



Groupe opérationnel n°2 du Predit 4  
« Gestion du trafic et des mobilités »



## CAPLOC

Combinaison de l'Analyse d'images  
et de la connaissance de la Propagation  
des signaux pour la LOCALISATION

Convention 10-MT-CVS-001



Rapport scientifique final du projet

Janvier 2015

Juliette MARAIS, Sébastien AMBELLOUIS, Christophe CONIGLIO, Amaury FLANCQUART,  
Cyril MEURIE, Sarab TAY

IFSTTAR

Yassine RUICHEK, Dhouha ATTIA

UTBM

Julien MOREAU

UTBM/IFSTTAR



## Identification

<b>Acronyme du projet</b>	CAPLOC
<b>Titre du projet</b>	Combinaison de l'Analyse d'images et de la connaissance de la Propagation des signaux pour la Localisation
<b>Coordinatrice du projet</b>	Juliette MARAIS (IFSTTAR/LEOST)
<b>Date de début de projet</b>	19 octobre 2010
<b>Date de fin de projet</b>	18 janvier 2015
<b>Partenaires</b>	IFSTTAR/LEOST, UTBM/IRTES-SET
<b>Type de document</b>	Livrable
<b>Date contractuelle de livraison</b>	18 janvier 2015
<b>Accessibilité</b>	La diffusion de ce document est limitée au programme PREDIT

## Gestion des versions

<b>Version</b>	<b>Date</b>	<b>Auteurs</b>	<b>Modifications</b>
1.0	01/12/2014	Juliette Marais	Plan du livrable
1.1	09/12/2014	Juliette Marais	Rédaction de l'introduction et du chapitre 4
1.2	14/01/2015	Yassine Ruichek	Rédaction du chapitre 6
1.3	15/01/2015	Sébastien Ambellouis	Rédaction des chapitres 3 et 4.2
1.4	15/01/2015	Cyril Meurie	Rédaction du chapitre 2
2.0	18/01/2015	Juliette Marais	Conclusions et approbation

## Table des matières

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Contexte . . . . .	4
1.2	Positionnement du projet par rapport à l'existant . . . . .	6
1.3	Organisation du rapport . . . . .	7
<b>2</b>	<b>Analyse d'images pour la détection et la caractérisation des obstacles perturbant la réception des signaux satellitaires.</b>	<b>8</b>
2.1	Détection des zones de ciel/non-ciel dans l'image . . . . .	8
2.1.1	Simplification par Reconstruction Géodésique par Dilatation . . . . .	9
2.1.2	Classification pixellaire par Fisher . . . . .	10
2.1.3	Post-traitement et évaluation de la qualité de la classification . . . . .	11
2.1.4	Repositionnement des satellites et identification des satellites LOS/NLOS . . . . .	13
2.2	Détection des zones de bâtiment/non-bâtiment dans l'image . . . . .	14
2.2.1	Détection de segments . . . . .	15
2.2.2	Création d'un maillage simplifié . . . . .	16
2.2.3	Carte de régions et carte de densité de segments . . . . .	17
2.2.4	Classification en bâtiment VS non-bâtiment . . . . .	17
<b>3</b>	<b>Construction d'un modèle 3D à partir d'une séquence d'images vidéo</b>	<b>19</b>
3.1	Description des architectures . . . . .	19
3.2	Procédure de calibrage . . . . .	20
3.3	Construction d'un nuage de points 3D . . . . .	22
3.4	Extraction de l'empreinte des bâtiments et de leur hauteur . . . . .	25
<b>4</b>	<b>Amélioration de la précision de la localisation</b>	<b>29</b>
4.1	Utilisation de la segmentation ciel/non-ciel . . . . .	29
4.1.1	Détection et exclusion des satellites reçus après réflexion . . . . .	29
4.1.2	Pondération des satellites dans le calcul de la position selon leur état de réception . . . . .	32
4.2	Utilisation du modèle 3D . . . . .	34
<b>5</b>	<b>Conclusions et perspectives</b>	<b>39</b>
<b>6</b>	<b>Valorisation</b>	<b>41</b>
	<b>Bibliographie</b>	<b>44</b>

# 1 Introduction

## 1.1 Contexte

Dans le cadre des objectifs globaux d'amélioration des systèmes de transports, un des objectifs vise à limiter l'impact de l'usage de la voiture, notamment en rendant plus attractifs et plus performants les systèmes de transports publics et en particulier urbains. Dans ce contexte, les nouveaux services de transport intelligents foisonnent. La plupart d'entre eux reposent de façon plus ou moins directe, sur une information de localisation. L'information peut être associée à une personne, à un véhicule, à une flotte. Le panel d'applications est vaste : navigation bien sûr, mais aussi éco-conduite, péage, assurance « pay as you drive », gestion des parkings, limitations de vitesse variables, priorité des bus aux feux... Techniquement, la solution la plus répandue aujourd'hui repose sur l'installation de puces GPS à bord des véhicules, agrémentée des équipements logiciels et matériels capables d'exploiter l'information de position pour délivrer un service : par exemple calcul d'itinéraire et recalage sur la carte pour le guidage automobile, système de communication lorsqu'il s'agit de suivi de flotte etc. Le système de radionavigation par satellites GPS offre en effet le meilleur compromis coût/simplicité/performance aujourd'hui. Il peut être complété par l'utilisation de signaux EGNOS, améliorant ainsi sa précision.

Certains services, tels que la navigation automobile s'accrochent aujourd'hui assez bien des performances obtenues. Le GPS est, la plupart du temps, fusionné à des informations de capteurs proprioceptifs (capteur de roue par exemple quand une mesure est disponible) ou à la carte numérique et permet le suivi du véhicule sur une route suffisamment précisément pour accompagner le conducteur. Les solutions satellitaires correspondent également bien aux demandes actuelles des opérateurs de flottes de transport (y compris transports guidés). Ceux-ci expriment fortement le besoin d'une réduction de la dépendance à l'infrastructure, en raison de l'augmentation des actes de vandalisme sur l'infrastructure et de la hausse des coûts de maintenance des systèmes. Malheureusement, si l'utilisateur urbain d'un système de localisation par satellites est le plus demandeur de précision, il dispose d'un service dégradé (figure 1). En effet, les conditions de propagation des signaux sont fortement liées à la densité des obstacles qui dégradent les performances optimales atteignables. En particulier, EGNOS étant composé de satellites géostationnaires, il est assez peu reçu en canyon urbain [1]. Il est alors difficile de répondre à des besoins exigeants tels que la signalisation d'un tramway, qui nécessite, par endroits, une position précise à 2m.

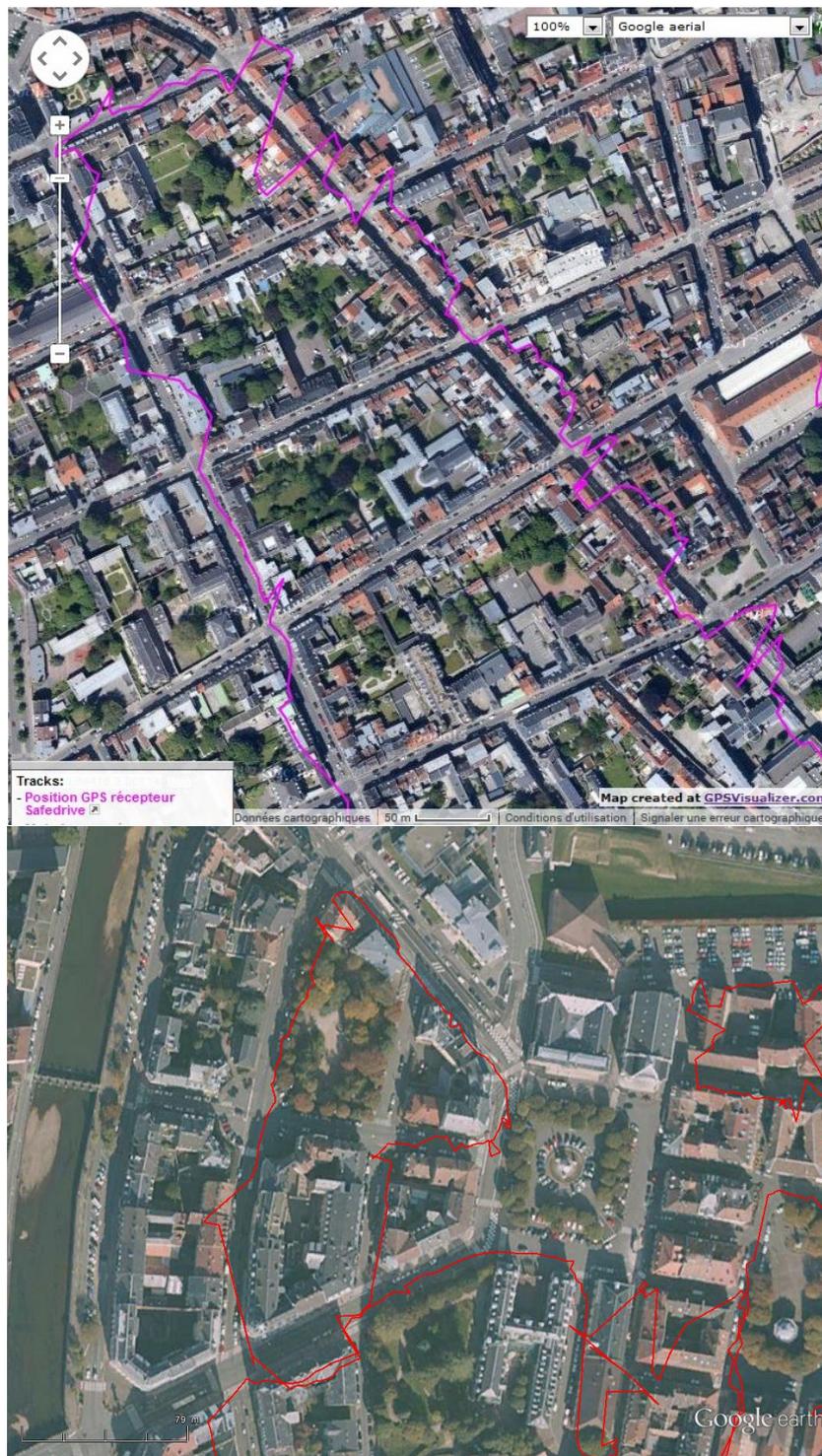


FIGURE 1 – Exemples de mesures GPS imprécises dans les centres de Lille et Belfort

Ainsi, l'objectif de CAPLOC est de délivrer une information de localisation précise, accessible en tous lieux, dans un contexte de réduction globale des coûts. CAPLOC est un projet de recherche amont qui apporte une brique technologique innovante de localisation précise. Il propose d'améliorer la précision d'un système de localisation grâce à la perception vidéo de l'environnement du véhicule et plus particulièrement l'environnement de propagation autour du récepteur GNSS. En effet, les problèmes de disponibilité et d'imprécision sont directement liés à la densité

des obstacles notamment en environnement urbain dense. Il s'agit de réduire l'effet des conditions de propagation des signaux GNSS, notamment les multi-trajets, sans dégrader la disponibilité du service, ce qui est souvent le cas lorsque, classiquement, des procédures d'exclusion de défauts sont mises en œuvre.

## 1.2 Positionnement du projet par rapport à l'existant

La plupart des solutions visant à rendre les solutions satellitaires de localisation plus précises et plus robustes reposent sur la multiplication des capteurs embarqués. Les solutions multicapteurs peuvent cependant s'avérer plus coûteuses si on multiplie le nombre de capteurs. Dans notre projet, avant de combler les lacunes du GNSS avec d'autres capteurs, nous avons souhaité explorer les performances du GNSS et les enrichir à l'aide de la connaissance de l'environnement de propagation.

Récemment, des approches basées sur la vision artificielle ont été proposées pour localiser un mobile dans son environnement de circulation. Il s'agit d'extraire un certain nombre d'amers visuels dans l'environnement du mobile pour localiser le véhicule et contrôler sa trajectoire [2, 3, 4]. Dans CAPLOC, l'objectif n'est pas de se localiser par l'image mais d'utiliser l'image pour qualifier les performances de localisation d'un récepteur GNSS et les améliorer. En effet, l'approche suivie, brevetée en 2008, repose une connaissance de l'environnement de propagation fournie par une ou plusieurs caméras et des traitements d'images capables de détecter l'état de réception des signaux satellitaires et d'apporter une connaissance des perturbations causées par les obstacles environnants.

On trouve dans la littérature plusieurs approches de détection et de mitigation des satellites reçus par multitrajets. Les multitrajets reçus en l'absence de trajet direct sont qualifiés de NLOS (Non Line Of Sight). En environnement contraint, tel qu'en urbain, les principaux perturbateurs sont les obstacles proches de l'antenne de réception (« proche » étant relatif face aux 20000km d'altitude des satellites). L'impact de ces obstacles est directement visible sur l'amplitude des signaux reçus, répercutée sur la valeur du rapport signal sur bruit (CN0). Dans les travaux de [5], une solution basée sur le seuillage de l'amplitude du signal GNSS reçu a été déjà proposée, ainsi que dans [6]. La difficulté de la détection des satellites LOS/NLOS sur ce critère réside dans le choix de la valeur du CN0 seuil. Ce seuil pourra être sensiblement différent d'un équipement à l'autre. De plus, il est fixé à partir de données statistiques et la dispersion des valeurs autour de ce seuil implique de mauvaises détections. C'est pourquoi d'autres techniques sont aujourd'hui étudiées. L'outil PREDISSAT du LEOST a été développé dans cet objectif en 2002 [7] et était fondé sur l'utilisation de caméras classiques et d'un processus de post-traitement. [8] a repris le concept avec une caméra Infra Rouge. Avec le développement des modèles 3D urbains, [9], [10], [11] ou encore [12] ont publié en 2011-2012 des résultats montrant comment ces modèles pouvaient également permettre de déterminer la présence ou non d'un trajet direct en comparant les positions des obstacles modélisés et les positions des satellites. Si ces propositions sont intéressantes, l'intérêt du dispositif CAPLOC n'en est pas diminué. En effet, il présente les avantages suivants :

- Les modèles doivent être mis à jour : les structures peuvent changer temporairement ou définitivement. La vidéo permet d'avoir une vue toujours à jour.
- Le modèle d'une ville représente un volume conséquent de données à stocker (offline), transférer (online). Avec un traitement temps réel, le stockage peut se contenter de contenir les quelques secondes précédant la mesure uniquement.
- Il n'est pas nécessaire d'estimer la position du récepteur dans le modèle. La caméra est placée sur le véhicule et la distance obstacles est immédiatement connue.

Dans les perspectives, nous avancerons des pistes pour l'évolution de notre approche dans ce

contexte.

*Le positionnement par rapport à l'état de l'art des propositions scientifiques plus spécifiques adressées dans le projet feront l'objet des introductions de chaque section.*

### **1.3 Organisation du rapport**

Ce rapport propose une synthèse des travaux scientifiques réalisés dans le projet.

Le projet CAPLOC s'articule autour de 3 tâches techniques. Les objectifs des deux premières tâches sont respectivement de développer une méthode de segmentation automatique du ciel et des bâtiments obstruant la réception directe des satellites et de construire un modèle 3D de l'environnement situé autour de l'antenne GNSS. La troisième tâche est dédiée à l'exploitation des sorties des deux tâches précédentes afin d'augmenter la précision de la localisation.

Dans le rapport, chaque tâche scientifique du projet aura une section dédiée. La première section fera ainsi un bilan des travaux liés à l'analyse d'images pour la détection et la caractérisation des obstacles perturbant la réception des signaux satellitaires. La deuxième sur la construction d'un modèle 3D à partir des données vidéos. La troisième sur l'apport des informations issues de ces deux tâches pour l'amélioration de la précision de la localisation. La conclusion générale reprendra les avancées majeures réalisées au cours du projet et les perspectives identifiées à court et moyen terme.

## 2 Analyse d'images pour la détection et la caractérisation des obstacles perturbant la réception des signaux satellitaires.

Dans cette section, liée à l'analyse d'images pour la détection et la caractérisation des obstacles perturbant la réception des signaux satellitaires, nous avons fait le choix de présenter, les stratégies de classification ciel/non-ciel et de bâtiments/non-bâtiment, décrites plus en détails dans les livrables des tâches 1.2 et 1.3 (ainsi que dans le manuscrit de thèse de Dhouha Attia [13] et le rapport de stage de projet de fin d'études d'ingénieur de Christophe Coniglio [14]) de sorte à pouvoir par soustraction, identifier les zones de végétation et donc de disposer d'une caractérisation fine (ciel, végétation, bâtiment) de l'environnement de réception des signaux satellitaires. Nous ne présenterons donc pas ici l'état de l'art sur l'extraction de descripteurs de couleur et de texture (présenté dans le livrable de la tâche 1.1) et ayant servi à la définition d'une méthode de segmentation combinant de manière adaptative et non paramétrique les informations de couleur et de texture d'une image (présenté dans le livrable de la tâche 1.3) écartée de la stratégie de classification bâtiment/non-bâtiment, mais qui est prometteuse et qui méritent donc d'être poursuivie au delà du projet CAPLOC, notamment pour caractériser et reproduire la texture des différentes surfaces à plaquer sur les plans du modèle 3D.

### 2.1 Détection des zones de ciel/non-ciel dans l'image

Dans cette section, nous présentons le synopsis global d'une stratégie de traitement d'images permettant de détecter les zones de « ciel » et de « non-ciel » d'une image fisheye acquise en mobilité. Ceci permet après repositionnement des satellites dans cette même image, d'exclure ou de pondérer l'impact de chacun d'entre eux dans le calcul d'une nouvelle position de localisation. Ce synopsis illustré sur la figure 2 est composé d'une brique d'acquisition de données, de cinq briques de traitement d'images, d'une brique d'évaluation et d'une dernière brique de calcul d'une position de localisation. La première brique concerne donc la phase d'acquisition des données de localisation (à l'aide des récepteurs GNSS embarqués) et images (à l'aide de la caméra couleur à objectif fisheye installée sur le toit du véhicule). Pour la partie traitement d'images, nous commençons par une phase de simplification de l'image afin de réduire les effets de luminosité, ombres, etc. Une classification pixellaire est ensuite réalisée afin d'affecter à chaque pixel de l'image, une classe d'appartenance parmi les deux classes possibles que sont le « ciel » et le « non-ciel ». Une étape de post-traitement morphologique est ensuite effectuée pour optimiser la phase de classification « ciel »/« non-ciel ». A ce moment de la stratégie, une évaluation des résultats est effectuée par rapport à une vérité terrain (c'est-à-dire par rapport à une classification faite à la main par un expert visant à identifier les régions de « ciel » des régions de « non-ciel » de l'image). Ceci permet d'évaluer le pourcentage de pixels correctement classés par notre approche par rapport à une classification de référence. Il est important de signaler que les algorithmes ont été testés sur la totalité des images de la base et évalués par rapport à une vérité terrain. L'étape suivante consiste à repositionner les satellites dans l'image afin d'identifier ceux se situant dans une zone de « ciel » (satellites Line-Of-Sight (LOS)) ou de « non-ciel » (satellites Non-Line-Of-Sight (NLOS)). Cette dernière étape de traitement d'images effectuée, il devient possible de calculer une nouvelle position de localisation en excluant/pondérant les différents satellites suivant leur classification dans l'image (pour ce dernier point le lecteur se référera à la tâche 3 du projet).

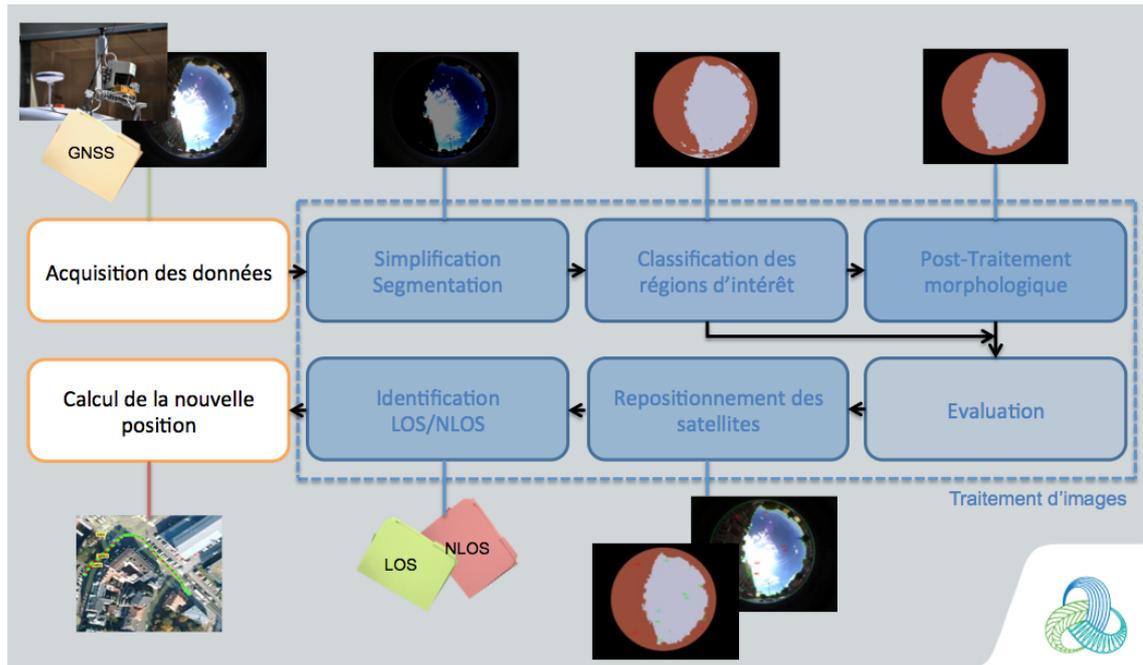


FIGURE 2 – Synopsis de la stratégie de détection de ciel proposée

### 2.1.1 Simplification par Reconstruction Géodésique par Dilatation

La simplification de l'image est une étape importante de notre stratégie de détection de « ciel » et de « non-ciel ». Elle permet d'atténuer les effets des changements brusques de luminosité, ombres, etc venant perturber la classification des pixels de l'image. Une importance toute particulière est donnée au temps de traitement de l'approche globale proposée. Nous avons dans le livrable de la tâche 1.2 décrit rapidement le principe d'un ensemble d'algorithmes de simplification (lissage) de la littérature (filtre médian, filtre gaussien, filtre moyenneur, filtre de Nagao, reconstruction géodésique par dilatation) et illustré leur comportement sur une image de la base d'images de test. Nous ne décrirons donc ici que la méthode de simplification qui a été retenue à savoir la reconstruction géodésique par dilatation (RGD), mais avant cela, il est nécessaire de rappeler quelques définitions basiques de morphologie mathématique. Rappelons que ces définitions sont présentées avec beaucoup plus de détails dans les travaux [15, 16].

**Définition 1** : Un maximum régional noté  $M$  est un plateau tels que les niveaux de gris des pixels voisins sont inférieurs à la valeur  $I(M)$ .

**Définition 2** : Un plateau en un pixel  $p$  d'une image  $I$ , noté  $Plt_p(I)$  est une composante connectée contenant  $p$  et ayant un niveau de gris constant égal à  $I(p)$ .

Notons  $Max_{reg}(I)$  l'ensemble des maximums régionaux de l'image  $I$ . En effet,  $M$  peut ne pas être un seul pixel mais une région de l'image. Par conséquent, un maximum régional  $M \subseteq \mathbf{IR}^2$ . Formellement,  $M$  est un maximum régional si et seulement si il existe un voisinage  $\mathcal{V}_M$  qui satisfait l'équation 2.

$$\forall p \in \mathcal{V}_M \setminus \{M\} \quad I(p) < I(M) \quad (1)$$

La reconstruction géodésique par dilatation est une technique des plus utilisées pour l'extraction des maximums régionaux [17]. En effet, cette technique fait partie des méthodes de

filtrage par lissage de l'image utilisant les opérateurs de la morphologie mathématique (dilatation, érosion, ouverture, fermeture, filtre alterné séquentiel, etc). La RGD utilise deux images en entrée : une image *masque*  $I$  et une image *marqueur*  $I'$ , telles qu'elles vérifient la condition  $\forall p, I'(p) < I(p)$ . Le principe de la RGD visant à extraire les maximums régionaux de l'image consiste à appliquer une suite finie de dilatation géodésique jusqu'à idempotence.

**Définition 3** : L'idempotence est une description de stabilité du résultat tel que le résultat obtenu à l'itération  $(n + 1)$  est identique au résultat obtenu à l'itération  $n$ .

Dans le cas usuel, la formule de la RGD est donnée par l'équation 2 relatée dans [18] :

$$I_{RGD} = \bigvee_{(n \geq 1)} \delta_I^{(n)}(I') \quad (2)$$

$$\delta_I^{(n)}(I') = \delta_I^{(1)} \circ \delta_I^{(1)} \circ \dots \circ \delta_I^{(1)}(I') \quad (3)$$

$$\delta_I^{(1)}(I') = (I' \oplus B) \wedge I \quad (4)$$

avec

$$\forall p \in I, \quad I'(p) < I(p) \quad (5)$$

où  $\oplus$  correspond à une dilatation morphologique avec l'élément structurant  $B$  et  $\delta_I^{(1)}(I')$  est le premier rang de dilatation géodésique.  $\delta_I^{(n)}(I')$  correspond au résultat de composition (notée par  $\circ$ ) des dilatations géodésiques jusqu'à stabilité (idempotence donnée par  $n^{eme}$  rang).

À partir de l'équation 2 et en prenant  $\forall p \in I, I'(p) = I(p) - 1$ , nous pouvons obtenir les maximums locaux  $Max_{reg}(I)$  par l'équation 6 :

$$Max_{reg}(I) = \{p \in I / (I - I_{RGD})(p) > 0\} \quad (6)$$

Les maximums locaux ont généralement une différence de niveaux de gris supérieure à 1. Soit  $H$  la hauteur d'un maximum régional (pic), l'image *marqueur*  $I'$  doit être obtenue en soustrayant une valeur égale à  $H$ . Autrement dit,  $\forall p \in I, I'(p) = I(p) - H$ . Nous pouvons facilement noter que la RGD a permis de supprimer les pics locaux de l'image en permettant d'avoir des zones plus homogènes en niveaux de gris en fonction de la valeur de  $H$  choisie.

### 2.1.2 Classification pixellaire par Fisher

Disposant d'une image simplifiée, nous avons maintenant un choix à effectuer entre, d'une part, segmenter l'image en régions d'intérêt pour effectuer ensuite une classification de ces régions en deux classes (« ciel » et « non-ciel ») et, d'autre part, procéder directement à une classification au niveau du pixel. La première solution est très utilisée mais présente l'inconvénient par rapport à la deuxième, d'ajouter une étape supplémentaire et donc d'augmenter le temps de traitement de l'approche globale. L'objectif ici n'étant que de distinguer les régions « ciel » des régions « non-ciel », la deuxième solution revenant à effectuer une classification pixellaire en deux classes nous semble suffisante. Nous avons dans le livrable de la tâche 1.2 décrits rapidement le principe des algorithmes de classification non-supervisée (KMlocal, Fuzzy C-means, Fisher, Statistical Region Merging) et supervisée (K-plus proches voisins, Bayes, SVM) que nous avons testés. Nous ne décrirons donc ici que le classifieur qui a été retenu pour la qualité de ses résultats et le temps d'exécution réduit, à savoir l'algorithme de Fisher. L'algorithme de Fisher fait partie

des algorithmes de classification non-supervisée et à la particularité de ne s'appliquer que sur des images en niveaux de gris [19, 20]. L'idée consiste à diviser l'histogramme de l'image en niveaux de gris en  $K$  classes en minimisant la somme des variances des classes. En effet, cela consiste à localiser des seuils représentant les séparateurs de la distribution des données. Soient  $s_k$  le seuil associé à la classe  $k$  et  $m_k$  la moyenne des intensités des pixels appartenant à la classe  $k$ . La répartition des pixels en  $K$  classes revient à localiser les seuils  $\{s_k\}_{k=1\dots K}$  comme des séparateurs de distributions en utilisant un critère qui minimise la somme des inerties associées aux différentes classes. Il est important de souligner que l'algorithme ne s'appliquant que sur des images en niveaux de gris, nous l'avons appliqué sur chaque composante de l'espace couleur RGB de sorte à ne retenir que celle qui maximise les résultats.

$$Histogram(s_k) = \min_{p \in ]m_k, m_{k+1}[} (Histogram(p)) \quad (7)$$

A chaque classe  $k$ , on associe une taille notée  $T_k$  (8) et une variance notée  $V_k$  (9)

$$T_k = \sum_{p \in D_k} \{Histogram(p)\} \quad (8)$$

$$V_k = \frac{1}{T_k} \sum_{p \in D_k} (p - m_k)^2 \cdot Histogram(p) \quad (9)$$

$D_k = ]s_{k-1}, s_k[$  est la gamme des niveaux de gris de la classe  $k$ . Le critère de minimisation est donné par l'équation (10) :

$$W = \sum_{k=1}^K T_k \cdot V_k \quad (10)$$

où  $m_k$  représente la moyenne pondérée donnée par l'équation (11)

$$m_k = \frac{1}{T_k} \cdot \sum_{p \in D_k} p \cdot Histogram(p) \quad (11)$$

### 2.1.3 Post-traitement et évaluation de la qualité de la classification

Une fois la classification pixellaire (en deux classes) réalisée, nous procédons à une étape de post-traitement. Celle-ci est réalisée à l'aide d'opérateurs issus de la morphologie mathématique et consiste à reclasser des pixels mal étiquetés (par exemple, des pixels classés ciel et situés à l'intérieur d'une zone de bâtiments, ou encore un/deux pixels dans une zone de végétation) en utilisant l'information de voisinage.

La méthode d'évaluation utilisée pour optimiser les différents paramètres des algorithmes et évaluer la qualité de la classification ciel/non-ciel obtenue a été adaptée à notre application et utilise une vérité-terrain. Ainsi, elle tient compte de cinq mesures : 1/ le pourcentage de pixels faisant partie de la région « ciel » dans l'image classée et dans l'image de référence ; 2/ le pourcentage de pixels classés comme « ciel » dans l'image classée et ne faisant pas partie de la région « ciel » dans l'image de référence ; 3/ le pourcentage de pixels classés comme « non-ciel » dans les deux images ; 4/ le pourcentage de pixels classés comme « non-ciel » dans l'image classée et qui ne font pas partie de la région « ciel » dans l'image de référence ; 5/ la somme de la première et de la troisième mesure. Cette dernière mesure traduit aussi le pourcentage de pixels correctement classés (appelée « taux de bonne classification » dans la suite de ce livrable).

		Filtres			
		Median		RGD	
		Performance	W	Performance	H
Classification non-supervisée					
Fisher	Rouge	82.27	8	89.60	100
	Vert	91.72	8	93.49	50
	Bleue	97.53	8	96.40	25
	Moyenne RGB	92.76	8	93.55	50
KMlocal	EZHybrid	96.39	8	95.20	25
	Hybrid	96.39	8	95.20	25
	Loyds	96.39	8	95.20	25
	Swap	96.06	8	95.44	25
Fuzzy Cmeans	m=1	96.44	8	95.25	25
	m=1.25	96.35	8	95.28	25
	m=1.5	96.16	8	95.36	25
	m=1.75	95.93	8	95.38	25
	m=2	95.73	8	95.38	25
SRM		73.18	6	97.02	100
Classification supervisée					
Bayes		94.61	8	96.02	100
Kppv	nbKppv = 4	96.68	6	97.83	25
	nbKppv = 6	96.83	8	97.99	25
SVM		88.81	6	97.07	25

TABLE 1 – Résultats de la classification ciel/non-ciel avec différents algorithmes, paramètres et méthodes de simplification, sur la base CAPLOC.

Dans le livrable de la tâche 1.2, nous avons présentés un grand nombre de résultats de classification ciel/non-ciel en fonction de la méthode de simplification, du classifieur et des différents paramètres utilisés, et ce, pour deux bases de données images (l'une acquise au préalable dans le cadre du projet LOCOPROL, et l'autre lors des expérimentations du projet CAPLOC). Nous ne présenterons donc ici, à travers les tableaux 1 et 2, qu'une synthèse de ces résultats. Nous pouvons alors conclure en s'appuyant que dans un cadre non-supervisé, une classification réalisée par l'algorithme de Fisher (sur la composante bleue de l'espace RGB) donne les meilleurs résultats (i.e. 97,53%) si elle est couplée à une simplification par un filtre médian et arrive en deuxième position (derrière SRM) si elle est associée à une simplification par RGD (i.e. 96,40%). Dans un cadre supervisé, l'utilisation d'une RGD offre toujours les meilleurs résultats, mais son utilisation couplée à l'algorithme des Kppv (avec un  $nb_{Kppv} = 6$ ) permet d'obtenir le meilleur taux de bonne classification à savoir 97.99%. Aussi, nous pouvons noter que le temps de traitement de l'approche proposée est peu affecté par la méthode de simplification utilisée, mais peut être jusqu'à 400 fois supérieur selon le classifieur employé (par exemple entre Fisher et Kppv). Afin de respecter la contrainte temps réel imposée pour la détection de la ligne d'horizon et des zones de ciel/non-ciel, nous retiendrons finalement l'algorithme de Fisher (et la composante bleue) comme meilleur classifieur. Une première préférence est donnée au filtre médian par rapport à la RGD mais reste très légère et donc à confirmer sur une base d'images plus importante qui pourra être réalisée au delà du projet CAPLOC.

-	Fisher	KMlocal	FuzzyCmeans	SRM	Bayes	Kppv	SVM
Median (en s)	0.035	0.41	3.30	0.35	0.34	15.3	5.0
RGD (en s)	0.037	0.45	3.34	0.47	0.47	15.4	5.04

TABLE 2 – Temps de traitement de l’approche en fonction du couple classifieur/simplification

#### 2.1.4 Repositionnement des satellites et identification des satellites LOS/NLOS

Une fois l’étape de classification pixellaire et de post-traitement effectuées, nous cherchons à déterminer le nombre de satellites dont les signaux sont reçus en trajets directs (respectivement réfléchis/bloqués). Pour ce faire, nous devons en premier lieu, repositionner les satellites dans l’image classée. La position d’un satellite dans le ciel étant définie par son élévation et son azimut relatif à un cap, positionner les satellites dans l’image nécessite d’extraire ces informations dans les trames NMEA puis de déterminer une équation de projection de coordonnées du monde réel sur l’image. L’objectif fisheye utilisé dans le cadre de nos expérimentations répond à un modèle de projection équisolide. Cette spécificité nous a permis d’établir l’équation de projection validée par une méthode expérimentale réalisée en laboratoire (qui sera détaillée dans une publication à venir) et ainsi reprojetter les satellites dans l’image fisheye acquise en mobilité. La figure 3 illustre ce repositionnement, les satellites se situant dans une région de « ciel » sont représentés par une couleur verte, tandis que ceux qui se situent dans une région de « non-ciel » sont représentés par une couleur rouge (pour la constellation GPS) ou bleue (pour la constellation EGNOS). Nous pouvons alors ensuite aisément identifier les satellites avec des trajets directs (en vert) et ceux dont les signaux sont bloqués ou réfléchis (en rouge ou bleu).

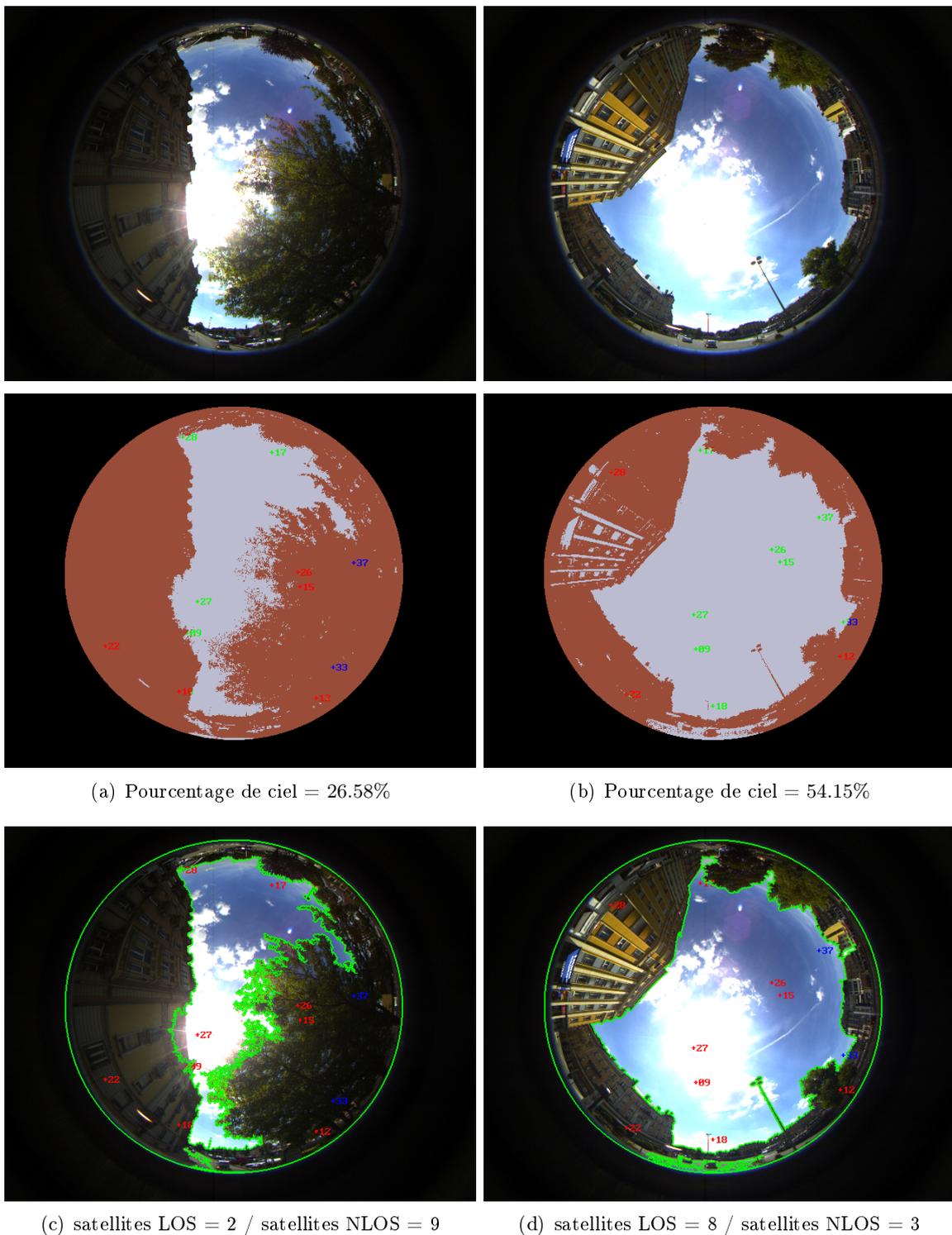


FIGURE 3 – Détection du ciel avec repositionnement des satellites dans l'image (de haut en bas : image initiale, avec une classification par Fisher après simplification par RGD).

## 2.2 Détection des zones de bâtiment/non-bâtiment dans l'image

Dans cette section, nous décrivons la stratégie de détection de bâtiment basée en grande partie sur une analyse de segments, et qui couplée à celle de la détection de ciel décrite ci-dessus, nous permettra par soustraction d'identifier les zones de végétation et donc de disposer d'une

caractérisation fine (ciel, végétation, bâtiment) de l'environnement de réception des signaux satellitaires. Cette stratégie décrite dans le livrable de la tâche 1.3 et détaillée dans le rapport de stage de projet de fin d'étude d'ingénieur de Christophe Coniglio, s'articule autour d'une détection de segments sur laquelle nous établissons deux hypothèses. La première est qu'un bâtiment, par sa constitution, contient un nombre important de segments. La deuxième est que le ratio entre le nombre de segments d'un bâtiment et le nombre de segments total de l'image est proche de 1. Nous avons donc statistiquement une probabilité importante de trouver un bâtiment dans une zone à forte concentration de segments. La stratégie de détection de bâtiments que nous avons développée est composée de cinq étapes décrites en détails ci-après : 1/ Détection des segments dans l'image ; 2/ Création d'un maillage simplifié ; 3/ Conversion en une carte de régions ; 4/ Création d'une carte de densité de segments ; 5/ Classification des régions en deux classes : bâtiments/non-bâtiment.

### 2.2.1 Détection de segments

L'étape de détection de segments est l'étape principale de la stratégie globale de détection de bâtiments, puisque le résultat final dépend en grande partie de la qualité de cette détection de segments. Comme introduit ci-dessus, nous faisons l'hypothèse qu'il existe une forte densité de segments dans une zone de bâtiments et que la proportion de segments détectés dans une zone de bâtiments est nettement plus importante que dans les autres zones réunies. De nombreuses méthodes de détection de segments existent dans la littérature, mais nombre d'entre elles voient leur paramètres optimisés pour une application visée. Dans le livrable de la tâche 1.3, nous avons analysé trois méthodes très utilisées (Transformée de Hough, Dseg, LSD) en fonction d'une de nos contraintes applicatives, à savoir le capteur optique utilisé : une caméra couleur avec un objectif fish-eye (offrant un champ de vision de plus de 180 degrés mais avec des distorsions importantes). Nous ne présenterons donc ici que celle qui offre les meilleurs résultats à savoir l'algorithme LSD.

La méthode LSD (qui signifie "a Line Segment Detector") est une technique de détection de segments qui se concentre sur l'orientation du gradient et une sélection stricte des segments détectés. La première étape de l'algorithme consiste à regrouper des pixels ayant une orientation de gradient proche. Chaque groupe de pixels supposé constitué un segment est validé comme tel après une série de tests capables de détecter des groupes de pixels correspondant à du bruit pour ensuite les éliminer de la sélection.



FIGURE 4 – Détection de segments par l'algorithme LSD (milieu) avec utilisation d'un redimensionnement de l'image (droite) sur une image acquise (gauche).

Les résultats de détection obtenus, avec cet algorithme, dans le cadre de notre application sont prometteurs puisque nous avons un grand nombre de segments détectés dans les zones de

bâtiments et peu dans les autres zones de l'image (végétation, etc). La série de tests réalisés met en avant que l'utilisation du gradient de Sobel en comparaison avec le gradient proposé dans LSD améliore significativement les résultats déjà très satisfaisants. Comme cela est suggéré dans la littérature, nous constatons également que les résultats de détection de segments par cet algorithme sont encore une fois améliorés lorsque un/plusieurs redimensionnements de l'image sont utilisés. En effet, comme l'illustre la figure 4, le fait d'effectuer une détection de segments sur une image redimensionnée, fusionnée à celle réalisée sur l'image initiale, permet d'augmenter le nombre total de segments détectés, ce qui renforce l'une de nos hypothèse de départ qui est d'obtenir davantage de segments dans les zones de bâtiments que dans les autres zones de l'image. En moyenne, nous détectons 85% de segments dans les zones de bâtiments avec le gradient de Sobel (VS 80% avec le gradient de Canny) et 15% de segments dans les autres zones de l'image (VS 20% avec le gradient de Canny). Notons aussi que le redimensionnement de l'image permet d'augmenter avec ce gradient, le nombre de segments détectés de 400 à 1300 en moyenne. Ces résultats qui confortent nos hypothèses de départ, sont donc très satisfaisants.

### 2.2.2 Création d'un maillage simplifié

La segmentation consiste à partitionner une image en régions homogènes c'est-à-dire en un ensemble de points de l'image ayant des propriétés communes. L'étape de segmentation (qui in fine regroupera les pixels pour former des régions) est primordiale pour notre application puisqu'elle influencera directement la qualité/précision de notre classification finale. Si nous nous concentrons sur les résultats de détection de segments ci-dessus, nous constatons que la majorité des contours des bâtiments sont détectés comme des segments (ce qui est dû à la forme géométrique rectiligne de ces derniers). Nous avons donc une forte densité de segments détectés dans les bâtiments mais aussi sur leurs contours. Il est donc utile de relier différents segments afin de créer des formes avec des contours fermés. Celles-ci seront ensuite étiquetées pour créer une carte de régions coïncidant avec des zones de bâtiments.

La connexion de segments s'apparente à la création d'un maillage. Les extrémités de segments compris dans un voisinage donné sont reliées par un nouveau segment créé. La distance maximale entre ces deux extrémités est calculée par rapport à la longueur du segment testé, de sorte à favoriser les plus grands segments. Ces grands segments étant généralement détectés sur les contours des bâtiments, les privilégier augmente les chances de fermer correctement ces contours (ce qui s'avère d'autant plus difficile lorsque la détection est de qualité moindre). Le résultat du maillage relève une concentration importante de régions de petites tailles dans les zones de bâtiments (dû à une forte densité de segments ne pouvant être reliés) contrairement aux autres zones de l'image (végétation) qui laissent apparaître peu de régions mais avec des surfaces plus importantes. Les amas de segments pouvant être provoqués peuvent aller jusqu'à dégrader la qualité de la segmentation de l'image en créant des frontières épaisses (supérieures à 3 pixels) entre les régions. Pour pallier ce problème, nous proposons d'élaguer le maillage réalisé à l'étape précédente en convertissant les pixels identifiés comme segments et reliés entre eux dans un voisinage en 4-connexité en pixels neutres. Néanmoins, cet élagage ne permet pas de supprimer complètement les segments détectés en dehors des zones de bâtiments. Les segments détectés dans les zones de végétation ne posent pas vraiment de problème mais ceux détectés dans les zones de ciel (et principalement à cause des nuages) sont plus problématiques. Nous proposons alors d'utiliser la classification ciel/non-ciel décrite précédemment comme un masque et ne considérer pour la détection de segments que les pixels non-ciel, et de supprimer les segments ayant une de leur extrémités présente dans une zone de ciel. La figure 5 illustre un exemple de maillage obtenu à partir d'une image de détection de segments. Le lecteur souhaitant davantage de détails sur cette partie pourra consulter le livrable de la tâche 1.3.



FIGURE 5 – Illustration d’une détection de segments (milieu), et de la création d’un maillage (droite) sur une image acquise en mobilité (gauche).

### 2.2.3 Carte de régions et carte de densité de segments

Une fois cette étape de création de maillage effectuée, nous disposons d’une segmentation de l’image utilisable. Il reste néanmoins à étiqueter les différentes zones détectées en différentes régions disposant chacune d’un numéro d’identification qui leur est propre. D’autre part, la classification des différentes régions bâtiment VS non-bâtiment, requiert une analyse de la densité de segments autour de chaque pixel. Pour ce faire, nous recherchons uniquement les pixels segments, puis incrémentons par pas de 1, tous ceux qui se situent dans un voisinage donné selon une distance euclidienne. Il est assez difficile de se faire une idée de la distance optimale pour chaque pixel segment. Néanmoins, les tests que nous avons réalisés dans le livrable de la tâche 1.3 nous permettent de conclure qu’une valeur de 42 pixels donnent les meilleurs résultats, soit environ 1/3 de la hauteur moyenne des bâtiments.

### 2.2.4 Classification en bâtiment VS non-bâtiment

L’étape de classification constitue l’étape finale de notre processus de détection de bâtiments. Il s’agit de définir, pour chaque pixel, à partir de la carte de régions et de la carte de densité de segments toutes deux calculées précédemment, sa classe d’appartenance parmi les deux classes possibles que sont : bâtiment et non-bâtiment. L’utilisation de la carte de densité de segments permet d’associer à chaque région, une densité moyenne de segments (en divisant la somme de densité par le nombre de pixels présents dans la région traitée). Pour optimiser la prise de décision sur la classe d’appartenance associée à chaque région de l’image, le processus de classification contient quatre étapes rapidement énoncées ci-après, mais décrite en détails dans le livrable de la tâche 1.3. Dans la première étape, nous commençons par considérer que les régions dont la densité de segments moyenne est significativement plus élevée que la densité de segments moyenne de l’image correspondent à des zones de bâtiments. Cette première étape de classification comporte logiquement de nombreuses erreurs, principalement dues aux regroupements de petites régions, qui de part leurs caractéristiques, ont une forte probabilité d’être classées comme bâtiment. La deuxième étape du processus de classification consiste alors à supprimer certaines erreurs en isolant les petites régions. Cela revient à réorganiser l’étiquetage de la carte de régions, en regroupant toutes les régions connexes de même classe (pixel classé comme bâtiment ou comme non-bâtiment). Dans la troisième étape, nous répétons la sélection réalisée dans la première étape en supprimant les nouvelles régions dont la moyenne de densité de segments est trop faible (ce qui a pour effet de supprimer les petites régions isolées) ainsi que les régions parasites. L’étape finale de notre processus de classification consiste à lever l’incertitude sur la présence de bâtiment ou non. Pour ce faire, nous comparons la moyenne de densité de segments de chaque région supposée être considérée comme bâtiment, à une constante dont la valeur a été déterminée suite à de nombreux tests expérimentaux. Si la moyenne en question est inférieure à cette constante, la zone

est alors considérée comme non-bâtiment. La figure 6 illustre le résultat final d'une détection de bâtiment sur un exemple d'image acquise en mobilité. Cette classification bâtiment/non-bâtiment réalisée et couplée à la classification ciel/non-ciel de la section précédente, nous permet par soustraction, d'identifier les zones de végétation et donc de disposer d'une caractérisation fine (ciel, végétation, bâtiment) de l'environnement de réception des signaux satellitaires.

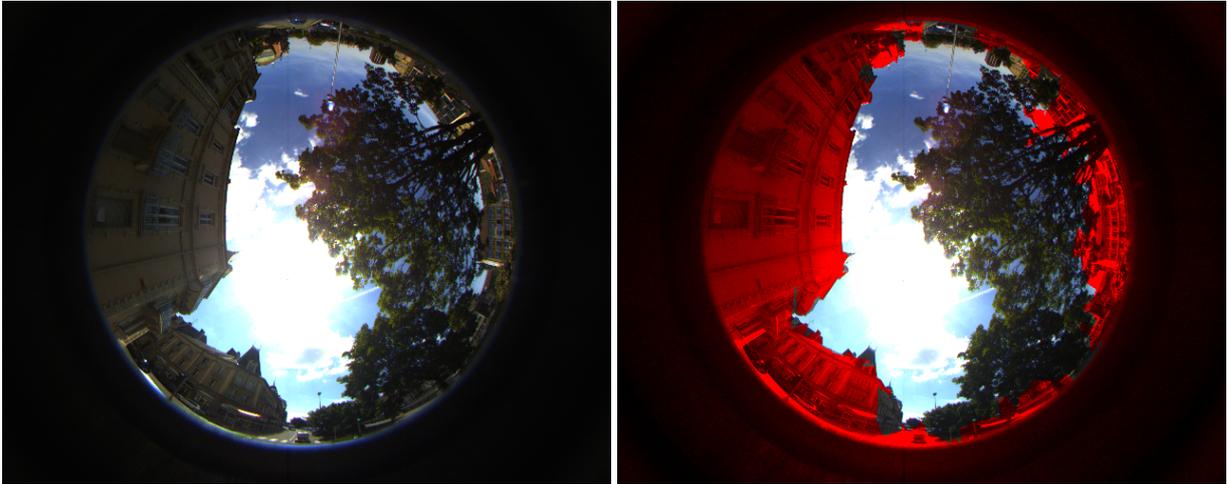


FIGURE 6 – Illustration de la détection de bâtiments (droite) sur une image acquise en mobilité (gauche).

### 3 Construction d'un modèle 3D à partir d'une séquence d'images vidéo

Cette section a pour but de présenter la synthèse des travaux réalisés en matière d'estimation de la structure 3D d'un environnement urbain à partir du traitement d'une séquence d'images multi-vues. Deux architectures de perception ont été définies et utilisées. La première est semblable à un système stéréoscopique composé de deux caméras munies d'un objectif fisheye dont les deux axes optiques sont parallèles et orientés verticalement vers le ciel. Les traitements proposés mettent en correspondance le contenu d'un couple d'images acquises simultanément.

La deuxième architecture exploite une seule caméra fisheye couleur dont l'axe optique est vertical et orienté vers le ciel. Elle est moins encombrante que l'architecture précédente et permet (1) d'estimer la structure 3D de l'environnement et (2) de renseigner directement (sans nécessité de calibrage) la nature des éléments de la scène urbaine à partir des résultats de la segmentation et de la classification ciel/bâtiment/végétation (outil développé dans la tâche 1). Les propriétés 3D des surfaces observées sont estimées à partir de la mise en correspondance du contenu d'images acquises à des instants successifs au cours du déplacement du mobile (depth from motion).

Dans la suite de cette section, après un rappel des deux architectures CAPLOC utilisées et leur spécificité, nous décrivons les étapes nécessaires à l'obtention du modèle 3D : la procédure de calibrage, la mise en correspondance "multi-vues" pour l'obtention d'un nuage de points et l'extraction de l'empreinte des bâtiments. Nous présentons une évaluation intermédiaire de la qualité du nuage de points en utilisant des images de synthèse qui nous permettent de générer des vérités terrain. Pour terminer, nous montrons qu'il est possible de positionner le mobile par rapport à l'empreinte des bâtiments.

#### 3.1 Description des architectures

Nous commençons par la description du système stéréoscopique retenu et sa mise en oeuvre. Les choix effectués quant aux éléments du système ont été guidés par la large analyse bibliographique qui a été menée et détaillée dans le livrable 2.1. Le système mis en place permet de travailler avec deux architectures pour lesquelles la méthodologie menant au nuage de points 3D est différente. L'une (Arch 1) est fondée sur la mise en correspondance du contenu de deux images acquises au même instant mais à deux positions différentes (stéréovision classique). Une phase de calibrage est réalisée à l'initialisation du système. Elle permet d'estimer la géométrie épipolaire pour faciliter la mise en correspondance stéréoscopique et pour retrouver le facteur d'échelle de la reconstruction. L'autre architecture (Arch 2) est réduite à une seule caméra et exploite le déplacement du mobile pour produire deux vues à deux positions différentes. Elle constitue une architecture multi-vues non rigide pour laquelle il est requis d'estimer le déplacement du mobile entre chaque vue pour estimer la géométrie épipolaire.

La figure 7 illustre le dispositif caméras installé sur le toit du camion laboratoire. Les trois caméras fisheye sont synchronisées et reliées à un système d'acquisition haut débit géoréférencé. Grâce à ce dispositif, il est possible de se placer dans l'une ou l'autre de l'architecture décrite précédemment simplement en exploitant une seule ou plusieurs caméras. Ce dispositif nous a permis de produire des bases de séquences d'images que nous avons utilisées pour le développement et l'évaluation des différents algorithmes proposés. La caméra centrale (caméra 2) du dispositif est utilisée pour réaliser la tâche de classification bâtiment/ciel/végétation des éléments de la scène (tâche 1). Les deux autres caméras (caméra 1 et 3) sont distantes d'un entraxe de  $2m$  et sont utilisées pour la reconstruction 3D de la scène. Lorsque le problème CAPLOC est placé sous la seconde architecture, seule la caméra centrale est utilisée simplifiant ainsi le système.

La figure 8 représente un couple d'images fournies par le dispositif embarqué. Il s'agit ici d'un couple stéréo obtenu à partir des caméras 1 et 3.



FIGURE 7 – Dispositif Caploc.

La tâche 2 comporte une action dont l'objectif est d'identifier la nature de l'objet auquel appartient chaque point 3D reconstruit i.e. si l'objet est un bâtiment, un élément de la végétation ou du ciel. Dans le cas de l'architecture mono-caméra, la correspondance est assez simple car la nature de l'objet est déterminée dans l'espace image de la caméra centrale et les points 3D de la scène sont calculés à partir des coordonnées de tous les pixels de cette même image. Lorsque nous utilisons trois caméras, il est nécessaire de connaître la matrice de passage entre le repère des caméras 1-3 et le repère de la caméra 2 pour assurer la correspondance. Le calibrage est donc une étape particulièrement importante dont la qualité influe directement sur la précision du nuage de points 3D et du modèle 3D calculés.

### 3.2 Procédure de calibrage

Le processus de mise en correspondance du contenu de plusieurs images nécessite la détermination des paramètres intrinsèques et extrinsèques de chaque caméra ayant servi à produire chaque image considérée. Ces paramètres sont estimés lors d'une phase de calibrage. Dans le cas de l'architecture à 3 caméras, ces paramètres sont déterminés une fois à l'installation du système puis ré-estimés ponctuellement au cours du déplacement du mobile afin de compenser l'effet des vibrations sur les parties mécaniques du système. Ce recalibrage ponctuel permet de garantir un niveau de précision suffisant et n'impacte globalement que les paramètres intrinsèques du système. Dans le cas de l'architecture mono-caméra, le calibrage est requis tout au long du trajet puisque la transformation entre les repères caméra pris à des positions successives n'est pas fixe. Seuls les paramètres intrinsèques de la caméra sont constants. Par ailleurs, il faut noter que, pour cette architecture, les acquisitions sont réalisées tous les  $2m$  en triggant la caméra. Cette spécificité permet d'estimer le facteur d'échelle à appliquer sur le modèle afin d'avoir une échelle 1 :1.

Les paramètres intrinsèques décrivent les paramètres internes à la caméra, c'est-à-dire la projection en une image et la relation entre le repère métrique centré sur le centre optique de la caméra (repère caméra 3D) et les coordonnées image en pixel (repère image 2D). Ils représentent les paramètres du modèle géométrique de la caméra. Les paramètres extrinsèques décrivent le lien entre un repère du monde réel et un repère centré sur la caméra. Ils définissent la transformation appliquée pour passer d'un repère à l'autre. Elle est composée d'une translation et d'une rotation.

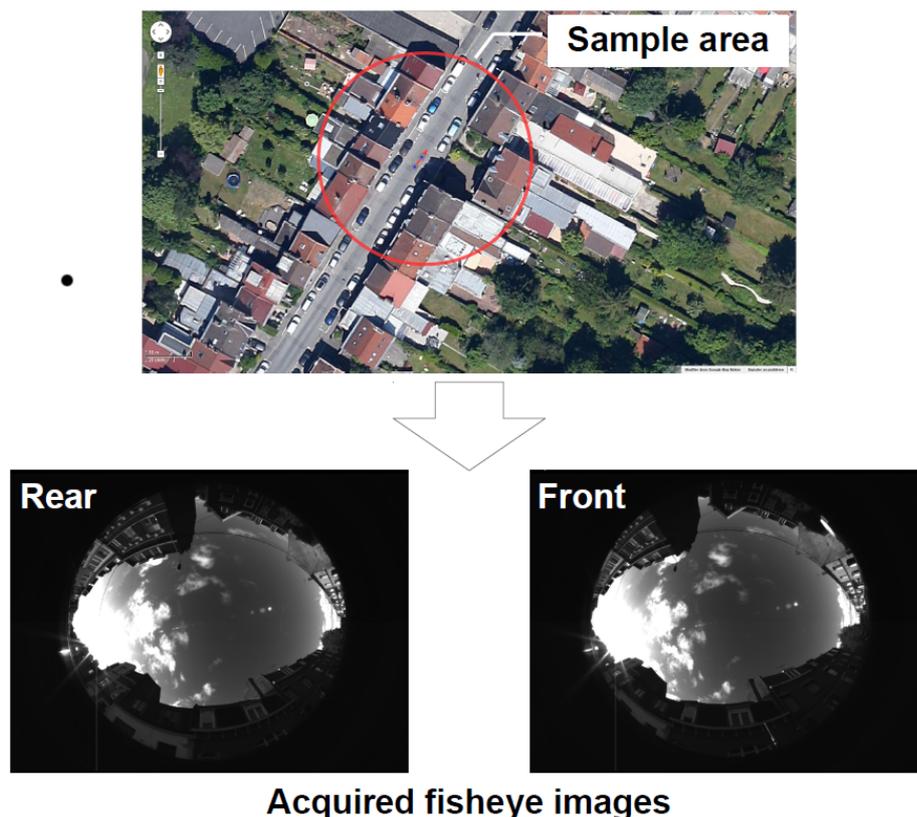


FIGURE 8 – Couple d'images fisheye fournies par le dispositif.

De nombreuses solutions ont été proposées dans la littérature pour des systèmes vidéo classique c'est à dire des systèmes utilisant des caméras dont la projection est perspective (modèle du sténopé). Avec l'apparition des systèmes omnidirectionnels, ces solutions algorithmiques ont dû être adaptées. Le système CAPLOC est basé sur des caméras munies d'objectifs fisheye dont la projection est sphérique. Dans [21], Mičušík formalise mathématiquement ce mode de projection ainsi que le problème du calibrage (notamment les paramètres intrinsèques). Il propose des résultats pour une projection de type "équidistante". Dans CAPLOC, nous utilisons des objectifs pour lesquels le type équidistant n'est pas adapté. Nous avons développé un processus de calibrage pour une projection sphérique de type "équisolide". La figure 9 illustre la projection fisheye, les caractéristiques de la courbe de droite dépendant du type de projection retenu (ici équisolide).

Nous avons mis en place une procédure de calibrage automatique. De nombreuses méthodes proposent de calibrer de tels systèmes à partir d'une mire présentée manuellement devant les caméras dans plusieurs positions et à différentes distances. Dans la cas d'Arch 1, l'entraxe entre les caméras est grande ce qui nécessiterait une mire très grande à déplacer dans un espace étendu. De plus, pour Arch 2, un calibrage devant être réalisé pour chaque couple, une procédure manuelle est à proscrire. L'objectif étant de déterminer la géométrie épipolaire du système, le processus proposé cherche à déterminer les valeurs de la matrice fondamentale en s'inspirant de la méthode proposée dans [22] et [21].

Notre méthode propose d'extraire automatiquement un ensemble de points d'intérêt dans un couple d'images (couple stéréo ou couple d'images acquises successivement) puis de les mettre en correspondance. L'estimation des valeurs de la matrice fondamentale est réalisée en minimisant l'erreur globale de mise en correspondance par l'algorithme RANSAC. L'erreur globale est calculée à partir des paramètres estimés avec  $n$  paires de points prises aléatoirement dans la liste des points d'intérêt mis en correspondances. Dans [21],  $n=9$  (modèle de caméra sphérique) et dans [23],

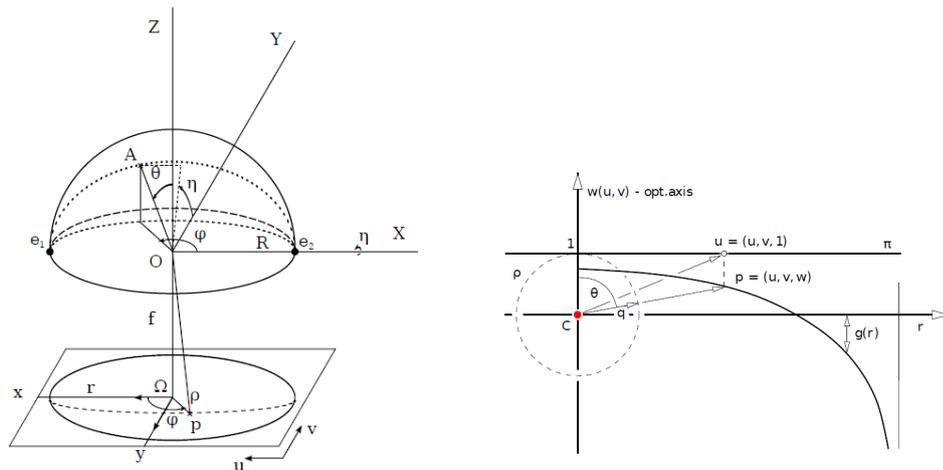


FIGURE 9 – Projection d'un objectif fisheye.

$n = 8$  ou  $7$  selon la méthode d'estimation (modèle de caméra sténopé classique).

Il est important de retrouver les mêmes points d'intérêt dans les images d'un couple afin de garantir l'existence de bons appariements. Nous avons fait une étude assez complète sur la robustesse des points d'intérêt à un changement de point de vue. Nous l'avons complété en étudiant avec quels descripteurs de la littérature nous devons les associer pour obtenir les meilleurs résultats. La particularité de notre système étant la projection sphérique des objectifs, nous avons également évalué des détecteurs définis directement dans cet espace. Nous en avons conclu que le détecteur SIFT associé au descripteur SIFT apportait les meilleurs résultats et les meilleurs compromis. Les détecteurs sphériques n'apportaient que très peu d'amélioration et augmentaient de temps de traitement de manière très importante. Par ailleurs, nous avons établis que le résultat était sensible à la distribution des  $n$  points d'intérêt dans les images.

Généralement le nombre d'itérations du ransac est élevé et offre une première estimation de la matrice qui est ensuite affinée. La matrice fondamentale permet de déterminer les deux épipoles et de déterminer les courbes épipolaires conjuguées le long desquelles le processus de mise en correspondance est réalisé. La figure 10 représente la géométrie épipolaire de deux prises de vue et la construction des courbes épipolaires. La figure 11 illustre le résultat obtenu par le calibrage sur un couple d'images fisheye. Nous rappelons que ces images ont été acquises soit simultanément à partir de deux caméras, soit au cours du déplacement du mobile par une unique caméra. Connaissant la matrice fondamentale, il est alors possible de déterminer toutes les courbes épipolaires conjuguées sur chacune d'entre elles.

### 3.3 Construction d'un nuage de points 3D

Nous avons utilisé deux méthodes d'appariement, toutes les deux fondées sur une technique d'optimisation globale. La première exploite la programmation dynamique et est inspirée de [24]. Elle a été adaptée afin de prendre en compte plusieurs couples de courbes épipolaires conjuguées. La deuxième méthode optimise une fonctionnelle par une technique de coupure de graphe.

Toutes deux sont appliquées sur les pixels situés le long des courbes épipolaires conjuguées. Aucune rectification n'est réalisée sur les images afin d'éviter les artefacts dus aux interpolations et elles permettent toutes deux d'effectuer une mise en correspondance dense des pixels des deux images. L'ensemble des courbes épipolaires conjuguées utilisées est défini avec pour contrainte que l'espace des courbes épipolaires passe par tous les pixels de l'image.

Il est alors possible de déterminer de manière analytique, et pour tout couple apparié, les

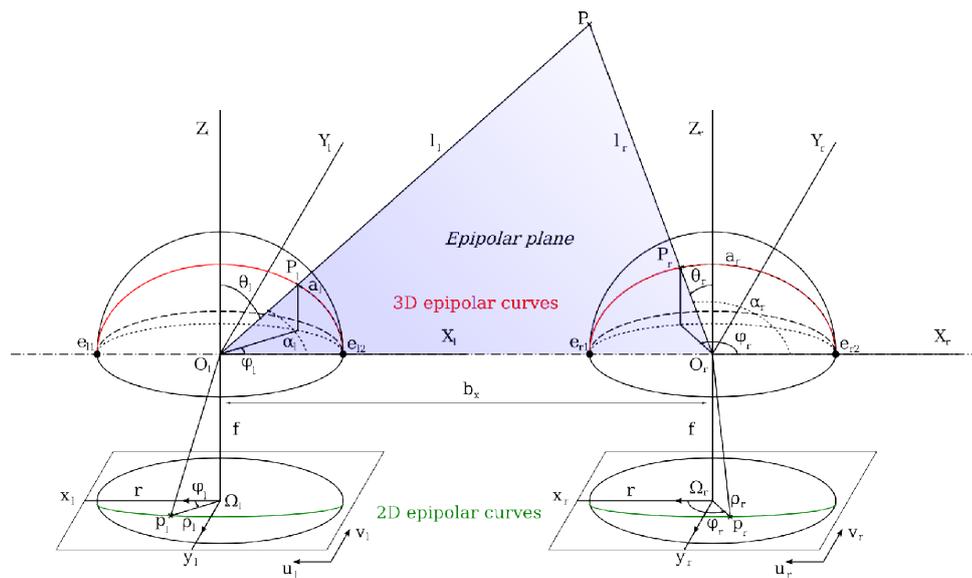


FIGURE 10 – Configuration épipolaire de prises de vue correspondantes à deux images fisheye.

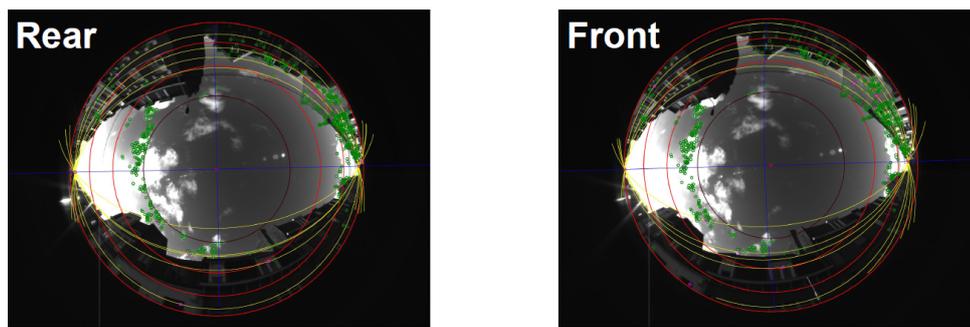


FIGURE 11 – Projection d'un objectif fisheye.

coordonnées du point de la scène qui s'est projeté en chaque pixel du couple considéré. La figure 12 montre une carte de distances estimée sur un couple d'images réelles stéréoscopiques par programmation dynamique. Des résultats similaires ont été obtenus en prenant un couple d'images à des instants consécutifs (Arch 2). Sur cette même figure est représenté le nuage de points produit à partir de la carte des distances obtenue par programmation dynamique. La figure 13 montre une carte de distances estimée sur un autre couple d'images réelles en appliquant cette fois une optimisation par coupure de graphe. La première ligne de cette figure montre les deux images du couple pour lesquelles nous avons aligné les courbes épipolaires avec les lignes horizontales des images. L'optimisation par coupure de graphe donne un nombre très réduit de faux appariements dans le ciel et respecte assez bien les discontinuités les contours des bâtiments. L'implantation actuelle n'étant pas optimale, nous l'avons testée sur des images de taille divisée par 4 et pour un nombre de disparités angulaires réduit. Par conséquent, l'image des distances obtenue souffre d'une quantification très importante des distances comme le montre la première image dans le bas de la figure 13 qui se manifeste sur le nuage de points par une distribution des points par palier.

Une évaluation de l'appariement par programmation dynamique a montré une erreur moyenne en distance inférieure à  $2.6m$  sur des objets situés à moins de  $50m$ . Cette erreur moyenne augmente de manière importante lorsque les objets sont situés à des distances plus importantes dans l'axe du véhicule. L'évaluation a été menée sur 30 couples d'images synthétisées d'une scène ur-

baine et pour lesquels une vérité terrain a pu être calculée. Chaque couple d'images a été produit à une position différente dans la scène urbaine. La figure 14 présente les résultats obtenus. Sur les images de la dernière colonne les pixels rouges sont les pixels pour lesquels la distance est erronée.

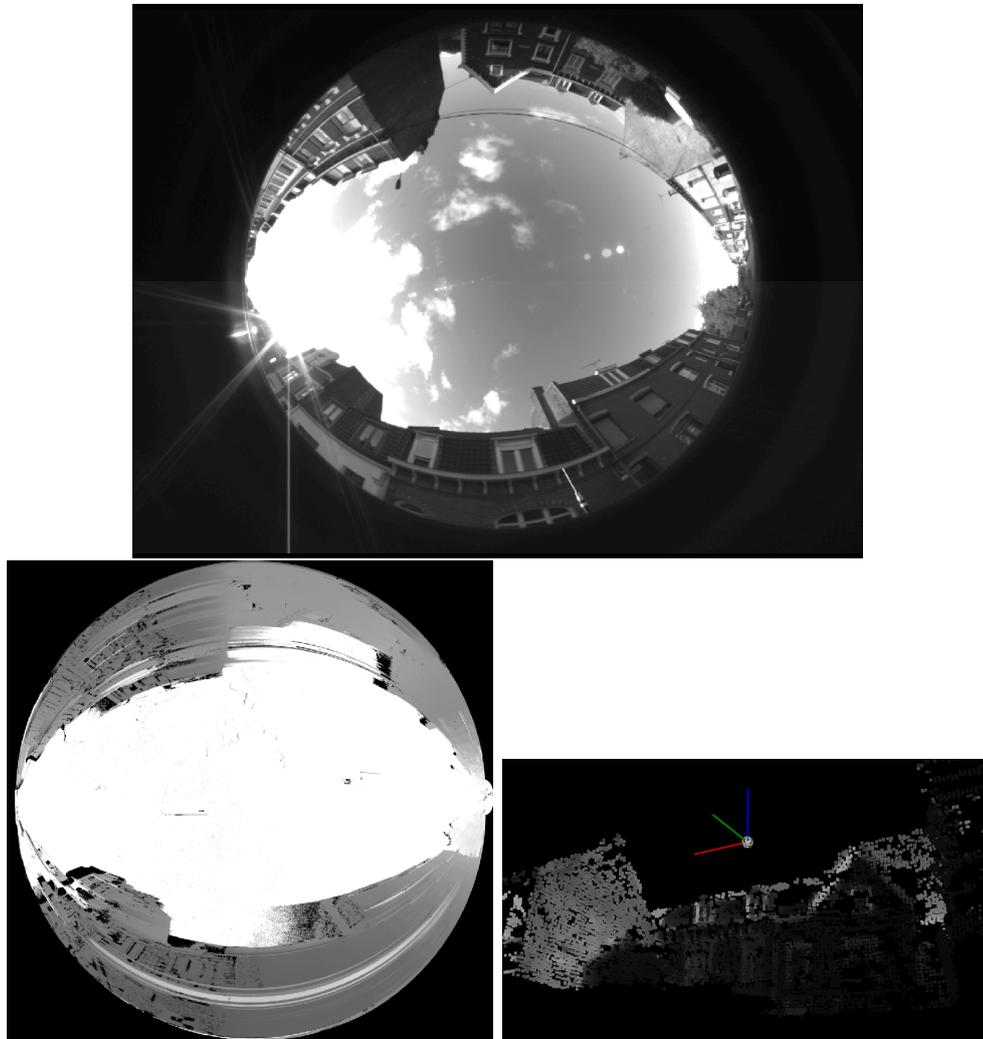


FIGURE 12 – (En haut) une image du couple d'images appariés. (En bas) une carte des distances obtenues par programmation dynamique et une vue partielle du nuage de points correspondant.

Un nuage de points est déterminé en un ensemble de positions du mobile sur son trajet. Une étape de fusion des nuages de points a été développée afin de combiner les nuages de points successivement lors du déplacement du mobile. Ainsi, à un instant donné, le nuage de points résultant est plus précis sur les parties où le mobile est déjà passé. La fusion que nous proposons exploite le processus de calibrage précédemment décrit. Que nous soyons dans le cas mono caméra ou dans le cas stéréoscopique, l'étape de calibrage permet d'estimer la matrice de passage du repère à l'autre de deux nuages de points consécutif. La figure 15 représente le résultat de la fusion de deux nuages de points. La figure 16 représente le résultat de la fusion de tous les nuages de points obtenus le long d'une trajectoire.

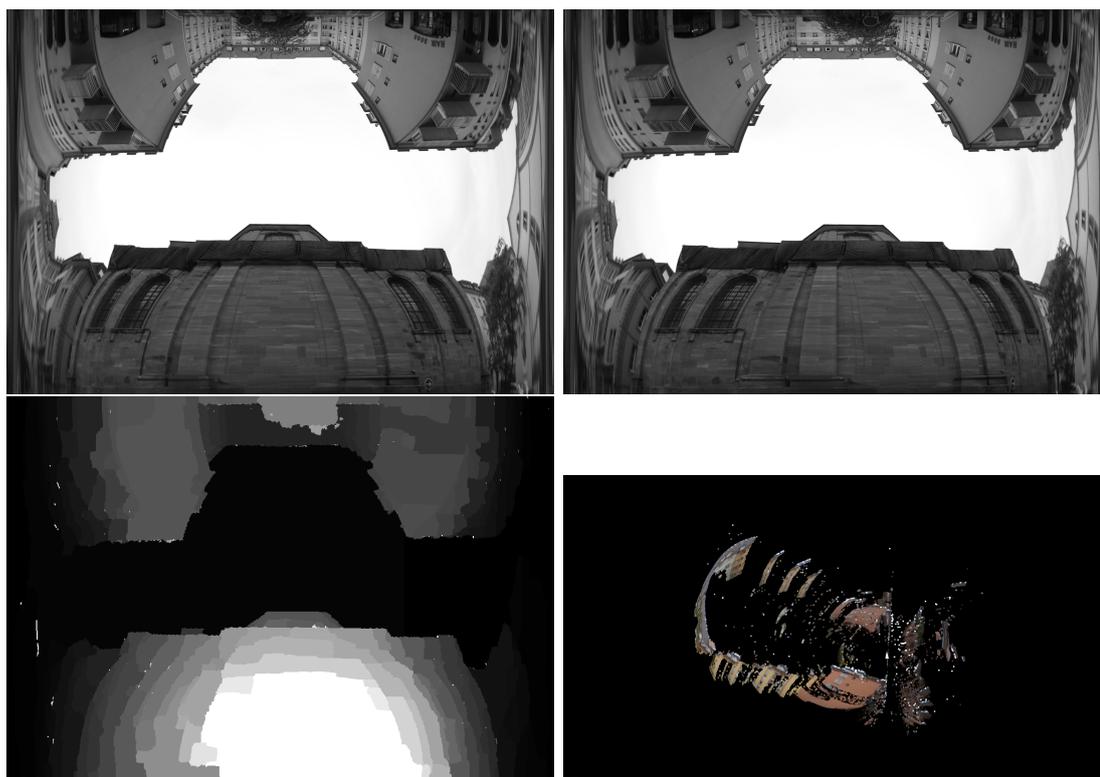


FIGURE 13 – (En haut) un couple d’images à appairer dont les courbes épilaires ont été alignées avec les lignes des images. (En bas) une carte des distances obtenues par coupure de graphe et une vue partielle du nuage de points correspondant.

### 3.4 Extraction de l’empreinte des bâtiments et de leur hauteur

L’objectif du projet CAPLOC est d’exploiter la structure du nuage de points que nous obtenons régulièrement le long de la trajectoire du mobile afin de corriger les erreurs de pseudo-distance dues aux réflexions multiples des signaux GNSS sur les objets et les façades autour du mobile. Pour cela, nous avons besoin (1) de connaître la position des surfaces réfléchissantes (principalement celles des bâtiments) et (2) de placer le mobile par rapport à ces surfaces. Nous avons développé une méthode d’extraction de l’empreinte des bâtiments à partir d’un nuage de points. Dans un premier temps, la méthode opère la projection le long de l’axe vertical et sur le plan horizontal de tous les points du nuage. Le résultat obtenu est similaire à un histogramme 2D du nuage. La projection exploite sur une grille régulière définie dans le plan horizontal en comptant, pour chaque élément de la grille, les points dont l’altitude est non nulle. Pour un élément de la grille, plus le nombre de points est élevé, plus il est certain de l’existence d’une structure à cet endroit. La figure 17 présente un nuage de points et l’empreinte des bâtiments correspondante. Dans cette figure, l’empreinte est représentée par des segments rouges détectés en appliquant une transformée de Hough probabiliste [25]. Une comparaison avec des mesures prises sur le terrain nous permet de conclure que l’empreinte des bâtiments est localisée avec une erreur maximale de  $30\text{cm}$  dans des configurations relativement simples. Dans des configurations plus complexes telles que les carrefours, les environnements présentant de la végétation ou les nuages de points présentant des problèmes d’appariements, l’erreur augmente et il est possible d’extraire des empreintes de bâtiments non existants (figure 18). Sur cette figure, l’image de droite montre le nuage de points reprojété dans l’image du couple servant de référence. Il apparaît clairement des points dans les nuages qui pourraient provoquer la fausse détection de segments et par conséquent l’extraction d’empreinte de bâtiments qui n’existent pas. Par ailleurs, cette

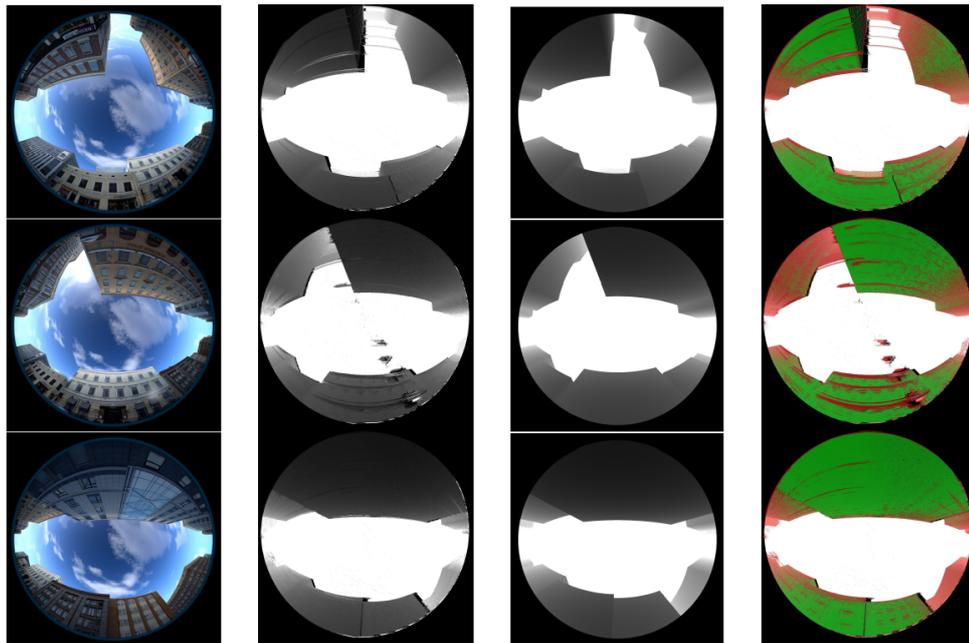


FIGURE 14 – Résultats de l'évaluation menée sur des images de synthèse d'une scène urbaine. De la première à dernière colonne : image de la caméra avant, carte des distances, carte des distances vérité terrain, carte des erreurs.

image présente un bâtiment arrondi pour lequel une transformée de Hough donne de mauvais résultats. La connaissance des matrices de projection estimées lors de l'étape de calibrage nous permet alors de localiser le mobile dans ce nuage et de calculer sa position par rapport à chaque surface de réflexion des signaux GNSS. Sur la figure 17 les deux caméras du stéréoscope sont représentées par deux cercles. L'une des caméras est estimée à une distance de  $13.98m$  et  $4.96m$  de deux bâtiments situés de part et d'autre de la rue.

L'utilisation du modèle est exposée dans le chapitre suivant.



FIGURE 15 – Fusion de plusieurs nuages de points consécutifs.



FIGURE 16 – Fusion des nuages de points obtenus le long d'une trajectoire complète.

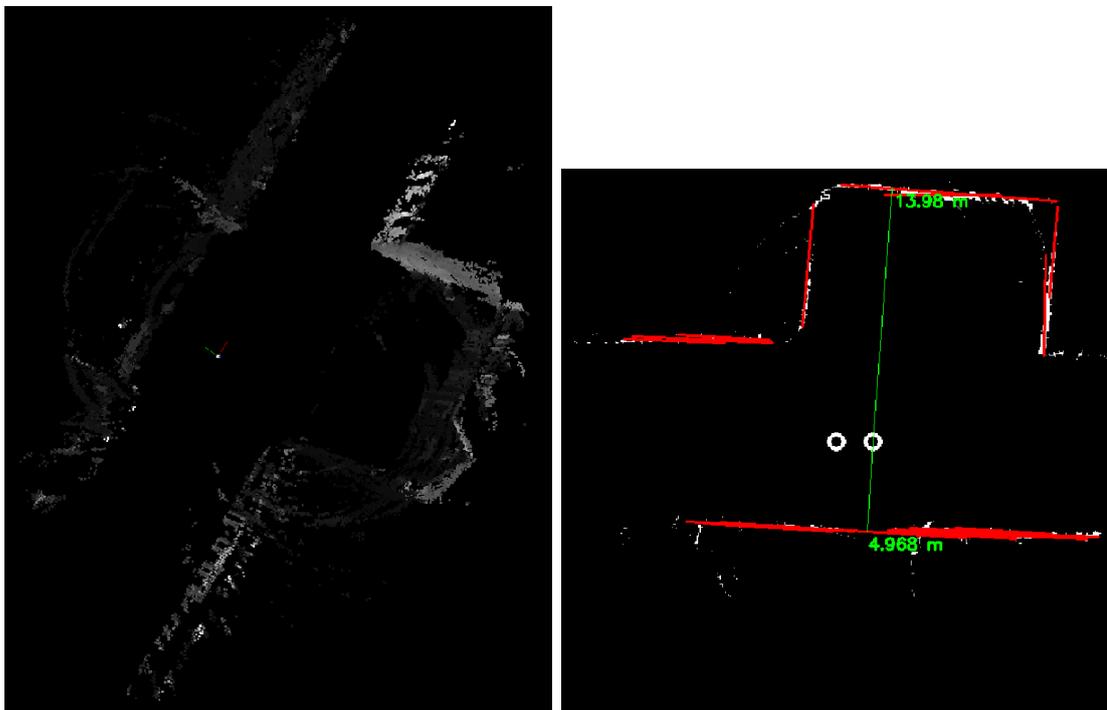


FIGURE 17 – Empreinte de bâtiments extrait d'un nuage de points. L'empreinte est présentée après une rotation.

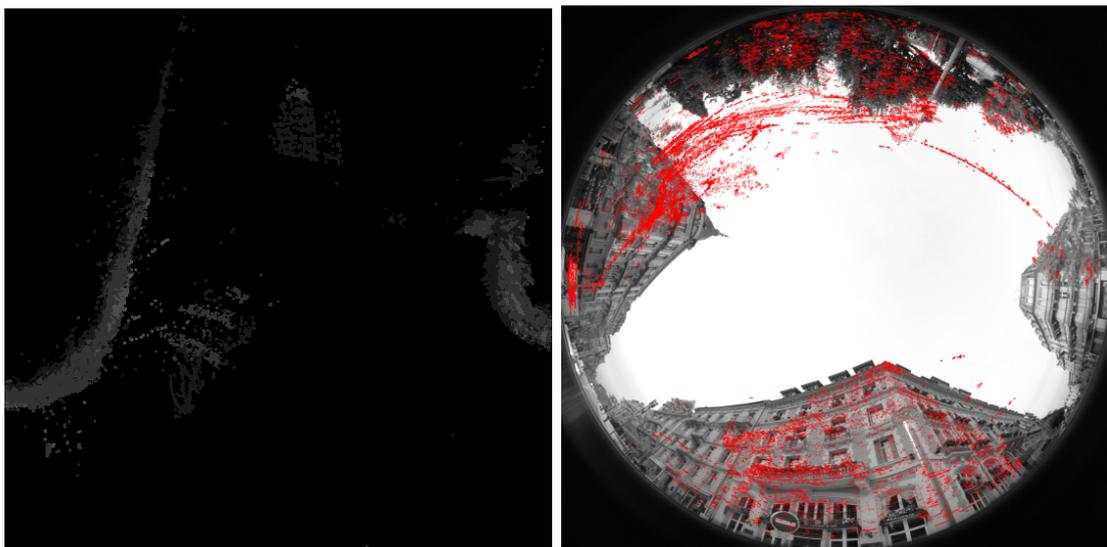


FIGURE 18 – Nuage de points obtenus sur un couple d'images présentant un carrefour avec végétation. Histogramme calculé à partir de ce nuage de points.

## 4 Amélioration de la précision de la localisation

La finalité de l'utilisation de l'image dans CAPLOC est l'amélioration de la précision de la localisation dans les environnements fortement perturbés par les multitrajets. Ce chapitre est composée de deux sections. La première montre comment utiliser la classification des signaux reçus en LOS (Line-Of-Sight, c'est à dire direct) et NLOS (Non Line-Of-Sight, soit les satellites indirects, reçus après réflexion) issue du traitement des images fish-eye. La deuxième illustrera l'utilisation du modèle 3D pour l'estimation des erreurs de pseudo-distance causée par ces réflexions.

### 4.1 Utilisation de la segmentation ciel/non-ciel

#### 4.1.1 Détection et exclusion des satellites reçus après réflexion

Il existe de nombreuses méthodes de mitigation des multitrajets tout au long de la chaîne de réception et de traitement des signaux GNSS [26] : traitement d'antennes, modélisation du canal de propagation, développement de nouveaux corrélateurs, algorithmes de navigation prenant en compte les bruits induits, ajouts de capteurs supplémentaires tels que capteurs inertiels ou même cartographies numériques. Dans le projet CAPLOC, le choix de la méthode s'est porté sur la détection et l'exclusion des satellites potentiellement perturbateurs de la qualité de la mesure.

Cette méthode a été réalisée dans le cadre de la thèse de Dhouha Attia et consiste, comme [11], à exclure les satellites classés en zone de non-ciel donc considérés en NLOS. La détection étant effectuée grâce au traitement d'images, nous montrons dans ce chapitre les résultats de la technique de détection et d'exclusion des satellites NLOS.

*Cette section est rédigée à partir du chapitre 5 de la thèse de D. Attia [27] (soutenue le 3 octobre 2013).*

Dans le but d'améliorer la position calculée à partir des signaux disponibles, nous appliquons une politique d'exclusion des satellites dont les mesures peuvent être considérées comme dégradées. Sont exclus les satellites se trouvant dans les régions « non-ciel » afin d'éliminer toutes sources de multitrajets. La technique est donc composée d'une première étape de segmentation de l'image en zones ciel et non ciel, suivie d'une étape de classification des signaux reçus selon leur état de réception (dans la zone ciel : LOS, dans la zone non ciel : NLOS). La position est ensuite calculée à partir d'un filtre de KALMAN étendu (EKF) classique implémenté par N. VIANDIER [5] avec les satellites LOS uniquement. La figure 19 illustre cette sélection.

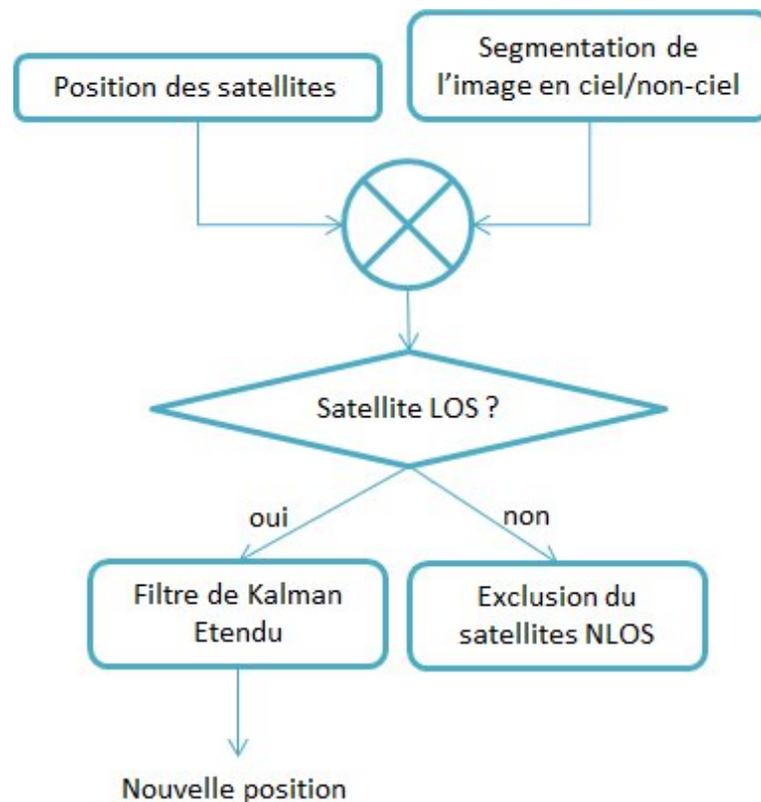


FIGURE 19 – Processus de sélection des satellites LOS

Dans les figures qui suivent, la trajectoire de référence, obtenue grâce à un récepteur RTK, est toujours tracée en rouge et permet d'évaluer le gain des différentes solutions testées. Nous avons appliqué cette méthode sur une portion de trajectoire de 171 points RTK dans le centre de la ville de Belfort. La trajectoire en bleue est constituée des positions estimées à partir des signaux reçus depuis tous les satellites disponibles en utilisant une technique basée sur le filtre de KALMAN étendu. Les pseudo-distances sont fournies par l'équipement Safedrive et utilisées dans le filtre. La trajectoire entière est formée de deux portions quasiment perpendiculaires reliées par un virage. La première ligne droite du trajet est longue d'environ 81 mètres et comprend 32 points de mesure, soit une vitesse moyenne inférieure à 10 km/heure, qui commence par un temps d'arrêt identifiable parce qu'il forme un amas de points. Sur la deuxième ligne droite, on compte 166 mètres et 26 points soit une vitesse moyenne de 23 km/heure.

Les résultats de l'exclusion ont montré qu'en éliminant tous les satellites NLOS, la précision était améliorée sur la première moitié du trajet ainsi que dans le virage où les positions estimées sont visiblement plus proches de la trajectoire de référence mais que la deuxième partie de la trajectoire était dégradée par rapport à la solution utilisant tous les satellites. Nous avons expliqué ceci par l'impact de la géométrie des satellites utilisés lors du calcul de la nouvelle position. Pour illustrer la relation entre la géométrie des satellites et la précision de la position estimée, nous utilisons le GDOP. Rappelons qu'une valeur de GDOP caractérisant une bonne géométrie est classiquement considérée inférieure à 3. A l'exception de quelques points de GDOP largement supérieur à 2, l'ensemble des points calculés avant  $t = 40s$  bénéficient d'une bonne géométrie, même lorsque les satellites NLOS sont exclus (points de couleur jaune et verte). La situation est clairement dégradée sur la période  $t = [40, 60]$  et l'impact sur la trajectoire est directement visible. Cette configuration montre les limites de notre stratégie : les satellites disponibles, même s'ils sont directs, ne sont pas disposés de façon optimale et la position est plus imprécise que lorsque tous les satellites sont utilisés.

Pour appliquer chaque stratégie de façon optimale, l'approche choisie consiste à alterner entre les deux solutions, c'est-à-dire à choisir la politique d'exclusion des satellites NLOS lorsque le DOP reste suffisamment bon (soit inférieur à un seuil choisi empiriquement - ici 3) et à garder tous les satellites dans le cas contraire. La figure 20 visualise l'ensemble de ces résultats : en rouge la trajectoire de référence ; en bleu la solution utilisant tous les satellites ; en vert les satellites directs uniquement et en violet la solution combinée.

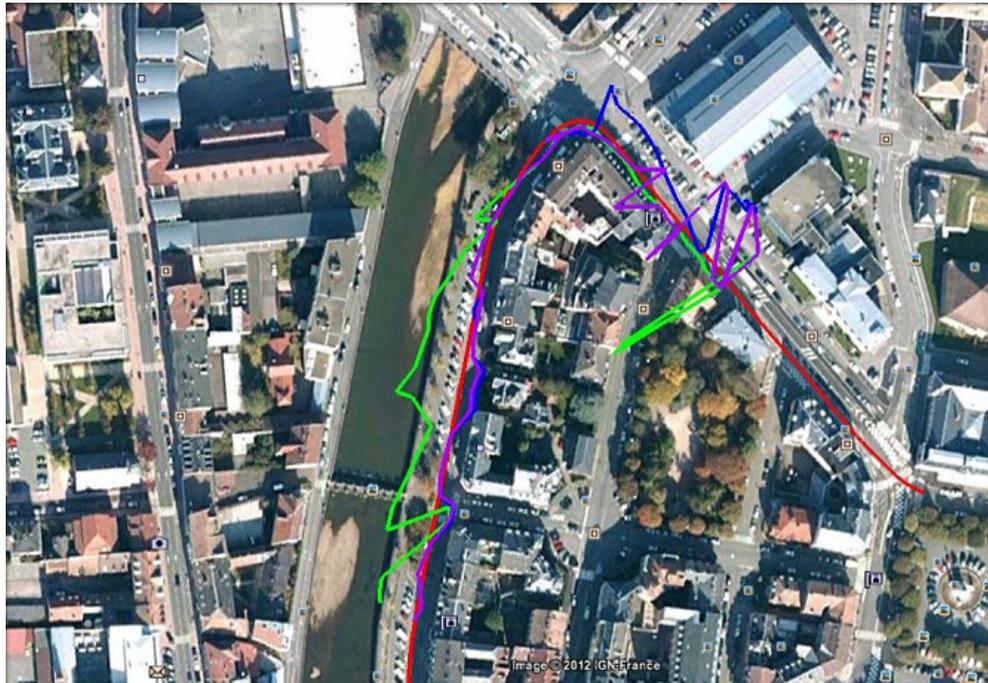


FIGURE 20 – Vue aérienne (Google Earth) des trajectoires obtenues avec les trois solutions (en rouge : la référence, en bleu : tous les satellites utilisés, en vert : les satellites LOS uniquement, en violet : la solution combinée).

L'erreur de localisation est calculée en projetant orthogonalement les points de mesures sur la trajectoire de référence RTK. Nous représentons ces erreurs sous la forme d'une fonction de répartition sur la figure 21 pour les trois solutions : avec tous les satellites, avec les satellites LOS et la solution combinée. Cette représentation met en évidence l'amélioration apportée par les différentes méthodes sur la précision d'origine. On constate par exemple que la médiane de l'erreur passe de 5m avec tous les satellites à 3 m avec le choix final soit une amélioration de 40% ou encore que 95% des positions sont inférieures à 24m contre 70% seulement des mesures utilisant tous les satellites.

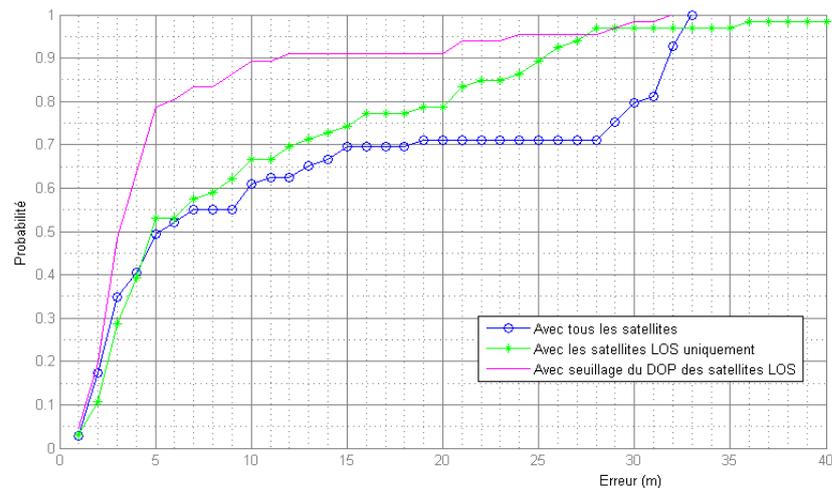


FIGURE 21 – Fonction de répartition de l'erreur de positionnement selon les trois méthodes.

Enfin, le tableau 3 reprend trois mesures calculées pour chacune des politiques testées : l'erreur moyenne, l'erreur maximale et l'écart type. Les politiques testées sont : la sélection des satellites LOS uniquement et la solution combinée DOP/tri. Ces résultats montrent que la stratégie basée sur le seuillage du DOP est toujours meilleure comparée aux deux autres stratégies (basée sur tous les satellites et basée sur l'exclusion des satellites NLOS). La précision des positions estimées est améliorée d'un facteur 2 en utilisant la politique du seuillage du DOP. D'autre part, quelle que soit la politique, l'intégration des satellites reçus à travers la végétation montre une augmentation de l'erreur moyenne. Par conséquent, cette mesure améliore la disponibilité mais dégrade légèrement les performances globales. Ainsi nous pouvons conclure que la classification binaire des images est suffisante à ce stade.

Stratégies	Mesures d'erreur (m)		
	Moyenne	Maximale	écart type
Tous les satellites	13.13	33.07	12.75
Tri de satellites LOS	10.5	59.23	12.14
Seuillage du DOP	5.73	32.92	7.66

TABLE 3 – Performance des différentes approches de calcul de position proposées.

#### 4.1.2 Pondération des satellites dans le calcul de la position selon leur état de réception

*Cette section est rédigée à partir de l'article [28] et le détail des travaux est rédigé dans le livrable 3.2.*

Dans les applications GNSS, le filtre de Kalman est souvent utilisé mais on trouve également les estimateurs par moindres carrés, et moindres carrés pondérés. L'intérêt des moindres carrés pondérés est qu'il permet d'ajuster le poids apporté par chacun des signaux en fonction de la confiance qui leur est accordée.

L'objectif des moindres carrés est de résoudre une équation du type  $A.x = b$  qui n'a pas de solution exacte [29].  $b$  est le vecteur d'observations (pseudo-distances), son nombre doit

être supérieur ou égal au nombre d'inconnues (4). Il correspond au nombre de satellites reçus. L'estimation de la position revient à :

$$\Delta\hat{X} = (H^t.H)^{-1}.H^t.\Delta\rho \quad (12)$$

Pondérer consiste à introduire un poids  $P$ , qui est à estimer en fonction des mesures ou en fonction d'un critère de confiance sur la mesure, par exemple l'élévation des satellites ou le rapport signal sur bruit. L'équation 12 devient alors :

$$\Delta\hat{X} = (H^t.P.H)^{-1}.H^t.P.\Delta\rho \quad (13)$$

Dans la littérature, les moindres carrés pondérés sont parfois utilisés pour optimiser l'usage des signaux reçus (ils utilisent tous les signaux disponibles) en attribuant une importance différente aux satellites en fonction de leur niveau de signal reçu et de leur élévation [30]. Le modèle le plus simple utilisant l'élévation exprime la variance comme l'inverse du sinus de l'élévation au carré [31], soit :

$$\sigma_i^2 = \sigma_0 / \sin(\theta_i)^2 \quad (14)$$

où  $\theta_i$  représente l'élévation du satellite  $i$  et  $\sigma_0$  est un paramètre du modèle qui dépend du type de récepteur et d'antenne et doit être déterminé à l'avance.

On trouvera également le modèle cité par [32] :

$$\sigma_i^2 = a + b \times \exp -\theta_i/\theta_0 \quad (15)$$

où  $\theta_i$  représente l'élévation du satellite  $i$  et  $a$ ,  $b$  et  $\theta_0$  sont des constantes du modèle.

D'autres auteurs, comme [33] considèrent qu'un modèle utilisant le rapport  $C/N_0$  permet de prendre en compte autant l'élévation que les conditions de propagation et proposent les modèles suivants. Le modèle *Sigma* -  $\epsilon$  [31] utilise la valeur du rapport  $C/N_0$  (valeur normalisée du rapport signal sur bruit  $SNR$ ) des signaux GPS pour estimer les poids à appliquer dans les moindres carrés.

$$\sigma_i^2 = c^2.10^{-CN_0i/10} \quad (16)$$

où  $c$  est une constante du modèle.

Wieser [34] a proposé une formule qui exprime la variance de phase  $\sigma^2$  comme une fonction des valeurs  $C/N_0$  :

$$\sigma_i^2 = v_i + c_i.10^{-0.1 \times CN_0i} \quad (17)$$

où  $v_i$  (en  $m^2$ ) et  $c_i$  (en  $m^2Hz$ ) sont les paramètres du modèle, estimés expérimentalement et  $i$  le satellite.

Afin de limiter les points d'indisponibilité rencontrés lors de l'application de notre technique d'exclusion des satellites NLOS, nous avons adapté ces solutions, en ajoutant une pré-étape de connaissance de l'état de réception des signaux.

Nous proposons d'adapter l'équation 16 en exprimant la variance comme suit :

$$\sigma_i^2 = k \times 10^{-0.1.CN_0i} / \sin(\theta_i)^2 \quad (18)$$

avec  $k$  variant en fonction de l'état de réception du signal :

- $k = 1$  lorsque le satellite est reçu en trajet direct (LOS),
- $k = 2$  lorsque le satellite est en NLOS.

La valeur de  $k$  a été déterminée de façon expérimentale.

Nous avons comparé notre proposition aux modèles de la littérature présentés ci-dessus pour quantifier son apport sur la précision à partir de la base de données expérimentale de Belfort. Les performances en termes de précision sont comparées sur la figure 22. L'erreur est estimée par projection des points calculés sur la courbe de référence obtenue grâce aux mesures RTK (courbe verte).

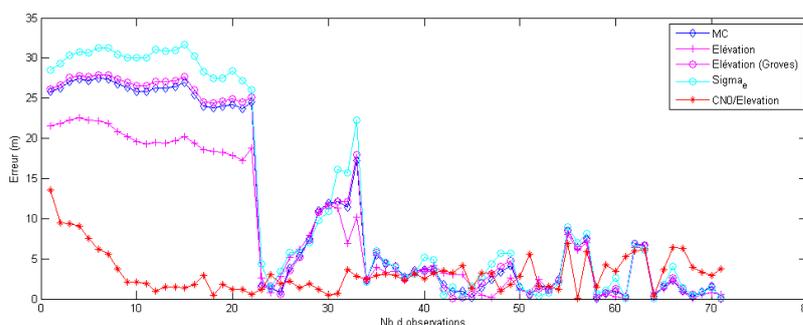


FIGURE 22 – Précisions obtenues avec les méthodes des moindres carrés ordinaires, les pondérations par des modèles d'élévation, Sigma-epsilon et par notre modèle

Les résultats présentent deux phases différentes dans la mesure. La première période de mesure ( $t = 0$  à  $25s$ ) est une zone de démarrage, dans une zone fortement bruitée en raison de l'environnement de réception et de l'état stationné du véhicule à  $t = 0$ . La deuxième partie de la trajectoire montre des erreurs inférieures à 10m essentiellement, et correspondent typiquement à des erreurs rencontrées par un récepteur GPS en environnement urbain.

Sur la première partie, les résultats diffèrent fortement selon les méthodes : l'erreur obtenue avec le modèle  $Sigma - \epsilon$  oscille en effet autour de 30m, contre 25 pour les moindres carrés classiques, 20 pour le modèle d'élévation et décroît très vite de 10 à quelques mètres avec notre méthode. Sur la deuxième partie de trajectoire les différences entre modèles sont moins visibles.

Le tableau 4.1.2 résume les performances des méthodes de la littérature avec la solution CAPLOC avec  $k = 2$ , en termes d'erreur moyenne, médiane, d'écart type et de moyenne quadratique (ou RMS - *Root Mean Square*) sur la globalité du trajet.

	M.C.	$Sigma - \epsilon$	El.	El.(Groves)	$CAPLOC_{k=2}$
Moyenne	10,57	12,22	8,59	10,78	<b>3,28</b>
Médiane	4,33	5,82	4,14	4,74	<b>2,88</b>
Ecart-type	10,8	12,38	8,24	11,09	2,5
RMS	15,07	17,34	11,86	15,41	<b>4,11</b>

TABLE 4 – Tableau des performances obtenues avec les différentes méthodes de pondération

Ces résultats montrent clairement l'intérêt de notre approche par rapport aux autres solutions qu'elles soient pondérées avec d'autres critères ou non pondérée. Ces résultats seront à confirmer à l'avenir sur d'autres trajectoires pouvant présenter des configurations différentes de visibilité et de propagation des signaux

## 4.2 Utilisation du modèle 3D

La tâche 2 du projet consiste à réaliser un modèle 3D de l'environnement à partir des images et d'en extraire les bâtiments représentant des surfaces de réflexion des signaux GNSS. Comme

nous l'avons présenté dans la section précédente, il est possible de positionner le mobile par rapport aux bâtiments situés autour de lui grâce à l'extraction de leurs empreintes dans le plan horizontal. Nous parvenons ensuite à faire une première estimation de la hauteur des bâtiments en analysant la distribution des points verticalement au dessus des empreintes détectées. Pour terminer, nous estimons la normale d'un plan que nous cherchons à ajuster sur l'ensemble des points ayant contribué à la détection de chaque segment composant l'empreinte des bâtiments. La figure 23 illustre le résultat obtenu et montre qu'il est alors possible de positionner les satellites reçus par rapport aux bâtiments présents dans la scène.

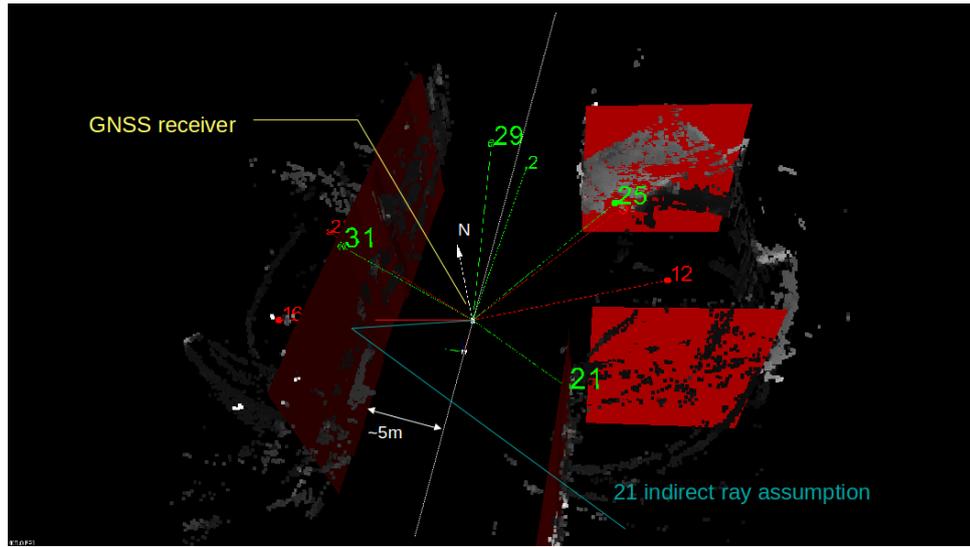


FIGURE 23 – Plane fitting sur le nuages de points à partir de l'empreinte au sol des bâtiments.

Pour cela la connaissance du nord est indispensable afin de correctement orienter le modèle. En l'absence de magnétomètre, nous avons choisi de faire une estimation initiale de manière manuelle en comparant la première image de la séquence avec différents outils de navigation à notre disposition tels que google street. Nous précéons alors au suivi du nord à partir de l'estimation des transformations réalisée au cours de chaque calibrage. Toutefois, des erreurs peuvent apparaître et se propager d'un couple d'images à l'autre. Une correction manuelle peut réduire ce risque.

Comme annoncé dans les objectifs du projet, l'utilisation de ce modèle doit permettre de corriger l'erreur des pseudo-distances des satellites NLOS reçus après une ou plusieurs réflexions.

En suivant les pistes développées dans [35], la piste envisagée est de calculer le retard  $r$  sur le signal propagé par une ou plusieurs réflexion afin de corriger l'estimation de la mesure  $\rho_{mesure}$  réalisée par le récepteur. La correction est conforme à l'équation 19.

$$\rho_{corrige} = \rho_{mesure} - r \quad (19)$$

La figure 24 est la représentation d'un retard lorsque le plan azimuthal du satellite est orthogonal aux bâtiments. La valeur du retard géométrique dépend à la fois de la configuration du canyon et de la position du satellite (azimut et élévation). On peut alors écrire  $r$  selon l'équation 20 dans laquelle  $el$  est l'élévation du satellite.

$$r = w_i \times \cos(el) \quad (20)$$

La figure 25 illustre comment déterminer ce retard pour un satellite dont l'angle d'azimut est quelconque. Les retards  $r$ , s'expriment alors de manière générale par les équations 21 confor-

mément aux notations de la figure et avec  $L$  représentant la largeur de la rue d'un bâtiment à l'autre.

Il est important de noter qu'il est possible de définir le nombre de réflexions qu'un satellite subira en comparant son élévation à trois seuils donnés par les équations 21. Dans ces équations,  $H$  représente la hauteur du bâtiment derrière lequel le satellite est positionné. Soit un satellite d'élévation  $el$ . Si  $el_{th1} < el < el_{th2}$  alors le satellite subit une seule réflexion. Si  $el_{th2} < el < el_{th3}$  alors le satellite subit deux réflexions. Si  $el < el_{th3}$  alors le satellite subit trois réflexions.

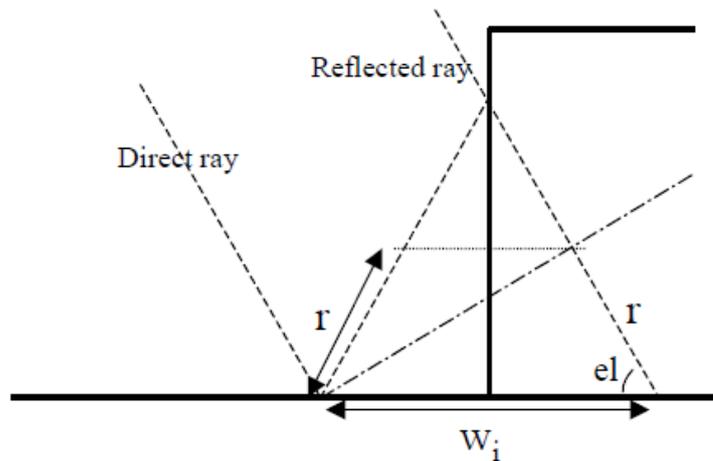


FIGURE 24 – Configuration d'un canyon urbain et retard  $r$  causé par la réflexion en 2D.

$$\begin{aligned}
 el_{th1} &= \arctan \left| \frac{H \times \cos(\alpha)}{w_i} \right| & r_{1 \text{ ref}} &= 2 \times w_i \times \cos(el) \times \cos(\alpha) \\
 el_{th2} &= \arctan \left| \frac{H \times \cos(\alpha)}{2 \times L - w_i} \right| & r_{2 \text{ ref}} &= 2 \times L \times \cos(el) \times \cos(\alpha) \\
 el_{th3} &= \arctan \left| \frac{H \times \cos(\alpha)}{2 \times L + w_i} \right| & r_{3 \text{ ref}} &= 2 \times (w_i + L) \times \cos(el) \times \cos(\alpha) \quad (21)
 \end{aligned}$$

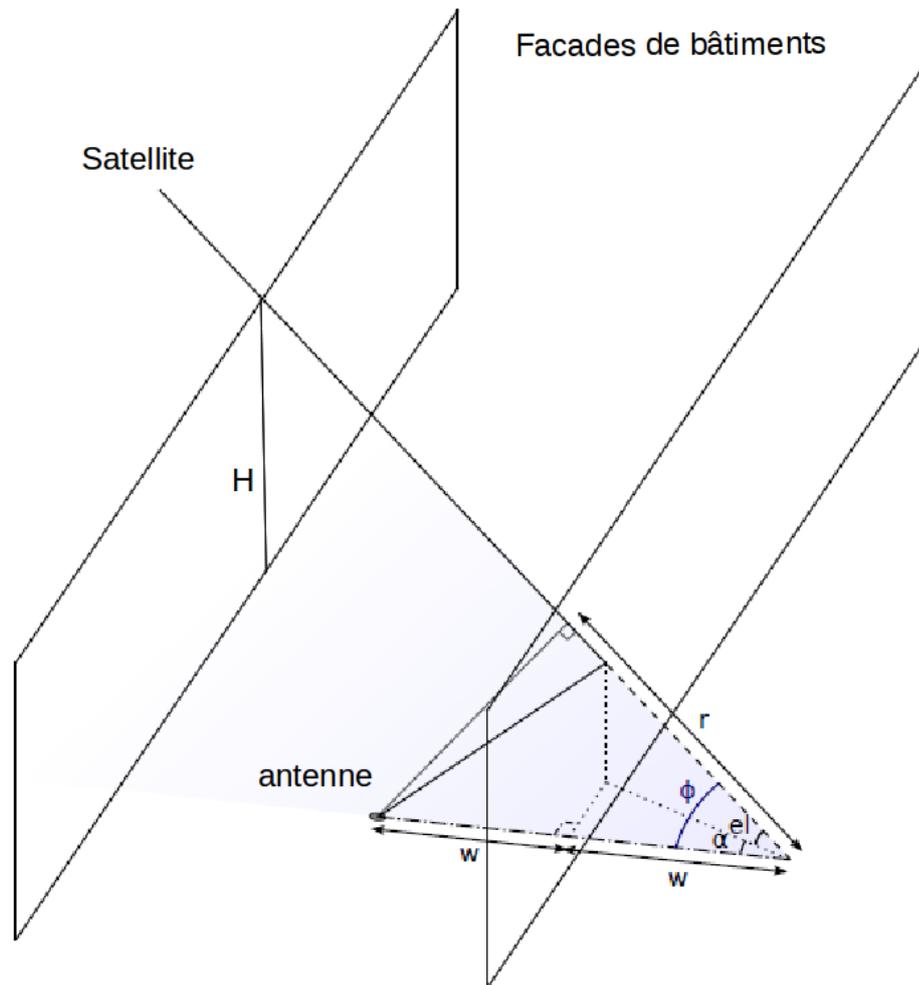


FIGURE 25 – Configuration d'un canyon urbain en 3D et retard d'un satellite NLOS dont l'azimut est quelconque. Dans ce schéma,  $w = w_i$

Ainsi, à chaque instant de réception des données GNSS fournies par le récepteur, il est possible de définir le retard associé à chaque satellite après avoir déterminé la nature de sa réflexion. Pour l'Arch 1, la cadence d'acquisition peut atteindre 15 images par seconde. Dans ce cas, le modèle considéré comme valide pour estimer les retards est le dernier modèle calculé. Pour l'Arch 2, un couple d'images est fourni tous les  $2m$ . Ainsi, lorsque la vitesse du mobile est faible, il est indispensable de prédire la position du mobile dans le dernier modèle valide jusqu'à ce qu'il soit mis à jour. Ainsi pour une vitesse du mobile inférieure à  $7.2km/h$  (seuil calculé sur la base d'une acquisition tous les  $2m$ ), le mobile peut recevoir plus d'une donnée GNSS avant de mettre à jour son modèle.

Nous avons validé cette méthode de correction sur des données acquises selon l'Arch 1 sur la ville de Lille et selon l'Arch 2 dans la ville de Belfort. La figure 26 montre le nuage de points en une position sur le trajet lillois du mobile sur lequel nous avons superposé les satellites reçus LOS et NLOS. Sur cette même figure nous représentons (1) la position fournie par le récepteur GNSS en considérant tous les satellites comme LOS (GPS position) et (2) la position après correction des satellites NLOS (PR rectification 2D). La position de référence est donnée par un récepteur RTK. Elle permet de conclure, dans ce cas de figure, que la correction apporte une amélioration.

Sur le trajet effectué à Belfort, nous avons été confronté à la présence importante d'arbres. Lors d'une première phase de correction nous n'avons pas tenu compte de la nature des obstacles

en présence et nous avons appliqué la correction en considérant ces obstacles comme des bâtiments. Nous avons ainsi montré que corriger un satellite reçu directement à travers la végétation induisait une erreur de position plus importante. La méthode de correction que nous proposons considère désormais la nature des obstacles renseignée par les outils développés dans la tâche 1.

Nous avons étudié la distribution des pseudo-distances fournies par le récepteur Safedrive. Nous avons constaté que, lorsque le mobile était à l'arrêt, l'écart-type de ces valeurs était bien supérieur à la valeur des corrections que nous pouvions appliquer aux signaux GNSS NLOS. Ce phénomène pourrait expliquer la mauvaise qualité des résultats que nous avons obtenus après correction. En mobilité, les résultats obtenus ne sont pas à la hauteur de ce que nous escomptions. La correction que nous appliquons aux pseudo-distances NLOS dégrade souvent la position calculée par le récepteur. La première conclusion serait que le phénomène corrigé par notre méthode n'est pas celui observé par le récepteur. La deuxième conclusion serait de considérer que les pseudo-distances fournies par le Safedrive étant filtrées, nous ne pouvons leur appliquer directement nos corrections.



FIGURE 26 – Amélioration de la localisation après correction du retard sur les pseudo-distances NLOS.

## 5 Conclusions et perspectives

L'objectif de CAPLOC était de délivrer une information de localisation précise, accessible en tous lieux, dans un contexte de réduction globale des coûts. CAPLOC est un projet de recherche amont. L'approche suivie, brevetée en 2008 et présentée dans cette synthèse, repose une connaissance de l'environnement de propagation fournie par une ou plusieurs caméras et des traitements d'images capables de détecter l'état de réception des signaux satellitaires et d'apporter une connaissance des perturbations causées par les obstacles environnants afin de réduire leur effet sur la localisation GNSS. Comme décrit également dans l'introduction, cette approche est aujourd'hui adaptée par plusieurs acteurs de la recherche dans des systèmes embarquant des cartes numériques 3D et avec des objectifs similaires. L'arrivée de ces travaux ne remet cependant pas en cause l'intérêt du projet. En effet, nous reviendrons sur la question de l'équipement à embarqué (volume et coût), mais l'utilisation de l'image présente le double avantage de ne pas dépendre d'un modèle à acquérir et mettre à jour et celui de ne pas devoir se positionner dans ce modèle.

Trois axes majeurs ont été développés dans le projet, tous illustrés avec des données réelles et valorisés par des publications scientifiques :

- Techniques de segmentation et de classification d'images fish-eye couleurs, comparées aux techniques de la littérature, permettant en temps réel, de détecter l'état de réception des signaux satellitaires reçus en fonction des obstacles rencontrés (bâtiment, végétation).
- Estimation de la structure 3D de l'environnement urbain à partir du traitement des images fournies par un système de perception mono et multi caméras fisheye. Positionnement du mobile relativement aux bâtiments pour estimer les retards dus aux réflexions multiples.
- L'utilisation de ces connaissances sur la propagation des signaux reçus pour une estimation plus précise de la position : par détection et exclusion du/des satellite(s) reçu(s) de façon indirecte ; par la pondération de ces satellites pour une solution plus optimale des signaux disponibles et par la correction de l'estimation de la pseudo-distance grâce au modèle 3D.

L'objectif final du projet était l'amélioration de la précision de la localisation, et, sur la trajectoire testée, la meilleure précision obtenue (avec la méthode des moindres carrés pondérés) a permis de localiser le véhicule à  $3,28m$  en moyenne contre  $10,57m$  avec une méthode dite classique  $\hat{z}$ . Les méthodes développées ouvrent encore des perspectives d'amélioration qu'il faudra poursuivre, en affinant le critère de pondération mais aussi en ouvrant la connaissance a priori de l'état de réception à d'autres algorithmes d'estimation dynamique de l'erreur par exemple. La correction des pseudo-distances n'offre pas encore de résultats significatifs en mobilité. Le concept a été validé sur certains points mais reste à confirmer sur un jeu de données dynamique. Il s'agira également d'étudier les passerelles entre ces méthodes et les méthodes utilisant des modèles 3D embarqués, qui se développent depuis 3 ans.

Les recherches effectuées sur les axes traitements d'images sont également riches de perspectives. La classification des pixels des images fisheye a intéressé les industriels de la simulation GNSS qui y ont vu un outil pour la modélisation de masquages réalistes ou pour le développement de techniques d'intégrité dans des environnements contraints. L'approche de segmentation d'images proposée et combinant de manière adaptative et non paramétriques les informations de couleur et de texture pourrait par exemple être utilisée dans d'autres applications. En effet, une segmentation fine de l'image faisant ressortir différentes zones d'intérêts est un atout pour le développement d'outils d'aide à la vidéosurveillance intelligente ou encore d'aide au diagnostic dans le domaine de la santé. De plus, les algorithmes de caractérisation fine (ciel/bâtiment/végétation) de l'environnement urbain pourraient être utilisés à d'autre domaine du transport, de la robotique, ou de la défense.

L'idée de corriger les erreurs de pseudo-distance dues aux réflexions des signaux GNSS sur les bâtiments a été validée dans la littérature sur la base de modèles numériques de terrain (MNT) embarqués à bord du mobile. Cette technique nécessite non seulement un MNT à jour et la possibilité de pouvoir se localiser assez précisément dans ce modèle à partir des signaux GNSS reçus. Les performances obtenues dépendent donc beaucoup de la précision de la carte et de la position GNSS initiale. CAPLOC propose non seulement d'estimer la structure 3D de l'environnement urbain en temps-réel et d'y positionner le mobile. Avec cela nous échappons à la nécessité de connaître avec précision la position initiale du mobile et nous faisons appel à des informations 3D actualisées et produites à partir d'un point de vue plus proche de celui de l'antenne GNSS. En fusionnant les modèles 3D de l'environnement obtenus au cours du déplacement, il est possible de calculer un modèle complet du trajet du mobile. Dans le cadre d'une flotte de véhicules munis de système de communication V2V et V2I, l'ambition pourrait être de concevoir de manière collaborative un modèle 3D toujours mise à jour à la dimension d'une ville. L'utilisation d'une seule caméra HD fisheye fait du système CAPLOC un système peu encombrant, facilement intégrable. Les algorithmes proposés permettent d'obtenir un modèle 3D assez précis sur lequel il est possible de renseigner la nature des objets numérisés comme le proposent les recherches décrites précédemment. La solution mono fisheye intégrée est utilisable dans d'autres contextes tels que celui de la robotique mobile et tout particulièrement celui du développement des robots d'assistance. Il est également envisagé de l'exploiter dans le cadre d'une flottille de drones pour des applications de surveillance d'infrastructures terrestres.

Le projet CAPLOC a montré, avec des résultats de recherche amont, que le concept utilisant l'image acquise en mobilité, pour une connaissance de la propagation des signaux GNSS permettait d'améliorer la précision de localisation dans des environnements urbains. À l'origine développé à la demande d'opérateurs de transports publics, le concept a été validé essentiellement sur véhicules routiers dans CAPLOC, et partiellement à partir d'images ferroviaires avec quelques images issues du projet européen SATLOC. Le système développé aujourd'hui comporte une caméra à objectif fisheye à placer verticalement sur le toit du véhicule, à proximité d'une antenne GNSS, un récepteur GNSS et un ordinateur pour le stockage et le traitement des acquisitions. Le système est tout à fait transposable à d'autres modes de transport. Les travaux futurs devront envisager la simplification ou la miniaturisation des équipements pour un embarquement plus facile dans tout type de véhicule. L'implémentation sur des processus parallèles pour optimiser la rapidité de traitement aidera à atteindre le temps réel, en particulier pour la construction du modèle 3D.

## 6 Valorisation

La dissémination autour du projet CAPLOC a été menée sous différentes formes. Tout d'abord un site internet a été réalisé pour communiquer sur les objectifs du projet et la démarche entreprise pour y parvenir (<http://caploc.ifsttar.fr/>). Le site affiche également les références des principales publications et communications effectuées dans le cadre du projet. La dissémination a ciblé aussi de nombreux événements nationaux et internationaux (PREDIT, CNIG, COST SaPPART, GIS ITS, etc.) pour présenter les objectifs, la démarche et les résultats intermédiaires du projet. Sur le plan scientifiques de nombreuses publications et communications nationales et internationales ont été réalisées pour mettre en avant les techniques et les méthodes développées pour les problématiques posées par le projet CAPLOC. Une communication sur le projet et ses résultats intermédiaires, présentée au 40<sup>me</sup> Congrès ATEXPO en 2013, a reçu le prix de la meilleure intervention et a été sélectionnée pour apparaître dans la revue TEC.

### Liste des publications

- 3D modelling of urban environment for enhancing GPS localization's accuracy, J. Moreau, S. Ambellouis, Y. Ruichek, TRA (Transport Research Arena) 2014, Paris, avril 2014
- Caractérisation et Structure 3D de l'environnement urbain pour une position GNSS plus précise, J. Marais, S. Ambellouis, C. Meurie, Y. Ruichek, Congrès ATEXPO, Paris, 30 janvier 2014
- Toward accurate localization in guided transport : Combining GNSS data and imaging information, J. Marais, C. Meurie, D. Attia, Y. Ruichek, A. Flancquart, Transportation Research Part C : Emerging Technologies, December 2013
- Projet CAPLOC : Combinaison de l'analyse d'images et la connaissance de la Propagation des signaux pour la Localisation, J. Marais, C. Meurie, S. Ambellouis, Y. Ruichek, TEC Transport Environnement Circulation, 218, 4p, 2013
- Weighting models for GPS Pseudorange observations for land transportation in urban canyons, S. Tay, J. Marais, 6<sup>th</sup> European Workshop on GNSS Signals and Signal Processing, Munich, December 5-6, 2013
- Equisolid fisheye stereovision calibration and point cloud computation, J. Moreau, S. Ambellouis, Y. Ruichek, ISPRS-SSG 2013, conference on Serving Society with Geoinformatics, Antalya, Turquie, 6p, novembre 2013
- Estimation des états de réception des satellites GNSS pour une position plus précise, J. Marais, S. Tay, Journées Nationales des Communications dans les Transports, Nevers, France, 2p, mai 2013
- Quantify and improve GNSS quality of service in land transportation by using image processing, J. Marais, C. Meurie, First CNES-ONERA Workshop on Earth-Space Propagation, Toulouse, France, 5p, janvier 2013
- Accurate Localisation Based on GNSS and Propagation Knowledge for Safe Applications in Guided Transport J. Marais, S. Ambellouis, A. Flancquart, S. Lefebvre, C. Meurie, Y. Ruichek, Transport Research Arena 2012, Athens, Greece, April 23-26, 2012
- 3D reconstruction of urban environments based on fisheye stereovision, J. Moreau, S. Ambellouis, Y. Ruichek, 8<sup>th</sup> International Conference on Signal Image Technology and Internet Based Systems, Naples, Italy, 2012
- Projet CAPLOC - Combinaison de l'Analyse d'images et la connaissance de la Propagation des signaux pour la LOCALISATION, J. Marais, C. Meurie, S. Ambellouis, Y. Ruichek, 40<sup>me</sup> congrès ATEXPO, Paris, 30-31 janvier 2013
- Counting of satellites with direct GNSS signals using Fisheye camera : A comparison of clustering algorithms, D. Attia, C Meurie, Y Ruichek, J. Marais, ITSC 2011, Washington

DC, USA, October 5-7, 2011

#### **Liste des communications**

Le projet et ses résultats intermédiaires ont été présentés au cours du projet dans les séminaires ou journées suivantes :

- Séminaire PREDIT Connaissance de la mobilité, 13 novembre 2014
- Commission Géo-Pos de la CNIG (Conseil National de l'Information Géographique), 16 octobre 2014
- Workshop du COST SaPPART, Londres, septembre 2014
- Séminaire interne du GIS ITS, Nantes, 4 avril 2012
- Séminaire de valorisation GO 2 - Gestion de trafic, Paris, 28 mars 2012
- Mi-parcours du Predit 4, Bordeaux, 10-12 mai 2011

**La thèse** de Julien Moreau intitulée Construction de modèles 3D à partir de données vidéo - Application à la localisation en milieu urbain financée dans le cadre du projet doit être soutenue en juin. Encadrement : Sébastien Ambellouis (IFSTTAR), Yassine Ruichek (UTBM)

## Références

- [1] Pavel KOVÁŘ, František VEJRAŽKA, Libor SEIDL et Pavel PURIČER : Assessment of the egnos signal availability for land mobile user. *12th Saint Petersburg International Conference on Integrated Navigation Systems. Saint Petersburg*, pages 148–156, 2005. 4
- [2] Cindy CAPPELLE : *Localisation de véhicules et détection d'obstacles : apport d'un modèle virtuel 3D urbain*. Thèse de doctorat, Lille 1, 2008. 6
- [3] Sergio NOGUEIRA, Yassine RUICHEK et Francois CHARPILLET : A self navigation technique using stereovision analysis. *Stereo Vision*, pages 287–298, 2008. 6
- [4] Eric ROYER, Maxime LHUILLIER, Michel DHOME, Thierry CHATEAU *et al.* : Towards an alternative gps sensor in dense urban environment from visual memory. *the 15th British Machine Vision Conference*, 2004. 6
- [5] Nicolas VIANDIER : *Modélisation et utilisation des erreurs de pseudodistances GNSS en environnement transport pour l'amélioration des performances de localisation*. Thèse de doctorat, Ecole Centrale de Lille, 2011. 6, 29
- [6] Juliette MARAIS, Si£;bastien AMBELLOUIS, Amaury FLANCQUART, Si£;bastien LEFEBVRE, Cyril MEURIE et Yassine RUICHEK : Accurate localisation based on gnss and propagation knowledge for safe applications in guided transport. *Procedia - Social and Behavioral Sciences*, 48(0):796 – 805, 2012. Transport Research Arena 2012. 6
- [7] Juliette MARAIS : *Localisation de mobiles terrestres par satellites : mise en œuvre d'outils permettant l'analyse de l'influence des conditions de propagation et des effets de masques sur la disponibilité du service offert*. Thèse de doctorat, Université de Lille 1, 2002. 6
- [8] J-I MEGURO, Taishi MURATA, J-I TAKIGUCHI, Yoshiharu AMANO et Takumi HASHIZUME : Gps accuracy improvement by satellite selection using omnidirectional infrared camera. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 1804–1810. IEEE, 2008. 6
- [9] Paul D GROVES : Shadow matching : A new gnss positioning technique for urban canyons. *Journal of Navigation*, 64(03):417–430, 2011. 6
- [10] Aude BOURDEAU, Mohamed SAHMOUDI et Jean-Yves TOURNERET : Constructive use of gnss nlos-multipath : Augmenting the navigation kalman filter with a 3d model of the environment. *Information Fusion (FUSION), 2012 15th International Conference on*, pages 2271–2276. IEEE, 2012. 6
- [11] Marcus OBST, Sven BAUER et Gerd WANIELIK : Urban multipath detection and mitigation with dynamic 3d maps for reliable land vehicle localization. *Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION*, pages 685–691. IEEE, 2012. 6, 29
- [12] Francois PEYRET, David BÉTAILLE et Florian MOUGEL : Non-line-of-sight gnss signal detection using an on-board 3d model of buildings. *ITS Telecommunications (ITST), 2011 11th International Conference on*, pages 280–286. IEEE, 2011. 6
- [13] D. ATTIA : *Segmentation d'images par combinaison adaptative couleur/texture et classification de pixels. Application a la caractérisation de l'environnement de réception des signaux GNSS*. Thèse de doctorat, Université de Technologie de Belfort-Montbéliard, Octobre 2013. 8
- [14] C. CONIGLIO : *Caratérisation d'environnement urbain par analyse de segments et de textures*. Thèse de doctorat, Université de Technologie de Belfort-Montbéliard, Septembre 2012. 8
- [15] J. SERRA : Image analysis and mathematical morphology. *Theoretical advances*, 2, 1988. 9
- [16] M. SHMITT et J. MATTIOLI : *Morphologie mathématique*. Masson, 1994. 9

- [17] Z. B. SBEH et L. D. COHEN : A new approach of geodesic reconstruction for drusen segmentation in eye funds images. *IEEE Transactions on medical imaging*, 20:1321–1333, 2001. 9
- [18] Luc VINCENT : Morphological gray scale reconstruction in image analysis : Applications and efficient algorithms. *IEEE Transactions on image processing*, 2:176–201, 1993. 10
- [19] R. FISHER : The use of multiple measurements in taxonomic problems. *Anal. of Eugenics*, 7:179–188, 1936. 11
- [20] W. FISHER : On grouping for maximum homogeneity. *JASA*, 53:789–798, 1958. 11
- [21] Branislav MIČUŠIK : *Two-View Geometry of Omnidirectional Cameras*. Thèse de doctorat, Czech Technical University in Prague, 2004. 21
- [22] Branislav MIČUŠIK et Tomáš PAJDLA : Estimation of omnidirectional camera model from epipolar geometry. *Conference on Computer Vision and Pattern Recognition*, 2003. 21
- [23] Richard HARTLEY et Andrew ZISSERMAN : *Multiple View Geometry in Computer Vision*. Cambridge University Press, second édition, 2004. 21
- [24] Sven FORSTMANN, Yutaka KANOU, Jun OHYA, Sven THURING et Alfred SCHMITT : Real-time stereo by using dynamic programming. *Computer Vision and Pattern Recognition Workshop*, 2004. 22
- [25] Jiri MATAS, C. GALAMBOS et J. KITTLER : Robust detection of lines using the progressive probabilistic hough transform. *Computer Vision and Image Understanding*, 78(1):119 – 137, 2000. 25
- [26] Michael LENTMAIER, Bernhard KRACH et Patrick ROBERTSON : Bayesian time delay estimation of gnss signals in dynamic multipath environments. *International Journal of Navigation and Observation*, 2008, 2008. 29
- [27] Dhouha ATTIA : *Segmentation d'images par combinaison adaptative couleur/texture et classification de pixels. Application à la caractérisation de l'environnement de réception des signaux GNSS*. Thèse de doctorat, Université de Belfort Montbéliard, 2013. 29
- [28] Sarab TAY et Juliette MARAIS : Weighting models for gps pseudorange observations for land transportation in urban canyons. Munich, 5-6 decembre 2013. 6<sup>th</sup> European Workshop on GNSS Signals and Signal Processing. 32
- [29] Kai BORRE : *A software-defined GPS and Galileo receiver : a single-frequency approach*. Birkhauser, 2007. 32
- [30] H HARTINGER et FK BRUNNER : Variances of gps phase observations : the sigma- $\tau$  model. *GPS solutions*, 2(4):35–43, 1999. 33
- [31] A WIESER : Weighting gnss observations and variations of gnss/ins integration. *Inside GNSS*, 2(1):26–33, 2007. 33
- [32] Paul D GROVES et Ziyi JIANG : Height aiding, c/n<sub>0</sub> weighting and consistency checking for gnss nlos and multipath mitigation in urban areas. *Journal of Navigation*, 66(05):653–669, 2013. 33
- [33] Jizhong LI et Muqing WU : The improvement of positioning accuracy with weighted least square based on snr. *Wireless Communications, Networking and Mobile Computing, 2009. WiCom'09. 5th International Conference on*, pages 1–4. IEEE, 2009. 33
- [34] Andreas WIESER et Fritz K BRUNNER : An extended weight model for gps phase observations. *EARTH PLANETS AND SPACE*, 52(10):777–782, 2000. 33
- [35] J MARAIS, M BERBINEAU et F NAHIMANA : Effect of alternate path on satellite positioning accuracy in an urban environment. *International Conference on Satellite and Space Communications, 2005. ICSCC 2005*, 2005. 35