Approche multi-noyaux pour le suivi temps-réel d'objet mobile

C. Teuliere

▶ **To cite this version:**

C. Teuliere. Approche multi-noyaux pour le suivi temps-réel d'objet mobile. ORASIS'09 - Congrès des jeunes chercheurs en vision par ordinateur, 2009, Trégastel, France, France. inria-00404649

HAL Id: inria-00404649
https://inria.hal.science/inria-00404649

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.

Approche multi-noyaux pour le suivi temps-réel d'objet mobile

Multiple kernel approach for real-time tracking of a moving object

C. Teulière

CEA, LIST, Service de Robotique Interactive, 18 route du Panorama, BP6, F- 92265 Fontenay aux Roses
INRIA Rennes-Bretagne Atlantique, IRISA, Lagadic, Rennes

Celine.Teuliere@cea.fr

Résumé

Les méthodes de suivi basé couleur ont fait la preuve de leur efficacité en terme de robustesse. La contrepartie de ce type de représentation par histogramme est qu'elle ne comprend pas ou peu d'information sur la configuration spatiale de l'objet, ce qui rend plus difficile l'analyse de déplacements complexes. Un moyen de surmonter ce problème consiste à utiliser des noyaux pour pondérer les pixels en fonction de leur position dans l'image. Cet article présente ainsi une configuration multi-noyaux développée à la fois dans le cadre des techniques d'optimisation déterministes et dans le cadre bayésien du filtrage particulaire. Les deux approches sont combinées pour concevoir un algorithme de suivi robuste, permettant de suivre la position, taille et orientation de l'objet. Des résultats comparatifs viennent valider l'approche proposée.

Mots Clef

Suivi visuel, histogrammes à noyaux, filtrage particulaire.

Abstract

Color-based tracking methods have proved to be efficient for their robustness qualities. The drawback of such representation of an object is the lack of information on its spatial configuration, making difficult the tracking of more complex motions. This issue can be overcome by using several kernels weighting pixels locations. In this paper a multiple kernels configuration is proposed and developed in both probabilistic and deterministic frameworks. The advantages of both approaches are combined to design a robust tracker allowing to track location, size and orientation of the object. Comparative results validate the proposed method.

Keywords

Visual tracking, kernel-based histograms, particle filtering.

1 Introduction

Le suivi en temps réel d'un objet dans une séquence d'images est une tâche essentielle pour un grand nombre d'applications utilisant des informations visuelles, telles que l'asservissement visuel, la surveillance, la réalité augmentée, etc. De façon générale, le principe du suivi visuel consiste à déterminer quelle partie de l'image courante "correspond le mieux" à un modèle de référence de l'apparence de l'objet que l'on cherche à suivre.

Pour cela il est nécessaire de définir des descripteurs de cette apparence, leur choix dépendant souvent de l'application visée. Lorsque l'objet à suivre est connu et que son apparence varie peu au cours du déplacement, l'utilisation de motifs [1] [2] [3] qui décrivent l'objet à l'échelle du pixel permet de conserver la structure spatiale de l'objet et donne de bons résultats pour suivre une grande variété de déplacements. Cependant, ce type de descripteurs est très sensible aux modifications de l'apparence de l'objet (flou, changements d'illumination, etc.) qui peuvent alors mettre en échec le suivi. À l'inverse, les descripteurs par densité de probabilité comme les histogrammes de couleur sont plus robustes, ce qui les rend intéressants pour traiter des tâches complexes.

Le suivi basé couleur a été principalement considéré selon deux approches. Dans l'algorithme du *mean shift* [4] la recherche de l'objet est effectuée en minimisant une distance entre histogrammes de couleur selon une méthode de type descente de gradient, et conduit à une bonne précision de suivi. Cependant, cette recherche étant déterministe elle ne permet pas d'être robuste aux occultations importantes, et l'algorithme peut échouer en présence d'un autre objet de couleurs similaires, ou dans le cas de grands déplacements. Dans un autre cadre, [5] et [6] utilisent le même type de critère de ressemblance entre histogrammes, mais l'intègrent dans un filtre particulaire. La densité de probabilité *a posteriori* de la position de l'objet est discrétisée en un ensemble de particules. L'évolution de ces particules et l'estimation de leur moyenne remplace ici la minimisation effectuée dans l'algorithme du *mean shift*. L'utilisation

de ce cadre probabiliste du filtrage particulaire induit une meilleure robustesse vis-à-vis des occultations ou de la présence d'objets similaires.

Dans les études ci-dessus, les descripteurs par histogrammes de couleur utilisés ne permettent pas ou peu de conserver des informations sur la configuration spatiale de l'objet car les pixels sont regroupés par couleur sans tenir compte de leur position, ce qui rend difficile le suivi de déplacements complexes plus complexes que les translations souvent étudiées.

Les méthodes basées sur l'utilisation de noyaux, comme [4], ajoutent une information de nature spatiale en donnant un poids plus élevés aux pixels proches du centre de l'objet. Cependant les noyaux symétriques utilisés sont peu sensibles aux rotations. Le problème de l'utilisation de descripteurs à noyaux pour le suivi de déplacements complexes est abordé dans [8]. En faisant le lien entre les méthodes à noyaux et les techniques d'optimisation, les auteurs montrent comment les adapter pour utiliser plusieurs noyaux. Cela permet de concevoir des algorithmes de suivi plus sensibles à des mouvements spécifiques. Des exemples de noyaux spécifiques sont donnés, sans toutefois proposer de configuration générique pour traiter une large variété de déplacements. Le problème peut également être formulé en terme de composition inverse [9].

Dans le cadre bayésien, il est nécessaire de définir une fonction de vraisemblance qui soit sensible aux déplacements considérés. L'idée de diviser l'objet en plusieurs parties a été introduite dans [5] afin d'améliorer la précision du suivi. Dans ce papier nous proposons une configuration multi-noyaux qui peut s'adapter au suivi de déplacements affines ou homographiques. Les deux approches ci-dessus sont développées en utilisant cette configuration et des résultats comparatifs pour le suivi de la position, orientation et taille d'un objet sont présentés. La meilleure sensibilité de l'approche multi-noyaux est mise en évidence, et une combinaison des deux approches est proposée. L'algorithme de suivi est utilisé dans le cadre d'une tâche d'asservissement visuel.

La suite de ce papier s'organise comme suit : la partie 2 formule le problème du suivi basé couleur et décrit la représentation multi-noyaux proposée. Cette configuration est présentée dans une approche déterministe dans la partie 3, et dans le cadre probabiliste du filtrage particulaire dans la partie 4. Une combinaison de ces deux approches est proposée en partie 5. Enfin des résultats expérimentaux viennent étayer nos propos.

2 Suivi basé couleur

2.1 Espace d'état

L'objectif d'un algorithme de suivi visuel est d'estimer l'état d'un objet ou d'une région d'intérêt dans chaque image d'une séquence vidéo. L'état désigne souvent la position, mais peut également contenir d'autres paramètres de déformation de la cible.

Les descripteurs par densité de probabilité, tels que les histogrammes de couleur, fournissant peu d'informations sur la configuration spatiale de l'objet, ils sont principalement utilisés pour suivre la position et parfois les changements d'échelle. Ce papier vise à adapter ce type de mesures pour suivre des déplacements plus complexes. Dans les expérimentations présentées ici, nous considérons des translations, des rotations et des changements d'échelle, ce qui permet d'utiliser ce suivi pour asservir visuellement quatre degrés de liberté d'un système robotique. Les méthodes développées peuvent cependant être adaptées pour une plus grande variété de déplacements, tels que les mouvements affines ou homographiques.

Dans la suite de cet article, sans perte en généralité, la région considérée est un rectangle de rapport fixe $r = \frac{h}{w}$, dont l'état est défini par sa position (x, y), son orientation θ et sa surface $A = hw$ (voir Figure 1). L'état $\mathbf{x}_k$ de l'objet dans l'image k est alors défini par $\mathbf{x}_k = (x_k, y_k, \theta_k, \frac{1}{\sqrt{A_k}})^t$. Dans la suite, $\{\boldsymbol{l}_i\}_{i=0..n}$ représente les positions pixelliques initiales de l'objet, qui est supposé être sélectionné ou détecté automatiquement dans la première image. Si $f_{\mathbf{x}}$ désigne la déformation entre l'état initial de l'objet et l'état $\mathbf{x}$, pour chaque position pixellique $\boldsymbol{l}_i$:

$$f_{\mathbf{x}}(\boldsymbol{l}_i) = f_{\mathbf{x}}(\mathbf{c}) + s\mathbf{R}_\theta(\boldsymbol{l}_i - \mathbf{c}) \quad (1)$$

où $\mathbf{R}_\theta$ est la matrice de rotation 2×2 d'angle $\theta - \theta_0$, $s = \sqrt{\frac{A}{A_0}}$ et $\mathbf{c} = (x, y)^t$ est le centre du rectangle.

2.2 Espace de mesure

Dans le suivi basé couleur, l'objet de référence est généralement caractérisé par son histogramme de couleur $\mathbf{q}^* = \{q_u^*\}_{u=1\ldots m}$, m étant le nombre de classes de l'histogramme. Plusieurs espace de couleurs peuvent être utilisés comme HSV [11] [5]. Dans notre cas nous avons retenu l'espace RVB avec $8 \times 8 \times 8$ classes. Pour chaque classe u :

$$q_u^* = \sum_{i=1}^{n} K(\boldsymbol{l}_i - \mathbf{c})\delta_u(b(\boldsymbol{l}_i)) \quad (2)$$

où $b(\boldsymbol{l}_i)$ est la classe correspondant à la couleur du pixel $\boldsymbol{l}_i$, δ désigne la fonction de Kronecker, et K est une fonction noyau centrée en $\mathbf{c}$, et pondérant les positions dans l'image. Sauf mention contraire K est supposée normalisée de sorte que $\sum_{u=1}^{m} q_u^* = 1$.

Quel que soit le cadre de suivi choisi (filtrage particulaire ou optimisation basée noyau), dans chaque image k, le but est de déterminer l'état candidat $\mathbf{x}_k$ pour lequel l'histogramme $\mathbf{q}(\mathbf{x}_k)$ est le plus "proche" de l'histogramme de référence $\mathbf{q}^*$. Pour cela, un critère de corrélation dans l'espace des histogrammes est fourni par le coefficient de Bhattacharyya :

$$\rho(\mathbf{x}_k) = \rho(\mathbf{q}^*, \mathbf{q}(\mathbf{x}_k)) = \sum_{u=1}^{m} \sqrt{q_u^* q_u(\mathbf{x}_k)} \quad (3)$$

$q_u(\mathbf{x}_k)$ étant calculé comme dans (2) avec les positions pixelliques correspondant à la cible en $\mathbf{x}_k$ *i.e.* $\{f_{\mathbf{x}_k}(\boldsymbol{l}_i)\}_{i=0..n}$.

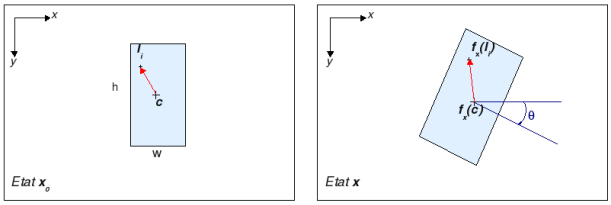


FIG. 1 – Représentation de l'état à suivre et déformation associée entre l'état $\mathbf{x}_0$ (gauche) et l'état $\mathbf{x}$ (droite).

Un état candidat $\mathbf{x}_k$ pour l'objet dans l'image k est alors comparé à l'objet de référence en utilisant la distance de Bhattacharyya :

$$d(\mathbf{x}_k) = d(\mathbf{q}^*, \mathbf{q}(\mathbf{x}_k)) = \sqrt{1 - \rho(\mathbf{x}_k)} \qquad (4)$$

Bien que cette représentation par un unique histogramme soit suffisante pour suivre la position et/ou l'échelle de l'objet, elle est très peu sensible aux rotations (voir Figure 3), en particulier lorsque le rapport r est proche de 1.
La partie suivante présente un moyen d'adapter ce modèle de représentation afin de le rendre plus sensible à des déplacements plus complexes.

2.3 Configuration multi-noyaux

Comme signalé dans [8], l'utilisation de plusieurs noyaux (et donc plusieurs histogrammes) augmente l'espace de mesure, ce qui rend le suivi sensible à une plus grande variété de déplacements. Un moyen de concevoir un algorithme de suivi sensible aux déplacements considérés ici, consiste à concevoir des noyaux spécifiques à chaque mouvement. Ainsi, [8] fournit des exemples de noyaux sensibles aux rotations. Cependant, étant donné que plusieurs histogrammes doivent être calculés sur l'objet entier, le temps de calcul est plus important, ce qui rend cette méthode difficile à utiliser dans le cadre du filtrage particulaire.
La configuration proposée dans ce papier consiste à utiliser plusieurs noyaux identiques, centrés comme illustré sur la Figure 2. La pondération des pixels selon ces noyaux leur donne une importance différente en fonction de leur position dans l'objet. Dans ce cas, le nombre de noyaux n'entraine pas une augmentation du temps de calcul dans la mesure où les histogrammes peuvent être calculés sur des sous-parties de l'objet. De plus, ce type de configuration pourrait être utilisé pour détecter un large panel de transformations.
Plus formellement, un état $\mathbf{x}_k$ est associé à un ensemble d'histogrammes (neuf ici) $\{\mathbf{q}_j(\mathbf{x}_k)\}_{j=1..9}$ calculés en utilisant les noyaux $K_j(\boldsymbol{l} - \mathbf{c}_j)$. Les distances $d_j(\mathbf{x}_k) = d(\mathbf{q}_j^*, \mathbf{q}_j(\mathbf{x}_k))$ sont définies comme dans (4).
Figure 3.b montre l'allure de la fonction de distance :

$$d_m(\mathbf{x}_k) = \frac{1}{9}\sum_{j=1}^{9} d_j(\mathbf{x}_k) \qquad (5)$$

calculée sur l'image de la Figure 7-b. La représentation multi-noyaux étant plus discriminatoire, la distance résultante varie plus vite qu'en utilisant un seul noyau. Par conséquent, la sensibilité de la méthode de suivi aux déplacements tels la rotation est améliorée. Il est bien sûr possible d'ajouter des noyaux pour augmenter la précision du suivi.
Les deux prochaines parties présentent deux façons différentes de considérer le problème de la recherche dans chaque image k de l'état $\mathbf{x}_k$ qui corresponde le mieux au modèle de référence, en utilisant des histogrammes de couleur comme mesures.

3 Cadre déterministe

Cette partie montre comment le formalisme de [8] s'applique à la configuration multi-noyaux présentée dans la section précedente (voir Figure 2).
Ici, un histogramme est calculé avec chaque noyau K_j à partir des n positions pixelliques de l'objet comme dans (2). Suivant le formalisme général proposé dans [8], les histogrammes sont à présent considérés sous leur forme vectorielle,
$\mathbf{q} = (q_1{}^{(1)}, .., q_m{}^{(1)}, ..., q_1{}^{(9)}, .., q_m{}^{(9)})^t$, où $q_u{}^{(j)}$ désigne la valeur de la classe u dans l'histogramme calculé avec K_j.
Pour rendre possibles les techniques d'optimisation, la distance dans l'espace des histogrammes pour un état candidat $\mathbf{x}$ est considérée ici sous la forme :

$$O(\mathbf{x}) = \|\sqrt{\mathbf{q}^*} - \sqrt{\mathbf{q}(\mathbf{x})}\|^2 = 2 - 2\rho(\mathbf{x}) \qquad (6)$$

où la racine carrée est appliquée à chaque composant des vecteurs. On note que les minima de (6) correspondent aux minima de (4).

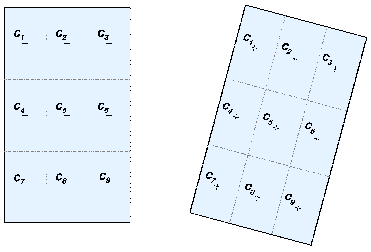
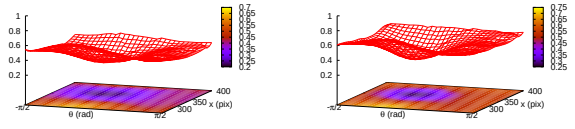


FIG. 2 – Configuration multi-noyaux.



(a) (b)

FIG. 3 – Fonction de distance en fonction de la rotation et de la position sur l'axe x, en utilisant une représentation de l'objet à un seul (a) ou plusieurs noyaux (b). Le minimum est beaucoup plus marqué dans le 2^e cas.

La définition de chaque fonction noyau K_j est étendue dans la suite pour inclure la déformation $f_{\mathbf{x}}$ de l'état $\mathbf{x}_0$ à l'état $\mathbf{x}$ en définissant pour chaque pixel $\boldsymbol{l}$:

$$K_j(\boldsymbol{l}, \mathbf{c}_j, \mathbf{x}) = CK(f_{\mathbf{x}}(\boldsymbol{l}) - \mathbf{c}_j) \tag{7}$$

$$C = \frac{1}{\sum_i K(f_{\mathbf{x}}(\boldsymbol{l}_i) - \mathbf{c}_j)} \tag{8}$$

où $\mathbf{c}_j$ est le centre du noyau K_j.
Le gradient du vecteur de noyau $\mathbf{K}_j(\mathbf{x}) = (K_j(\boldsymbol{l}_1, \mathbf{c}_j, \mathbf{x}), ..., K_j(\boldsymbol{l}_n, \mathbf{c}_j, \mathbf{x}))^t$ est donnée par :

$$\mathbf{J}_{\mathbf{K}j} = \left[\frac{\partial \mathbf{K}_j}{\partial x}, \frac{\partial \mathbf{K}_j}{\partial y}, \frac{\partial \mathbf{K}_j}{\partial \theta}, \frac{\partial \mathbf{K}_j}{\partial s}\right]$$

La matrice $\mathbf{U}$ définie par $\mathbf{U} = \{u_{i,j}\}_{i=1..n, j=1..m}$ avec $u_{i,j} = \delta_j(b(\boldsymbol{l}_i))$, lie les n points avec la classe de couleur du pixel correspondant dans l'image considérée (voir [8]).
Alors, la minimisation de (6) en utilisant la méthode de Newton conduit à

$$\Delta \mathbf{x} = -2\mathbf{J}_{\mathbf{U}}^{+}\mathbf{e} \tag{9}$$

où $\mathbf{e} = \sqrt{\mathbf{q}(\mathbf{x})} - \sqrt{\mathbf{q}^*}$, $\mathbf{J}_{\mathbf{U}} = d(\mathbf{q}(\mathbf{x}))^{-\frac{1}{2}} \begin{bmatrix} \mathbf{U}^t\mathbf{J}_{\mathbf{K}1} \\ ... \\ \mathbf{U}^t\mathbf{J}_{\mathbf{K}9} \end{bmatrix}$, $\mathbf{J}_{\mathbf{U}}^{+}$ désigne la pseudo-inverse de $\mathbf{J}_{\mathbf{U}}$ et $d(\mathbf{q}(\mathbf{x}))$ représente la matrice de diagonale $\mathbf{q}$.

Dans notre étude, des noyaux d'Epanechnikov ont été utilisés. Pour chaque pixel $\boldsymbol{l}$:

$$K_j(\boldsymbol{l} - \mathbf{c}_j) = \begin{cases} C\left(1 - \frac{\|\boldsymbol{l}-\mathbf{c}_j\|^2}{h^2}\right) & \text{si } \|\frac{\boldsymbol{l}-\mathbf{c}_j}{h}\| \leqslant 1 \\ 0 & \text{sinon} \end{cases} \tag{10}$$

h représente l'étendue du noyau.
D'après (7), la dépendance du noyau vis-à-vis des paramètres de l'état apparaît en écrivant pour chaque pixel $\boldsymbol{l}$:

$$K_j(\boldsymbol{l}, \mathbf{c}_j, \mathbf{x}) = C\left(1 - \frac{\|f_{\mathbf{x}}(\boldsymbol{l}) - \mathbf{c}_j)\|^2}{h^2}\right)$$

où $f_{\mathbf{x}}(\boldsymbol{l})$ est défini dans (1).
La procédure d'optimisation permet une bonne précision, mais le suivi échoue en cas d'occultation ou en présence de plusieurs modes dans la fonction de coût. En particulier, pour des mouvements rapides dans lesquels des effets de flou (voir Figure 4) lissent l'apparence de l'objet, l'optimisation tend plus souvent à converger vers un mauvais minimum, le rectangle rétrécissant vers une portion plus petite de l'objet comme illustré Figure 5.
Les résultats obtenus avec cette méthode sont présentés dans la deuxième ligne de la Figure 7 pour un objet mobile de taille fixe.
La partie suivante présente l'approche par filtrage particulaire considérée pour la configuration multi-noyaux choisie.

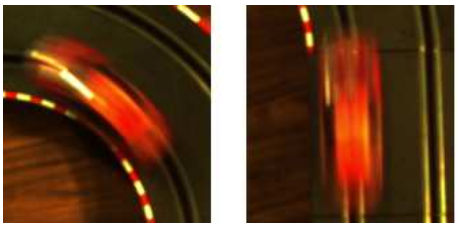


FIG. 4 – Flou dû au mouvement rapide de l'objet.

4 Cadre bayésien

4.1 Principe du filtrage particulaire

L'utilisation du cadre bayésien, et plus particulièrement du filtrage particulaire, dans les applications reposant sur la vision a été largement décrit dans la littérature (ex. [10]). Les principales idées sont rappelées ici.
La question de l'estimation de l'état $\mathbf{x}_k$ étant données des observations $\mathbf{z}_{1:k}$, peut être considérée de façon équivalente comme l'estimation de la fonction de densité de probabilité $p(\mathbf{x}_k \mid \mathbf{z}_{1:k})$, où $\mathbf{z}_{1:k}$ représente $(\mathbf{z}_1, ..., \mathbf{z}_k)$. Supposant que ces observations sont indépendantes et que le système est Markovien, on a :

$$\begin{aligned} p(\mathbf{x}_k \,|\, \mathbf{z}_{1:k}) &\propto p(\mathbf{z}_k \,|\, \mathbf{x}_k) p(\mathbf{x}_k \,|\, \mathbf{z}_{1:k-1}) \\ &\propto p(\mathbf{z}_k \,|\, \mathbf{x}_k) \int p(\mathbf{x}_k \,|\, \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \,|\, \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}. \end{aligned}$$

De là, les étapes du filtre non linéaire optimal récursif sont déduites :

$$p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) \overset{\text{évolution}}{\longrightarrow} p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) \overset{\text{mise à jour}}{\longrightarrow} p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \tag{11}$$

– **Evolution :** en utilisant un modèle d'évolution de l'état, une densité *a priori* $p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})$ est déduite de la densité précédente $p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1})$

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}. \tag{12}$$

– **Mise à jour :** la densité *a posteriori* est calculée en utilisant la densité *a priori* et la vraisemblance $p(\mathbf{z}_k \mid \mathbf{x}_k)$ de la nouvelle observation $\mathbf{z}_k$

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k \mid \mathbf{z}_{1:k-1})}. \tag{13}$$

Dans le cas d'un système linéaire et de densités gaussiennes, les équations ci-dessus admettent une solution explicite donnée par le filtre de Kalman. Cependant, la solution analytique de (12) et (13) reste inconnue dans le cas général, et des méthodes d'approximation sont nécessaires. Une approximation possible consiste à linéariser le système et ainsi obtenir un filtre de Kalman étendu. Cette solution n'est pas appropriée pour l'application visée ici, car le système peut être non linéraire ou comporter plusieurs modes.
L'idée du filtrage particulaire (également connu sous le terme CONDENSATION [10]) est d'approcher $p(\mathbf{x}_k \mid$

$\mathbf{z}_{1:k}$) en la représentant par un ensemble fini $\{s_k^{(i)}\}_{i=1}^N$ de N échantillons, ou particules, associés à des poids $\{w_k^{(i)}\}_{i=1}^N$:

$$p(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \simeq p^N(\mathbf{x}_k \mid \mathbf{z}_{1:k}) \triangleq \sum_{i=1}^{N} w_k^{(i)} \delta_{s_k^{(i)}}(\mathbf{x}_k) \quad (14)$$

où $w_k^{(i)} \geq 0$ et $\sum_{i=1}^N w_k^{(i)} = 1$. Chaque particule représente ainsi un état possible de l'objet.

Les étapes récursives du filtre (12) et (13) ci-dessus conduisent maintenant à l'algorithme suivant :

Etant donné l'ensemble de N particules $\left\{(s_{k-1}^{(i)}, \frac{1}{N})\right\}_{i=1..N}$ au rang $k-1$:

- **Evolution** des particules suivant un modèle de mouvement, donnant un nouvel ensemble : $\left\{(s_k^{'(i)}, \frac{1}{N})\right\}_{i=1..N}$.
- **Mise à jour :** en utilisant l'observation $\mathbf{z}_k$ le poids de chaque particules prédite est calculé $w_k^{(i)} \propto p(\mathbf{z}_k \mid \mathbf{x}_k = s_k^{(i)})$. Un nouveau jeu de particules et alors obtenu : $\left\{(s_k^{'(i)}, w_k^{(i)})\right\}_{i=1..N}$, avec $\sum_{i=1}^N w_k^{(i)} = 1$.
- **Tirage aléatoire pondéré** de N particules de $\left\{(s_k^{'(i)}, w_k^{(i)})\right\}_{i=1..N}$, menant à un nouvel ensemble : $\left\{(s_k^{(i)}, \frac{1}{N})\right\}_{i=1..N}$.

Dans les expérimentations, un modèle à vitesse constante a été utilisé pour l'étape d'évolution (ou prédiction). La sortie du filtrage particulaire considérée ici est l'estimateur de l'espérance : $E[\mathbf{x}_k] = \frac{1}{N}\sum_{n=1}^N s_k^{(i)}$. D'autres estimateurs peuvent être utilisés, comme la meilleure particule (en terme de vraisemblance), ou la moyenne dans une portion réduite de l'espace d'état. Cependant, le premier conduit à un bruit plus important et le deuxième est plus long à calculer. La fonction de vraisemblance est décrite dans la partie suivante.

4.2 Fonction de vraisemblance

Dans le cadre du filtrage particulaire, la fonction de vraisemblance est évaluée pour chaque particule, et l'ensemble des particules constitue un échantillonnage de la fonction de densité de probabilité à estimer. Par conséquent, cette méthode ne requiert pas le calcul du gradient comme c'est le cas pour les méthodes d'optimisation classiques ([4], [8]). Dans cette partie, *K* est fixé égal à une constante, de sorte que les histogrammes soient normalisés. L'information spatiale est obtenue en calculant neuf histogrammes dans les neuf sous-régions définies dans la Figure 2 (et donc utilisant chacun un neuvième des n pixels de l'objet) comme dans [5].

Plus formellement, un état $\mathbf{x}_k$ se voit associer les histogrammes $\{\mathbf{q}_j(\mathbf{x}_k)\}_{j=1..9}$ correspondants. La vraisemblance de l'état $\mathbf{x}_k$ est alors définie par :

$$p(\mathbf{z}_k \mid \mathbf{x}_k) \propto exp(-\lambda {d_m}^2(\mathbf{x}_k)) \quad (15)$$

où d_m est donné par (5) et λ est un paramètre constant réglé empiriquement (une valeur typique est $\lambda = 20$).

La Figure 5 présente les résultats comparatifs entre un filtre particulaire à noyau unique (1^e ligne), l'approche déterministe multi-noyaux (2^e ligne) et le filtre particulaire multi-noyaux (dernière ligne) pour une séquence avec un mouvement rapide de l'objet. Comme attendu, l'approche multi-noyaux est plus précise que celle à noyau unique. Dans cette séquence, le manque d'information dû à l'effet de flou fait converger la méhode d'optimisation vers un mauvais minimum, tandis que le filtrage particulaire multi-noyaux permet une bonne localisation.

5 Combinaison des deux approches

Arrivé à ce point, il est intéressant d'observer les aspects complémentaires des deux approches considérées. Dans l'optimisation multi-noyaux déterministe, quelques itérations permettent d'atteindre le minimum de la fonction de coût avec une bonne précision. Cependant, la nature déterministe de l'approche la rend inapte en cas d'occultation totale. De plus, l'algorithme peut tomber dans un minimum local.

Le filtrage particulaire permet bien sûr le filtrage temporel de l'état, et donc de tirer parti des informations antérieures. Indépendemment de cet aspect, il permet également d'estimer la densité de probabilité sans aucune hypothèse préalable sur sa forme, ce qui permet notamment de garder plusieurs modes possibles. Cependant, la précision dépend linéairement du nombre de particules (et donc du temps de calcul). Il est donc nécessaire de trouver le bon équilibre entre précision et temps de calcul.

Pour tirer parti des avantages de chacune de ces méthodes, il est intéressant d'étudier un moyen de les combiner. Dans [7], un filtrage particulaire avec *mean shift* intégré est proposé, dans lequel l'optimisation par *mean shift* est appliquée à chaque particule pour générer des particules dans les maxima locaux de la fonction de densité de probabilité. Cela permet d'utiliser moins de particules en conservant une bonne précision de localisation. Cependant, dans le cas du suivi multi-noyaux, réaliser l'optimisation à partir de chaque particule serait trop long. Par conséquent, nous avons choisi d'appliquer la procédure d'optimisation directement sur l'estimée fournie par le filtre particulaire. Ainsi, la diminution du nombre de particules introduit un manque de précision compensé par la procédure d'optimisation qui trouve le mode de la fonction de distance. Bien sûr, pour obtenir des résultats satisfaisants il est nécessaire que l'estimée du filtre soit suffisamment proche du mode en question, mais les expériences ont montré que cette condition pouvait être obtenue en pratique avec un nombre réduit de particules.

Pour éviter les échecs dus aux occultations temporaires, un seuil sur la fonction de distance permet de les détecter et de changer pour l'estimée du filtre particulaire. Si $\overline{\mathbf{x}}_k{}^{PF}$, et $\overline{\mathbf{x}}_k{}^{PFSSD}$ représentent respectivement l'estimée donnée par le filtre particulaire et l'algorithme combiné dans l'image k, la procédure peut être formulée comme suit :

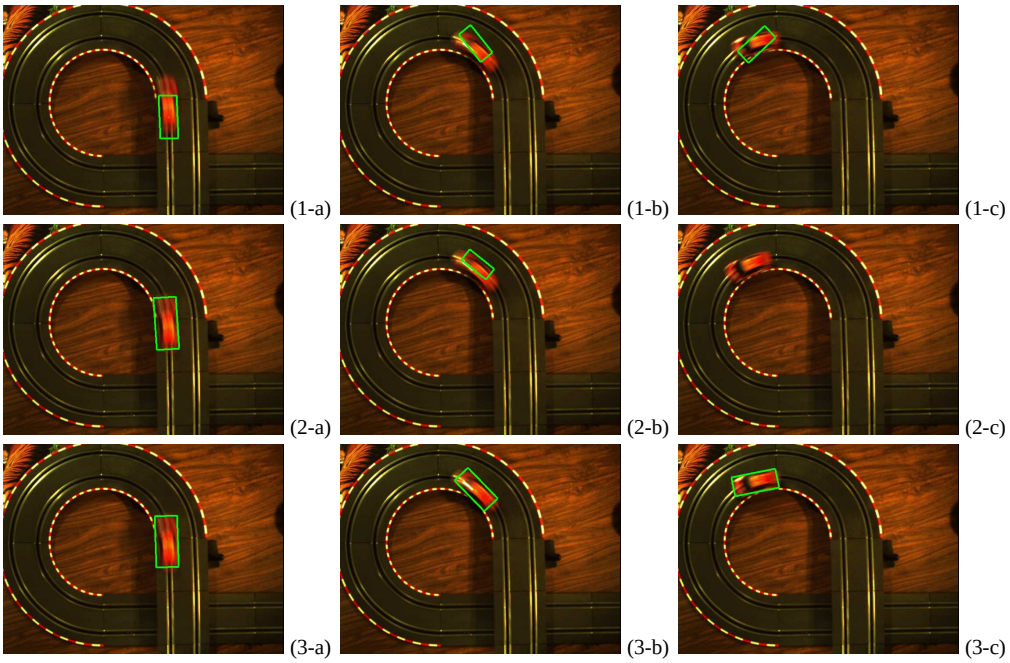


FIG. 5 – Résultats comparatifs des différentes approches avec des translations, rotations et changement d'échelle. 1e ligne : filtre particulaire avec 1 noyau, et 500 particules. 2e ligne : approche déterministe multi-noyaux [8]. 3e ligne : filtre particulaire avec 9 noyaux constants, et 500 particules.

Si $(d_m\left(\overline{\mathbf{x}}_k{}^{PF}\right) \geq \epsilon)$ **Alors**
  calculer $\overline{\mathbf{x}}_k{}^{PFSSD}$ à partir de $\overline{\mathbf{x}}_k{}^{PF}$ en utilisant l'optimisation multi-noyaux
Sinon
  $\overline{\mathbf{x}}_k{}^{PFSSD} \leftarrow \overline{\mathbf{x}}_k{}^{PF}$
Fin Si

ϵ étant un seuil à régler.
Tandis que l'optimisation seule échoue en cas de large déplacements, la combiner avec un filtrage particulaire lui permet de tirer profit de l'estimée comme étant une approximation proche de l'état réel de l'objet. En outre, cette estimée préalable permet de réduire le nombre d'itérations requises par la minimisation de type Newton.

6 Résultats comparatifs

Les figures 7 et 6 présentent les résultats comparatifs des différentes approches sur une séquence avec un déplacement relativement rapide (1-2ms^{-1}) comportant une occultation complète, pour un objet de taille fixée. Un filtrage particulaire à noyau unique est testé, donnant des résultats insuffisants notamment dans l'estimation de la rotation de l'objet (voir Figure 6 (1) et première ligne de la Figure 7). La méthode d'optimisation seule est précise, mais de manière prévisible, elle échoue lors de l'occultation (Figure 7 (2-c), ce qui correspond à la rupture qui apparaît dans l'image 46 de la Figure 6 (2)).
Le filtre particulaire avec 75 particules ((4) dans les Figures 7 et 6) est robuste à l'occultation mais moins précis qu'avec

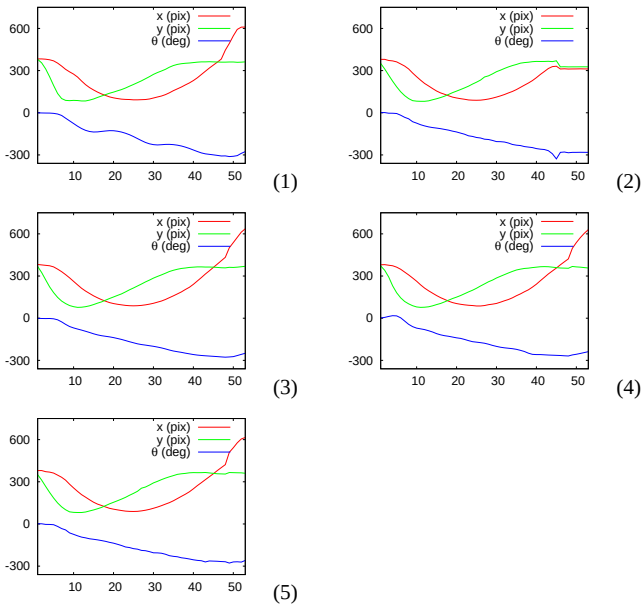


FIG. 6 – Position et orientation de l'objet dans la séquence de la Figure 7 obtenue en utilisant : (1) un filtre particulaire à noyau unique avec 500 particules, (2) l'approche déterministe multi-noyaux [8], (3) un filtre particulaire multi-noyaux avec 500 particules, (4) un filtre particulaire multi-noyaux avec 75 particules, (5) l'approche combinée avec 75 particules. L'abscisse indique le numéro de l'image dans la séquence.

500 particules (3). Les résultats expérimentaux montrent que l'approche combinée (5) atteint la même précision que le filtre particulaire seul (3), en utilisant beaucoup moins de particules, ce qui la rend plus efficace du point de vue du temps de calcul puisqu'elle tourne 3 fois plus vite environ en pratique.

7 Remerciements

Ce travail a été réalisé dans le cadre du projet ANR SCUAV (ANR Psirob 06_174032 SCUAV ref ANR-06-ROBO-0007-02).

8 Conclusions

Pour suivre des objets en mouvement, dont l'apparence peut varier au cours du déplacement, les représentations d'objets sous forme de densités de probabilité telles que les histogrammes de couleurs constituent un choix attractif. Cependant, dans le cas de mouvements complexes, les méthodes à noyau unique se révèlent insuffisantes. L'utilisation d'une représentation multi-noyaux de l'objet est dès lors nécessaire pour rendre le suivi sensible à un large panel de mouvements. En combinant les approches multi-noyaux d'optimisation et de filtrage particulaire, nous avons proposé ici un algorithme de suivi robuste permettant d'estimer la position, l'orientation et la taille d'un objet en mouvement. L'approche proposée peut faire face à des mouvements rapides en utilisant environ 75% de particules en moins.

Références

[1] G. Hager, P. Belhumeur. "Efficient Region Tracking With Parametric Models of Geometry and Illumination", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20(10), pp. 1025-1039, 1998.

[2] F. Jurie, M. Dhome. "Hyperplane Approximation for Template Matching", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24(7), pp. 996-1000, 2002.

[3] E. Malis, S. Benhimane. "Homography-based 2D Visual Tracking and Servoing", *The International Journal of Robotics Research*, vol. 26(7), pp. 661-676, 2007.

[4] D. Comaniciu, V. Ramesh, and P. Meer, "Real-Time tracking of non-rigid objects using mean shift',' *Conference on Computer Vision and Pattern Recognition*, pp. 142-149, May 2000.

[5] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-Based Probabilistic Tracking", *7th European Conference on Computer Vision*, pp. 661-675, May 2002.

[6] K. Nummiaro, E. Loller-Meier, and L. Van Gool. "An adaptive color-based particle filter", *Image and Vision Computing*, vol.21, pp. 99-110, 2003.

[7] C. Shan, T. Tan, and Y. Wei, "Real-time hand tracking using a mean shift embedded particle filter", *Pattern Recognition*, vol. 40, pp. 1958-1970, 2007.

[8] G. Hager, M. Dewan, and C. Stewart, "Multiple kernel tracking with SSD", *Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 790-797, 2004.

[9] R. Megret, M. Mikram, and Y. Berthoumieu, "Inverse Composition for Multi-kernel Tracking", *In proc. of Computer Vision, Graphics, and Image Processing, LNCS 4338, Springer*, pp. 480-491, 2006.

[10] M. Isard, and A. Blake, "CONDENSATION : conditional density propagation for visual tracking", *Journal of Computer Vision*, vol. 29, pp. 5-28, 1998.

[11] Z. Zivkovic, B. Krose, "An EM-like algorithm for color-histogram-based object tracking", *Computer Vision and Pattern Recognition*, vol.1, pp. 798-803, 2004.

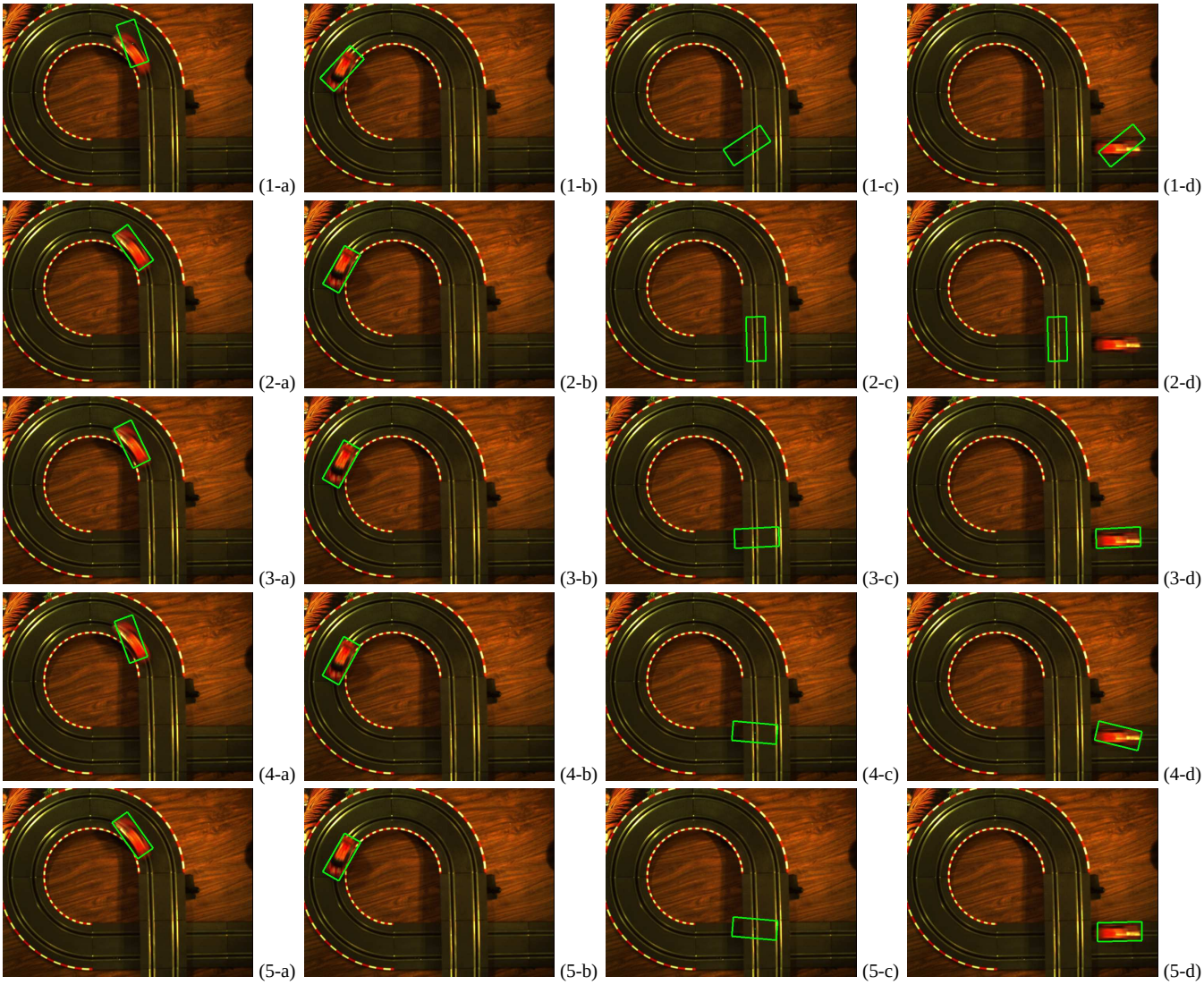


FIG. 7 – Résultats comparatifs des différentes approches avec des translations et des rotations. 1e ligne : filtre particulaire à noyau unique, avec 500 particules. 2e ligne : approche déterministe multi-noyaux [8]. 3e ligne : filtre particulaire multi-noyaux, avec 500 particules. 4e ligne : filtre particulaire multi-noyaux, avec 75 particules. 5e ligne : approche combinée avec 75 particules.