

Ego-localisation de véhicules en milieu urbain par fusion stéréovision/GPS

Lijun WEI, Cindy CAPPELLE, Yassine RUICHEK

Université de Technologie de Belfort-Montbéliard (UTBM), Laboratoire Systèmes et Transports (SeT)
13 rue E. Thierry-Mieg, 90010 BELFORT Cedex, France
lijun.wei@utbm.fr, cindy.cappelle@utbm.fr, yassine.ruichek@utbm.fr

Résumé – Dans cet article, une méthode de reconstruction 3D de l’environnement et de localisation de véhicules utilisant un système de stéréovision et un récepteur GPS est présentée. La stéréovision permet de reconstruire en trois dimensions les amers naturels de la scène et de prédire la position (x, y) et le cap θ de la caméra. Les mesures GPS « cohérentes » (selon le test NIS - Normalized Innovation Square) sont intégrées pour corriger la pose du véhicule à l’aide d’un filtre de Kalman sans parfum (UKF - Unscented Kalman Filter). L’approche a été testée sur des données réelles acquises en milieu urbain par le véhicule électrique et instrumenté du laboratoire SeT. Les résultats obtenus montrent que la localisation par stéréovision permet de pallier les absences ou erreurs de GPS telles que les sauts de positionnement.

Abstract – In this article, a 3D environment reconstruction and vehicle localization method using a stereoscopic system and a GPS receiver is presented. The stereovision system permits to reconstruct the natural landmarks in three dimensional space, and to predict the camera position (x, y) and orientation θ . The “coherent” GPS measurements (according to the NIS test - Normalized Innovation Square) are integrated to correct the vehicle pose by an Unscented Kalman Filter (UKF). The proposed approach has been tested on real data acquired in urban environment by the electrical vehicle equipped by SeT laboratory. The results show that the stereovision based localization permits to correct the GPS absences and errors such as GPS jumps.

1 Introduction

La navigation autonome d’un véhicule requiert notamment une estimation très précise de la position et de l’orientation du véhicule en tout temps. La localisation absolue utilisant des systèmes satellitaires tels que le GPS est souvent utilisée à cette fin. Cependant, les signaux peuvent être réfléchis ou bloqués notamment dans les canyons urbains ou encore dans les tunnels. La visibilité satellitaire limitée ou le problème de multi-trajet réduisent alors la précision de la localisation voire rendent indisponible le service GPS. Les mesures GPS sont donc souvent fusionnées avec des données proprioceptives acquises par des centrales inertielle, des gyromètres ou encore des codeurs incrémentaux montés sur les roues du véhicule (odométrie) [1]. Dans cet article, l’odométrie visuelle est choisie, elle est utilisée pour la localisation du véhicule d’une part et la cartographie de l’environnement d’autre part.

Une alternative au GPS pour la localisation de véhicules en milieu urbain basée sur la stéréovision est présentée. Les deux aspects suivants sont étudiés : 1/ Le problème des objets mobiles : le mouvement 3D et le flot optique 2D sont calculés pour supprimer les points caractéristiques appartenant à des objets non-statique ou non-rigides, ensuite le mouvement de la caméra est estimé avec uniquement les amers considérés comme statiques. 2/ Un filtre de Kalman sans parfum (UKF, Unscented Kalman Filter) fusionne les informations de l’odométrie visuelle par stéréovision et les données GPS. Les me-

sure GPS corrigent (étape de correction de l’UKF) la trajectoire du véhicule et la cartographie 3D prédites par l’approche basée sur la vision. Lors de ce processus, uniquement les mesures GPS considérées comme fiables et cohérentes (selon un test NIS, Normalized Innovation Square) sont prises en compte pour la mise-à-jour de la prédiction. Autrement les positions GPS sont rejetées et seule la méthode basée sur la stéréovision est utilisée.

La suite du papier est structuré comme suit. La méthode proposée est décrite dans la section 2. Puis la section 3 présente les résultats expérimentaux obtenus avec des données réelles acquises par le véhicule électrique SeTCar du laboratoire. Des conclusions et des perspectives sont finalement proposées dans la section 4.

2 Approche proposée

2.1 Odométrie visuelle par stéréovision

Les principales étapes de l’algorithme d’odométrie visuelle (OV) [2, 3] sont les suivantes :

2.1.1 Détection et mise en correspondance des amers

Les points SURF [4] sont détectés dans les images gauche et droite et mis en correspondance en utilisant les descripteurs. Afin d’éliminer les faux appariements, les contraintes épipo-

laire, de disparité minimale et maximale, d'unicité ainsi que le score ZNCC (Zero-mean Normalized Cross Correlation) entre les voisinages des points sont utilisés pour vérifier les appariements trouvés avec les descripteurs SURF.

2.1.2 Reconstruction 3D des amers

Dès lors que le problème de mise en correspondance est résolu, la position 3D des points (dans le repère de la caméra gauche) est obtenue par triangulation. La matrice de covariance des points 3D reconstruits est également calculée.

La matrice Jacobienne Q' (3×4) des points 3D Q par rapport aux coordonnées $\{q_x, q_y\}$ et $\{q'_x, q'_y\}$ des pixels (dans l'image gauche et droite respectivement) est estimée en utilisant la méthode décrite dans [5]. Les incertitudes sur les positions des points caractéristiques $\{q_x, q_y\}$ et $\{q'_x, q'_y\}$ dans les images sont modélisées en considérant des bruits blancs Gaussiens de moyennes nulles et non-corrélés. La matrice de covariance 2×2 s'écrit alors :

$$H_{left} = \begin{bmatrix} \delta_x^2 & 0 \\ 0 & \delta_y^2 \end{bmatrix}, H_{right} = \begin{bmatrix} \delta'_x{}^2 & 0 \\ 0 & \delta'_y{}^2 \end{bmatrix} \quad (1)$$

avec $\delta_x^2, \delta'_x{}^2, \delta_y^2$ et $\delta'_y{}^2$ les variances des pixels de coordonnées q_x, q'_x, q_y et q'_y . La matrice de covariance de la paire stéréoscopique H_{pair} est une matrice 4×4 diagonale dont les éléments sur la diagonale sont $\{\delta_x^2, \delta'_x{}^2, \delta_y^2, \delta'_y{}^2\}$. La matrice de covariance des points 3D reconstruits est alors $p = Q' H_{pair} Q'^T$.

2.1.3 Sélection et mis-à-jour de la paire de référence

Les points caractéristiques 2D sont suivis sur la suite de la séquence d'images (algorithme de Kanade-Lucas-Tomasi [6]) jusqu'à ce que leur nombre soit inférieur à un seuil prédéfini. Dès lors que cette condition n'est plus respectée, la paire courante est choisie comme nouvelle paire d'images de référence.

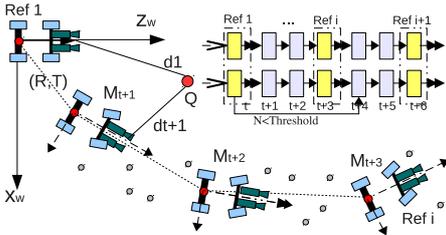


FIGURE 1 – Sélection et réinitialisation de la paire de référence

2.1.4 Elimination des outliers et des points mobiles

Lors de l'estimation du déplacement du véhicule, il est important d'éliminer les points correspondants à des objets mobiles ou non-rigides et les points mal appariés lors de l'étape de suivi.

Soient $\{Q_i^{t+1}\}$ et $\{Q_i^r\}$ les deux ensembles de points mis en correspondance ($i = 1 : N$, avec N le nombre de points appar-

iés). Afin d'estimer la transformation $[R, T]$ optimale qui transforme l'ensemble des points à l'instant précédent avec ceux à l'instant courant, l'erreur résiduelle doit être minimisée :

$$\varepsilon^2 = \sum_{i=1}^N \|Q_i^{t+1} - RQ_i^r - T\| \quad (2)$$

avec R une matrice de rotation 3×3 composée des angles de rotation de lacet, tangage et roulis (θ, ψ, ϕ) , $T = (T_x, T_y, T_z)$ le vecteur de translation.

Trois contraintes géométriques sont appliquées pour éliminer les erreurs de suivi : 1/ La distance SAD (Sum of Absolute Differences) entre les voisinages des points 2D suivis doit être inférieure à 500 ; 2/ Les points caractéristiques suivis doivent être dans le plan image courant ; 3/ La contrainte épipolaire entre les images gauche et droite doit être respectée.

Lorsque le véhicule se déplace, tous les objets statiques (considérés comme le "background") doivent avoir le même mouvement 3D par rapport au véhicule. Pour les points 3D $\{Q_i^{t+1}\}$ et $\{Q_i^r\}$ aux instants respectifs $t + 1$ et de *reference* ($i = 1 : M$, avec M le nombre de points), la vitesse de déplacement peut être obtenue comme suit [7] :

$$\widetilde{\Delta Q}_i = \begin{bmatrix} \widetilde{X}_i^r \\ \widetilde{Y}_i^r \\ \widetilde{Z}_i^r \end{bmatrix} = \begin{bmatrix} \tau_x \\ \tau_y \\ \tau_z \end{bmatrix} + \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \begin{bmatrix} X_i^r \\ Y_i^r \\ Z_i^r \end{bmatrix} \quad (3)$$

avec : $\omega = [\omega_x, \omega_y, \omega_z]^T$ la vitesse de rotation et $\tau = [\tau_x, \tau_y, \tau_z]^T$ la vitesse de translation. Les projections 2D des points Q_i^r dans le plan image gauche peuvent s'écrire : $\{x_i = f * X_i^r / Z_i^r, y_i = f * Y_i^r / Z_i^r\}$.

Le flot optique 2D $[u_i, v_i]^T$ est calculé par dérivation de $\{x_i, y_i\}$. Pour chaque trois couples de points suivis, l'équation du flot optique 2D peut être résolue afin d'obtenir le vecteur de déplacement. L'erreur moyenne $\bar{\epsilon}$ et la variance σ^2 des erreurs sont utilisées pour éliminer les points suivis dont la variance est supérieure à $2\sigma^2$. En utilisant l'algorithme de RANSAC [8], l'ensemble des points contenant le plus grand nombre d'inliers est utilisé pour estimer le déplacement du véhicule (eq.2 de la solution d'Arun [9]). La position de la première caméra est considérée comme l'origine du repère global W . Les positions dans le repère global de la caméra à chaque instant et des amers 3D peuvent ainsi être obtenus à partir des paires stéréoscopiques de référence.

2.2 Fusion OV/GPS

Les informations de l'odométrie visuelle et les mesures GPS sont fusionnées à l'aide d'un filtre de Kalman sans parfum (UKF, Unscented Kalman Filter) [10]. L'UKF est utilisé pour l'estimation des systèmes non-linéaires. Alors que le filtre de Kalman étendu (EKF) linéarise la fonction non-linéaire autour de l'estimation courante, l'UKF utilise la transformation sans parfum dont le principe est d'approximer la densité de probabilité gaussienne de la variable par un nombre minimum de points appelés « sigma-points ».

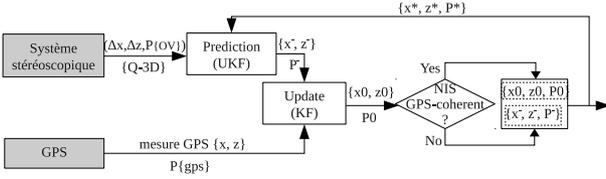


FIGURE 2 – Description de la méthode de localisation

Le déplacement $(\Delta x, \Delta z)$ entre deux instants t et $t + 1$ et la variation de cap ω obtenus par l'odométrie visuelle sont utilisés dans l'étape de prédiction alors que les mesures GPS considérées comme cohérentes selon le test NIS (Normalized Innovation Square) sont utilisées dans l'étape de correction de l'UKF. Notons que le modèle de mesure étant linéaire, les équations classiques du filtre de Kalman sont utilisées dans l'étape de correction comme indiqué dans la figure 2.

2.2.1 Prédiction par odométrie visuelle

La transformation sans parfum choisit un nombre fixé de sigma-points de la distribution originale, puis propage ces sigma-points à travers la fonction non-linéaire et enfin estime la moyenne et la covariance. Soit $\{M_r^t\}$ le vecteur de déplacement qui transforme la position du véhicule de l'instant *reference* à l'instant courant t .

$$M_r^t = f(Q_i^r, Q_i^t) = [R|T], Q_i^t = [R|T] \begin{bmatrix} Q_i^r & 1 \end{bmatrix}^T \quad (4)$$

avec $i = 1, \dots, N$, où N est le nombre d'amers utilisés pour estimer le déplacement. On suppose que chaque amer Q_i^r suit une distribution Gaussienne de moyenne Q_i^r et de covariance p_i^r , $Q_i^r \sim N(Q_i^r, p_i^r)$ et que les amers Q_i^t suivent une distribution Gaussienne de moyenne Q_i^t et de covariance p_i^t , $Q_i^t \sim N(Q_i^t, p_i^t)$. Par propagation des erreurs [11], M_r^t doit suivre une distribution Gaussienne de moyenne $\{d_r^t, \theta_r^t\}$ et de covariance P_r^t estimée comme suit :

$$P_r^{t-1} = S^T p^{-1} S = \sum_i (S_i^T p_i^{-1} S_i) \quad (5)$$

Soit S la matrice Jacobienne (eq. 4) par rapport à d et θ , P est l'incertitude sur les amers, S_i est la matrice Jacobienne du i -th point par rapport à d et θ , et p_i est l'incertitude de l'amer i à l'instant *reference* et à t , $p_i = p_i^r + p_i^t$. L'ensemble des $(2n + 1)$ sigma-points X_{t-1} et leurs poids associés $W_m^{(i)}$ et $W_c^{(i)}$ sont d'abord calculés. Enfin, la moyenne et la covariance de la prédiction sont :

$$m_t^- = \hat{X}(t|t-1) [W_m^{(0)} \dots W_m^{(2n)}]^T \quad (6)$$

$$P_t^- = \left(\sum_{i=1}^{2n} W_c^{(i)} [\hat{X}(t|t-1)^{(i)} - m_t^-] [\hat{X}(t|t-1)^{(i)} - m_t^-]^T \right) + P_r^t \quad (7)$$

2.2.2 Correction par GPS

Les positions GPS sont converties du système WGS84 au repère cartésien puis translatées et rotées dans le repère global

(de l'odométrie visuelle). L'observation GPS (linéaire) s_t peut s'écrire comme suit :

$$s_t = \begin{bmatrix} x_t^{gps} \\ z_t^{gps} \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \tilde{x}_t \\ \tilde{z}_t \\ \theta_t \end{pmatrix} + Q_t^{gps} \quad (8)$$

L'observation GPS (x_t^{gps}, z_t^{gps}) est fournie par la mesure GPS $(\tilde{x}_t, \tilde{z}_t)$, et Q_t^{gps} est le bruit de mesure estimé en utilisant la trame NMEA GST [12].

L'innovation est donnée par $v_t = s_t - C m_t^-$, et le gain de Kalman est :

$$K_t = P_t^- C^T (C P_t^- C^T + Q_t^{gps})^{-1} \quad (9)$$

avec $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, la moyenne et la covariance corrigées sont données par :

$$m_t = m_t^- + K_t * v_t \quad (10)$$

$$P_t = P_t^- - K_t (C P_t^- C^T + Q_t^{gps}) K_t^T \quad (11)$$

Avant d'utiliser une mesure GPS pour corriger la prédiction, la cohérence de cette mesure est vérifiée par un test NIS (Normalized Innovation Squared) [13] avec une distribution χ^2 . La distance d_m entre l'observation GPS et l'état prédit est calculée en utilisant le vecteur d'innovation v_t comme suit :

$$d_m = v_t^T (I_v)^{-1} v_t \quad (12)$$

avec I_v la covariance de l'innovation. Si la distance d_m est inférieure à un seuil (déterminé par la loi χ^2), la mesure est considérée comme cohérente et est alors utilisée pour corriger la prédiction.

3 Résultats expérimentaux

L'approche a été testée sur des données acquises grâce à la plateforme expérimentale du SeT composée d'un véhicule électrique GEM instrumenté entre autres d'un système de stéréovision Bumblebee XB3 et d'un GPS-RTK (Fig. 3). La trajectoire de référence mesurée par le GPS-RTK est d'une longueur de 674.5m (Tab. 1). En utilisant l'odométrie visuelle seule, la distance parcourue estimée est de 650.25m (erreur de 3.60%). Avec la méthode d'intégration GPS-RTK/OV, la distance parcourue estimée est de 685.58m (erreur de 1.64%). On constate que l'approche vision est précise à court mais dérive à plus long terme. La localisation par stéréovision est néanmoins utile pour pallier les pannes ou mauvaises précisions GPS de courtes durées. Dès lors que la mesure GPS est précise, elle permet de corriger la dérive de l'estimation par la stéréovision.

Dans la figure 4(a), des erreurs de positionnement GPS sont relevées (cercle de droite sur la figure). Dans cette situation, la méthode de localisation par vision présentée précédemment permet seule d'assurer le service de localisation du véhicule et de pallier la défaillance du GPS. Dans le cercle de gauche, un masquage d'environ 100 mesures GPS a été simulé. De nouveau, l'odométrie visuelle permet d'assurer l'estimation de la position durant cette période de panne GPS (simulée).



FIGURE 3 – Gauche : en rouge la trajectoire et en vert les amers 3D reconstruits par OV ; Droite (haut) : paire stéréoscopique et points appariés ; Droite (bas) : véhicule instrumenté

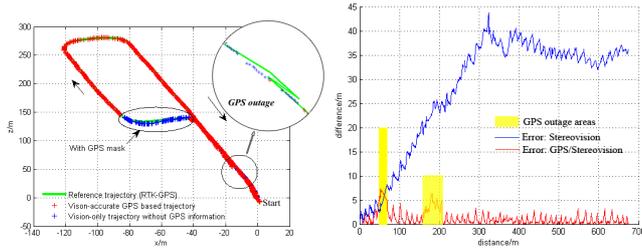


FIGURE 4 – Résultat de la méthode de localisation proposée (gauche : trajectoire, droite : erreur de positionnement)

Les erreurs des méthodes de localisation basées sur la vision seule et sur le couplage GPS/Vision sont comparées sur la figure 4(b). Les zones indiquées en jaune correspondent aux périodes de défaillance GPS.

Méthode	GPS-RTK Dis.	Dis. estimée	Erreur Moyenne	Ecart-type
OV	674.50	650.25	3.60 %	11.71
GPS-RTK/OV	674.50	685.58	1.64 %	0.80

4 Conclusion

Dans cet article, nous avons présenté une méthode de localisation de véhicule et de reconstruction 3D de l'environnement utilisant la stéréovision et un récepteur GPS. La stéréovision sert à la reconstruction 3D et à l'estimation du déplacement. Les mesures GPS sont utilisées pour corriger la prédiction de la pose du véhicule en utilisant un filtre de Kalman sans parfum. L'approche proposée a été testée sur des données acquises en environnement urbain. Comme toute méthode odométrique, la localisation par odométrie visuelle utilisant la stéréovision dérive à long terme. En effet, les erreurs d'estimation s'accroissent avec le temps. Néanmoins, elle permet de pallier efficacement l'indisponibilité du GPS (nombre de satellites visibles insuffisant) ou son manque de précision (dû par exemple au multi-trajet des signaux) à court terme. Les perspectives à court terme de ce travail sont d'intégrer d'autres capteurs tels qu'une centrale inertielle, un télémètre laser. L'estimation du

mouvement de la caméra en considérant les 6 degrés de liberté est également à l'étude afin de lever l'hypothèse d'un sol plan.

Références

- [1] <http://canbuskit.com/what.php>.
- [2] R.I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. pages 117–121, Cambridge University Press, ISBN : 0521540518, second edition, 2004.
- [3] D. Nister, O. Naroditsky, and J. Bergen. Visual odometry. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)*, 2004.
- [4] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3) :346–359, 2008.
- [5] Y. Cheng, M. Maimone and L. Matthies. Visual Odometry on the Mars Exploration Rovers. *IEEE Robotics and Automation Magazine*, 13(2) :54–62. 2006.
- [6] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *IJCAI'81*, 674–679, 1981.
- [7] M. Shih and B. Fu. Robust Moving Object Detection on Moving Platforms. *Advances in Image and Video Technology Lecture Notes in Computer Science*, 1319 :591–600, 2006.
- [8] M. A. Fischler and R. C. Bolles. Random Sample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24 :381–395, 1981.
- [9] K.S. Arun, T.S. Huang and S.D. Blostein. Least-squares fitting of two 3-d point sets. *IEEE Trans. Pattern Anal.*, 9(5) :698–700, 1987.
- [10] S.J. Julier and J.K. Uhlmann. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3) : 401–422, 2004.
- [11] F. A. Moreno, J. L. Blanco, J. Gonzalez. An efficient closed-form solution to probabilistic 6D visual odometry for a stereo camera. *ACIVS*, 932–942, 2007.
- [12] M.E. E. Najjar and P. Bonnifait. A road-matching method for precise vehicle localization using belief theory and kalman filtering. *Autonomous Robots*, 19(2) :173–191, 2005.
- [13] M. Jabbour, Ph. Bonnifait, and V. Cherfaoui. Management of Landmarks in a GIS for an Enhanced Localisation in Urban Areas. *IEEE Intelligent Vehicle Symposium*, Tokyo, Japan, June, 2006.