



Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D

Clément Deymier, Céline Teulière, Thierry Chateau

► **To cite this version:**

Clément Deymier, Céline Teulière, Thierry Chateau. Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D. *Traitement du Signal*, Lavoisier, 2015, 32 (2-3), pp.121-145. <10.3166/ts.32.121-145 >. <hal-01316416>

HAL Id: hal-01316416

<https://hal.archives-ouvertes.fr/hal-01316416>

Submitted on 18 May 2016

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.

Étalonnage automatique d'un système d'acquisition Caméras - Centrale inertielle - Lidar 3D

Clément Deymier, Céline Teulière, Thierry Chateau

Clermont Université, Université Blaise Pascal, Institut Pascal
BP 10448, F-63000 CLERMONT-FERRAND, France

prenom.nom@univ-bpclermont.fr

RÉSUMÉ. Cet article présente une méthode entièrement automatique d'étalonnage de systèmes d'acquisition complexes comprenant une ou plusieurs caméras, une centrale inertielle et un lidar 3D. Le principe consiste à estimer les paramètres intrinsèques et extrinsèques en mettant en correspondance des primitives détectées dans les images des caméras avec le nuage de points 3D fourni par le télémètre. Ce travail propose une formalisation mathématique unifiant les trois types de capteurs au sein d'une même fonction de vraisemblance, une stratégie pour l'évaluation rapide de contraintes entre les données images et les données télémétriques, et enfin, l'utilisation d'un algorithme de minimisation à quatre familles de paramètres qui permet une estimation simultanée de tous les paramètres d'étalonnage. Des expériences réalisées sur des systèmes d'acquisition synthétiques et réelles évaluent le domaine de convergence de l'approche proposée ainsi que ses performances en termes de précision et de robustesse en présence de bruit.

ABSTRACT. This article presents a fully automated calibration method suitable for complex acquisition systems made with one or more cameras, an inertial measurement unit and a 3D lidar. The principle consists in estimating the intrinsic and extrinsic parameters by matching features detected in the camera images with the 3D point cloud provided by the rangefinder. This work proposes a mathematical formalization for the unification of the three types of sensors within the same likelihood function and a minimization algorithm with three families of parameters that enables the simultaneous estimation of all calibration parameters. Experiments conducted on synthetic acquisition systems and real sequences are presented and assess the area of convergence of the proposed approach and its performance in terms of accuracy, correctness and robustness in the presence of noise.

MOTS-CLÉS: Étalonnage automatique, Optimisation, Ajustement de faisceau, Centrale inertielle, Télémètre 3D.

KEYWORDS: Automatic calibration, Optimisation, Bundle-adjustment, IMU, 3D Lidar.

1. Introduction

Les systèmes d'acquisition de données mobiles, développés notamment dans un objectif de cartographie de l'environnement, utilisent actuellement différents capteurs : télémètres, caméras, GPS, centrale inertielle, etc. Chaque capteur fournit des informations relatives à son référentiel propre. Pour construire un modèle tridimensionnel cohérent il est alors nécessaire de recalibrer les différentes données dans un référentiel commun. Ceci implique d'une part de disposer d'une fonction de localisation qui estime le déplacement d'un repère de référence du système d'acquisition par rapport à l'environnement, et d'autre part de connaître la position relative des différents capteurs par rapport à ce repère de référence, i.e. leurs paramètres extrinsèques. Ainsi dans le cas du télémètre, l'estimation du déplacement combiné à la connaissance des paramètres extrinsèques permet le recalage des différents points d'impact dans un nuage de points 3D cohérent.

De la même façon, la détermination des différentes poses (orientation et translation) associées aux images de caméras permet de recalibrer ces images avec le nuage de points 3D obtenu par le capteur télémétrique et effectuer des traitements tels que la coloration du nuage de points, la détection des objets mobiles... Là encore, un algorithme de localisation peut fournir ces informations dès lors que l'on dispose d'un étalonnage extrinsèque et intrinsèque des caméras.

Les centrales inertielles de haute qualité fournissent ce recalage grâce à leur fréquence élevée ainsi que leur précision. Leur principe repose sur l'utilisation simultanée de deux organes de mesure différents : (i) un accéléromètre qui mesure l'accélération du système et (ii) un gyromètre mesurant les vitesses angulaires. Les mesures sont ensuite intégrées deux fois afin de restituer les poses 3D de la centrale inertielle. Ces systèmes appartiennent à la catégorie des capteurs proprioceptifs et, comme l'odométrie, ils dérivent au cours du temps à cause de biais et d'incertitude de mesure. De ce fait, ces outils sont communément couplés à un système de localisation extéroceptif, un GPS centimétrique par exemple.

Dans cet article, nous présentons une méthode d'auto-étalonnage estimant à la fois les paramètres extrinsèques (les transformations géométriques entre les capteurs) et les paramètres intrinsèques des capteurs d'un système d'acquisition comportant une centrale inertielle couplée à un GPS cinématique, un capteur télémétrique et une ou plusieurs caméras. Nous dressons un bref état de l'art dans la section suivante. Puis, nous formalisons le problème de l'auto-étalonnage multi-capteurs dans la section 3 sous la forme d'une fonction de vraisemblance, dont la minimisation à l'aide d'un algorithme d'optimisation (section 4.4) permet l'obtention des valeurs d'étalonnage. Les résultats obtenus sont analysés en termes de précision et de robustesse dans la section 5.

2. État de l'art

Cette section synthétise quelques travaux pertinents du domaine de l'étalonnage multi-capteurs, en se focalisant sur 1) l'étalonnage entre deux capteurs extéroceptifs et 2) l'étalonnage entre un capteur extéroceptif et une centrale inertielle.

La littérature présente déjà de nombreuses solutions d'étalonnage entre un télémètre et une caméra. La plupart des travaux se basent sur l'observation conjointe d'une mire plane possédant un motif damier par les deux capteurs (? ; ? ; ? ; ?). L'approche consiste à détecter le plan, puis, après une formulation théorique des équations de projection, à effectuer la minimisation d'un critère (figure 1) visant à contraindre les mesures télémétriques sur la surface plane. Ces solutions contraignantes nécessitent que les capteurs soient statiques pendant l'étalonnage. Il est donc impossible d'étalonner à posteriori une séquence déjà acquise. Des techniques plus automatiques ont vu le jour, l'une d'elles (? ; ?) se base sur l'information mutuelle entre la luminance mesurée par l'image d'une caméra et la réflectivité mesurée par un télémètre laser 3D. Il existe en effet un lien entre les variations de texture de l'image d'une caméra et la variation d'intensité de retour du laser. Il existe en effet un lien entre les variations de texture de l'image d'une caméra et la variation d'intensité de retour du laser. La maximisation de cette quantité d'information permet de remonter aux paramètres extrinsèques liant télémètre et caméra sans la nécessité de présenter une mire au système d'acquisition.

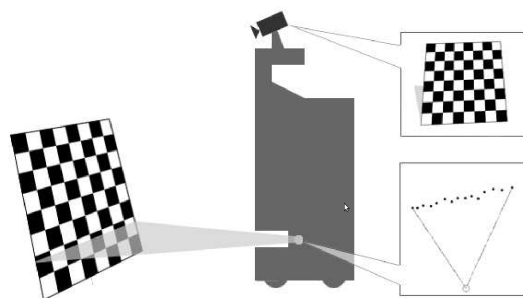


Figure 1. Illustration du processus d'étalonnage caméra-télémètre utilisant une mire plane. Le système d'acquisition, ici un robot, perçoit par son télémètre des informations de profondeur appartenant à un plan 3D et la caméra détecte le damier. Les deux informations sont intégrées au sein d'une fonction objectif tandis qu'un algorithme de minimisation retrouve le paramétrage extrinsèque du système.

L'étalonnage entre un système de caméras et une centrale inertielle est un problème peu traité. La méthode présentée dans (? ; ?) estime l'étalonnage extrinsèque entre ces deux capteurs. La rotation est obtenue en se basant sur l'estimation de la verticale perçue à la fois par la caméra par l'observation d'une mire et par la centrale inertielle dans la gravité. La translation est obtenue à partir d'un ensemble de poses dont les positions sont contraintes par l'utilisation d'un plateau tournant. Au final, l'algorithme fournit la transformation affine estimée. Cette solution présente le désavantage d'être

entièrement supervisée. De plus, elle nécessite une mire et reste spécifique à cette configuration.

L'étalonnage entre un capteur télémétrique et une centrale inertielle a aussi été effectué en utilisant l'algorithme ICP (Iterative Closest Point) (? ; ? ; ?) pour recalculer les données télémétriques et trouver les paramètres extrinsèques liant la trajectoire de la centrale inertielle et celle du télémètre. L'article (? ; ?) estime ces mêmes paramètres grâce à l'extraction d'un plan parmi les mesures télémétriques constituant le sol supposé plan. En plus de faire l'hypothèse de planéité du sol, le principe rend impossible l'observation d'un angle et de deux translations de la transformation. L'article contourne cette difficulté en détectant des poteaux dans la scène.

Aucune des approches citées ne permet une unification réelle de tous les capteurs au sein d'un même espace de travail de manière automatique et sans l'utilisation d'une mire ou d'une configuration spécifique. L'objectif de nos travaux est de développer une méthode d'auto-étalonnage caméra-système inertiel-télémètre flexible et ne posant aucun a priori sur la scène observée. La méthode que nous présentons ici est de plus capable de traiter de manière égale les données d'un capteur télémétrique à balayage quelle que soit sa configuration d'acquisition : qu'il observe plusieurs fois la scène ou qu'il ne l'observe qu'une seule fois comme les télémètres dont le balayage s'effectue de manière perpendiculaire au déplacement du système d'acquisition.

3. Formalisation

Cette section présente la formalisation mathématique du problème d'auto-étalonnage multi-capteurs sous la forme d'une estimation de maximum de vraisemblance.

Dans la suite nous considérons un système d'acquisition composé d'un télémètre, d'une centrale inertielle et de m caméras. Nous supposons que nous disposons d'une localisation globale précise du système par rapport à l'environnement sous la forme de poses notées T_{imu_i} . Cette hypothèse se justifie dans la réalité lors de l'utilisation d'une centrale inertielle de haute qualité couplée à un GPS cinématique.

L'objectif de notre approche est alors d'estimer : T_{imu}^t la transformation géométrique permettant le passage de la centrale inertielle au repère du télémètre, $T_{imu}^{c_j}$ les transformations de la centrale inertielle aux m caméras $\mathcal{C} = \{c_j\}_{j=1..m}$ ainsi que les paramètres intrinsèques θ_j des caméras (voir figure 2). La composition $T_{imu}^{c_j} T_{imu}^t^{-1}$ permet de définir une transformation entre la caméra et le capteur télémétrique directement. Afin d'obtenir une méthode fonctionnant sans l'aide d'une mire, d'une configuration particulière ou d'une trajectoire singulière, il nous faut définir une fonction de contrainte liant les informations caméra avec les informations télémétriques. Pour cela, nous proposons une approche basée primitive consistant à détecter des points d'intérêts dans les images puis à les trianguler dans l'espace pour finalement les comparer avec le nuage de mesures télémétriques 3D afin d'en extraire un étalonnage. On suppose dans ce travail disposer de données télémétriques denses et d'une

connaissance à priori des valeurs approximatives des paramètres d'étalonnage. Celles-ci peuvent être obtenues à partir de données constructeurs et de mesures grossières.

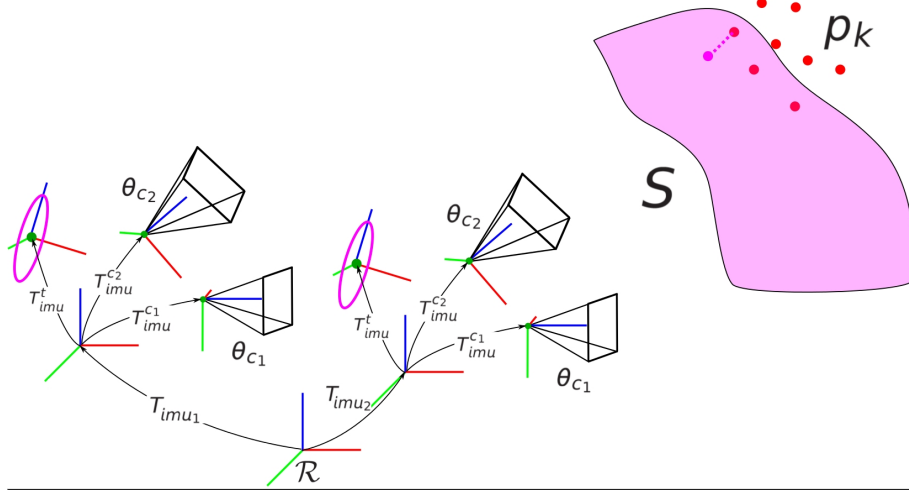


Figure 2. Illustration du problème d'étalonnage caméra-système inertiel-téléporteur. La position d'un système d'acquisition est représentée à deux instants différents. T_{imu_1} et T_{imu_2} sont les poses du système inertiel à ces deux moments, par rapport à un repère fixe \mathcal{R} . $T_{imu}^{c_1}$ et $T_{imu}^{c_2}$ sont les paramètres extrinsèques à retrouver et transformant la pose de la centrale inertielle vers le repère des caméras. \mathcal{P} l'ensemble des points 3D issus de la géométrie reconstruite grâce aux images et S la surface engendrée par les données télémétriques. L'objectif est de minimiser les erreurs de projection des p_k dans les images ainsi que l'erreur entre p_k et la surface S afin de restituer toutes les valeurs d'étalonnage.

Soit $F_{(k,j,r)}$ le r -ième point d'intérêt détecté au sein de l'image d'index k provenant de la caméra d'index j . Soit \mathcal{P} l'ensemble des points 3D p de l'espace correspondant à la triangulation des observations du repère image. On peut alors exprimer la vraisemblance L_1 des observations caméra en fonction des paramètres θ_j , $T_{imu}^{c_j}$ et des poses de la centrale inertielle notées T_{imu_i} sous la forme :

$$L_1 = \prod_{k=1}^n \prod_{j=1}^m \prod_{r=1}^l f(F_{(k,j,r)}; \theta_j, T_{imu}^{c_j}, T_{imu_k}), \quad (1)$$

où f est une densité de probabilité, n est le nombre d'images de la séquence, m le nombre de caméras, l le nombre de primitives présentes dans l'image et \mathcal{F} l'ensemble des primitives présentes dans la séquence. Sous l'hypothèse que le bruit du détecteur de primitives de l'image est gaussien d'écart type σ_g la maximisation de la log-vraisemblance équivaut à une minimisation des moindres carrés des erreurs de reprojection :

$$-\log(L_1) = \frac{1}{2\sigma_g^2} \sum_{k=1}^n \sum_{j=1}^m \sum_{r=1}^l \|\pi(\theta_j, T_{imu}^{c_j}, T_{imu_k}, p(k, j, r)) - F_{(k,j,r)}\|^2 \quad (2)$$

avec π la fonction de projection dans la caméra j dépendante des paramètres θ_j et $p(k, j, r)$ le point 3D associé à l'observation $F_{(k,j,r)}$.

À ce stade, les données télémétriques n'ont pas été insérées dans la formalisation. La minimisation de cette fonction et l'obtention des paramètres optimaux seront abordées dans la section 4.4 et font intervenir des algorithmes d'optimisation non linéaire.

L'intégration des données télémétriques s'effectue de la même manière. Les points télémétriques acquis au cours du déplacement sont recalés dans un même nuage de points. Ce nuage constitue un échantillonnage de la surface réelle des objets. Par ailleurs les primitives issues des images correspondent à des objets de la scène et sont donc localisées à la surface de ces objets. Ces primitives doivent donc appartenir à la surface représentée discrètement par les données télémétriques. Nous allons donc contraindre par l'ajout d'un critère supplémentaire les primitives 3D issues des images à être présentes sur la surface définie par les données télémétriques. En reprenant l'équation (1) et en admettant que la surface possède un bruit gaussien σ_s on obtient le nouveau critère :

$$-\log(L_2) = -\log(L_1) + \sum_{p \in \mathcal{P}} \frac{\|p - S(T_{imu}^t, T_{imu})\|^2}{2\sigma_s^2} \quad (3)$$

où $S(T_{imu}^t, T_{imu})$ représente la surface engendrée par les mesures télémétriques. Cette surface est dépendante des poses de la centrale inertielle T_{imu} , mais aussi de la transformation entre la centrale inertielle et le capteur télémétrique T_{imu}^t qui est le paramètre à estimer. Il est donc nécessaire de remettre à jour la surface à chaque fois qu'une modification de ces paramètres est effectuée. La section suivante décrit la mesure de distance proposée entre la primitive et la surface $\|p - S(T_{imu}^t, T_{imu_k})\|$.

3.1. Distance à la surface engendrée

La surface S n'est pas directement observable. En réalité, nous ne la connaissons que par l'intermédiaire des points de mesures télémétriques. Afin de définir totalement le score d'optimisation exprimé par l'équation (3), il est nécessaire d'estimer une distance entre la primitive 3D et la surface déterminée par les données télémétriques.

La distance considérée est la distance du point p au point de la surface S le plus proche :

$$\|p - S(T_{imu}^t, T_{imu_s})\| = \min_{(a,b)} \|p - S(a, b)\| \quad (4)$$

dans laquelle a et b sont une paramétrisation de la surface à deux dimensions.

Dans le cas où la surface est échantillonnée de manière dense, nous pouvons remplacer cette distance à la surface par la distance à la mesure télémétrique la plus proche, ce qui se formalise par :

$$\|p - S(T_{imu}^t, T_{imu_s})\| = \min_{\mathcal{Z}} \|p - z_i\| \quad (5)$$

où $\mathcal{Z} \doteq \{z_k\}_{k=1, \dots, n}$ représente l'ensemble des points d'impact télémétriques exprimés dans le repère monde \mathcal{R} (voir Figure 2). La minimisation de la fonction de vraisemblance nécessitera d'effectuer ce calcul pour chaque primitive et à chaque itération avec un nombre très important de données télémètre. La solution employée devra ainsi être étendue à la résolution de problèmes considérant au minimum 10^5 primitives et 10^8 mesures télémétriques, ce qui correspond à une trajectoire d'étalonnage de taille standard. Nous détaillerons dans la partie applicative le calcul rapide de cette distance ainsi que l'estimation de σ_s et σ_g .

3.2. Mouvements singuliers

La fonction objectif employée laisse apparaître plusieurs mouvements singuliers ne permettant pas la détermination complète de tous les paramètres à optimiser. Nous devons bannir les trajectoires rectilignes pour lesquelles la translation latérale du système par rapport à la trajectoire de la centrale inertielle laisse le critère invariant. En effet, en ajoutant simultanément une translation à $T_{imu}^{c_j}$ et T_{imu}^t les primitives et la surface se décaleraient d'autant et laisseraient inchangé le score d'optimisation. De même, les trajectoires planes ne permettent pas de déterminer la composante verticale de $T_{imu}^{c_j}$ et T_{imu}^t . Toute translation selon la normale au plan n'induirait pas de variation de coût. Une bonne trajectoire d'étalonnage possède aux moins deux zones de courbures différentes (virages) ainsi qu'une variation de hauteur ou de tangage du système d'acquisition. Une analyse poussée des mouvements singuliers des systèmes multi-caméras a été effectuée dans (?).

4. Méthode de résolution proposée

Cette section détaille les solutions proposées pour résoudre le problème d'auto-étalonnage.

4.1. Principe général

Le problème dont la formulation théorique a été présentée dans la partie précédente est résolu en minimisant la fonction objectif intégrant les différents paramètres intrinsèques et extrinsèques. Cette méthode se déroule en 5 étapes successives. Premièrement, des primitives de type point sont extraites des images grâce à un détecteur présenté dans la section 4.3. Dans un second temps, des descripteurs locaux de texture sont extraits autour de chaque point afin de pouvoir les identifier et les comparer

aux autres. Les primitives sont associées entre les différentes images ce qui permet d'estimer la position tridimensionnelle de chaque point par triangulation.

Une fois ces trois phases réalisées, un algorithme de minimisation non linéaire permet d'optimiser tous les paramètres à estimer : paramètres intrinsèques des caméras, paramètres extrinsèques des caméras et du télémètre, ainsi que la position des points 3D triangulés. L'algorithme se déroule en deux parties majeures qui sont répétées de manière itérative :

(i) On effectue l'estimation des erreurs (aussi appelées résidus) qui sont composées des erreurs vision, calculées en reprojectant les primitives dans les images ainsi que des erreurs à la surface télémétrique, calculées grâce à la distance entre les primitives et la surface comme défini à la section 3.1. (ii) Un incrément est estimé et corrige le vecteur de paramètres, étape présentée en détail dans la section 4.4. La minimisation s'arrête lorsque les erreurs de reprojection et les erreurs de distance entre les points triangulés et la surface télémétrique sont minimales. Une vision schématique du processus est représentée par la figure 3.

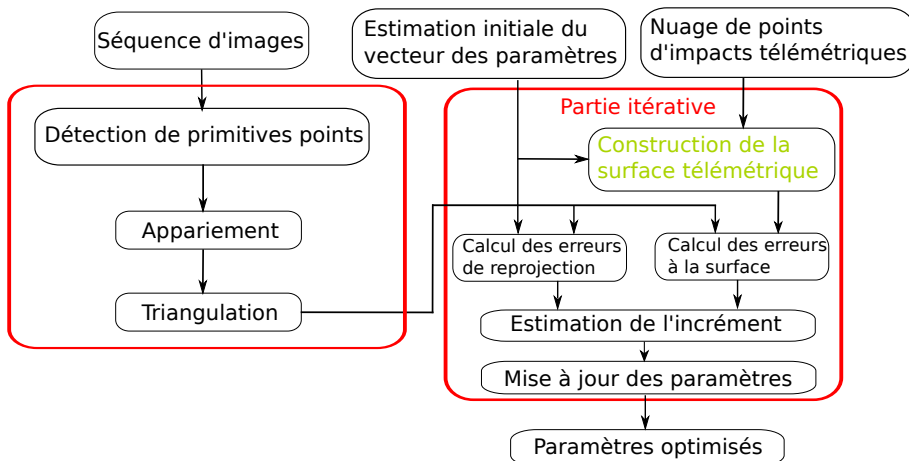


Figure 3. Principe général de la méthode d'étalonnage.

4.2. Modélisation du système

Cette section décrit les différents modèles paramétriques choisis dans la formalisation géométrique du problème d'auto-calibration.

4.2.1. Modèle de caméra unifié

Les caméras sont modélisées à l'aide du modèle unifié (?). Il s'est avéré être capable de modéliser avec le même formalisme mathématique autant les caméras perspectives que les caméras grands-angles de type Fish-Eye. Contrairement au modèle sténopé dont l'étape de projection s'effectue directement sur le plan focal, le modèle

unifié passe par une projection sur la sphère puis une seconde phase de projection sur un plan. Cette solution a la propriété d'être une excellente modélisation des phénomènes physiques donnant lieu à la distorsion géométrique. Une illustration du fonctionnement de ce modèle et des différents repères et notations est disponible sur la figure 4. Ce modèle de caméra possède cinq paramètres intrinsèques (f_x, f_y, u_0, v_0, ξ).

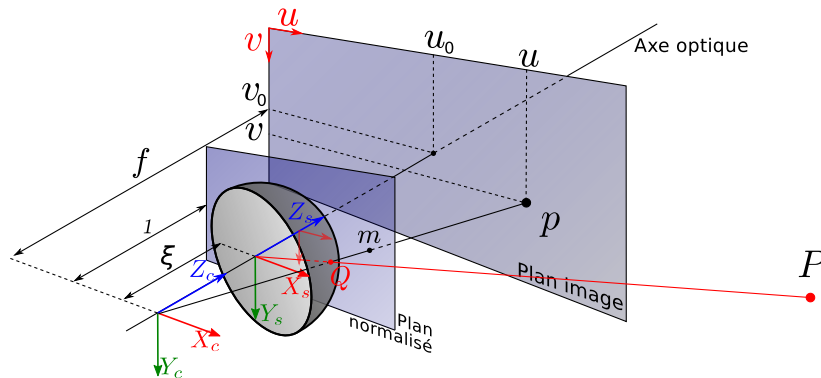


Figure 4. Illustration du principe du modèle unifié.

4.2.2. Paramétrisation extrinsèque

La paramétrisation extrinsèque, c'est-à-dire les transformations géométriques sont représentées par trois paramètres de translation exprimés par leurs coordonnées cartésiennes et trois paramètres de rotation. Ceux-ci sont exprimés grâce à un atlas de carte exponentielle locale comme expliqué dans les travaux de (? , ?).

4.3. Détection d'amer, appariement et triangulation

Les primitives de type points sont détectées en utilisant le détecteur de point d'intérêt de Harris et Stephen (? , ?). L'appariement de celles-ci est réalisé à l'aide de descripteurs locaux de type *ZNCC* permettant l'invariance à la luminosité de la scène, de manière similaire aux travaux de (? , ?). Pour chaque caméra, la recherche d'appariement s'effectue de façon incrémentale, entre les images successives I_k et I_{k+1} de la séquence. Puisque des paramètres extrinsèques et intrinsèques approximatifs sont supposés connus, nous connaissons approximativement la pose de chaque image. Cette information permet l'utilisation de la contrainte épipolaire pour réduire la zone de recherche des appariements, tout en conservant une marge d'erreur sur la distance à la droite pour tenir compte de l'erreur dans les paramètres initiaux.. Finalement pour chaque primitive 3D nous possédons la listes des images où elle est vue.

La restitution en 3D des primitives points s'effectue donc directement par triangulation à partir des points d'intérêt à l'aide de la méthode du point du milieu adapté au modèle unifié. La position des points 3D sera optimisée au même titre que les paramètres de calibration lors de l'étape de minimisation décrite ci-dessous.

4.4. Minimisation

La minimisation de la fonction de vraisemblance est effectuée à l'aide d'un algorithme d'optimisation non-linéaire. La problématique est donc l'optimisation conjointe de l'ensemble des paramètres extrinsèques, des paramètres intrinsèques des caméras ainsi que la position de tous les points 3D triangulés. Ce cas d'utilisation est très proche de la notion d'ajustement de faisceaux classiquement utilisé en "*Structure from Motion*". La seule différence réside dans le fait que les poses des images ne sont pas optimisées. Pour cette raison, nous avons utilisé un algorithme classique : Levenberg-Marquardt (? , ?) qui se formalise par un pas et une direction d'optimisation définis par :

$$X_{k+1} = X_k - (J^T J + \lambda I)^{-1} J^T \epsilon \quad (6)$$

Dans lequel J est la jacobienne du système, λ le coefficient d'amortissement, X le vecteur comprenant l'ensemble des paramètres à optimiser aux itérations k et $k + 1$. ϵ représente le vecteur des résidus. En pratique nous avons ici quatre familles de paramètres à estimer : 5 paramètres intrinsèques par caméra, 6 paramètres extrinsèques par caméra, 6 pour le télémètre, et 3 paramètres de position pour chaque point 3D. La taille importante de la jacobienne ne permet pas la résolution par inversion de la matrice $H = J^T J$. L'implémentation de la résolution utilise la technique du complément de Schur (? , ?) et tire parti de la structure creuse de la matrice H .

4.4.1. Estimation rapide des résidus

Les résidus sont constitués de deux parties, l'une correspondant aux erreurs de reprojection des primitives et l'autre aux distances entre les primitives et la surface télémétrique. Les premières sont calculées de la manière suivante : pour chaque primitive point 3D p de la scène et chaque observation ν fournie par le détecteur de Harris dans une image, la primitive est projetée dans le repère image grâce à la fonction de projection du modèle unifié. Le vecteur des résidus pour cette primitive dans cette image est défini par $\vec{e}(p, \nu) = \pi(p, X) - \nu$, où X est le vecteur de paramètres de l'optimisation contenant les paramètres extrinsèques et intrinsèques de la caméra. Pour les résidus liés à la distance entre la primitive 3D et la surface télémétrique, il est nécessaire de régénérer la surface télémétrique pour la valeur des paramètres extrinsèques du télémètre fournie par l'optimisation. À ces fins, nous effectuons un changement de repère afin d'exprimer les données télémétriques dans le repère local au capteur dans un repère global à l'aide de la centrale inertielle. Cette transformation se formalise par $T_{imu_i}^t = T_{imu}^t T_{imu_i}$ où T_{imu}^t représente les paramètres extrinsèques du télémètre (présent dans X) et T_{imu_i} la i -ème pose de la centrale inertielle correspondant à la même date que la donnée télémétrique à recalculer. La fréquence d'acquisition du capteur télémétrique étant souvent supérieure à la fréquence de rafraîchissement de la centrale inertielle, nous ne disposons pas directement de la pose inertielle correspondant à la date de chaque point. Nous interpolons donc linéairement (dans l'algèbre de Lie pour la partie rotation) les transformations inertielles à la date de chaque point télémétrique pour pouvoir les recalculer dans un même repère. Comme présenté dans la section 3.1 la distance à la surface est approchée par la distance minimale entre chaque

primitive et les points télémétriques. Une recherche de voisinage est ainsi mise en oeuvre pour extraire la donnée télémétrique la plus proche de chaque primitive. Cependant, face au volume très important de mesures télémétriques et au grand nombre de primitives détectées par image, nous exploitons un arbre de partitionnement pour accélérer ces requêtes. Les points d'impact des mesures télémétriques sont insérés dans un Octree (?, ?) à chaque itération de l'algorithme d'optimisation, et pour chaque primitive, nous recherchons le point q le plus proche grâce à un algorithme de voisinage classique. Le résidu fourni à la minimisation est : $\vec{\epsilon}(p, q) = p - q$.

4.4.2. Évaluation des écart-types des mesures

La formulation de la fonction de coût (3) fait intervenir les écart-types des bruits de mesures, σ_g pour le bruit des détections dans le repère image et σ_s pour l'écart type relatif au bruit de la surface télémétrique utilisée pour le calcul des résidus. σ_s représentant le bruit de la surface, il est donc choisi égal à l'écart type de mesure du capteur télémétrique. Cette donnée est disponible dans les documents techniques joints au capteur. σ_g représentant l'écart-type sur la détection des points de Harris, nous employons la valeur $2px$ communément utilisée comme approximation par la communauté scientifique. Notons que la matrice de covariance du détecteur de Harris peut être calculée de manière analytique par $\Sigma = \sigma_p^2 H(\nu)^{-1}$ où σ_p représente le bruit photogrammétrique pixellique de la caméra, $H(\nu)$ la matrice de Harris estimée au coin détecté ν .

5. Validation expérimentale

Cette section présente les expérimentations mises en oeuvre pour évaluer la méthode d'auto-étalonnage proposée.

5.1. Méthodologie

Afin d'analyser les performances de cette méthode d'étalonnage extrinsèque caméra - système inertiel - télémètre, nous avons réalisé des séquences réelles et synthétiques. Le système d'acquisition considéré pour cette validation est composé d'une caméra, d'un télémètre RiegI et d'une centrale inertielle figure 5.

5.1.1. Séquences de synthèse

Pour avoir une vérité terrain nous avons utilisé dans un premier temps un simulateur réaliste embarquant tous les capteurs dans un véhicule d'acquisition virtuel (4D-Virtualiz (?, ?)). Dans cet environnement le véhicule virtuel est conduit manuellement à l'aide d'un joystick. La caméra est installée sur le toit du véhicule et perpendiculairement au sens d'avancement, le télémètre RiegI est implanté à l'arrière du toit et tel que son plan de rotation soit perpendiculaire au sens d'avancement du véhicule tandis que la centrale inertielle est disposée au centre de l'essieu arrière. Ces capteurs simulés possèdent les mêmes propriétés que les capteurs réels en termes de bruit d'acquisition.

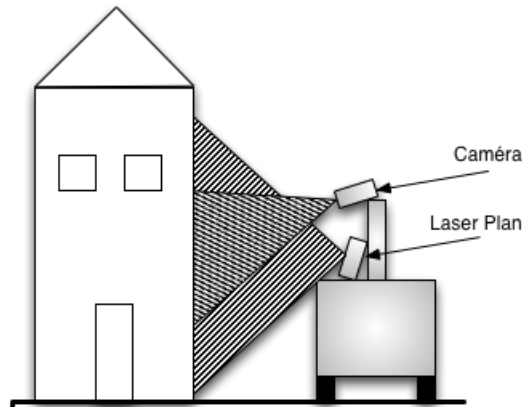


Figure 5. Schéma de positionnement des capteurs extéroceptifs utilisés sur le véhicule.

Dans le cadre des séquences synthétiques, les paramètres extrinsèques et intrinsèques approximatifs sont obtenus par l'ajout d'un biais aux paramètres exacts. Les valeurs employées sont consignées dans le tableau 2. Dans un dernier temps, nous employons des déplacements beaucoup plus importants dans le but de déterminer le bassin de convergence¹ de notre approche. Celui-ci permettra de définir le cadre d'utilisation de cette méthode ainsi que sa robustesse aux mauvaises initialisations.

5.1.2. Séquences réelles

Les séquences réelles sont acquises sur le véhicule d'acquisition Stéréopolis (?). Ce véhicule est équipé de multiples capteurs rendant possible l'acquisition de données cartographiques précises (voir Figure 6). Dans nos expérimentations nous avons utilisé les images provenant de la caméra gauche du réseau de caméras, les données télémétriques du Riegl ainsi que les poses géoréférencées de la centrale inertielle.

Télémètre

Le télémètre utilisé est un Riegl du modèle $VQ - 250$. Ce capteur extrêmement rapide permet l'acquisition de plus de 3.10^5 mesures par seconde tout en effectuant des rotations avec une fréquence de $100Hz$. La portée maximale atteint $500m$ tout en conservant une erreur centimétrique. Il est aussi doté d'un système de synchronisation avec le GPS et d'une horloge permettant une datation précise de chaque rayon de mesure.

1. Sous-ensemble de l'espace des paramètres pour lequel l'optimisation converge.



Figure 6. Vue d'ensemble du véhicule d'acquisition Stéréopolis.

Réseau de caméras

Le véhicule comporte 14 caméras HD. Celles-ci fournissent des images couleur haute résolution 1920×1080 dont le déclenchement est réalisé par un signal ("trigger") externe. Les caméras disposent toutes d'un objectif similaire et d'un temps d'exposition identique garantissant que les images acquises par chaque caméra aient la même luminosité. Les expériences présentées n'utilisent ici qu'une seule de ces caméras, orientée de la même façon que pour les séquences synthétiques, mais la formalisation proposée s'applique de la même façon pour plusieurs caméras.

Centrale inertielle

Afin de recalibrer les mesures télémétriques dans un repère commun, il est nécessaire de posséder un système de localisation précis et intègre du véhicule. Pour cela, Stéréopolis a été doté d'une centrale inertielle Applanix fournissant des poses géoréférencées précises grâce au couplage avec les données GPS (? , ?).

5.1.3. Critères d'évaluation

Le critère d'évaluation considéré pour les séquences de synthèse est l'erreur entre les résultats obtenus et la vérité terrain. Nous comparons les paramètres intrinsèques des caméras par leur différence :

$$\Delta(u_0, v_0, f_x, f_y, \xi) = (u_0, v_0, f_x, f_y, \xi)_e - (u_0, v_0, f_x, f_y, \xi)_r$$

où les indices e et r signifient Estimé et Réel. Nous faisons de même pour les translations présentes dans les transformations extrinsèques. Pour les rotations, nous calculons les paramètres dans l'algèbre de Lie associée à l'espace des rotations pour la matrice de passage entre l'estimation et la vérité qui s'écrit : $\omega = \log(R_r^T R_e)$. Pour plus de lisibilité, on calcule aussi l'angle θ de rotation dans la représentation axe-angle entre les deux rotations donnée par $|\theta| = \frac{\|\log(R_r^T R_e)\|_F}{\sqrt{2}}$. Ces données sont présentées sous la forme d'un tableau pour plusieurs étalonnages et plusieurs séquences.

Par construction, les séquences réelles ne disposent pas de vérité terrain exacte. Dans ce cas, la précision du résultat est appréciée à l'aide de la répartition des erreurs

de reprojection post-optimisation ainsi qu'un histogramme des erreurs à la surface télémétrique.

5.2. Propriétés des séquences utilisées

Pour conduire nos expériences, nous utilisons quatre séquences synthétiques nommées $S1$, $S2$, $S3$ et $S4$ ainsi que deux séquences réelles $R1$, $R2$. Ci-après, le détail des scénarios ainsi que leurs objectifs.

5.2.1. Séquences synthétiques

- $S1$ est une trajectoire d'étalonnage courte longeant un bâtiment puis tournant à une intersection. Cette séquence vise à démontrer les capacités d'estimation de la méthode et le bassin de convergence dans un cas simple. La trajectoire est volontairement choisie singulière (plane) afin d'en mesurer l'impact sur les paramètres extrinsèques obtenus.

- $S2$ et $S3$ représentent des séquences plus conséquentes. Celles-ci parcourent une centaine de mètres le long de la voie de tramway du modèle numérique de la place de Jaude, Clermont-Ferrand, France. Elles comportent plusieurs virages tandis que le véhicule virtuel fait l'acquisition d'un grand nombre de mesures télémétriques et d'images comme le montre le tableau 1. Les données télémétriques sont acquises à grande distance et donc extrêmement dépendantes de l'étalonnage angulaire du capteur. Ces séquences visent à démontrer que l'algorithme converge malgré la déformation de la surface télémétrique originale par des paramètres approximatifs.

- La trajectoire $S4$ est une séquence massive de millions de mesures télémétriques couvrant une grande longueur et un corpus massif d'images à traiter. Elle est composée de plusieurs virages et de grandes courbes afin d'illustrer la dépendance de la précision du résultat à la trajectoire choisie.

5.2.2. Séquences réelles

Afin de valider l'algorithme d'étalonnage en conditions réelles, nous étudions deux séquences $R1$, $R2$ provenant du véhicule d'acquisition Stéréopolis. Ces séquences ont été capturées durant une campagne d'acquisition dans le centre de Paris, France. Leurs longueurs et leurs propriétés sont recensées dans le tableau 1. Toutes deux sont non singulières et se réfèrent à des contextes que l'on retrouve en cartographie : carrefours, lignes droites et virages.

Tableau 1. Statistiques concernant les séquences utilisées pour l'étalonnage.

Séquence	$S1$	$S2$	$S3$	$S4$	$R1$	$R2$
Nb mesures télémétriques (10^3) :	217	262	263	627	14730	15851
Nb points 3D triangulés :	8314	52014	43511	75247	24761	12638
Longueur de la séquence (m)	20	80	70	250	260	200
Temps de calcul (s)	15	20	20	65	120	120

Tableau 2. Biais ajouté aux paramètres exacts des séquences synthétiques.

	S1	S2 - 1	S2 - 2	S3	S4
Paramètres intrinsèques caméra					
Δu_0 (px)	0	-5.0	0.0	7.0	5.0
Δv_0 (px)	0	5.0	0.0	7.0	5.0
Δf_x (px)	0	0.0	0.0	-10.0	10.0
Δf_y (px)	0	0.0	0.0	-10.0	10.0
$\Delta \xi$	0.1	0.1	0.0	0.05	0.05
Paramètres extrinsèques caméra					
Δt_x (m)	0.2	0.05	-0.1	0.2	0.15
Δt_y (m)	0.2	0.05	-0.1	0.1	0.20
Δt_z (m)	0.2	0.0	-0.1	0.5	0.15
$\Delta \omega_x$ (°)	3.0	0.0	0.0	3.0	1.0
$\Delta \omega_y$ (°)	3.0	0.0	0.0	3.0	-1.0
$\Delta \omega_z$ (°)	3.0	0.0	0.0	3.0	1.0
Paramètres extrinsèques télémètre					
Δt_x (m)	0.0	0.05	-0.05	-0.15	-0.1
Δt_y (m)	0.0	0.05	-0.05	0.05	-0.05
Δt_z (m)	0.3	0.0	-0.2	0.05	0.05
$\Delta \omega_x$ (°)	0.0	3.0	-5.0	2.0	2.0
$\Delta \omega_y$ (°)	0.0	3.0	-3.0	0.0	-2.0
$\Delta \omega_z$ (°)	0.0	3.0	-5.0	0.0	2.0

5.3. Résultats qualitatifs

Les figures suivantes, 7, 8, 9 et 10 illustrent les résultats obtenus avec les séquences synthétiques tandis que 11 et 12 représentent ceux obtenus avec les séquences réelles.

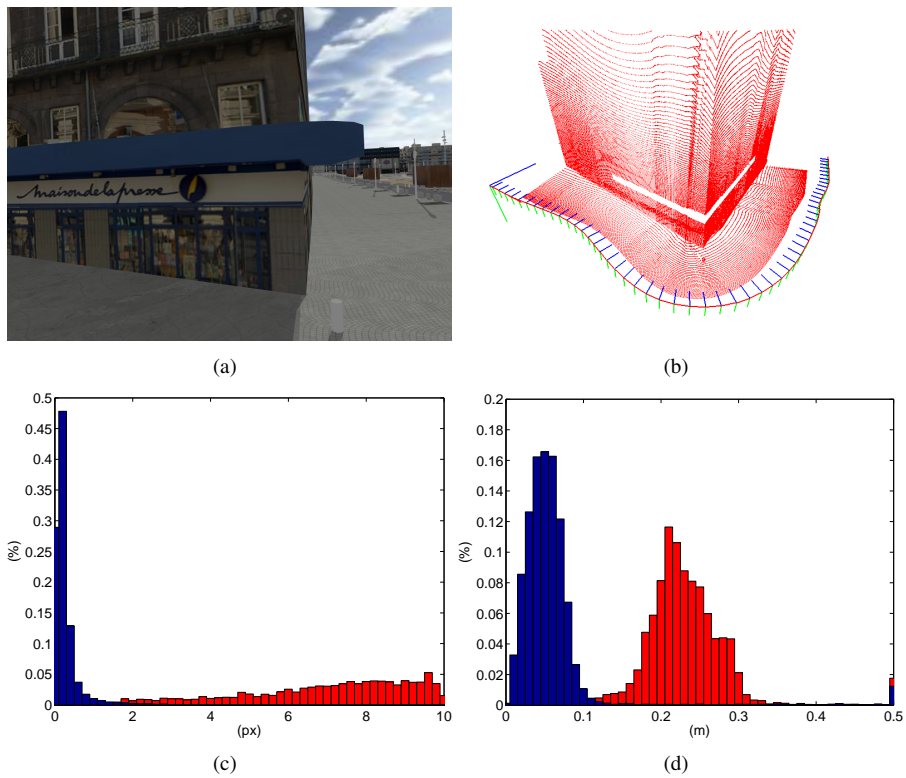


Figure 7. Résultats d'étalonnage obtenus sur la séquence $S1$. (a) présente l'une des images acquises par la caméra du véhicule d'acquisition virtuel durant sa trajectoire. (b) illustre le nuage de points 3D reconstruit à l'aide des paramètres approximatifs ainsi que la trajectoire de la centrale inertielle (représentée par les repères tricolores). (c) et (d) sont des histogrammes de la norme des résidus associés respectivement aux erreurs de reprojection et aux erreurs métriques à la surface. Dans ceux-ci, on trouvera en rouge la répartition avant optimisation et en bleu après.

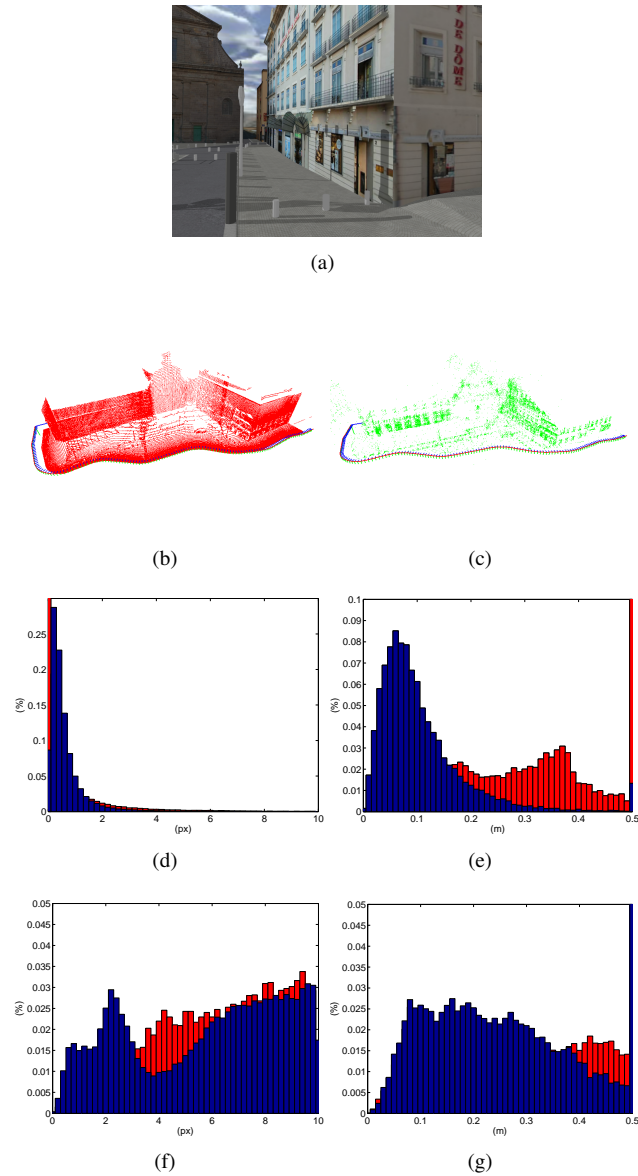


Figure 8. Résultats d'étalonnage obtenus avec la séquence $S2$. Au sommet, (a), est une des images caméras utilisées, (b) et (c) présentent respectivement le nuage de mesures télémétriques en rouge et le nuage de points 3D triangulés en vert. La trajectoire de la centrale inertielle est représentée par la succession de repères tricolores. (d) et (e) illustrent la norme des résidus de reprojection pour un petit biais ($S2 - 1$) appliqué aux paramètres tandis que (f) et (g) montrent les répartitions lors d'un biais important ($S2 - 2$) conduisant à la convergence vers un minimum local.

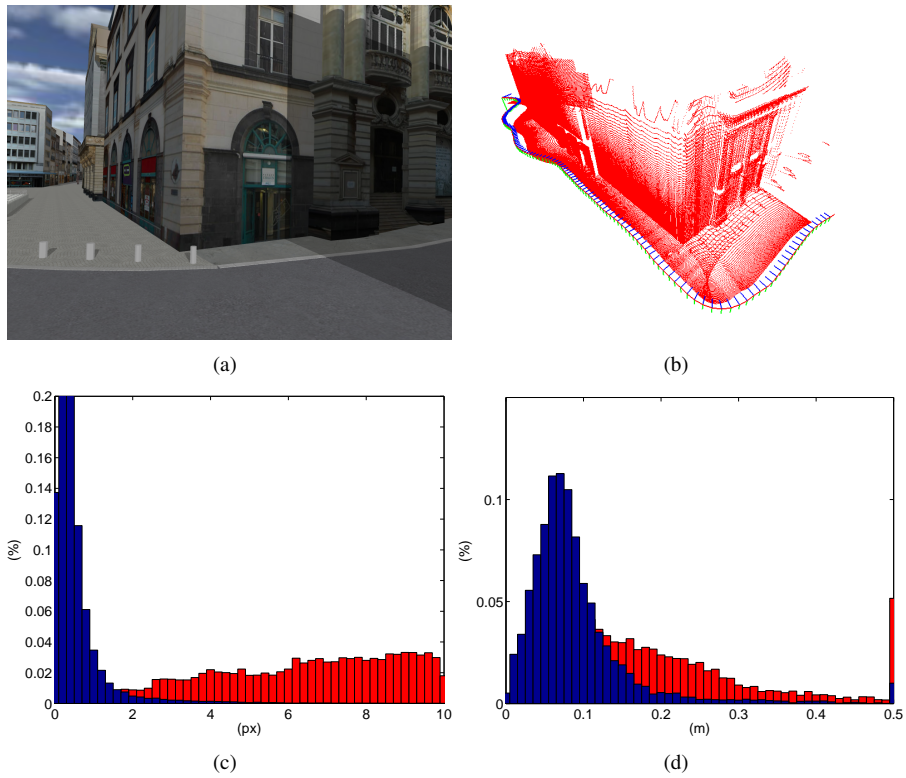


Figure 9. Résultats d'étalonnage obtenus avec la séquence S3. (a) est l'une des images acquises par la caméra, (b) le nuage de points recalé à l'aide des paramètres approximatifs tandis que (c) et (d) représentent les histogrammes de la norme des résidus de reprojection et à la surface avant (en rouge) et après (en bleu) optimisation.

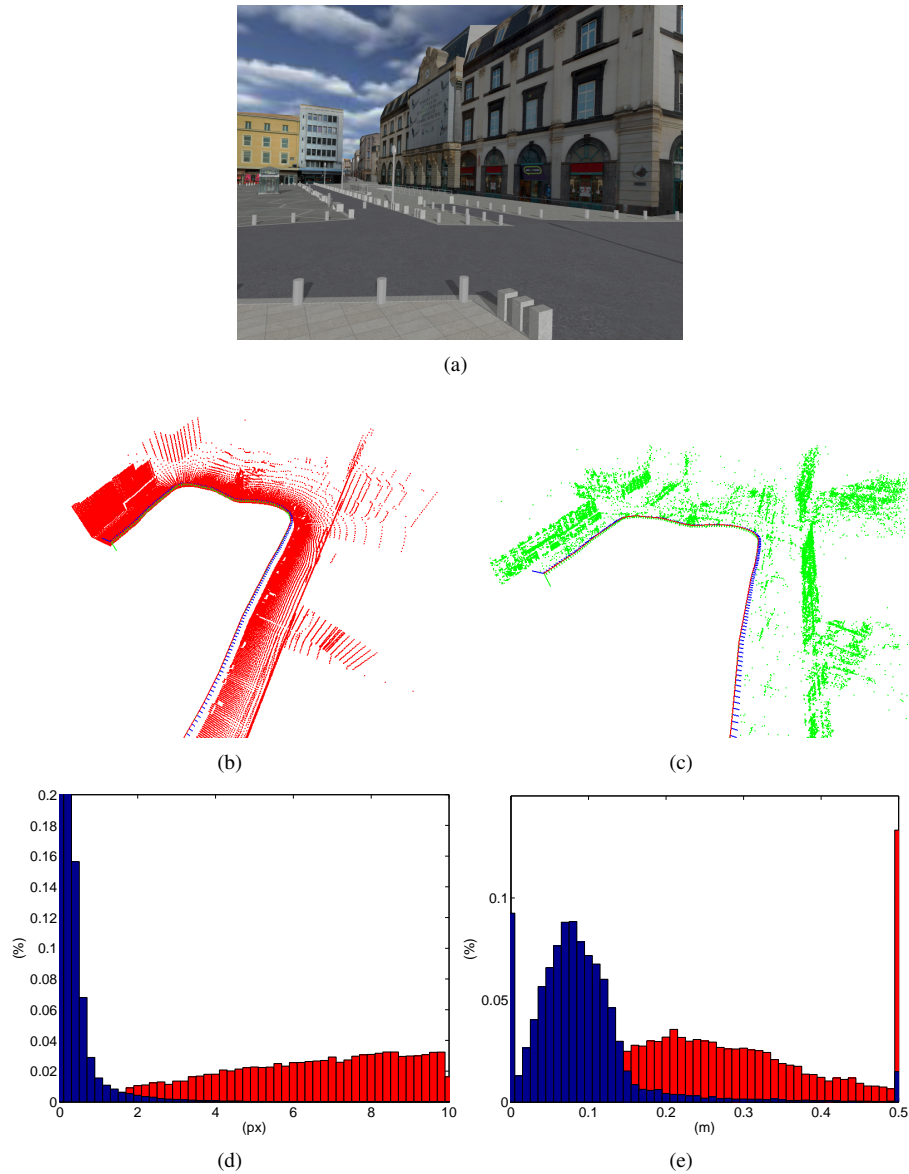


Figure 10. Résultats d'étalonnage obtenus avec la séquence S_4 . (a) est l'une des images acquises par la caméra, (b) le nuage de points recalé à l'aide des paramètres approximatifs tandis que (c) illustre le nuage de points triangulés. Les images (d) et (e) représentent respectivement l'histogramme des erreurs de reprojection et l'histogramme de la norme des résidus de reprojection et à la surface avant (en rouge) et après (en bleu) optimisation.

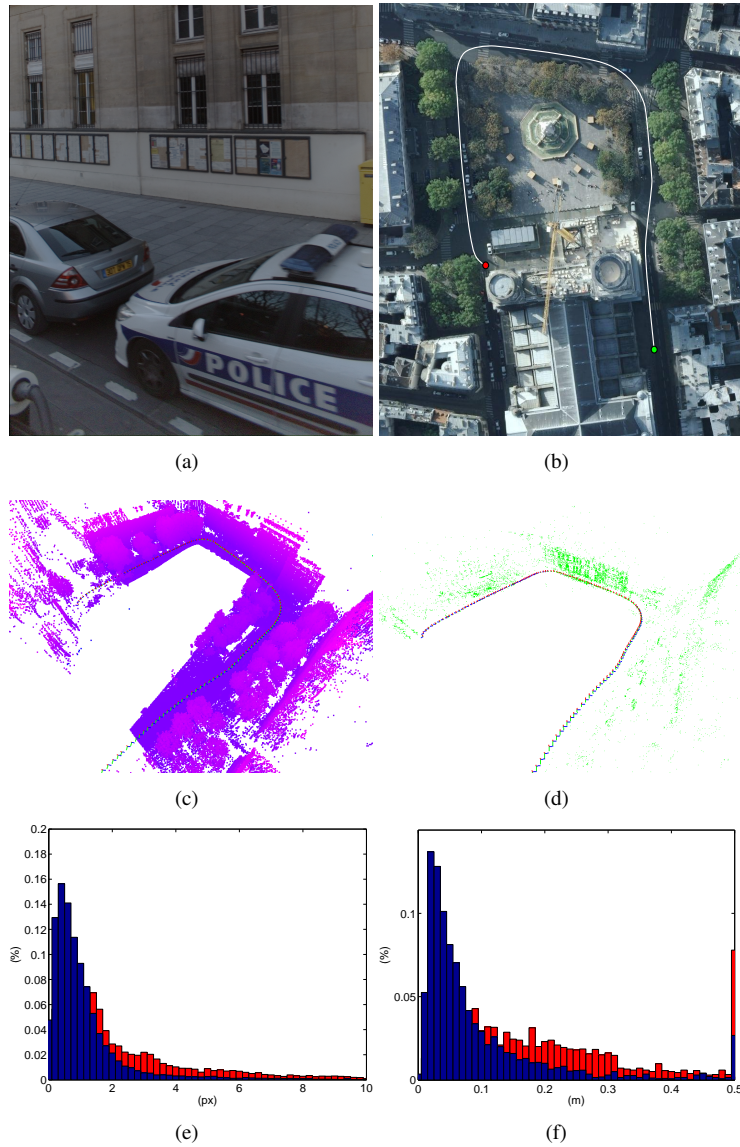


Figure 11. Résultats d'étalonnage obtenus avec la séquence R1. (a) est l'une des images acquises par la caméra, (b) la trajectoire superposée à une image satellite de la zone d'expérimentation. (c) le nuage de points recalé à l'aide des paramètres approximatifs tandis que (d) illustre le nuage de points triangulés. Les images (e) et (f) représentent respectivement l'histogramme des erreurs de reprojection et l'histogramme de la norme des résidus de reprojection et à la surface avant (en rouge) et après (en bleu) optimisation.

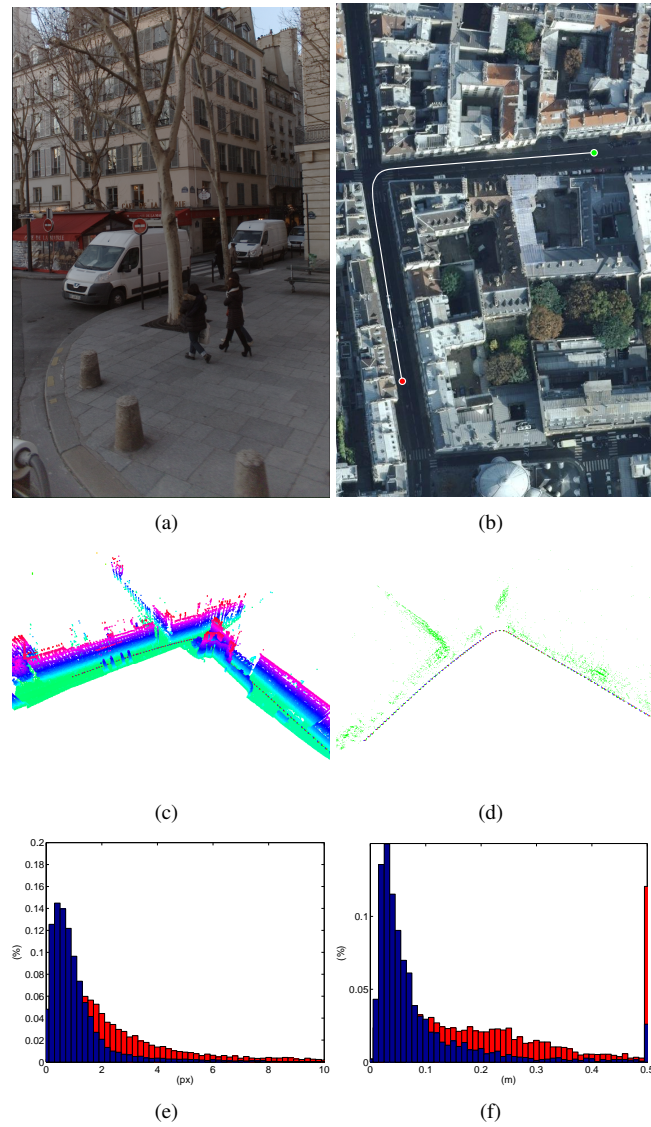


Figure 12. Résultats d'étalonnage obtenus avec la séquence R2. Résultats d'étalonnage obtenus avec la séquence R2. (a) est l'une des images caméra, (b) illustre la trajectoire superposée à une image satellite de la zone d'expérimentation. (c) est le nuage de points recalé à l'aide des paramètres approximatifs tandis que (d) montre le nuage de point triangulé. Les images (e) et (f) représentent respectivement l'histogramme des erreurs de reprojection et l'histogramme de la norme des résidus de reprojection et à la surface avant (en rouge) et après (en bleu) optimisation.

5.4. Résultats quantitatifs

Le tableau 3 montre les différences entre les paramètres intrinsèques et extrinsèques provenant de la vérité terrain et ceux estimés à l'aide de la maximisation de la vraisemblance pour les séquences synthétiques.

Tableau 3. Erreur d'étalonnage sur les séquences synthétiques (* n'a pas convergé).

	S1	S2 - 1	S2 - 2	S3	S4
Paramètres intrinsèques caméra					
Δu_0 (px)	0.390	-1.651	0.000*	-2.664	0.398
Δv_0 (px)	1.729	0.355	0.000*	-1.472	1.356
Δf_x (px)	-0.055	-0.548	0.000*	-0.263	-0.827
Δf_y (px)	-0.510	1.420	0.000*	-0.403	0.418
$\Delta \xi$	7.10^{-4}	0.003	0.000*	0.003	0.002
Paramètres extrinsèques caméra					
Δt_x (m)	-0.010	0.027	-0.097*	-0.013	-0.008
Δt_y (m)	0.002	0.019	-0.100*	-0.010	-0.019
Δt_z (m)	0.039	-0.040	-0.099*	-0.029	0.009
$\Delta \omega_x$ (°)	-0.004	0.006	0.012	0.006	0.006
$\Delta \omega_y$ (°)	0.009	-0.109	0.073	1.10^{-4}	-0.002
$\Delta \omega_z$ (°)	0.101	0.193	0.006	-0.031	0.017
$ \theta $ (°)	0.071	0.157	0.052	0.022	0.0181
Paramètres extrinsèques télémètre					
Δt_x (m)	-0.448	0.124	-0.050	-0.021	0.024
Δt_y (m)	0.018	0.083	-0.051	-0.018	0.006
Δt_z (m)	-0.019	-0.031	-0.202	-0.034	0.011
$\Delta \omega_x$ (°)	0.782	0.495	-4.843	-1.567	-0.221
$\Delta \omega_y$ (°)	-4.051	0.473	-2.217	-0.769	-1.949
$\Delta \omega_z$ (°)	-0.102	0.182	-4.815	-0.088	-0.089
$ \theta $ (°)	2.919	0.501	5.067	1.236	1.963

5.5. Analyse et critique de la méthode

Tout premièrement, on remarque dans le tableau 2 les valeurs adoptées comme biais ajouté aux paramètres exacts des séquences synthétiques. Ceux-ci agissent autant sur les paramètres intrinsèques de la caméra que sur ses paramètres extrinsèques et ceux du télémètre. Les valeurs employées restituent ce qui aurait pu être obtenu à l'aide d'un étalonnage approximatif du système à l'aide d'un instrument de mesure peu fiable. Les erreurs de translation des paramètres extrinsèques ont une moyenne de 5cm et peuvent atteindre 20cm tandis que les erreurs commises sur les rotations seront bornées à 5 degrés.

La figure 7 illustre les résultats obtenus sur une petite séquence d'étalonnage. On remarque en (d) et (e) les histogrammes des erreurs de projections et des erreurs à la

surface. En rouge, la dispersion avant optimisation et en bleu, la réparation après optimisation. Ces deux distributions sont représentatives de la réussite ou de l'échec de l'optimisation et par conséquent, de l'étalonnage. En effet, plus les résidus sont centrés autour de l'origine et plus leur valeur est faible donc la vraisemblance est maximisée. On constate sur cette première séquence que les résidus liés à la vision ont été rassemblés alors qu'ils étaient étalés sur une plage de 10 *px* d'amplitude. De la même manière, les erreurs liées à la surface télémétrique ont vu leurs valeurs moyennes être divisées par quatre. De ce fait, l'optimisation a convergé.

La séquence *S2*, plus longue, a subi deux biais d'amplitudes différentes (*S2 - 1* et *S2 - 2* dans le tableau 2). Pour le premier jeu de paramètre la convergence a pu être obtenue, ce que l'on constate sur (*d*) et (*e*) de la figure 8 car les résidus ont été rassemblés similairement à la première séquence. Toutefois, le second jeu de paramètres présente une erreur importante des paramètres extrinsèques du télémètre et l'algorithme d'optimisation ne converge pas efficacement ((*f*) et (*g*)), il tombe dans un minimum local. En effet, une grande variation d'angle du télémètre induit une rotation de ses mesures. Or, lorsque les mesures télémétriques sont très distantes, elles subissent un déplacement important qui déforme la surface télémétrique. Sachant que les résidus sont calculés grâce à la distance minimale entre un point 3D *P* et toutes les mesures télémétriques, il y a, dans ce cas, une grande probabilité que la mesure la plus proche soit erronée. Par exemple, si *P* est un point triangulé appartenant à une façade, et que les paramètres approximatifs entraînent une déformation telle qu'il soit proche d'un pylône alors le résidu obtenu tendra à faire converger *P* vers le pylône. L'algorithme minimise ces erreurs ce qui empêche la convergence vers le minimum global. Les séquences *S3* et *S4* ont été exploitées dans les figure 9 et 10 et présente une convergence efficace de la méthode de minimisation.

Le tableau 3 référence les erreurs d'étalonnage commises par l'algorithme appliqué aux séquences synthétiques. On remarque une bonne précision des résultats pour les paramètres extrinsèques et intrinsèques de la caméra. En effet, le nombre très important de contraintes et de primitives détectées dans les images conduit à une précision de moins d'un dixième de degré en rotation et de moins de deux centimètres en translation. De la même façon, les paramètres intrinsèques sont estimés avec précision comme le coefficient de distorsion ξ qui a été recouvré avec une erreur moyenne de 0.002. Toutefois, comme les paramètres extrinsèques du télémètre sont moins contraints, les erreurs sont plus importantes. On observe une erreur angulaire d'environ trois degrés et une erreur de translation avoisinant les trois centimètres. Cependant, la séquence *S4*, la plus longue de toute, permet une meilleure estimation des paramètres, car l'erreur en translation ne dépasse pas 2.4 *cm* et 2 degrés en rotation.

Nous avons également utilisé les séquences réelles *R1* et *R2* provenant du véhicule Stéréopolis et calculé l'histogramme des résidus. Contrairement aux précédentes expérimentations, nous ne disposons pas d'un étalonnage "vérité terrain" pour ce dispositif d'acquisition, de ce fait, nous avons pris comme étalonnage initial, celui effectué en interne par l'IGN. Le résultat de l'optimisation fait apparaître une réduction

des résidus ; ce qui peut s'interpréter comme une convergence de l'algorithme, et donc une amélioration de l'étalonnage initial.

Le domaine de convergence de notre méthode est limité par deux étapes critiques. En effet, si les paramètres initiaux de la caméra ne sont pas assez précis, l'étape de triangulation échoue et l'optimisation ne peut avoir lieu. De plus, si les paramètres initiaux du télémètre sont erronés, la surface télémétrique est trop distante des points triangulés et l'optimisation échoue dans un minimum local, ou l'ensemble des résidus est considéré comme aberrant. Pour une séquence de taille importante comme *S3* ou *S4* on observe des valeurs maximales de translation de $70cm$ et 4° pour les rotations.

6. Conclusion

Nous avons présenté dans cet article une méthode d'étalonnage automatique d'un système d'acquisition possédant une ou plusieurs caméras, une centrale inertielle et un capteur télémétrique de type Lidar 3D. Nous avons défini une formalisation permettant l'unification de tous les paramètres d'étalonnage, extrinsèques et intrinsèques au sein d'une unique fonction de vraisemblance. La minimisation de celle-ci par un algorithme d'optimisation fournit une estimation précise des paramètres d'étalonnage. Les expérimentations ont montré que cette méthode est performante et permet, pour un grand domaine de convergence, de restituer avec exactitude les paramètres recherchés. Dans le cadre de ces travaux nous avons supposé disposer des paramètres approximatifs permettant la triangulation, l'appariement de primitive puis l'optimisation. Une extension possible consisterait à utiliser des méthodes de "*Structure from motion*" telles que la décomposition de la matrice fondamentale, l'ajustement de faisceau incrémental afin de les déterminer, agrandissant d'autant le cadre d'utilisation de cette méthode.

Remerciements

Ce travail est financé dans le cadre des projets ANR Contenus et Interactions (CONTINT) 2010 : iSpace&Time.