

Trajectographie des véhicules en vision monoculaire

Yann GOYAT¹, Thierry CHATEAU², Laurent MALATERRE², Laurent TRASSOUDAIN²

¹LCPC

route de Bouaye, 44341 Bouguenais, France

²LASMEA, université Blaise Pascal

24, av. des landais, 63177, Aubière cedex, France

(1)yann.goyat@lcpc.fr, (2)prenom.nom@lasmea.univ-bpclermont.fr

Résumé – Cet article présente une méthode de suivi précis de trajectoires de véhicules légers par vision monoculaire. Elle s'insère dans un projet de recherche SARI destiné à mieux comprendre les comportements des conducteurs pour aider les gestionnaires à sécuriser leur réseau routier. Cette article focalise sur les contributions développées dans la méthode de suivi. Ce dernier est réalisé par un filtre temporel stochastique non linéaire (filtre à particules). Nous proposons d'utiliser un modèle de véhicule simple comme état du filtre, pour contraindre l'évolution des particules dans un espace réaliste. De plus, nous introduisons une fonction de vraisemblance basée sur la comparaison entre la projection d'un modèle 3D du véhicule dans la caméra et une carte binaire fond-forme issue des observations. Un algorithme de calcul rapide de cette vraisemblance est présenté. Les performances de la méthode proposée ont été quantifiées à l'aide d'une campagne d'essais sur site réel pour laquelle un GPS cinématique (précis au centimètre) a servi de vérité terrain.

Abstract – This article presents a probabilistic method of accurate vehicles tracking by monocular vision. The main contributions of this work are the development of a background subtraction algorithm, the use of a kinematic model close to reality, known as bicycle, and the definition of an observation function based on the reprojection of simplified models 3D of vehicles. Many trajectories of a vehicle equipped with a kinematic GPS were measured on real sites, with a video system developed for this project. These grounds truths make it possible to validate in experiments the result obtained with this algorithm. The final goal of this work is intended for better understanding the drivers behaviors to help the road managers to make safe their network.

1 Introduction

Cet article présente une méthode de suivi de trajectoires de véhicules par vision monoculaire. Cette dernière s'appuie sur la littérature abondante présente autour du suivi visuel. Le but de ce système est d'estimer, de manière précise, la trajectoire d'un véhicule dans un virage. Pour cela, un capteur, composé de trois caméras, sans recouvrement, qui observent respectivement l'entrée, le milieu et la sortie de virage a été développé. Actuellement, de nombreux systèmes de mesure de trajectoires sont déjà industrialisés et installés au-dessus de nos routes. Malheureusement, ces systèmes, bien qu'utilisant les contraintes d'évolution fortes offertes par le code de la route (conduite à droite, un seul sens de circulation, . . .), ont été développés pour couvrir un grand linéaire et donc aboutissent à une précision relativement faible (trajectoire moyenne, comp- tage, . . .).

La méthode de suivi développée prend en compte les spécificités de l'application. Les caméras étant statiques, une extraction de fond est effectuée afin de faciliter la phase de suivi. La méthode de suivi utilisée est composée d'un filtre à particules classique, pour lequel nous proposons d'injecter les contraintes liées à la non holonomie des trajectoires (véhicules) à travers un vecteur d'état de type modèle bicyclette. Enfin, l'observation est réalisée à travers une fonction de vraisemblance originale, utilisant un modèle 3D simplifié du véhicule. Nous proposons une méthode rapide de calcul de cette vraisemblance basée sur l'approximation de sa projection dans l'image par son enve-

loppe convexe et l'utilisation d'une image intégrale.

Cet article est divisé en deux parties principales. La méthode développée est décrite dans la première partie, en insistant sur chacune des deux contributions : le modèle d'état non holo- nome utilisé et la fonction de vraisemblance développée. La deuxième partie présente la campagne de mesure menée afin de quantifier la précision et la robustesse de la méthode. Un grand nombre de trajectoires ont été estimées. Un GPS cinématique a été utilisé pour fournir la vérité terrain associée à chaque trajectoire.

2 Méthode

Cette section expose la méthode proposée pour le suivi de trajectoires par vision. Après une partie résumant le principe de l'approche retenue, le filtre temporel, coeur du système, est présenté.

2.1 Principe de la méthode

L'étude et la mise en oeuvre d'un algorithme de suivi précis de trajectoires de véhicules par vision est le principal but du système. Deux modes de fonctionnement peuvent être envisagés. Le premier repose sur un traitement en ligne des flux vidéos (trois caméras VGA¹ à la cadence de 30 images par seconde) et nécessite un algorithme dont la cadence d'exécution

¹480 lignes par 640 colonnes

est compatible avec celle des capteurs. Le deuxième mode de fonctionnement est basé sur un enregistrement des flux vidéos sur disque et un traitement différé des données. Dans ce cas, les contraintes de temps d'exécution ne sont plus présentes. Nous avons opté pour ce mode de fonctionnement, tout en recherchant à optimiser les temps de calcul des différentes briques de la méthode développée. Ainsi, il sera possible, moyennant une architecture adaptée, de migrer par la suite vers le premier mode de fonctionnement.

L'estimation de la trajectoire peut être réalisée, soit par un recalage de la totalité des données sur un modèle global de trajectoire, soit par un filtre temporel. Nous avons choisi la deuxième solution car le recalage global ne peut pas être réalisé dans une stratégie de traitement en ligne des données.

Le coeur de la méthode proposée (voir figure 1) est composé d'un filtre récursif, formalisé de manière stochastique (filtre à particules). L'état du véhicule est représenté par son modèle bicyclette. La fonction de prédiction fait l'hypothèse d'une vitesse et d'un angle au volant constant. Un modèle géométrique 3D du véhicule, projeté dans l'image est confronté aux données images décrites par une carte binaire d'appartenance à la forme.

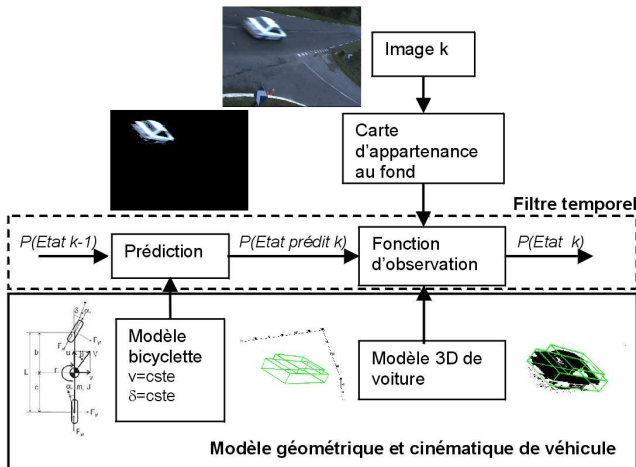


FIG. 1 – Synoptique de la méthode proposée

2.2 Extraction de fond

Un grand nombre de méthodes [1] de suivi d'objets dans une scène statique utilisent une étape d'extraction de fond. Ces dernières consistent, dans un cas binaire, à associer à chaque pixel de l'image une étiquette « Fond/Forme ». Dans [2], nous proposons d'utiliser une approche non paramétrique, basée sur la modélisation discrète de la densité de probabilité d'appartenance de chaque pixel au fond, nommé *Vu-mètre*. Cette méthode permet de construire une image binaire I_f , dans laquelle les pixels « Formes » ont la valeur 1 et les pixels « Fonds » la valeur 0.

2.3 Filtrage particulière

La trajectoire du véhicule est estimée de façon récursive par un filtre non linéaire. La résolution de ce dernier s'effectue par une méthode stochastique courante en vision : un filtre à particules. Ce choix est guidé par la nature non linéaire du sys-

tème. Le filtrage particulière [3, 4] est basé sur l'estimation de la densité de probabilité a posteriori $p(\mathbf{X}_t | \mathbf{Z}_{0:t})$ de l'état \mathbf{X}_t conditionné par l'historique des mesures $\mathbf{Z}_{0:t}$, à l'instant t , par un ensemble de N particules pondérées $\{(\mathbf{X}_t^n, \pi_t^n)\}_{n=1}^N$ avec des poids associés (π_t^n) .

La mise en oeuvre d'un filtre à particules nécessite la définition de trois modèles : (1) un modèle d'état qui définit les caractéristiques de localisation de l'objet à suivre ; (2) un modèle d'évolution qui définit l'état d'un objet à un instant donné en fonction de l'état à l'instant précédent ; (3) un modèle d'observation qui définit une mesure entre une hypothèse sur l'état et les observations. La suite de cette section est consacrée à la définition de ces trois modèles.

2.3.1 Modèle d'état

Le modèle d'état définissant les caractéristiques de la trajectoire à suivre doit intégrer les contraintes liées à la cinématique du véhicule. Nous proposons d'utiliser un modèle de type bicyclette et de rendre compte de nombreuses propriétés du comportement et de la stabilité. Ce modèle de type bicyclette, comporte deux degrés de liberté :

- β : angle de dérive du véhicule.
- r ou ψ : vitesse de lacet,

Soient (x, y) les coordonnées d'un point du véhicule dans un référentiel absolu. Les relations cinématiques qui régissent ce modèle peuvent donc s'écrire : $\dot{x} = v \cdot \cos \beta$, $\dot{y} = v \cdot \sin \beta$, $\dot{\psi} = \frac{v}{L} \cdot \tan \delta$.

Le vecteur d'état du système peut s'écrire :

$$\mathbf{X}_t \doteq (\mathbf{P}_t, \beta_t, \delta_t, v_t)^t \quad (1)$$

avec $\mathbf{P}_t \doteq (x_t, y_t)$ représentant la position du véhicule dans un repère monde plan, β_t l'angle de dérive au centre de gravité, δ_t l'angle de braquage de la roue avant et v_t la vitesse du véhicule dans le référentiel Lambert II étendu, selon l'axe du véhicule.

2.3.2 Modèle de prédiction

Le modèle cinématique de type bicyclette appliqué à chaque particule évolue non linéairement de la manière suivante :

$$\begin{aligned} x_{t+1} &= x_t + T \cdot v_t \cdot \cos(\beta_t) \\ y_{t+1} &= y_t + T \cdot v_t \cdot \sin(\beta_t) \\ \beta_{t+1} &= \beta_t + T \cdot \frac{v}{L} \cdot \tan \delta_t \\ \delta_{t+1} &= \delta_t + T \cdot b_\delta \\ v_{t+1} &= v_t + T \cdot b_a \end{aligned} \quad (2)$$

avec $b_\delta \sim \mathcal{N}(0, \sigma_\delta)$ et $b_a \sim \mathcal{N}(0, \sigma_a)$. Les deux termes δ et a sont distribués de manière aléatoire (selon une gaussienne). Les termes σ_δ et σ_a peuvent être estimés comme étant respectivement les plages d'écarts de vitesse d'angle de braquage de la roue avant et d'accélération, effectués par un véhicule standard, pendant le temps d'échantillonnage T .

2.3.3 Modèle d'observation

La fonction de vraisemblance que nous proposons utilise un modèle géométrique tridimensionnel simplifié du véhicule à suivre, représenté sur la figure 2. Il s'agit d'un modèle constitué de deux parallélépipèdes emboîtés. Dans un cas général, le modèle peut être plus complexe et constitué de P_M parallélépipèdes. Soit $\mathcal{M}^{(R_0)} = \{M_i^{(R_0)}\}_{i=1, \dots, N_M}$ l'ensemble des

sommets des cubes du modèle ($N_M = 8 \times C_M$) exprimé dans un repère associé au modèle R_0 . Ce repère est choisi afin que les 3 axes soient dans la même direction que ceux du repère associé à la scène R_w . Pour une particule \mathbf{X}_t^n donnée, le calcul de la vraisemblance (poids) est défini par la différence entre le nombre de points formes et le nombre de points fonds situés à l'intérieur de la projection du modèle du véhicule dans l'image.

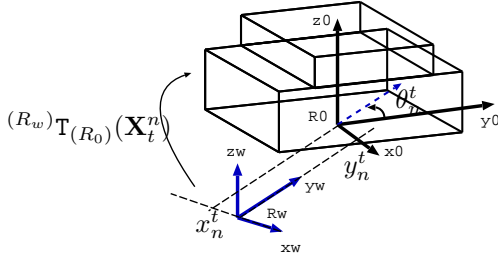


FIG. 2 – Exemple de modèle géométrique tridimensionnel simple utilisé pour un véhicule. Il est composé de deux cubes. Le repère associé au cube et le repère associé à la scène sont liés par une translation pure. Le plan (Oxy) du repère monde et les axes associés sont confondus avec le repère GPS.

Ce calcul effectué pour chaque particule est consommateur de temps processeur. Nous proposons une méthode de calcul rapide de la vraisemblance basée sur l'approximation de la projection du modèle 3D dans l'image par son enveloppe convexe. Chaque point du modèle se projette dans l'image par l'équation suivante :

$$\tilde{\mathbf{m}}_i \propto C \cdot (R_w)T_{(R_0)}(\mathbf{X}_t^n) \cdot \tilde{\mathbf{M}}_i^{(R_0)} \quad (3)$$

avec $\tilde{\mathbf{M}}$ coordonnées homogènes associées au point M ; C matrice de projection caméra, et $(R_w)T_{(R_0)}(\mathbf{X}_t^n)$ matrice de transformation homogène entre le repère monde et le repère associé au modèle 3D (cf. figure 2).

L'ensemble $\mathcal{M}^{(R_0)} = \{\mathbf{m}_i\}_{i=1, \dots, N_M}$ est alors construit à partir de la projection des points du modèle 3D dans l'image.

Soit $\mathcal{E}(\mathcal{M}^{(R_0)}; \mathbf{X}_t^n) = \{e_i\}_{i=1, \dots, N}$ ($e_i = (x_i^e, y_i^e)$, coordonnées de e_i dans le plan image), la liste des points de l'enveloppe convexe². Nous définissons $\mathcal{E}_t^n \doteq \mathcal{E}(\mathcal{M}^{(R_0)}; \mathbf{X}_t^n)$ afin d'alléger les notations. Le calcul de la vraisemblance (différence entre le nombre de points « Formes » et le nombre de points « Fonds » situés à l'intérieur de l'enveloppe convexe) peut être effectué de manière efficace en utilisant une image intégrale par ligne :

$$I_I(x, y) = \sum_{i=1}^x I_f(i, y) \quad (4)$$

où I_f est l'image binaire fond-forme calculée dans la section 2.2. Les points e_i sont classés par couples de même ordonnée, tels que :

$$\mathcal{E}_t^n = \left\{ \begin{array}{l} (x_1^e, y^e), (x_2^e, y^e), \\ (x_3^e, y^e + 1), (x_4^e, y^e + 1), \dots \\ (x_{N-1}^e, y^e + N/2), (x_N^e, y^e + N/2) \end{array} \right\} \quad (5)$$

Le codage de l'enveloppe convexe dans l'ensemble \mathcal{E}_t^n nécessite une discrétisation de la forme selon les lignes de l'image.

²Nous ne développons pas le calcul de l'enveloppe convexe dans cet article, qui est effectué à l'aide d'un algorithme dont la complexité est en $O(n \cdot \log n)$

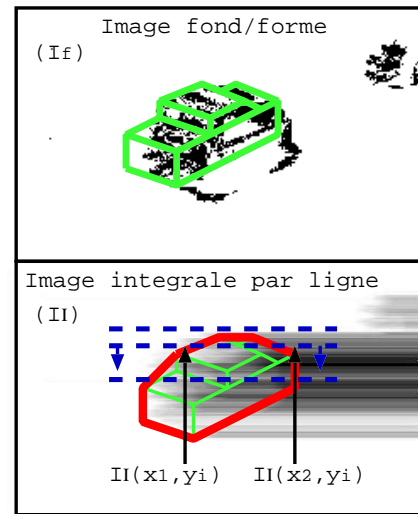
De plus, un intérêt particulier doit être porté au codage des extrémités hautes et basses. Par contre, il n'est pas nécessaire de trier les points situés sur la même ligne. Une mesure d'adéquation associée à une enveloppe convexe est calculée dans l'image intégrale par la relation suivante :

$$a(\mathcal{E}_t^n, I_I) = \sum_{i=1}^{N/2} l_i \quad (6)$$

avec

$$l_i = [2 \cdot (I_I(e_{2i}) - I_I(e_{2i-1})) - (x_{2i}^e - x_{2i-1}^e)] \quad (7)$$

l_i représente la contribution d'une ligne de l'image à la mesure d'adéquation. La figure 3 décrit le principe de la méthode de calcul de vraisemblance par l'image intégrale. Un balayage par ligne est effectué.



$$l_i = 2 \cdot (I_I(x_1, y_i) - I_I(x_2, y_i)) - (x_2 - x_1)$$

FIG. 3 – Illustration du calcul de la vraisemblance. Le modèle 3D du véhicule (en vert) est approximée par son enveloppe convexe (en rouge). La valeur de la vraisemblance est donnée par la différence entre le nombre de points « Formes » et le nombre de points « Fonds » situés dans l'enveloppe convexe. ($\sum l_i$)

Le poids associé à chaque particule est directement lié à $a(\mathcal{E}_t^n, I_I)$:

$$\pi_t^n \propto p(\mathbf{Z}_t | \mathbf{X}_t = \mathbf{X}_t^n) \doteq \max(0, a(\mathcal{E}_t^n, I_I)) \quad (8)$$

L'initialisation du filtre à particules est effectuée de manière automatique par une méthode de détection d'amas décrite dans [5].

3 Validation expérimentale

Cette section présente les expérimentations effectuées pour valider la méthode de suivi de trajectoires.

Le système vidéo développé dans le cadre du projet SARI (www.sari.prd.fr) permet d'enregistrer en continu flux vidéos couleur 640x480, cadencés à 30 i/s, pendant une dizaine d'heures (en mode non compressé).

Pour évaluer la précision des mesures, un véhicule de type Peugeot 406, équipé d'un GPS cinématique, précis au centimètre près, a été utilisé. La position de l'antenne GPS étant

connue, celle-ci a pu être prise en compte dans chacun des modèles géométriques. Les trajectoires calculées ont ensuite été comparées aux trajectoires estimées par l'approche vision, en calculant la distance entre chaque point calculé et la droite passant par les deux points GPS les plus proches. Pour chaque vitesse, au moins cinq passages de véhicules ont été effectués dans un virage, permettant de faire une statistique très grossière sur les mesures relevées.

Trois études métrologiques ont été menées afin de déterminer la précision du système en fonction de la variation : (1) de trois modèles de véhicules différents, (2) du nombre de particules utilisé dans le filtre et (3) du modèle tridimensionnel de véhicule utilisé. L'ordre de grandeur de l'erreur moyenne, sur un linéaire de 100 m, entre la trajectoire calculée et la trajectoire mesurée est d'environ 30 cm.

Les résultats de la première étude montrent qu'il semble plus intéressant de sous-estimer les dimensions du modèle 3D. Le tableau 1 montre les résultats de la deuxième étude. Il semble que l'augmentation du nombre de particules atteint une asymptote au niveau de la justesse et de la fidélité (moyenne et écart type). A 50 particules, un passage n'a pas pu être suivi et les deux indicateurs sont relativement forts. Le temps de calcul étant proportionnel au nombre de particules, il est intéressant d'identifier cette valeur à utiliser pour le filtrage et pour laquelle la précision atteint son maximum. Ici, la valeur de 100 particules semble être un bon compromis.

TAB. 1 – Erreur d'estimation de la trajectoire en fonction du nombre de particules : le modèle utilisé est le modèle deux cubes de type 406.

Vitesse de passage (km/h)	Nombre de particules					
	50		100		200	
	moyenne	écart-type	moyenne	écart-type	moyenne	écart-type
50	0.43	0.17	0.39	0.11	0.36	0.16
60	0.38	0.18	0.37	0.14	0.38	0.17
70	0.40	0.20	0.32	0.16	0.39	0.39
80	0.29	0.18	0.26	0.11	0.30	0.30
90	0.44	0.17	0.37	0.15	0.38	0.22
100	0.71	0.63	0.30	0.15	0.25	0.12
110	0.39	0.18	0.40	0.20	0.34	0.18
moyenne	0.43	0.24	0.34	0.15	0.34	0.22

Enfin, la dernière étude a montré qu'en utilisant le modèle le plus simple possible (1 seul cube), la précision est aussi bonne qu'en utilisant un modèle composé de deux cubes (cf. tableau 2). Cela peut s'expliquer par la qualité médiocre de l'image binaire fond-forme extraite à partir des données brutes, qui est la principale cause de détérioration de la précision.

TAB. 2 – Erreur d'estimation de la trajectoire en fonction de la complexité du modèle 3D choisi.

Vitesse de passage (km/h)	Modèles géométriques			
	2 cubes		1 cube	
	moyenne	écart-type	moyenne	écart-type
50	0.39	0.11	0.37	0.16
60	0.37	0.14	0.38	0.17
70	0.32	0.16	0.39	0.18
80	0.26	0.11	0.28	0.19
90	0.37	0.15	0.35	0.19
100	0.30	0.15	0.26	0.15
110	0.40	0.20	0.34	0.14
moyenne	0.34	0.15	0.34	0.17

4 Conclusion - Perspectives

Ce travail a permis de démontrer que l'utilisation de la vision peut permettre de mesurer précisément la trajectoire des véhicules, en conditions réelles. Ce constat est impulsé par le fait que les conditions de mesure sont fortement contraintes, puisque le sens de circulation, la gamme de vitesse et la forme générale de l'objet à suivre sont assez bien connus à l'avance. Toutes ses contraintes sont intégrées de manière naturelle dans la méthode proposée.

Certains verrous ont été levés, tels les modèles d'état et d'évolution choisis pour le filtre, permettant de se rapprocher de la cinématique réelle des véhicules.

Les prochains travaux se focaliseront sur la robustesse du suivi, en utilisant, en complément à la vision, un autre système optique : un télémètre laser à balayage. Ce capteur permettra d'une part de prendre le relais lorsque la vision n'est plus adaptée, comme par exemple la nuit, et d'autre part de se substituer à la vision si les mesures ne sont plus "fiabiles". Cette phase de l'étude fera appel aux notions de fusion de données.

Deux nouveaux sites réels vont être instrumentés en 2007 et permettront de valider ces prochains travaux.

Références

- [1] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, Aug. 2000.
- [2] Y. Goyat, T. Chateau, L. Malaterre, and L. Trassoudaine, "Vehicle trajectories evaluation by static video sensors," *9th International IEEE Conference on Intelligent Transportation Systems*, Sept. 2006.
- [3] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002. [Online]. Available : cite-seer.nj.nec.com/arulampalam02tutorial.html
- [4] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *IJCV : International Journal of Computer Vision*, vol. 29, no. 1, pp. 5–28, 1998.
- [5] Y. Goyat, T. Chateau, L. Malaterre, and L. Trassoudaine, "Estimation précise de la trajectoire d'un véhicule par vision monoculaire," *ORASIS*, Juin 2007.