



**HAL**  
open science

## Localisation à base d'amers visuels: Cartographie et mise en correspondance de marquages au sol et intégration dans LBA

Bahman Soheilian, Xiaozhi Qu, Mathieu Brédif

### ► To cite this version:

Bahman Soheilian, Xiaozhi Qu, Mathieu Brédif. Localisation à base d'amers visuels: Cartographie et mise en correspondance de marquages au sol et intégration dans LBA. *Reconnaissance des Formes et Intelligence Artificielle (RFIA)*, Jun 2016, Clermont-Ferrand, France. hal-02551465

**HAL Id: hal-02551465**

**<https://hal.science/hal-02551465>**

Submitted on 23 Apr 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Localisation à base d’amers visuels : Cartographie et mise en correspondance de marquages au sol et intégration dans LBA

B. Soheilian

X. Qu

M. Brédif

Université Paris-Est, IGN, SRIG, MATIS, 73 avenue de Paris, 94160 Saint Mandé, France.  
bahman.soheilian@ign.fr

## Résumé

La localisation précise en milieux urbains denses est un verrou important pour les systèmes de transport intelligent et cartographie mobile. Nous proposons une stratégie pour utiliser des marquages au sol comme des amers visuels pour la localisation basée vision. La première étape consiste à reconstruire une carte de marquages au sol en 3D. Un système de cartographie mobile est utilisé pour numériser l’environnement en 3D et générer une ortho-photo de route. L’algorithme RJ-MCMC est utilisé pour détecter un ensemble extensible de modèles de marquage. Les objets détectés sont reconstruits en 3D en utilisant les informations de hauteur des points 3D. Une caméra optique calibrée et une antenne GPS de bas coût sont embarquées sur un véhicule pour la localisation. L’ajustement de faisceaux local (LBA) est appliqué pour estimer la trajectoire du véhicule. Pour réduire la dérive de localisation, les images acquises par la caméra sont fréquemment appariées sur les marquages au sol géoréférencés. Cet appariement est initialisé en utilisant la pose estimée (LBA) et optimisée par un algorithme MCMC. Cela produit des points d’appui qui seront intégrés dans le LBA. La méthode est évaluée sur des données réelles acquises en milieux urbains denses.

## Mots Clef

Localisation, amers visuels, marquages au sol, RJ-MCMC, MCMC, ajustement de faisceaux local.

## Abstract

Precise localization in dense urban areas is a challenging task for both mobile mapping and driver assistance systems. The paper proposes a strategy to use road markings as localization landmarks for vision based systems. The first step consists in reconstructing a map of road marks. A mobile mapping equipped with precise georeferencing devices is applied to scan the scene in 3D and generating an ortho-image of the road surface. A RJMCMC method is applied to detect road mark patterns. The detected objects are reconstructed in 3D using the height information obtained from 3D points. A calibrated camera and a low cost GPS receiver are embedded on a vehicle and used as localization devices. Local bundle adjustment is applied to estimate

the trajectory of the vehicle. In order to reduce the drift of the trajectory, images are matched with the reconstructed road marks frequently. The matching is initialized by the initial poses estimated by LBA and optimized by a MCMC algorithm. The matching provides ground control points that are integrated in the LBA in order to refine the pose parameters. The method is evaluated on a set of images acquired in a real urban area and compared to a precise ground-truth.

## Keywords

Localization, landmark, road-marking, RJ-MCMC, MCMC, local bundle adjustment.

## 1 Introduction

La localisation précise en milieux urbains denses est un verrou important pour plusieurs applications comme la cartographie mobile, l’aide à la conduite et la navigation autonome. La localisation basée GNSS souffre des problèmes de multi-trajets et des masques de signaux et aboutit à des précisions non suffisantes. Des centrales à inerties peuvent permettre de combler les manques de localisation GNSS [1]. Cependant, pour des centrales à inertie de coût abordable, la dérive de localisation due à l’accumulation des erreurs pour les longues interruptions de signaux GNSS peut aboutir à des erreurs importantes.

Des solutions de localisation plus abordables sont proposées par les communautés de vision par ordinateur et de robotique. Nistér *et al.* utilise l’estimation de pose relative entre les images successives dans une séquence d’images mono ou stéréo [2]. Une solution plus optimale consiste à affiner les paramètres de pose en utilisant un ajustement de faisceaux local [3]. Cependant, comme tout système de localisation relative, l’accumulation des erreurs aboutit à des dérives importantes pour les longs trajets. La technique de fermeture de boucle proposée par des approches de type SLAM (simultaneous localization and mapping) permet de limiter ces dérives [4]. Cependant, d’une part, beaucoup de trajectoires ne contiennent pas de boucle et d’autre part, l’ajustement de grandes boucles est souvent très coûteux en temps de calcul [5].

D’autres auteurs ont proposé de fusionner la localisation basée vision et GNSS pour limiter la dérive [6]. L’effica-

citée de cette stratégie dépend de la qualité de la localisation GNSS et de la durée d'interruption des signaux.

Des systèmes à deux phases de cartographie hors ligne puis localisation en ligne ont également été investigués. Pendant la phase de cartographie, des primitives bas niveaux de type points [7] et patches texturés [8] sont reconstruits en 3D. La phase de localisation utilise ces primitives comme amers visuels pour recalculer la trajectoire. Les inconvénients de ce type de carte sont dans la difficulté de mise à jour partielle et la grande quantité d'espace de stockage nécessaire.

Des données cartographiques plus classiques de type modèle 3D de bâtiment [9], modèle numérique de terrain (MNT) [10, 11] ou ortho-photo [12] ont été également utilisées comme amers visuels par des systèmes basés vision. L'amélioration de la localisation reste limitée par la précision de ce type de cartes qui sont souvent de qualité métrique.

Des objets sémantiques comme des poteaux [13], panneaux de signalisation [14] et marquages au sol [15] ont également été utilisés comme amers visuels pour la localisation. L'avantage de l'utilisation des objets sémantiques comme amers visuels est triple :

- Compacité : facilite le transfert et le stockage de carte sur un robot.
- Précision : leurs formes simples permettent des appariements géométriquement précis.
- Facilité d'appariement : l'ambiguïté d'appariement est réduite grâce à l'information sémantique très discriminante.

Nous présentons dans cet article une approche de localisation utilisant les marquages au sol comme amers visuels.

## 2 Travaux liés et notre stratégie

La détection de marquage est utilisée dans un grand nombre de systèmes de transport intelligent pour maintenir le véhicule sur une même voie [16]. Peu d'auteurs ont utilisé les marquages comme amers visuels pour la localisation absolue.

Dans un système proposé par Pink [17], les marquages sont d'abord détectés dans des images aériennes par une méthode semi-automatique. En supposant que la surface de route est sur un plan connu, les objets détectés sont projetés dans l'espace objet et produisent une carte de marquages au sol. Dans la phase de localisation en ligne, les marquages sont détectés par filtrage dans les images acquises par une caméra embarquée. La méthode de V-disparité est utilisée pour estimer le plan de la route et les marquages sont projetés sur celui-ci. L'appariement de marquages est effectué par une méthode d'ICP initialisée par la localisation GPS. L'estimation d'une transformation 2D (composée d'une translation et d'une orientation) permet d'affiner la localisation de véhicule.

Dans une autre méthode [15], les marquages sont saisis sur les images aériennes. Comme la méthode précédente, le passage dans l'espace objet est effectué en projetant les objets sur un plan approximant la surface de route. La localisation

en ligne est initialisée par le GPS. Cette localisation est affinée en maximisant la similarité entre l'image vue de la caméra embarquée et la vue virtuelle de la carte de marquages. L'intérêt de cette méthode est dans l'appariement direct de l'image sur la carte sans l'étape intermédiaire de détection de marquages lors de l'étape de localisation en ligne.

Dans un système appelé LaneLoc [18], les marquages sont reconstruits à partir des images optiques et des nuages de points 3D acquis au niveau de la route par un véhicule terrestre. Dans la phase de localisation, les marquages sont détectés par un filtrage d'image. La localisation est effectuée en estimant une transformation 2D appliquée à la pose (translation et lacet) minimisant la distance entre les marquages de la carte projetés et ceux détectés dans l'image. La même approche est utilisée dans un autre système pour appairer les marquages [19]. Ces appariements sont ensuite intégrés dans un ajustement de faisceaux local. Cette approche permet d'affiner non seulement la pose des images recalculées mais aussi ces voisins dans la séquence d'image.

Deux types d'approximations sont appliqués dans toutes les méthodes mentionnées ci-dessus.

- Surface plane de route : Dans les deux étapes de cartographie et localisation, la surface de route est approximée par un plan de position et d'orientation connues.
- Correction de pose en 2D : Dans la phase d'estimation de pose seulement trois degrés de liberté sont pris en compte au lieu de six.

Ces approximations introduisent des erreurs dans les deux phases de cartographie et d'estimation de pose. Dans la méthode que nous proposons, aucune de ces approximations ne sont appliquées. Nous utilisons un véhicule de cartographie mobile équipé d'un système de géoréférencement précis (GPS/INS/odomètre) et d'un scanner laser pour acquérir des nuages de points 3D précis et denses [20]. Ces données sont utilisées pour générer une ortho-image de route. Nous utilisons notre travail précédent sur la détection de marquages pour générer une carte de marquages en 3D (section 3). Nous présentons une approche basée MCMC pour appairer ces marquages 3D avec les images vues par une caméra embarquée sur un véhicule à localiser (section 5). La mise en correspondance d'un objet de marquage avec son apparition sur une image, génère un point d'appui que nous intégrerons dans un ajustement de faisceaux local (section 4). La méthode proposée utilise l'incertitude des poses initiales pour guider l'appariement de marquage. Il intègre l'incertitude des objets 3D dans le calcul de pose. Nous présenterons des résultats d'expérimentation sur des données réelles (section 6).

## 3 Génération de carte de marquages

Nous utilisons la méthode proposée par Hervieu *et al.* pour la détection de marquages au sol sur une ortho-image [21]. Nous synthétisons d'abord cette méthode dans la section

3.1 et ensuite expliquons dans la section 3.2 comment les marquages détectés sont extrudés en 3D.

### 3.1 Détection de marquages par géométrie stochastique

D'abords, les nuages de points 3D géoréférencés acquis par un système de cartographie mobile sont projetés verticalement sur un plan horizontal pour générer une image de deux carnaux (intensité et altitude). Des filtres morphologiques sont utilisés pour remplissage des trous (cf. Fig 1). La forme géométrique et la taille des marquages

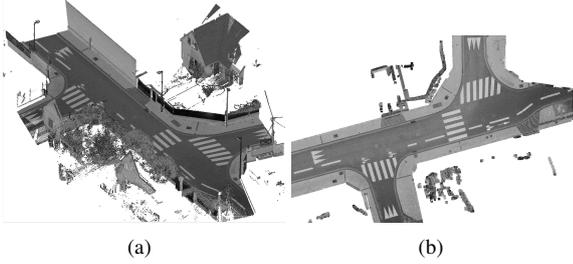


FIGURE 1 – (a) Nuage de point 3D , (b) ortho-image  $I$  ( $GSD = 2cm$ ).

au sol sont gouvernées par des spécifications strictes. Cette propriété permet de formuler le problème de détection de marquages comme recherche de l'occurrence transformée (translation, rotation, changement d'échelle) des modèles instanciés d'un bibliothèque de marquages (cf . Fig 3). L'extraction de marquages se résume à trouver un ensemble d'objet  $X = (\ell_i, x_i, y_i, \theta_i, \lambda_i)$  paramétrés par un type  $\ell$ , une translation  $(x, y)$ , une rotation  $\theta$  et une échelle  $\lambda$ . Le type de marquages définit un modèle  $I_\ell$  qui peut être échantillonné (cf. Fig 2) dans la géométrie de l'ortho-photo par une similitude  $T_{x,y,\theta,\lambda}$  (noté  $T_{X_i}$  par souci de simplicité). Hervieu *et al.* [21] formalisent le problème d'ex-

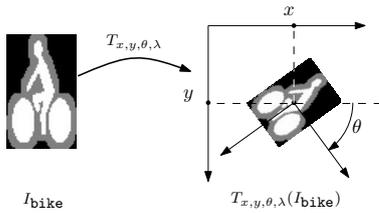


FIGURE 2 – L'objet  $i$  avec les paramètres ( $\ell_i = \text{bike}, x_i, y_i, \theta_i, \lambda_i$ ).



FIGURE 3 – La bibliothèque de marquages.

traction de marquages comme un problème de minimisation d'énergie sur un espace de recherche de dimension variable comme défini ci-dessus. Pour un ensemble de marquages  $X = (X_i)_{i=1\dots n}$ , l'énergie est définie comme suit :

$$U(X) = \sum_{i=1}^n u_1(X_i) + \sum_{i < j} u_2(X_i, X_j)$$

$$u_1(X_i) = f^0 - \max(0, ZMNC(I_{\ell_i}, T_{X_i}^{-1}(I)))$$

$$u_2(X_i, X_j) = \beta \frac{|S(X_i) \cap S(X_j)|}{\min(|S(X_i)|, |S(X_j)|)}$$

$ZMNC(I, I')$  désigne le score de corrélation centré normalisé entre les images  $I$  et  $I'$  et  $S(X) = T_X(I_{\ell_X})$  est l'image re-échantillonnée du modèle. Les opérateurs  $|\cdot|$  et  $\cap$  désigne l'aire et l'intersection des pixels. Le coefficient  $\beta$  permet de contrôler le compromis entre les deux termes d'énergies. Cette énergie est minimisée en utilisant un échantillonneur MCMC à sauts réversibles (RJ-MCMC) couplé avec un algorithme de recuit simulé permettant de gérer les espaces de recherche de dimensions variables (le nombre d'objet à détecter étant inconnu). Les noyaux classiques de RJ-MCMC (naissance-mort, perturbation) ainsi que des noyaux plus avancés permettant de biaiser l'échantillonnage aléatoire vers la bonne solution et ainsi d'accélérer la convergence sont décrits dans [21].

### 3.2 Génération de marquages au sol en 3D

Une fois que la détection de marquages sur une ortho-photo effectuée, les objets détectés sont plongés en 3D en utilisant le modèle numérique de terrain (MNT) calculé et enregistré dans le canal d'altitude de l'ortho-photo lidar. Chaque sommet de boîte englobante de marquage est plongé en 3D en associant l'altitude de celui-ci dans le MNT. Notre application cible dans cet article est la génération d'amers visuels pour la localisation. Nous avons réglé les paramètres de l'énergie (compromis entre les attaches aux données et les interactions) pour limiter le nombre des fausses détections au prix d'une sous-détection de certains marquages. Cela aboutit à une détection avec certaines sous-détections mais un nombre de sur-détections très limité. Par souci de précision de cette base de données, les objets extraits ont été validés manuellement pour supprimer le peu de faux objets restants. Pour limiter et accélérer cette intervention manuelle, les objets sont triés en utilisant leurs scores d'attache aux données. L'opérateur vérifie simplement les objets ayant des évidences les plus faibles. Cela fournit une base de données d'amers visuels géoréférencés.

## 4 Système de localisation

Nous utilisons une caméra calibrée et une antenne GPS à bas coût comme capteur de localisation. Une observation GPS et une boussole sont utilisées au début du trajet pour initialiser la localisation et la direction de véhicule. L'ajustement de faisceaux local (LBA) et la propagation d'incertitudes sont utilisés pour estimer les paramètres de pose et

les incertitudes associées. La figure 4 montre le principe de LBA sur une séquence d'images. L'ajustement est effectué sur une fenêtre glissante de  $N$  images, avançant d'un pas de  $n$  ( $n < N$ ). Le processus commence par un ajustement de faisceaux classiques sur les  $N$  premières images (étape 1 dans la Fig. 4). Ensuite, la fenêtre de calcul avance de  $n$  images (étape 2 sur la figure 4). A partir de la deuxième étape, les poses de  $N - n$  images de l'étape courante sont déjà calculées dans l'étape précédente ( $C_p$ ) et seulement  $n$  nouvelles poses sont inconnues ( $C_n$ ). Supposons que la fonction  $F$  renvoie les re-projection des points 3D  $X_t$  dans les images  $C_n$  et  $C_p$  et que  $m_t$  soit un vecteur contenant les positions 2D observées de ces points dans les images, les erreurs de re-projection de ces points suivent l'équation suivante :

$$v_t = F(C_p, C_n, X_t) - m_t \quad (1)$$

Les valeurs estimées pour les poses  $C_p$  à l'étape précédente ( $C_p^0$ ) sont utilisées pour contraindre la solution à l'étape courante en ajoutant l'équation qui suit :

$$v_p = C_p - C_p^0 \quad (2)$$

Le système est résolu en minimisant l'erreur quadratique pondérée :

$$[\hat{C}_p, \hat{C}_n, \hat{X}_t] = \arg \min_{[C_p, C_n, X_t]} \frac{1}{2} (v_t^T Q_t^{-1} v_t + v_p^T Q_{C_p}^{-1} v_p) \quad (3)$$

$Q_t$  et  $Q_{C_p}$  représentent respectivement la matrice de variance-covariance des observations des points de liaisons dans les images et celle des poses estimées à l'étape précédente. Pour intégrer les points d'appui dans le système, nous ajoutons deux équations à ce système. La première d'entre elle contrôle les erreurs de re-projection des points 3D connus  $X_G$  sur leurs positions observées  $m_g$ . La deuxième équation contraint la solution de ne pas s'éloigner des valeurs connues des points 3D ( $X_G^0$ ).

$$v_g = F_g(C_p, C_n, X_G) - m_g \quad (4)$$

$$v_G = X_G - X_G^0 \quad (5)$$

Le système est ensuite résolu en minimisant l'erreur quadratique pondérée :

$$[\hat{C}_p, \hat{C}_n, \hat{X}_t, \hat{X}_G] = \arg \min_{[C_p, C_n, X_t, X_G]} \frac{1}{2} (v_t^T Q_t^{-1} v_t + v_p^T Q_{C_p}^{-1} v_p + v_g^T Q_g^{-1} v_g + v_G^T Q_G^{-1} v_G) \quad (6)$$

$Q_G$  et  $Q_g$  correspondent respectivement à la matrice de variance-covariance des points d'appui en 3D et celle des positions observées dans les images.

Après la résolution de système, la matrice de variance-covariance des paramètres est obtenue en utilisant l'inversion partielle de la matrice normale. Nous avons utilisé le détecteur de point SIFT pour générer automatiquement des points de liaisons. La taille de la fenêtre glissante est fixée à 7 images ( $N = 7$ ). Un pas d'avancement d'une image est choisi ( $n = 1$ ).

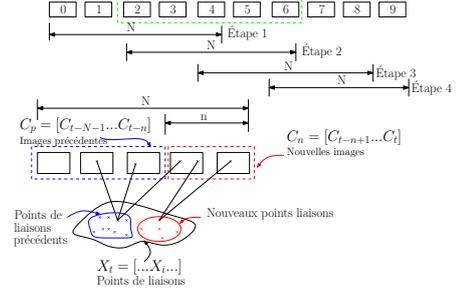


FIGURE 4 – Schéma d'ajustement de faisceaux local.

## 5 Appariement de marquage

L'ajustement de faisceaux local (LBA) sur une séquence d'images permet d'estimer les poses et leurs incertitudes. A cause de l'accumulation des erreurs, cette localisation dérive avec le temps. Donc, les positions projetées des amers visuels 3D (marquages au sol géoréférencés) dans les poses estimées ne superposent pas les apparitions des marquages. La figure 5(a) montre un exemple de projection d'une ligne blanche discontinue dans une image en utilisant la pose estimée (en rouge). Les ellipses dessinées en couleur jaune montrent les incertitudes (99%) des quatre coins du marquage. Elles sont calculées suivant l'équation suivante :

$$Q_{\bar{x}_G} = \begin{bmatrix} \frac{\partial F}{\partial C} & \frac{\partial F}{\partial X_G} \end{bmatrix} \begin{bmatrix} Q_C & 0 \\ 0 & Q_G \end{bmatrix} \begin{bmatrix} \frac{\partial F}{\partial C} \\ \frac{\partial F}{\partial X_G} \end{bmatrix} \quad (7)$$

L'objectif de cette étape est d'aligner l'objet projeté avec

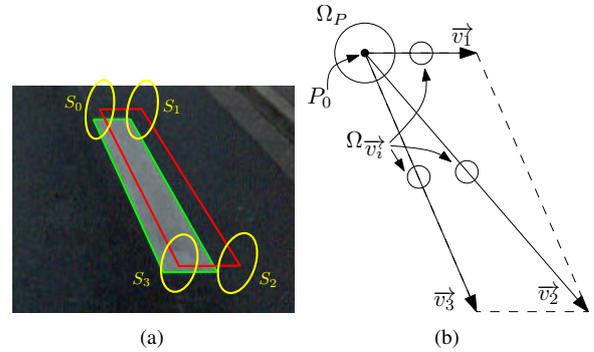


FIGURE 5 – (a) La définition de l'espace de recherche en utilisant le LBA. (b) Paramétrisation de marquage.

l'image. Une fois que les positions optimales des quatre coins du marquage dans l'image sont connues, elles sont intégrées comme des points d'appui dans le LBA et les poses sont recalculées. Cela permet de recalibrer la trajectoire aussi souvent qu'il y a des marquages appariés. L'appariement pourrait être effectué en utilisant l'ensemble des objets qui sont dans le champ de vision de la vue courante. Cela générerait plusieurs points d'appui dans chaque image. Dans notre mise en œuvre nous avons utilisé uniquement l'objet le plus proche de la pose courante. En effet, les objets les plus proches de la caméra ont plus de

chance de rester visible (ne pas être caché par des piétons et des voitures) et apportent plus de précision. Une sélection plus optimale peut être obtenue par des approches dites «carte active» permettant de sélectionner les objets d'appariement en fonction de leur apport en précision, ambiguïté et de leur probabilité de détection [22].

## 5.1 Fonction-objectif

Les quatre coins projetés des marquages dans la vue courante, définissent une homographie entre le plan 3D d'objet et le plan d'image. La qualité de cette projection peut être évaluée en calculant le score de corrélation centré et normalisé. Nous formulons notre problème comme la recherche des quatre coins dans leurs régions d'incertitude de façon à ce que le score de corrélation soit maximum. Pour éviter les cas dégénérés quand les zones d'incertitudes s'agrandissent, nous imposons aux quatre coins une forme polygonale convexe.

## 5.2 Optimisation MCMC

Contrairement à l'étape d'extraction de marquage, cette optimisation est définie dans une dimension fixe à savoir les huit coordonnées des quatre coins de marquages. L'optimisation MCMC s'adapte bien à ce problème.

**Paramétrisation de marquage et initialisation .** Etant donnée la corrélation forte entre les erreurs des quatre coins, nous proposons une modélisation d'objet avec un point d'insertion ( $P_0$ ) et trois vecteurs ( $\vec{v}_1, \vec{v}_2, \vec{v}_3$ ). Ces derniers définissent les trois autres coins de celui-ci à partir de point d'insertion (cf. la figure 5(b)). Pour initialiser ces paramètres, nous utilisons la projection de marquage dans l'image (cf. Fig 5(a)). Le point  $P_0$  est initialisé à la position du point  $S_0$  et les trois vecteurs :  $\vec{v}_i = S_i - S_0$ .

**Définition de l'espace de recherche et des noyaux de perturbations.** L'espace de recherche pour le point  $P_0$  est un cercle centré sur le point  $S_0$  d'un rayon égal à un demi grand axe de l'ellipse d'incertitude de 99% ( $\Omega_P$ ). Les espaces de recherche des corrections des vecteurs ( $\vec{v}_i$ ) sont des cercles ( $\Omega_{\vec{v}_i}$ ) plus petits (85% de demi grand axe de l'ellipse d'incertitude de points). Les deux noyaux de transformations suivants sont utilisés :

- Transformation rigide de l'objet en appliquant une translation au point d'insertion  $P_0$ .
- Modification d'un des vecteurs  $\vec{v}_1, \vec{v}_2, \vec{v}_3$  permettant d'ajuster les coins de l'objet.

Comme dans la phase de la génération de la carte de marquages, l'optimisation est effectuée en couplant un échantillonneur MCMC avec un recuit simulé.

# 6 Expérimentation

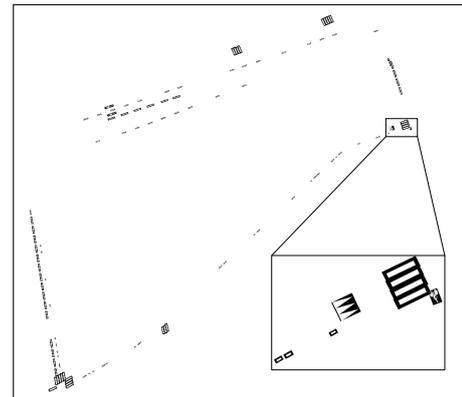
## 6.1 Mise en place de donnée

Nous avons choisi une trajectoire de test de longueur de 500m (cf. la figure 6(a)) pour évaluer l'approche proposée. Le système Stereopolis [20] a été utilisé à la fois pour acquérir les nuages de points 3D nécessaires pour la

phase de cartographie de marquage et la séquence d'image requise pour la phase de localisation. Grâce au système de géoréférencement précis de Stereopolis (intégrant deux antennes GPS, une centrale à inertie de précision et un odomètre) nous disposons d'une vérité terrain précise des poses des images. Les images sont de résolution  $1920 \times 1080$  pixels et le champ de la caméra est de  $70^\circ$  (cf. la figure 7(f)). L'algorithme de reconstruction 3D de marquages présenté dans la section 3 est appliqué sur l'ensemble du site de test. La carte 3D obtenue a été ensuite révisée manuellement pour supprimer les fausses détections. La figure 6(b) montre le résultat sur l'ensemble du site de test.



(a)



(b)

FIGURE 6 – (a) La trajectoire de test (500m). (b) Environ 200 objets de marques sont reconstruits par l'étape de la cartographie et 109 ont été conservés après la modification manuelle.

## 6.2 Résultats

Pour montrer l'apport de l'intégration des amers visuels, nous avons calculé la trajectoire avec et sans l'intégration des marquages. La figure 7 montre les résultats obtenus ainsi que les incertitudes des poses. La trajectoire verte correspond à la vérité terrain. La séquence d'image contient 205 images. Au total 255 appariements ont été effectués avec succès (1,25 appariement par image). Nous constatons que quand les marquages ne sont pas intégrés dans le système, la trajectoire commence à dériver dès le début de

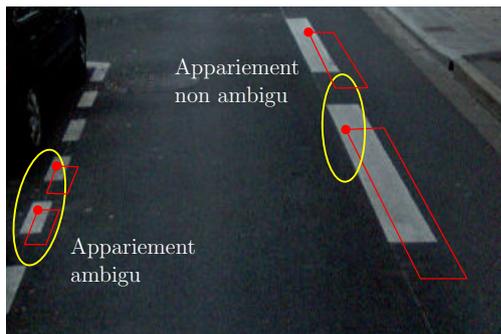


FIGURE 8 – Robustesse de l'appariement de marquage

la trajectoire alors que l'intégration des marquages évite ces erreurs sur une grande partie de la trajectoire. Cependant, vers la fin du trajet, l'estimation de pose dérive et les tailles des ellipsoïdes d'incertitude augmentent (cf. la figure 7(b)). Cela est dû à la répartition non uniforme des marquages sur la trajectoire. Le taux d'appariement de marquages sur cette deuxième partie est de 0,4 par image contre 1,6 sur la première partie. La figure 7(c) montre l'erreur absolue de localisation pour les deux scénarios. L'erreur est gardée autour de 10cm sur une grande partie de la trajectoire avec une pointe à 40cm. Cette erreur est augmentée jusqu'à 4m si les amers ne sont pas intégrés.

### 6.3 Temps de calcul

La majorité du temps de calcul est consacré à l'appariement de marquages. Il dépend de l'incertitude des quatre coins. Dû à certaines erreurs non gaussiennes, notre estimation d'incertitude est parfois sous-estimée. C'est pourquoi, en pratique, nous agrandissons la zone de recherche et augmentons le nombre d'itérations de MCMC pour garantir la convergence. Le temps de calcul moyen est entre 10s et 40s par appariement.

### 6.4 Discussion sur la robustesse

Dans notre expérimentation, la densité élevée de marquages permet de compenser la dérive de localisation en intégrant des amers dans l'ajustement de faisceaux et ainsi restreindre la taille des ellipsoïdes d'incertitude. Cependant, en absence d'appariement de marquage pendant une longue période de temps ces régions d'incertitude agrandissent et peuvent générer des problèmes d'ambiguïté d'appariement (quand la région d'incertitude contient plusieurs marquages). Comme montré sur la figure 8, ces ambiguïtés peuvent être prédites et ainsi éviter les appariements ambigus.

## 7 Conclusions et Perspectives

Nous avons présenté une approche de localisation basée vision utilisant une carte 3D précise des marquages au sol comme amers visuels. L'expérimentation a montré qu'avec peu d'objets sémantiques précis, le système est capable d'obtenir une localisation de précision de l'ordre de

10 cm en absolu. Le système est actuellement loin d'être en temps réel. Il est correctement adapté pour l'utilisation hors ligne. Une application intéressante se trouve dans les systèmes de cartographies mobiles à très bas coût. Cependant, l'étape d'appariement peut être accélérée en modifiant le terme d'énergie. Une piste intéressante serait l'utilisation de critères de similarité basés sur les contours ou les coins des marquages au lieu du score de corrélation. Une meilleure adaptation des paramètres de l'optimisation (nombre d'itérations, température initiale et descente de température) pour chaque appariement permettrait d'éviter des itérations inutiles. Une autre alternative pour accélérer l'étape d'appariement est l'utilisation de méthode d'appariement de modèle par approximation d'hyperplan [23].

## Références

- [1] D. Lapucha, K. Schwarz, M. Cannon, and H. Martell, "The use of ins/gps in a highway survey system," in *Position Location and Navigation Symposium, 1990. Record. The 1990's - A Decade of Excellence in the Navigation Sciences. IEEE PLANS '90.*, Mar 1990, pp. 413–420.
- [2] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the CVPR*, vol. 1. IEEE, 2004, pp. 652–659.
- [3] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Real time localization and 3d reconstruction," in *Proceedings of the CVPR*, vol. 1. IEEE, 2006, pp. 363–370.
- [4] A. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *Proceedings of the ICCV*, Oct 2003, pp. 1403–1410 vol.2.
- [5] J. Aulinas, Y. R. Petillot, J. Salvi, and X. Lladó, "The slam problem : a survey." in *CCIA*, 2008, pp. 363–371.
- [6] M. Lhuillier, "Incremental fusion of structure-from-motion and gps using constrained bundle adjustments," *TPAMI*, vol. 34, no. 12, pp. 2489–2495, 2012.
- [7] E. Royer, M. Lhuillier, M. Dhome, and J.-M. Lavest, "Monocular vision for mobile robot localization and autonomous navigation," *IJCV*, vol. 74, no. 3, pp. 237–260, 2007.
- [8] B. Charmette, E. Royer, and F. Chausse, "Efficient planar features matching for robot localization using gpu," in *CVPRW*, June 2010, pp. 16–23.
- [9] C. Arth, C. Pirchheim, J. Ventura, D. Schmalstieg, and V. Lepetit, "Instant outdoor localization and slam initialization from 2.5d maps," in *Proceedings of the ISMAR*, 2015.
- [10] D. Larnaout, S. Bourgeois, V. Gay-Bellile, and M. Dhome, "Towards bundle adjustment with gis constraints for online geo-localization of a vehicle in urban center," in *Proceedings of the 3DIMPVT. IEEE*, 2012, pp. 348–355.

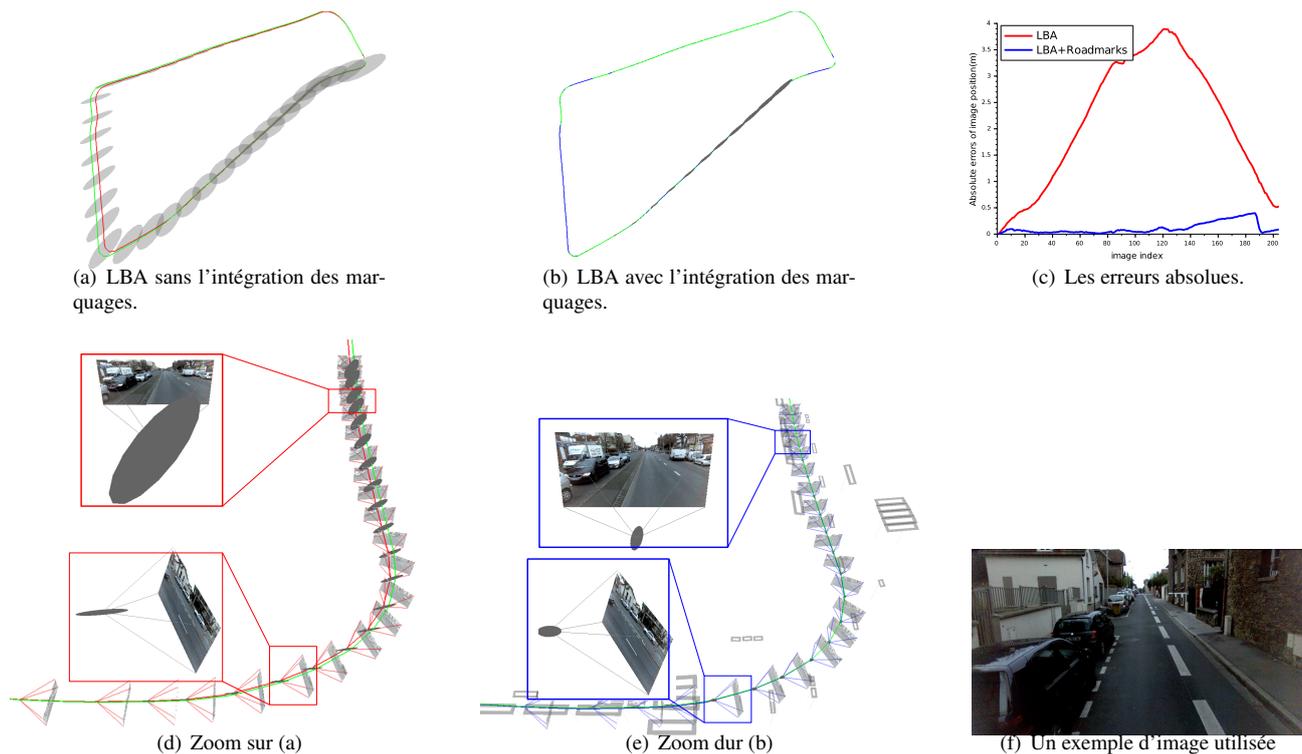


FIGURE 7 – Comparaison de la localisation avec et sans marquages . Les ellipsoïdes d'erreur sont exagérées 10 fois. La trajectoire de référence est affichée en vert.

- [11] P. Lothe, S. Bourgeois, F. Dekeyser, E. Royer, and M. Dhome, "Towards geographical referencing of monocular slam reconstruction using 3d city models : Application to real-time accurate vision-based localization," in *Proceedings of the CVPR*. IEEE, 2009, pp. 2882–2889.
- [12] M. Jaud, R. Rouveure, P. Faure, and M.-O. Monod, "Methods for fmcw radar map georeferencing," *IJ-PRS*, vol. 84, pp. 33–42, 2013.
- [13] A. Schlichting and C. Brenner, "Localization using automotive laser scanners and local pattern matching," in *Proceedings of the IVS*, June 2014, pp. 414–419.
- [14] X. Qu, B. Soheilian, and N. Paparoditis, "Vehicle localization using mono-camera and geo-referenced traffic signs," in *Proceedings of the IVS*, June 2015, pp. 605–610.
- [15] N. Mattern, R. Schubert, and G. Wanielik, "High-accurate vehicle localization using digital maps and coherency images," in *Proceedings of the IVS*. IEEE, 2010, pp. 462–469.
- [16] M. Bertozzi and a. Broggi, "GOLD : a parallel real-time stereo vision system for generic obstacle and lane detection." *IEEE TIP*, vol. 7, no. 1, pp. 62–81, jan 1998.
- [17] O. Pink, "Visual map matching and localization using a global feature map," in *CVPRW*. IEEE, June 2008, pp. 1–7.
- [18] M. Schreiber, C. Knoppel, and U. Franke, "Laneloc : Lane marking based localization using highly accurate maps," in *Proceedings of the IVS*. IEEE, 2013, pp. 449–454.
- [19] A. Ranganathan, D. Ilstrup, and T. Wu, "Light-weight localization for vehicles using road markings," in *IROS*. IEEE, Nov 2013, pp. 921–927.
- [20] N. Paparoditis, J.-P. Papelard, B. Cannelle, A. Devaux, B. Soheilian, N. David, and E. HOUZAY, "Sterepolis ii : A multi-purpose and multi-sensor 3d mobile mapping system for street visualisation and 3d metrology," *RFPT*, no. 200, pp. 69–79, 2012.
- [21] A. Hervieu, B. Soheilian, and M. Bredif, "road marking extraction using model&data-driven RJ-MCMC," *ISPRS Annals*, vol. II-3/W4, pp. 47–54, 2015.
- [22] C. Aynaud, C. Bernay-Angeletti, R. Chapuis, R. Aufrère, C. Debain, and N. Karam, "Real-time vehicle localization by using a top-down process," in *Proceedings of the ICIF*, 2014.
- [23] F. Jurie and M. Dhome, "Hyperplane approximation for template matching," *TPAMI*, vol. 24, no. 7, pp. 996–1000, Jul 2002.