

Guidage du filtrage particulaire par une étude du mouvement

Elodie BICHOT, Laurent MASCARILLA, Pierre COURTELLEMONT

Laboratoire d'Informatique – Image – Interaction
Université de La Rochelle
Av. Michel Crépeau, 17042 La Rochelle Cedex, FRANCE
elodie.bichot@univ-lr.fr, laurent.mascarilla@univ-lr.fr
pierre.courtellemont@univ-lr.fr

Résumé – Nous proposons d’exploiter le résultat d’une segmentation par étude du mouvement afin d’améliorer le suivi par filtrage particulaire. Nous générons d’abord des hypothèses du filtrage particulaire qui correspondent aux régions dont le mouvement est similaire au déplacement de la cible à l’instant précédent. Ensuite, après avoir identifié la région de mouvement correspondant à la cible, nous apprenons son modèle qui est intégré comme référence lors de la prochaine itération du suivi. Cette approche améliore les performances du suivi dans des situations complexes en comparaison avec un filtre Bootstrap conventionnel comme le montrent les expérimentations menées sur des séquences d’images d’aquarium.

Abstract – In this article, we propose to use results of a motion based segmentation to improve tracking by particle filtering. We generate firstly particle filtering hypotheses corresponding to regions of similar motion to target displacement at previous time step. Then, after identifying target motion region, target model is updated and integrated as reference for the next tracking iteration. This approach improves performances in complex situations compared with Bootstrap filter as shown our experiments on fish tank sequences.

1 Introduction

Dans le cadre d’une application automatique de reconnaissance de poissons appelée projet Aqu@thèque [5], nous devons suivre des poissons en mouvement jusqu’à ce que le système de reconnaissance puisse les identifier. Notre problème consiste donc à suivre des objets non rigides et animés de mouvements complexes en milieu perturbé. Pour ce faire, nous proposons d’exploiter le résultat d’une segmentation par analyse du mouvement en propageant les particules du filtre particulaire vers les régions de mouvement similaire à celui de la cible et en apprenant un modèle qui est ensuite intégré comme référence lors de la prochaine itération du suivi. Les résultats obtenus montrent que le suivi est maintenu dans des situations complexes d’occultations et de variations d’illumination en comparaison avec un filtre Bootstrap conventionnel. Nous présentons l’approche proposée dans la partie 2 puis analysons les résultats obtenus dans la partie 3 et finalement dressons une conclusion dans le dernier paragraphe.

2 L’approche proposée

2.1 Vue générale de l’approche de suivi proposée

Le filtrage particulaire consiste en une exploration de l’espace d’état de la cible par des particules pondérées. Les particules représentent des hypothèses sur l’état de la cible et les poids associés les probabilités des hypothèses. Dans le filtre Bootstrap, elles sont propagées suivant le modèle dynamique de la cible, pondérées selon la vraisemblance des observations conditionnellement à l’état prédit et finalement, sélectionnées proportionnellement à leur poids. La somme des fonctions de Dirac pondérées par les poids des particules et centrées en les

particules constitue une approximation de la distribution de probabilité *a posteriori* sur l’état de la cible [3]. Si la prédiction est précise et si le modèle de représentation permet de discriminer la cible de son environnement, alors cette méthode donne des résultats satisfaisants. Cependant, dans le cas de scènes réelles où il est difficile de déterminer un modèle dynamique précis et où la représentation de la cible varie en fonction de son orientation et de l’illumination, elle peut être mise en échec. Ceci justifie notre approche : nous utilisons le résultat de la segmentation pour générer des particules du filtre particulaire correspondant aux régions de mouvement similaire au déplacement de la cible et améliorer ainsi la prédiction. Ensuite, nous apprenons, à partir de la région de mouvement identifiée comme étant la cible, un modèle précis et représentatif de la cible qui est intégré comme référence dans l’étape de correction à la prochaine itération du suivi. Le schéma proposé est représenté sur la figure 1.

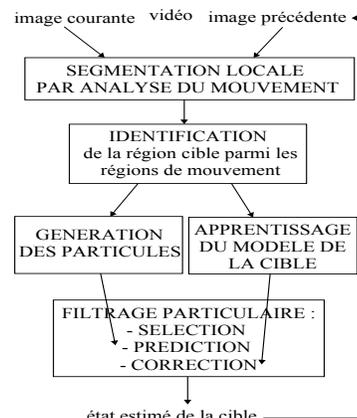


FIG. 1 – Schéma de suivi proposé.

2.2 Segmentation par analyse du mouvement

Afin de segmenter l'image au sens du mouvement, le flot optique a d'abord été estimé par une procédure de mise en correspondance sur des images couleurs. Ensuite, nous avons appliqué l'approche de classification relationnelle des C-médoïds flous sur le champ des vecteurs de mouvement et pour cela introduit une mesure de similarité entre vecteurs géométriques fondée sur l'égalité de deux vecteurs [1]. Nous obtenons une liste de $Nb_regions_mvt$ régions homogènes en mouvement :

$$Region_mvt_t^{(k)} = \left\{ BB_t^{(k)}, M_t^{(k)}, v_t^{(k)} \right\}_{k=1}^{Nb_regions_mvt}$$

- $BB_t^{(k)}$ est le rectangle optimal englobant la région k ,
- $M_t^{(k)}$ désigne le masque des pixels appartenant à la forme pour la région k ,
- $v_t^{(k)}$ est le vecteur déplacement de la région k .

Parmi ces régions, nous retenons, pour l'étape d'échantillonnage d'importance du filtrage particulaire, celles dont le mouvement maximise la similarité avec le déplacement de la cible à l'instant $t - 1$ par

$$s = \arg \max_k R(v_cible_{t-1}, v_t^{(k)}) \quad (1)$$

où v_cible_{t-1} est le déplacement de la cible à l'instant $t - 1$ et $R(., .)$ la mesure de similarité entre deux vecteurs mouvements introduite dans [1].

2.3 Filtrage particulaire fondé sur le mouvement et la couleur

Nous renforçons un filtre Bootstrap conventionnel par la conduite des particules vers les régions d'intérêt de l'espace d'état et l'apprentissage du modèle de référence de la cible à partir du résultat de la segmentation. Notre approche suppose que d'une image à l'autre le mouvement et l'apparence de la cible n'évoluent que très peu.

2.3.1 Espace d'état

L'espace d'état est défini par le rectangle englobant l'objet d'intérêt. $X_t = \{x_t, y_t, largeur_t, hauteur_t\}$ correspond à une particule, c'est-à-dire à une hypothèse de rectangle englobant. Nous désignons par

$$\left\{ X_t^{(i)}, w_t^{(i)} \right\}_{i=1}^{Nb_total} \quad (2)$$

la liste des Nb_total particules $X_t^{(i)}$ et leur poids associé $w_t^{(i)}$.

2.3.2 Modèle dynamique

Le modèle dynamique s'appuie sur les positions et vitesses estimées de la cible dans les images précédentes, la vitesse étant mise à jour à partir des positions précédentes de la cible.

$$x_t^{(i)} = x_{t-1}^{(i)} + \dot{x}_{t-1}^{(i)} + \sigma_{dyn}\omega \quad (3)$$

$$\dot{x}_t = \bar{x}_{t-1} - \bar{x}_{t-2} \quad (4)$$

$$largeur_t^{(i)} = largeur_{t-1}^{(i)} \pm 10\%largeur_{t-1}^{(i)} \quad (5)$$

où σ_{dyn} est l'écart-type du bruit de dynamique et ω un bruit blanc Gaussien de moyenne nulle. \bar{x}_t est la moyenne des positions $x_t^{(i)}$ à l'instant t . Les composantes y et \dot{y} suivent respectivement les mêmes lois que x et \dot{x} . La *hauteur* évolue de même

que la *largeur* mais indépendamment de celle-ci. Une partie des particules est propagée suivant ce modèle.

2.3.3 Echantillonnage d'importance

L'autre partie des particules est générée selon le principe d'échantillonnage d'importance. Il s'agit de diriger la recherche en utilisant une connaissance supplémentaire décrivant les parties de l'espace d'état qui contiennent le plus d'informations sur la distribution de probabilité *a posteriori* de la cible. Dans notre cas, nous utilisons le résultat de la segmentation par étude du mouvement. Plus précisément, nous générons les particules à partir des régions de mouvement cohérent avec le déplacement de la cible (et retenues après l'étape de segmentation par l'équation (1)) et d'une hypothèse de bruit blanc Gaussien pour les positions et de bruit uniforme pour les largeur et hauteur (équation 5). Nous limitons ainsi la recherche aux régions de forte vraisemblance de l'espace d'état et améliorons la prédiction. Toutes les particules, qu'elles soient propagées par le modèle dynamique ou générées selon l'échantillonnage d'importance, sont pondérées suivant le modèle de mesures que nous décrivons maintenant.

2.3.4 Modèle de mesures

Le modèle d'observation s'appuie sur la couleur dans l'espace rgb normalisé. Le choix de cet espace se justifie par son invariance à l'intensité d'illumination. L'objectif de l'étape de correction est de favoriser les histogrammes candidats similaire à l'histogramme modèle de la cible. Nous définissons alors le modèle de la vraisemblance couleur par

$$w_t^{(i)} \propto \exp(-D^2(h_{X_t^{(i)}}, h_{cible_{t-1}})/2\sigma_{obs}^2) \quad (6)$$

- $h_{X_t^{(i)}}$ est l'histogramme calculé sur le rectangle donné par l'hypothèse $X_t^{(i)}$,
- $h_{cible_{t-1}}$ est l'histogramme de la cible à l'instant $t - 1$,
- $D(h1, h2) = (1 - \sum_{j=1}^b \sqrt{h1(j)h2(j)})^{1/2}$ est la distance de Bhattacharyya entre les histogrammes $h1$ et $h2$,
- b désigne le nombre de classes de l'histogramme,
- σ_{obs} est l'écart-type du bruit d'observation.

Une fois pondérées, les particules sont rééchantillonnées proportionnellement à leur poids. L'état de la cible est, finalement, estimé par la moyenne des particules

$$\hat{X}_t = \frac{1}{Nb_total} \sum_{i=1}^{Nb_total} X_t^{(i)} \quad (7)$$

2.3.5 Mise à jour du mouvement et de l'apparence de la cible

Afin de préparer l'étape suivante de sélection des régions de mouvement d'intérêt, nous mettons à jour le déplacement de la cible selon l'équation

$$v_cible_t = v_t^{(s)} \quad (8)$$

où s désigne les régions retenues par l'équation (1). Puis, nous recherchons parmi les régions de mouvement la région dont

l'apparence maximise la similarité avec celle de la cible à l'instant $t - 1$:

$$u = \arg \min_k D(h_{cible_{t-1}}, h_{M_t}^{(k)}) \quad (9)$$

où D désigne la distance de Bhattacharyya et $h_{M_t}^{(k)}$ l'histogramme construit à partir du masque de la région k . Nous mettons alors à jour l'histogramme couleur de la cible à partir du masque de la forme identifiée ci-dessus :

$$h_{cible_t} = h_{M_t}^{(u)} \quad (10)$$

En construisant l'histogramme à partir du masque de la forme, nous n'intégrons pas de faux pixels (pixels appartenant au fond de l'image ou à d'autres objets) et bénéficions ainsi d'un modèle représentatif de la cible pour l'étape suivante.

3 Expérimentations

Nous appliquons notre approche et un filtre Bootstrap conventionnel sur des séquences d'aquarium. L'initialisation du suivi s'effectue manuellement selon le projet Aqu@thèque [5] à partir des images 219 pour la première cible et 163 pour la seconde. L'écart-type du bruit de dynamique σ_{dyn} est de 5 pixels et celui du bruit σ_{obs} du modèle d'observation fondé sur la distance de Bhattacharyya est égal à 0.1. Dans le filtre Bootstrap, 150 particules sont utilisées. Dans notre approche, 100 particules sont propagées suivant le modèle dynamique et 50 générées à partir de la carte de segmentation. Différents vidéos résultats sont disponibles à l'adresse http://13iexp.univ-lr.fr/magnoliarico/aquathequedemo/prb_scientifique/suivi2.html.

Pour la première séquence représentée en figure 3, nous nous intéressons au poisson porc encadré dans l'image 223. Ce poisson est occulté par un chirurgien bleu de l'image 243 à l'image 265. Au moment de l'occultation se trouve dans son voisinage un autre poisson porc de même espèce donc présentant une distribution couleur similaire. Le suivi effectué par le filtre Bootstrap conventionnel dérive du poisson porc cible, au moment de l'occultation par le chirurgien bleu, pour se concentrer sur l'autre poisson porc. La cible est définitivement perdue comme le montrent les images de la première ligne figure 3. En revanche, avec notre approche, les particules générées à partir de la segmentation et représentées par les hypothèses de rectangles dessinés en blanc (dernière ligne figure 3) dirigent la recherche vers la région de mouvement cohérent à l'évolution de la cible. Et, le suivi du poisson d'intérêt est maintenu pendant l'occultation. De plus, son état est correctement estimé tout au long de la séquence. L'information de couleur ne discrimine pas les deux poissons de la même espèce comme l'illustrent les deux modes de l'image de la vraisemblance couleur de la figure 2. Par ailleurs, une segmentation par détection du mouvement comme dans [2] ou [4] ne nous permet pas de suivre le poisson d'intérêt car elle ne lève pas l'ambiguïté entre les régions de mouvement correspondant aux poissons porcs qui se déplacent (image de droite de la figure 2). En revanche, l'étude du mouvement offre la possibilité de distinguer ces trois régions car elles sont animées de mouvements différents (seconde ligne figure 3).

L'objectif de la seconde expérimentation (figure 4) est de suivre le chirurgien bleu qui évolue d'abord rapidement puis se retourne et finalement se déplace plus lentement dans une

zone faiblement illuminée (empêchant en outre sa détection par différence d'images (image de droite figure 2)). Avec un filtre Bootstrap conventionnel, l'état de la cible n'est pas toujours estimé correctement comme sur les images 208 et 264 même si elle n'est jamais définitivement perdue. En revanche, avec notre approche, les positions et tailles sont évaluées de manière satisfaisante quelles que soient les conditions d'illumination grâce à la mise à jour du modèle de référence. La trajectoire estimée par notre approche, de par sa régularité, illustre également l'amélioration apportée en comparaison avec celle donnée par le filtre Bootstrap.

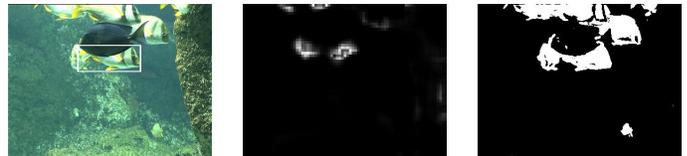


FIG. 2 – De gauche à droite : image 251 de la séquence Aqu@thèque analysée avec le poisson d'intérêt encadré en blanc. Vraisemblance couleur représentée sur une grille sous échantillonnée en fonction de la position : les régions ayant une forte vraisemblance de correspondre à la cible sont représentées en blanc. Détection du mouvement par différence d'images avec une image de référence [5].

4 Conclusion

Notre contribution consiste à guider le suivi en exploitant le résultat d'une segmentation par analyse du mouvement en deux temps : nous générons des particules dans les régions de forte vraisemblance de l'espace d'état et nous mettons à jour le modèle d'apparence de la cible afin qu'il reflète les variations liées à l'environnement. L'approche proposée améliore le suivi en présence d'occultations et de variations d'illumination comparé à un filtre Bootstrap conventionnel. Dans le contexte du projet Aqu@thèque et en cas d'échec de la reconnaissance, nous envisageons de suivre le poisson d'intérêt et de ne lancer l'identification que lorsque l'orientation et l'illumination du poisson sont favorables.

Références

- [1] E. Bichot, L. Mascarilla and P. Courtellemont. Motion based Segmentation to improve Tracking of Non Rigid Objects, *Proceedings of the 3rd IEEE Conference on Sciences of Electronic, Technologies of Information and Telecommunications*, Tunisie, 2005.
- [2] D. J. Bullock and J. S. Zelek. Real-Time Tracking for Visual Interface Applications in Cluttered and Occluding Situations, *Image and Vision Computing* 22, pages 1083-1091, 2004.
- [3] C. Hue. Méthodes séquentielles de Monte-Carlo pour le filtrage non linéaire multi-objets dans un environnement bruité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D, *Thèse*, 2003.
- [4] P. Pérez, J. Vermaak and A. Blake. Data Fusion for Visual Tracking With Particles, *Proceedings of the IEEE*, vol. 92, no 3, 2004.
- [5] D. Semani, T. Bouwmans, C. Frélicot and P. Courtellemont. Automatic Fish Recognition in Interactive Live Videos, *Int. Workshop on Interactive Video between Research challenges and Industrial Applications (IVRCIA)*, XIV :94-99, 2002.

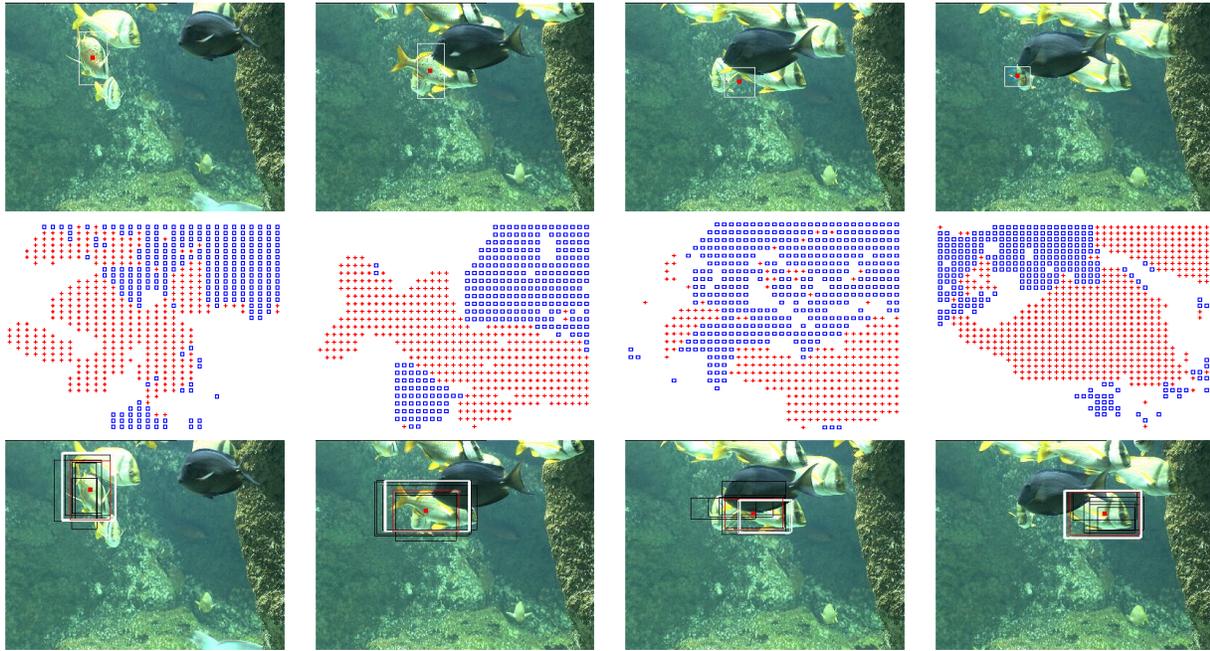


FIG. 3 – Images 223, 243, 251, 258 d’une séquence Aqu@thèque. Première ligne : résultats du suivi avec un filtre Bootstrap conventionnel. Le rectangle blanc représente l’état estimé de la cible et le carré rouge la position estimée. Deuxième ligne : carte de segmentation par étude du mouvement. La cible est représentée par des croix rouges et les autres objets en mouvement par des carrés bleus. Dernière ligne : résultat de notre schéma de suivi. En blanc : rectangles correspondant aux hypothèses générées à partir de la carte de segmentation. En noir : rectangles correspondant aux hypothèses propagées suivant le modèle dynamique. En rouge : rectangle correspondant à l’état estimé de la cible. Les particules générées à partir de la carte de segmentation dirigent la recherche vers la région cible (image 251). Le suivi est maintenu et l’état de la cible correctement estimé. Pour faciliter l’interprétation, nous n’avons représenté que quelques hypothèses sélectionnées aléatoirement.



FIG. 4 – Images 180, 208, 264, 306. Première ligne : résultats du suivi avec un filtre Bootstrap conventionnel. Dernière ligne : résultat de notre schéma de suivi. A droite : trajectoires du chirurgien bleu estimées avec l’approche Bootstrap (en haut) et notre approche (en bas).