Dans sa thèse, Royer [95] développe une méthode de localisation basée sur la vision monoculaire. Pour cela, une carte tri-dimensionnelle de la zone parcourue est construite pendant une phase d'apprentissage. Cette étape de cartographie repose sur un algorithme de reconstruction 3D à partir d'une caméra en mouvement. La carte est un modèle 3D de la scène observée constitué d'un nuage de points. La carte ainsi construite est ensuite utilisée pour se localiser, en recalant l'image courante sur le modèle 3D.

On constate donc l'évolution de la cartographie du 2D puis 2D-1/2 ou 2D augmenté jusqu'au 3D. C'est donc dans ce contexte d'évolution de la cartographie que nos travaux se situent. En effet, grâce aux technologies de la réalité virtuelle, des bases de données, . . . , une nouvelle source d'information, à savoir les modèles virtuels 3D de l'environnement (cf §II.1), se développe de plus en plus. Ainsi dans l'approche que nous avons développée, plutôt que de construire un modèle de l'environnement dans une phase d'apprentissage ou simultanément à la localisation, nous intégrons dans le processus de localisation un modèle virtuel 3D urbain.

Notre approche est alors la suivante. Une localisation à l'estime basée sur un modèle odométrique et des mesures acquises par des codeurs incrémentaux et un gyromètre permet d'obtenir une estimation quasi continue de la pose du véhicule. Cette estimation est recalée à l'aide d'un GPS. Par ailleurs, afin de limiter la dérive de la localisation à l'estime lors d'une longue absence du GPS ou de pallier les effets du multi-trajet, une autre observation absolue de la pose est construite. Cette observation appelée par la suite observation cartographique 3D est élaborée par la mise en correspondance entre le modèle virtuel 3D urbain et des images acquises par une caméra embarquée.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Chapitre II

Modèle virtuel 3D urbain et SIG-3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Résumé du chapitre

Ce chapitre s'intéresse à la source d'information *a priori* utilisée, à savoir un modèle virtuel 3D urbain. Un modèle virtuel 3D urbain est une représentation photo-réaliste en trois dimensions de l'environnement d'évolution du véhicule. Par ailleurs, l'outil permettant de manipuler la base de données du modèle 3D est également présenté. Il est qualifié de système d'information géographique 3D (SIG-3D). On détaille alors d'une part les entrées du SIG-3D c'est-à-dire la base de données du modèle 3D et les paramètres de la caméra virtuelle à configurer et d'autre part les différentes sorties du SIG-3D. Afin d'utiliser le modèle virtuel 3D urbain pour nos applications de navigation, une étape essentielle est le géo-référencement du modèle 3D.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sommaire

II.1 Modèle virtuel 3D urbain 46
II.1.1 Introduction 46
II.1.2 Le modèle Tecnomade de Nancy 50
II.2 Système d'information géographique 3D 55
II.2.1 Introduction 55
II.2.2 Le SIG-3D développé 56
II.2.2.1 Les entrées du SIG-3D 57
II.2.2.2 Les sorties du SIG-3D 59
II.2.2.2.a Image virtuelle 59
II.2.2.2.b Fichier de profondeur 60
II.2.2.2.c Image de profondeur 61
II.2.2.2.d Fichier des segments 3D 61
II.2.2.2.e Image des segments 3D visibles 63
II.2.2.3 Le géo-référencement du modèle 3D 64
II.2.2.3.a Transformation WGS84 vers Lambert I 64
II.2.2.3.b Transformation Lambert I vers le repère local du modèle 3D 64
II.3 Conclusion 65

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Deux types de sources d'information sont utilisées pour la localisation du véhicule dans le cadre de cette thèse : des informations acquises par des capteurs embarqués (GPS, caméra, odomètres et gyromètre) d'une part et une information *a priori* sur l'environnement d'autre part. L'information *a priori* sur l'environnement est un modèle 3D virtuel urbain manipulé par un système d'information géographique 3D (SIG-3D).

Dans ce chapitre, nous présentons le modèle virtuel 3D urbain utilisé (cf § II.1) dans nos expérimentations ainsi que le SIG-3D développé (cf § II.2) pour la navigation [81, 88, 128, 127].

II.1 Modèle virtuel 3D urbain

II.1.1 Introduction

Le nombre d'applications générant ou traitant de l'information géographique est depuis quelques années en constante augmentation [94]. L'éventail des secteurs d'activités impliqués dans cet essor est à la fois vaste et diversifié : la défense, l'aménagement, le tourisme ou le jeu utilisent tous aujourd'hui des données géographiques 3D. Dans [50], quelques exemples d'utilisation de ces données sont recensés en le regroupant par thèmes d'applications :

- **Aménagement**
  - étude d'impact paysager (zone rurale, milieu urbain, volet paysager des permis de construire)
  - analyse pour l'insertion de projets dans un site (ouvrage d'art, bâtiment, station de ski ...)
  - analyse pour des études d'urbanisme
  - visualisation de l'évolution de la végétation sur plusieurs années
  - parcours d'une future route et évaluation de son impact
  - aide à la cartographie des servitudes d'utilité publique (SUP) tels que les plans de servitude aéronautique
- **Risques et pollutions**
  - cartographie 3D du bruit et analyse de l'impact d'aménagements (écran anti-bruit ...)
  - cartographie et étude de pollutions atmosphériques

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

  - simulation d'inondations, application de modèles d'écoulement des eaux (mise au point de plans de prévention des risques PPR)
  - évaluation des populations en zone d'aléa : répartition de la population par volumétrie des bâtiments

- **Sous-sol, sous-marin**
  - analyse de mesures d'exploration pétrolière
  - aide à la pose de pipe-lines sous-marins
  - cartographie et analyse géologique
  - spéléologie (représentation de réseaux souterrains)
- **Tourisme**
  - présentation de villes : villes virtuelles
  - présentation de régions touristiques, d'un département
  - reconstitution virtuelle d'un patrimoine historique (village, cathédrale, grotte ...)
  - mise à disposition d'informations par Internet
- **Télécoms**
  - aide à l'implantation d'antennes (téléphone, télévision)
  - calcul d'inter-visibilité entre émetteurs
- **Transports**
  - simulateurs de poste de conduite (avion, train, bateau)
  - simulation de trafic pour évaluer les capacités d'un nouvel aménagement routier
  - vision haute pour les pilotes d'avions
- **Jeux, Films**
  - réalisation et visualisation de paysages virtuels
- **Géomarketing**
  - visualisation de zones d'influence (dans ce cas, la 3D n'est pas indispensable, mais améliore la visualisation et aide à l'analyse des zones d'influences)
- **Défense**
  - entraînement, simulateurs de conduite (chars, avions, navires)
  - préparation de mission
  - vision haute pour les pilotes d'avions

L'attrait pour les modèles virtuels 3D vient du fait qu'ils permettent une meilleure compréhension des relations spatiales et sont par conséquent idéaux à des fins de démonstrations ou de présentations, comme le montre le projet allemand GEIST

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[97]. Le système développé dans GEIST informe les personnes sur un site historique en offrant une visite personnalisée et animée sans guide touristique humain. Lorsque les utilisateurs s'approchent d'un endroit qui a un lien historique avec une personne ou un fait, ils peuvent interagir avec un personnage virtuel qui raconte les faits. A l'aide d'un système de lunette, l'utilisateur peut voir les objets réels auxquels ont été superposés des éléments virtuels (par exemple des bâtiments) montrant la situation historique.

Au cours des quinze dernières années, différentes approches ont été proposées pour la reconstruction 3D de sites urbains [106]. Ces approches sont très variées, tant sur le plan des données nécessaires en entrée (cf Figure II.1) que sur celui des types de zones supportés (par exemple : relief plat ou accidenté) et du degré d'automatisation de la méthode. Les méthodes dépendent aussi du degré de précision désirée. Parfois, on peut se contenter d'une reconstruction grossière alors que dans d'autres cas, on veut reproduire plus fidèlement un petit secteur.

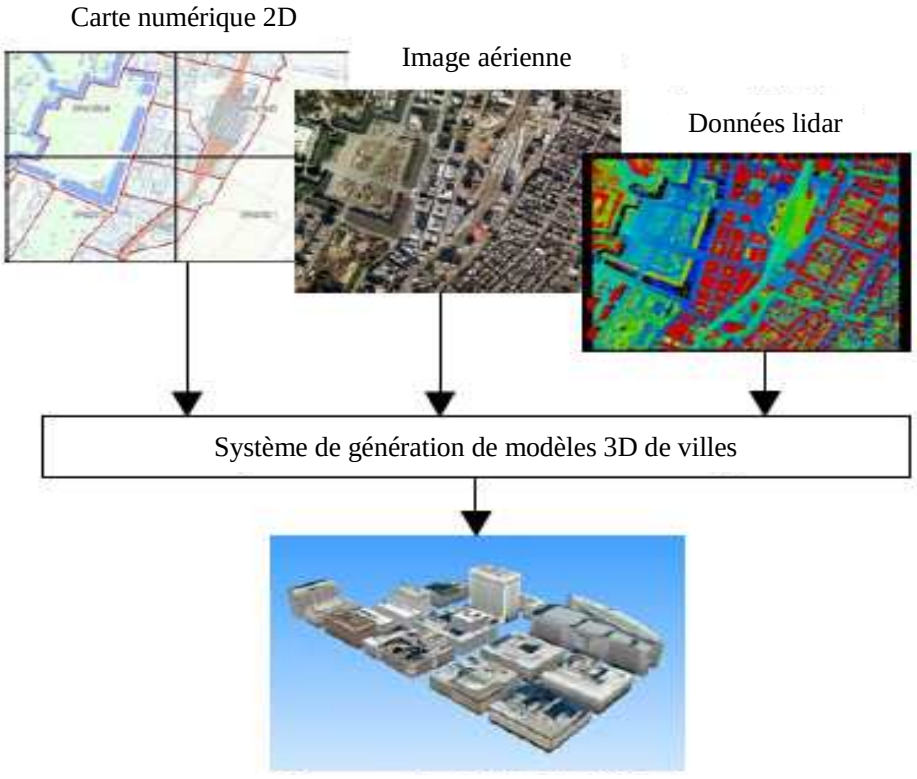


FIGURE II.1 – Génération de modèles virtuels 3D urbains

Dans la majorité des approches développées, le principe de base consiste généralement à diviser le problème de reconstruction en plusieurs sous-problèmes.

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- Localiser les bâtiments sur une carte ;
- Calculer le relief du terrain ;
- Trouver la hauteur des bâtiments ;
- Trouver la forme des bâtiments (principalement les toitures) ;
- Augmenter le réalisme de la scène (ex. : utilisation de textures).

Les données les plus utilisées dans la reconstruction virtuelle en 3D de zones urbaines sont :

- des plans de cadastre
- des données aériennes (photographies, images radar ...)
- des images satellitaires
- des photos / vidéos prises au sol
- des mesures de positionnement de type DGPS
- des modèles numériques de terrain (MNT)
- des modèles numériques d'élévation (MNE)
- des cartes numériques 2D
- des données lidar
- ...

Dans le chapitre 1 de sa thèse, Sourimant [101] dresse un état de l'art de méthodes de reconstruction de zones urbaines. Les méthodes de reconstruction avec des données aériennes [41, 64, 103, 104, 116] permettent de couvrir une large zone géographique. Cependant, ce point de vue macroscopique des zones urbaines limite les possibilités de mesure des textures des façades. Ces approches comprennent généralement deux étapes : la détection des bâtiments et le calcul de la hauteur des bâtiments par stéréovision. Les méthodes de reconstruction avec des données au sol [28, 29, 34, 98, 108, 109, 109, 113, 115, 122, 121] ont l'avantage de donner accès aux informations photométriques (textures des bâtiments, etc) et géométriques (position et forme des portes, fenêtres, etc) de façon précise. Cependant cette seconde approche est très localisée. Reconstruire entièrement une ville de cette façon requiert donc énormément de travail du point de vue acquisition et traitement des données. En vue d'obtenir des modèles 3D à grande échelle mais localement précis, des approches hybrides [46, 47, 48, 64] de reconstruction sont souvent mises en œuvre. Elles combinent à la fois des données acquises au sol et des données aériennes.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

On trouve sur le marché différents logiciels aidant à la création de modèles 3D. Parmi ceux-ci, citons : Maya, 3D Studio Max, Blender3D.

Par ailleurs, notons que la norme CityGML semble rencontrer un consensus et devrait vraisemblablement être adoptée comme norme officielle dans les mois à venir. L'objectif de cette norme est de définir un format lisible pour stocker et représenter les villes à trois dimensions [78]. La norme CityGML (http://www.citygml.org/) a été lancée par un groupe d'Allemands (Geodata Infrastructure North-Rhine Westphalia). Cette initiative réunit plus de 70 entreprises, municipalités et laboratoires de recherche qui coopèrent au développement et à l'exploitation commerciale de modèles 3D interopérables. La norme CityGML définit les classes et les relations des objets trouvés dans les villes de plusieurs points de vue, géométrique, topologique, sémantique et d'apparence. Le format CityGML est un modèle de données ouvert construit sur le format XML permettant le stockage et l'échange de modèles virtuels 3D urbain.

Une illustration de l'essor des modèles virtuels 3D est la modélisation de toutes les grandes villes japonaises en trois dimensions depuis 2002 ; modélisation mise à jour tous les 6 mois. Par ailleurs le projet Bati3D de l'IGN (Institut Géographique National) a pour objectif de développer d'ici 2010 une maquette numérique urbaine 3D du territoire français.

II.1.2 Le modèle Tecnomade de Nancy

Le modèle virtuel 3D utilisé pour nos expérimentations a été développé par la société Tecnomade (http://www.tecnomade.fr). Celle-ci a été créée en 2003 et provient du programme de transfert technologique de l'INRIA Lorraine sur la visualisation interactive d'objets par Internet. Elle fournit des outils d'aide à la décision et à la communication de projets d'architecture et d'urbanisme aux collectivités locales. Elle réalise des maquettes numériques 3D réalistes de morceaux de villes, manipulables en 3D temps-réel.

L'efficacité de ce modèle repose sur la technologie « temps-réel » interactive en 3D (représentation de la hauteur, de la largeur et de la profondeur dans l'espace pour augmenter le niveau de réalisme virtuel), qui permet de visualiser instantanément n'importe quel point de vue à la demande (vue aérienne, automobile ou piéton) et de se déplacer de manière fluide dans toutes les directions. Par inter-activité, on entend le fait de pouvoir se déplacer librement dans le modèle 3D, dans n'importe

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

quelle direction pour visualiser immédiatement telle ou telle partie à la demande de l'utilisateur, à l'inverse des animations pré-calculées, qui défilent sans intervention possible.

Le modèle 3D mis à notre disposition cartographie le centre-ville de Nancy, et plus particulièrement la place Stanislas. La Figure II.2 propose quelques vues de ce modèle.

FIGURE II.2 – Quelques Vues du modèle 3D de la place Stanislas à Nancy

La base de données du modèle 3D Tecnomade utilisé a une taille de 60 Mo et est constituée de plusieurs éléments :

- Un répertoire intitulé *Textures* qui contient un ensemble d'images au format JPEG ou PNG, ce sont les textures utilisées pour l'affichage du modèle 3D (Figure II.3).
- Un fichier XML intitulé *Geometrie.xml* qui contient la définition des polygones et des formes géométriques constituant la scène 3D. Pour chaque élément géométrique, il est indiqué son nom, son type et les informations concernant les différents vertices qui constituent cet élément, c'est-à-dire les coordonnées du vertex, les coordonnées de la normale à la surface formée par ce vertex, et les coordonnées de la texture en ce vertex (Figure II.4).

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- Un fichier XML intitulé *SceneGraph.xml* qui contient la constitution des nœuds du graphe de scène. Pour cela, les objets géométriques définis dans le fichier *Geometrie.xml* sont utilisés et les différentes transformations de ces objets et des nœuds du graphe sont indiquées (Figure II.5).
- Un fichier XML intitulé *PlaceStan.xml* qui contient des informations sur le modèle 3D telles que le numéro de la version et la date de création ; il inclut les deux autres fichiers XML (*Geometrie.xml* et *SceneGraph.xml*).

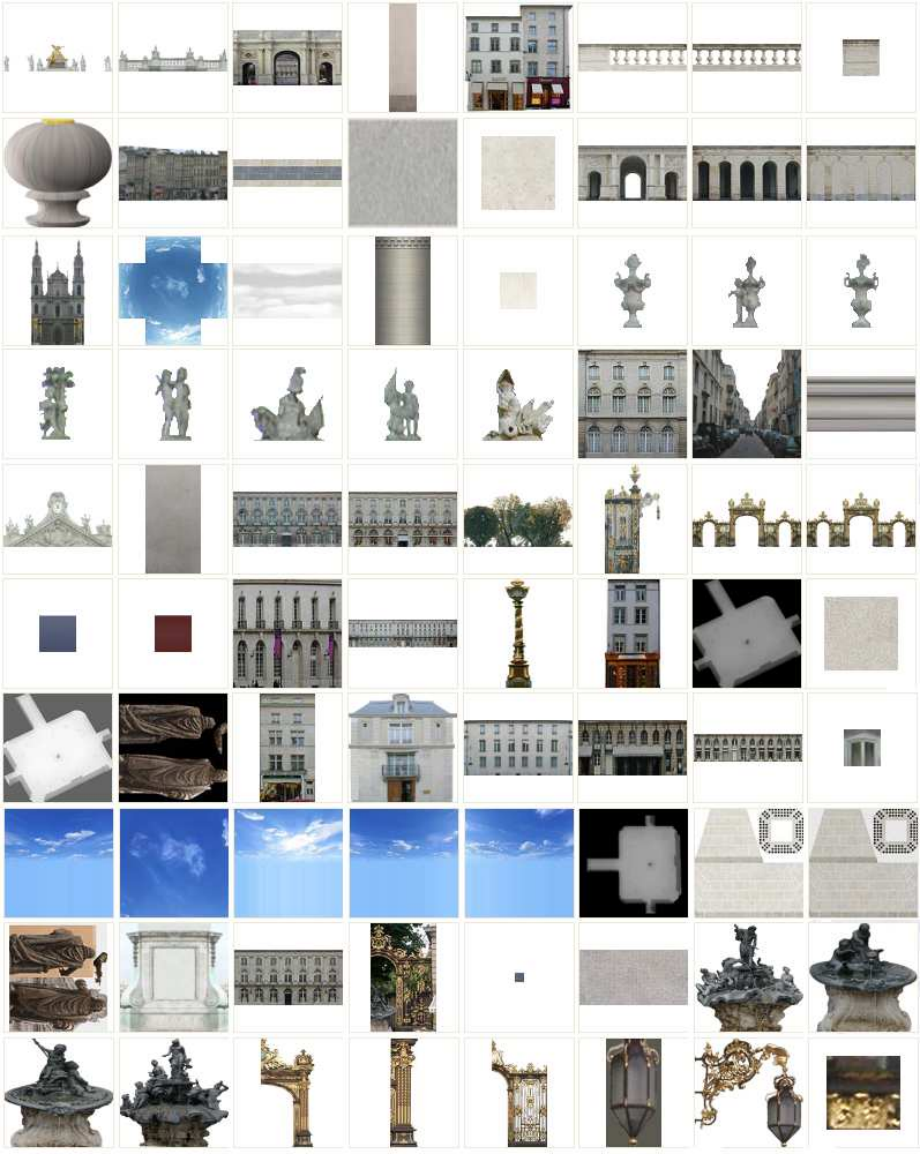


FIGURE II.3 – Quelques textures du modèle 3D de la place Stanislas à Nancy

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

```
<Definitions>
 <Geometrie Nom="Geom_Line41_630" Type="triangles_bruts">
 </Geometrie>
 <Geometrie Nom="Geom_Rectangle42_631" Type="triangles_bruts">
  <Point3D Vertex="-59.1352 -22.416 17.7403" Normale="0.448328 -0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-56.4535 -21.0709 17.7403" Normale="0.448328 -0.893869 0.0" Tex1="0.995809 0.000499547"/>
  <Point3D Vertex="-56.4535 -21.0709 19.4422" Normale="0.448328 -0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-59.1352 -22.416 17.7403" Normale="0.448328 -0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-56.4535 -21.0709 19.4422" Normale="0.448328 -0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-59.1352 -22.416 19.4422" Normale="0.448328 -0.893869 0.0" Tex1="0.299503 0.9995"/>
  <Point3D Vertex="-62.6214 -24.1645 17.7403" Normale="0.448328 -0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-59.9398 -22.8195 17.7403" Normale="0.448328 -0.893869 0.0" Tex1="0.995809 0.000499547"/>
  <Point3D Vertex="-59.9398 -22.8195 19.4422" Normale="0.448328 -0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-72.8053 2.83156 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-78.9731 -0.261995 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-76.2915 1.08303 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-76.2915 1.08303 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.000499547"/>
  <Point3D Vertex="-78.9731 -0.261995 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-78.9731 -0.261995 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.9995"/>
  <Point3D Vertex="-76.2915 1.08303 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-82.459 -2.01035 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-79.7773 -0.665323 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.9995"/>
  <Point3D Vertex="-79.7773 -0.665323 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.000499547"/>
  <Point3D Vertex="-82.459 -2.01035 17.7403" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.000499964"/>
  <Point3D Vertex="-82.459 -2.01035 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.299503 0.9995"/>
  <Point3D Vertex="-79.7773 -0.665323 19.4422" Normale="-0.448328 0.893869 0.0" Tex1="0.995809 0.9995"/>
    ...
 </Geometrie>
 <Geometrie Nom="Geom_Line42_632" Type="triangles_bruts">
  <Point3D Vertex="-56.1222 -19.5385 6.56499" Normale="-0.448328 0.893869 0.0" Tex1="-0.0165866 0.385341"/>
  <Point3D Vertex="-56.1222 -19.5385 7.43999" Normale="-0.448328 0.893869 0.0" Tex1="-0.0124131 0.385341"/>
    ...
 </Geometrie>
   ...
</Definitions>
```

FIGURE II.4 – Extrait du fichier *Geometrie.xml* du modèle 3D de la place Stanislas

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

```
<SceneGraph>
  <Objet Nom="Line41">
    <InsererGeometrie Ref="Geom_Line41_630"/>
  </Objet>
  <Objet Nom="Rectangle42">
    <Culling Etat="0"/>
    <Texture Canal="1" Fichier="balustre3.png" TexMode="modulate"/>
    <TrierZ/>
    <Materiel Ambiante="1.0 1.0 1.0" Diffuse="1.0 1.0 1.0" Speculaire="1.0 1.0 1.0" Brillance="0.0"/>
    <Transparence Facteur="1.0"/>
    <InsererGeometrie Ref="Geom_Rectangle42_631"/>
  </Objet>
  <Objet Nom="Line42">
    <Texture Canal="1" Fichier="musee2.JPG" TexMode="modulate"/>
    <Materiel Ambiante="1.0 1.0 1.0" Diffuse="1.0 1.0 1.0" Speculaire="1.0 1.0 1.0" Brillance="0.0"/>
    <Transparence Facteur="1.0"/>
    <InsererGeometrie Ref="Geom_Line42_632"/>
  </Objet>
  <Objet Nom="Plane122">
    <Culling Etat="0"/>
    <Texture Canal="1" Fichier="Deco06.PNG" TexMode="modulate"/>
    <TrierZ/>
    <Materiel Ambiante="1.0 1.0 1.0" Diffuse="1.0 1.0 1.0" Speculaire="1.0 1.0 1.0" Brillance="0.0"/>
    <Transparence Facteur="1.0"/>
    <InsererGeometrie Ref="Geom_Plane122_656"/>
  </Objet>
  <Objet Nom="Box06">
    <Texture Canal="1" Fichier="Gris2.jpg" TexMode="modulate"/>
    <Materiel Ambiante="1.0 1.0 1.0" Diffuse="1.0 1.0 1.0" Speculaire="0.9 0.9 0.9" Brillance="0.25"/>
    <Transparence Facteur="1.0"/>
    <InsererGeometrie Ref="Geom_Box06_657"/>
  </Objet>

  ...
</SceneGraph>
```

FIGURE II.5 – Extrait du fichier *SceneGraph.xml* du modèle 3D de la place Stanislas

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

II.2 Système d'information géographique 3D

II.2.1 Introduction

Afin de manipuler en temps-réel la base de données du modèle virtuel 3D, un système d'information géographique trois dimensions (SIG-3D) est requis. Un SIG est un système informatique capable d'intégrer, d'enregistrer, d'éditer, d'analyser, de partager et d'afficher des informations géo-référencées.

Les trente dernières années, de nombreuses nouvelles technologies sont apparues pour le développement des systèmes d'information géographique (SIG). Un SIG classique fournit uniquement une représentation 2D des entités spatiales et utilise des primitives simples telles que des points, des lignes et des polygones. Au début des années 1980, des technologies de visualisation et d'analyse 2D de terrains ont été introduites. Du milieu des années 1980 à aujourd'hui, des technologies relatives aux SIG tri-dimensionnels ont été implémentées et ont permis le développement de la visualisation et de l'analyse de terrains en trois dimensions. L'évolution des technologies des SIG est illustrée sur la Figure II.6 [75].

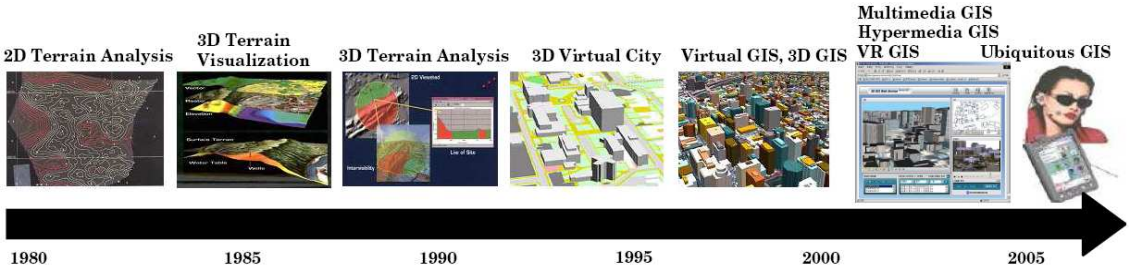


FIGURE II.6 – Evolution des systèmes d'information géographique

Le SIG-3D développé se base sur l'utilisation d'un moteur 3D. Un moteur 3D est un logiciel qui permet la représentation d'un objet dans un espace tridimensionnel virtuel et de l'afficher, entre autres, sur l'écran d'un ordinateur. Un moteur 3D est une bibliothèque dont les principales fonctions permettent :

- de gérer des objets dans un espace tridimensionnel virtuel, appelé la scène ;
- d'afficher à l'écran cette scène à partir d'un point de vue spécifié.

L'affichage, qui est une projection de la scène, est appelé le rendu. Plusieurs méthodes de rendu 3D existent, parmi lesquelles :

- le **lancer de rayons** : la principe est de simuler le parcours inverse de la lumière de la scène vers l'oeil. Pour chaque pixel de l'image à afficher, le moteur

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

tire un rayon partant de la caméra, calcule ce qui lui arrive (intersection avec les surfaces d'objets, réflexions, réfractions ...) et affiche le pixel de l'image avec la couleur ainsi déterminée en fonction des sources de lumière, des matériaux ...

- le **rendu de polygones** : les objets sont représentés sous forme de polygones (des triangles dans notre cas) auxquels le moteur 3D fait subir différentes transformations géométriques (rotation et translation dans l'espace puis projection) puis il les synthétise en fonction de l'éclairement, de la couleur ou texture.

Le moteur 3D manipule les objets stockés sous forme de maillage (*mesh*), c'est-à-dire d'ensembles de polygones, souvent triangulaires. Chaque triangle est formé de 3 points, appelés vertex. Les vertex sont représentés par :

- les trois composantes X, Y et Z, représentant les coordonnées du vertex dans l'espace tridimensionnel ;
- les trois composantes V_x, V_y et V_z du vecteur normal permettant de calculer la réflexion de la lumière ;
- une couleur.

Les transformations géométriques permettent de connaître l'emplacement de chacun des points du maillage à l'écran. On obtient alors un affichage « fil de fer » reliant uniquement les sommets entre eux. Afin de « remplir » ce maillage, il faut procéder à l'ombrage des faces (*shading*). L'ombrage des faces permet :

- de prendre en compte le modèle d'illumination ;
- de déterminer les faces affichées afin de ne pas écraser une face visible par une face non-visible ;
- de déterminer les paramètres de visualisation de chacune des faces : couleur, texture ...

Pour accéder aux services offerts par un moteur 3D, on utilise généralement une interface de programmation ou API. Le moteur 3D fourni par Tecnomade est basé sur l'API OpenGL initié par Silicon Graphics (SGI), utilisable sur un grand nombre d'architectures.

II.2.2 Le SIG-3D développé

Le moteur 3D permet donc de naviguer dans le modèle 3D à l'aide de la souris et du clavier et de visualiser à l'écran la scène. Pour nos applications, nous avons besoin d'autres fonctions que nous avons ajoutées au moteur 3D pour créer un outil

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

que nous avons qualifié de système d'information géographique en trois dimensions, SIG-3D (Figure II.7).

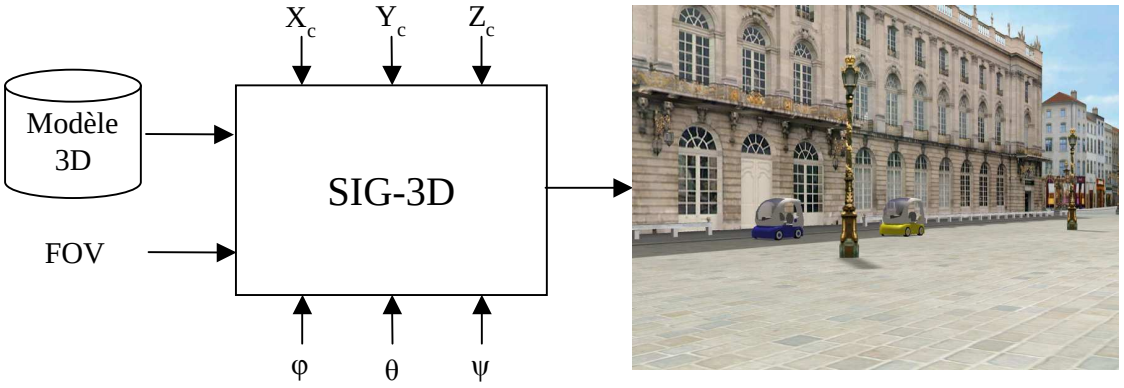


FIGURE II.7 – Système d'information géographique 3D

Les entrées du SIG-3D sont décrites dans le paragraphe II.2.2.1.

Les fonctions du SIG-3D sont le géo-référencement du modèle 3D d'une part et la possibilité de sauvegarder des données (sous forme d'une image ou d'un fichier texte ou XML) de chaque affichage de la scène à l'écran. Le géo-référencement du modèle 3D (§ II.2.2.3) permet de naviguer dans le modèle 3D à partir de coordonnées exprimées par exemple dans le repère *WGS84* du GPS. Les types de données sauvegardées, décrites dans les paragraphes II.2.2.2.a à II.2.2.2.e sont :

- l'image virtuelle (§ II.2.2.2.a)
- le fichier texte de profondeur (§ II.2.2.2.b)
- l'image de profondeur (§ II.2.2.2.c)
- le fichier XML des segments 3D (§ II.2.2.2.d)
- l'image des segments 3D visibles (§ II.2.2.2.e)

II.2.2.1 Les entrées du SIG-3D

Les entrées du SIG-3D développé sont les suivantes :

- la base de données du modèle, c'est-à-dire les vertex, les textures, les paramètres de configuration ...
- le paramètre intrinsèque de la caméra virtuelle
- les paramètres extrinsèques de la caméra virtuelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

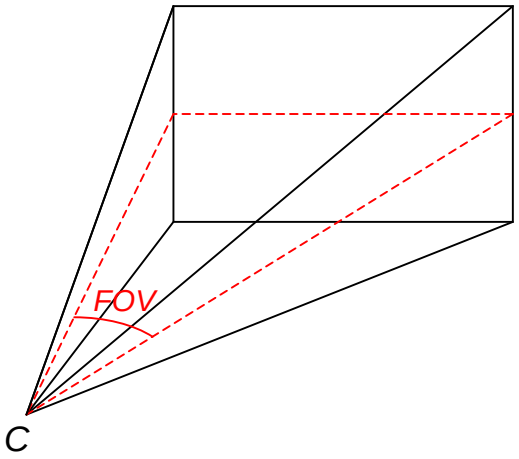


FIGURE II.8 – FOV de la caméra virtuelle

Le paramètre intrinsèque de la caméra virtuelle (Figure II.8) est l'angle de champ de vision horizontal FOV (Field Of View).

Les paramètres extrinsèques de la caméra virtuelle sont les six degrés de liberté par rapport au repère attaché au modèle 3D : trois paramètres de position (X_C, Y_C, Z_C) et trois paramètres d'orientation (ϕ, θ, ψ).

Le repère $\mathcal{R}_M(O_M, X_M, Y_M, Z_M)$ attaché au modèle 3D est défini comme suit (cf Figure II.9 :

- l'origine O_M du repère est la base de la statue Stanislas située au centre de la place,
- l'axe X_M du repère est dans la direction de l'Est,
- l'axe Y_M du repère est dans la direction du Nord,
- l'axe Z_M du repère pointe vers le haut (verticale)

Le repère du modèle 3D est donc un repère ENU (East - North - Up) centré sur la zone cartographiée.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

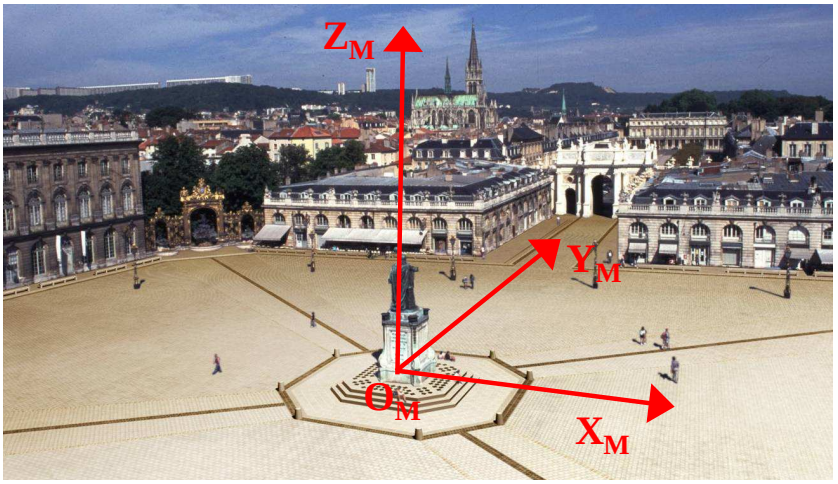


FIGURE II.9 – Repère attaché au modèle 3D

II.2.2.2 Les sorties du SIG-3D

II.2.2.2.a Image virtuelle

L'image virtuelle (Figure II.10) est une image au format Bitmap 24 bits qui représente la vue de la scène capturée par la caméra virtuelle placée et paramétrée selon les paramètres d'entrées du SIG-3D. C'est une capture du rendu (avec textures) synthétisé à l'écran par le moteur 3D.

FIGURE II.10 – Image virtuelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

II.2.2.2.b Fichier de profondeur

Le fichier de profondeur (Figure II.11) est un fichier texte contenant les distances en mètre entre la caméra virtuelle et les points 3D correspondants à chacun des pixels de l'image virtuelle, avec une précision de 24 bits minimum.

```
8.8784628e+000  8.8781366e+000 ...
8.8778124e+000  8.8779469e+000 ...
8.8776264e+000  8.8773079e+000 ...
8.8769913e+000  8.8771334e+000 ...
8.8768196e+000  8.8765087e+000 ...
8.8766565e+000  8.8763485e+000 ...
8.8760431e+000  8.8757391e+000 ...
8.8758941e+000  8.8755941e+000 ...
8.8752966e+000  8.8750000e+000 ...
8.8747053e+000  8.8748703e+000 ...
8.8745794e+000  8.8742905e+000 ...
8.8740034e+000  8.8741760e+000 ...
8.8738928e+000  8.8736115e+000 ...
8.8733320e+000  8.8730545e+000 ...
8.8732367e+000  8.8729630e+000 ...
8.8726902e+000  8.8724203e+000 ...
8.8726101e+000  8.8706694e+000 ...
8.8704271e+000  8.8701878e+000 ...
                ...
```

FIGURE II.11 – Fichier de profondeur

L'information de profondeur est extraite à partir du Z-Buffer de la carte vidéo, également appelé « *Depth Buffer* ». L'algorithme du Z-Buffer est un algorithme d'élimination des parties cachées. Il est utilisée pour la visualisation d'une scène, il permet en effet de supprimer l'affichage de tout objet ou morceau d'objet masqué par un autre objet.

L'idée directrice de la méthode du Z-Buffer est de maintenir à jour une mémoire tampon contenant la profondeur du point observé en chaque pixel de l'image ; cela de la même manière que la mémoire vidéo contient l'information « couleur » du point observé en chaque pixel. Lors du dessin d'une primitive, la profondeur du pixel à afficher est remise à jour si et seulement le point correspondant à ce pixel sur la primitive présente une profondeur inférieure à celle déjà présente dans la mémoire tampon. Cette remise à jour dans le tampon de profondeur entraîne alors une remise à jour de l'information « couleur » dans la mémoire vidéo.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

En lisant le Z-Buffer correspondant à l'image virtuelle, on a donc bien la profondeur de chacun des pixels de l'image virtuelle.

II.2.2.2.c Image de profondeur

L'image de profondeur (Figure II.12) est une image au format Bitmap 24 bits. C'est une représentation, en niveaux de gris, de la profondeur des pixels. Plus un objet de la scène est loin de la caméra, plus les pixels correspondants dans l'image sont clairs, et inversement, plus un objet de la scène est près de la caméra, plus les pixels correspondants dans l'image sont foncés.

FIGURE II.12 – Image de profondeur

II.2.2.2.d Fichier des segments 3D

Le fichier des segments 3D est un fichier XML (Figure II.13). Il contient, pour tous les segments 3D de la base de données, les coordonnées des extrémités et le statut de visibilité de ce segment (c'est-à-dire, si le segment est complètement visible, partiellement visible ou invisible par la caméra virtuelle).

La description pour chacun des segments est comme suit :

- Segment
  - Statut de visibilité
    - hors champ : non visible

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008


```
<Segments>
 <Segment>
   <Visible status="hors champ"/>
  <Point numero="1">
    <Coordonnees x="-24.276800" y="39.234901" z="-0.069327"/>
  </Point>
  <Point numero="2">
    <Coordonnees x="-41.521400" y="30.152100" z="0.006949"/>
  </Point>
 </Segment>
 <Segment>
   <Visible status="cadre"/>
   <FacteurVisibilite coef="100%"/>
  <Point numero="1">
    <Coordonnees x="36.050701" y="60.231899" z="0.749043"/>
  </Point>
  <Point numero="2">
    <Coordonnees x="36.050701" y="60.231899" z="6.049240"/>
  </Point>
 </Segment>
 <Segment>
   <Visible status="cadre partiellement"/>
   <FacteurVisibilite coef="100%"/>
  <Point numero="1">
    <Coordonnees x="31.873800" y="56.756001" z="-0.112608"/>
  </Point>
  <Point numero="2">
    <Coordonnees x="19.197201" y="60.775101" z="-0.112608"/>
  </Point>
 </Segment>
 <Segment>
   <Visible status="hors champ"/>
  <Point numero="1">
    <Coordonnees x="-22.656799" y="38.718300" z="-0.219362"/>
  </Point>
  <Point numero="2">
    <Coordonnees x="-40.925201" y="29.102501" z="-0.143087"/>
  </Point>
 </Segment>
 ...
</Segments>
```

FIGURE II.13 – Fichier XML des segments 3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

  - cadré : totalement visible
  - partiellement cadré : une portion du segment est visible
- Facteur de visibilité
  - Coefficient : 20 %, 40 %, 60 %, 80 %, 100 % (pourcentage visible du segment)
- Points (nombre : 2)
  - Numéro : 1 ou 2
  - Coordonnées : x, y, z

II.2.2.2.e Image des segments 3D visibles

L'image des segments 3D visibles (Figure II.14) est une image Bitmap des segments de la base de données qui sont visibles à partir de la position et de l'orientation de la caméra. C'est donc une représentation graphique du fichier XML des segments 3D.

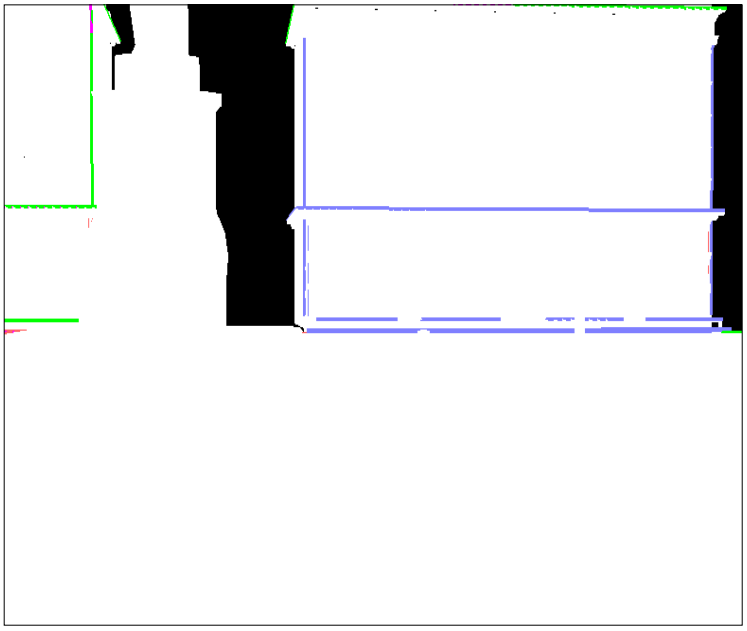


FIGURE II.14 – Images des segments 3D visibles

Trois couleurs différentes sont utilisées pour représenter les segments :

- Bleu : le segment est déterminé comme visible à 100 % et est complétement inscrit dans l'écran
- Vert : le segment est visible à 100 % mais une partie est hors de l'écran
- Rouge : le segment n'est que partiellement visible

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

II.2.2.3 Le géo-référencement du modèle 3D

Une fonction essentielle du SIG-3D est de pouvoir naviguer dans le modèle 3D à partir de positions fournies par un GPS. Il est donc nécessaire de géo-référencer le modèle 3D. Cela consiste à déterminer la transformation rigide entre le système géodésique WGS84 du GPS et le repère local attaché au modèle 3D. On peut alors suivre en temps-réel le panorama vu par un véhicule dont la trajectoire est donnée par un GPS.

Le géo-référencement du modèle 3D se fait en deux temps. Il faut d'abord transformer la pose désirée de la caméra virtuelle exprimée en coordonnées géographiques dans le système géodésique *WGS84* en coordonnées planes en projection *Lambert I*. Ceci se fait à l'aide de fonctions classiques de géodésie (cf § II.2.2.3.a). Ensuite, les coordonnées planes en projection Lambert sont converties pour obtenir la pose de la caméra dans le repère attaché au modèle 3D (cf § II.2.2.3.b).

II.2.2.3.a Transformation WGS84 vers Lambert I

Le système de référence géodésique auquel les projections *Lambert I*, *Lambert II*, *Lambert III* et *Lambert IV*, utilisées pour la France métropolitaine, sont associées s'appelle *NTF* (Nouvelle Triangulation de la France). Pour Nancy, la zone *Lambert* utilisée est *Lambert I*.

La transformation des coordonnées géographiques dans le système géodésique *WGS84* en coordonnées planes en projection *Lambert I* se fait en plusieurs étapes.

- Les coordonnées géographiques *WGS84* sont tout d'abord transformées en coordonnées cartésiennes *WGS84*
- Les coordonnées cartésiennes *WGS84* sont ensuite transformées en coordonnées cartésiennes *NTF* en appliquant trois paramètres de translation définis par l'IGN
- Les coordonnées cartésiennes *NTF* sont ensuite transformées en coordonnés géographiques *NTF*
- Les coordonnés géographiques *NTF* sont finalement projetées dans la zone Lambert appropriée, on obtient les coordonnées planes en projection Lambert dans le système *NTF*

II.2.2.3.b Transformation Lambert I vers le repère local du modèle 3D

Afin d'obtenir les coordonnées de la caméra virtuelle dans le repère local du mo-

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

dèle 3D à partir de ses coordonnées planes en projection *Lambert I*, il faut appliquer la translation suivante :

$$\begin{aligned} T_X &= 883065.147455 \\ T_Y &= 117584.647213 \\ T_Z &= 230.273523 \end{aligned} \tag{II.1}$$

Les paramètres de cette translation correspondent aux coordonnées de l'origine du repère attaché au modèle 3D. Dans notre cas, il s'agit la statue de Stanislas située au centre de la place. Les coordonnées de ce point ont été mesurées à l'aide d'un GPS RTK centimétrique.

II.3 Conclusion

L'attrait incontestable pour les modèles virtuels 3D urbains nous a amené à étudier l'apport de tels modèles dans le domaine des transports. Les modèles virtuels 3D urbains ont tout d'abord un intérêt du point de vue visualisation. En effet, ils reproduisent l'environnement de la rue « à l'identique ». Depuis son GPS, le conducteur peut ainsi visualiser les commerces et les façades des bâtiments qui l'entourent. Les volumes du bâti, les toitures et la hauteur des édifices sont respectés. Cette précision permet alors de donner des points de repères supplémentaires indéniables au conducteur, ce qui facilite la conduite et le changement de direction. En 2005, Sony a lancé le premier GPS en 3D (modèle NVXYZ777). Tele Atlas a présenté une interface de navigation GPS en vraie 3D lors du congrès 3GSM World de Barcelone. Navteq a annoncé au CeBIT de Hanovre, le lancement prochain de son système analogue prévu pour améliorer les usages de la navigation pour automobiles.

Nous pensons que les modèles virtuels 3D urbains peuvent apporter bien plus à la navigation qu'uniquement l'aspect visualisation. Par exemple, dans [85], les auteurs utilisent un modèle 3D de l'environnement pour estimer la visibilité satellitaire et le multi-trajet des ondes GPS. Dans cette thèse, nous proposons deux concepts d'utilisation de tels modèles virtuels 3D, l'un concerne la localisation des véhicules (traité dans les chapitres III et IV) et l'autre concerne la détection et la géo-localisation d'obstacles (traité dans le chapitre V). En effet, à l'aide du modèle 3D, on peut obtenir une vue virtuelle de la scène. Cette vue virtuelle peut alors être comparée à une vue réelle de la scène, par exemple acquise par une caméra embarquée à bord du véhicule. Un recalage entre les deux vues peut permettre de localiser le véhicule.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Chapitre III

Méthode de localisation proposée

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Résumé du chapitre

Ce chapitre présente la méthode développée pour la localisation de véhicules en milieu urbain. Il est composé de deux parties principales. La première traite de la construction de l'observation cartographique 3D, c'est-à-dire de l'estimation de la pose du véhicule en recalant un modèle virtuel 3D urbain sur des images acquises par une caméra embarquée. La seconde partie traite de la fusion multi-capteurs. Après avoir rappelé les équations du filtre de Kalman, du filtre de Kalman étendu et du filtre de Kalman sans parfum dans le cadre général, l'application du filtre de Kalman sans parfum à notre cas sera présentée.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sommaire

III.1 Observation cartographique 3D 72
III.1.1 Introduction 72
III.1.2 Modélisation et calibration de la caméra réelle 75
III.1.2.1 Modèle sténopé de la caméra 77
III.1.2.1.a Modèle externe 78
III.1.2.1.b Modèle interne 80
Projection perspective 80
Transformation caméra / image 81
Modèle interne complet 82
III.1.2.2 Calibration de la caméra 83
Implémentation 85
III.1.3 Paramétrage de la caméra virtuelle 85
III.1.3.1 Paramètres intrinsèques de la caméra virtuelle . 85
III.1.3.2 Paramètres extrinsèques de la caméra virtuelle . 87
III.1.4 Détection des primitives de l'image 87
III.1.4.1 Introduction 87
III.1.4.2 Points de Harris 88
III.1.4.3 Algorithme du détecteur de Harris 90
III.1.5 Mise en correspondance 2D/2D 91
III.1.6 Calcul des points 3D correspondants aux points de Harris de l'image virtuelle 94
III.1.7 Calcul de la pose de la caméra réelle à partir des appariements 2D/3D de points 97
III.1.7.1 Introduction 98
III.1.7.1.a Estimation par minimisation algébrique 98
Direct Linear Transformation 99
POSIT 99
III.1.7.1.b Estimation par minimisation géométrique 100
III.1.7.2 Méthode de calcul de la pose choisie 100
III.1.8 Incertitude associée à l'observation cartographique 3D . . 101
Améliorations envisagées 101
III.2 Fusion de données multi-capteurs 102
III.2.1 Introduction 103
III.2.1.1 Filtre de Kalman 103

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.2.1.2 Filtre de Kalman Etendu 105
III.2.1.3 Filtre de Kalman sans parfum 106
III.2.1.3.a Transformée sans parfum 106
III.2.1.3.b Algorithme du filtre de Kalman sans parfum 108
III.2.1.4 Autres méthodes de fusion de données pour la localisation de véhicules 110
III.2.1.5 Conclusion 111
III.2.2 Méthode de fusion multi-capteurs choisie pour la localisation de véhicules 111
III.2.2.1 Etape de prédiction : modèle d'évolution odométrique 111
III.2.2.2 Etape de mise à jour : observation GPS 114
III.3 Conclusion 116

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Le principe général de la méthode de localisation développée est décrit dans la Figure III.1. Les odomètres et le gyromètre dont est équipé le véhicule expérimental, permetent d'estimer la position relative du véhicule. Cette position relative est recalée à l'aide de mesures GPS, lorsqu'elles sont disponibles et que leur qualité est suffisante. Lorsque les mesures GPS ne peuvent pas être utilisées, une autre estimation de la pose absolue est utilisée afin de recaler la position relative : l'observation cartographique 3D. Cette observation cartographique 3D est basée sur la mise en correspondance entre des images vidéos acquises par une caméra embarquée à bord du véhicule et le modèle virtuel 3D urbain décrit dans le chapitre précédent (II).

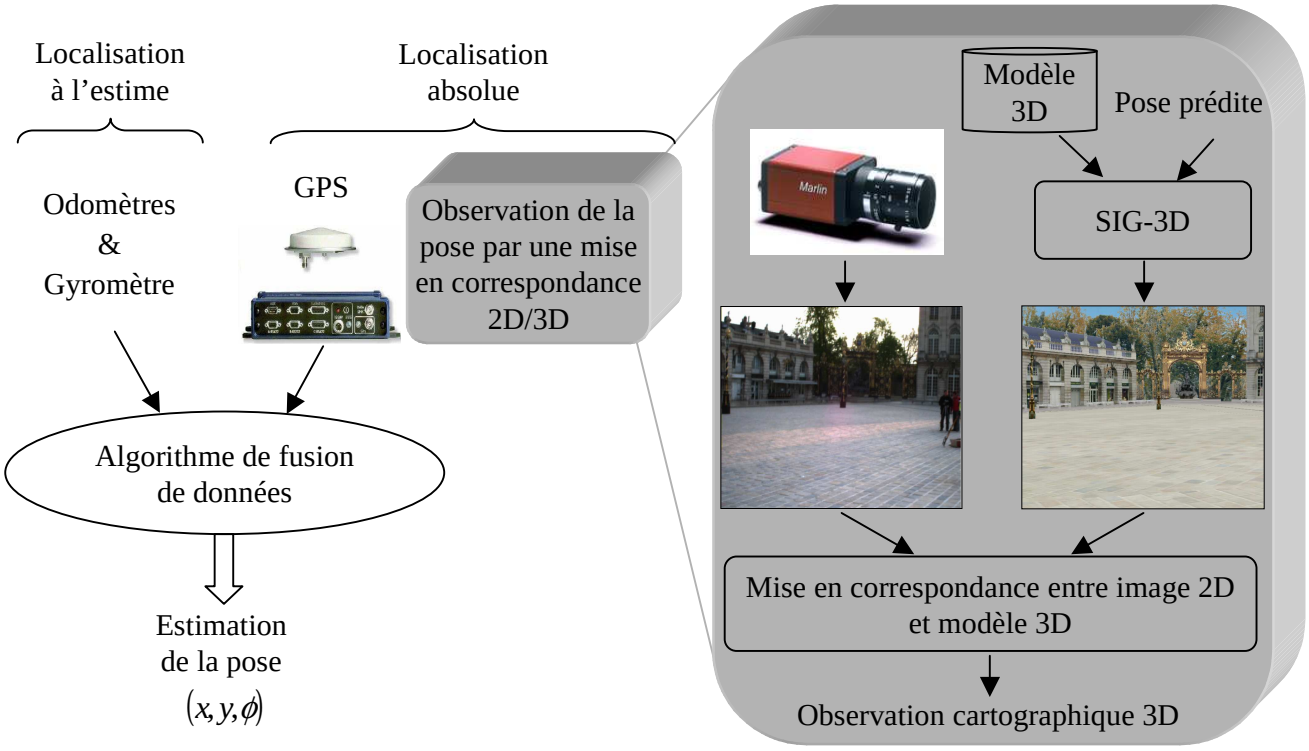


FIGURE III.1 – Principe général de la méthode de localisation

Ce chapitre se décompose en deux parties. Dans un premier temps, nous présentons la méthode développée pour construire l'observation cartographique 3D, c'est-à-dire l'estimation de la pose à partir du modèle 3D et de la vision (cf § III.1). Dans un deuxième temps, nous présentons l'algorithme de fusion de données mis en œuvre et basé sur un filtre de Kalman sans parfum (cf § III.2).

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.1 Observation cartographique 3D

III.1.1 Introduction

L'observation cartographique 3D est l'estimation de la pose du véhicule en recalant l'image courante acquise par une caméra embarquée à bord du véhicule sur le modèle 3D de la scène.

Le principe de la méthode de construction de l'observation cartographique est décrite dans la figure III.2.

L'idée est de calculer la pose de la caméra embarquée à partir d'appariements de primitives extraites dans 2 images, l'image courante acquise par la caméra embarquée, appelée « image réelle » d'une part et l'image extraite à l'aide d'un SIG-3D, appelée « image virtuelle » d'autre part.

La première étape de l'algorithme concerne donc l'obtention de ces deux images. L'image réelle est l'acquisition faite par la caméra réelle au préalable calibrée (cf § III.1.2 pour la calibration de la caméra réelle). Afin d'extraire l'image virtuelle correspondante, le SIG-3D doit paramétrer la caméra virtuelle (cf § III.1.3). Le SIG-3D a donc besoin des paramètres intrinsèques et extrinsèques de la caméra virtuelle à configurer. Le paramètre intrinsèque de la caméra virtuelle est le FOV (cf § II.2.2.1). La détermination du FOV de la caméra virtuelle à partir des paramètres intrinsèques de la caméra réelle est décrit dans le paragraphe III.1.3.1. Les paramètres extrinsèques de la caméra virtuelle, c'est-à-dire sa position et son orientation, sont déterminés à partir de la pose *a priori*. La pose *a priori* est la pose estimée dans l'étape de prédiction de l'algorithme de fusion (cf § III.2.2.1) puis convertie dans le repère du modèle 3D.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

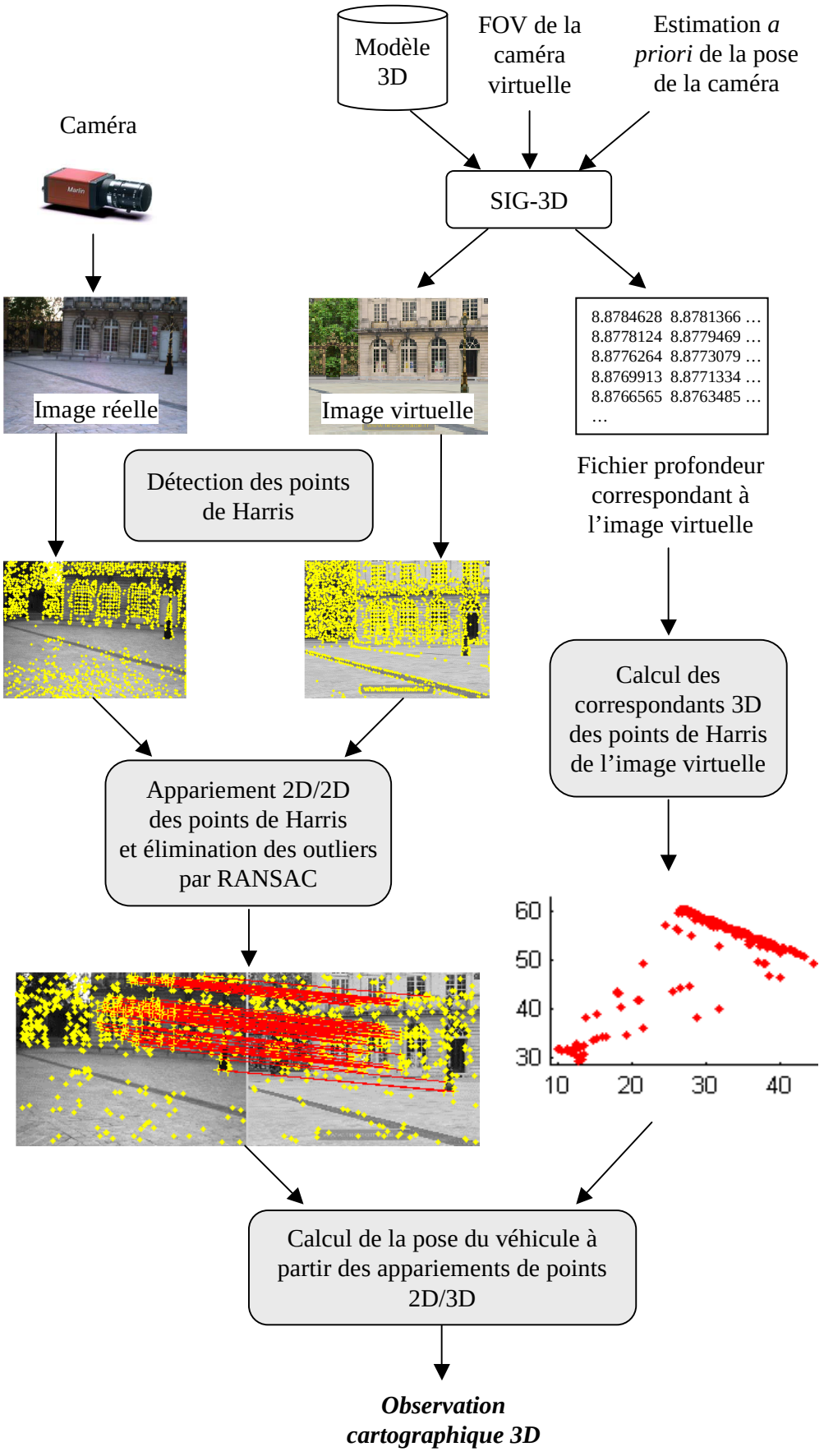


FIGURE III.2 – Synoptique de la construction de l'observation cartographique 3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Dès lors que l'image réelle et l'image virtuelle sont à disposition, on en extrait les points de Harris (cf Figure III.3a et § III.1.4). Puis, l'ensemble des points de Harris de l'image réelle est apparié avec l'ensemble des points de Harris de l'image virtuelle (cf Figure III.3b et § III.1.5). Ensuite, les coordonnées des correspondants 3D des points de Harris appariés de l'image virtuelle sont calculées (cf Figure III.3c et § III.1.6). Les correspondants 3D sont les points 3D dont les projetés sont les pixels de l'image virtuelle. A partir des appariements 2D/3D ainsi établis, on peut calculer la pose de la caméra réelle (cf Figure III.3d et § III.1.7). On propose finalement dans le paragraphe III.1.8 quelques réflexions pour l'estimation de l'erreur de l'observation cartographique 3D.

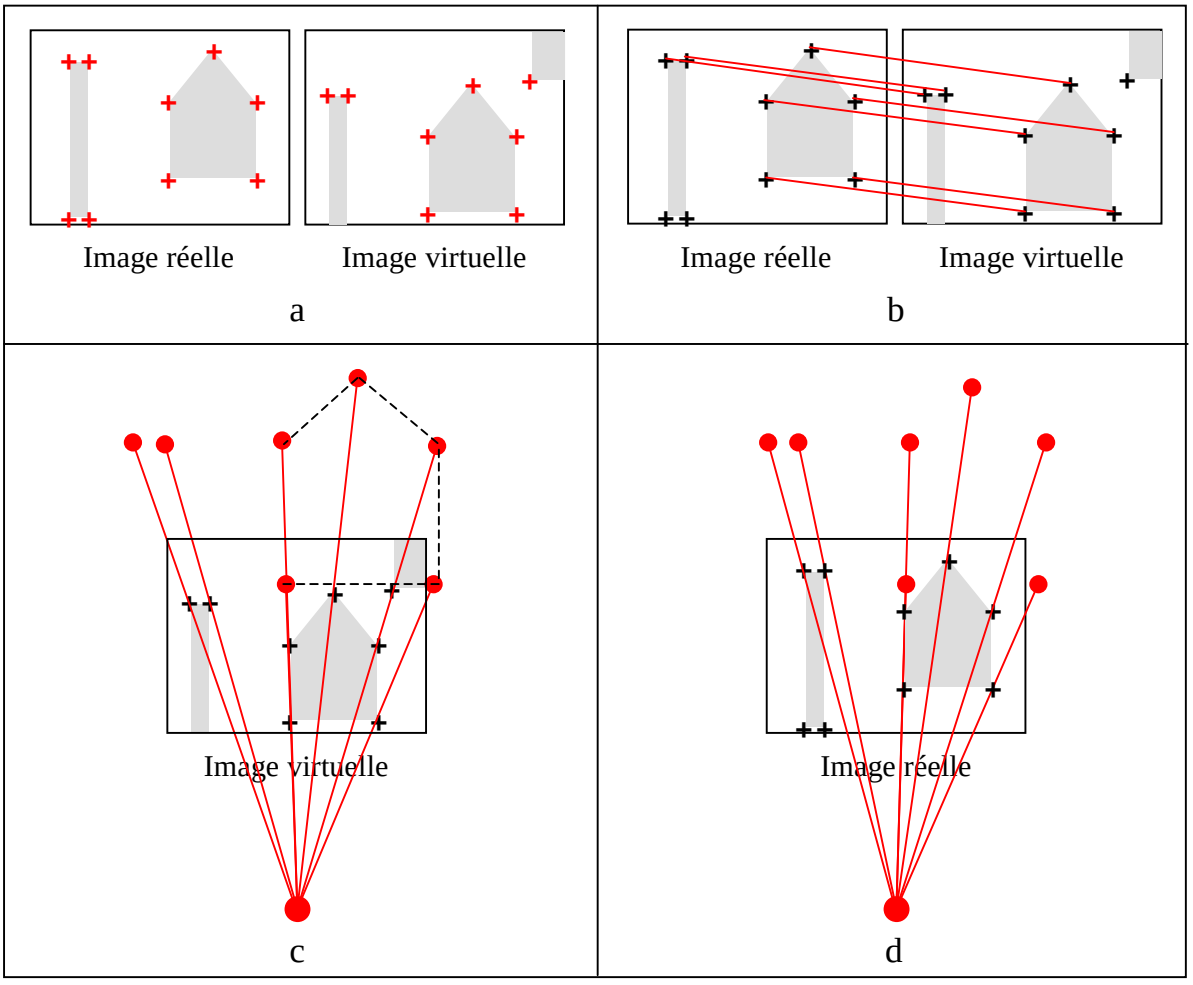


FIGURE III.3 – (a) Détection des points de Harris de l'image réelle et de l'image virtuelle - (b) Appariement des points de Harris - (c) Calcul des correspondants 3D - (d) Calcul de la pose de la caméra réelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Comme indiqué précédemment et résumé sur la Figure III.4, chaque étape de la construction de l'observation cartographique est détaillée dans les paragraphes III.1.2 à III.1.8.

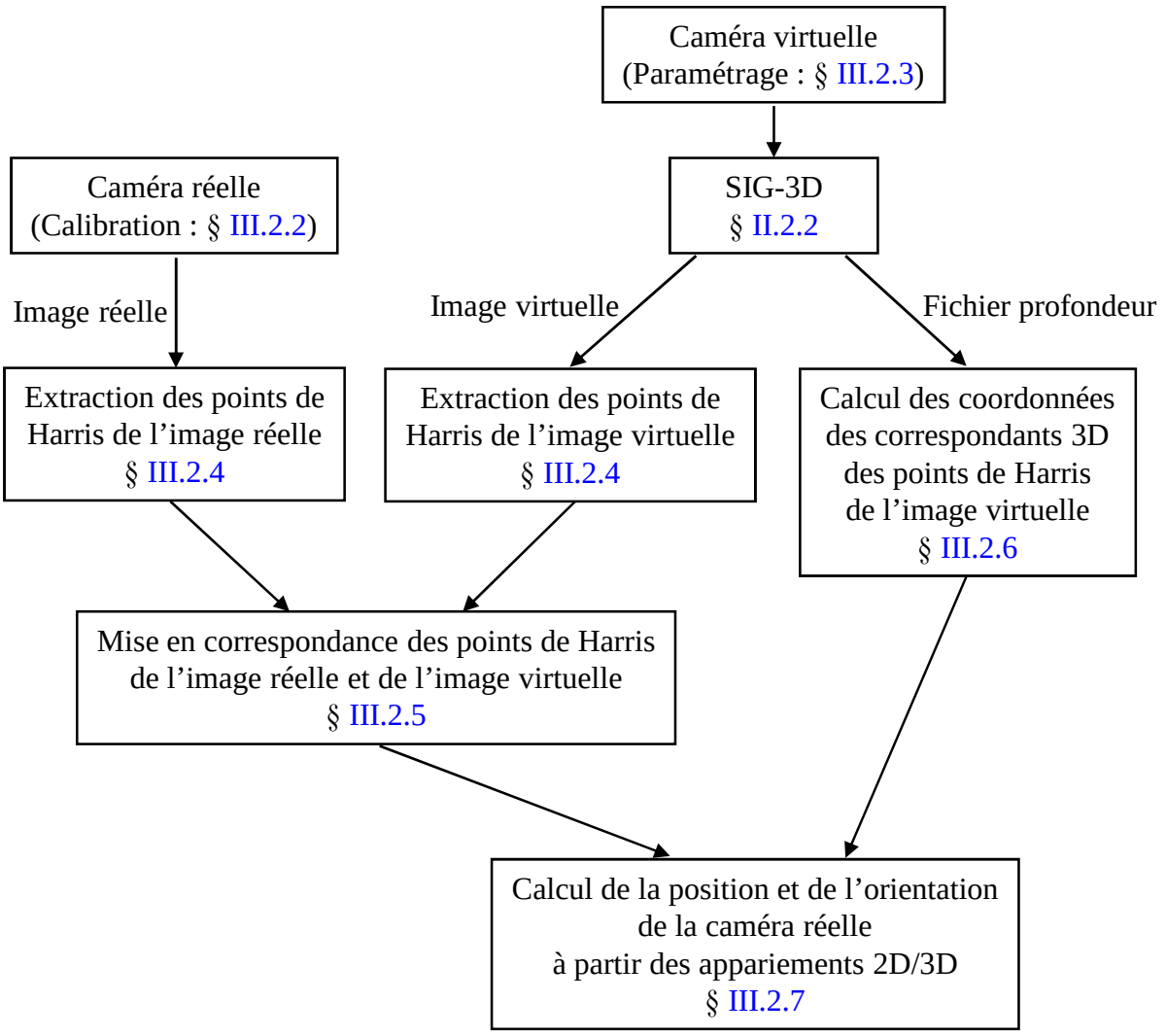


FIGURE III.4 – Construction de l'observation cartographique 3D

III.1.2 Modélisation et calibration de la caméra réelle

Le modèle mathématique d'une caméra est l'ensemble des lois géométriques décrivant la façon dont un point de l'espace à trois dimensions se projette sur le plan image à deux dimensions. Il existe différents modèles permettant de représenter géométriquement une caméra [42, 65]. Le modèle sténopé ou projection centrale est le modèle de caméra le plus couramment utilisé. Le modèle sténopé consiste à simplifier l'ensemble des lentilles qui composent le système optique par un point où convergent

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

tous les rayons lumineux pour aller se projeter sur le plan image. L'image d'un objet est alors obtenue par une projection perspective. C'est ce modèle que nous allons utiliser pour écrire les équations qui définissent la projection d'un point de l'espace sur l'image.

Un modèle est caractérisé par un certain nombre de paramètres permettant de calculer, à partir des coordonnées cartésiennes d'un point de l'espace, les coordonnées du pixel qui lui correspond sur l'image acquise par la caméra. On distingue les paramètres intrinsèques et les paramètres extrinsèques d'une caméra. Les paramètres intrinsèques d'une caméra modélisent ces caractéristiques internes et ne dépendent donc pas de sa position ni de son orientation dans l'espace. Ce sont par exemple la distance focale, la taille des pixels, la position de l'axe de visée dans l'image. Les paramètres extrinsèques, quant à eux, relient le système de coordonnées de la caméra au système de coordonnées du monde. Ils représentent donc la position et l'orientation de la caméra par rapport à un repère fixe. La projection d'un point de l'espace sur une image dépend donc de sa localisation par rapport à la caméra d'une part et des caractéristiques physiques du canal d'acquisition (objectif de la caméra) d'autre part.

Dès lors qu'un type de modèle est choisi, il faut identifier ses paramètres. L'estimation de la valeur de ces paramètres pour une caméra est réalisée par calibrage. C'est une étape nécessaire pour toute application orientée vision. De nombreuses techniques ont été développées. Elles peuvent être classées en deux catégories :

- Calibrage basé sur des objets 3D de référence
  Cette technique utilise l'observation d'objets 3D de coordonnées connues. Les objets de calibrage (mire) sont généralement des points répartis sur des plans orthogonaux ou sur un plan translaté dans la direction de sa normale. Le calcul peut alors être effectué de façon relativement simple [90, 65, 125, 5].
- Calibrage automatique (ou auto-calibrage)
  Le mouvement connu de la caméra filmant une scène statique est utilisé pour poser des contraintes sur les paramètres intrinsèques prenant en compte la rigidité des objets filmés en utilisant uniquement les informations de l'image [38, 61]. Bien que très flexible cette deuxième approche n'est pas encore assez développée pour obtenir des résultats satisfaisants.

Dans ce qui suit, le modèle de projection sténopé est détaillé ainsi que la méthode de calibrage de la caméra réelle.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.1.2.1 Modèle sténopé de la caméra

Soit un point O du plan image appelé point principal. La normale au plan image passant par O est appelé axe optique ou axe principal. Le point situé sur l'axe optique à une distance f du plan image est noté C. Le point C est le centre de projection et f est la distance focale. On peut placer le centre de projection devant ou derrière le plan image. Si on place le centre de projection devant le plan image, on a alors une image inversée. Dans ce qui suit, nous allons placer le centre de projection derrière le plan image, comme sur la Figure III.5

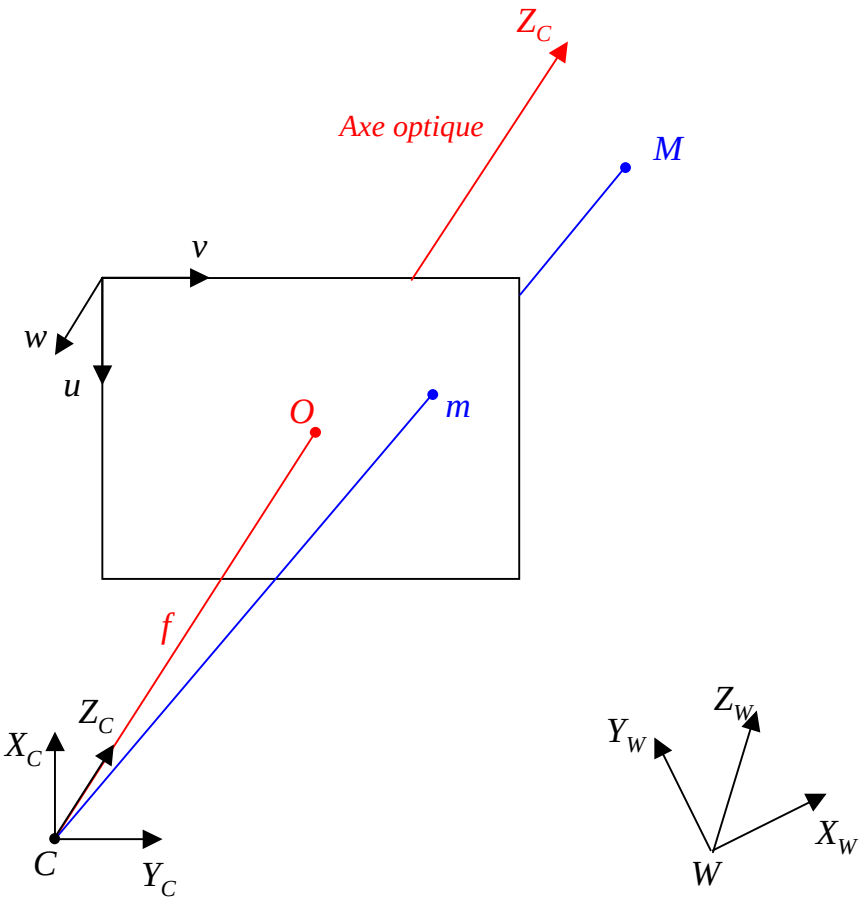


FIGURE III.5 – Modèle sténopé d'une caméra

Tout point M de l'espace admet une projection m sur le plan image par l'intersection du plan image avec le rayon optique CM.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Soient les systèmes de coordonnées suivants :

- Le repère monde ou repère objet est un repère orthonormé de l'espace et est noté $\mathcal{R}_W(W, X_W, Y_W, Z_W)$. Les coordonnées d'un point M dans ce repère sont notés $M^W = (x_M^W, y_M^W, z_M^W)^T$
- Le système de coordonnées attaché à la caméra est noté $\mathcal{R}_C(C, X_C, Y_C, Z_C)$. Le plan $X_C - Y_C$ de ce repère est parallèle au plan image et l'axe Z_C est confondu avec l'axe optique. L'origine C de ce repère est le centre de projection. Les coordonnées d'un point M dans ce repère sont notés $M^C = (x_M^C, y_M^C, z_M^C)^T$
- Le repère image est noté $\mathcal{R}_I(O, u, v, w)$. Dans ce repère, les coordonnées sont exprimées en pixels. L'équation du plan $(u - v)$ de l'image dans le repère de la caméra est $z = f$ où f est la distance focale de la caméra.

La projection d'un point de l'espace M exprimé dans le repère $\mathcal{R}_W$ sur le plan image en un point m se décompose en trois parties distinctes :

- un changement de repère qui vise à exprimer le point M dans le repère $\mathcal{R}_C$,
- une projection perspective qui transforme le point M en m exprimé dans $\mathcal{R}_C$,
- une mise à l'échelle qui transforme les coordonnées métriques exprimées dans le repère lié à la caméra $\mathcal{R}_C$ à des coordonnées pixelliques dans le repère lié à l'image $\mathcal{R}_I$.

Le changement de repère constitue le modèle externe et est fonction des paramètres extrinsèques de la caméra, alors que la projection perspective et la mise à l'échelle constituent le modèle interne, fonction des paramètres intrinsèques.

III.1.2.1.a Modèle externe

Le changement de repère est une notion fondamentale en robotique, qui permet d'exprimer un même point M de l'espace dans plusieurs repères différents. Un point M de la scène est défini par ses coordonnées $M^W = (x_M^W, y_M^W, z_M^W)^T$ dans le repère monde $\mathcal{R}_W$ et est également défini par ses coordonnées $M^C = (x_M^C, y_M^C, z_M^C)^T$ dans le repère caméra $\mathcal{R}_C$. Les coordonnées de M dans le repère $\mathcal{R}_C$ sont obtenus à partir des coordonnées de M dans $\mathcal{R}_W$ en introduisant les paramètres extrinsèques de la caméra. Le changement de repère est représenté par un vecteur de translation T et une matrice de rotation R, les composantes de R et de T étant les paramètres extrinsèques.

Le vecteur de translation $T = (t_x, t_y, t_z)^T$ représente la position de la caméra dans le repère monde $\mathcal{R}_W$.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

La matrice de rotation R du repère $\mathcal{R}_W$ vers le repère $\mathcal{R}_C$ exprime l'orientation de la caméra. R est une matrice orthogonale 3×3 dont les colonnes sont les coordonnées des vecteurs (X_W, Y_W, Z_W) de $\mathcal{R}_W$ exprimées dans le repère $\mathcal{R}_C$. La matrice R est formée de trois rotations distinctes qui décrivent le roulis (rotation $R_{x,\alpha}$ d'angle α autour de l'axe X_C), le tangage (rotation $R_{y,\beta}$ d'angle β autour de l'axe Y_C) et le lacet (rotation $R_{z,\gamma}$ d'angle γ autour de l'axe Z_C) avec :

$$R_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \end{bmatrix} \tag{III.1}$$

$$R_{y,\beta} = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \tag{III.2}$$

$$R_{z,\gamma} = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{III.3}$$

La matrice R est obtenue par la multiplication des trois matrices :

$$\begin{aligned} R &= R_{z,\gamma}.R_{y,\beta}.R_{x,\alpha} \\ &= \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \\ &= \begin{bmatrix} \cos\beta\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma \\ \cos\beta\sin\gamma & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma \\ -\sin\beta & \sin\alpha\cos\beta & \cos\alpha\cos\beta \end{bmatrix} \end{aligned} \tag{III.4}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

La transformation relative au changement de repère est :

$$M^C = R.M^W + T \tag{III.5}$$

En utilisant la notation homogène, on définit la matrice E de transformation homogène du repère monde $\mathcal{R}_W$ vers le repère caméra $\mathcal{R}_C$ telle que :

$$\begin{pmatrix} x_M^C \\ y_M^C \\ z_M^C \\ 1 \end{pmatrix} = E. \begin{pmatrix} x_M^W \\ y_M^W \\ z_M^W \\ 1 \end{pmatrix} \tag{III.6}$$

avec

$$E = \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{III.7}$$

E est la matrice du modèle externe, elle contient donc les paramètres extrinsèques de la caméra.

III.1.2.1.b Modèle interne

Projection perspective

Le point 3D M de l'espace se projette en m sur le plan image. Les coordonnées de m exprimées dans le repère $\mathcal{R}_C$ de la caméra sont notés $m^C = (x_m^C, y_m^C, z_m^C)^T$. Ils sont déduits à partir des coordonnées de M dans $\mathcal{R}_C$ et de la distance focale f grâce au théorème de Thalès :

$$\begin{cases} x_m^C = & f\frac{x_M^C}{z_M^C} \\ y_m^C = & f\frac{y_M^C}{z_M^C} \\ z_m^C = & f \end{cases} \tag{III.8}$$

Sous forme matricielle et avec la notation homogène, on a alors :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

$$\begin{pmatrix} s.x_m^C \\ s.y_m^C \\ s.z_m^C \\ s \end{pmatrix} = P. \begin{pmatrix} x_M^C \\ y_M^C \\ z_M^C \\ 1 \end{pmatrix} \tag{III.9}$$

avec

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \tag{III.10}$$

Transformation caméra / image

Les points de l'image sont mesurés en *pixels* dans le repère $\mathcal{R}_I$ associé à l'image. Soit (u_m, v_m, w_m) les coordonnées pixelliques du point m dans le repère image $\mathcal{R}_I$. La transformation des coordonnées de m du repère de la caméra $\mathcal{R}_C$ au repère image $\mathcal{R}_I$ s'écrit :

$$\begin{cases} u_m = & -k_u x_m^C + u_0 \\ v_m = & k_v y_m^C + v_0 \\ w_m = & 0 \end{cases} \tag{III.11}$$

où (u_0, v_0) sont les coordonnées du point principal O dans le repère image $\mathcal{R}_I$ (mesurées en pixels), k_u est le facteur d'échelle vertical (pixels/mm) et k_v est le facteur d'échelle horizontal.

C'est une transformation affine représentant un changement d'échelle, une rotation et une translation.

$$\begin{pmatrix} u_m \\ v_m \\ w_m \end{pmatrix} = \begin{pmatrix} -k_u & 0 & 0 \\ 0 & k_v & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_m^C \\ y_m^C \\ z_m^C \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \tag{III.12}$$

La composante w_m étant toujours nulle, on peut ignorer la troisième ligne. L'écriture matricielle avec les coordonnées homogènes prend alors la forme suivante :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

$$\begin{pmatrix} u_m \\ v_m \\ 1 \end{pmatrix} = K. \begin{pmatrix} x_m^C \\ y_m^C \\ z_m^C \\ 1 \end{pmatrix} \tag{III.13}$$

avec

$$K = \begin{pmatrix} -k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{III.14}$$

Modèle interne complet

Les équations du modèle interne complet (projection perspective suivie de la transformation caméra / image) reliant les coordonnées de M dans $\mathcal{R}_C$ aux coordonnées pixels de son projeté m dans l'image s'écrivent donc :

$$\begin{cases} u_m = & -k_u f \frac{x_M^C}{z_M^C} + u_0 \\ v_m = & k_v f \frac{y_M^C}{z_M^C} + v_0 \end{cases} \tag{III.15}$$

Soit sous forme matricielle :

$$\begin{aligned} \begin{pmatrix} su_m \\ sv_m \\ s \end{pmatrix} &= K.P. \begin{pmatrix} x_M^C \\ y_M^C \\ z_M^C \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -k_u & 0 & u_0/f & 0 \\ 0 & k_v & v_0/f & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} x_M^C \\ y_M^C \\ z_M^C \\ 1 \end{pmatrix} \end{aligned} \tag{III.16}$$

$K.P$ est la matrice du modèle interne, elle contient donc les paramètres intrinsèques de la caméra.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Modèle sténopé complet

En appliquant le modèle externe aux coordonnées d'un point M de la scène, exprimées dans le repère du monde $\mathcal{R}_W$, on obtient ses coordonnées dans le repère attaché à la caméra $\mathcal{R}_C$. Puis en appliquant le modèle interne aux coordonnées de M dans $\mathcal{R}_C$, on obtient les coordonnées du projeté m de M dans l'image.

$$\begin{aligned}\begin{pmatrix} su_m \\ sv_m \\ s \end{pmatrix} &= K.P.E \begin{pmatrix} x_M^W \\ y_M^W \\ z_M^W \\ 1 \end{pmatrix} \\ &= \Pi \begin{pmatrix} x_M^W \\ y_M^W \\ z_M^W \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \pi_{14} \\ \pi_{21} & \pi_{22} & \pi_{23} & \pi_{24} \\ \pi_{31} & \pi_{32} & \pi_{33} & \pi_{34} \end{pmatrix} \begin{pmatrix} x_M^W \\ y_M^W \\ z_M^W \\ 1 \end{pmatrix}\end{aligned} \tag{III.17}$$

Remarque

En pratique, les vecteurs du repère image ne sont pas tout à fait perpendiculaires. On introduit alors un paramètre supplémentaire : l'angle θ entre l'axe u et l'axe v. On définit un nouveau repère image $(u' - v')$. La relation entre le repère orthonormé $(u - v)$ et le nouveau repère $(u' - v')$ s'écrit :

$$\begin{cases} u' = u \sin\theta + v \cos\theta \\ v' = v \end{cases} \tag{III.18}$$

III.1.2.2 Calibration de la caméra

Calibrer une caméra consiste à identifier les valeurs pratiques de ses paramètres constituant le modèle théorique. Il s'agit plus spécifiquement des paramètres intrinsèques car ils constituent la géométrie du capteur lui-même (objectif, barette CCD). Un changement de mise au point ou de la focale implique bien entendu que le calibrage doit être refait.

Les paramètres intrinsèques de la caméra peuvent être exprimés en fonction des coefficients π_{ij} de la matrice globale Π du modèle de la caméra [65]. Le problème

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

de calibrage revient alors en pratique à déterminer Π. Plusieurs techniques ont été développées pour résoudre ce problème [42, 114]. La méthode utilisée dans notre approche est celle de Zhang [125].

La procédure de calibrage est la suivante :

- La première étape consiste à capteur des images d'une cible sous différentes poses de la caméra. Le modèle de cible utilisée est une mire plane sur laquelle a été imprimé un pattern de 4 par 5 carrés noirs (Fig.III.6).
- La deuxième étape consiste à extraire les points représentant les coins de chaque carré de chaque cible.
- La troisième étape consiste à calculer les paramètres intrinsèques de la caméra et les paramètres extrinsèques pour chaque pose à l'aide de l'algorithme de Zhang [125]. Le principe de cet algorithme est de calculer une solution analytique initiale (*closed-form solution*) suivie d'une technique d'optimisation basée sur un critère de maximum de vraisemblance. Notons par ailleurs que la méthode de Zhang tient compte des distorsions dans la minimisation.

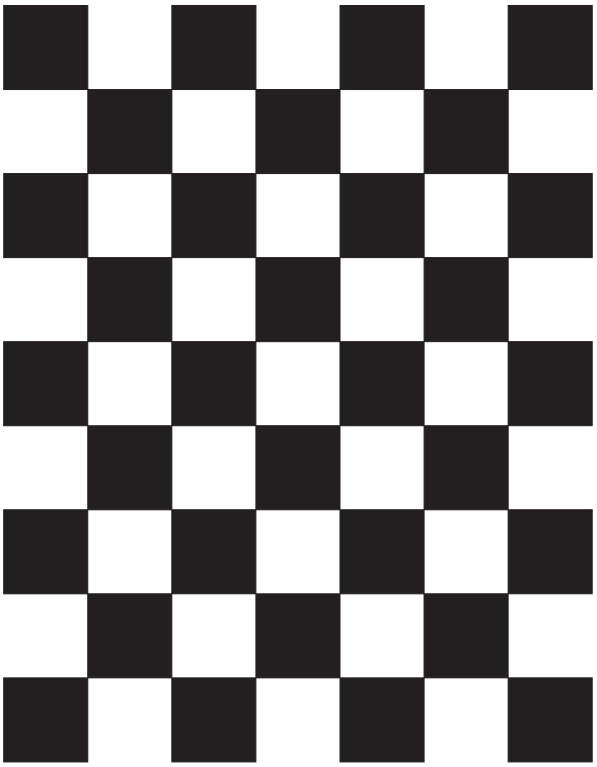


FIGURE III.6 – Mire utilisée pour la calibration de la caméra

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Implémentation

Afin de calibrer notre caméra, nous avons utilisé la Toolbox Matlab développée par Jean-Yves Bouguet [14]. Cette boîte à outil implémente la méthode de Zhang. Elle permet d'estimer les distances focales en u et v et exprimées en pixels, c'est-à-dire, $f_u = k_u f$ et $f_v = k_v f$, les coordonnées du centre optique, l'angle θ entre les axes u et v, et un vecteur de dimension 5 contenant les paramètres de distorsions radiale et tangentielle [63].

III.1.3 Paramétrage de la caméra virtuelle

Le paramétrage de la caméra virtuelle consiste à en fixer ses paramètres intrinsèques et extrinsèques. Cette étape permet donc de caractériser l'objectif de la caméra virtuelle similairement à celui de la caméra réelle d'une part et de positionner et d'orienter la caméra virtuelle dans le modèle virtuel 3D d'autre part.

III.1.3.1 Paramètres intrinsèques de la caméra virtuelle

Comme il a été indiqué lors de la description du SIG-3D (cf. II.2.2.1), le paramètre intrinsèque de la caméra virtuelle qui doit être fixé est l'angle d'ouverture horizontal, noté FOV (Field Of View).

Afin d'obtenir une image virtuelle équivalente à l'image réelle, le FOV de la caméra virtuelle doit être déterminé à partir des paramètres intrinsèques de la caméra réelle calibrée. L'influence du paramètre FOV sur l'image virtuelle obtenue est illustrée dans la Figure III.7. Les trois images correspondent à trois rendus obtenus en observant la même scène, pour la même pose de la caméra virtuelle mais avec trois FOV différents.

La relation permettant de déterminer FOV à partir des paramètres intrinsèques de la caméra réelle est (cf Figure III.8)

$$FOV = 2 \arctan \frac{l/2}{f'} \tag{III.19}$$

avec :

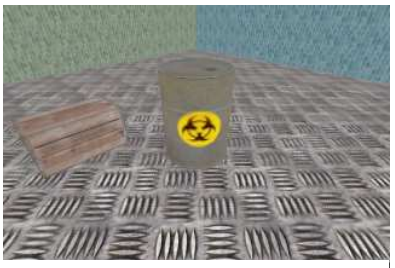
- l : la largeur de l'image réelle en nombre de pixels
- f' : la distance focale de la caméra réelle exprimée en pixels, c'est-à-dire

$$f' = f.k_v$$

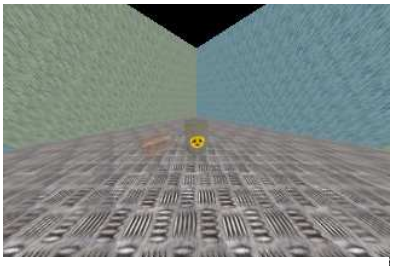
© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

FOV = 40°



FOV = 70°



FOV = 130°

FIGURE III.7 – Illustration de l'influence du paramètre FOV de la caméra virtuelle

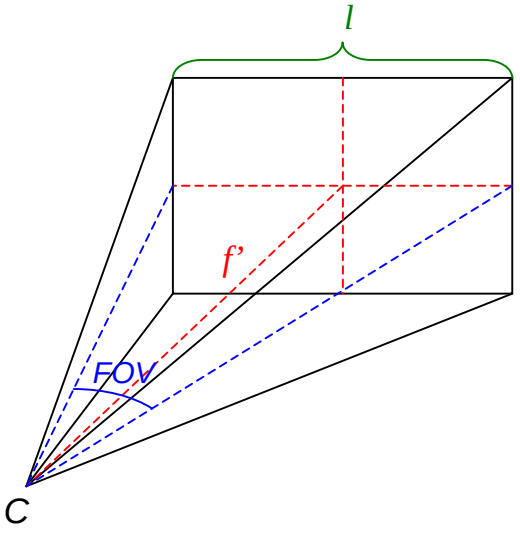


FIGURE III.8 – Calcul du paramètre intrinsèque FOV de la caméra virtuelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

D'autre part, la taille des images virtuelles produites par le SIG-3D est configurable. Nous avons choisi la même taille que celle des images réelles acquises par la caméra.

III.1.3.2 Paramètres extrinsèques de la caméra virtuelle

Afin d'obtenir l'image virtuelle correspondante à l'image réelle, il faut également positionner la caméra virtuelle dans le modèle 3D. Il faut donc connaître une valeur approchée de la pose de la caméra. Pour cela, nous utilisons l'estimation *a priori* de la pose prédite lors de l'étape de prédiction de l'algorithme de fusion multi-capteurs (cf § III.2.2.1). C'est la pose estimée à partir de la pose précédente, du modèle odométrique et des données délivrées par les odomètres et le gyromètre.

La pose *a priori* exprimée dans le système *WGS84* est ensuite convertie dans le repère local attaché au modèle 3D selon la procédure de géo-référencement du modèle virtuel 3D décrite dans le paragraphe II.2.2.3.

Le SIG-3D peut alors extraire les informations qui correspondent au paramètre intrinsèque et aux paramètres extrinsèques de la caméra virtuelle ainsi fixés. Pour la méthode de localisation ici présentée, les informations qui seront utilisées sont l'image virtuelle et le fichier de profondeur.

III.1.4 Détection des primitives de l'image

III.1.4.1 Introduction

Différents types de primitives sont utilisés dans le domaine de la vision : des contours, des points ou des régions. La nature de la scène permet de déterminer les primitives les plus adéquates. Ainsi, en présence de scènes structurées, telles que les environnements d'intérieur, les primitives privilégiées sont généralement les lignes ou contours, c'est-à-dire du point de vue du signal, des primitives caractérisées par une forte variation de l'intensité lumineuse. C'est par exemple le cas dans les travaux de Bourgeois [15], Carmichael et Hébert [27] ou Mikolajczyk *et al* [83].

Les environnements d'extérieur sont quant à eux souvent texturés, les primitives susceptibles d'être observées sont donc souvent assimilées à des coins, et correspondent à une forte variation bi-dimensionnelle du signal. D'autre part, détecter des

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

points est moins sensible aux occultations que détecter des primitives qui occupent une part significative de l'image.

Enfin, la détection de régions nécessite souvent des algorithmes plus coûteux en temps de calcul et est très sensible au bruit, aux occultations et aux changements de points de vues. Cependant, une méthode permettant d'extraire des régions invariantes aux transformations affines est proposée dans [117] afin d'apparier des images acquises depuis des points de vues éloignés.

L'application visée traitant de scènes extérieures, les primitives détectées sont des points d'intérêt. Il existe une large variété de points d'intérêts dans la littérature, pouvant être détectés par trois grandes familles de méthodes :

- Les méthodes qui s'appuient sur l'extraction préalable de contours. Une fois les contours extraits, les points de courbures maximales le long des chaînes de contour sont recherchés ou une approximation polygonale des contours est réalisée pour localiser des points particuliers tels que des intersections ou des inflexions.
- Les méthodes qui travaillent directement sur le signal en niveau de gris. Un point d'intérêt est alors caractérisé comme une forte variation bi-directionnelle du signal dans son voisinage.
- Les méthodes qui utilisent un modèle théorique de la primitive à détecter.

Dans [99], Schmid *et al* dressent une étude comparative de différents extracteurs de points d'intérêt. Le critère d'évaluation utilisé est la répétabilité. Ce critère vérifie si la détection est invariante aux changements de conditions de prises de vue : changement de luminosité, changement de position et d'orientation de la caméra, changement des paramètres intrinsèques de la caméra. Parmi les méthodes étudiées, le détecteur de Harris [60] donne souvent les meilleurs résultats.

Les primitives retenues pour notre méthode sont donc les points de Harris.

III.1.4.2 Points de Harris

Beaudet [7] et Moravec [84] ont présentés dans leurs travaux les premiers détecteurs de points d'intérêt. Le détecteur de Harris est une amélioration de la formulation de base de Moravec qui emploie une mesure d'auto-corrélation du signal. En effet, un point d'intérêt est caractérisé dans l'image par une zone où le signal varie fortement. Il correspond alors à un maximum local d'une mesure d'auto-corrélation.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Le calcul d'auto-corrélation consiste à sommer la différence de niveaux de gris pixel à pixel entre une fenêtre de référence $\mathcal{W}$ centrée sur le pixel p considéré et une autre fenêtre décalée.

Soient $p = (u, v)$ un point de l'image $\mathcal{I}$, $\mathcal{I}(u, v)$ le niveau de gris en ce point et $(\Delta u, \Delta v)$ un décalage. L'auto-corrélation est alors définie comme suit :

$$Corr(p) = \sum_{u_i, v_i \in \mathcal{W}} \left[\mathcal{I}(u_i, v_i) - \mathcal{I}(u_i + \Delta u, v_i + \Delta v)\right]^2 \tag{III.20}$$

où (u_i, v_i) désignent les points dans la fenêtre $\mathcal{W}$ centrée sur p.

Le développement de Taylor limité à l'ordre 1 de $\mathcal{I}(u_i + \Delta u, v_i + \Delta v)$ est :

$$\mathcal{I}(u_i + \Delta u, v_i + \Delta v) \approx \mathcal{I}(u_i, v_i) + \left[I_u(u_i, v_i) I_v(u_i, v_i)\right] \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} + \varepsilon(u, v) \tag{III.21}$$

où $I_u(\cdot, \cdot)$ et $I_v(\cdot, \cdot)$ désignent les dérivées partielles de l'image $\mathcal{I}$ selon u et v respectivement.

En négligeant le résidu $\varepsilon(u, v)$, et en substituant l'équation III.21 dans l'équation III.20, on obtient :

$$\begin{aligned}
Corr(p) &= \textstyle\sum_{u_i, v_i \in \mathcal{W}} \left[\mathcal{I}(u_i, v_i) - \mathcal{I}(u_i + \Delta u, v_i + \Delta v)\right]^2 \\
&= \textstyle\sum_{u_i, v_i \in \mathcal{W}} \left(\mathcal{I}(u_i, v_i) - \mathcal{I}(u_i, v_i) - \left[I_u(u_i, v_i) I_y(u_i, v_i)\right] \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \right)^2 \\
&= \textstyle\sum_{u_i, v_i \in \mathcal{W}} \left(- \left[I_u(u_i, v_i) I_y(u_i, v_i)\right] \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \right)^2 \\
&= \textstyle\sum_{u_i, v_i \in \mathcal{W}} \left(\left[I_u(u_i, v_i) I_y(u_i, v_i)\right] \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \right)^2 \\
&= \textstyle\sum_{u_i, v_i \in \mathcal{W}} \left[\Delta u \ \Delta v\right] \begin{bmatrix} I_u^2(u_i, v_i) & I_u(u_i, v_i) I_v(u_i, v_i) \\ I_u(u_i, v_i) I_v(u_i, v_i) & I_v^2(u_i, v_i) \end{bmatrix} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \\
&= \left[\Delta u \ \Delta v\right] \left(\textstyle\sum_{u_i, v_i \in \mathcal{W}} \begin{bmatrix} I_u^2(u_i, v_i) & I_u(u_i, v_i) I_v(u_i, v_i) \\ I_u(u_i, v_i) I_v(u_i, v_i) & I_v^2(u_i, v_i) \end{bmatrix} \right) \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \\
&= \left[\Delta u \ \Delta v\right] C \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix}
\end{aligned} \tag{III.22}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

avec :

$$C = \left(\sum_{u_i, v_i \in \mathcal{W}} \begin{bmatrix} I_u^2(u_i, v_i) & I_u(u_i, v_i) I_v(u_i, v_i) \\ I_u(u_i, v_i) I_v(u_i, v_i) & I_v^2(u_i, v_i) \end{bmatrix} \right) \quad \text{(III.23)}$$

Le détecteur de Harris est basé sur l'hypothèse qu'un point d'intérêt se situe aux maxima locaux de la fonction d'auto-corrélation. Or, les valeurs propres de la matrice d'auto-corrélation C correspondent aux courbures principales de la fonction d'auto-corrélation $Corr$. Les valeurs propres de la matrice C permettent donc de caractériser le point étudié. Soient λ_1 et λ_2 les valeurs propres de C telle que $\lambda_1 \geq \lambda_2$. Trois cas sont envisageables :

- Si λ_1 et λ_2 sont faibles alors le point appartient à une région homogène
- Si λ_1 est élevée et λ_2 est faible, alors le point appartient à un contour : le vecteur propre associé à la valeur propre λ_1 dominante est perpendiculaire à ce contour
- Si les deux valeurs propres λ_1 et λ_2 sont élevées alors le point considéré est un point d'intérêt car il se situe à l'intersection entre deux arrêtes : il représente un maximum local, caractéristique d'un coin.

Le calcul des valeurs propres de C étant lourd, Harris propose l'expression suivante, dans laquelle les valeurs propres n'interviennent pas explicitement :

$$R = Det(C) - \alpha Trace^2(C) \quad \text{(III.24)}$$

Harris conseille de régler le scalaire positif α entre $[0,04; 0,06]$. Si la fonction R dépasse un seuil fixé par l'utilisateur, alors on est en présence d'un point d'intérêt.

III.1.4.3 Algorithme du détecteur de Harris

Notons que, l'image étant souvent bruitée, l'auto-corrélation est multipliée par une fonction de lissage Gaussienne.

L'algorithme du détecteur de Harris est alors le suivant :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- Calcul des gradients de l'image : I_u et I_v
- A partir de I_u et I_v, calculer I_u^2, I_v^2 et $I_u I_v$
- Lissage de I_u^2, I_v^2 et $I_u I_v$ à l'aide d'un filtre gaussien F ($F_{u,v} = e^{-\frac{u^2+v^2}{2\sigma^2}}$)
- Pour chaque pixel :
  - Calcul de la matrice C de Harris :

$$C = F \begin{bmatrix} I_u^2 & I_u I_v \\ I_u I_v & I_v^2 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}$$

  - Calcul de la trace et du déterminant de la matrice (sans calculer les valeurs propres)
    - $Trace(C) = \lambda_1 + \lambda_2 = C_{11} + C_{22}$
    - $Determinant(C) = \lambda_1 \lambda_2 = C_{11}C_{22} - C_{12}C_{21}$
  - Calcul de la réponse R du détecteur :
    - $R = Det(C) - \alpha Trace^2(C)$
    - α est un paramètre à régler, typiquement $\alpha = 0,04$
- Extraction des maxima locaux positifs de la réponse R
- Extraction des n meilleurs points de Harris ($n = 500$ dans notre cas)

Remarque sur le choix du nombre n de points de Harris retenus :

Sur une image, on obtient généralement un très grand nombre de maxima locaux de R (de l'ordre de 5000). Cependant, on ne peut en retenir qu'une partie pour des raisons de temps de calcul. Après plusieurs essais, un bon compromis entre temps d'exécution et précision est $n = 500$ avec les données traitées et le matériel utilisé. Si on diminue excessivement n, la pose calculée perd alors en précision.

III.1.5 Mise en correspondance 2D/2D

Une fois que les points de Harris sont détectés dans l'image réelle ainsi que dans l'image virtuelle, l'étape suivante consiste à apparier ces deux ensembles de points.

On cherche donc à former des appariements (p^R, p^V, s) avec p^R point d'intérêt de l'image réelle I^R, p^V point d'intérêt de l'image virtuelle I^V et s un score de corrélation. Le calcul du score de corrélation s se fait sur un voisinage V_{15} des pixels p^R et p^V (fenêtre 15 x 15 pixels centrée sur p^R et p^V). s est d'autant plus élevé que le voisinage des points p^R et p^V est semblable.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Remarque sur le choix de la taille du voisinage retenue (15 x 15 pour des images 640 x 480) :

Si on choisit une fenêtre trop grande, alors d'une part le temps de calcul est prohibitif et d'autre part des appariements corrects sont éliminés. *A contrario*, si on choisit une fenêtre trop petite, alors il y a beaucoup de faux appariements, car le voisinage n'est pas assez discriminant. Pour les images traitées et le matériel utilisé, une taille de 15 x 15 est un bon compromis.

Pour chaque point p^R de l'image réelle, on définit une zone de recherche W rectangulaire centrée sur p^R dans l'image virtuelle (cf Figure III.9). On forme tous les appariements possibles (p^R, p^V, s) avec p^V point de Harris de l'image virtuelle détecté dans la zone de recherche W. Seuls les appariements ayant un score de corrélation supérieur à un seuil fixé (0,75) sont retenus. Puis, pour satisfaire la contrainte d'unicité (un point ne peut pas être apparié à deux points différents), l'appariement (p^R, p^V, s) avec le score s maximal est extrait, et tous les autres couples qui contiennent p^R ou p^V sont éliminés. Cette opération est répétée jusqu'à ce qu'il n'y ait plus de candidats possibles.

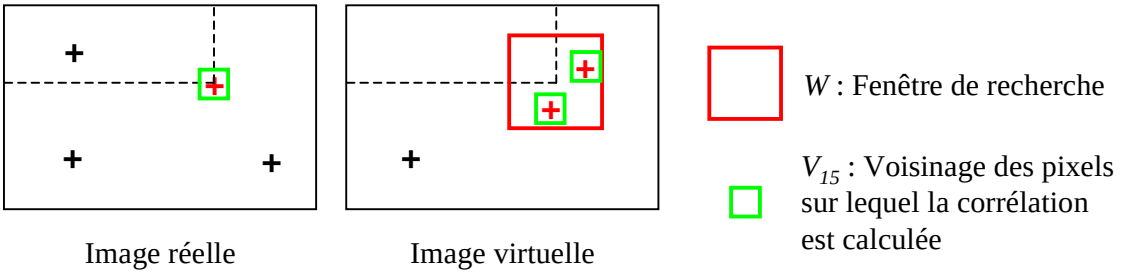


FIGURE III.9 – Un point de Harris détecté dans l'image réelle (Gauche) et la zone de recherche de son correspondant dans l'image virtuelle (Droite)

Remarque sur le choix du seuil du score de corrélation :

Après plusieurs essais, il s'avère que 0,75 est un bon compromis pour le choix du seuil du score de corrélation. Si on choisit un score trop bas, alors de nombreux faux appariements sont retenus. *A contrario*, si on choisit un seuil trop élevé, alors des appariements corrects sont éliminés, et il reste trop peu d'appariements pour calculer avec précision la pose de la caméra.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

La mesure de similarité utilisée est le score de corrélation croisée centrée et normalisée (ZNCC, *Zero mean Normalized Cross Correlation*) :

$$ZNCC\left(p^R, p^V\right) = \frac{\sum_{d \in V_{15}} \left(I^R\left(p^R + d\right) - \overline{I^R}\left(p^R\right)\right)\left(I^V\left(p^V + d\right) - \overline{I^V}\left(p^V\right)\right)}{\sqrt{\sum_{d \in V_{15}} \left(I^R\left(p^R + d\right) - \overline{I^R}\left(p^R\right)\right)^2} \sqrt{\sum_{d \in V_{15}} \left(I^V\left(p^V + d\right) - \overline{I^V}\left(p^V\right)\right)^2}} \tag{III.25}$$

avec :

$$\overline{I^R}\left(p^R\right) = \frac{1}{15^2} \sum_{d \in V_{15}} I^R\left(p^R + d\right) \tag{III.26}$$

et

$$\overline{I^V}\left(p^V\right) = \frac{1}{15^2} \sum_{d \in V_{15}} I^V\left(p^V + d\right) \tag{III.27}$$

Avec cette méthode, la présence d'outliers (c'est-à-dire de faux appariements) dans l'ensemble de correspondance est quasiment inévitable. Or, pour que le calcul de pose soit précis, tous ces outliers doivent être détectés et éliminés. Pour cela, nous utilisons l'algorithme d'estimation robuste RANSAC [43]. La mise en correspondance robuste basée sur RANSAC permet de sélectionner plus ou moins aléatoirement un nombre réduit de paires de points de Harris, de calculer la transformation pour le recalage candidat, d'apparier le plus possible de paires de primitives modèle-image et au bout du compte de retenir le recalage dont la transformation calculée rassemble le maximum d'appariements (l'ensemble consensuel). Cette technique permet d'éliminer de mauvais appariements parmi les appariements hypothétiques.

L'algorithme est alors le suivant :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

1. Sélectionner aléatoirement quatre paires de points points appariés
2. Calculer l'homographie H
3. Compter le nombre d'inliers, c'est-à-dire le nombre de paires vérifiant

$$distance(p^R, Hp^V) < \varepsilon$$

(avec p^R et p^V les points de Harris détectés respectivement dans l'image réelle et dans l'image virtuelle)

4. Retourner à l'étape 1

Finalement, on garde l'ensemble d'inliers le plus important.

III.1.6 Calcul des points 3D correspondants aux points de Harris de l'image virtuelle

Pour calculer les positions 3D correspondantes aux points de Harris détectés dans l'image virtuelle, une première solution consisterait à calculer, dans le repère du modèle 3D, les intersections entre le modèle 3D visualisé et les lignes de vue passant par le centre de la caméra virtuelle et les points 2D de Harris de l'image virtuelle par lancer de rayons. Avec un grand nombre de points, cette solution peut cependant s'avérer coûteuse en terme de temps de calcul. Nous avons donc opté pour une méthode basée sur le Z-Buffering. Pour afficher la projection du modèle 3D, le moteur OpenGL calcule pour tous les pixels de l'image des informations de profondeur dans le repère attaché à la caméra virtuelle (principe du Z-Buffering déjà décrit dans le paragraphe II.2.2.2.b). Connaissant la position de la caméra virtuelle dans le repère attaché au modèle 3D, la distance de chaque pixel projeté à la caméra virtuelle (donnée par le Z-Buffer et stockée dans le « fichier de profondeur »), et la matrice de projection de la caméra, on connaît potentiellement pour chaque pixel de l'image virtuelle sa position correspondante dans le repère attaché au modèle virtuel.

Les relations suivantes décrivent le calcul des coordonnées dans le repère $\mathcal{R}_M$ attaché au modèle 3D d'un point 3D de la scène, à partir des coordonnées du pixel correspondant dans le plan de l'image virtuelle, du fichier de profondeur (qui contient les distances entre la caméra virtuelle et les points 3D projetés sur l'image virtuelle) et la position de la caméra virtuelle dans le modèle 3D.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Considérons p un pixel de l'image virtuelle (dont les points de Harris). Soient (u, v) les coordonnées de p dans le plan image. Notons P le point 3D correspondant à p, c'est-à-dire le point du modèle 3D qui se projette en p sur l'image virtuelle. L'objectif est de déterminer les coordonnées de P dans le repère $\mathcal{R}_M$ attaché au modèle 3D.

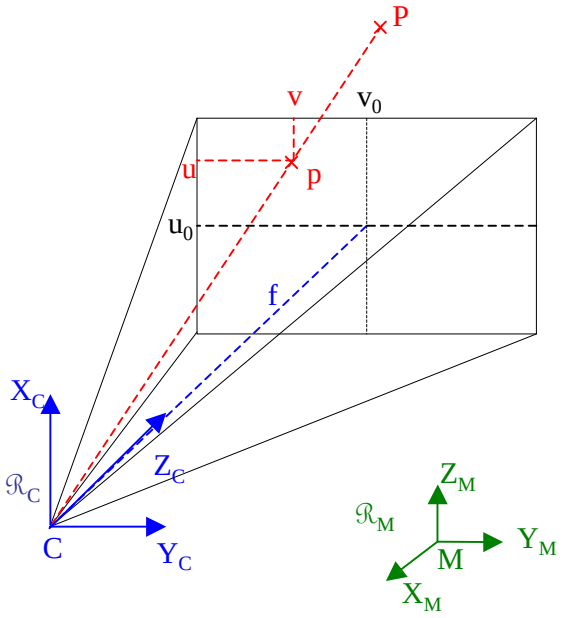


FIGURE III.10 – Les coordonnées de P dans $\mathcal{R}_M$ se déduisent de son correspondant p, de la pose de la caméra C dans $\mathcal{R}_M$ ainsi que de la distance d lue dans le Z-Buffer

Définissons quelques notations (cf Figure III.10) :

- $P^M = (x_P^M, y_P^M, z_P^M)^T$: coordonnées du point P dans le repère $\mathcal{R}_M$ attaché au modèle 3D
- $P^C = (x_P^C, y_P^C, z_P^C)^T$: coordonnées du point P dans le repère $\mathcal{R}_C$ de la caméra
- $p = (u, v)^T$: coordonnées du pixel p correspondant de P dans le repère image $\mathcal{R}_I$

Les coordonnées P^C de P dans $\mathcal{R}_C$ sont déterminées à partir des coordonnées (u, v) de son correspondant p dans l'image à l'aide de la relation suivante (modèle de projection inverse) :

$$\begin{cases} z_P^C = \frac{d.f}{m} \\ x_P^C = -z_P^C \frac{u-u_0}{f} \\ y_P^C = z_P^C \frac{v-v_0}{f} \end{cases} \quad \text{(III.28)}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

avec :

- f : la distance focale exprimée en pixels
- $(u_0, v_0)^T$: les coordonnées du point principal
- d : la distance en mètres entre le point 3D P et la caméra virtuelle C donnée par le Z-Buffer et stockée dans le fichier profondeur
- m : la distance en pixels entre la caméra virtuelle et le pixel p

$$m = \sqrt{f^2 + (u - u_0)^2 + (v - v_0)^2} \tag{III.29}$$

Les coordonnées P^C de P dans $\mathcal{R}_C$ étant déterminées, les coordonnées P^M de P dans $\mathcal{R}_M$ sont calculées comme suit :

$$\left(x_P^M, y_P^M, z_P^M\right)^T = R.\left(x_P^C, y_P^C, z_P^C\right)^T + T \tag{III.30}$$

avec R et T la matrice de rotation et le vecteur de translation correspondants à la transformation rigide entre le repère attaché à la caméra virtuelle $\mathcal{R}_C$ et le repère du modèle 3D $\mathcal{R}_M$. R et T sont déterminés à partir de la pose connue de la caméra virtuelle dans le modèle 3D d'une façon analogue à celle présentée dans le paragraphe III.1.2.1.a.

Ces relations permettent donc de calculer les coordonnées 3D, dans le repère attaché au modèle 3D, des points projetés sur les pixels de l'image virtuelle et en particulier sur les points de Harris.

Dans le paragraphe III.1.5, des correspondances 2D/2D entre les points de Harris détectés dans l'image réelle et les points de Harris détectés dans l'image virtuelle ont été établies (flèches rouges de la Figure III.11), puis dans le présent paragraphe (III.1.6), la position 3D des points de Harris de l'image virtuelle a été calculée (flèches bleues de la Figure III.11), ce qui permet d'obtenir des correspondances 2D/3D entre les points 2D de l'image réelle courante et les points 3D du modèle virtuel (flèches vertes de la Figure III.11). L'étape suivante de l'algorithme consiste à calculer la pose de la caméra réelle à partir de ces appariements 2D/3D de points (cf § III.1.7).

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

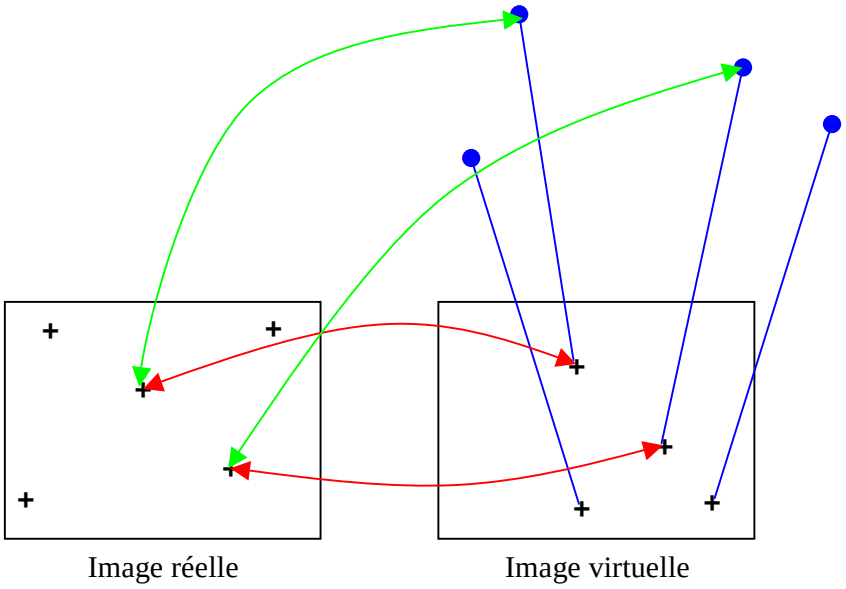


FIGURE III.11 – Appariements 2D/3D

III.1.7 Calcul de la pose de la caméra réelle à partir des appariements 2D/3D de points

Ce paragraphe traite de l'estimation des paramètres extrinsèques de la caméra réelle, c'est-à-dire de l'estimation de sa pose, en connaissant des correspondances entre des points 3D exprimés dans le repère du modèle 3D et leur projection 2D dans le plan de l'image réelle (Figure III.12).

Quelques méthodes couramment utilisées pour estimer la pose de la caméra à partir d'appariements 2D/3D sont d'abord données en introduction (§ III.1.7.1), puis la méthode choisie est présentée (§ III.1.7.2).

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

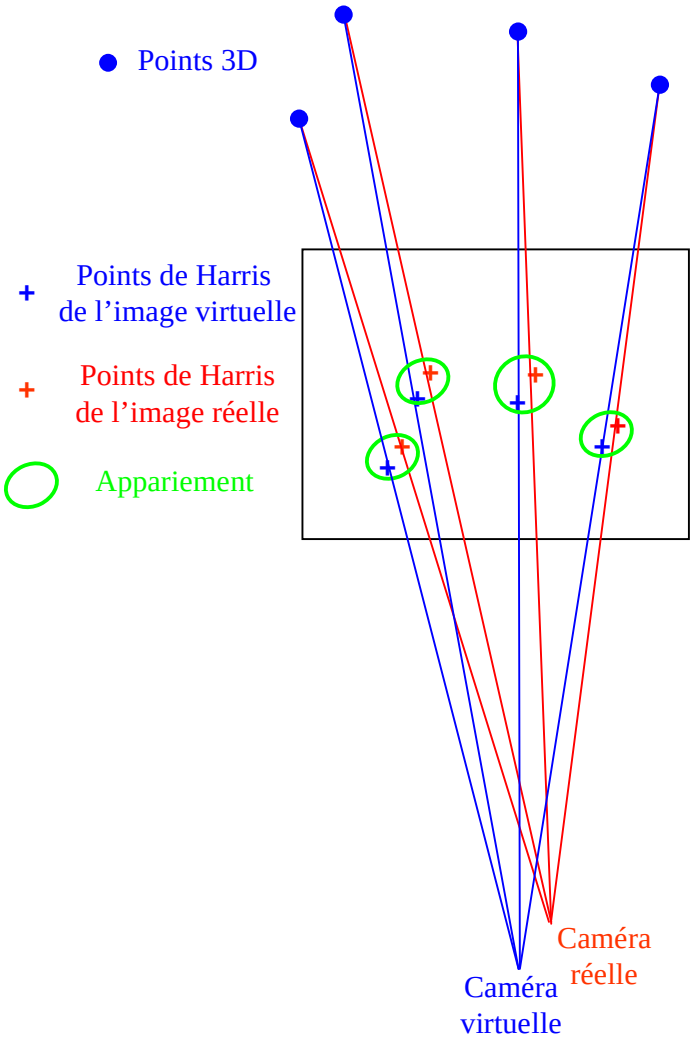


FIGURE III.12 – Calcul de la pose de la caméra réelle à partir des appariements 2D/3D

III.1.7.1 Introduction

III.1.7.1.a Estimation par minimisation algébrique

Considérons n correspondances entre des points 3D P_i et leurs projections p_i. L'objectif est de déterminer la matrice de projection Π projetant P_i sur p_i. On recherche donc Π telle que :

$$\forall i, p_i \approx \Pi P_i \tag{III.31}$$

On rappelle (cf § III.1.2.1) que Π est de la forme $\Pi = K.P.E$ avec $K.P$ définissant les paramètres intrinsèques de la caméra et E définissant les paramètres extrinsèques.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Direct Linear Transformation

La DLT (*Direct Linear Transformation*) a été introduite dans la communauté vision [42, 62] pour estimer entièrement Π à partir des correspondances entre $P_i = (x_i, y_i, z_i, 1)^T$ et $p_i = (u_i v_i 1)^T$.

En utilisant l'équation III.31, chaque correspondance $P_i \longleftrightarrow p_i$ donne deux équations linéairement indépendantes sur les entrées de Π :

$$\frac{\pi_{11}x_i + \pi_{12}y_i + \pi_{13}z_i + \pi_{14}}{\pi_{31}x_i + \pi_{32}y_i + \pi_{33}z_i + \pi_{34}} = u_i \tag{III.32}$$

$$\frac{\pi_{21}x_i + \pi_{22}y_i + \pi_{23}z_i + \pi_{24}}{\pi_{31}x_i + \pi_{32}y_i + \pi_{33}z_i + \pi_{34}} = v_i \tag{III.33}$$

Ces équations peuvent être réécrites sous la forme d'un système de la forme $A\pi = 0$ où π est un vecteur composé des coefficients de Π. La solution à ce système est déduite de la décomposition en valeurs singulières de A comme étant le vecteur propre associé à la plus petite valeur propre.

Si $K.P$ est connue, les paramètres extrinsèques peuvent être déduits de Π à un facteur d'échelle près :

$$[R|T] \approx (KP)^{-1}\Pi \tag{III.34}$$

La matrice R trouvée n'est pas forcément une matrice de rotation, mais elle peut être corrigée en ce sens [125]. En conditions réelles, les mesures p_i sont généralement bruitées. L'erreur minimisée par la DLT étant algébrique, les paramètres de la caméra estimés avec cette méthode doivent généralement être affinés par une phase d'optimisation non linéaire de l'erreur de reprojection.

POSIT

POSIT [35] est un algorithme très utilisé pour l'estimation de pose dans le cas où $n \geq 4$ (avec au moins quatre points non coplanaires). Le processus est itéré jusqu'à convergence, et chaque itération se décompose en deux étapes :

- POS (*Pose from Orthography and Scaling*), qui approche la projection perspective de la position et de l'orientation de la caméra en résolvant un système linéaire
- POSIT (*POS with ITerations*), où le facteur d'échelle de la projection orthographique est mis à jour itérativement afin de faire converger la projection

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

orthographique vers la projection perspective, permettant ainsi de déterminer la « vraie » pose de l'objet visualisé.

Cette technique est simple à mettre en œuvre, mais est relativement sensible au bruit. Une version modifiée est présentée dans [32] pour traiter à part le cas des points coplanaires.

III.1.7.1.b Estimation par minimisation géométrique

L'avantage des techniques d'estimation par minimisation algébrique décrites précédemment est qu'elles sont rapides et ne requièrent pas d'initialisation de la pose. Elles sont cependant relativement sensibles au bruit, et de ce fait manquent de précision. Afin d'obtenir de meilleurs résultats, il est donc préférable de minimiser une erreur géométrique, à savoir la distance euclidienne (plus précisément son carré) entre les points de Harris 2D mesurés dans l'image réelle et les points 3D projetés correspondants. On a alors :

$$\left[\tilde{R}|\tilde{T}\right] = \underset{[R|T]}{\operatorname{argmin}} \left(\sum_{i=1}^{n} \text{distance}^2(\Pi P_i, p_i) \right) \tag{III.35}$$

La pose estimée est optimale sous la condition que les erreurs de mesures soient indépendantes et suivent une distribution Gaussienne. Elle est généralement calculée itérativement par une méthode de minimisation des moindres carrés (Gauss-Newton, Levenberg-Marquardt, ...), qui nécessite une initialisation de la pose proche de la solution pour ne pas converger vers un minimum local. Au prix de calculs plus lourds, ce problème peut être palier en utilisant des méthodes telles que le recuit simulé.

III.1.7.2 Méthode de calcul de la pose choisie

A partir des n correspondances établies entre les points 2D de l'image réelle et les points 3D du modèle 3D, il est possible de déterminer la pose de la caméra. La méthode choisie pour cela se base sur un algorithme de minimisation itérative à partir d'une valeur approchée de la pose. La valeur approchée de la pose utilisée dans notre cas est la pose *a priori* prédite dans l'étape de prédiction de l'algorithme de fusion multi-capteurs (cf § III.2.2). La fonction à minimiser est l'erreur de reprojection des appariements des points 3D dans l'image réelle en faisant varier les paramètres qui définissent la pose de la caméra réelle [2]. La minimisation est faite par l'algorithme de Levenberg-Marquardt, qui est une méthode qui combine l'algorithme de Gauss-Newton et une descente de gradient.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Soient C les paramètres extrinsèques de la caméra réelle, $P = \{P_1, \ldots, P_i, \ldots, P_n\}$ l'ensemble des points 3D appariés et $p = \{p_1, \ldots, p_i, \ldots, p_n\}$ l'ensemble des points de Harris correspondants mesurés dans l'image réelle. La fonction de coût à minimiser en fonction de C est la somme des erreurs de reprojection des n points 3D appariés :

$$f(C) = \sum_{P_i \in P} \varepsilon_i^2 \tag{III.36}$$

avec $\varepsilon_i^2 = d^2(p_i, \Pi P_i)$ la distance euclidienne entre ΠP_i la projection du point 3D P_i à travers la caméra C et l'observation correspondante mesurée (points de Harris de l'image réelle) p_i.

On rappelle que Π est la matrice de projection de la caméra réelle composée des ses paramètres extrinsèques C et de ses paramètres intrinsèques connus (cf § III.1.2.1).

III.1.8 Incertitude associée à l'observation cartographique 3D

Comme il a été vu dans le paragraphe précédent (III.1.7.2), l'estimation de la pose de la caméra réelle est donnée après convergence d'un processus itératif minimisant la somme des distances des points 3D du modèle aux points de Harris appariés de l'image réelle. La matrice de covariance des paramètres de position et d'orientation ne peut donc pas être calculée analytiquement. Cette matrice est alors estimée après la dernière itération du processus d'optimisation.

Pour estimer la matrice de covariance de l'observation cartographique 3D, il faudrait alors inverser la matrice $J^T J$, où J est la matrice Jacobienne de la fonction associant au vecteur des paramètres extrinsèques de la caméra, le vecteur de mesure des erreurs de projection.

Améliorations envisagées

La méthode d'estimation de l'erreur de l'observation cartographique décrite précédemment ne tient compte que des incertitudes sur les mesures (liées à l'image réelle). Quelques pistes de réflexions sont cependant envisagées afin de l'améliorer.

Parmi les sources d'erreurs négligées et dont il faudrait tenir compte, citons :

- l'erreur du modèle 3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- l'erreur au niveau de la calibration de la caméra : prise en compte de l'incertitude sur l'estimation des paramètres intrinsèques de la caméra réelle
- l'incertitude sur la transformation supposée rigide entre le véhicule et la caméra
- l'influence de l'estimation de la pose *a priori* utilisée

Afin de prendre en compte l'erreur du modèle 3D, une possibilité serait de réécrire le problème de minimisation avec pour inconnues non seulement la pose de la caméra mais également la position 3D des points du modèle autour de leur position calculée. C'est l'approche suivie par Royer dans sa thèse [95, 96]. On pourrait également envisager d'ajouter de la même façon l'incertitude sur l'estimation des paramètres intrinsèques de la caméra.

L'influence de l'estimation de la pose *a priori* utilisée intervient essentiellement sur l'apparition de faux-appariements. Ainsi, on considère que si l'incertitude sur cette estimée *a priori* est trop grande, alors le nombre de faux-appariements pourraient être trop important pour permettre un calcul de l'observation cartographique 3D satisfaisant. Il serait alors intéressant d'instaurer un test de consistance basé sur l'incertitude de l'estimation *a priori*, avant de calculer l'observation cartographique 3D .

III.2 Fusion de données multi-capteurs

Dans un premier temps (§ III.2.1), le problème du filtrage dans le cas général est exposé. Différentes approches classiquement utilisées sont présentées, notamment les méthodes analytiques telles que le filtre de Kalman pour le cas linéaire et ses extensions au cas non-linéaire : le filtre de Kalman étendu et le filtre de Kalman sans parfum. De nombreuses autres approches existent également comme le filtre d'information, le filtrage par mélange de gaussiennes, le filtrage particulaire, les approches ensemblistes ... Celles-ci n'étant cependant pas considérées dans cette thèse, elles ne seront que très brièvement abordées.

Nous avons choisi le filtre de Kalman sans parfum pour notre application. Nous présentons donc dans un deuxième temps (§ III.2.2) l'application de ce filtre dans le cadre de la fusion multi-capteurs pour la localisation de véhicules.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.2.1 Introduction

Dans cette partie, nous présentons le filtre de Kalman, le filtre de Kalman étendu et le filtre de Kalman sans parfum dans le cadre général.

Soient les modèles d'évolution et de mesure suivants :

$$x_{t+1} = f_t(x_t, v_t) \quad \text{(III.37)}$$

$$z_t = h_t(x_t) + w_t \quad \text{(III.38)}$$

où v_t et w_t sont des bruits affectant les processus d'évolution et de mesure.

Dans le cas où la fonction f_t est linéaire en x_t et v_t, la fonction h_t est linéaire en x_t et les bruits v_t et w_t sont gaussiens, on parle de modèle linéaire à bruits additifs gaussiens ou plus simplement de modèle linéaire gaussien. Pour ce type de modèles, le filtre de Kalman [71] est une solution au filtrage optimal.

Dans le cas où les fonctions f_t ou h_t sont non linéaires, ou si les bruits v_t ou w_t ne sont pas gaussiens, des filtres non optimaux déterministes basés sur l'algorithme du filtre de Kalman linéaire sont couramment utilisés. Le filtre de Kalman étendu (*Extended Kalman Filter*, EKF) procède à une linéarisation locale des fonctions non linéaire d'évolution et de mesure ; le filtre de Kalman sans parfum (*Unscented Kalman Filter*, UKF) procède à une approximation de la densité de probabilité par un nuage de points générés de façon déterministe, en utilisant la transformée sans parfum (*Unscented Transform*, UT).

III.2.1.1 Filtre de Kalman

On suppose que le modèle constitué des équations III.37 et III.38 peut se mettre sous la forme des modèles linéaires d'état de de mesure suivants :

$$x_{t+1} = F_t x_t + C_t u_t + G_t v_t \quad \text{(III.39)}$$

$$z_t = H_t x_t + w_t \quad \text{(III.40)}$$

où x_t est le vecteur d'état à l'instant t, v_t est le bruit d'évolution additif et gaussien, u_t est une entrée connue, F_t est la matrice d'évolution de l'état, C_t est la matrice de transition de l'entrée et G_t est la matrice de transition du bruit. z_t est

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

la mesure réalisée au temps t, w_t est le bruit de mesure additif et gaussien et H_t est la matrice de mesure. Les bruits d'évolution et de mesure sont supposés gaussiens, décorrélés et centrés, de matrices de covariance respectives :

$$\mathbb{E}\left[v_i v_j^T\right] = \delta_{i,j} Q_i \quad \text{(III.41)}$$

$$\mathbb{E}\left[w_i w_j^T\right] = \delta_{i,j} R_i \quad \text{(III.42)}$$

avec :

$$\delta_{i,j} = \begin{cases} 1 \text{ si } i = j \\ 0 \text{ sinon} \end{cases} \quad \text{(III.43)}$$

A l'instant $t = 0$, x_0 est supposé distribué selon une densité de probabilité gaussienne de moyenne et matrice de covariance connues :

$$x_0 \approx \mathcal{N}(\hat{x}_0, \Sigma_{0|0}) \quad \text{(III.44)}$$

Ces hypothèses correspondent à la définition de l'*a priori* initial du filtre de Kalman.

Le filtre de Kalman peut s'écrire sous la forme de relations de récurrence :

$$p(x_{t-1}|z_{1:t-1}) = \mathcal{N}(x_{t-1}; \hat{x}_{t-1|t-1}, \Sigma_{t-1|t-1}) \quad \text{(III.45)}$$

$$p(x_t|z_{1:t-1}) = \mathcal{N}(x_t; \hat{x}_{t|t-1}, \Sigma_{t|t-1}) \quad \text{(III.46)}$$

$$p(x_t|z_{1:t}) = \mathcal{N}(x_t; \hat{x}_{t|t}, \Sigma_{t|t}) \quad \text{(III.47)}$$

avec :

$$\hat{x}_{t|t-1} = F_{t-1}\hat{x}_{t-1|t-1} + C_{t-1}u_{t-1} \quad \text{(III.48)}$$

$$\Sigma_{t|t-1} = F_{t-1}\Sigma_{t-1|t-1}F_{t-1}^T + G_{t-1}Q_{t-1}G_{t-1}^T \quad \text{(III.49)}$$

$\hat{x}_{t|t-1}$ et $\Sigma_{t|t-1}$ étant respectivement la prédiction de l'état et la matrice de covariance de l'erreur de prédiction.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t \nu_t \quad \text{(III.50)}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t S_t K_t^T \quad \text{(III.51)}$$

$\hat{x}_{t|t}$ et $\Sigma_{t|t}$ étant respectivement l'estimée de l'état et la matrice de covariance de l'erreur d'estimation.

K_t est le gain de Kalman défini à l'instant t par :

$$K_t = \Sigma_{t|t-1} H_t^T \left[H_t \Sigma_{t|t-1} H_t^T + R_t \right]^{-1} \quad \text{(III.52)}$$

ν_t est l'innovation, c'est-à-dire la différence entre la mesure et la prédiction de la mesure donnée par :

$$\nu_t = z_t - H_t \hat{x}_{t|t-1} \quad \text{(III.53)}$$

de matrice de covariance :

$$S_t = H_t \Sigma_{t|t-1} H_t^T + R_t \quad \text{(III.54)}$$

III.2.1.2 Filtre de Kalman Etendu

Lorsque les fonctions f_t et h_t sont non linéaires, il n'est plus possible d'écrire les équations III.37 et III.38 sous les formes III.39 et III.40. Cependant, lorsque ces fonctions sont faiblement non linéaires, une linéarisation locale peut être une approximation satisfaisante de la non-linéarité. Le filtre de Kalman étendu est une approximation du filtrage optimal qui se base sur cette linéarisation. Il fait l'hypothèse que la densité de probabilité *a posteriori* $p(x_t|z_{1:t})$, qui n'est dans ce cas plus gaussienne, peut être approchée par une densité gaussienne selon les récurrences suivantes :

$$p(x_{t-1}|z_{1:t-1}) \approx \mathcal{N}(x_{t-1}; \hat{x}_{t-1|t-1}, \Sigma_{t-1|t-1}) \quad \text{(III.55)}$$

$$p(x_t|z_{1:t-1}) \approx \mathcal{N}(x_t; \hat{x}_{t|t-1}, \Sigma_{t|t-1}) \quad \text{(III.56)}$$

$$p(x_t|z_{1:t}) \approx \mathcal{N}(x_t; \hat{x}_{t|t}, \Sigma_{t|t}) \quad \text{(III.57)}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

avec :

$$\hat{x}_{t|t-1} = f_{t-1}\hat{x}_{t-1|t-1} \quad \text{(III.58)}$$

$$\Sigma_{t|t-1} = \hat{F}_{t-1}\Sigma_{t-1|t-1}\hat{F}_{t-1}^T + G_{t-1}Q_{t-1}G_{t-1}^T \quad \text{(III.59)}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t\left[z_t - h_t(\hat{x}_{t|t-1})\right] \quad \text{(III.60)}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t S_t K_t^T \quad \text{(III.61)}$$

$$K_t = \Sigma_{t|t-1}H_t^T\left[H_t\Sigma_{t|t-1}H_t^T + R_t\right]^{-1} \quad \text{(III.62)}$$

$\hat{F}_t$ et $\hat{H}_t$ sont les linéarisations locales des fonctions f_t et h_t respectivement autour de l'estimée et de la prédiction :

$$\hat{F}_t = \left.\frac{\partial f_t}{dx_t}\right|_{x_t=\hat{x}_{t|t}} \quad \text{(III.63)}$$

$$\hat{H}_t = \left.\frac{\partial h_t}{dx_t}\right|_{x_t=\hat{x}_{t|t}} \quad \text{(III.64)}$$

III.2.1.3 Filtre de Kalman sans parfum

Le filtre de Kalman sans parfum (*Unscented Kalman Filter, UKF*) , proposé par [68], procède à une approximation de la densité *a posteriori* par une gaussienne comme dans le cas du filtre de Kalman étendu. Mais plutôt que de faire une approximation des fonctions non linéaires du modèle d'évolution et du modèle de mesure, il réalise une approximation de la densité de probabilité par un ensemble de points pondérés convenablement choisis de façon déterministe. Ces points sont transformés par les fonctions non linéaires d'évolution et de mesure afin d'obtenir une nouvelle densité de probabilité. Cette approximation est appelée la transformation sans parfum (*Unscented Transform*)

III.2.1.3.a Transformée sans parfum

La transformée sans parfum est une méthode permettant de calculer les statistiques d'une variable aléatoire qui subit une transformation non linéaire [119, 70]. Elle est fondée sur l'idée qu'il est plus facile d'estimer une distribution gaussienne que d'approcher une fonction non linéaire.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Considérons le système non linéaire suivant :

$$y = f(x) \tag{III.65}$$

avec x une variable aléatoire de moyenne $\overline{x}$ et de covariance P_{xx} et y une variable aléatoire de statistiques à déterminer. Un ensemble de points est choisi de manière déterministe telle que sa moyenne et sa covariance soient respectivement $\overline{x}$ et P_{xx}. Ces points capturent la forme de la densité de probabilité de x. La fonction non linéaire f est appliquée à chacun de ces points afin d'obtenir un nuage de points transformés de moyenne $\overline{y}$ et de covariance P_{yy}.

La densité de probabilité de la variable aléatoire x de dimension n, de moyenne $\overline{x}$ et de covariance P_{xx} est approchée par $2n+1$ points pondérés donnés par :

$$\begin{array}{ll} \mathcal{X}_0 = \overline{x} & W_0 = \frac{\kappa}{n+\kappa} \\ \mathcal{X}_i = \overline{x} + \left(\sqrt{(n+\kappa)\,P_{xx}}\right)_i & W_i = \frac{1}{2(n+\kappa)} \\ \mathcal{X}_{i+n} = \overline{x} - \left(\sqrt{(n+\kappa)\,P_{xx}}\right)_i & W_{i+n} = \frac{1}{2(n+\kappa)} \end{array} \tag{III.66}$$

où $\kappa \in \Re$, $\left(\sqrt{(n+\kappa)\,P_{xx}}\right)_i$ est la $i^{ème}$ ligne ou colonne de la matrice racine carrée de $(n+\kappa)\,P_{xx}$ et W_i est le poids associé au $i^{ème}$ point.

La procédure de transformation est alors la suivante :

1. Transformer chaque point $\mathcal{X}_i$ par la fonction non linéaire f afin d'obtenir l'ensemble des points transformés

$$\mathcal{Y}_i = f(\mathcal{X}_i) \tag{III.67}$$

2. La moyenne $\overline{y}$ est donnée par la moyenne pondérée des points transformés

$$\overline{y} = \sum_{i=0}^{2n} W_i \mathcal{Y}_i \tag{III.68}$$

3. La matrice de covariance P_{yy} est donnée par :

$$P_{yy} = \sum_{i=0}^{2n} W_i (\mathcal{Y}_i - \overline{y})(\mathcal{Y}_i - \overline{y})^T \tag{III.69}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.2.1.3.b Algorithme du filtre de Kalman sans parfum

On considère les modèles d'évolution et de mesure définis comme suit :

$$x_{t+1} = f_t(x_t, v_t) \tag{III.70}$$

$$z_t = h_t(x_t) + w_t \tag{III.71}$$

où v_t et w_t sont des bruits qui affectent les processus d'évolution et de mesure, ils sont supposés décorrélés et centrés, de matrices de covariance respectives :

$$E\left[v_i v_j^T\right] = \delta_{i,j} Q_i \tag{III.72}$$

$$E\left[w_i w_j^T\right] = \delta_{i,j} R_i \tag{III.73}$$

On suppose par ailleurs que :

$$E\left[v_i v_j^T\right] = 0 \quad \forall i, j \tag{III.74}$$

On note :

$$\mathcal{X}^a = \left[(\mathcal{X}^x)^T (\mathcal{X}^v)^T\right] \tag{III.75}$$

et

$$n = \dim(x_y) + \dim(v_t) \tag{III.76}$$

L'algorithme du filtre de Kalman sans parfum est le suivant :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

1. **Calcul des points d'approximation**

$$\mathcal{X}^a_{t-1} = \left[\hat{x}^a_{t-1} \quad (\hat{x}^a_{t-1} \dots \hat{x}^a_{t-1}) + \sqrt{(n+\kappa)\Sigma^a_{t-1}} \quad (\hat{x}^a_{t-1} \dots \hat{x}^a_{t-1}) - \sqrt{(n+\kappa)\sum_{t-1}^{a}}\right] \tag{III.77}$$

avec

$$\hat{x}^a_{t-1} = \left[\hat{x}^T_{t-1|t-1} \quad 0^T_{dim(v_t-1)}\right] \tag{III.78}$$

et

$$\Sigma^a_{t-1} = \begin{pmatrix} \Sigma_{t-1|t-1} & 0 \\ 0 & Q_{t-1} \end{pmatrix} \tag{III.79}$$

2. **Calcul des poids associés**

Pour calculer les poids associés, utiliser l'équation III.66.

3. **Prédiction**

$$\mathcal{X}^x_{t|t-1} = f_{t-1}(\mathcal{X}^x_{t-1}, \mathcal{X}^v_{t-1}) \tag{III.80}$$

$$\hat{x}_{t|t-1} = \sum_{i=0}^{2n} W_i \mathcal{X}^x_{i,t|t-1} \tag{III.81}$$

$$\Sigma_{t|t-1} = \sum_{i=0}^{2n} W_i (\mathcal{X}^x_{i,t|t-1} - \hat{x}_{t|t-1})(\mathcal{X}^x_{i,t|t-1} - \hat{x}_{t|t-1})^T \tag{III.82}$$

$$\mathcal{Z}_{t|t-1} = h_t(\mathcal{X}^x_{t|t-1}) \tag{III.83}$$

$$\hat{z}_{t|t-1} = \sum_{i=0}^{2n} W_i \mathcal{Z}^x_{i,t|t-1} \tag{III.84}$$

4. **Estimation**

$$S_t = P_{v_t v_t} = R_t + \sum_{i=0}^{2n} W_i (\mathcal{Z}_{i,t|t-1} - \hat{z}_{t|t-1})(\mathcal{Z}_{i,t|t-1} - \hat{z}_{t|t-1})^T \tag{III.85}$$

$$P_{x_t z_t} = \sum_{i=0}^{2n} W_i (\mathcal{X}^x_{i,t|t-1} - \hat{x}_{t|t-1})(\mathcal{Z}^x_{i,t|t-1} - \hat{z}_{t|t-1})^T \tag{III.86}$$

$$K_t = P_{x_t z_t} S_t^{-1} \tag{III.87}$$

$$\hat{x}_{t|t} = \hat{x}_{t|t-1} + K_t(z_t - \hat{z}_{t|t-1}) \tag{III.88}$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t S_t K_t \tag{III.89}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Le filtre de Kalman sans parfum fournit une alternative intéressante au filtre de Kalman étendu et donne dans plusieurs cas de meilleurs résultats que ce dernier, pour une complexité équivalente [120]. Le filtre de Kalman sans parfum montre cependant ses limites dans le cas de systèmes fortement non linéaires et/ou non gaussiens. [69]

III.2.1.4 Autres méthodes de fusion de données pour la localisation de véhicules

Comme mentionné précédemment, d'autres méthodes de fusion de données existent, par exemple :

- Le **filtre d'information** [8, 55, 102] est une forme informationnelle du filtre de Kalman. L'état estimé $\hat{x}_{k/k}$ et sa covariance $P_{k/k}$ sont remplacés par le vecteur d'information $\hat{y}_{k/k}$ et la matrice d'information de Fisher $Y_{k/k}$.
- Le **filtrage particulaire** [3, 9, 57] est une approche statistique particulièrement adaptée aux systèmes non linéaires. En effet, elle permet d'estimer l'état d'un système non linéaire sans restriction sur la nature des non-linéarités ou sur la distribution des bruits de dynamiques et d'observation. Le principe général du filtrage particulaire repose sur une approximation de type Monte Carlo. Elle permet d'estimer l'état à l'aide d'une représentation de l'espace d'état constituée de N mesures de Dirac dont les supports X_k^i et les poids p_k^i sont conditionnés par les mesures. L'évolution des particules dans l'espace de probabilité est donnée par la loi de transition et leur pondération est évaluée par la loi d'observation.
- Dans sa thèse, Jabbour [66] a utilisé un **filtre par mélange gaussien** pour une approche de localisation sur carte multi-hypothèses. A l'approche d'une zone d'ambiguïté (par exemple un rond point), la gaussienne représentant la pose du véhicule se divise en plusieurs gaussiennes, chacune suivant une des routes de la zone d'ambiguïté. Un poids est associé à chacune des hypothèses.
- Dans [100], une méthode de fusion multi-capteurs pour la localisation précise de véhicules et le map-matching utilisant un **réseau dynamique bayésien** est présenté. Les réseaux dynamiques bayésiens permettent dans ce cas la gestion multi-hypothèses.
- Les **approches ensemblistes** sont également des approches utilisées pour la localisation de véhicules [51, 52, 67, 72, 82].

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

III.2.1.5 Conclusion

Le filtre de Kalman donne des résultats rapides et satisfaisants dans le cadre de systèmes linéaires ou faiblement non-linéaires pour le filtre de Kalman étendu. Lorsque le système devient fortement non-linéaire, ses performances se dégradent. Le filtre de Kalman sans parfum permet de pallier ce problème sous la contrainte de distributions de probabilité *a priori* et *a posteriori* gaussiennes. C'est pourquoi, dans le cadre de cette thèse, l'approche choisie pour la fusion multi-capteurs est le filtre de Kalman sans parfum.

III.2.2 Méthode de fusion multi-capteurs choisie pour la localisation de véhicules

Afin de construire l'observation cartographique 3D présentée dans le paragraphe III.1, il est nécessaire de connaître une valeur approchée de la pose du véhicule. Pour cela, un prototype de localisation par GPS hybridé a été développé. La localisation est effectuée par un suivi de pose basé sur le filtrage de Kalman sans parfum (cf § III.2.1.3). Les capteurs proprioceptifs (odomètres et gyromètre) sont utilisés dans l'étape de prédiction (cf § III.2.2.1) et cette solution obtenue est ensuite corrigée par l'information GPS (cf § III.2.2.2), si celle-ci est disponible et cohérente (le test de cohérence de la mesure GPS est décrit dans le paragraphe IV.2.2).

L'état X_k à estimer est la pose 2D du véhicule, c'est-à-dire la position 2D (x_k, y_k) et le cap θ_k :

$$X_k = \begin{pmatrix} x_k \\ y_k \\ \theta_k \end{pmatrix} \tag{III.90}$$

III.2.2.1 Etape de prédiction : modèle d'évolution odométrique

L'odométrie est une technique de localisation qui permet de déterminer la position et le cap (x, y, θ) d'un véhicule par rapport à un repère lié à sa configuration initiale, en intégrant des translations et rotations élémentaires. Il existe différentes techniques pour les mesurer et différents algorithmes pour les intégrer [39]. Nous avons retenu une configuration géométrique simple, qui consiste à utiliser les codeurs incrémentaux directement montés sur les deux roues arrières du véhicule.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Entre deux instants d'échantillonnage k et $k+1$, on mesure les distances élémentaires parcourues, notées Δard_k et Δarg_k, correspondant respectivement aux roues arrières droite et gauche.

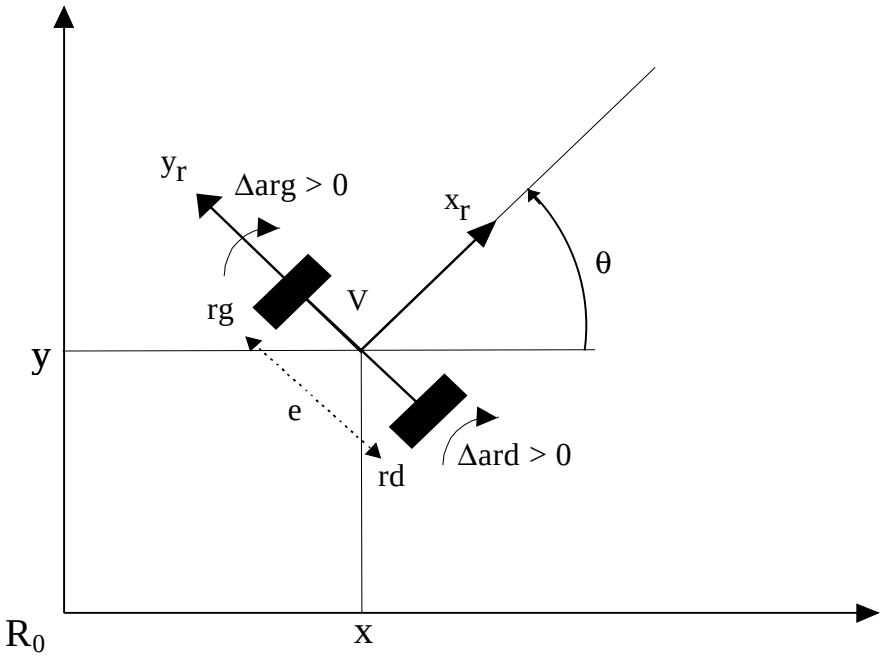


FIGURE III.13 – Modélisation 2D du train arrière du CyCab

La distance séparant les deux points de contacts des roues avec le sol est notée e (cf Figure III.13). Le centre de l'essieu (noté V) est le centre du repère R_V lié au véhicule que l'on cherche à localiser. Ce point remarquable se situe sur la droite reliant les points de contact des roues avec le sol. Ces contacts avec le sol, s'il n'y a ni patinage ni dérapage (c'est-à-dire pas de glissement), font que la vitesse du point V est perpendiculaire à l'axe reliant les deux roues. Celle-ci est portée par l'axe (V, xr). Quand le mobile se déplace, s'il n'y a pas glissement, la translation élémentaire Δ_k et la rotation élémentaire w_k du point M sont données par :

$$\begin{cases} \Delta_k = \frac{\Delta ard_k + \Delta ard_k}{2} \\ w_k = \frac{\Delta ard_k - \Delta ard_k}{2e} \end{cases} \quad \text{(III.91)}$$

Cependant, dans le cas de notre véhicule CyCab, nous n'avons pas accès directement à Δard_k et Δarg_k, mais uniquement à Δ_k. Nous ne pouvons donc pas utiliser la deuxième relation de III.91 pour calculer la rotation élémentaire. Notre véhicule étant également équipé d'un gyromètre, la rotation élémentaire sera alors donné par celui-ci.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

L'« entrée du filtre » sera notée :

$$U_k = \begin{bmatrix} \Delta_k \\ w_k \end{bmatrix} \tag{III.92}$$

Si on divise ces mesures par la période d'échantillonnage T_e et si T_e est « petite », on obtient les vitesses linéaire et angulaire du véhicule. Cependant, l'odomètre est bien un capteur de déplacement et non pas un capteur de vitesse. On cherche à exprimer de manière récurrente la position et le cap du véhicule à l'instant $k+1$ en fonction de la position et du cap à l'instant k et des déplacements élémentaires mesurés. Sous l'hypothèse d'un mouvement circulaire, on a :

$$\Delta = \rho w \tag{III.93}$$

où ρ représente le rayon de courbure (cf Figure III.14)

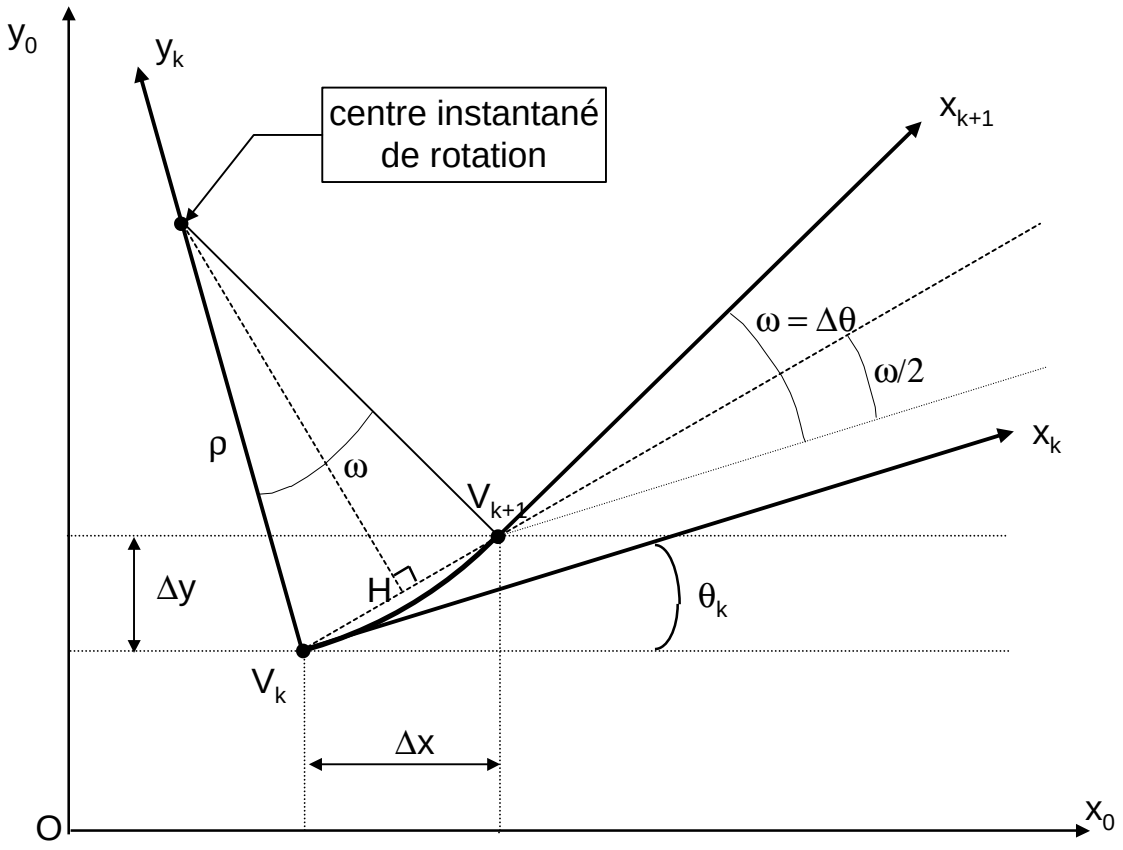


FIGURE III.14 – Déplacement circulaire entre deux instants d'échantillonnage

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

On a :

$$\|V_k H\| = \|H V_{k+1}\| \approx \rho \sin(w/2) \tag{III.94}$$

et

$$\|V_k V_{k+1}\| \approx 2\rho \sin(w/2) = 2\frac{\Delta}{w}\sin(w/2) = \Delta\frac{\sin(w/2)}{w/2} \tag{III.95}$$

Les variations Δx et Δy sont définies par le vecteur $V_k V_{k+1}$, dont l'angle avec l'horizontale est donné par :

$$(\widehat{\overrightarrow{V_k V_{k+1}}, \overrightarrow{OX_0}}) = \theta_k + w/2 \tag{III.96}$$

On obtient alors :

$$\begin{cases} x_{k+1} &= x_k + \frac{\sin(w_k/2)}{w_k/2}\Delta_k \cos(\theta_k + w_k/2) \\ y_{k+1} &= y_k + \frac{\sin(w_k/2)}{w_k/2}\Delta_k \sin(\theta_k + w_k/2) \\ \theta_{k+1} &= \theta_k + w_k \end{cases} \tag{III.97}$$

Si de plus w_k est petit, ce qui suppose que la période d'échantillonnage est suffisamment « petite »par rapport aux dynamiques du véhicule, alors nous obtenons les relations suivantes à l'ordre deux :

$$\begin{cases} x_{k+1} &= x_k + \Delta_k \cos(\theta_k + w_k/2) \\ y_{k+1} &= y_k + \Delta_k \sin(\theta_k + w_k/2) \\ \theta_{k+1} &= \theta_k + w_k \end{cases}$$

III.2.2.2 Etape de mise à jour : observation GPS

La fusion des informations issues des capteurs proprioceptifs avec les informations issues du récepteur GPS est effectuée par un filtre de Kalman sans parfum qui utilise le mécanisme du prédicteur/estimateur. Dans la phase de prédiction, le véhicule évolue en utilisant les équations III.98 et la covariance de l'erreur est estimée.

Quand une mesure GPS est disponible et cohérente, une correction de la pose prédite est effectuée (le test de cohérence de la mesure GPS est décrit dans le paragraphe IV.2.2). Dans les milieux urbains, le GPS souffre de multi-trajets et d'une

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

mauvaise constellation de satellites (canyons urbains). L'erreur GPS n'est pas stationnaire. Elle peut être estimée en temps réel en utilisant la trame NMEA « GST ». Cette information contient les paramètres qui définissent un ellipsoïde d'imprécision qui détermine une zone qui a une vraisemblance de 40% de contenir la vraie solution sous une hypothèse gaussienne. Soit $Q_k^{\beta,gps}$ la matrice de covariance de l'erreur GPS. $Q_k^{\beta,gps}$ est reconstruite à partir des données de la trame GST.

Les mesures du GPS sont relativement simples à manipuler parce que ce sont des points dans un système de coordonnées connu. L'équation d'observation linéaire est de la forme :

$$Y_k = \begin{bmatrix} x_k^{gps} \\ y_k^{gps} \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ \theta_k \end{pmatrix} + \beta_{gps} \tag{III.98}$$

La matrice de covariance $Q_k^{\beta,gps}$ du GPS est de la forme :

$$Q_k^{\beta,gps} = \begin{pmatrix} \sigma_{x,gps}^2 & Q_{xy,gps} \\ Q_{xy,gps} & \sigma_{y,gps}^2 \end{pmatrix} \tag{III.99}$$

Afin d'intégrer l'observation cartographique 3D, notée $(x_k^{carto3D}, y_k^{carto3D})$, dans l'algorithme de fusion, il faudrait modifier l'équation d'observation :

$$Y_k = \begin{bmatrix} x_k^{gps} \\ y_k^{gps} \\ x_k^{carto3D} \\ y_k^{carto3D} \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ \theta_k \end{pmatrix} + \beta_m \tag{III.100}$$

Soit $Q_k^{\beta m}$ la matrice de covariance de l'erreur d'observation. Nous supposons que la position fournie par le GPS et l'observation cartographique 3D ont des perturbations non corrélées. Alors $Q_k^{\beta m}$ peut s'écrire sous la forme d'une matrice diagonale par blocs :

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

$$Q_k^{\beta,m} = \begin{pmatrix} \sigma^2_{x,gps} & Q_{xy,gps} & 0 & 0 \\ Q_{xy,gps} & \sigma^2_{y,gps} & 0 & 0 \\ 0 & 0 & \sigma^2_{x,carto3D} & Q_{xy,carto3D} \\ 0 & 0 & Q_{xy,carto3D} & \sigma^2_{y,carto3D} \end{pmatrix} = \begin{pmatrix} Q_k^{\beta,gps} & 0_{2\times2} \\ 0_{2\times2} & Q_k^{\beta,carto3D} \end{pmatrix} \quad \text{(III.101)}$$

avec :

- $Q_k^{\beta,gps}$: la matrice de covariance de l'erreur de la mesure GPS,
- $Q_k^{\beta,carto3D}$: la matrice de covariance de l'erreur de l'observation cartographique (cf § III.1.8), $Q_k^{\beta,carto3D}$ est une matrice 2x2 correspondant à l'incertitude sur la position du véhicule.

III.3 Conclusion

Ce chapitre a présenté une approche pour localiser un véhicule. Lorsque les informations GPS sont disponibles et cohérentes alors la pose est estimée par un filtre de Kalman sans parfum hybridant le GPS et les capteurs proprioceptifs (odomètres et gyromètre). Cependant, en milieu urbain, les données sont parfois peu précises voire indisponibles, en raison de la disponibilité satellitaire, du multi-trajets ...

Pour pallier à la dérive des capteurs proprioceptifs lorsque le GPS est masqué, nous avons proposé une nouvelle estimation de la pose du véhicule : l'observation cartographique 3D. Celle-ci est construite en se basant sur des amers naturels détectés dans les images 2D acquises par une caméra embarquée. Comme toute méthode de localisation sur amers, il est nécessaire de connaître une carte/un modèle des amers. La source d'information utilisée à cet effet est un modèle virtuel 3D urbain. Des points de Harris sont détectés dans l'image réelle et dans l'image virtuelle, qui correspond à une projection 2D du modèle virtuel 3D. Les points de Harris détectés dans les deux images sont mis en correspondance. Grâce à l'information de profondeur disponible au travers du modèle virtuel 3D, les coordonnées des correspondants 3D des points de Harris peuvent être calculées. Nous avons alors des appariements 2D/3D. A partir de ceux-ci, la pose de la caméra peut être estimée en minimisant l'erreur de reprojection.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Chapitre IV

Mise en œuvre expérimentale et résultats

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Résumé du chapitre

Ce chapitre est dédié à la mise en œuvre et à la validation expérimentale des méthodes proposées dans le chapitre précédent. Des essais ont été effectués sur la place Stanislas de Nancy (dont on possède le modèle virtuel 3D) afin de valider le concept de l'observation cartographique 3D.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sommaire

IV.1 Moyens expérimentaux 122
IV.2 Localisation par fusion multi-capteurs 126
IV.2.1 Résultat sur une séquence 127
IV.2.2 Test de cohérence 129
IV.3 Observation cartographique 3D 131
IV.3.1 Transformation rigide entre le GPS et la caméra 131
IV.3.2 Résultat sur une séquence 133
IV.4 Conclusion 149

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

IV.1 Moyens expérimentaux

Le véhicule CyCab

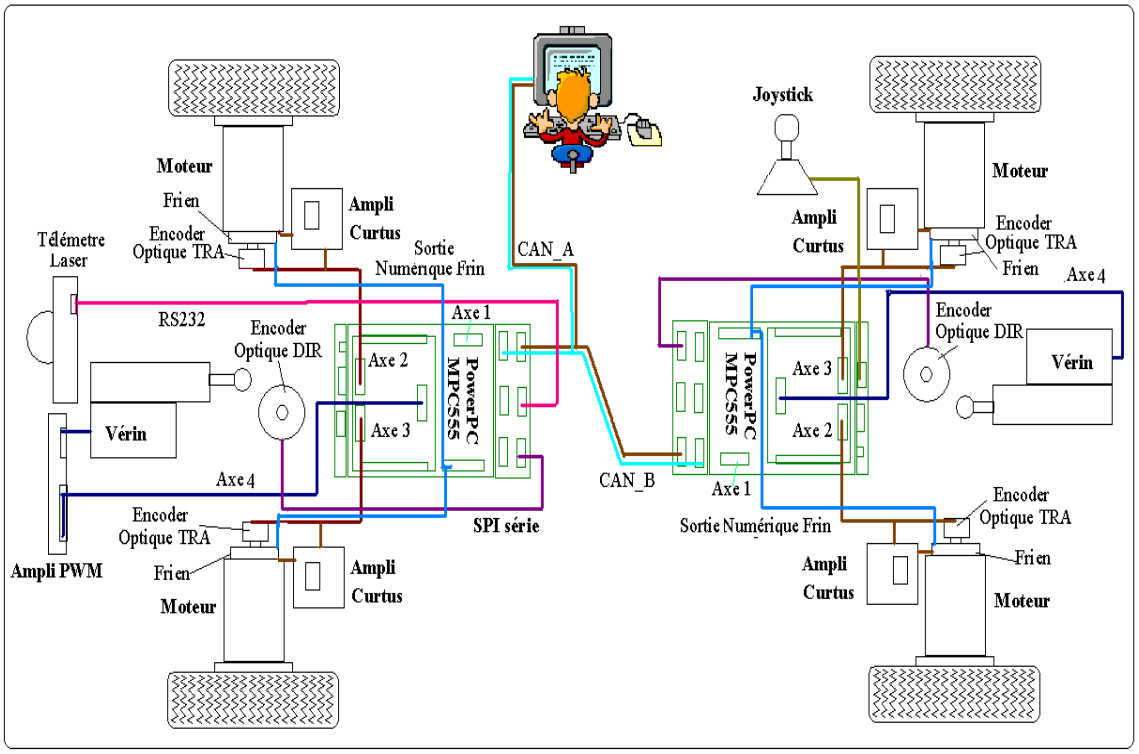
Le véhicule sur lequel ont été réalisées les expérimentations décrites dans cette thèse est un véhicule de type CyCab (cf Figure IV.1. Les CyCabs sont développés par la société Robosoft (http://www.robosoft.fr). Les CyCabs sont équipés de quatre roues motrices indépendantes, de deux vérins de direction (un pour le train avant l'autre pour le train arrière), d'un PC embarqué, de deux micro-contrôleurs Motorola MPC555 avec variateurs de puissance à commande PWM, un joystick et un écran.

FIGURE IV.1 – CyCab

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Les lois de commande sont calculées par le PC embarqué, sur lequel est installé Linux avec un noyau temps-réel RTAI. Les capteurs et actionneurs sont reliés aux deux micro-contrôleurs (un pour le train avant et l'autre pour le train arrière et le joystick). Un bus CAN sert de média de communication (cf Figure IV.2).



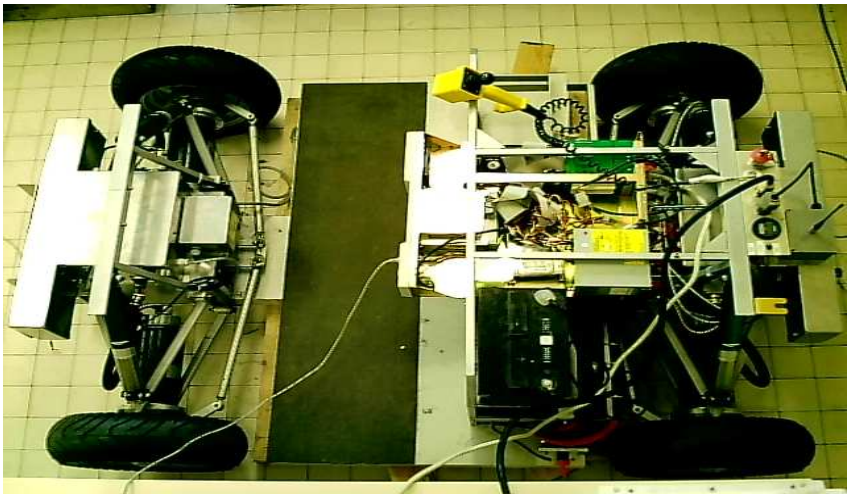


FIGURE IV.2 – Architecture distribuée du CyCab

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

La vitesse maximale du véhicule est de 18 km/h avec 2 heures d'autonomie. Deux passagers peuvent voyager à bord du véhicule. Le logiciel d'aide à l'implantation d'applications distribuées en temps-réel SYNDEX est utilisé pour la mise en oeuvre de la commande du véhicule. La force de ce logiciel est de permettre la programmation rapide et sûre des applications complexes fonctionnant de manière optimisée sur un ensemble de processeurs.

Le CyCab est livré avec des codeurs incrémentaux optiques de roues et de direction. Pour notre application, nous avons ajouté d'autres capteurs : un GPS, une caméra. Un télémètre laser a également été installé pour la détection d'obstacles (cf Chapitre V). Afin de ne pas surcharger, le PC embarqué gérant la commande temps-réel du véhicule, nous avons connecté ces capteurs à un autre PC. La configuration de ce PC, dédié à l'acquisition des données, fonctionne sous Windows.

Pour chacun des capteurs, nous avons développé un driver en C++. La plate-forme d'acquisition développée permet d'acquérir, de dater et d'enregistrer les données issues du GPS, de la caméra, des odomètres, du gyromètre. Les données sont synchronisées en utilisant l'horloge du PC dédié à l'acquisition.

Le récepteur GPS

Le récepteur GPS utilisé est le GPS Sagitta02 de Thales Navigation (cf Figure IV.3). Le récepteur Sagitta 02 est un récepteur bi-fréquence L1/L2. Il est connecté au PC par liaison série RS232. Il a une cadence de 10Hz pour les données brutes et de 20Hz pour les données calculées. Il dispose de plusieurs modes : WAAS/EGNOS, DGPS, RTK, ...Le mode RTK (*Real Time Kinematic*) dont il dispose lui permet de fournir un positionnement centimétrique, en temps réel avec une couverture radio jusqu'à 40 kilomètres. Dans ce mode, une base fixe est utilisée, elle envoie ses corrections au récepteur mobile par UHF.

FIGURE IV.3 – GPS Thales Sagitta02

Les équipements du récepteur Sagitta 02 sont donc les suivants : une station mobile, une station fixe (qui diffuse les corrections vers la station mobile), un émetteur/récepteur radio UHF, un écran/clavier (TRM 100), une antenne pour les signaux satellites GPS et une antenne UHF.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

La caméra

La caméra utilisée est une AVT MARLIN F-046C, connectée au port FIREWIRE IEEE 1394 (cf Figure IV.4). Elle délivre des images 640x480 pixels à la fréquence d'acquisition de 30Hz.

FIGURE IV.4 – Caméra AVT MARLIN F-046C

Le SIG-3D

La méthode utilise le modèle virtuel 3D urbain et le SIG-3D décrits dans le chapitre II. La Figure IV.5 montre l'interface graphique du SIG-3D.

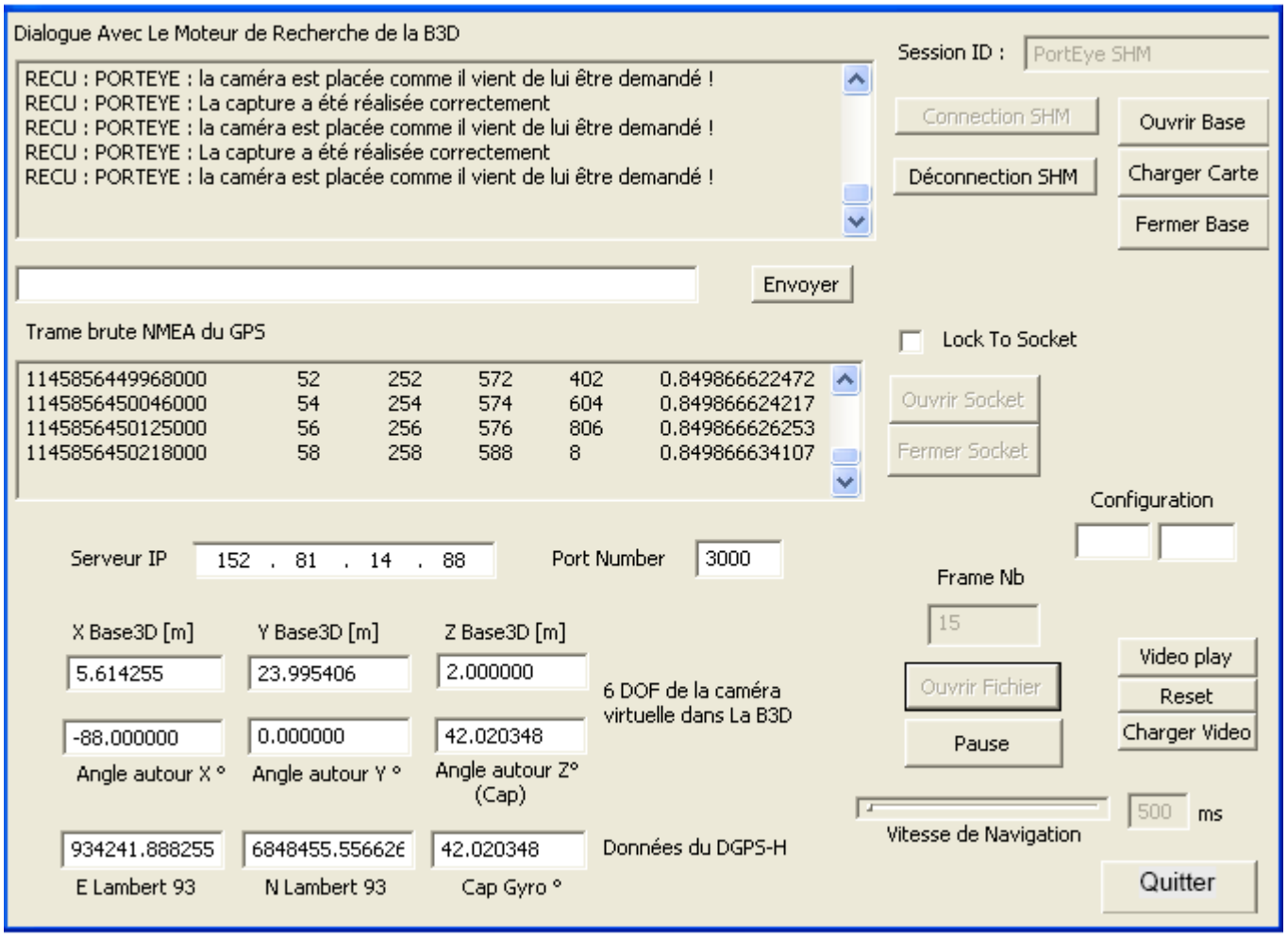


FIGURE IV.5 – Interface du SIG-3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

IV.2 Localisation par fusion multi-capteurs

Les résultats présentés ont été obtenus avec des données acquises sur la place Stanislas de Nancy. En effet, le modèle virtuel 3D urbain dont nous disposons cartographie le centre-ville de Nancy. Le CyCab est conduit manuellement sur le parcours tracé sur la Figure IV.6.

La courbe bleue représente la trajectoire acquise par le récepteur GPS Sagitta 02 en mode RTK. Le point de départ, l'arrivée et le sens de parcours sont indiqués. La croix bleue (+) située en $(0, 0)$ représente le centre du modèle virtuel 3D, c'est-à-dire la statue Stanislas.

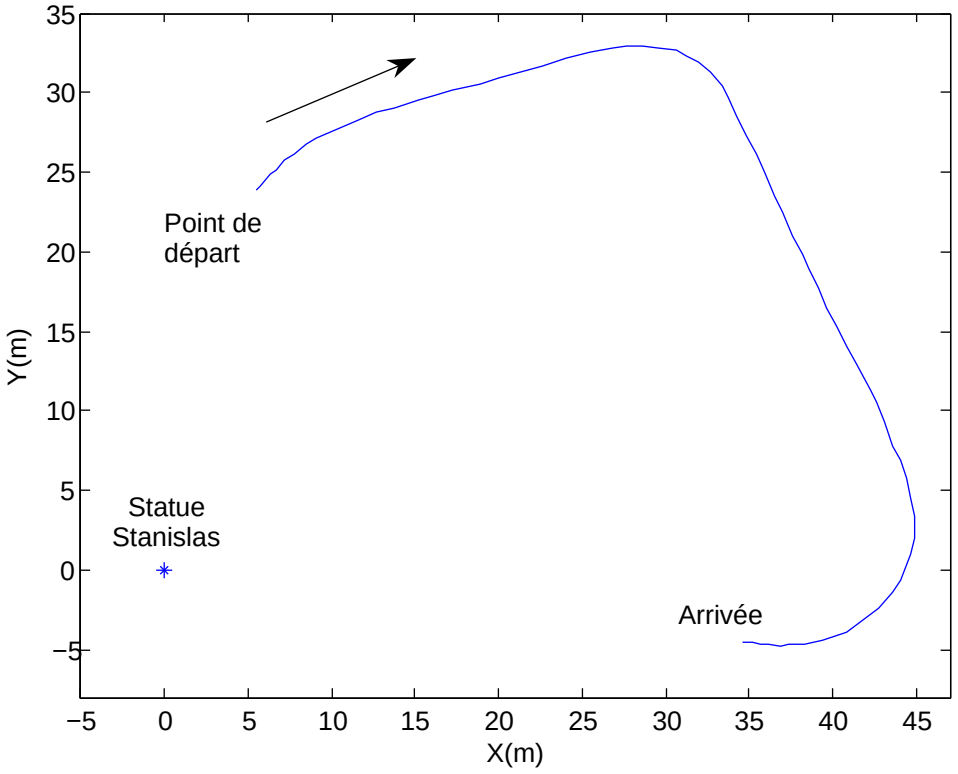


FIGURE IV.6 – Trajectoire de référence du parcours (données GPS RTK)

Comme mentionné précédemment, les informations issues du récepteur GPS RTK Sagitta 02, de la caméra AVT Marlin, des odomètres et du gyromètre sont acquises, datées et enregistrées sur un PC dédié à cet effet. La vitesse du CyCab lors de cet essai est indiquée sur la Figure IV.7.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

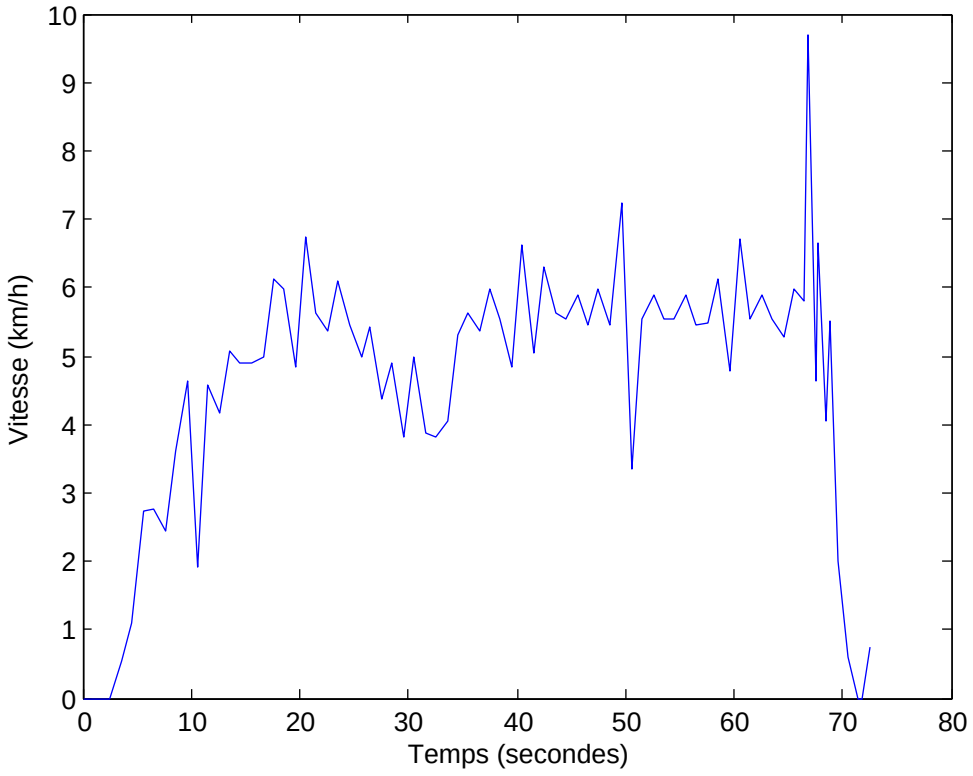


FIGURE IV.7 – Vitesse du CyCab pendant le test

IV.2.1 Résultat sur une séquence

Dans ce paragraphe, nous illustrons les résultats de la localisation par fusion multi-capteurs.

Sur la Figure IV.8, nous avons tracé en vert la trajectoire de référence acquise par le GPS RTK centimétrique. Nous avons simulé deux périodes de masquage de l'information GPS. Entre les instants 25 et 35 puis entre les instants 45 et 60, les mesures GPS ne sont donc pas utilisées pour estimer la pose du véhicule. Dans ces deux cas, la pose est alors estimée uniquement à l'aide des informations issues des capteurs proprioceptifs (odomètres et gyromètre).

La pose estimée est représentée par une croix bleue (+) lorsque les observations GPS sont utilisées et par une croix rouge (+) lorsqu'elles ne sont pas utilisées (un masquage total de l'information GPS est simulé). Quand l'information GPS est absente, nous constatons la dérive de la localisation à l'estime.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

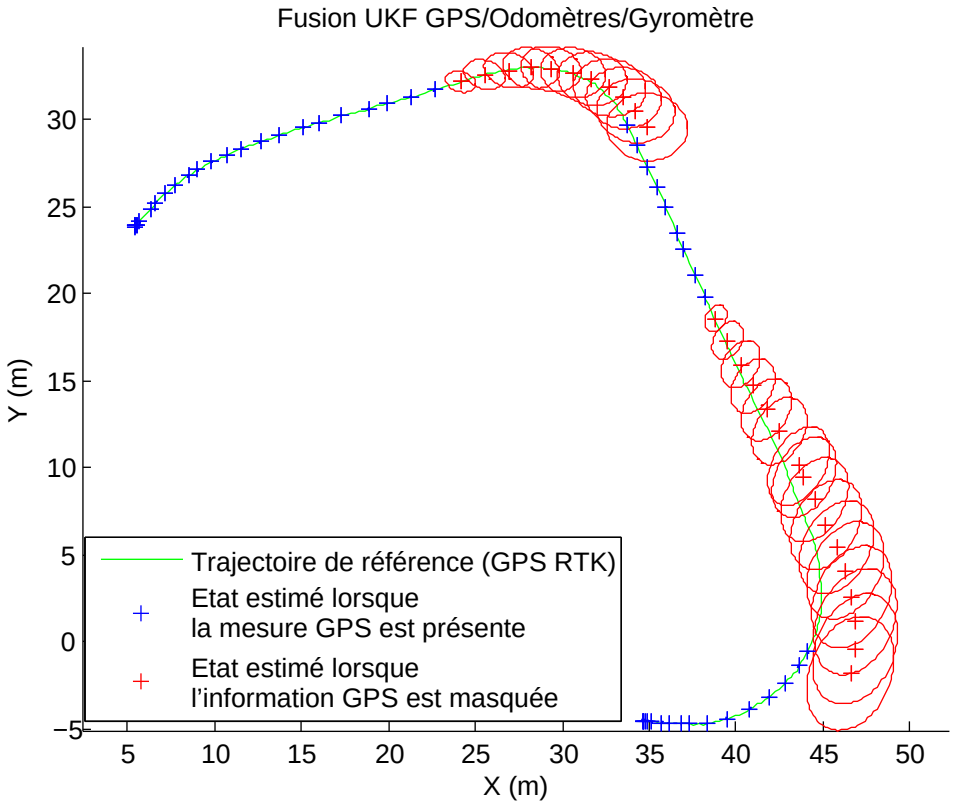


FIGURE IV.8 – Pose estimée par le filtre de Kalman sans parfum

L'ellipse d'incertitude de la position estimée est également tracée. Lorsque la mesure GPS est présente, cette ellipse est très petite, donc à peine visible sur la Figure. En effet, le GPS RTK a une précision centimètrique. Lorsque nous simulons un masquage de l'information GPS, l'ellipse d'incertitude de la pose estimée augmente progressivement.

Sur la Figure IV.9, l'erreur d'estimation de la pose est tracée. Les bandes grisées correspondent aux instants où les informations GPS sont masquées. Une nouvelle fois, nous constatons la dérive de la localisation à l'estime lorsque les données GPS sont absentes. Après 10 secondes, cette dérive est d'environ 1,7 mètres.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

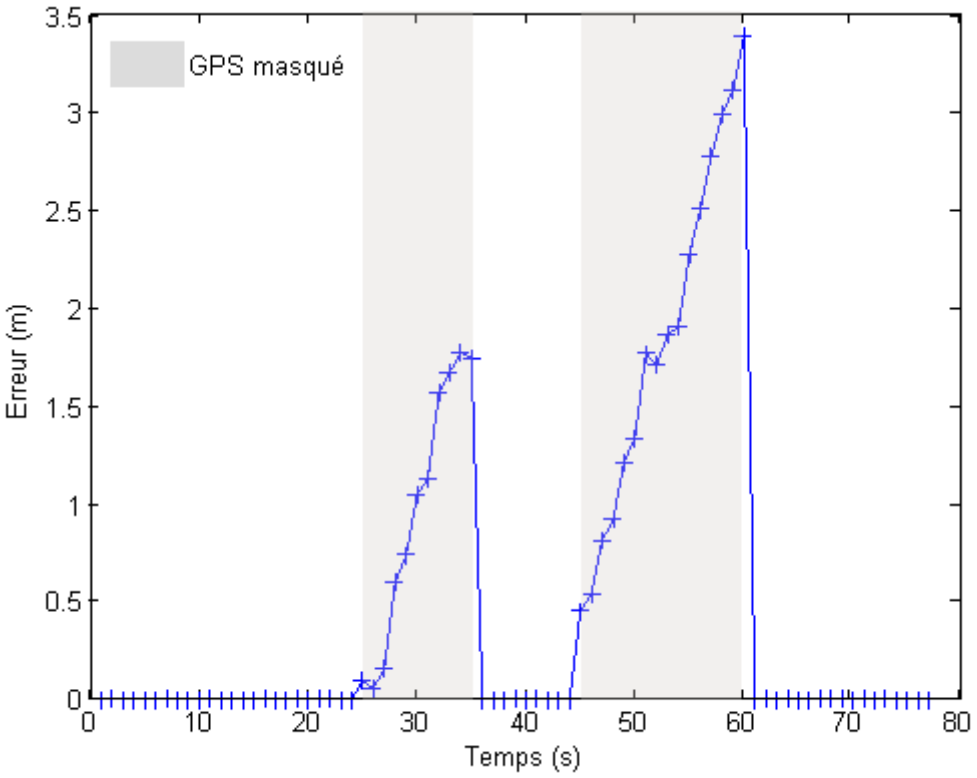


FIGURE IV.9 – Erreur de la pose estimée

IV.2.2 Test de cohérence

Comme il a été dit précédemment dans ce manuscrit, le GPS souffre de multi-trajets et d'une mauvaise visibilité satellitaire en milieu urbain. Avant de corriger la prédiction de la pose, nous avons donc vérifié la cohérence de la mesure GPS. Pour cela, le test NIS (*Normalized Innovation Squared*) qui suit une distribution du χ^2 est utilisé.

Loi du Chi-2 : Soient $X_1, \ldots, X_k$, k variables aléatoires indépendantes, normales, centrées et réduites. Alors par définition, la variable $\sum_{i=1}^{k} X_i^2$ suit une loi du χ^2 à k degrés de liberté, notée χ_k^2. L'espérance mathématique de χ_k^2 vaut k et sa variance vaut $2k$.

Quand une mesure GPS est reçue, le test NIS est appliqué. Ce test consiste à calculer la distance Δ_k entre l'observation GPS et l'état prédit :

$$\Delta_k = \nu_k S_k^{-1} \nu_k \tag{IV.1}$$

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

avec S_k la matrice de covariance de l'innovation ν_k.

La mesure GPS est cohérente, si la distance NIS est inférieure à un seuil $\chi^2_{k,1-\alpha}$ avec k la dimension de l'état, et α le degré de risque. Dans notre cas, $k = 3$ et $\alpha = 0.05$.

Donc, si la distance Δ_k est inférieure $\chi^2(0.05, 3)$ alors la mesure GPS est supposée fiable et est utilisée pour corriger la pose prédite. Sinon, la mesure GPS est rejetée et la pose prédite par le modèle d'évolution est conservée.

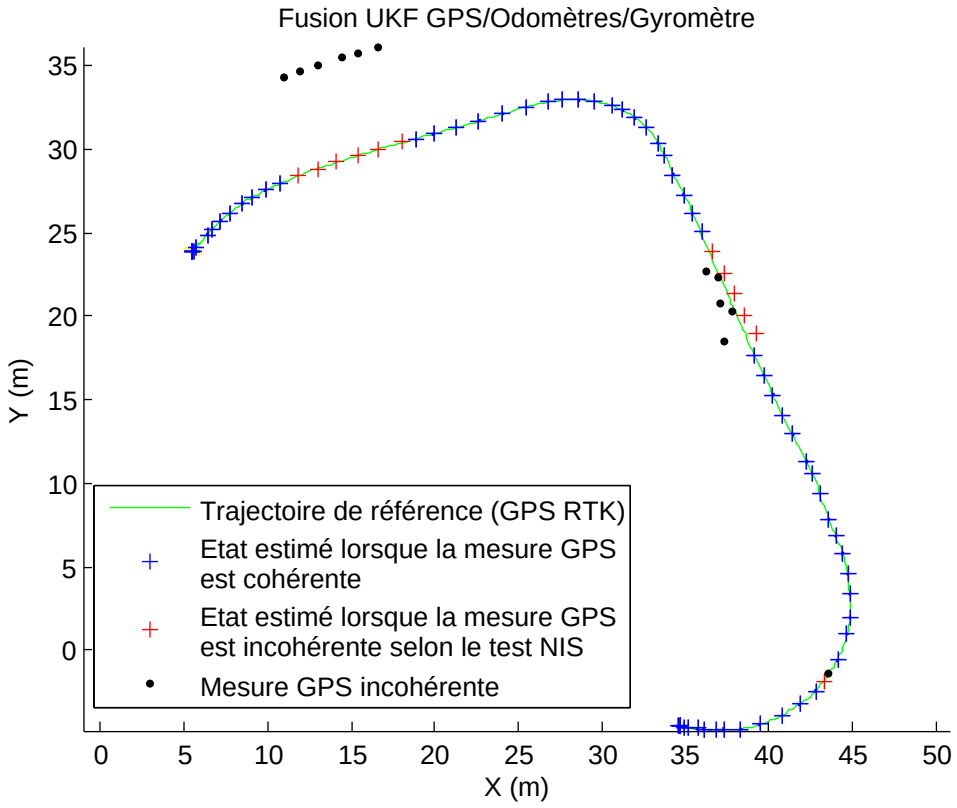


FIGURE IV.10 – Pose estimée par le filtre de Kalman sans parfum avec test de cohérence

Les résultats de ce test de cohérence sont illustrés sur les Figures IV.10 et IV.11. Nous avons volontairement dégradé la mesure GPS à deux endroits de la trajectoire, pour simuler un saut de GPS ou des mesures aberrantes. On constate que le test de cohérence a détecté l'incohérence des mesures GPS, et les a donc rejeté. L'état prédit n'est donc pas corrigé par la mesure GPS. Les croix bleues représentent les poses estimées lorsque les mesure GPS sont cohérentes. Les croix rouges représentent les poses estimées lorsque les mesure GPS sont incohérentes ; dans ce cas, les mesures GPS sont indiquées par un point noir. Le rejet des mesures GPS incohérentes permet de garder une bonne estimation de la pose, si toutefois cette période n'est pas

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

trop longue. Dans le cas contraire, l'estimation de la pose diverge puisque seule les informations issues des capteurs proprioceptifs sont utilisées.

Sur la Figure IV.11, nous avons tracé la valeur de la distance Δ_k en fonction du temps. Les valeurs de Δ_k supérieures au seuil indiqué en rouge correspondent aux mesures GPS rejetées.

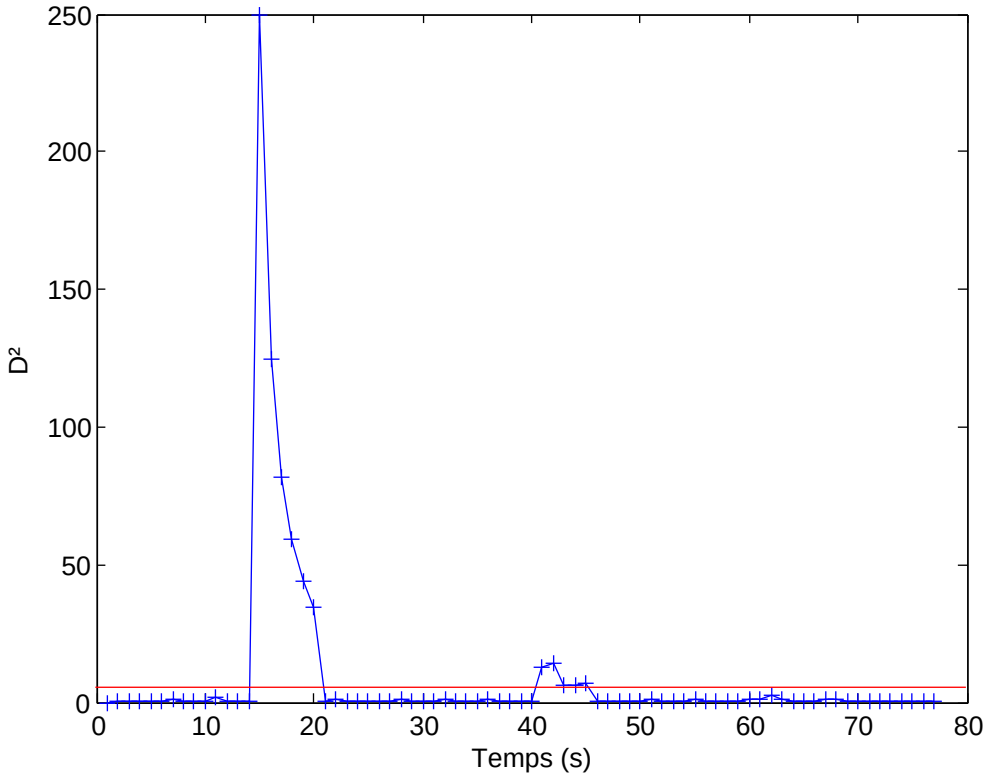


FIGURE IV.11 – Test de cohérence des mesures GPS

IV.3 Observation cartographique 3D

Dans ce paragraphe, les résultats de l'estimation de la pose à l'aide de la caméra et du modèle virtuel 3D urbain sont présentés.

IV.3.1 Transformation rigide entre le GPS et la caméra

La caméra et le récepteur GPS n'étant pas placés au même endroit sur le CyCab (cf Figure IV.12), il faut calculer le changement de repère entre le repère de la caméra et les coordonnées GPS. En effet, ils n'enregistrent pas la même trajectoire.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

L'antenne GPS est située à l'aplomb du milieu de l'essieu arrière. La caméra a été placée à l'avant pour que le toit du CyCab ne vienne pas masquer une partie de la scène dans l'image.

FIGURE IV.12 – Positionnement des capteurs sur le CyCab

Le changement de repère entre repère caméra et repère robot a été mesuré manuellement approximativement. La Figure IV.13 permet d'introduire les notations qui définissent les différents repères utilisés. Le repère monde est noté $R_W = (O_W, X_W, Y_W, Z_W)$. L'axe Z_W est vertical et dirigé vers le haut. Le CyCab évolue dans le repère R_W, sa position est définie comme la position du point milieu de l'essieu arrière noté O_V. Le repère attaché au CyCab permet de définir la position des capteurs sur le véhicule. Ce repère est appelé repère véhicule et est noté $R_V = (O_V, X_V, Y_V, Z_V)$. Son origine est aussi le milieu de l'essieu arrière. L'axe Y_V est donné par la direction d'avancement du véhicule. L'antenne GPS est placée à la verticale de l'origine du repère robot. Le repère caméra R_C est défini de façon classique en vision. L'origine de R_C est au centre optique de la caméra O_C et il est orienté pour avoir Z_C selon l'axe optique de la caméra. Les axes X_C et Y_C correspondent aux axes de la matrice CCD.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

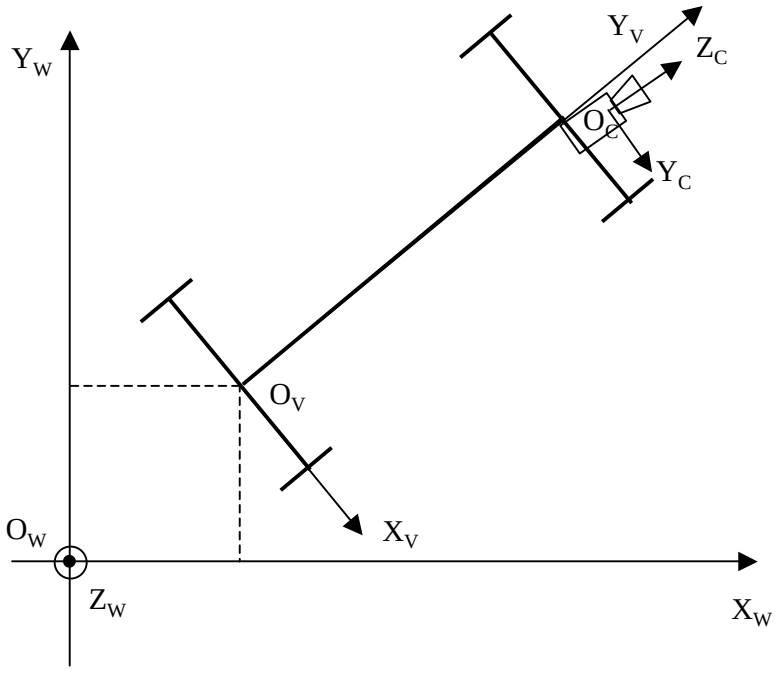


FIGURE IV.13 – Les différents repères des capteurs

A ce stade, on dispose de la trajectoire du même point du véhicule enregistré par les capteurs chacun dans leur repère propre. L'échantillonnage des capteurs est différent. Cependant, lorsque les données des capteurs (GPS, Odomètre, Gyromètre, Caméra) sont enregistrées, elles sont datées (par rapport à l'horloge PC) et une interpolation linéaire sur la trajectoire GPS est suffisante pour obtenir des données synchronisées.

IV.3.2 Résultat sur une séquence

Nous illustrons dans ce paragraphe les résultats du calcul de l'observation cartographique. Ces résultats correspondent au même essai que celui du paragraphe précédent (IV.2).

Rappelons brièvement la méthode de construction de l'observation cartographique 3D. La première consiste à acquérir l'image réelle acquise par la caméra embarquée préalablement calibrée ainsi que l'image virtuelle. Puis les points de Harris sont détectés dans chacune de ces deux images. Ensuite ces points de Harris sont appariés en utilisant la corrélation croisée normalisée centrée. Afin d'éliminer les

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

faux-appariements, l'algorithme RANSAC est appliqué. Parallèlement, les coordonnées des correspondants 3D des points de Harris de l'image virtuelle sont calculées. Finalement, la pose est calculée en minimisant l'erreur de reprojection des appariements 2D/3D ainsi obtenus.

Sur les Figures IV.14 à IV.24, nous illustrons les résultats obtenus après chaque étape de l'algorithme pour quelques instants d'échantillonnage de la séquence.

Sur ces 11 Figures, nous avons tracé :

- sur la première ligne :
  - à gauche : l'image réelle
  - à droite : l'image virtuelle correspondante

- sur la deuxième ligne :
  - à gauche : l'image réelle avec les points de Harris détectés en surimpression
  - à droite : l'image virtuelle avec les points de Harris détectés en surimpression

- sur la troisième ligne :
  - à gauche : les appariements 2D/2D entre les points de Harris de l'image réelle et les points de Harris de l'image virtuelle en surimpression de l'image réelle. L'image réelle est affichée et en surimpression, nous affichons les appariements 2D/2D sous la forme de segments bleus. Une extrémité du segment correspond à la position du point apparié dans l'image réelle, l'autre extrémité correspond la position de son correspondant dans l'image virtuelle.
  - à droite : la reconstruction des points 3D de la scène (correspondants à des points de Harris de l'image virtuelle) à partir de l'image virtuelle, de la pose de la caméra virtuelle dans le modèle 3D et de l'information de profondeur

- sur la quatrième ligne :
  - les appariements 2D/2D après élimination des outliers (faux-appariements) par l'algorithme de RANSAC. La partie gauche correspond à l'image réelle, la partie droite correspond à l'image virtuelle. Nous avons relié les points appariés de l'image réelle à leurs correspondants dans l'image virtuelle par des segments rouges.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

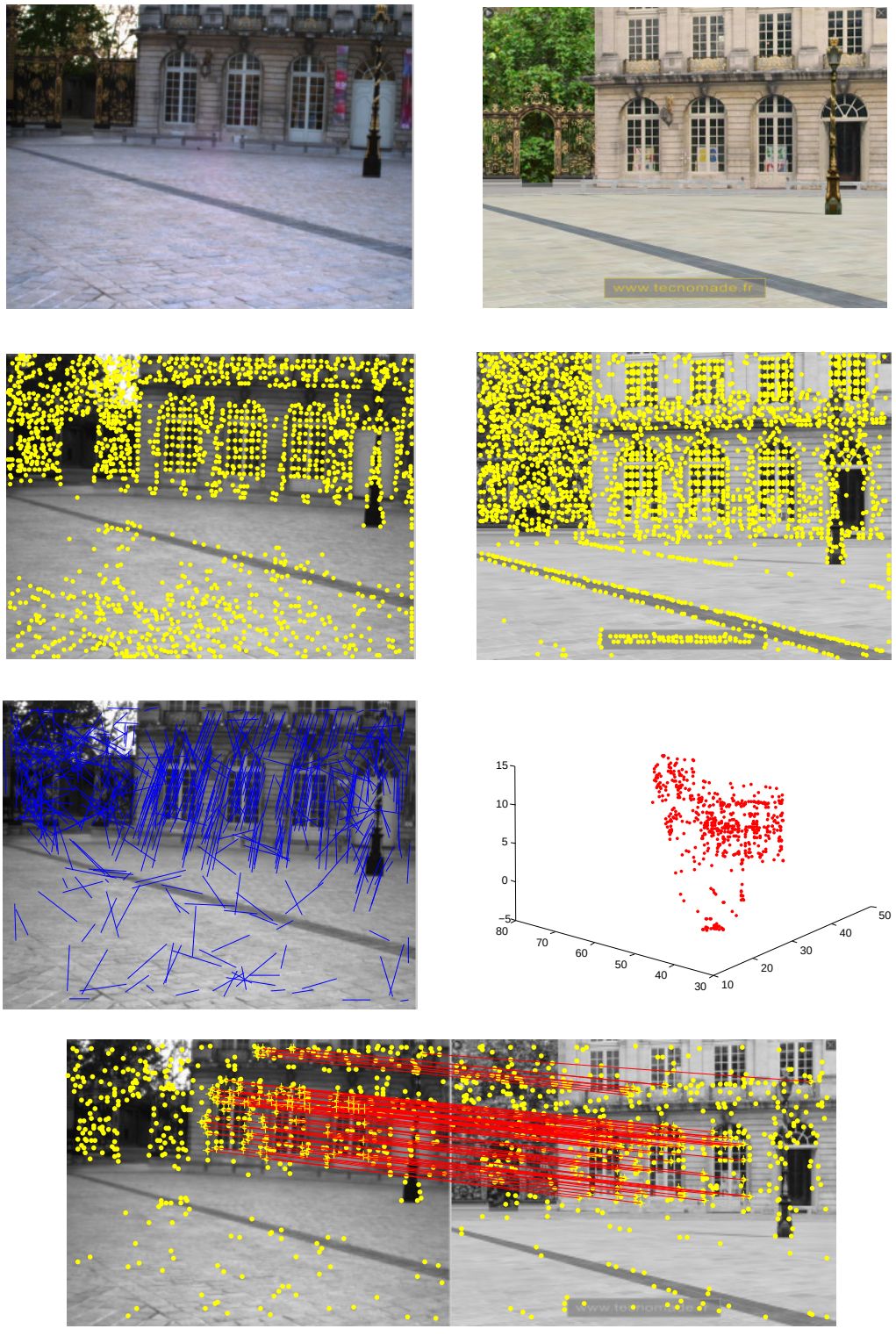


FIGURE IV.14 – Observation cartographique 3D - Résultat 1

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

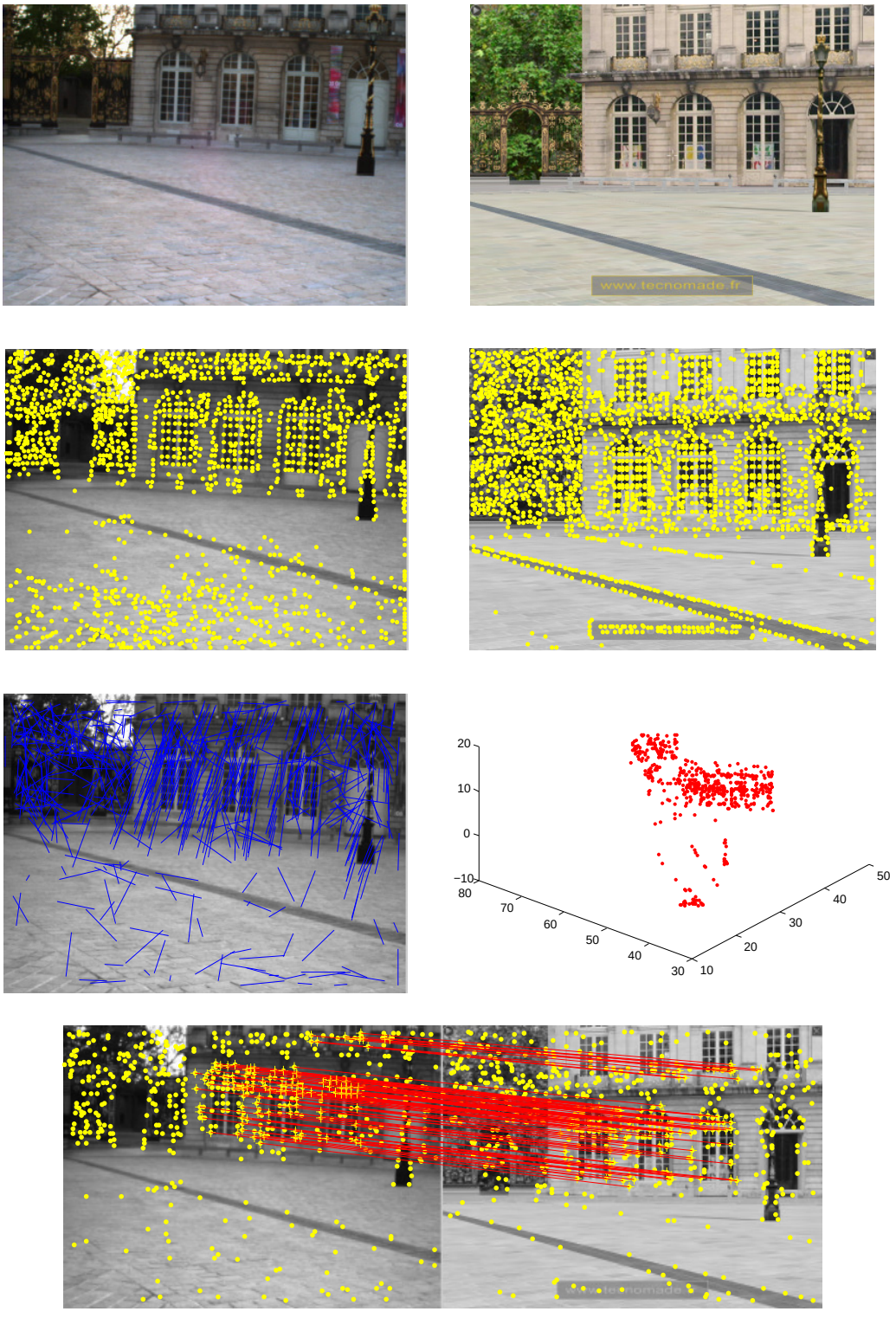


FIGURE IV.15 – Observation cartographique 3D - Résultat 2

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

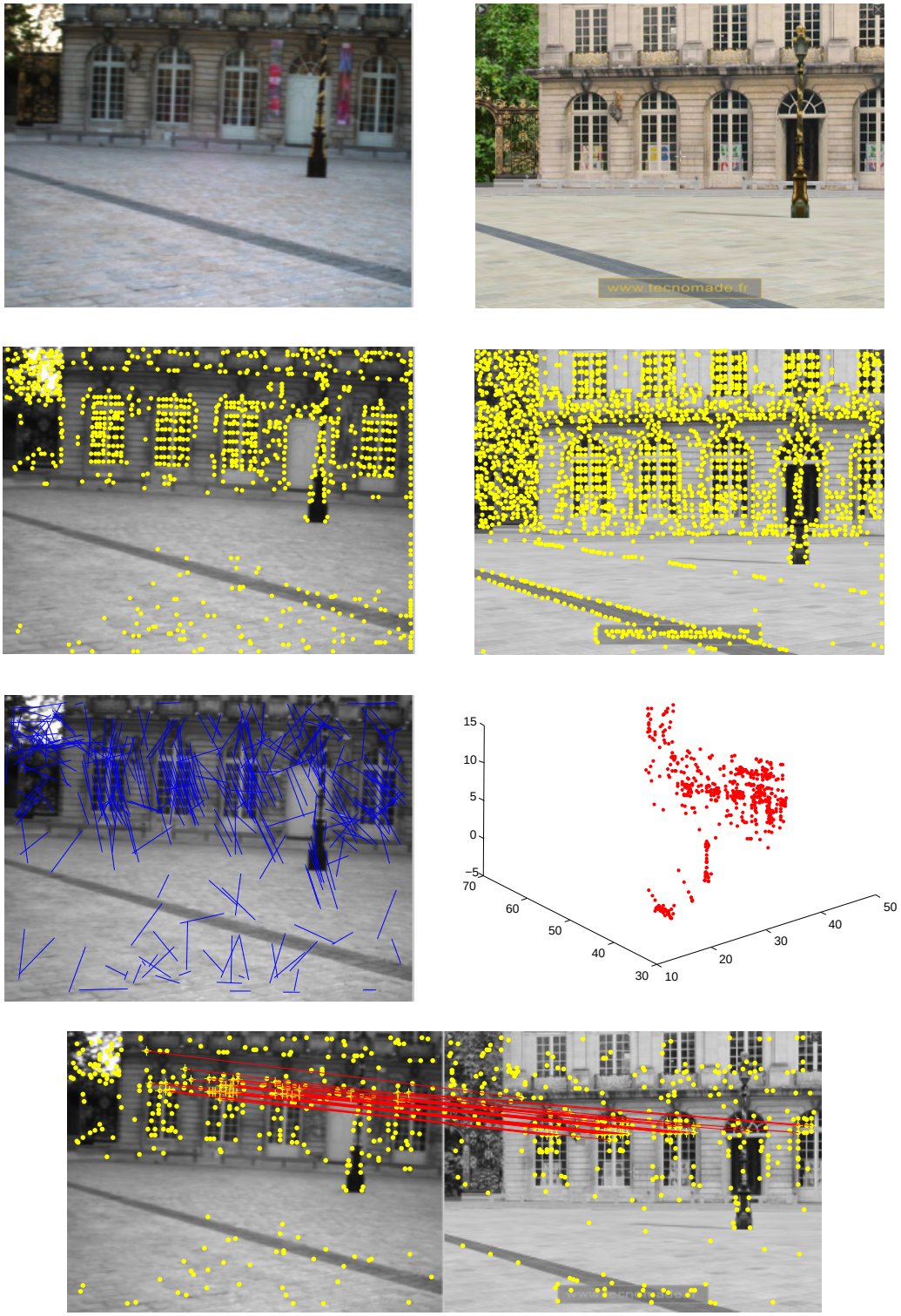


FIGURE IV.16 – Observation cartographique 3D - Résultat 3

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

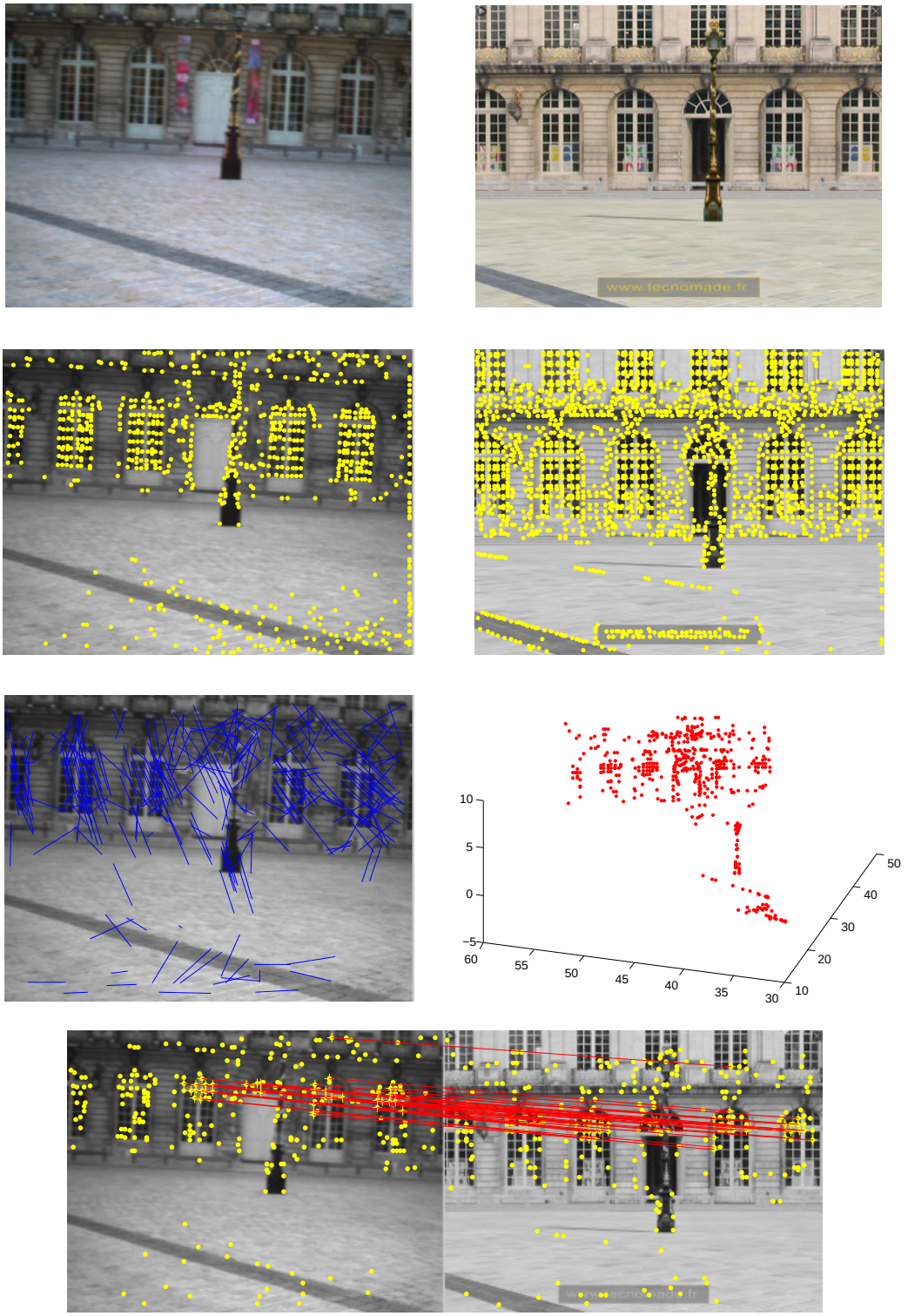


FIGURE IV.17 – Observation cartographique 3D - Résultat 4

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

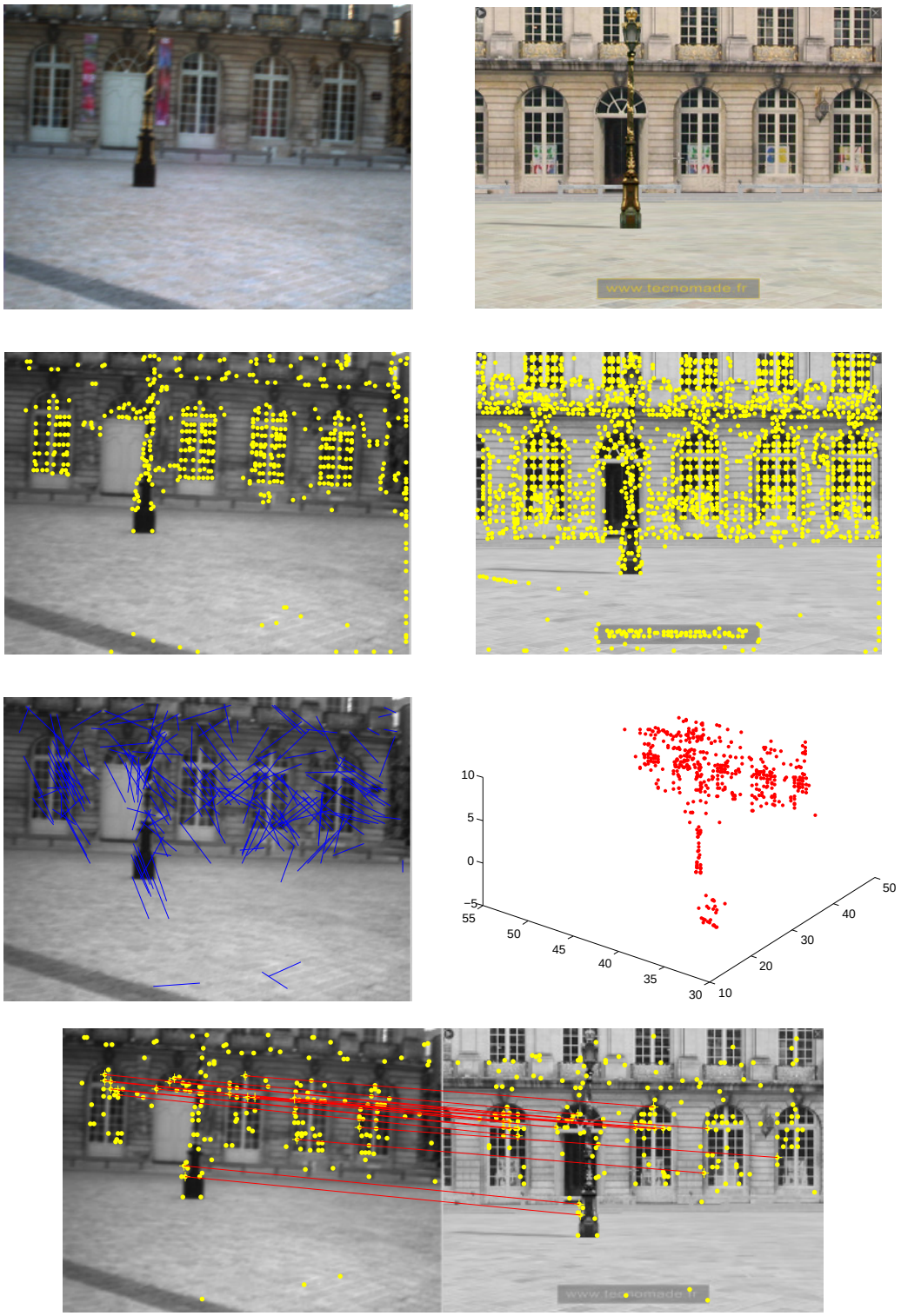


FIGURE IV.18 – Observation cartographique 3D - Résultat 5

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

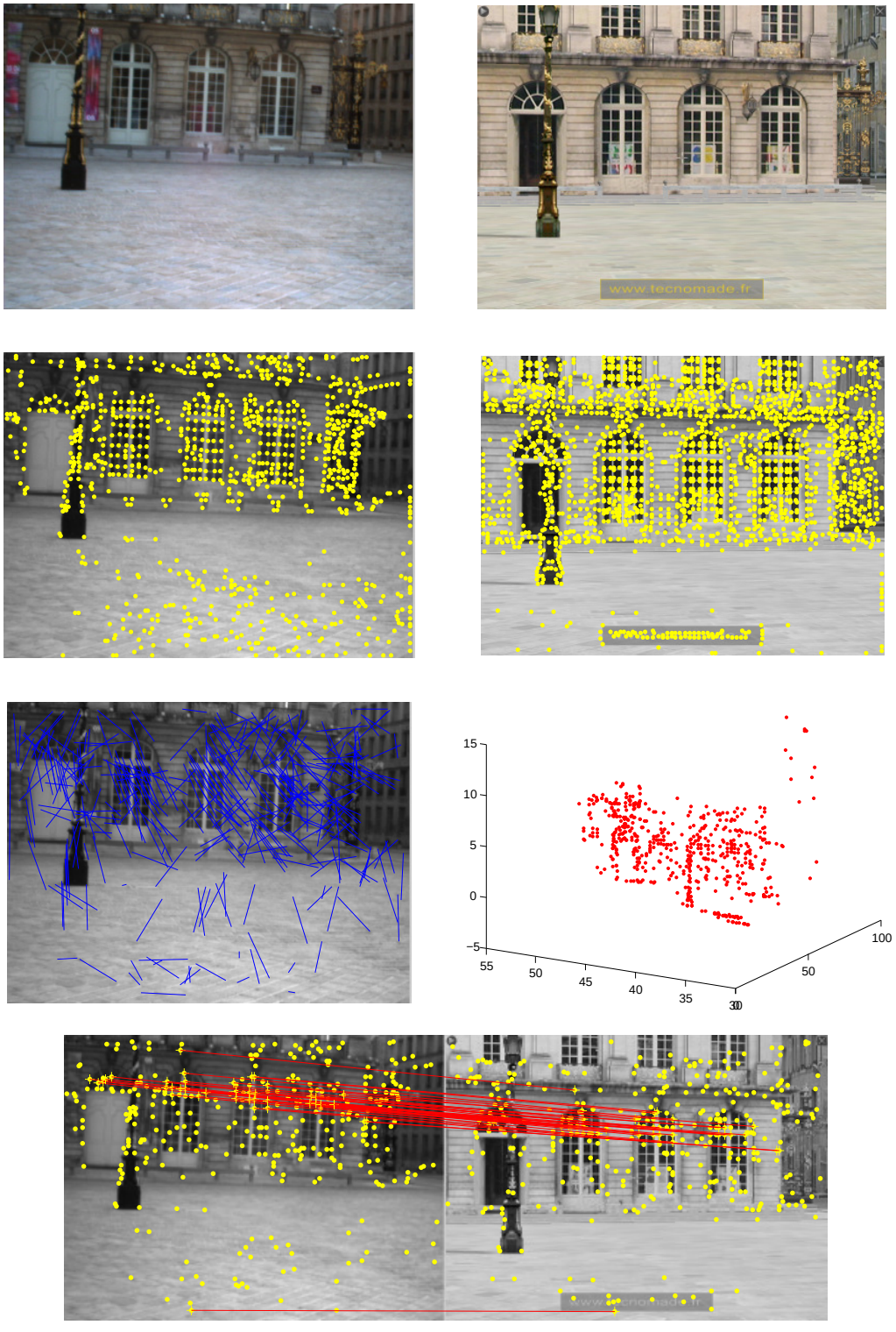


FIGURE IV.19 – Observation cartographique 3D - Résultat 6

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

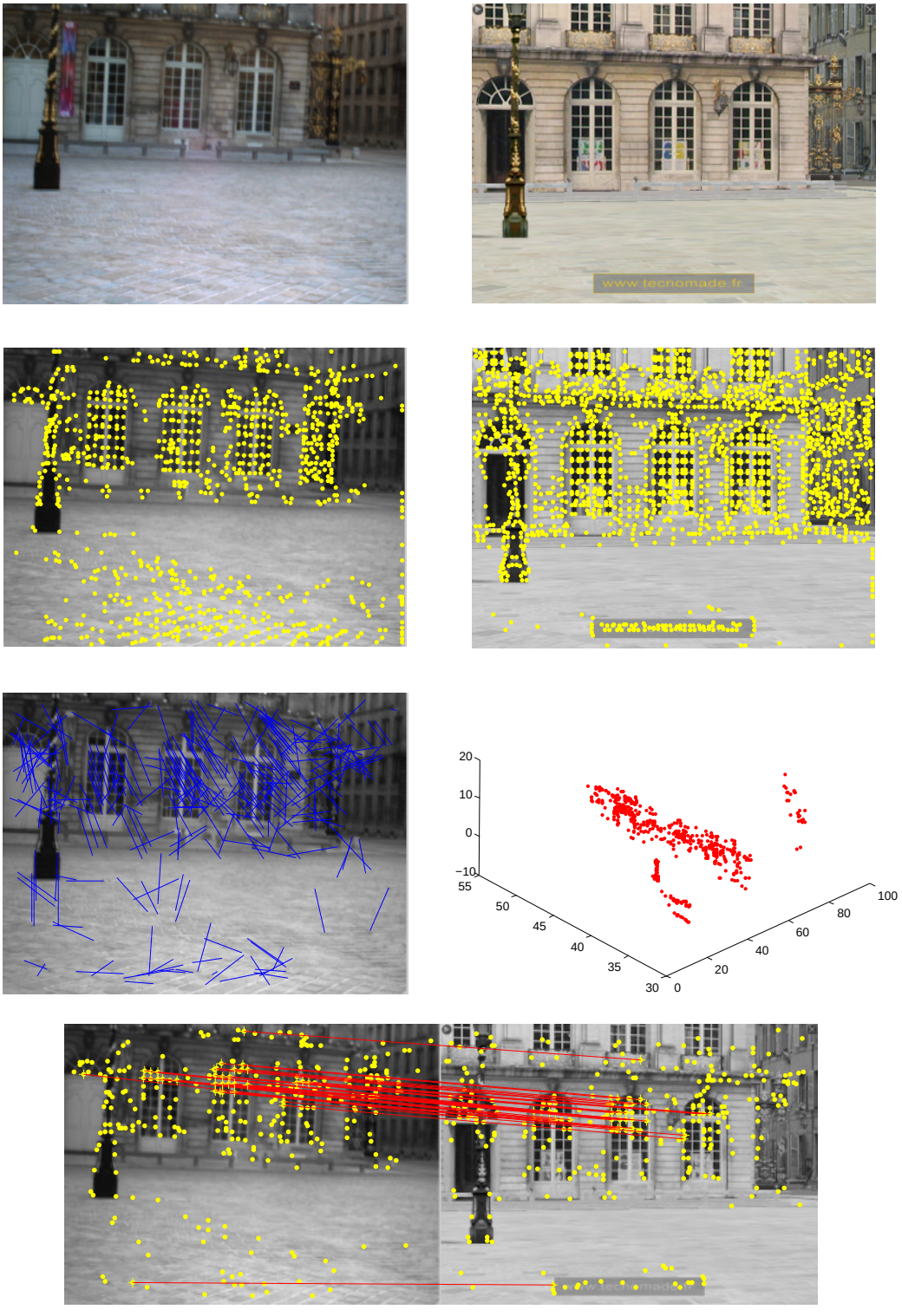


FIGURE IV.20 – Observation cartographique 3D - Résultat 7

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

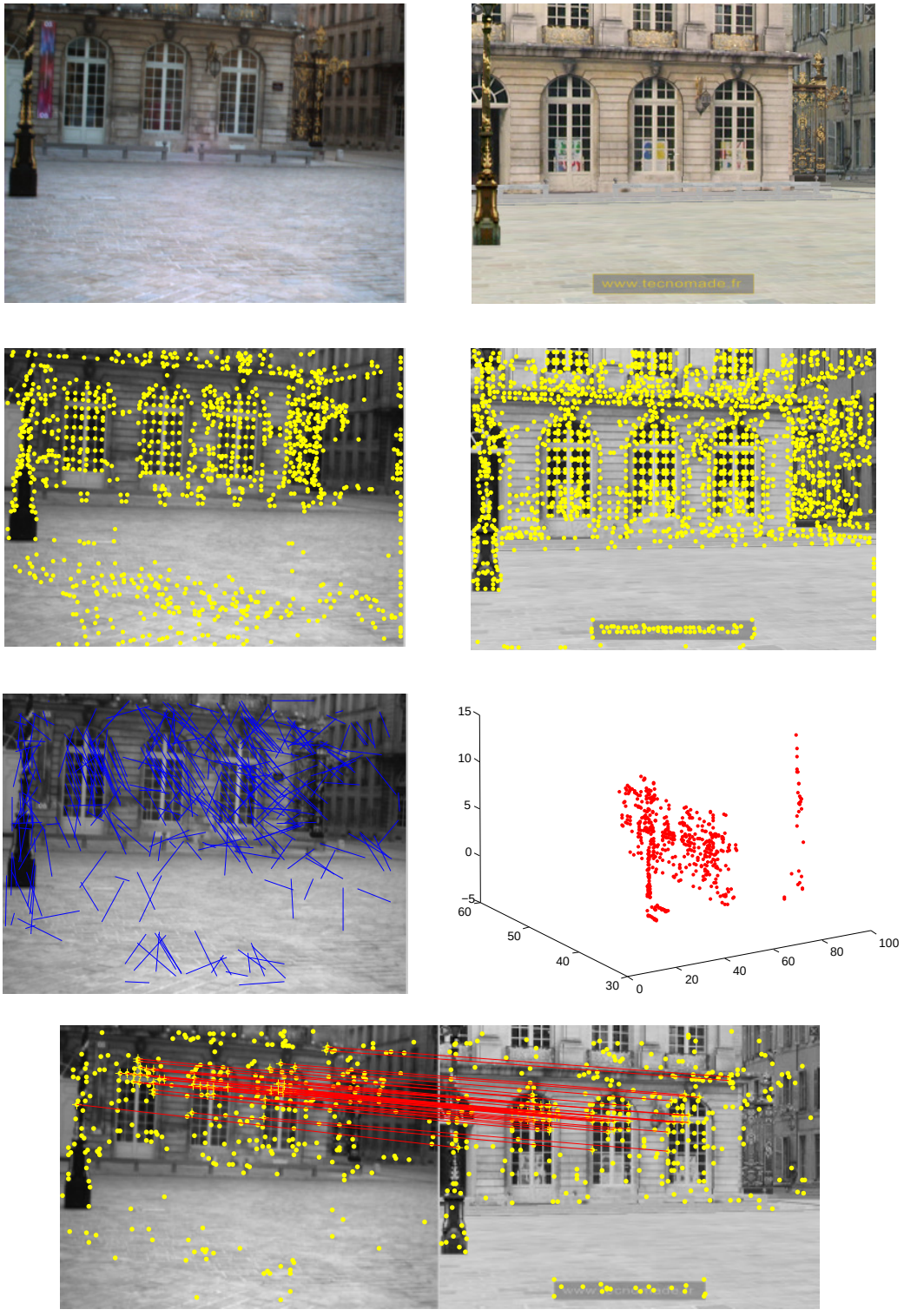


FIGURE IV.21 – Observation cartographique 3D - Résultat 8

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

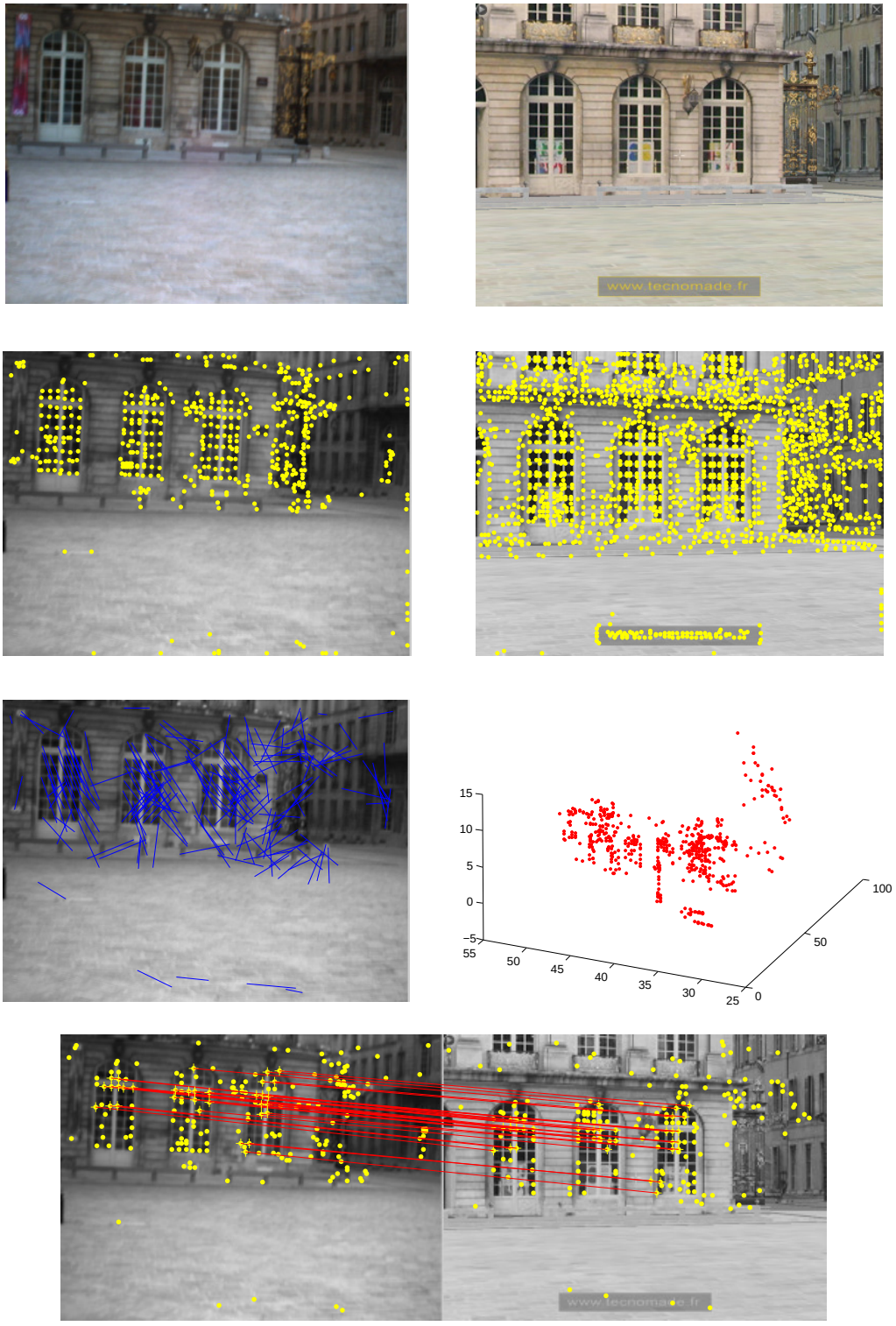


FIGURE IV.22 – Observation cartographique 3D - Résultat 9

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

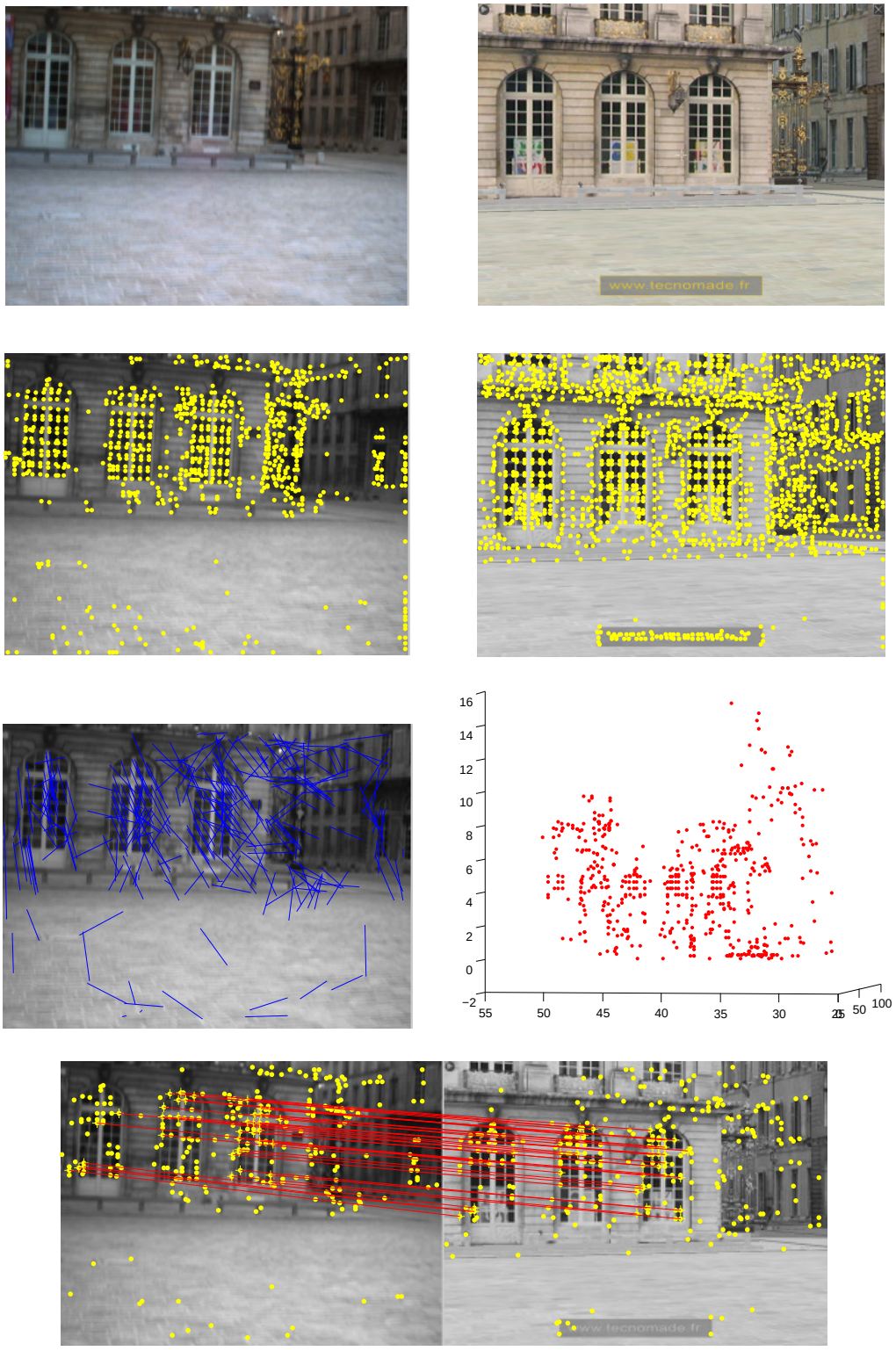


FIGURE IV.23 – Observation cartographique 3D - Résultat 10

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

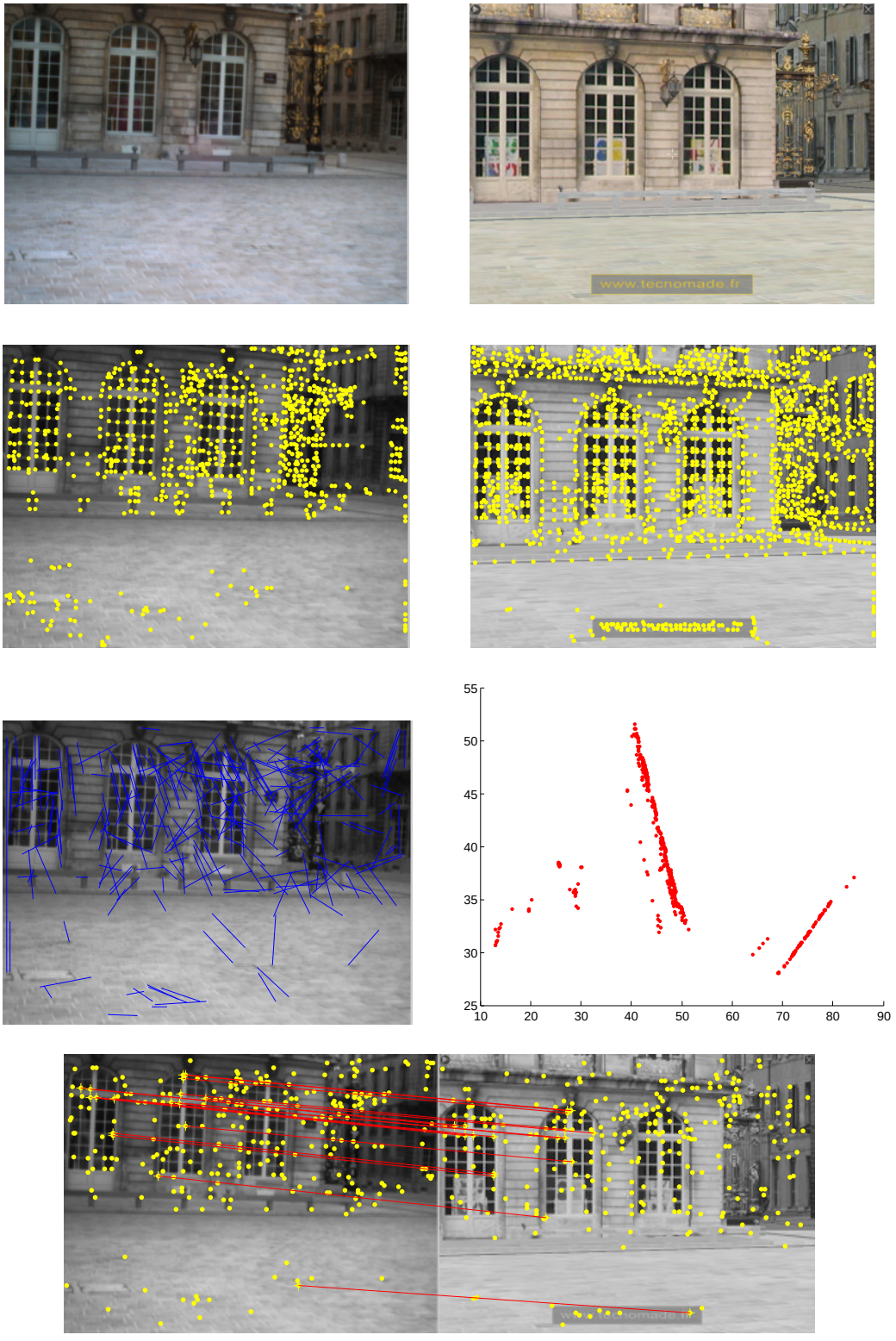


FIGURE IV.24 – Observation cartographique 3D - Résultat 11

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Nous pouvons remarquer sur les Figures précédentes, l'importance de l'algorithme de RANSAC pour éliminer les faux-appariements. Cependant, l'algorithme de RANSAC est parfois inefficace. La Figure IV.25 illustre le problème des appariements erronés non détectés. Si l'estimation *a priori* de la pose est trop éloignée de la pose réelle, alors la majorité des appariements sont faux, et l'algorithme RANSAC d'élimination des *outliers* ne fonctionnent pas. Dans les deux cas présentés, nous avons volontairement décalés la pose *a priori*, on constate alors que les appariements sont décalés d'une fenêtre. Nous avons donc instauré un test : si l'incertitude sur la pose *a priori* utilisée pour le calcul de l'observation cartographique 3D est trop grande, alors on ne calcule pas cette observation cartographique 3D, le risque de faux appariements non détectés par RANSAC étant en effet trop grand.

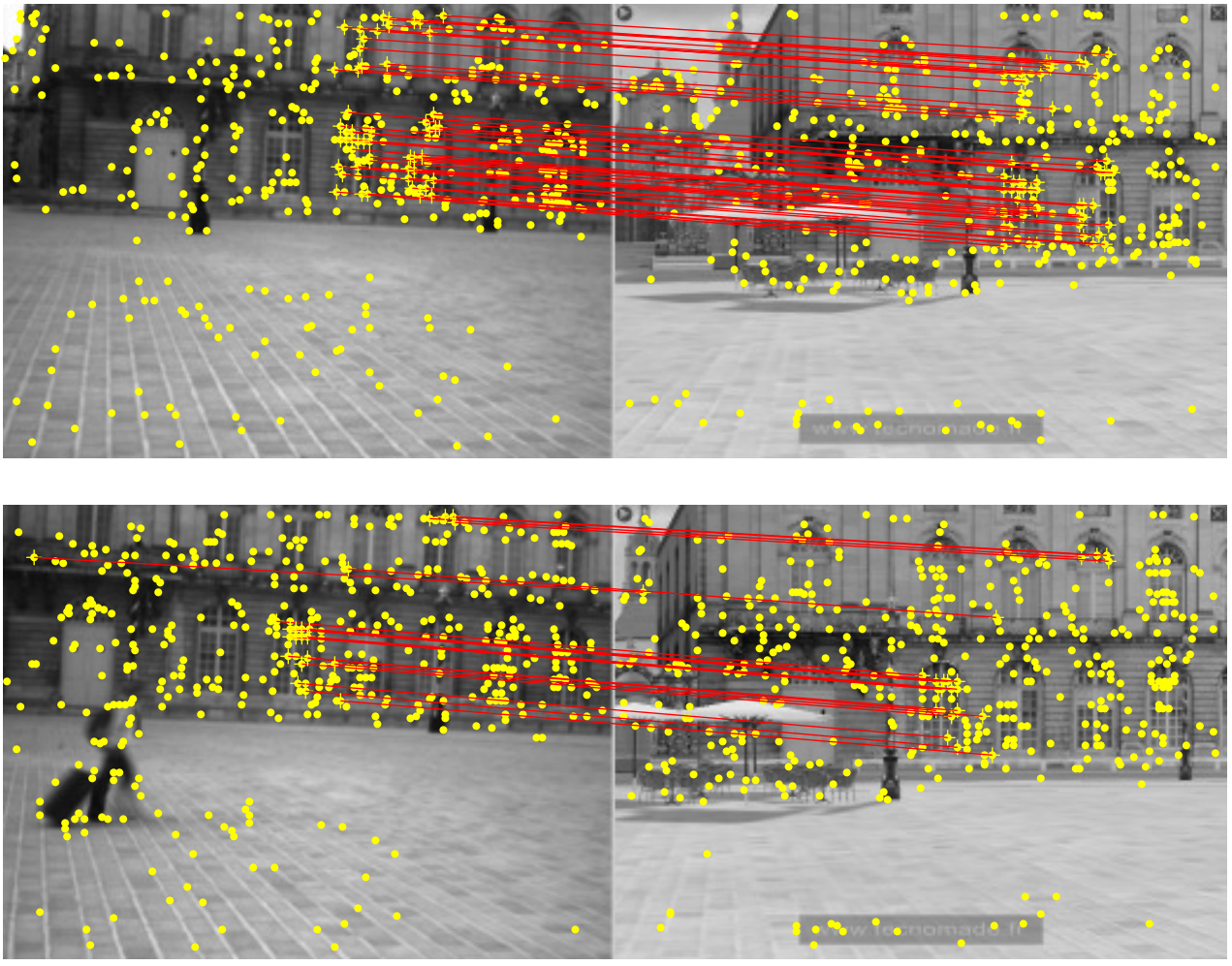


FIGURE IV.25 – Problème d'appariements 2D/2D si la pose *a priori* est trop éloignée

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

La Figure IV.26 illustre les résultats de l'observation cartographique 3D pour la séquence détaillée précédemment. Les croix rouges représentent les positions vraies du véhicule (acquise par le GPS RTK centimétrique), les croix bleues représentent les observations cartographiques 3D calculées. Un trait vert relie les observations GPS et cartographiques 3D correspondantes.

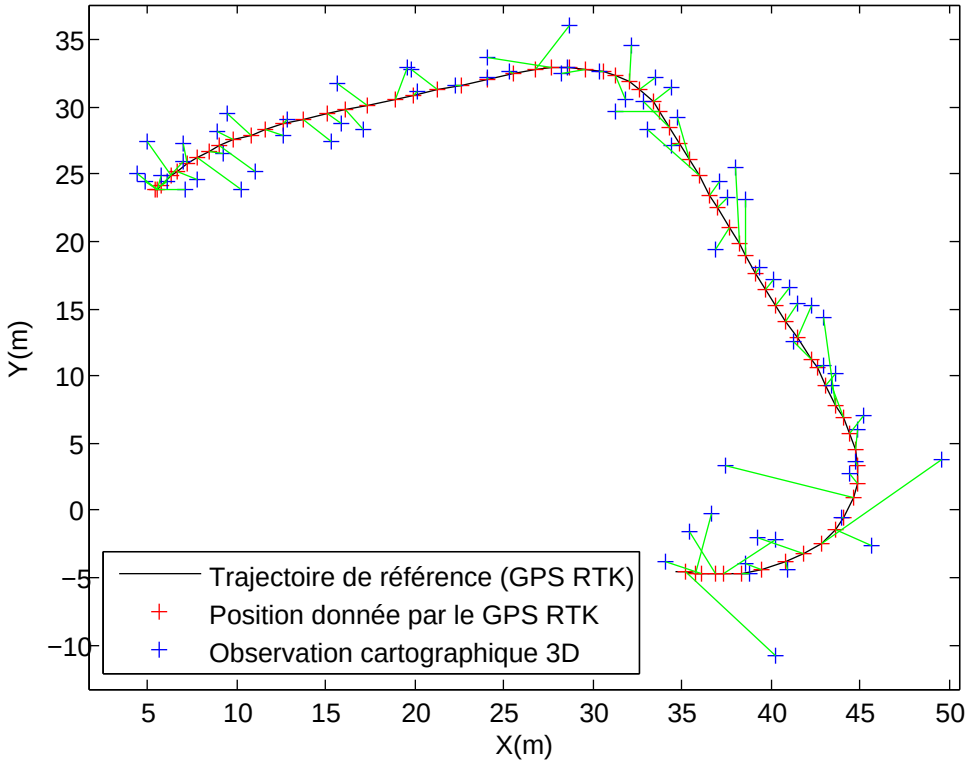


FIGURE IV.26 – Estimation de la pose donnée par l'observation cartographique 3D

Les écarts selon les axes X et Y entre les positions GPS et cartographiques 3D sont données respectivement dans les Figures IV.27 et IV.28.

Soient E_X l'erreur selon l'axe X et E_Y l'erreur selon l'axe Y. On note que pour cet essai l'erreur maximale est d'environ 11,3 mètres selon l'axe X et 6 mètres selon l'axe Y. La valeur moyenne de $|E_X|$ (valeur absolue de E_X) est 1,3 mètres et la valeur moyenne de $|E_Y|$ est 1,5 mètres. L'écart type de $|E_X|$ est 1,77 ; celui de $|E_Y|$ est 1,55. La précision atteinte est donc du même ordre que la précision d'un GPS en mode DGPS WAAS/EGNOS.

L'estimation de la pose par cette méthode est très sensible aux erreurs d'estimation des paramètres intrinsèques de la caméra. Nous envisageons donc dans des travaux à venir, d'intégrer cette incertitude dans le calcul de la pose. A ce jour, la minimisation effectuée afin de calculer la pose de la caméra, est réalisée en fonction

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

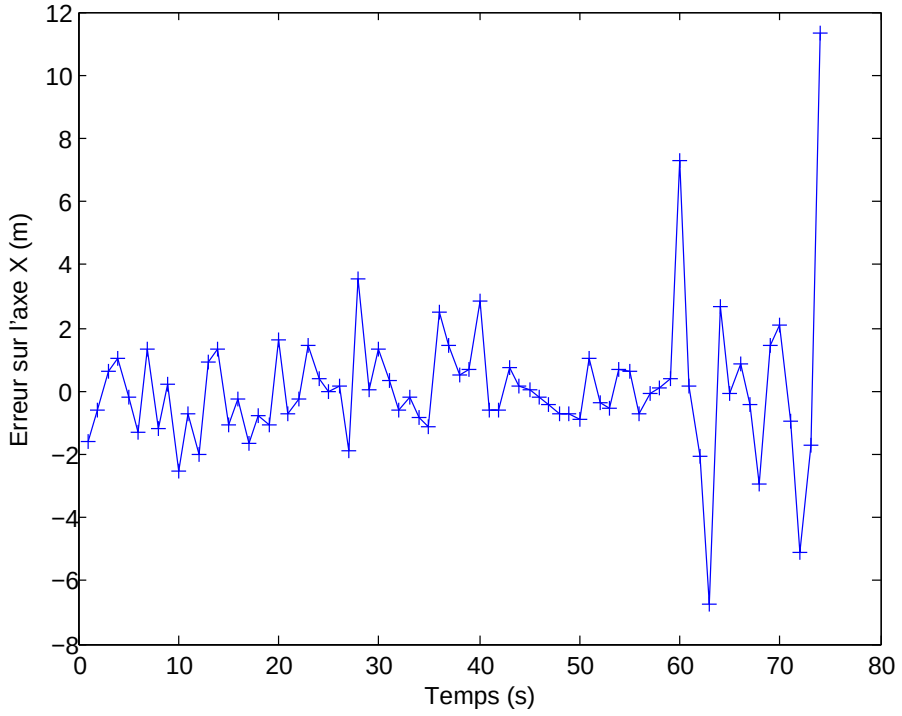


FIGURE IV.27 – Erreur sur l'axe X de l'observation cartographique 3D par rapport à un GPS RTK centimétrique

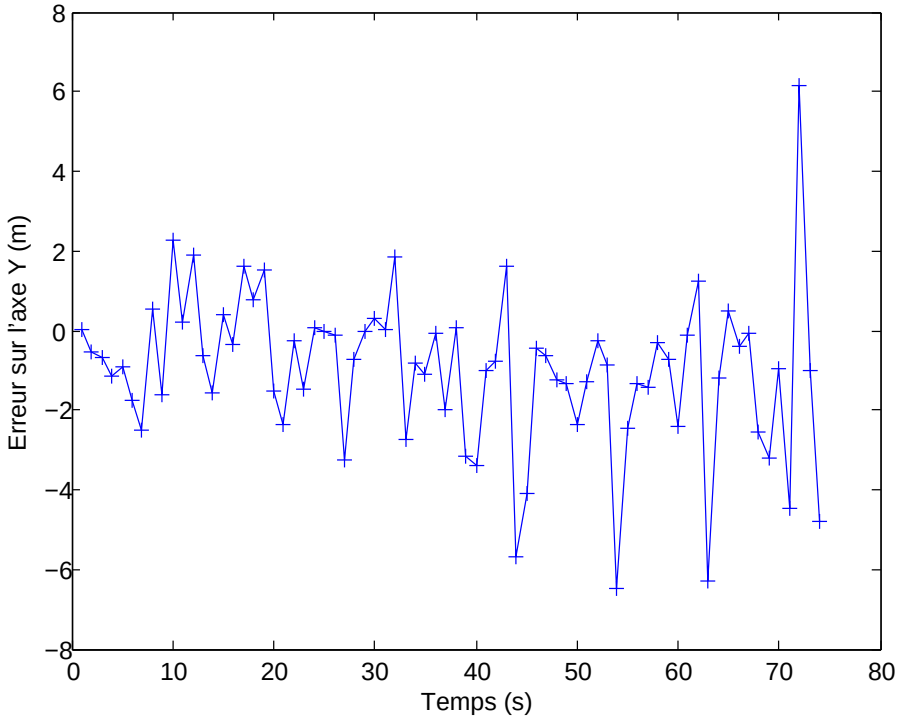


FIGURE IV.28 – Erreur sur l'axe Y de l'observation cartographique 3D par rapport à un GPS RTK centimétrique

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

des seuls paramètres extrinsèques. Nous pensons qu'en ajoutant les paramètres intrinsèques et en particulier la distance focale, les résultats pourraient être améliorés.

IV.4 Conclusion

Ce chapitre a été dédié à la présentation des résultats des méthodes de localisation présentées dans le chapitre III. D'une part, nous avons présenté les résultats du filtre de Kalman sans parfum. Cette méthode de localisation fusionnent les informations GPS et les informations issues des capteurs proprioceptifs (odomètres et gyromètre). Lorsque les informations GPS sont disponibles alors l'estimation de la pose repose essentiellement sur celles-ci. Le GPS est utilisé en mode centimétrique RTK afin de connaître la vérité terrain. Ainsi, lorsque les informations GPS sont utilisées pour corriger la prédiction, la pose estimée repose surtout sur celles-ci ; l'ellipse d'incertitude des informations GPS étant dans ce cas très petite. Nous avons simulé des masquages GPS, pour constater la dérive progressive de la localisation à l'estime seule. Un test de cohérence des informations GPS a également été instauré afin de détecter des mesures aberrantes, dues au problème du multi-trajets, à la visibilité satellitaire réduite en milieu urbain ... Sur une courte durée, on peut continuer à estimer la pose en l'absence de données GPS, mais à plus long terme, une autre estimation de la pose est nécessaire pour recaler la localisation à l'estime.

Nous avons donc ensuite présenté les résultats de la méthode de localisation se basant sur l'utilisation d'une caméra et du modèle virtuel urbain 3D. On constate que la pose estimée par cette méthode permet de maintenir une estimation de la pose du véhicule lorsque les informations GPS sont indisponibles. La précision atteinte par cette méthode est de quelques mètres, c'est-à-dire du même ordre que la précision d'un GPS en mode EGNOS. Cette approche est particulièrement adaptée à la navigation en milieu urbain, où le GPS fonctionne mal, mais où les amers naturels sont nombreux. Malheureusement, nous avons noté que dans certaines situations, de nombreux appariements sont erronés. Ceci vient du fait que la mise en correspondance se base sur un score de corrélation donc sur l'information colorimétrique. Les résultats seraient sans doute meilleurs en utilisant une mise en correspondance 3D/3D utilisant un système de stéréovision.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Chapitre V

Détection d'obstacles

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Résumé du chapitre

Ce chapitre traite de l'utilisation d'un modèle virtuel 3D urbain pour une autre application relative à la navigation : la détection et la localisation des obstacles. Deux types d'obstacles sont distingués : d'une part les obstacles de l'infrastructure, c'est-à-dire en particulier les bâtiments et d'autre part les obstacles n'appartenant pas à l'infrastructure tels que les piétons, les véhicules ...

En ce qui concerne les obstacles de l'infrastructure, on peut considérer que ceux-ci sont détectés et localisés dès lors que le modèle 3D de l'environnement est connu et que le véhicule est localisé dans ce modèle 3D.

Concernant l'autre type d'obstacles, l'approche suivie consiste à comparer l'image réelle acquise par la caméra embarquée avec l'image virtuelle extraite du modèle 3D en considérant que ce type d'obstacles est présent dans l'image réelle mais absent de l'image virtuelle.

A partir de l'information de profondeur disponible grâce au modèle 3D, les obstacles détectés sont ensuite géo-localisés.

Les résultats expérimentaux obtenus sont comparés et validés grâce à un télémètre laser.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Sommaire

V.1 Méthode développée 154
V.1.1 Présentation 154
V.1.2 Détection des obstacles 157
V.1.3 Extraction de la distance entre le véhicule et les obstacles 159
V.1.4 Calcul de la géo-position des obstacles détectés 160
V.2 Résultats expérimentaux 161
V.2.1 Moyens expérimentaux 161
V.2.2 Interface graphique 163
V.2.3 Validation des résultats à l'aide d'un télémètre laser à balayage 166
V.3 Conclusion 168

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Après avoir proposé une méthode montrant l'intérêt des modèles virtuels urbains 3D pour la géo-localisation de véhicules, nous présentons dans ce chapitre une utilisation de la modélisation urbaine 3D pour une autre application liée au domaine des transports intelligents : la détection et la localisation d'obstacles. L'idée de la méthode est de comparer l'image réelle acquise par la caméra embarquée avec l'image virtuelle fournie par le SIG-3D.

L'environnement est constitué de deux types d'obstacles que nous qualifions d'obstacles « structurels » d'une part et d'obstacles « ponctuels » d'autre part.

Les obstacles structurels sont les éléments constitutifs de l'architecture de l'environnement tels que les bâtiments, les constructions urbaines (monuments, poteaux, bordures de trottoirs, barrières, panneaux, ...), ils sont considérés invariants avec le temps. Ces obstacles, fixes et permanents, sont donc dans la base de données du modèle 3D. Une fois localisé dans le modèle, on peut donc facilement les identifier.

A contrario, les obstacles ponctuels évoluent provisoirement dans l'environnement, ils n'appartiennent pas à l'infrastructure, ce sont par exemple les piétons, les véhicules, ... Ils ne sont donc pas intégrés dans le modèle 3D. C'est cette propriété que nous allons utiliser afin de les détecter. Les obstacles ponctuels peuvent en effet être détectés en comparant l'image réelle qui contient les éventuels obstacles ponctuels avec l'image virtuelle d'où ils sont absents. Dans la suite de cette étude, c'est à ce type d'obstacles que nous ferons référence.

Une fois l'obstacle détecté, la distance entre celui-ci et le véhicule peut être calculée à partir de la position du pixel représentant l'obstacle dans l'image et de l'information de profondeur extraite par le SIG-3D. Finalement, les obstacles sont géo-localisés à l'aide de la distance précédemment calculée et de l'estimation de la pose du véhicule.

V.1 Méthode développée

V.1.1 Présentation

Le principe général de la méthode développée afin de détecter et localiser les obstacles est décrit dans la Figure V.1.

Cette méthode requiert le modèle virtuel 3D urbain, une caméra et une estimation

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

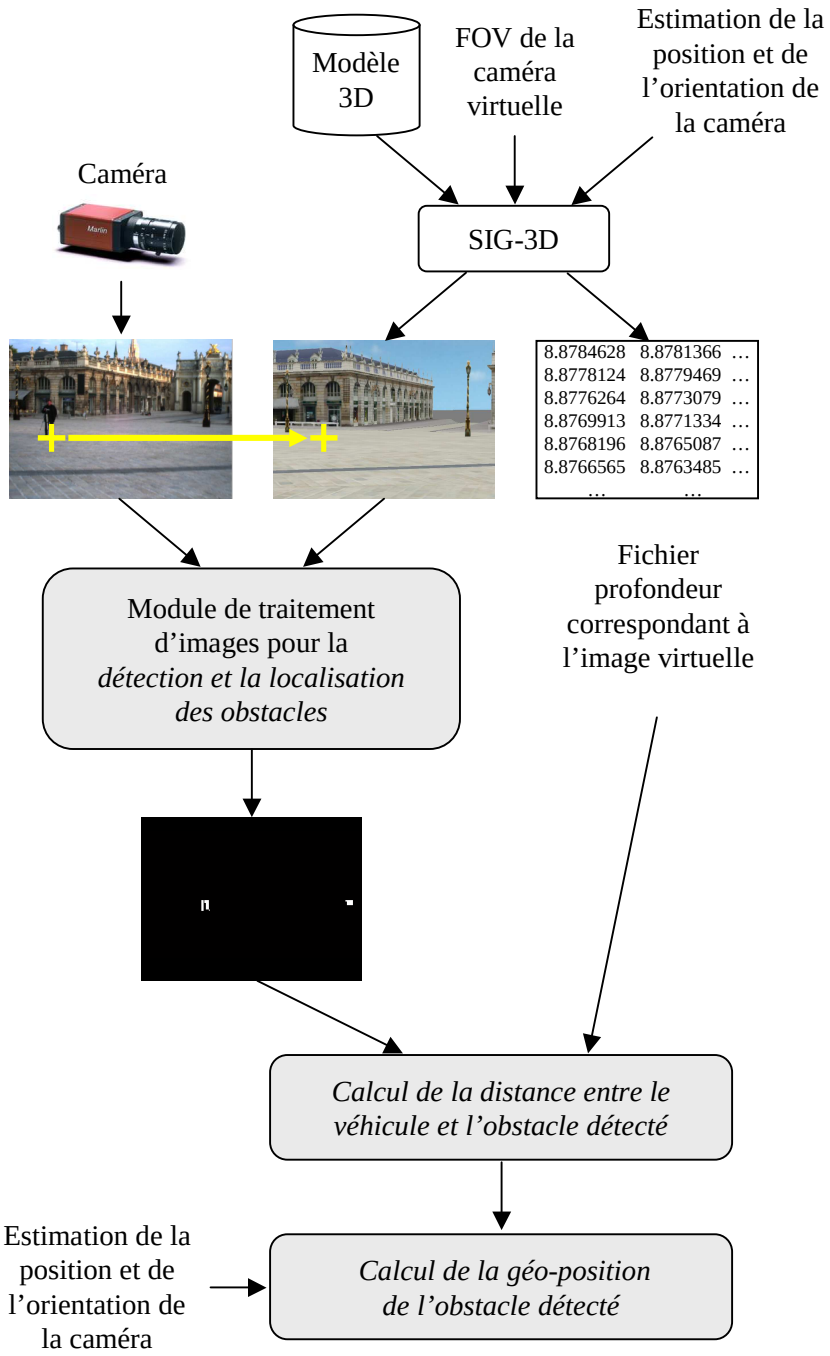


FIGURE V.1 – Principe de la méthode de détection et localisation d'obstacles

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

de la position et de l'orientation du véhicule que l'on suppose connue. Comme dans le cas de la localisation de véhicules, une première étape s'effectue au préalable de l'algorithme et hors-ligne : il s'agit de la détermination du paramètre intrinsèque FOV (Field Of View) de la caméra virtuelle. Rappelons (cf. § III.1.3) que la caméra réelle ayant été calibrée, on peut déterminer le FOV horizontal de la caméra virtuelle à partir des paramètres intrinsèques de la caméra réelle avec la relation suivante :

$$FOV = 2 \arctan \frac{l/2}{f'} \tag{V.1}$$

avec (cf Figure V.2) :

- FOV : le paramètre Field Of View horizontal
- l : la largeur en pixels de l'image
- f' : la distance focale en pixels de la caméra réelle obtenue grâce à la calibration

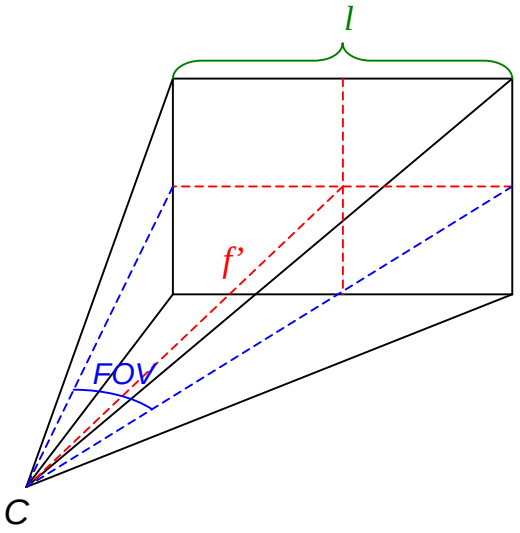


FIGURE V.2 – Calcul du paramètre intrinsèque FOV de la caméra virtuelle

La caméra virtuelle étant paramétrée similairement à la caméra réelle, la partie en ligne de l'algorithme peut s'effectuer. A chaque instant k, on considère que l'on dispose de l'image réelle courante acquise par la caméra embarquée à bord du véhicule ainsi que l'estimation de la position et de l'orientation du véhicule dans le repère géographique *WGS84* du GPS. La pose est convertie dans le système *Lambert 93* spécifique à la France, puis dans le repère local du modèle 3D. A partir des paramètres extrinsèques de la caméra virtuelle ainsi déterminés, du paramètre intrinsèque FOV et de la base de données du modèle, le SIG-3D peut extraire la vue de la scène par la caméra virtuelle, c'est-à-dire l'image virtuelle. Un module de traitement d'images, basé sur une soustraction, traite alors l'image réelle et l'image virtuelle afin de détecter les éventuels obstacles de la scène. Par ailleurs, le SIG-3D

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

extrait non seulement l'image virtuelle mais fournit également un fichier contenant une information sur la profondeur des pixels de l'image virtuelle, c'est-à-dire la distance entre la caméra virtuelle et les points 3D représentés par chacun des pixels de l'image virtuelle. Cette information, en plus de la position de l'obstacle dans l'image, est utilisée pour calculer la distance entre l'obstacle detecté et la caméra. Enfin, connaissant cette distance et la pose de la caméra, la géo-position de l'obstacle peut être calculée.

Après avoir détaillé les 3 étapes principales de la méthode, c'est-à-dire la détection des éventuels obstacles (cf. § V.1.2), puis la détermination de la distance entre le véhicule et l'obstacle (cf. § V.1.3) et enfin le calcul de la géo-position de l'obstacle détecté (cf. § V.1.4). Nous présenterons également la validation de nos résultats en utilisant un télémètre laser à balayage (cf. § V.2.3).

V.1.2 Détection des obstacles

A chaque instant k, deux images sont disponibles (Fig.V.3) :
- l'image réelle capturée par la caméra embarquée,
- l'image virtuelle extraite par le SIG-3D.

La méthode de détection d'obstacles a été développée pour être implémentable en temps-réel. Par conséquent, un algorithme de traitement d'image simple est requis. La première étape est une étape de pré-traitement des images qui consiste principalement en une conversion d'image « couleur » en image « niveau de gris » , puis en une binarisation.

Après ce pré-traitement, un module de traitement d'images basé sur le calcul de la différence entre l'image virtuelle binarisée et l'image réelle binarisée est mis en œuvre. Afin d'éviter de détecter des obstacles contenus dans le modèle 3D mais absents de la scène réelle, un ordre dans la soustraction doit être respecté. Finalement, un post-traitement est réalisé afin d'éliminer le bruit. Le résultat de ces différentes étapes est illustré sur la Figure V.4.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

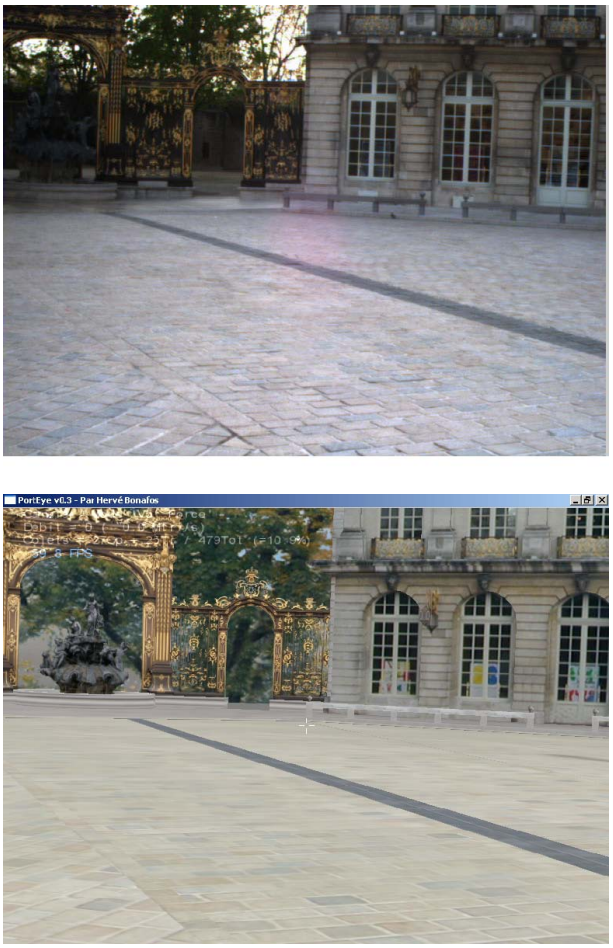


FIGURE V.3 – Image réelle (haut) et image virtuelle (bas) à un instant k

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

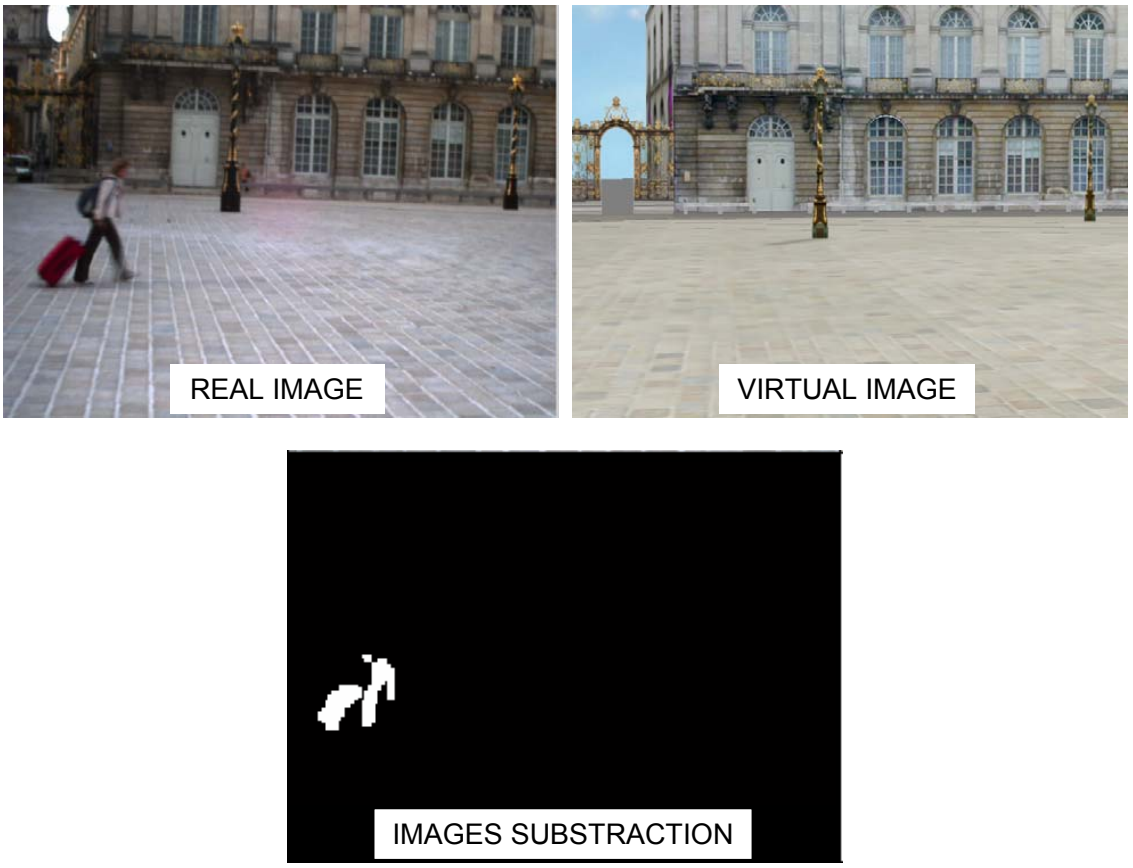


FIGURE V.4 – Image résultat pour la détection des obstacles de la scène

V.1.3 Extraction de la distance entre le véhicule et les obstacles

L'obstacle étant détecté, la distance entre le véhicule et l'obstacle peut être calculée en utilisant l'information de profondeur extraite par le SIG-3D. Le SIG-3D fournit en effet un fichier contenant la profondeur de chaque pixel de l'image virtuelle, c'est-à-dire la distance entre la caméra virtuelle et le point 3D correspondant au pixel considéré. La profondeur du pixel correspondant à l'obstacle détecté est alors recherchée dans ce fichier. C'est la distance entre la caméra et l'obstacle, car nous considérons que les obstacles sont au contact du sol. Cette hypothèse restrictive provient du fait que dans le cas d'un objet qui n'est pas en contact avec le sol (ex : oiseau en vol), le pixel correspondant dans l'image virtuelle ne représente pas l'obstacle mais un point 3D plus lointain. La profondeur correspond alors à ce point 3D du décor et non à l'obstacle. Mais si on considère le contact obstacle/sol, il y a correspondance entre le pixel représenté et la position de l'obstacle. L'intersection

© 2009 Tous droits réservés.
http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

entre le sol et l'obstacle est le pixel le plus bas détecté dans l'image « soustraction ». C'est ce pixel que nous utilisons pour caractériser la distance à l'obstacle.

V.1.4 Calcul de la géo-position des obstacles détectés

La distance entre le véhicule et la caméra ainsi que la pose du véhicule étant connues, l'obstacle peut être géo-localisé. Dans la mesure où la hauteur de l'obstacle n'est pas nécessaire, la scène peut être projetée sur le plan XY, qui est supposé être le plan du sol.

L'obstacle appartient alors à la droite passant par la caméra et d'équation :

$$y = \coth(\beta).(x - X_C) + Y_C \tag{V.2}$$

avec :

- X_C et Y_C , les coordonnées de la caméra dans le système *Lambert93* (donné par le GPS RTK),
- β, l'orientation de l'obstacle dans le système de coordonnées *Lambert93*, c'est-à-dire l'angle entre l'obstacle et l'axe Nord.

Puisque la distance entre la caméra et l'obstacle est connue, l'obstacle est géo-localisé.

L'orientation β est telle que :

$$\beta = head - \alpha \tag{V.3}$$

avec *head*, l'orientation du véhicule c'est-à-dire l'angle entre l'axe de la caméra et l'axe Y (Nord) du repère *Lambert93* et α, l'angle entre l'axe de la caméra et l'obstacle. L'angle α est donc l'angle formé par l'axe orthogonal au plan de l'image et au centre de l'image (segment D dans la Figure V.5) et le segment d_{ob} reliant la caméra à l'obstacle. Dans le même plan de profondeur, l'angle α peut être calculé avec l'équation trigonométrique suivante (cf Fig V.5) :

$$\tan\alpha = \frac{A}{width/2}\tan\left(O_h/2\right) \tag{V.4}$$

où :

- O_h est le FOV horizontal (Field Of View),

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- *width* est le nombre de pixels dans la largeur de l'image,
- A représente le nombre de pixels entre le pixel du centre de l'image et la projection P_{ob} de l'obstacle dans le plan P dont la profondeur est la profondeur du pixel central de l'image.

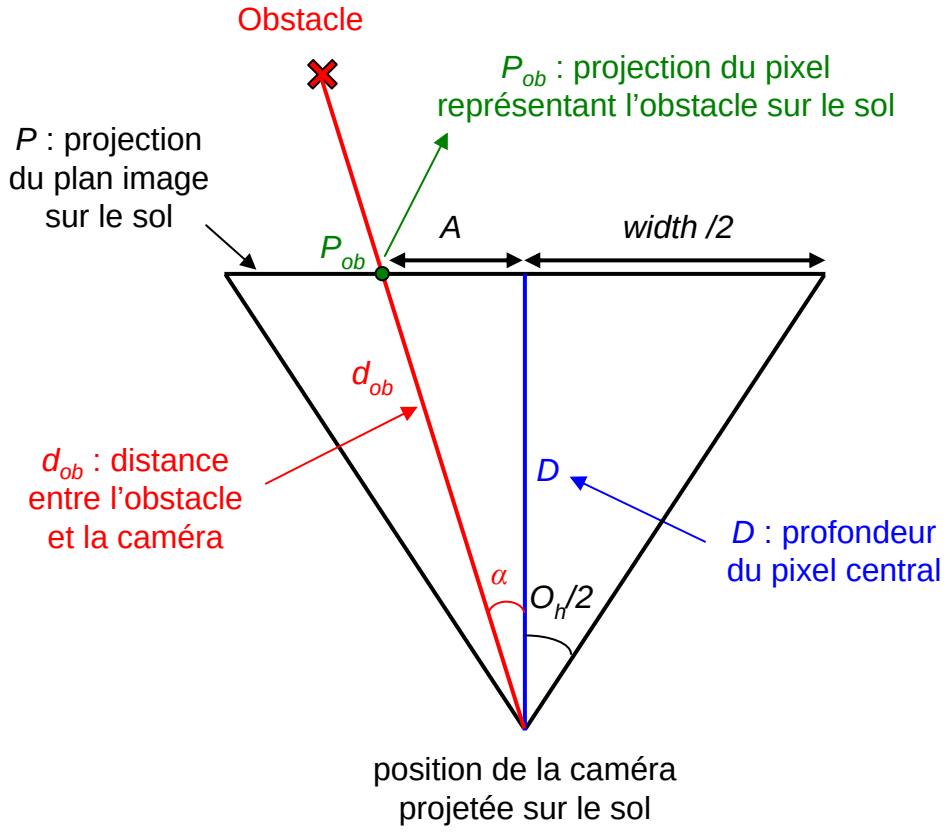


FIGURE V.5 – Calcul de la géo-position de l'obstacle détecté

V.2 Résultats expérimentaux

V.2.1 Moyens expérimentaux

Les moyens expérimentaux utilisés dans le cadre de la détection d'obstacle sont les mêmes que pour la localisation de véhicules, à savoir le véhicule CyCab, la caméra Marlin, le GPS Sagitta, le modèle 3D. Un télémètre laser est ajouté afin de valider la méthode proposée. Le télémètre laser est un SICK LMS 291 (Fig V.6). Son angle d'ouverture est de 180°, la résolution angulaire utilisée est de 1° et la portée du capteur est de 80 mètres, sa fréquence est de 75 Hz. Les capteurs sont fixés rigidement à l'avant du véhicule comme indiqué sur la Figure V.7.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

FIGURE V.6 – Télémètre laser utilisé : LMS SICK 291

FIGURE V.7 – Fixation des capteurs sur le CyCab

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

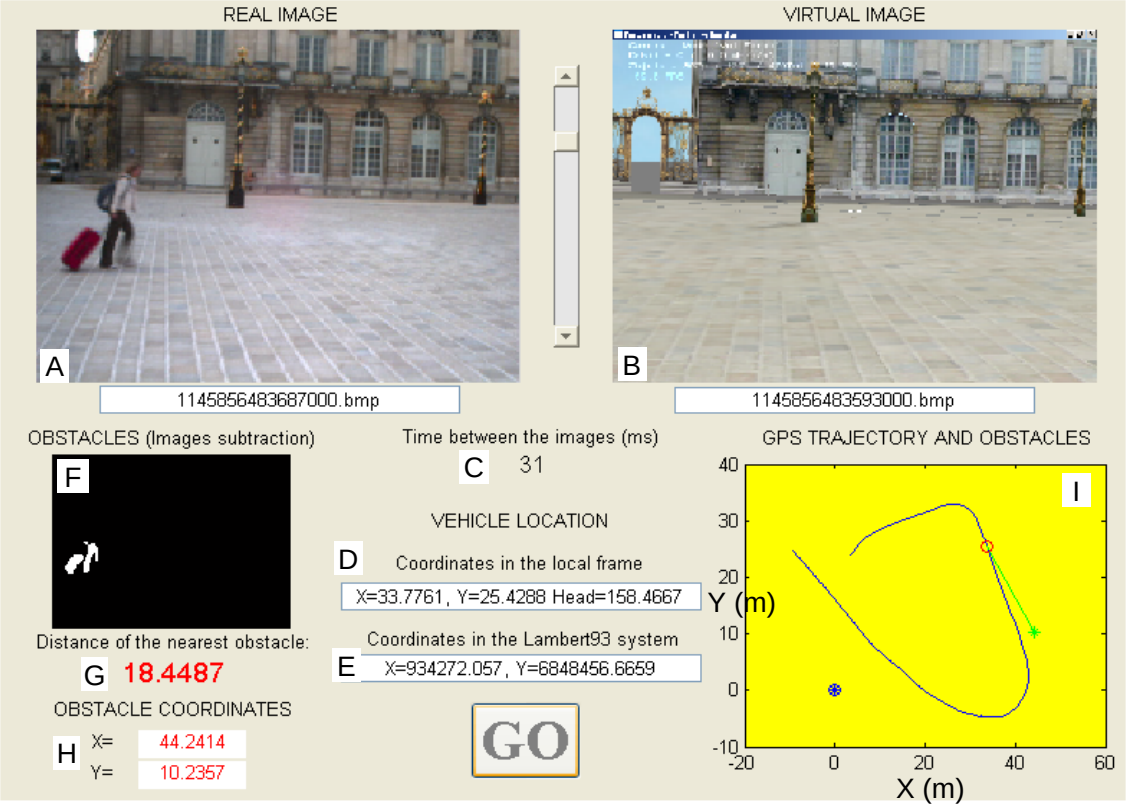


FIGURE V.8 – Interface graphique du détecteur d'obstacles

V.2.2 Interface graphique

Afin de présenter nos résultats, une interface graphique a été développée. Celle-ci est présentée dans la Figure V.8.

Afin de décrire cette interface graphique, des légendes de A à I sont utilisées comme suit :

- Partie A : image réelle acquise par la caméra embarquée ainsi que sa date d'acquisition (temps écoulé depuis le 1er Janvier 1970) au bas de l'image

- Partie B : image virtuelle extraite par le SIG-3D et sa date d'acquisition, qui est en fait la date de l'estimation de la pose de la caméra virtuelle

- Partie C : différence entre les dates d'acquisition de l'image réelle et de l'image virtuelle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

- Partie D : estimation de la position et de l'orientation de la caméra dans le repère local attaché au modèle 3D
- Partie E : estimation de la position et de l'orientation de la caméra en Lambert93
- Partie F : résultat du module de traitement d'images pour la détection des obstacles (les pixels blancs représentent les obstacles)
- Partie G : distance calculée entre l'obstacle et le véhicule
- Partie H : géo-position, dans le repère local, de l'obstacle détecté
- Partie I :
  - la ligne continue bleue est la trajectoire suivie par le véhicule, c'est-à-dire les positions délivrées par le GPS RTK converties dans le repère local. L'origine de ce repère local (la statue de Stanislas, au milieu de la place) est représentée par une étoile bleue sur ce même graphe
  - le cercle rouge sur la trajectoire est la position courante du véhicule
  - en guise d'alarme, la couleur de fond de la partie I devient jaune lorsqu'un obstacle est détecté
  - l'étoile verte est l'obstacle détecté
  - le segment vert relie le véhicule à l'obstacle, c'est-à-dire l'estimation de la distance et de l'angle entre le véhicule et l'obstacle

Quelques résultats sont présentés sur la Figure V.9. Dans chacun des 7 cas présentés, pour plus de clarté, nous n'affichons que l'image réelle (haut-gauche), l'image virtuelle (haut-droite), l'image des obstacles détectés (bas-gauche) et le graphe de la partie I de l'interface graphique (bas-droite).

Sur la première ligne de la Figure V.9, nous avons illustré le cas, où il n'y a pas d'obstacle. Puis, sur les lignes suivantes, différents obstacles sont détectés : des personnes, un oiseau, un lampadaire. La méthode mise en œuvre permet donc de détecter et de localiser des obstacles de taille et de nature variées.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

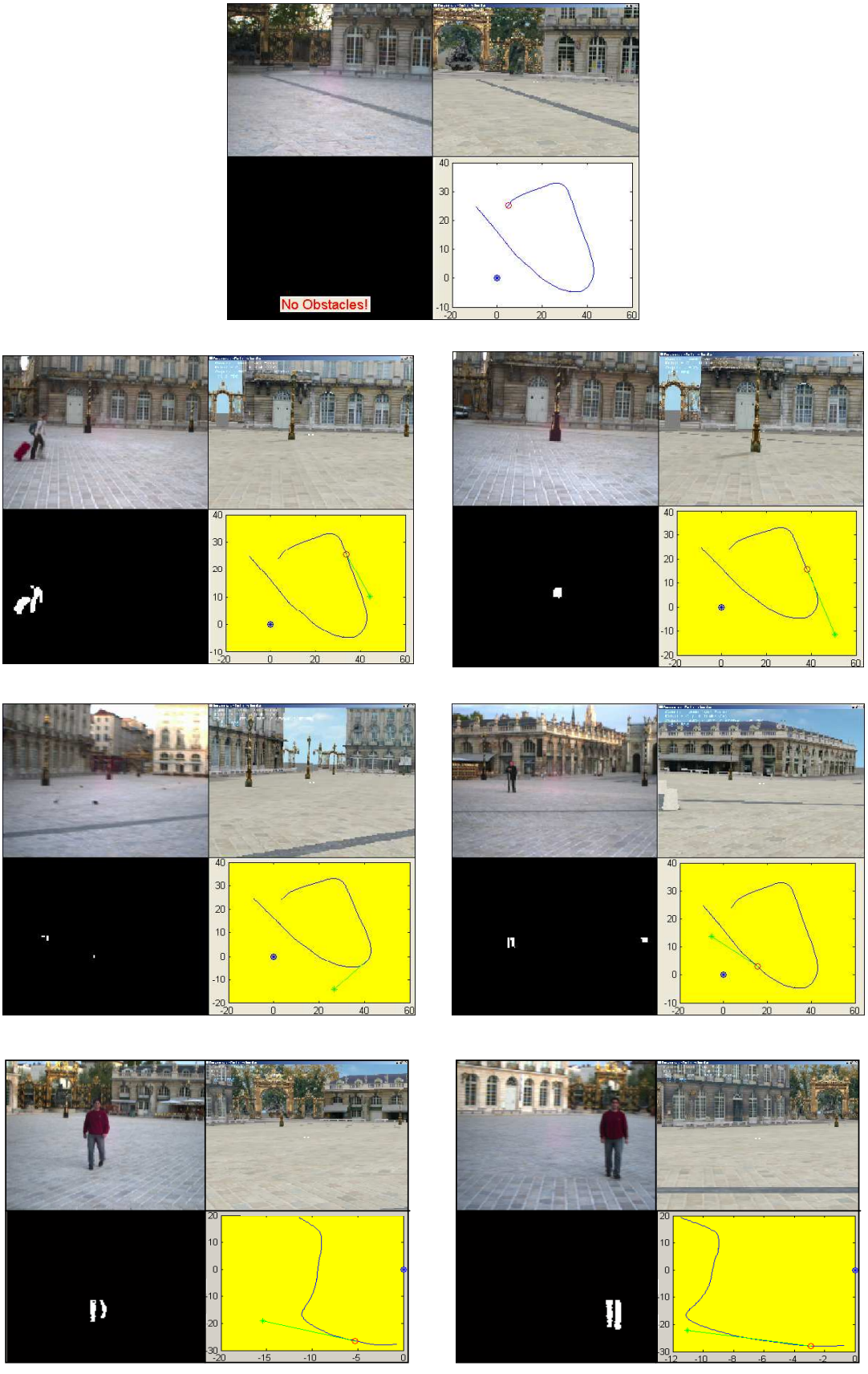


FIGURE V.9 – Résultats du détecteur d'obstacle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

V.2.3 Validation des résultats à l'aide d'un télémètre laser à balayage

Afin de valider la distance calculée entre le véhicule et l'obstacle avec notre méthode, nous avons comparé les résultats obtenus avec la distance mesurée par un télémètre laser.

La télémétrie laser, aussi appelée Lidar (Light Detection and Ranging), est une technique active qui émet de la lumière visible ou infrarouge (Laser IR) de façon contrôlée. Comme pour une caméra, un télémètre laser a un angle de vue sur la scène et en collecte des informations. Alors qu'une caméra acquiert une information colorimétrique sur la scène, le télémètre laser retourne une information de distance entre les premiers objects détectés et le capteur lui-même. A chaque point de la scène visible par le télémètre laser est donc associés une distance et un angle de vue, qui peuvent être convertis en coordonnées spatiales cartésiennes. La technique de mesure de distance est qualifiée de mesure par temps de vol. On mesure le temps Δt que met l'onde électromagnétique pour faire l'aller-retour entre le télémètre laser et l'objet. Connaissant la vitesse c de propagation de l'onde, on en déduit la distance D qui sépare le capteur de la cible sur laquelle l'onde s'est réfléchie :

$$D = c \cdot \Delta t/2 \tag{V.5}$$

Le capteur balaye donc un plan horizontal situé devant le véhicule et renvoie, tous les degrès, la distance entre le capteur et le premier obstacle rencontré par le rayon laser. Les distances fournies par le télémètre laser sont comparées aux distances calculées entre la caméra et les obstacles par notre méthode. L'objectif étant ici la validation de notre approche, nous n'avons pas automatisé la recherche de l'obstacle parmi les données issues du télémètre laser. Nous avons sélectionné manuellement les données issues du télémètre laser qui correspondent aux obstacles détectés par notre approche.

Puisqu'on cherche ici à valider notre approche, l'extraction de la distance, qui correspond à l'obstacle détecté par notre méthode, parmi celles envoyées par le télémètre laser a été fait manuellement sur quelques essais uniquement.

Les Figures V.10 et V.11 montrent les résultats obtenus.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

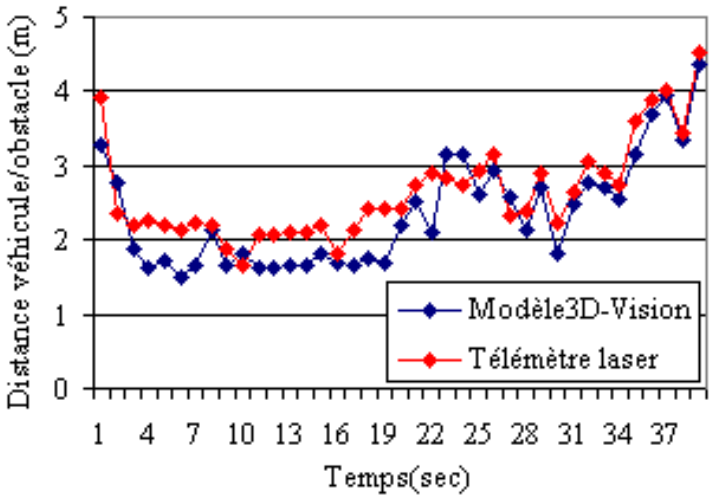


FIGURE V.10 – Distance entre le véhicule et l'obstacle donnée par notre méthode et par le télémètre laser

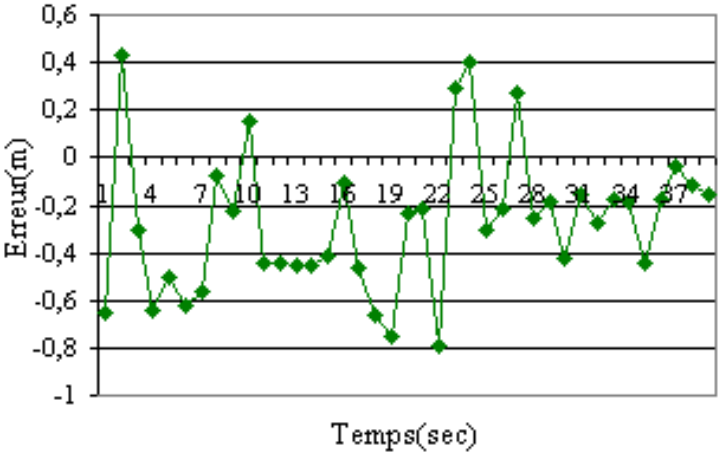


FIGURE V.11 – Ecart entre les distances données par notre méthode et par le télémètre laser

Sur la Figure V.10, la distance entre le véhicule et l'obstacle donnée par notre méthode est tracée en bleu alors que la distance mesurée par le télémètre laser est indiqué en rouge. La courbe tracée dans la Figure V.11 indique l'écart entre ses 2 distances.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Pour un obstacle (une personne) passant devant le véhicule à une distance de celui-ci comprise entre 1,50 mètres et 4,50 mètres, l'écart maximal entre la distance calculée par notre méthode et celle mesurée par le télémètre laser est de 0,8 mètres. L'écart moyen est de 0,35 mètres.

V.3 Conclusion

Dans ce chapitre, nous avons proposé une méthode de détection et de localisation d'obstacles à l'aide d'un modèle virtuel 3D urbain et d'une caméra. Cette approche nécessite de connaître une estimation très précise de la pose du véhicule. Par ailleurs, notons que comme toute méthode basée sur l'utilisation de la vision, cette méthode n'est pas utilisable sous certaines conditions climatiques (obscurité, brouillard ...).

Les résultats obtenus, avec un traitement d'images simple, sont satisfaisants. Cependant, de nombreuses améliorations pourraient être apportées afin de rendre le détecteur d'obstacles plus performant. Par exemple, l'utilisation d'un système de stéréovision au lieu de la caméra mono-vision, permettrait de travailler sur les images de profondeur fournies par la stéréovision d'une part et par le SIG-3D d'autre part plutôt que sur les images couleur. Ainsi, de nombreuses difficultés, telle que la gestion des ombres, seraient palliées.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Conclusion générale et perspectives

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

Conclusions

Dans cette thèse, nous avons traité le problème de la localisation de véhicules en milieu urbain et de la détection et la localisation des obstacles. Nous avons proposé deux concepts répondant à ces deux applications utilisant une nouvelle source d'information, à savoir un modèle virtuel 3D urbain.

Nous avons tout d'abord commencé par une introduction sur les méthodes de la localisation de véhicules routiers existantes. Nous avons ensuite présenté la modélisation urbaine en trois dimensions, et détaillé le modèle virtuel 3D urbain utilisé dans nos expérimentations. Le système d'information géographique 3D développé a également été présenté. Cet outil informatique permet de manipuler la base de données du modèle 3D. Les paramètres à fixer de la caméra virtuelle ont été décrits. Le géo-référencement du modèle 3D a également été exposé, c'est une étape essentielle pour l'utilisation du modèle 3D dans les applications dédiées aux véhicules intelligents. Les différentes sorties disponibles du SIG-3D ont également été détaillés. Il s'agit notamment du rendu de la projection du modèle 3D sous la forme d'une image virtuelle, ainsi que l'information de profondeur correspondante. Ces deux informations sont nécessaires pour les méthodes de localisation de véhicules et de détection d'obstacles présentées dans les deux chapitres suivants.

La méthode de localisation de véhicules s'effectue en deux temps. Tout d'abord une hybridation « GPS - Odomètre - Gyromètre »au travers un filtre de Kalman sans parfum permet d'obtenir une estimation continue de la pose du véhicule. Cette estimation est ensuite utilisée pour construire l'observation cartographique 3D. L'observation cartographique 3D est basée sur le recalage entre le modèle virtuel 3D et l'image réelle c'est-à-dire l'image courante acquise par la caméra embarquée. Pour cela, on extrait les points de Harris de l'image réelle et de l'image virtuelle. L'image virtuelle est l'image extraite par le SIG-3D en positionnant la caméra virtuelle selon la pose estimée précédemment. Les deux ensembles de points de Harris ainsi obtenus sont ensuite appariés. Les pixels de l'image virtuelle sont ensuite reconstruits en 3D grâce au fichier de profondeur. On dispose alors d'appariements 2D/3D, à partir desquels on peut calculer la pose de la caméra réelle.

La méthode de détection d'obstacles est basée sur un traitement simple des images réelles et virtuelles. Le principe repose sur le constat que les obstacles (n'appartenant pas à l'infrastructure) sont présents dans l'image réelle mais absent de l'image virtuelle. Grâce à l'information de profondeur disponible grâce au modèle

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

virtuel 3D, les obstacles détectés ont été géo-localisés.

Les méthodes développées ont été testées sur des données réelles acquises par les différents capteurs utilisés (GPS, Odomètres, Gyromètre, Caméra) et embarqués sur le véhicule CyCab. Nous retiendrons que les performances obtenues au niveau de l'application de localisation de véhicules sont du même ordre que celle d'un DGPS. Les résultats obtenus dans l'application de localisation d'obstacles ont été comparés à des données fournies par un télémètre laser.

Perspectives

De nombreuses perspectives des travaux présentés sont envisagées afin d'améliorer les résultats obtenus et / ou de les compléter.

Tout d'abord, au niveau applicatif, une implémentation temps-réel de la méthode de localisation présentée est en cours de développement.

D'autre part, concernant l'observation cartographique 3D, la technique de traitement d'images est certainement perfectible. En particulier, il serait intéressant de travailler avec un système de stéréovision plutôt qu'en vision monoculaire afin de tirer pleinement bénéfice de la modélisation en trois dimensions. En effet, à l'heure actuelle la stéréovision en temps réel est un acquis. L'appariement stéréovision/modèle 3D permettrait certainement d'améliorer les résultats tant pour la localisation de véhicules que pour la détection des obstacles.

Ensuite, les améliorations proposées pour la modélisation de l'erreur cartographique 3D doivent être développées afin d'intégrer cette observation à l'algorithme de fusion de données.

Le modèle virtuel 3D urbain peut également être envisagé comme une source de données afin de faire du couplage serré GPS-Modèle 3D. La fusion des pseudo-distances du GPS avec des distances entre quelques amers géo-référencés du modèle 3D et le véhicule pourrait en effet permettre d'estimer la pose dans une zone de visibilité GPS réduite.

Enfin, afin de disposer d'un SIG-3D plus complet, il serait intéressant d'intégrer les fonctionnalités, attributs et requêtes disponibles dans les SIG-2D usuels.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Bibliographie

[1] M. Aldon. Capteurs et méthodes pour la localisation des robots mobiles. *Techniques de l'Ingénieur, traité Informatique Industrielle*, S-7852 :1–19, 2001.

[2] H. Araùjo, R. L. Carceroni, and C. M. Brown. A fully projective formulation to improve the accuracy of lowe's pose-estimation algorithm. *Computer Vision and Image Understanding*, 70(2) :227–238, 1998.

[3] S. Arulampalam, S. Maskell, and N. Gordon. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50 :174–188, 2002.

[4] N. M. Barbour, J. M. Elwell, and R. H. Setterlund. Inertial instruments - Where to now ? *American Institute of Aeronautics and Astronautics : Guidance, Navigation and Control Conference, Technical Papers (A92-55151 23-63)*, pages 566–574, Aug. 1992.

[5] J. Batista, H. Araujo, and A. De Almeida. Iterative multi-step explicit camera calibration. *Sixth International Conference on Computer Vision (ICCV'98)*, pages 709–714, 1998.

[6] P. A. Beardsley, P. H. S. Torr, and A. Zisserman. 3D Model Acquisition from Extended Image Sequences. *4th European Conference on Computer Vision (ECCV'96), London, UK*, II :683–695, 1996.

[7] P. Beaudet. Rotationally invariant image operators. *International Joint Conference on Pattern Recognition*, pages 579–583, 1978.

[8] D. Berdjag and D. Pomorski. DGPS/INS Data Fusion for Land Navigation. *Seventh International Conference on Information Fusion (Fusion'04), Stockholm, Sweden*, II :881–887, Jun 2004.

[9] N. Bergman. *Recursive Bayesian Estimation : Navigation and Tracking Applications*. PhD thesis, Linkoping University, Sweden, 1999.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[10] D. Bernstein and A. Kornhauser. Map matching for personal navigation assistants. *The Transportation Research Board - 77th Annual Meeting, Washington, D.C*, 1998.

[11] P. Bonnifait. *Localisation précise en position et en attitude des robots mobiles d'extérieur à évolution lente*. PhD thesis, Université de Nantes, Ecole Centrale de Nantes, France, 1997.

[12] J. Borenstein, H. R. Everett, L. Feng, and D. Wehe. Mobile Robot Positioning : Sensors and Techniques. *Journal of Robotic Systems*, 14(4) :231–249, 1997.

[13] J. Borenstein and L. Feng. Measurement and correction of systematic odometry errors in mobile robots. *IEEE Transactions on Robotics and Automation*, (12), 1996.

[14] J.-Y. Bouguet. Camera calibration toolbox for matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html.

[15] S. Bourgeois, S. Naudet-Colette, and M. Dhome. Recalage d'un modèle cao à partir de descripteurs locaux de contours. *Congrès francophone de Reconnaissance des Formes et d'Intelligence Artificielle (RFIA'06)*, 2006.

[16] D. Bouvet, G. Garcia, B. Gorham, and D. Betaille. Precise 3-d localization by automatic laser theodolite and odometer for civil-engineering machines. *IEEE InternationalConference on Robotics and Automation (ICRA'01)*, 2 :2045–2050, 2001.

[17] C. Brenneke, O. Wulf, and B. Wagner. Using 3D laser range data for SLAM in outdoor environments. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'03)*, 1, 2003.

[18] C. Cappelle, M. El Badaoui El Najjar, F. Charpillet, and D. Pomorski. Outdoor obstacle detection and localisation with monovision and 3d geographical database. *IEEE Intelligent Transportation Systems Conference (ITSC'07), Seattle, Washington, USA*, pages 1102–1107, 30 Sept - 3 Oct 2007.

[19] C. Cappelle, M. El Badaoui El Najjar, F. Charpillet, and D. Pomorski. Obstacle detection and localization method based on 3d model : distance validation with ladar. *IEEE International Conference on Robotics and Automation (ICRA'08), Pasadena, Californie, USA*, 19 - 23 Mai 2008.

[20] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Localisation in urban environment using gps and ins aided by monocular vision system and 3d geographical model. *IEEE Intelligent Vehicles Symposium (IV'07)*, pages 811–816, 13 - 15 Juin 2007.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[21] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Mise en correspondance d'images réelles avec un modèle 3d géographique pour la localisation terrestre. *Traitement et Analyse de l'Information : Méthodes et Applications (TAIMA'07), Hammamet, Tunisie*, pages 523–528, 23 - 26 Mai 2007.

[22] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Détection, suivi et géo-localisation d'obstacles à l'aide d'un modèle 3d géo-référencé, une caméra et un gps : validation avec un lidar. *Conférence Internationale Francophone d'Automatique (CIFA'08), Bucarest, Roumanie*, 3-5 Septembre 2008.

[23] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Multisensor data fusion for geo-localisation in urban areas. *European Navigation Conference - Global Navigation Satellite Sytems (ENC-GNSS'08), Toulouse, France*, 23 - 25 Avril 2008.

[24] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Multisensors data fusion using dynamic bayesian network for robotised vehicle geo-localisation. *IEEE International Conference on Information Fusion (Fusion'08), Cologne, Allemagne*, 1-3 Juillet 2008.

[25] C. Cappelle, M. El Badaoui El Najjar, D. Pomorski, and F. Charpillet. Intelligent geo-localisation in urban areas using gps, 3d-gis and vision. *Journal of Intelligent Transportation Systems edited by Francis and Taylor*, to appear.

[26] C. Cappelle, D. Pomorski, and Y. Yanqin. Gps/ins data fusion for and vehicle localization. *IMACS Multiconference on Computational Engineering in Systems Applications (CESA'06)*, pages 21–27, October 4 - 6 2006.

[27] O. Carmichael and M. Hebert. Object recognition by a cascade of edge probes. *British Machine Vision Conference*, 2002.

[28] S. Coorg and S. Teller. Automatic extraction of textured vertical facades from pose imagery. Technical report, MIT LCS TR-729, Cambridge, MA, USA, 1998.

[29] S. Coorg and S. Teller. Extracting textured vertical facades from controlled close-range imagery. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99), Los Alamitos, CA, USA*, 1 :625–632, 1999.

[30] N. Cornelis, K. Cornelis, and L. Van Gool. Fast Compact City Modeling for Navigation Pre-Visualization. *IEEE Computer Society Conference on Com-*

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

puter Vision and Pattern Recognition (CPRV'2006), Los Alamitos, CA, USA, 02 :1339–1344, 2006.

[31] I. Cox. Blanche : position estimation for an autonomous robot vehicle. *Autonomous robot vehicles, Springer-Verlag New York, Inc.*, pages 221–228, 1990.

[32] P. David, D. DeMenthon, R. Duraiswami, and H. Samet. SoftPOSIT : Simultaneous Pose and Correspondence Determination. *International Journal of Computer Vision*, 59(3) :259–284, September 2004.

[33] A. Davison. Mobile robot navigation using active vision. Technical report, Department of Engineering Science, University of Oxford, 1998.

[34] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs : A hybrid geometry- and image-based approach. *Conference on Computer Graphics and Interactive Technique*, pages 11–20, 1996.

[35] D. F. Dementhon and L. S. Davis. Model-based object pose in 25 lines of code. *International Journal of Computer Vision*, 15 :123–141, 1995.

[36] M. G. Dissanayake, P. Newman, S. Clark, and H. Durrant-Whyte. A Solution to the Simultaneous Localization and Map Building (SLAM) Problem. *IEEE Transactions on Robotics and Automation*, 17(3) :229–241, Juin 2001.

[37] A. Doucet, S. Godsill, and C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, (10), 2000.

[38] F. Du and M. Brady. Self-calibration of the intrinsic parameters of cameras for activevision systems. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'93)*, pages 477–482, 1993.

[39] M. El Badaoui El Najjar. *Localisation dynamique d'un véhicule sur une carte routière numérique pour l'assistance à la conduite*. PhD thesis, Université de Technologie de Compiègne, France, 2003.

[40] M. El Badaoui El Najjar and P. Bonnifait. A road-matching method for precise vehicle localization using belief theory and Kalman filtering. *Autonomous Robots Journal*, 19(2) :173–191, 2005.

[41] A. Elaksher, J. Bethel, and E. Mikhail. Reconstructing 3d building wireframes from multiple images. *International Symposium on Photogrammetric Computer Vision (ISPRS'02), Graz, Austria*, pages 91–96, September 2002.

[42] O. Faugeras. *Three-Dimensional Computer Vision : A Geometric Viewpoint*. MIT Press, 1993.

[43] M. A. Fischler and R. C. Bolles. Random sample consensus : a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6) :381–395, 1981.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[44] C. Fouque and P. Bonnifait. Tightly coupled gis data in gnss fix computation with integrity test. *International Journal of Intelligent Information and Database Systems*, 2(2) :167–186, 2008.

[45] D. Fox, S. Thrun, F. Dellaert, and W. Burgard. Particle filters for mobile robot localization. *Sequential Monte Carlo in Practice*, 2001.

[46] C. Frueh and A. Zakhor. Automated reconstruction of building facades for virtual walk-thrus. *ACM SIGGRAPH Conference on Sketches & Applications*, 2003.

[47] C. Frueh and A. Zakhor. Constructing 3d city models by merging ground-based and airborne views. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03), Los Alamitos, CA, USA*, 2 :562–569, 2003.

[48] C. Frueh and A. Zakhor. An automated method for large-scale, ground-based city model acquisition. *International Journal of Computer Vision*, 60(1) :5–24, October 2004.

[49] J.-M. Gaudin. *Conception d'un système hybride GPS - Estime pour la localisation en condition de visibilité satellitaire réduite.* PhD thesis, Université de Montpellier II, France, 8 Décembre 2000.

[50] P. Gerbe. Outils, données et applications de l'information géographique en 3d. *GeoEvenement*, Avril 2002.

[51] A. Gning. *Localisation garantie d'automobiles. Contribution aux techniques de satisfaction de contraintes sur les intervalles.* PhD thesis, Université de Technologie de Compiègne, France, 2006.

[52] A. Gning, F. Abdallah, and P. Bonnifait. A new estimation method for multisensor fusion by using interval analysis and particle filtering. *IEEE International Conference on Robotics and Automation (ICRA'07)*, pages 3844–3849, 10-14 April 2007.

[53] A. Gning and P. Bonnifait. Constraints propagation techniques on intervals for a guaranteed localization using redundant data. *Automatica*, Volume 42(Issue 7) :1167–1175, July 2006.

[54] J. Greenfeld. Matching gps observations to locations on a digital map. *81th Annual Meeting of the Transportation Research Board, Washington, DC*, 2002.

[55] B. Grocholsky. *Information-Theoretic Control of Multiple Sensor Platforms.* PhD thesis, Australian Centre for Filed Robotics, 2002.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[56] J. Guivant, E. Nebot, and S. Baiker. Autonomous navigation and map building using laser range sensors in outdoor application. *Journal of Robotic Systems*, 17(10) :565–583, Octobre 2000.

[57] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P. Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, 50(2) :425–437, 2002.

[58] E. Halbwachs. *Estimation à erreur bornée pour la génération d'hypothèses multiples de localisation d'un véhicule*. PhD thesis, Université de Technologie de Compiègne, France, 1997.

[59] C. Harris. Geometry from visual motion. *Active vision, MIT Press, Cambridge, MA, USA*, pages 263–284, 1993.

[60] C. Harris and M. Stephens. A combined corner and edge detection. *Alvey Vision Conference*, pages 147–152, 1988.

[61] R. Hartley. Self-calibration from multiple views with a rotating camera. *3rd European Conference on Computer Vision (ECCV'94), Stockholm, Sweden*, pages 471–478, 1994.

[62] R. Hartley and A. Zisserman. Multiple view geometry in computer vision. *Cambridge University Press*, 2004.

[63] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'97)*, pages 1106–1112, 1997.

[64] A. Hofmann, H. Maas, and A. Streilein. Knowledge-based building detection based on laser scanner data and topographic map information. *International Archives of Photogrammetry and Remote Sensing, Commission III Symposium : Photogrammetic Computer Vision, Graz, Austria*, 34(PartA) :169–174, 9-13 Sept 2002.

[65] R. P. Horaud and O. Monga. *Vision par ordinateur : outils fondamentaux*. 'Editions Hermès, Paris, 1995.

[66] M. Jabbour. *Localisation de véhicules en milieu urbain à l'aide d'un lidar et d'une base de données navigable*. PhD thesis, Université de Technologie de Compiègne, France, Novembre 2007.

[67] L. Jaulin. *Solution globale et garantie de problèmes ensemblistes : application à l'estimation non linéaire et à la commande robuste*. PhD thesis, Université Paris XI Orsay, France, 1994.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[68] S. Julier and J. Uhlmann. A new extension of the Kalman filter to nonlinear systems. *International Symposium on Aerospace/Defense Sensing, Simulation and Controls, Orlando, Floride, USA*, 1997.

[69] S. Julier and J. Uhlmann. The scaled unscented transformation. *IEEE American Control Conference (ACC'02), Anchorage AK, USA*, pages 4555–4559, 8-10 May 2002.

[70] S. Julier, J. Uhlmann, and H. Durrant-Whyte. A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3) :477–482, 2000.

[71] R. E. Kalman. A new approach to linear filtering and prediction problems. *Transaction of the AMSE - Journal of Basic Engineering*, (82) :34–45, 1960.

[72] M. Kieffer. *Estimation ensembliste par analyse par intervalles : Application à la localisation de véhicules*. PhD thesis, Université de Paris Sud, France, 1999.

[73] J. Kim, J. Lee, T. Kang, W. Lee, and Y. Kim. Node based map matching algorithm for car navigation system. *29 th ISATA Symposium*, pages 121–126, 1996.

[74] S. Kim and J. Kim. Adaptive fuzzy-network-based c-measure mapmatching algorithm for car navigation system. *IEEE Transactions on Industrial Electronics*, (48) :432–441, 2001.

[75] S. Kim, K. Kim, S. Lee, and J. Lee. Video navigation system using the geographic hypermedia. *High Resolution Earth Imaging for Geospatial Information, ISPRS Hannover Workshop*, 2005.

[76] A. King. Inertial navigation : forty years of evolution. *General Electric Company Review*, (13), 1998.

[77] L. Kleeman and R. Kuc. Mobile Robot Sonar for Target Localization and Classification. *International Journal of Robotics Research*, 14(4) :295–318, 1995.

[78] R. Laurini and S. Servigne. Panorama des potentialités SIG en 3 dimensions : vers des modèles virtuels 3D de villes. *Revue XYZ (Spécial 4ème forum de la topologie)*, (14) :18–22, 2008.

[79] D. G. Lowe. Object recognition from local scale-invariant features. *International Conference on Computer Vision (ICCV'99)*, 2 :1150–1157, 1999.

[80] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2(60) :91–110, 2004.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[81] R. Malaka, K. Schneider, and U. Kretschmer. Stage-based augmented edutainment. *4th International Symposium on Smart Graphics*, May 23-25 2004.

[82] D. Meizel, O. Leveque, L. Jaulin, and E. Walter. Initial localization by set inversion. *IEEE Transactions on Robotics and Automation*, 18(6) :966–971, Dec 2002.

[83] K. Mikolajczyk, A. Zisserman, and C. Schmid. Shape recognition with edge-based features. *British Machine Vision Conference, Norwich, UK*, 2003.

[84] H. Moravec. Rover visual obstacle avoidance. *7th International Joint Conference on Artificial Intelligence*, pages 785–790, 1981.

[85] D. F. Nahimana, E. Duflos, and J. Marais. Reception state estimation of GNSS satellites in urban environment using particle filter. *IEEE International Conference on Information Fusion (Fusion'08), Cologne, Allemagne*, 1-3 Juillet 2008.

[86] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6) :756–777, 2004.

[87] W. Y. Ochieng, M. Quddus, and R. B. Noland. Map-matching in complex urban road networks. *Brazilian Journal of Cartography*, 55 :1–18, 2004.

[88] S. Persa and P. Jonker. Real-time computer vision system for mobile robot. *SPIE Intelligent Robots and Computer Vision XX : Algorithms, Techniques, and Active Vision*, 4572 :105–114, October 2001.

[89] M. Pollefeys, R. Koch, M. Vergauwen, and L. V. Gool. Visual modeling with a hand-held acquisition of 3d models with a video camera. *IEEE Second International Conference on 3D Digital Imaging and Modeling (3DIM'99)*, pages 14–23, 1999.

[90] P. Puget and T. Skordas. Calibrating a mobile camera. *Image and Vision Computing*, 8(4) :341–348, 1991.

[91] M. A. Quddus, R. B. Noland, and W. Y. Ochieng. A high accuracy fuzzy logic based map matching algorithm for road transport. *Journal of Intelligent Transportation Systems : Technology, Planning, and Operations*, 10(3) :103–115, 2006.

[92] M. A. Quddus, W. Y. Ochieng, and R. B. Noland. Current map-matching algorithms for transport applications : State-of-the art and future research directions. *Transportation Research Part C : Emerging Technologies*, 15(5) :312–328, October 2007.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[93] M. A. Quddus, W. Y. Ochieng, L. Zhao, and R. B. Noland. A general map matching algorithm for transport telematics applications. *GPS Solutions Journal*, (7) :157–167, 2003.

[94] F. Ramos. *Modélisation et validation d'un système d'information géographique 3D*. PhD thesis, Université de Marne La Vallée, 2003.

[95] E. Royer. *Cartographie 3D et localisation par vision monoculaire pour la navigation autonome d'un robot mobile*. PhD thesis, Université Blaise Pascal - Clermont II, France, Septembre 2006.

[96] E. Royer, M. Lhuillier, M. Dhome, and J. Lavest. Monocular vision for mobile robot localization and autonomous navigation. *International Journal of Computer Vision*, 74(3) :237–260, September 2007.

[97] A. Schilling and D. Holweg. 3D GIS for virtual City Models, exploring foreign Cities and reconstructed Buildings. *Computer Graphik Topics*, 2004.

[98] K. Schindler and J. Bauer. A model-based method for building reconstruction. *First IEEE International Workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis (HLK'03)*, page 74, 2003.

[99] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2) :151–172, 2000.

[100] C. Smaili, M. El Badaoui El Najjar, and F. Charpillet. Multi-sensor fusion method using dynamic bayesian network for precise vehicle localization and road matching. *19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'07)*, 1 :146–151, 29-31 Oct. 2007.

[101] G. Sourimant. *Reconstruction de scènes urbaines à l'aide de fusion de données de type GPS, SIG et Vidéo*. PhD thesis, Université de Rennes 1, France, 2007.

[102] S. Sukkarieh. *Low cost, high integrity, aided inertial navigation systems for autonomous land vehicles*. PhD thesis, Australian Centre for Filed Robotics, Department of Mechanical and Mechatronic Engineering, University of Sydney, Australie, 2000.

[103] I. Suveg and G. Vosselman. 3d reconstruction of building models. *XIXth ISPRS Congress, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33-B3 :538–545, 2000.

[104] I. Suveg and G. Vosselman. Localization and generation of building models. *ISPRS Journal of Photogrammetry and Remote Sensing*, 34 :356–361, 2002.

[105] S. Syed and M. Cannon. Fuzzy logic based-map matching algorithm for vehicle navigation system in urban canyons. *ION National Technical Meeting*, pages 982–993, 2004.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[106] Y. Takase, N. Sho, A. Sone, and K. Shimiya. Generation of digital city model. *Journal of the Visualization Society of Japan*, 23(88) :21–27, 6 December 2003.

[107] J. Tardos, J. Neira, P. Newman, and J. Leonard. Robust Mapping and Localization in Indoor Environments Using Sonar Data. *International Journal of Robotics Research*, 21(4) :311, 2002.

[108] S. Teller. Automatic acquisition of hierarchical, textured 3d geometric models of urban environments : Project plan. *DARPA'97*, pages 767–770, 1997.

[109] S. Teller. Automated urban model acquisition : Project rationale and status. *DARPA'98*, pages 455–462, 1998.

[110] S. Thrun. Robotic mapping : A survey. In G. Lakemeyer and B. Nebel, editors, *Exploring Artificial Intelligence in the New Millenium*. Morgan Kaufmann, 2002.

[111] S. Thrun, A. Bücken, W. Burgard, D. Fox, T. Fröhlinghaus, D. Henning, T. Hofmann, M. Krell, and T. Schmidt. Map learning and high-speed navigation in RHINO. In D. Kortenkamp, R. Bonasso, and R. Murphy, editors, *AI-based Mobile Robots : Case Studies of Successful Robot Systems*. MIT Press, 1998.

[112] S. Thrun, D. Fox, W. Burgard, and F. Dellaert. Robust monte carlo localization for mobile robots. *Artificial Intelligence Journal*, (128) :99–141, 2001.

[113] P. H. S. Torr, A. R. Dick, and R. Cipolla. Layer extraction with a bayesian model of shapes. *6th European Conference on Computer Vision (ECCV '00)*, II :273–289, 2000.

[114] E. Trucco and A. Verri. *Introductory Techniques for 3-D Computer Vision*. Prentice Hall, 1998.

[115] J. Tsay. A concept and algorithm for 3d city surface modeling. *International Archives of Photogrammetry and Remote Sensing, Commission III Symposium : Photogrammetic Computer Vision, Graz, Austria*, PartB :283–286, 2002.

[116] F. Tupin and M. Roux. Detection of building outlines based on the fusion of SAR and optical features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1) :71–82, June 2003.

[117] T. Tuytelaars and L. Van Gool. Matching widely separated views based on affine invariant regions. *International Journal of Computer Vision*, 59(1) :61–85, 2004.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Thèse de Cindy Cappelle, Lille 1, 2008

[118] J. Vaganay, J. Bellingham, and J. Leonard. Comparison of fix computation and filtering for autonomous acoustic navigation. *International Journal of Systems Science*, 29(10) :1111–1122, 1998.

[119] E. Wan and R. Van der Merwe. The Unscented Kalman Filter for Nonlinear Estimation. *IEEE Symposium Adaptive Systems for Signal Processing, Commununication and Control (AS-SPCC 2000), Lake Louise, Alberta, Canada*, pages 153–158, October 2000.

[120] E. Wan and R. Van Der Merwe. Chapter 7 : The Unscented Kalman Filter. *Kalman Filter and Neural Networks*, pages 221–280, 2001.

[121] T. Werner and A. Zisserman. Model selection for automated reconstruction from multiple views. *British Machine Vision Conference (BMVC'02)*, pages 53–62, 2002.

[122] T. Werner and A. Zisserman. New techniques for automated architectural reconstruction from photographs. *7th European Conference of Computer Vision - Part II (ECCV'02), London, UK*, pages 541–555, May 2002.

[123] C. E. White, D. Bernstein, and A. L. Kornhauser. Some map matching algorithms for personal navigation assistants. *Transportation Research Part C : Emerging Technologies*, 8(1-6) :91–108, 2000.

[124] O. J. Woodman. An introduction to inertial navigation. Technical Report UCAM-CL-TR-696, University of Cambridge, Computer Laboratory, Aug. 2007.

[125] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 :1330–1334, 2000.

[126] H. Zhao and R. Shibasaki. A vehicle-borne urban 3-d acquisition system using single-row laser range scanners. *IEEE Transactions on Systems, Man, and Cybernetics (Part B)*, 33(4) :658–666, August 2003.

[127] S. Zlatanova, A. Rahman, and M. Pilouk. Present status of 3d gis. *GIM International*, 16 :41–43, June 2002.

[128] S. Zlatanova and F. Van den Heuvel. 3d gis for outdoor ar applications. *Third International Symposium on Mobile Multimedia Systems & Applications*, pages 117–124, 6 December 2002.

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Résumé

Localisation de véhicules et détection d'obstacles - Apport d'un modèle virtuel 3D urbain

Cette thèse s'intéresse à l'apport d'un modèle virtuel 3D urbain pour l'égo-localisation de véhicules intelligents ainsi que la détection et la localisation d'obstacles.

La localisation du véhicule s'appuie sur plusieurs sources d'information : un GPS, des capteurs proprioceptifs (odomètres et gyromètre), une caméra et un modèle virtuel 3D de l'environnement. Les capteurs proprioceptifs permettent d'obtenir une estimation quasi continue de la pose relative du véhicule. Cette estimation de la pose est recalée par des mesures GPS lorsque celles-ci sont disponibles. Afin de pallier la dérive de la localisation à l'estime lors de longues indisponibilités des informations GPS, on construit une observation cartographique 3D. Celle-ci est basée sur le recalage entre le modèle virtuel 3D urbain et les images acquises par la caméra. Des travaux expérimentaux illustrent l'approche développée.

Par ailleurs, l'apport d'un modèle virtuel 3D urbain est également étudié pour la détection et la localisation des obstacles. Une fois localisé dans le modèle 3D, les obstacles de l'infrastructure tels que les bâtiments sont connus et localisés. Pour détecter les obstacles n'appartenant pas à l'infrastructure (véhicules, piétons ...), on compare l'image réelle et l'image virtuelle en considérant que ce type d'obstacles est présent dans l'image réelle mais absent de l'image virtuelle. A partir de l'information de profondeur disponible grâce au modèle 3D, les obstacles détectés sont ensuite géo-localisés. Les résultats expérimentaux obtenus sont comparés et validés grâce à un télémètre laser.

Mots-clés : Localisation de véhicules intelligents, Détection d'obstacles, Modèle virtuel 3D urbain, GPS, Caméra

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl

Abstract

Vehicle localization and obstacles detection with virtual 3D city model

This thesis deals with ego-localization of intelligent vehicles and obstacles detection with virtual 3D city model.

Vehicle localization uses several sources of information : a GPS receiver, proprioceptive sensors (odometers and gyrometer), a video camera and a virtual 3D city model. The proprioceptive sensors allow to continuously estimate the dead-reckoning position and orientation of the vehicle. This dead-reckoning estimation of the pose is corrected by GPS measurements. Moreover, a 3D geographical observation is constructed to compensate the drift of the dead-reckoning localisation when GPS measurements are unavailable for a long time. The 3D geographical observation is based on the matching between the virtual 3D city model and the images acquired by the camera. Experimental results illustrate the developed approach.

Moreover, the contribution of a virtual 3D city model is also studied for the detection and the localization of obstacles. Once the vehicle is localized in the 3D model, the obstacles of the infrastructure like buildings are known and localized. In order to detect the other obstacles as vehicles, pedestrians, ...the real image acquired by the camera and the virtual image extracted from the virtual 3D model are compared, by considering that this kind of obstacles are in the real image but are absent from the virtual image. With the detph information available from the 3D model, the detected obstacle are then localized. Experimental results are compared with Lidar measurements.

Key-words : Intelligent vehicle localization, Obstacles detection, Virtual 3D city model, GPS, Camera

© 2009 Tous droits réservés.

http://www.univ-lille1.fr/bustl