



HAL
open science

Suivi et Catégorisation Visuels en Temps Réel d'un Nombre Variable d'Objets : application au suivi de véhicules

François Bardet, Thierry Chateau, Datta Ramadasan

► **To cite this version:**

François Bardet, Thierry Chateau, Datta Ramadasan. Suivi et Catégorisation Visuels en Temps Réel d'un Nombre Variable d'Objets : application au suivi de véhicules. ORASIS'09 - Congrès des jeunes chercheurs en vision par ordinateur, 2009, Trégastel, France, France. inria-00404619

HAL Id: inria-00404619

<https://inria.hal.science/inria-00404619>

Submitted on 16 Jul 2009

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.

Suivi et Catégorisation Visuels en Temps Réel d'un Nombre Variable d'Objets : application au suivi de véhicules

François Bardet¹

Thierry Chateau¹

Datta Ramadasan¹

¹ LASMEA, Université Blaise Pascal
24, avenue des Landais
63177 -Aubière cedex -France
{bardet,chateau,ramadasan}@lasmea.univ-bpclermont.fr

Résumé

Ce papier expose un système automatique de suivi et catégorisation en temps réel d'un nombre variable d'objets génériques. La configuration multi-objet est estimée par un Filtre Particulaire par Chaîne de Markov (Markov Chain Monte-Carlo Particle Filter ou MCMC PF). Ce système a précédemment été validé sur du suivi de piétons, et nous présentons ici des applications au suivi de plusieurs classes de véhicules. Comme principale contribution, nous proposons de mener conjointement la catégorisation et le suivi au sein du Filtre Particulaire. Nous présentons également une mesure de distance anisotrope entre véhicules, basée sur leurs modèles géométriques 3D. Les résultats sont présentés et discutés sur des séquences de suivi de véhicules dans le trafic, impliquant différents types d'objets tels que motos, voitures, camionnettes et poids lourds.

Mots Clef

Suivi Visuel Multi-Objets, Catégorisation, Filtre Particulaire MCMC, Temps Réel.

Abstract

This paper addresses real-time automatic visual tracking and categorization of a variable number of generic objects. Multi-object configuration is estimated through a Markov Chain Monte-Carlo Particle Filter (MCMC PF). As the system yet is validated for pedestrian tracking, we demonstrate its ability to track multiple classes of vehicles in traffic. As a main contribution, this paper proposes to jointly track and categorize objects within the filter as a single processing step. This paper also addresses a vehicle anisotropic distance measurement based on 3D geometric model. Real time tracking results are shown and discussed on road sequences involving various types of vehicles such as motorcycles, cars, light trucks and heavy trucks.

Keywords

Multi-Object Visual Tacking, Object Categorization, Particle Filter, MCMC, Real Time.

1 Introduction

Le suivi visuel en temps réel d'un nombre variable d'objets est d'un grand intérêt pour de nombreuses applications telles que la surveillance et l'analyse comportementale de personnes, d'animaux ou de véhicules... Dans le domaine des Véhicules Intelligents, la surveillance du trafic est demandée par les applications d'assistance à la conduite et d'anticollision [9]. Dans les Systèmes de Transport Intelligents, le suivi visuel est d'un grand intérêt pour mesurer les flux de véhicules aux intersections, dans le but d'améliorer l'efficacité des transports urbains et routiers [2].

Nous avons choisi la vision pour sa grande étendue de mesure et ses faibles coûts et contraintes d'installation, mais la méthode présentée ci-dessous permet la fusion avec d'autres capteurs. Dans le système de suivi présenté ci-dessous, les objets (véhicules ou piétons) sont modélisés par des parallélépipèdes. Le suivi monoculaire d'une voiture a été réalisé en 1993 [11] et de nombreux travaux ont ensuite été publiés. Les occultations causées par la faible élévation de la caméra doivent être modélisées [8]. En outre, dans les scènes réelles de circulation cohabitent différentes catégories d'usagers de dimensions très différentes, tels que camions, camionnettes, voitures, motos, cycles et piétons. Le système doit aussi modéliser ces cibles de tailles très différentes.

Les Filtres Particulaires appartiennent aux méthodes de Monte-Carlo, plus précisément à la classe des filtres Bayésiens récursifs, et leur usage est très répandu dans le domaine du suivi d'objets, parce qu'ils gèrent les non-linéarités et multi-modalités induites par les occultations et les arrière-plans désordonnés. La clé de voûte des Filtres Particulaires est le choix judicieux des échantillons (les "particules") au fil du temps. Comme l'état du système évolue au cours du temps, les échantillons doivent être déplacés vers les régions de l'espace d'état les plus porteuses d'information. C'est la méthode séquentielle de ré-échantillonnage adaptatif connue sous le nom de l'algorithme SIR (Sequential Importance Resampling) [12]. Un suivi monoculaire multi-objet basé sur le SIR a été proposé [7], et de nombreux autres ouvrages ont suivi. L'inconvénient du SIR, montré par de nombreux auteurs [7, 16], est

qu'il est inefficace dans les espaces d'état de grande dimension, car le nombre de particules requises croît en exponentielle de la dimension de l'espace d'état. Ainsi, un Filtre Particulaire SIR ne peut suivre plus de 2 ou 3 objets.

Le Filtre Particulaire partitionné [13] a été proposé pour surmonter cette limitation, mais il n'échantillonne pas équitablement les objets [16]. Une approche hors ligne non-séquentielle, quand le temps réel n'est pas nécessaire, fait appel à une chaîne de Markov pour explorer l'espace des trajectoires d'objets sur toute une séquence [18].

Pour le suivi en temps réel, les Filtres Particulaires par Chaîne de Markov (MCMC PF) ont démontré leur efficacité, permettant de suivre 2 à 4 piétons [16]. Plus d'objets (jusqu'à 20 fourmis) [10] ont été suivis avec succès dans un cas où on évite les occultations (vue du dessus de fourmis marchant sur un sol plan). Nous associons cette méthode avec une fonction d'observation présentée dans [1], permettant de suivre les objets sur de grandes étendues de mesure, et de gérer leurs occultations. Après segmentation binaire de l'image en pixels d'arrière-plan et en pixels d'avant plan, le Filtre Particulaire propose, évalue, et propage des configurations multi-objets candidates. L'évaluation est basée sur la coïncidence des projections des modèles géométriques 3D des objets, avec les pixels d'avant-plan.

Le suivi de véhicules dans le trafic réel présente une difficulté supplémentaire : plusieurs catégories de géométrie et de dynamiques très différentes coexistent, telles que motos, voitures, poids-lourds... Une première réponse à ce problème, opérant avec un unique modèle géométrique d'objet, consiste à appliquer aux formes des objets candidats une forte dynamique, afin de les adapter aux formes des objets cibles. Cette approche est inopérante car elle ne peut empêcher plusieurs objets candidats de petite taille d'expliquer un objet cible de grande taille. Une stratégie en deux étapes - suivi puis catégorisation - est donc inadaptee. Nous intégrons donc la catégorisation au sein du filtre, la traitant en variable aléatoire comme les autres variables d'état de l'objet. Cette approche présente en outre la possibilité de générer des prédictions dynamiques spécifiques à chaque catégorie d'objet. Le suivi d'un objet par Filtre Particulaire intégrant plusieurs dynamiques a été proposé par [6], et appliqué au suivi d'un objet obéissant à des dynamiques très hétérogènes (rebond d'une balle), ou à la catégorisation d'actions d'un même objet (une main au repos, dessinant, ou coloriant). Nous l'appliquons ici à la catégorisation des objets.

Ce papier est organisé ainsi : en section 2, nous étendons la méthode MCMC PF pour suivre et catégoriser simultanément un nombre variable d'objets génériques. La fonction d'observation est décrite en section 3, avec un accent sur les interactions entre objets. Le système ayant été précédemment validé sur du suivi de piétons, les résultats expérimentaux présentés et discutés en section 4 sont focalisés sur du suivi de véhicules.

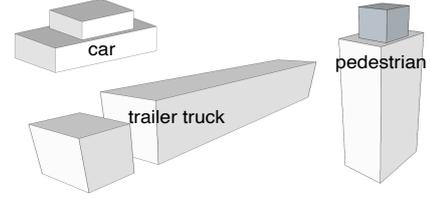


FIGURE 1 – Modèles géométriques d'une voiture, d'un camion semi-remorque et d'un piéton par 2 parallélépipèdes.

2 Suivi multi-objet par MCMC PF

Le suivi d'objets consiste à estimer leur configuration, ou état noté \mathbf{X}_t à chaque instant t . Notons $p(\mathbf{X}_t|\mathbf{Z}_{1:t})$ la densité de probabilité *a posteriori* que le système soit dans l'état \mathbf{X}_t à l'instant t , connaissant la séquence d'observations $\mathbf{Z}_{1:t}$. Les Filtres Particulaires propagent un nombre N de particules, pour approximer $p(\mathbf{X}_t|\mathbf{Z}_{1:t})$ comme une somme de distributions de Dirac, telles que : $p(\mathbf{X}_t|\mathbf{Z}_{1:t}) \approx \frac{1}{N} \sum_{n=1}^N \delta(\mathbf{X}_t - \mathbf{X}_t^n)$ où \mathbf{X}_t^n est le n -ième échantillon d'état, ou particule, à l'instant t . Les Filtres Particulaires par Chaîne de Markov génèrent ces échantillons itérativement par un processus Markovien d'ordre 1.

2.1 Espace d'état

En suivi d'objets, l'état encode la configuration des objets perceptibles : $\mathbf{X}_t^n = \{I_t^n, \mathbf{x}_t^{j,n}\}$, $j \in \{1, \dots, I_t^n\}$, où I_t^n est le nombre d'objets proposés par l'hypothèse n à l'instant t , $n \in \{1, \dots, N\}$ où N est le nombre de particules, et $\mathbf{x}_t^{j,n}$ est un vecteur encodant l'état de l'objet j , tel que $\mathbf{x}_t^{j,n} = \{c_t^{j,n}, \mathbf{p}_t^{j,n}, \mathbf{v}_t^{j,n}, \mathbf{a}_t^{j,n}, \mathbf{s}_t^{j,n}\}$. La catégorie de l'objet j à l'itération n est définie par $c_t^{j,n}$, une variable aléatoire discrète dont l'état peut prendre l'une des valeurs d'un sous ensemble de $\mathcal{C} = \{\text{piéton}, \text{moto}, \text{voiture}, \text{camionnette}, \text{poids lourd}\}$. Le sol est supposé plan. Ainsi la position absolue de l'objet j de la particule n au temps t , est définie par $\mathbf{p}_t^{j,n} = (x_t^{j,n}, y_t^{j,n}, \psi_t^{j,n})$, où $x_t^{j,n}$ et $y_t^{j,n}$ est la position absolue du centre de gravité et $\psi_t^{j,n}$ est l'angle de cap. La vitesse absolue de l'objet est $\mathbf{v}_t^{j,n}$, un vecteur décrit par son intensité et son orientation. La forme de l'objet peut être modélisée par un nombre choisi de parallélépipèdes définis par des dimensions et positions relatives. La figure 1 représente la modélisation géométrique d'une voiture, d'un camion semi-remorque et d'un piéton par l'union de deux parallélépipèdes. Nous choisissons une modélisation plus simple, par un seul parallélépipède. Ainsi $\mathbf{s}_t^{j,n}$ regroupe la longueur, la largeur et la hauteur de l'objet. Dans le cas d'un véhicule, la commande du conducteur au temps t est définie par $\mathbf{a}_t^{j,n} = \{\alpha_t^{j,n}, \gamma_t^{j,n}\}$, où $\alpha_t^{j,n}$, est l'angle de braquage et $\gamma_t^{j,n}$, l'accélération longitudinale.

2.2 Dynamique des objets

Afin d'améliorer l'efficacité du suivi, nous limitons l'espace d'état à l'étendue que peuvent physiquement atteindre les objets suivis. A cette fin, chaque objet est contrôlé par le

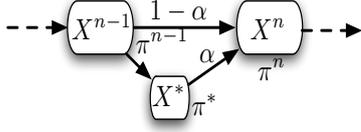


FIGURE 2 – Une itération du filtre MCMC PF produit une proposition \mathbf{X}^* , qui sera acceptée comme prochaine itération \mathbf{X}^n avec la probabilité α . Si elle est rejetée alors on duplique la précédente : $\mathbf{X}^n = \mathbf{X}^{n-1}$. Le taux d'acceptation α est calculé à partir de π^* , vraisemblance de \mathbf{X}^* , et de π^{n-1} , vraisemblance de \mathbf{X}^{n-1} .

modèle dynamique correspondant à sa catégorie. Si l'objet est supposé être un piéton, sa dynamique est décrite par les équations (1) et (2) :

$$x_t^{j,n} = x_{t-1}^{j,n} + T.v_{t-1}^n \cdot \cos(\psi_{t-1}^{j,n}), \quad (1)$$

$$y_t^{j,n} = y_{t-1}^{j,n} + T.v_{t-1}^n \cdot \sin(\psi_{t-1}^{j,n}), \quad (2)$$

où $v_{t-1,i}^n$ est la norme de la vitesse de l'objet au temps $t-1$, et T est la fréquence d'acquisition vidéo. Si l'objet est supposé être un véhicule, sa dynamique obéit également aux équations (1) et (2), et à deux équations supplémentaires modélisant le roulement sans glissement d'un véhicule à roues avant directrices, connu sous le nom de modèle bicycle [3, 4]. Les équations (3) et (4) modélisent respectivement la dynamique longitudinale et transversale du véhicule d'empattement L :

$$v_t^{j,n} = v_{t-1}^{j,n} + T.\gamma_{t-1}^{j,n}, \quad (3)$$

$$\psi_t^{j,n} = \psi_{t-1}^{j,n} + T.v_{t-1}^{j,n} \cdot \tan(\alpha_{t-1}^{j,n}) / L, \quad (4)$$

2.3 Suivi multi-objet par MCMC PF

Le suivi d'un nombre variable d'objets par Filtre Particulaire MCMC introduit et décrit dans [10], est basé sur la règle d'acceptation de Metropolis-Hastings [14]. Notre implémentation, détaillée dans [1], suit jusqu'à 12 véhicules ou piétons au temps réel vidéo (25 images par seconde). La figure 2 détaille la transition markovienne de la particule \mathbf{X}^{n-1} à la particule \mathbf{X}^n . Le point clé de la méthode est le choix de propositions de changement d'état marginalisées : la transition de la particule \mathbf{X} à la suivante \mathbf{X}^* , est conditionnée par la densité de proposition $q(\mathbf{X}^*|\mathbf{X})$, changeant toutes les dimensions d'un seul objet à la fois (choisi au hasard), les autres objets restant inchangés.

2.4 Nombre variable d'objets

Pour permettre aux objets d'entrer et sortir de la scène, les auteurs ont présenté le RJMCMC (Reversible Jump Markov Chain Monte Carlo) [5]. Comme le nombre d'objets visibles peut changer, la dimension de l'espace d'état peut aussi changer, l'état peut "sauter" d'un sous espace d'état vers un autre sous espace d'état de plus grande ou plus petite dimension si un nouvel objet entre ou si un objet quitte

TABLE 1 – dimensions des objets

catégorie	longueur (m)	largeur (m)	hauteur (m)
piéton	0.25 à 0.4	0.3 à 0.5	1.4 à 1.9
moto	1.5 à 1.8	0.4 à 0.6	1.2 à 1.5
voiture	3.5 à 5.0	1.4 à 2	1.2 à 1.9
camionnette	6.0 à 8.0	2 à 2.3	2.6 à 3.4
poids lourd	12.0 à 18.0	2.5 à 2.6	3.5 à 4.2

la scène. Pour empêcher la chaîne de recherche de se bloquer dans un minimum local, les sauts entre les sous-espace doivent être réversibles. Pour cette raison, les auteurs ont proposé la paire de mouvements discrets réversibles {entrée, sortie} [10, 16].

2.5 Mouvements proposés

Entrée : propose une nouvelle configuration jointe $\mathbf{X}^* = \{\mathbf{X}^{n-1}, \mathbf{x}_{I^{n-1}+1}\}$, ajoutant un nouvel objet $\mathbf{x}_{I^{n-1}+1}$ à la précédente configuration jointe \mathbf{X}^{n-1} , où I^{n-1} est le nombre d'objets porté par \mathbf{X}^{n-1} . On donne à l'objet entrant une vitalité initiale et des dimensions *a priori* conformément à la table 1. Le processus de recherche "saute" vers un sous-espace de plus grande dimension.

Sortie : le mouvement inverse propose l'exclusion de l'objet j de \mathbf{X}^{n-1} : propose $\mathbf{X}^* = \{\mathbf{X}^{n-1} \setminus \mathbf{x}^{j,n-1}\}$, $j \in \{1, \dots, I^n\}$, où I^n est le nombre d'objets dans l'hypothèse \mathbf{X}^{n-1} , et $\{s \setminus e\}$ est l'ensemble s privé de l'élément e . Le processus de recherche "saute" vers un sous-espace de plus petite dimension.

Mise à jour objet : dans ce mouvement, on tire au hasard un objet j de la particule \mathbf{X}_t^{n-1} à mettre à jour. On tire au hasard dans la chaîne précédente $\{\mathbf{X}_{t-1}^r\}_{r=1}^N$ une particule \mathbf{X}_{t-1}^r , $r \in \{1, \dots, N\}$, qui contient une instance de l'objet j à $t-1$ notée $\mathbf{x}_{t-1}^{j,r}$. On prédit tout d'abord sa catégorie c^* selon la loi suivante :

$$P(c^* = i | c_{t-1}^{j,r} = j) = \Theta_{ij}, \quad (5)$$

où Θ_{ij} est la probabilité de transition de la catégorie i à la catégorie j . Les catégories sont classées par ordre croissant de taille. Par exemple si $\mathcal{C} = \{moto, voiture, camionnette, poids lourd\}$, $c = 1$ pour une moto, etc... La matrice de probabilités de transitions Θ n'autorise que le passage vers les catégories les plus proches en termes de dimensions. Avec l'exemple précédent où $|\mathcal{C}| = 4$, Θ est définie par (6), où t_c est la probabilité de transition vers une des catégories les plus proches en termes de dimensions.

$$\Theta = \begin{pmatrix} 1 - 2.t_c & 2.t_c & 0 & 0 \\ t_c & 1 - 2.t_c & t_c & 0 \\ 0 & t_c & 1 - 2.t_c & t_c \\ 0 & 0 & 2.t_c & 1 - 2.t_c \end{pmatrix} \quad (6)$$

Si c^* propose que l'objet j soit un véhicule, on tire une

TABLE 2 – Empattement des véhicules et écarts-types des commandes conducteur.

catégorie véhicule	empattement L (m)	écart type accélération $\sigma_l (m.s^{-2})$	écart type braquage $\sigma_a (deg)$
moto	1.2	4	2
voiture	2.5	4	2
camionnette	4.5	2	2
poids lourd	10	1	1

proposition de commande du conducteur, selon (7) :

$$q(\mathbf{a}^* | \mathbf{a}_{t-1}^{j,r}) = \mathcal{N}(0, \text{diag}([\sigma_l^2, \sigma_a^2])), \quad (7)$$

où σ_l est l'écart-type des accélérations longitudinales pratiquées par les conducteurs, σ_a est l'écart-type des angles de braquage pratiqués par les conducteurs. Les équations du bicycle 3 et 4 sont alors appliquées à l'objet j pour mettre à jour sa vitesse et son angle de lacet. Les commandes pratiquées par les conducteurs dépendent de la catégorie de véhicule proposée, et sont données tableau 2. Si c^* propose que l'objet j soit un piéton, on tire une proposition de vitesse selon (8), où σ_m et σ_a sont les écarts types respectifs de norme et d'orientation de la vitesse.

$$q(\mathbf{v}^* | \mathbf{v}_{t-1}^{j,r}) = \mathcal{N}(\mathbf{v}_{t-1}^{j,r}, \text{diag}([\sigma_m^2, \sigma_a^2])) \quad (8)$$

Dans tous les cas, on applique enfin les équations de la dynamique (1) et (2) pour mettre à jour la position $\mathbf{p}_t^{j,n}$ de l'objet j . La forme de l'objet est mise à jour selon (9) :

$$q(\mathbf{s}^* | \mathbf{s}_{t-1}^{j,r}) = \mathcal{N}(\mathbf{s}_{t-1}^{j,r}, \sigma_s^2 I_3), \quad (9)$$

où σ_s est l'écart type de forme. La vitalité de l'objet j est mise à jour de façon incrémentale, l'incrément étant calculé à chaque itération en fonction de la vraisemblance de l'observation selon chaque objet.

2.6 Suivi et Catégorisation Jointe

L'intérêt de mener le suivi et la catégorisation conjointement au sein du MCMC PF est double. Il permet d'une part de maintenir et propager des hypothèses appartenant à différentes catégories, celle-ci étant traitée comme une variable aléatoire (discrète) au même titre que les autres attributs (continus) de l'objet. Il permet d'autre part de générer des prédictions dynamiques spécifiques à chaque catégorie d'objet, permettant à la dynamique temporelle des objets de contribuer à leur catégorisation au même titre que leur géométrie. Cela est d'autant plus bénéfique que les objets ont des dynamiques différentes, telles que poids lourd/voiture sur route sinueuse, et surtout dans le cas où le modèle géométrique n'est pas suffisamment discriminant (piéton/cycliste).

3 Fonctions de vraisemblance

Pour chaque particule proposée \mathbf{X}^* , la vraisemblance globale est calculée à partir de π_{fore}^* , vraisemblance de l'avant plan, et de π_{inter}^* , vraisemblance d'interaction inter objets :

$$\pi^* = \pi_{fore}^* \cdot \pi_{inter}^* \quad (10)$$

3.1 Vraisemblance d'avant plan

La fonction de vraisemblance de l'avant plan exprime la vraisemblance de la segmentation binaire arrière plan/avant plan, connaissant la particule proposée \mathbf{X}^* . A partir de l'image courante (figure 3-a), et d'un modèle de fond (figure 3-b), une image binaire d'avant plan $\mathbf{I}_F(g)$ est calculée (figure 3-d), par l'algorithme $\Sigma - \Delta$ [15]. Cet algorithme calcule efficacement en ligne une approximation de la médiane et de la variance temporelles de l'image de fond. Les deux sont remises à jour à chaque image, ce qui permet de faire face aux changements d'illumination et aux bruits pour un faible coût de calcul. D'autre part, on calcule pour chaque particule \mathbf{X} une image binaire du type de la figure 3-c), notée $\mathbf{I}_M(\mathbf{X}, g)$, où le pixel g vaut 1 s'il est à l'intérieur d'au moins une des enveloppes convexes des projections des parallélépipèdes d'objets de la particule dans l'image, sinon 0. On calcule enfin l'image de similarité $\mathbf{I}_S(\mathbf{X}, g)$, selon (11) :

$$\mathbf{I}_S(\mathbf{X}, g) = \begin{cases} 1 & \text{si } \mathbf{I}_F(g) = \mathbf{I}_M(\mathbf{X}, g) \\ 0 & \text{sinon} \end{cases} \quad \forall g \in \{1, \dots, G\}. \quad (11)$$

Les caméras de surveillance routière ne sont pas toujours placées en un point très élevé, ce qui cause de fortes occultations entre objets. En conséquence, la fonction de vraisemblance doit comparer équitablement les objets, qu'ils soient occultés ou non. De plus, la grande étendue de mesure (plusieurs centaines de mètres) provoque des variations considérable de la surface apparente des objets du fait de la projection. Nous avons proposé dans [1] un calcul de π_{fore}^* compatible avec ces contraintes. Il dépend de la surface de l'objet, et est basé sur $\mathbf{I}_S(\mathbf{X}, g)$, assurant à l'algorithme de Metropolis-Hastings un taux d'acceptation constant, indépendamment de la surface de l'objet.

3.2 Vraisemblance d'interaction

Comme la fonction de vraisemblance d'avant plan permet à un objet totalement occulté de survivre, nous devons l'empêcher de rester "caché" derrière un autre objet. Pour modéliser le fait que les objets se déplacent habituellement à distance les uns des autres, [10] affecte les densités de proposition de mouvements d'objets d'un potentiel d'interaction isotrope, basé sur la distance inter-objets. Pour suivre des personnes, [17] remplace ce potentiel par un potentiel anisotrope, basé sur la distance de Mahalanobis entre objets. Nous adoptons cette distance anisotrope, indispensable en suivi de véhicule, car leur longueur est beaucoup plus importantes que leur largeur (environ 8 fois pour les poids lourds). Cependant nous ne l'utilisons pas comme potentiel de proposition de mouvements d'objets,

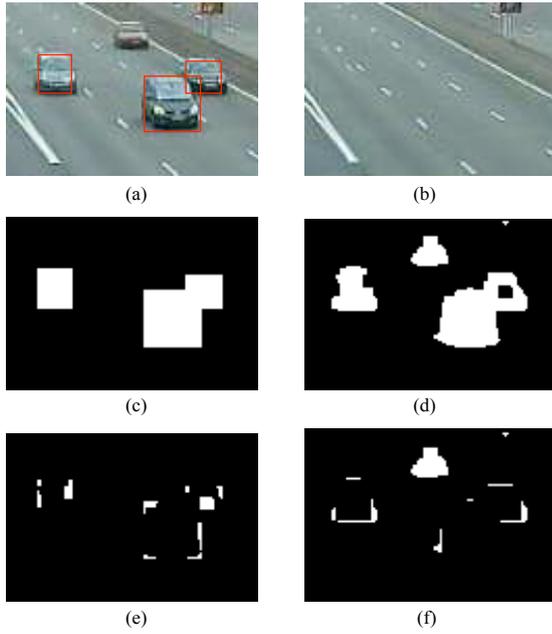


FIGURE 3 – Segmentation arrière plan/avant plan et images des résidus. (a) : images couleur avec les objets en rouge. Pour des raisons de lisibilité, les projections des objets sont approximées par des rectangles. (b) : modèle de fond. (c) : image binaire hypothèse. (d) : image binaire d'avant-plan. (e) : image binaire de "faux avant plan", pixels couverts par la projection d'au moins un objet, mais classés comme arrière-plan. (f) : image binaire de "faux arrière plan", pixels non couverts par la projection d'aucun objet hypothèse, mais classifiés comme avant-plan.

afin de maintenir des propositions marginalisées objet par objet. Nous la faisons donc intervenir au stade du calcul de vraisemblance. La distance entre les centres de gravité de deux véhicules roulant côte à côte sur deux voies voisines est d'environ 3 mètres, tandis que la même distance entre deux véhicules roulant l'un derrière l'autre sur la même voie est matériellement impossible. En outre, l'interaction entre les deux véhicules dépendra de leurs dimensions. Pour être compatible avec ces deux conditions, le calcul de vraisemblance d'interaction π_{inter}^* est basé sur la distance anisotrope entre chaque paire de véhicules, définie comme la distance de Mahalanobis entre deux distributions gaussiennes bidimensionnelles approximant chaque véhicule (chaque matrice de covariance est constituée des moments de masse d'ordre deux du véhicule) :

$$d_{ij} = (\Delta_{ij}^T \cdot (C_i \cdot C_j)^{-1} \cdot \Delta_{ij})^{1/2} \quad (12)$$

où Δ_{ij} est le vecteur de différence de position des véhicules i et j , C_i et C_j leurs matrices de covariance respectives. La sigmoïde (13), génère alors des vraisemblances tendant vers 1 pour les objets lointains et vers 0 pour les distances

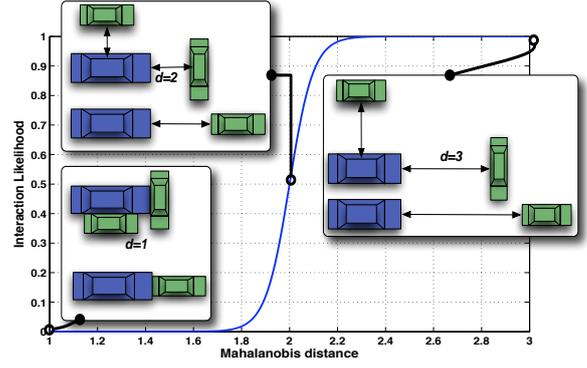


FIGURE 4 – Fonction sigmoïde définissant la probabilité d'interaction à partir de la distance anisotrope intervéhicule d . Quelques configurations de véhicules correspondant à trois valeurs de d , sont dessinées dans le cas de deux véhicules. Le véhicule bleu est nettement plus grand que le vert pour montrer l'indépendance de la distance à la taille du véhicule. d vaut 1 si les deux véhicules sont en contact, quelles que soient leurs faces en contact et leurs dimensions respectives. La sigmoïde est tracée pour les valeurs choisies $k = 20$ et $d_s = 2$, qui correspond à l'imagette en haut à gauche, permettant à deux véhicules de rouler côte à côte sur deux voies adjacentes, tout en pénalisant les configurations trop proches (imagette en bas à gauche).

inter-objets ne respectant pas les contraintes matérielles.

$$\pi_{ij} = \left(1 + e^{-k_s \cdot (d_{ij} - d_s)}\right)^{-1}, \quad (13)$$

où d_s est la distance "seuil" et k_s un coefficient permettant d'ajuster la pente de la sigmoïde autour de d_s . Nous choisissons $d_s = 2$ et $k_s = 20$ dans nos expérimentations, comme représenté sur la figure 4. La vraisemblance d'interaction de la particule \mathbf{X} contenant O objets vaut alors :

$$\pi_{inter}^* = \prod_{i=1}^O \prod_{j=i+1}^O \pi_{ij} \quad (14)$$

3.3 Propositions guidées par l'observation

En plus de l'image de similarité décrite en section 3.1, deux images résiduelles sont calculées pour chaque particule, utilisées pour guider les propositions *entrée* et *sortie*. Les propositions d'*entrée* sont guidées par le taux de "faux arrière plan" (Figure. 3-f), taux de pixels non couverts par aucun des objets de la particule mais classifiés avant-plan par la segmentation fond-forme. En l'occurrence, la particule courante est une hypothèse à trois objets, laissant un amas important sur l'image de "faux arrière plan" (Figure. 3-f). Cet amas sera utilisé pour proposer l'entrée d'un 4e objet. Les propositions de *sortie* de l'objet j sont guidées par sa vitalité, une variable continue incrémentée ou décrémentée à chaque itération en fonction du taux de "faux avant plan" (Figure 3-e), c'est à dire le taux de pixels contenus dans l'enveloppe convexe de la projection de l'objet

j , mais classifiés comme avant-plan par la segmentation fond-forme. Ce processus de mise à jour itérative de la vitalité maintient la cohérence temporelle de la présence et de l'identité des objets suivis : les propositions de *sortie* de l'objet j augmentent quand sa vitalité baisse, c'est à dire quand celui-ci a été mal extrait de l'arrière-plan depuis plusieurs images.

4 Expérimentations et résultats

Ce système de suivi et catégorisation multi-objets a été conçu pour être aussi générique que possible, utilisant aussi peu d'a priori que possible. Les expériences présentées ci-après ont été menées sur différents scénarios, impliquant des motos, véhicules légers, des camionnettes et poids lourds. Le respect du temps réel vidéo (25 images par seconde sur des images 480x640) par l'algorithme a été démontré dans [1], en utilisant la bibliothèque *NT² C++*¹ sous Linux sur un PC bi-processeur 3GHz Intel E6850, avec 3,25Go de RAM.

Dans les mêmes conditions, les expériences ci-dessous visent à évaluer la capacité de suivi de plusieurs véhicules appartenant à diverses catégories. Le suivi a été mené sur trois scénarios. Le premier vise à évaluer la précision du suivi de trajectoire. Le véhicule (une moto) est équipé d'un GPS RTK (précision centimétrique) pour fournir une vérité terrain. Une caméra couleur 480x640 vise l'entrée de courbe, et une autre la sortie de courbe, comme indiqué sur la Fig. 5. Le second scénario est une route sinueuse sur laquelle circulent voitures, motos et camions. Le troisième scénario est la sortie d'une barrière de péage d'autoroute, avec des véhicules légers, des camionnettes et des poids lourds, capturé avec une webcam de faible résolution (320x240), faible fréquence (15 images par seconde) (voir figure 6). Sur ces deux figures, les fines lignes noires définissent les limites de la zone de suivi. Lorsqu'un nouvel objet entre dans ce domaine, il est étiqueté avec un identifiant unique, et la projection de son estimation est tracée.

4.1 Précision du suivi

L'acquisition vidéo et la vérité terrain par GPS centimétrique ont été réalisées par le LCPC². Le suivi a été répété 10 fois sur chacune des 4 séquences. La moyenne et la variance de l'erreur d'estimation de trajectoire par rapport au GPS sont indiquées dans le tableau 3. L'erreur est beaucoup plus élevée pour les séquences d'entrée courbe que pour les séquences de sortie de courbe. Les causes sont :

- l'initialisation est beaucoup plus précise pour les objets proches de la caméra,
- l'ombre de la moto biaise nettement plus la position sur les séquences d'entrée, du fait de la position du soleil.

Pour l'analyse des trajectoires de véhicules, l'objectif est une précision d'environ 20 cm sur la position du véhicule. Nos résultats répondent à cette exigence sur les séquences de sortie de courbe. Mais l'erreur des séquences d'entrée



FIGURE 5 – Gauche : entrée de courbe. Droite : sortie de courbe. Enveloppe estimée et trajectoire de la moto. Haut : images couleur. Bas : images binaires d'avant-plan.

TABLE 3 – Erreur moyenne de position par rapport à la mesure GPS

séquence	moyenne (m)	écart-type (m)
entrée de courbe	0.99	0.067
sortie de courbe	0.15	0.039

est beaucoup trop élevée. En conséquence, pour obtenir la précision de 20 cm, nous aurons à modéliser l'ombre des objets suivis dans nos prochains travaux. En outre, dans ces expériences, le roulis (la rotation autour de l'axe longitudinal) de la moto n'est pas modélisé. Il faudrait l'ajouter à l'espace d'état et le prendre en compte dans un modèle dynamique de moto plus réaliste. Le processus d'initialisation d'objet peut également être amélioré.

4.2 Suivi et catégorisation sur autoroute

Pour évaluer la performance, le suivi des véhicules a été mené sur une séquence de 13.500 images (15 minutes à 15 images par seconde), impliquant 439 véhicules appartenant à 3 catégories : 330 véhicules légers, 58 camionnettes et 51 poids lourds. La scène, située après une barrière de péage, est illustré sur la Figure 6, avec zone de suivi de 180 x 18 mètres tracée en lignes noires. On a choisi des conditions de circulation moyennes, ainsi que des conditions météo nuageuses, pour minimiser les ombres portées. Un véhicule est considéré comme correctement suivi si :

- est suivi sur au moins 75% de la zone de suivi,
- erreur d'estimation de position inférieure à 2 mètres.

La vérité terrain (suivi et catégorisation) a été établie manuellement. Le système est testé dans trois configurations :

- Deux catégories, distance inter-objet anisotrope,
- Trois catégories, distance inter-objet anisotrope,
- Trois catégories, distance inter-objet isotrope.

1. <http://nt2.sourceforge.net>

2. www.lcpc.fr



FIGURE 6 – Enveloppe estimée et trajectoire de camions et voitures sur autoroute (la zone de suivi est délimitée par les lignes noires).

TABLE 4 – Suivi et catégorisation en 2 catégories sur autoroute. Distance inter-vehicule anisotrope.

	voiture	camionnette	poids lourds	total
Vérité	330	58	51	439
Suivis	304 (92%)	52 (90%)	50 (98%)	406 (92%)
Classifiés	302 (91%)	X	39 (76%)	341 (78%)
Fantômes	53 (16%)	X	5 (10%)	58 (13%)

Deux catégories de véhicules avec distance anisotrope : Véhicules légers et poids lourds. Les résultats du tableau 4 montrent de bons résultats en suivi, tandis que plus de 20 % des véhicules sont mal catégorisés. Cela est principalement dû à l'absence de classe de véhicules moyens (camionnettes). La dernière ligne du tableau met en évidence un nombre élevé de véhicules "fantômes" la classe des camionnettes étant inactive, le système entretient des hypothèses tentant d'expliquer la silhouette d'une camionnette par deux silhouettes de voitures.

Trois catégories de véhicules avec distance anisotrope : Véhicules légers, camionnettes et poids lourds. Les résultats du tableau 5 montrent que l'activation de trois modèles de véhicules améliore nettement les performances de suivi et de catégorisation, les particules expliquant mieux l'avant-plan. En conséquence, les véhicules "fantômes" apparaissent moins fréquemment (dernière ligne du tableau).

Trois catégories de véhicules avec distance isotrope : Véhicules légers, camionnettes et poids lourds. Nous avons choisi comme distance isotrope la distance euclidienne entre les centres de gravité des véhicules avec $d_s = 3$ mètres (la largeur d'une voie). Les résultats du tableau 6 montrent

TABLE 5 – Suivi et catégorisation en 3 catégories sur autoroute. Distance inter-vehicule anisotrope.

	voiture	camionnette	poids lourds	total
Vérité	330	58	51	439
Suivis	318 (96%)	58 (100%)	51 (100%)	427 (97%)
Classifiés	315 (95%)	30 (52%)	51 (100%)	396 (90%)
Fantômes	27 (8%)	2 (3.5%)	0	29 (6%)

TABLE 6 – Suivi et catégorisation en 3 catégories sur autoroute. Distance inter-vehicule isotrope.

	voiture	camionnette	poids lourds	total
Vérité	330	58	51	439
Suivis	320 (97%)	58 (100%)	51 (100%)	429 (98%)
Classifiés	320 (97%)	34 (57%)	47 (92%)	401 (91%)
fantômes	31 (9%)	20 (34%)	4 (8%)	55 (12%)

que cette distance génère beaucoup plus de véhicules "fantômes", en particulier sur les camions. Malgré une très légère amélioration du suivi des cibles et de leur catégorisation, en raison de la diminution des contraintes de suivi, ces résultats montrent que notre distance anisotrope est plus efficace pour la modélisation des interactions entre véhicules.

4.3 Suivi et catégorisation sur route sinueuse

Les expériences sur route sinueuse visent à montrer que l'algorithme peut traiter tout type de séquence routière, sans aucune adaptation *ad-hoc* des modèles dynamique des véhicules. Le suivi des véhicules a été menée sur une séquence de 5.500 images (480x640 enregistré à 25 images par seconde), impliquant 37 véhicules : 20 véhicules légers, 15 motos et 2 camionnettes. La scène est illustrée sur la Fig. 7, les lignes noires délimitant la zone de suivi de 160 x 6 mètres. Sur cette séquence, 100% des véhicules ont été correctement suivis et classés.

5 Conclusion et perspectives

Nous avons proposé un cadre unifié pour suivre et catégoriser simultanément plusieurs véhicules appartenant à différentes catégories en temps réel. Le système est entièrement intégré au sein d'un Filtre Particulaires MCMC. Pour ce faire, la classe de l'objet est considérée comme une variable aléatoire. Les expériences montrent que mener simultanément



FIGURE 7 – Enveloppe estimée et trajectoires motos et voitures sur route sinueuse.

ment le suivi et la catégorisation améliore le suivi. Cette approche unifiée est également d'un grand intérêt car elle exploite la dynamique des objets au même titre que leurs formes pour les catégoriser. Cette fonctionnalité peut être utilisée pour améliorer le suivi et la catégorisation sur des routes sinueuses ou des zones urbaines, où la dynamique du véhicule diffère d'une classe à l'autre. En outre, elle pourrait aider à classer les cyclistes et les piétons grâce à la différence de leurs dynamiques. Toutes ces expériences ont été menées par temps nuageux, ce qui minimise les ombres portées des véhicules. Celles-ci n'étant pas encore modélisées, elles perturbent fortement la capacité de suivi. Nous allons remédier à ce problème dans nos prochains travaux.

Références

- [1] François Bardet and Thierry Chateau. MCMC particle filter for real-time visual tracking of vehicles. In *International IEEE Conference on Intelligent Transportation Systems*, pages 539 – 544, 2008.
- [2] T. Bragatto, G. Ruas, V. Benso, M. Lamar, D. V. Aldigueri, G.L. Teixeira, and Y. Yamashita. A new approach to multiple vehicle tracking in intersections using harris corners and adaptive background subtraction. In *Int. IEEE Intelligent Vehicle Symposium*, pages 548–553, 2008.
- [3] T. Gillespie. *Fundamentals of Vehicle Dynamics*. Society of Automotive Engineers, 1992.
- [4] Yann Goyat, Thierry Chateau, Laurent Malaterre, and Laurent Trassoudaine. Vehicle trajectories evaluation by static video sensors. In *9th International IEEE Conference on Intelligent Transportation Systems Conference (ITSC 2006)*, 2006.
- [5] Peter J. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 4(82) :711–732, 1995.
- [6] M. Isard and A. Blake. A mixed-state condensation tracker with automatic model-switching. In *Computer Vision, 1998. Sixth International Conference on*, pages 107–112, 1998.
- [7] Michael Isard and John MacCormick. Bramble : A bayesian multiple-blob tracker. In *Proc. Int. Conf. Computer Vision*, vol. 2 34-41, 2001.
- [8] S. J. Kanhere, N. K. Pundlik and S. T. Birchfield. Vehicle segmentation and tracking from a low-angle off-axis camera. In *CVPR, Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1152–1157, 2005.
- [9] Bernd Roessler Kay Ch. Fuerstemberg. Advanced intersection safety - the ec project intersafe. In *International IEEE Intelligent Vehicle Symposium*, pages 89–93, 2006.
- [10] Z. Khan, T. Balch, and F. Dellaert. MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 :1805 – 1918, 2005.
- [11] D. Koller, K. Daniilidis, and H. H. Nagel. Model-based object tracking in monocular image sequences of road traffic scenes. *International Journal of Computer Vision*, 10(3) :257–281, 1993.
- [12] M. Isard and A. Blake. Condensation – conditional density propagation for visual tracking. *IJCV : International Journal of Computer Vision*, 29(1) :5–28, 1998.
- [13] John MacCormick and Andrew Blake. A probabilistic exclusion principle for tracking multiple objects. In *Int. Conf. Computer Vision*, 572-578, 1999.
- [14] David MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.
- [15] Antoine Manzanera and Julien Richefeu. A robust and computationally efficient motion detection algorithm based on sigma-delta background estimation. In *Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP'04)*, pages 46–51, 2004.
- [16] Kevin Smith. *Bayesian Methods for Visual Multi-Object Tracking with Applications to Human Activity Recognition*. PhD thesis, EPFL, Lausanne, Suisse, 2007.
- [17] Jian Yao and Jean-Marc Odobez. Multi-camera multi-person 3D space tracking with mcmc in surveillance scenarios. In *European Conference on Computer Visionworkshop on Multi Camera and Multi-modal Sensor Fusion Algorithms and Applications (ECCV-M2SFA2)*, 2008.
- [18] Q. Yu, G. Medioni, and I. Cohen. Multiple target tracking using spatio-temporal markov chain monte carlo data association. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1 – 8, 2007.