

Filtrage conditionnel pour la trajectographie dans des séquences d'images - Application au suivi de points

Conditional filters for image sequence tracking - Application to point tracker

E. Arnaud¹

E. Mémin¹

B. Cernuschi-Frías²

¹ IRISA, Université de Rennes 1

² LIPSIRN, Universidad de Buenos Aires, Argentina

IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France
earnaud@irisa.fr

Résumé

Dans cet article, nous proposons une nouvelle formulation conditionnelle des méthodes classiques de filtrage, dédiée à la trajectographie dans des séquences d'images. Ces filtres conditionnels permettent de considérer un modèle d'état et un modèle de mesure qui dépendent de la séquence. Dans ce cadre, deux filtres ont été construits pour la trajectographie de points. Ces méthodes permettent de suivre des trajectoires qui subissent des changements abrupts et des occlusions. Elles combinent une dynamique construite sur la contrainte de flot optique, et des mesures fournies par une méthode de corrélation. Le premier filtre est linéaire, particulièrement bien adapté aux séquences présentant un mouvement dominant. Il est dérivé de l'utilisation d'un nouvel estimateur appelé estimateur linéaire conditionnel de variance minimale. Le second filtre est non linéaire, construit à partir d'un filtre particulière. Ce dernier permet de suivre des points dont le mouvement ne peut être décrit que localement.

Mots Clef

vision dynamique, filtrage, trajectographie de points, méthode séquentielle de Monte Carlo.

Abstract

In this paper, we propose a new conditional formulation of classical filtering methods dedicated to image sequence based tracking. These conditional filters allow to consider a state model and a measure model which both depend on the image sequence data. On this basis, we derive two filters for the point tracking problem, which authorize to cope with trajectories exhibiting abrupt changes and occlusions. They combine a dynamic relying on the optical flow constraint and measures provided by a matching technique. The first tracker is linear, well-suited to image sequences exhibiting global dominant motion. This filter is deduced through the use of a new estimator called

the conditional linear minimum variance estimator. The second one is a nonlinear tracker, implemented from a particle filter. This latter allows to track points whose motion may only be locally described.

Keywords

dynamic vision, filtering, point tracking, sequential Monte Carlo method.

1 Introduction

Le suivi de points dans une séquence d'images constitue un problème basique mais essentiel dans de nombreuses applications en vision par ordinateur. Paradoxalement, ce problème, qui consiste à reconstruire la trajectoire d'un point le long de la séquence, est fondamentalement difficile. En effet, contrairement au suivi de formes structurées, les seules informations sur lesquelles on peut s'appuyer sont des caractéristiques locales de luminance ou géométriques. Il est également très difficile de construire au préalable un modèle dynamique, sans avoir une connaissance *a priori* du mouvement de l'objet environnant. Ces difficultés ont amené les chercheurs à mettre en œuvre des techniques locales, basées sur des invariants géométriques et photométriques. Deux types de méthodes ont ainsi été définies.

Les premières sont des méthodes intuitives basées sur des critères de corrélation [4]. De telles techniques sont utilisées dans de nombreux domaines pour le suivi de points mais aussi pour estimer les déplacements de formes hautement déformables comme les nuages en imagerie météorologique [20] ou les écoulements fluides en imagerie PIV [1]. Malgré le fait que les fonctions de similarité utilisées ne soient pas invariantes aux transformations géométriques (changement d'échelle, rotation, distorsion perspective) et photométriques de l'image, ces méthodes restent très employées pour leur simplicité. Cependant, en cas de transformations trop importantes, ou en cas d'occlusions, l'efficacité de ces méthodes décroît significativement.

Les méthodes différentielles représentent le second type d'approches. Elles sont construites à partir d'une formulation différentielle d'un critère de corrélation. L'exemple le plus représentatif et le plus connu est l'algorithme de Shi-Tomasi-Kanade [21], dérivé à partir de l'hypothèse de conservation de luminance dans un voisinage du point. Cependant, ces méthodes de suivi restent sensibles aux changements d'illumination. Pour résoudre ce problème, la solution la plus répandue consiste à inclure des paramètres photométriques et/ou de contrastes [13]. D'autres adaptations ont été proposées pour améliorer la qualité des résultats et pour évaluer si le point a été suivi avec succès, telles que l'utilisation de règles de rejets robustes [23].

Un moyen classique pour augmenter la robustesse des méthodes de suivi aux points aberrants et aux occlusions consiste à avoir recours à des techniques de filtrage stochastique. On peut citer par exemple le filtre de Kalman dans le cas linéaire gaussien [18, 15, 17], ou des méthodes séquentielles d'approximation de Monte Carlo dans le cas non linéaire [11].

Le problème de filtrage est formulé comme suit. On considère un processus d'état caché de Markov discret $\mathbf{x}_{0:n} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$, d'équation de transition $p(\mathbf{x}_k | \mathbf{x}_{k-1})$. Soit $\mathbf{z}_{1:n} = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n\}$ la suite de mesures de l'état, de distributions marginales $p(\mathbf{z}_k | \mathbf{x}_k)$. On suppose ces mesures indépendantes conditionnellement à la séquence d'états. Les filtres stochastiques donnent alors des procédures pour estimer l'état du système conditionnellement à l'ensemble des mesures passées. Plus précisément, à chaque instant, le problème de filtrage consiste à construire une approximation précise de la densité de probabilité a posteriori de l'état $\mathbf{x}_k$ connaissant l'ensemble des mesures $\mathbf{z}_{1:k}$. Ce problème peut être résolu de façon exacte en deux étapes par une solution bayésienne récursive appelée filtre optimal :

- L'étape de prédiction permet d'obtenir une première approximation du nouvel état, conditionnellement à l'ensemble des informations disponibles. En supposant $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1})$ connu, on obtient :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1}$$

- La prise en compte de la nouvelle mesure $\mathbf{z}_k$ permet de corriger cette première approximation en utilisant l'équation d'observation. Cette étape est appelée étape de mise à jour :

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1})}{\int p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) d\mathbf{x}_k}$$

Le problème de la mise en œuvre de cette solution optimale est l'obtention d'une formulation calculable de ces deux intégrales. Dans le cas de modèles linéaires gaussiens, le filtre de Kalman [2] fournit la solution optimale en terme d'équations récursives de la moyenne et de la covariance de la distribution gaussienne $p(\mathbf{x}_k | \mathbf{z}_{1:k})$. Dans le cas non linéaire, une évaluation directe (sans linéarisation)

de l'espérance conditionnelle est souvent impossible. Les méthodes séquentielles de Monte Carlo (ou filtrage particulaire) [3, 7, 8] fournissent alors des approximations précises. Ces techniques consistent à approximer $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ par une somme pondérée finie de diracs centrées en des éléments de l'espace d'état appelés particules. Les poids correspondant sont définis en utilisant le principe d'échantillonnage pondéré. Le nuage de particules pondérées est alors mis à jour de façon récursive.

De notre point de vue, reconstruire la trajectoire d'un attribut dans une séquence d'images peut nécessiter, dans certains cas, une formulation différente du problème de filtrage. En particulier pour le suivi d'entités ponctuelles ou non connues à l'avance, il peut être intéressant d'avoir un cadre de filtrage qui permette la prise en compte d'un modèle entièrement estimé sur la séquence. Celui-ci sera présenté dans la section suivante. Deux filtres seront alors définis, pour traiter d'une part les systèmes linéaires et les systèmes non linéaires d'autre part. Ces deux méthodes de trajectographie ont été appliquées au suivi de points dans des séquences d'images. Enfin, la dernière section présente les résultats obtenus pour cette application et les compare à des algorithmes existants.

2 Filtrage pour la trajectographie dans des séquences d'images

2.1 Motivations

Comme nous l'avons mentionné en introduction, reconstruire la trajectoire d'un attribut dans une séquence d'images peut nécessiter, dans certains cas, une formulation différente du problème classique de filtrage. Un premier problème vient du choix du modèle d'observation. En effet, dans la mesure où l'ensemble des informations est contenu dans la séquence d'images, le cas idéal serait de considérer chaque image comme une mesure. Cependant, ces dernières ont des structures trop grandes et trop complexes pour être utilisées directement. Un choix raisonnable consiste alors à s'appuyer sur une information condensée obtenue à partir de la séquence. L'extraction d'une telle mesure est souvent faite par une fonction hautement non linéaire par rapport aux pixels de l'image.

Une autre source de difficultés vient de la définition d'un modèle dynamique approprié. Ces modèles sont la plupart du temps définis *a priori* [12, 19] ou obtenus par apprentissage [11]. Dans le cas de suivi d'attributs dont les trajectoires subissent des changements abrupts, des occlusions ou obéissent à des lois d'évolution trop complexes, ces modèles s'avèrent limités. De plus, un tel modèle n'est pas toujours disponible [24]. En particulier, ceci est le cas pour la trajectographie d'entités ponctuelles ou arbitraires. Dans ce contexte, la seule possibilité consiste à s'appuyer sur un modèle dynamique estimé à partir de la séquence d'images. De tels choix conduisent au final à un problème de filtrage pour lequel le système complet (mesure, modèle d'observation et modèle d'état) dépend des images. Pour ce cas

particulier, nous proposons dans cet article une formulation modifiée des approches classiques de filtrage. Ces filtres, appelés filtres *conditionnels* sont des extensions des filtres classiques, qui incluent un conditionnement par rapport à la séquence. Ce formalisme est en fait une généralisation du problème de trajectographie dans des séquences d'images. En effet, les systèmes construits sans aucune information *a priori* mais également les systèmes comprenant une dynamique connue au préalable peuvent être résolus dans ce cadre.

Dans ce cadre, nous proposons deux filtres. Le premier est un filtre linéaire construit à partir d'un estimateur déduit de l'estimateur linéaire de variance minimale. Ce nouvel estimateur est appelé l'*estimateur linéaire conditionnel de variance minimale*, puisqu'il introduit un conditionnement par rapport à la séquence. Le second filtre est un filtre non linéaire construit sur le principe du filtrage particulaire.

2.2 Notations

Les notations utilisées sont les suivantes : on note $\mathbf{I}_k$ une image obtenue au temps k . La séquence $\mathbf{I}_{0:n}$ représente la suite finie de variables aléatoires $\{\mathbf{I}_k, k = 0, \dots, n\}$. Connaissant une réalisation de $\mathbf{I}_{0:k}$, notre problème de trajectographie est modélisé par les équations dynamique et de mesure suivantes :

$$\begin{aligned}\mathbf{x}_k &= f_k^{\mathbf{I}_{0:k}}(\mathbf{x}_{k-1}, \mathbf{w}_k^{\mathbf{I}_{0:k}}), \\ \mathbf{z}_k &= h_k^{\mathbf{I}_{0:k}}(\mathbf{x}_k, \mathbf{v}_k^{\mathbf{I}_{0:k}}).\end{aligned}$$

A chaque temps k , une réalisation de $\mathbf{z}_k$ est fournit par un processus d'estimation basé sur la séquence d'images $\mathbf{I}_{0:k}$. Aucune hypothèse n'est faite sur la forme des fonctions $f_k^{\mathbf{I}_{0:k}}$ et $h_k^{\mathbf{I}_{0:k}}$. Ces fonctions peuvent être non linéaires et estimées sur $\mathbf{I}_{0:k}$. Le bruit d'état $\mathbf{w}_k^{\mathbf{I}_{0:k}}$ et le bruit de mesure $\mathbf{v}_k^{\mathbf{I}_{0:k}}$ peuvent également dépendre de $\mathbf{I}_{0:k}$, et ne sont pas nécessairement gaussiens. Les distributions de probabilité sont supposées être de la forme suivante :

$$\begin{aligned}p(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:n}) &= p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:n}), \\ p(\mathbf{z}_k | \mathbf{x}_{0:k}, \mathbf{z}_{1:k-1}, \mathbf{I}_{0:n}) &= p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:n}).\end{aligned}$$

Par analogie avec le cadre classique du filtrage, conditionnellement à la séquence, l'hypothèse markovienne ainsi que l'hypothèse d'indépendance conditionnelle des observations sont maintenues. De plus, on suppose que l'état et la mesure au temps k sont indépendantes de $\mathbf{I}_{k+1:n}$. Les équations du filtre optimal peuvent être appliquées au modèle proposé. La distribution *a posteriori* recherchée est alors $p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. En supposant $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1})$ connu, la solution bayésienne récursive s'écrit naturellement:

$$\begin{aligned}&p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) \\ &= \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) d\mathbf{x}_{k-1},\end{aligned}$$

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$$

$$= \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})}{\int p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) d\mathbf{x}_k}.$$

Pour résoudre ce problème de filtrage avec un système qui dépend de la séquence, les filtres standards doivent être dérivés dans une version conditionnelle. Nous détaillerons deux instances de ce problème : une version linéaire (§3) et une version non linéaire (§4).

3 Filtre linéaire conditionnel

Dans le cas linéaire, le problème de filtrage auquel nous devons faire face est modélisé par le système suivant :

$$\mathbf{x}_k = A_k^{\mathbf{I}_{0:k}} \mathbf{x}_{k-1} + \mathbf{b}_k^{\mathbf{I}_{0:k}} + \mathbf{w}_k^{\mathbf{I}_{0:k}} \quad (1)$$

$$\mathbf{z}_k = H_k^{\mathbf{I}_{0:k}} \mathbf{x}_k + \mathbf{v}_k^{\mathbf{I}_{0:k}}. \quad (2)$$

Rappelons que les matrices A_k , H_k et le vecteur $\mathbf{b}_k$ peuvent dépendre de $\mathbf{I}_{0:k}$ (par souci de clarté, l'exposant $\mathbf{I}_{0:k}$ sera omis par la suite). Les variables $\mathbf{w}_k$, $\mathbf{v}_k$ sont supposées être des bruits blanc indépendants de moyenne nulle (éventuellement non gaussiens), conditionnellement à $\mathbf{I}_{0:k}$. Les variances conditionnelles correspondantes sont supposées connues (éventuellement estimées sur $\mathbf{I}_{0:k}$) et sont notées respectivement Q_k et R_k . Le filtre linéaire conditionnel (FLC) est déduit d'une extension de l'estimateur linéaire de variance minimale, adaptée à notre contexte. Cet estimateur est appelé l'*estimateur linéaire conditionnel de variance minimale*. Il permet d'obtenir une estimation des deux premiers moments de $p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. Cette description est évidemment optimale dans le cas gaussien. Nous commencerons par introduire cet estimateur ainsi que ces propriétés.

3.1 Estimateur linéaire conditionnel de variance minimale

Définition 1 Soit X, Z, W trois variables aléatoires jointes. $E_W^*[X|Z]$ dénote le meilleur estimateur de X , linéaire à Z , conditionnellement à W : $E_W^*[X|Z] = AZ + B$, avec A et B telles que $E[\|X - AZ - B\|^2 | W]$ est minimale. $E_W^*[X|Z]$ est appelé l'*estimateur linéaire conditionnel de variance minimale*.

Il doit être noté que $E_W^*[X|Z]$ n'est pas une espérance. En notant $\Sigma_{X,Z|W} = E[XZ^t | W] - E[X|W]E[Z|W]^t$, l'important résultat suivant est obtenu (voir annexe) :

$$E_W^*[X|Z] = E[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - E[Z|W]).$$

Cet estimateur a des propriétés possédées par l'estimateur linéaire de variance minimale [2], dans leur version conditionnelle. Les propriétés les plus importantes, utilisées dans la construction du FLC sont listées ci-dessous¹.

Théorème 1 (propriété à être non biaisé) Soit X, Z, W des variables aléatoires jointes, alors

$$E[X - E_W^*[X|Z] | W] = \mathbf{0}.$$

1. Les preuves, qui ne sont pas données ici par manque de place peuvent être facilement écrites.

Théorème 2 (quantités conditionnellement non corrélées)
Soit $X, Z_1, Z_2, \dots, Z_k, W$ des variables aléatoires jointes avec $Z_1, Z_2, \dots, Z_k$ non corrélées conditionnellement à W , alors

$$E_W^*[X|Z_{1:k}] = \sum_{i=1}^k E_W^*[X|Z_i] - (k-1) E[X|W].$$

Théorème 3 (changement de variables conditionnelles)
Soit X, Z, W des variables aléatoires jointes, et soit $M = C Z + D$ (avec C non singulière), alors

$$E_W^*[X|Z] = E_W^*[X|M].$$

Théorème 4 (principe d'orthogonalité) *Soit X, Z, W des variables aléatoires jointes, alors*

$$E[(X - E_W^*[X|Z]) Z^t | W] = 0.$$

3.2 Équations du filtre linéaire conditionnel

On s'intéresse à un modèle linéaire de la forme décrite précédemment (1, 2). On note $\hat{\mathbf{x}}_{k+1|k} = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1} | \mathbf{z}_{1:k}]$ et $\Sigma_{k+1|k}$ la covariance d'erreur associée. En considérant avec attention les expressions conditionnelles induites par E^* , des formulations récursives de $\hat{\mathbf{x}}_{k+1|k}$ et $\Sigma_{k+1|k}$ peuvent être construites par le biais de manipulations similaires au cas gaussien (filtre de Kalman). On obtient les équations suivantes :

$$\begin{aligned} \hat{\mathbf{x}}_{k+1|k} &= A_{k+1} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1} + \tilde{K}_k (\mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1}), \\ \Sigma_{k+1|k} &= (A_{k+1} - \tilde{K}_k H_k) \Sigma_{k|k-1} (A_{k+1} - \tilde{K}_k H_k)^t \\ &\quad + Q_{k+1} + \tilde{K}_k R_k \tilde{K}_k^t, \end{aligned}$$

où la matrice $\tilde{K}_k$ est définie en utilisant le gain de Kalman K_k :

$$\begin{aligned} \tilde{K}_k &= A_{k+1} K_k \\ &= A_{k+1} (\Sigma_{k|k-1} H_k^t) (H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1}. \end{aligned}$$

Il est possible de séparer ces équations pour distinguer l'étape de prédiction de l'étape de mise à jour.

De plus, afin de limiter le coût de mise en œuvre, il peut se révéler utile de définir une région de recherche restreinte, où le processus d'estimation de la mesure $\mathbf{z}_k$ est appliqué. Une telle région, appelée région de validation est définie comme étant la région de l'espace de mesure où la probabilité de trouver la future observation est très élevée [5]. Elle est construite à partir de la connaissance de $p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})$. Une région elliptique de concentration de probabilité est alors définie telle que :

$$valid_k = \{\mathbf{z}_k : \epsilon_k = \tilde{\mathbf{z}}_{k|k-1}^t S_k^{-1} \tilde{\mathbf{z}}_{k|k-1} \leq \gamma\},$$

où $\hat{\mathbf{z}}_{k|k-1} = H_k \hat{\mathbf{x}}_{k|k-1}$ est la mesure prédite, $\tilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$ l'innovation et $S_k = H_k \Sigma_{k|k-1} H_k^t + R_k$ sa covariance conditionnelle associée.

4 Filtre non linéaire conditionnel

Dans le cas où l'équation d'état ou l'équation de mesure est non linéaire, le problème classique de filtrage n'admet pas de solution de dimension finie : il n'est plus possible

d'avoir une expression récursive calculable de l'espérance conditionnelle de l'état à l'instant k connaissant l'ensemble des mesures jusqu'à k . Pour surmonter cette difficulté, les techniques de filtrage particulière proposent de mettre en œuvre de façon récursive une approximation de la densité a posteriori (voir [3, 7] pour une description détaillée de ces méthodes). L'algorithme de filtrage particulière consiste à propager un nuage de N particules. Chaque particule $\mathbf{x}_{0:k}^{(i)}$ correspond à une trajectoire réalisable de l'état initial $\mathbf{x}_0$. Un poids attribué à chaque trajectoire est calculé en utilisant la vraisemblance des observations jusqu'à k . La trajectoire optimale est alors obtenue par pondération du nuage de particules.

Comme dans le cas linéaire, on se concentrera sur la construction d'un filtre non linéaire conditionnel (FNLC). Il est déduit d'une méthode de filtrage particulière. La dépendance de l'ensemble du système à la séquence d'images est prise en compte via un conditionnement par rapport aux données images. Les équations correspondantes sont données par la suite.

En supposant la connaissance des distributions $p(\mathbf{x}_0)$, et $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$, $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) \forall k \geq 1$, on cherche le meilleur estimateur de la trajectoire d'un point $\mathbf{x}_{0:k}$ connaissant les données images et l'ensemble des mesures, i.e., on cherche à estimer l'espérance conditionnelle :

$$T = E[\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}] = \int \mathbf{x}_{0:k} p(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) d\mathbf{x}_{0:k}.$$

Une telle intégrale est impossible à calculer en raison de sa trop grande dimension. L'utilisation de l'échantillonnage pondéré ("importance sampling") permet d'approximer cette intégrale en introduisant une distribution de probabilité appelée fonction de proposition, pour laquelle on sait obtenir des réalisations. Elle est notée $\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$.

A partir d'un ensemble $\{\mathbf{x}_{0:k}^{(i)}\}$ de N particules tirées selon la fonction de proposition, la connaissance des N poids normalisés associés $\tilde{w}_k^{(i)}$ permet d'approximer l'espérance conditionnelle par une somme finie :

$$\hat{T} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{0:k}^{(i)} \tilde{w}_k^{(i)}.$$

Les poids non normalisés sont donnés par :

$$w_k = \frac{p(\mathbf{z}_{1:k} | \mathbf{x}_{0:k}, \mathbf{I}_{0:k}) p(\mathbf{x}_{0:k} | \mathbf{I}_{0:k})}{\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}.$$

Afin de construire une expression récursive de l'espérance conditionnelle, on suppose une forme récursive de la fonction de proposition. Cette formulation introduit une hypothèse de causalité sur les données (images et mesures) :

$$\begin{aligned} \pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k}) \\ = \pi(\mathbf{x}_{0:k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1}) \pi(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k}). \end{aligned} \quad (3)$$

Une telle expression conduit à l'équation récursive des poids :

$$w_k = w_{k-1} \frac{p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})}{\pi(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_{1:k}, \mathbf{I}_{0:k})}.$$

Néanmoins, cette approximation conduit à une augmentation de la variance des poids dans le temps [14]. Cela a pour effet, en pratique, de faire décroître rapidement le nombre de particules significatives. Pour limiter cette dégénérescence, deux méthodes ont été proposées. Nous les présentons ici dans le cadre du filtrage conditionnel.

Une première solution consiste à choisir une fonction de proposition qui minimise la variance des poids w_k conditionnellement à $\mathbf{x}_{0:k-1}$, $\mathbf{z}_{1:k}$ et $\mathbf{I}_{0:k}$. Il est alors possible de démontrer que la fonction de proposition optimale est $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$, ce qui amène à une nouvelle formulation récursive de w_k :

$$w_k = w_{k-1} p(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}). \quad (4)$$

Le problème de cette approche est qu'elle induit d'avoir une expression de $p(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$, et de pouvoir tirer selon la fonction de proposition optimale $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k})$. Pour la plupart des applications de vision, cette fonction optimale est inconnue et est identifiée au modèle d'état (i.e. $\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1})$ [11, 8]), ce qui exclut toute prise en compte des mesures pendant le processus de diffusion des particules. Nous verrons que ce n'est pas le cas dans notre application de trajectographie de points.

Une deuxième solution repose sur une méthode de rééchantillonnage qui consiste à supprimer les trajectoires ayant un faible poids normalisé et à ajouter des copies des trajectoires associées à un poids fort (voir [9] pour une description des méthodes de rééchantillonnage existantes).

Il est évident que ces deux solutions peuvent être couplées pour augmenter l'efficacité de l'algorithme.

5 Application au suivi de points

Après avoir proposé un cadre de filtrage spécifique au problème de trajectographie dans des séquences d'images, et décrit les filtres permettant de résoudre des systèmes estimés sur les images, nous abordons dans cette partie le problème du suivi de points dans une séquence d'images. Nous nous plaçons dans un contexte le plus général possible, où aucune information sur le mouvement de l'objet environnant n'est disponible. Comme nous l'avons évoqué précédemment, la seule possibilité consiste alors à s'appuyer sur une dynamique extraite à partir des images. Ce problème sera donc résolu dans le cadre du filtrage conditionnel, présenté dans les paragraphes précédent. Pour le système étudié, $\mathbf{x}_k$ représente la position du point à l'instant k dans l'image $\mathbf{I}_k$. L'état est supposé observable par une mesure $\mathbf{z}_k$, également estimée sur la séquence. On peut remarquer que, à la fois la mesure et le système sont construits sur $\mathbf{I}_{0:k}$. En effet, l'approche proposée combine une mesure apportée par une mesure de corrélation et un modèle d'état construit au moyen d'une estimation de mouvement par une méthode différentielle.

5.1 Équation d'état conditionnelle

La distribution $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$, que nous allons décrire ici, décrit le mouvement d'un point $\mathbf{x}_{k-1}$ entre les images

$k-1$ et k . Dans le but d'être réactif à chaque changement de vitesse et de direction du point, cette distribution est définie à partir d'une méthode robuste d'estimation paramétrique de mouvement [6]. Cette méthode permet d'estimer de façon fiable un modèle paramétrique 2D représentant les champs de vitesses dominant apparent sur un support donné $\mathcal{R}$. L'utilisation d'une telle méthode sur un support approprié autour de $\mathbf{x}_{k-1}$, à partir des images $\mathbf{I}_{k-1}$ et $\mathbf{I}_k$, fournit une estimation du vecteur vitesse au point $\mathbf{x}_{k-1}$. Ce vecteur s'écrit :

$$\mathbf{U}(\mathbf{x}_{k-1}) = P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_k,$$

où $P(\mathbf{x}_{k-1})$ est une matrice $2 \times p$. $\boldsymbol{\theta}_k$ est un vecteur de taille p , qui représente un vecteur paramétrique estimé à partir d'une régression robuste :

$$\boldsymbol{\theta}_k = \arg \min_{\boldsymbol{\theta}} \int_{\mathcal{R}} \rho \left(\nabla \mathbf{I}_k^t(\mathbf{s}) P(\mathbf{s}) \boldsymbol{\theta} + \frac{\partial \mathbf{I}_k}{\partial t}(\mathbf{s}) \right) d\mathbf{s},$$

ρ est une fonction de coût robuste permettant de gérer les points aberrants (zone d'occlusion, mouvement secondaire, etc.). L'équation d'état peut alors être écrite comme suit :

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + \mathbf{U}(\mathbf{x}_{k-1}) + \mathbf{w}_k \\ &= \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_k + \mathbf{w}_k. \end{aligned}$$

où $\mathbf{w}_k$ est un bruit blanc de covariance Q_k . Le principal problème d'une telle approche est de définir un support d'estimation approprié $\mathcal{R}$ pour un point donné $\mathbf{x}_{k-1}$. deux cas peuvent être distingués.

- Quand le mouvement du point à suivre correspond au mouvement dominant global, le support d'estimation est défini comme la grille d'image complète. En conséquence, l'estimation des paramètres de mouvement ne dépend pas de la position de $\mathbf{x}_{k-1}$ et l'équation dynamique est linéaire. Pour un modèle de mouvement affine, elle est exactement de la forme (1) avec A_k la matrice relative à la rotation, la divergence et le cisaillement, et $\mathbf{b}_k$ le vecteur de translation. La variable de bruit $\mathbf{w}_k$ rend compte des erreurs liées au modèle *global* de vitesse et est vraisemblablement non gaussienne. Néanmoins, $\mathbf{w}_k$ est supposé être un bruit blanc de moyenne nulle et de covariance Q_k , conditionnellement à $\mathbf{I}_{0:k}$.

- Quand le mouvement du point suivi peut être uniquement décrit par un modèle paramétrique local, le support $\mathcal{R}$ est fixé à un petit domaine centré en $\mathbf{x}_{k-1}$. Ainsi, $\boldsymbol{\theta}_k$ dépend de $\mathbf{x}_{k-1}$ et l'équation d'état considérée devient non linéaire par rapport à $\mathbf{x}_{k-1}$. Dans ce cas, la variable de bruit $\mathbf{w}_k$ représente les erreurs de modélisation *locale* du mouvement. Il sera supposé par la suite que le bruit d'état suit une distribution gaussienne centrée conditionnellement à $\mathbf{I}_{0:k}$, i.e. $p(\mathbf{w}_k | \mathbf{I}_{0:k}) = \mathcal{N}(0, Q_k)$. L'équation d'état non linéaire s'écrit alors :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_k, Q_k) \quad (5)$$

5.2 Équation de mesure conditionnelle

Alors que l'équation d'état décrit l'évolution d'un point entre deux images successives, l'équation de mesure conditionnelle $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k})$ nous permet l'obtention d'une information entre l'image initiale et l'image courante. Cette information pourra éviter une éventuelle déviation (problème de "feature drift"). Au temps k , on suppose que $\mathbf{x}_k$ est observable via un processus de mise en correspondance du point $\mathbf{x}_0$ entre les images $\tilde{\mathbf{I}}_0$ et $\tilde{\mathbf{I}}_k$. Le résultat de ce processus constitue notre mesure $\mathbf{z}_k$. $\tilde{\mathbf{I}}_0$ représente l'image de référence. Elle est définie comme étant l'image initiale $\mathbf{I}_0$, éventuellement mise à jour par recalage en cas de déformations géométriques et/ou photométriques trop importantes autour du point à suivre le long de la séquence. De nombreux critères de corrélation peuvent être utilisés pour quantifier la similarité entre le point cible et les points candidats. L'hypothèse de conservation de la luminance a simplement amené à considérer la somme des différences au carré. La mesure $\mathbf{z}_k$ est donc définie telle que :

$$\mathbf{z}_k = \arg \min_{\mathbf{z}} \underbrace{\sum_{\mathbf{y} \in \mathcal{W}} [\tilde{\mathbf{I}}_0(\mathbf{x}_0 + \mathbf{y}) - \mathbf{I}_k(\mathbf{z} + \mathbf{y})]^2}_{r_k(\mathbf{z})}. \quad (6)$$

$r_k(\mathbf{z})$ dénote le résidu calculé sur un petit voisinage $\mathcal{W}$ autour de $\mathbf{z}$. On suppose que cette mesure transporte suffisamment d'information pour pouvoir écrire que $\mathbf{z}_k = \mathbf{x}_k$, à un bruit blanc gaussien près. Cette variable de bruit modélise une erreur locale d'estimation et rend compte d'une déviation par rapport à la mesure idéale. L'équation d'observation est alors modélisée par la très simple expression suivante :

$$p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_k, R_k). \quad (7)$$

Estimation de la covariance du bruit. Une bonne estimation de la covariance du bruit de mesure R_k est essentielle pour rendre la méthode de suivi plus robuste aux mauvaises observations due à des occlusions, du bruit d'acquisition ou à des grands changements géométriques et photométriques. Dans ce but, une surface de corrélation autour de $\mathbf{z}_k$ est construite à partir des résidus $r_k(\mathbf{z})$. Comme dans [16], notre approche reprend l'idée de Singh et Allen [22], qui fait une analogie entre la distribution de probabilité d'avoir la vraie mesure et la surface de corrélation transformée en une distribution réponse :

$$\mathcal{D}_k(\mathbf{z}) = \exp(-c r_k(\mathbf{z})), \quad (8)$$

où c est un facteur de normalisation tel que $\sum_{\mathbf{z} \in \mathcal{W}(\mathbf{z}_k)} \mathcal{D}_k(\mathbf{z}) = 1$, où $\mathcal{W}(\mathbf{z}_k)$ est un support centré autour de la mesure $\mathbf{z}_k$, de taille $n \times n$. La matrice de covariance R_k associée à la mesure $\mathbf{z}_k$ est alors construite à partir de la distribution réponse (8) :

$$R_k = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix}, \quad (9)$$

avec $\sigma_{uv} = \sum_{\mathbf{z} \in \mathcal{W}(\mathbf{z}_k)} \mathcal{D}_k(\mathbf{z})(u - u_k)(v - v_k)$, $\mathbf{z}_k = (x_k, y_k)^t$ et $\mathbf{z} = (x, y)^t$ appartenant au support autour de $\mathbf{z}_k$. Une telle modélisation nous permet de définir une ellipse adaptative d'incertitude autour de l'observation. Néanmoins, on peut remarquer que les termes de variance ne peuvent pas excéder n , la taille du support. Se limiter au processus décrit ci-dessus peut amener à deux cas problématiques.

— Tout d'abord, en cas de séquences d'images très bruitées, la surface de corrélation peut présenter plusieurs pics d'amplitude inférieure au niveau du bruit. Dans ce cas, la différence entre les pics n'est pas significative. On se propose de détecter de tels pics et de les égaliser pour ne pas en favoriser un en particulier. Pour cela, on se base sur une approximation de la distribution des résidus. On suppose la conservation de la luminance entre deux points mis en correspondance, à un bruit blanc près, i.e. :

$$\mathbf{I}_k(\mathbf{z}) - \tilde{\mathbf{I}}_0(\mathbf{x}_0) \rightsquigarrow \mathcal{N}(0, \sigma),$$

où σ correspond à une variance de bruit d'acquisition fixée pour la séquence. On en déduit que pour deux régions de taille n^2 mises en correspondance, $r_k(\mathbf{z})/\sigma$ est distribué selon une loi du chi-deux à n^2 degrés de liberté. En considérant un support d'estimation suffisamment grand, on peut faire l'hypothèse suivante :

$$\sqrt{2 * r_k(\mathbf{z})/\sigma} - \sqrt{2 * n^2} \rightsquigarrow \mathcal{N}(0, 1). \quad (10)$$

Si plusieurs points de la surface de corrélation suivent cette loi (10), leur valeur sur la surface de corrélation associée est seuillée au même niveau (i.e. la valeur correspondant au résidu le plus faible). Cela signifie qu'il leur est attribué la même probabilité d'être la vraie position de mise en correspondance. Cette étape nous permet ainsi d'estimer une meilleure matrice de covariance dans le cas de surfaces de corrélation présentant plusieurs pics de faibles amplitudes.

— Le second problème intervient dans le cas où approcher la distribution de corrélation étudiée par une loi gaussienne 2D n'a pas de sens. La surface de corrélation présente alors plusieurs pics significatifs, d'amplitudes différentes. Ce cas apparaît lors d'occlusion ou de mise en correspondance de points qui se trouvent dans des zones fortement texturées. La surface correspondante $\mathcal{D}_k$ peut alors être très lisse et mieux approchée par une distribution uniforme. Afin d'éviter une mauvaise approximation, un test d'ajustement du chi-deux est réalisé. Il permet de tester si la distribution d'intérêt est mieux représentée par une loi normale ou une loi uniforme. Dans ce dernier cas, la matrice R_k est supposée diagonale, avec ses coefficients de variance fixés à l'infini.

Procédure d'observation. Pour résumer, le processus d'observation suit les étapes suivantes : la mesure $\mathbf{z}_k$ est d'abord estimée selon (6), une surface de corrélation est construite sur un support $\mathcal{W}$ centré autour de $\mathbf{z}_k$. Une partie de cette

surface est éventuellement seuillée selon (10). La distribution réponse $\mathcal{D}_k$ est alors déduite selon (8). Un test d'ajustement du chi-deux est réalisé pour tester l'adéquation de la loi uniforme sur cette distribution. Finalement, la matrice de covariance de la mesure R_k est estimée (9) dans le cas d'une adéquation à une loi gaussienne.

5.3 Algorithme final

L'algorithme complet de trajectographie d'un point est organisé selon les étapes suivantes. Tout d'abord, un test de décision pour savoir si le point initial peut être suivi avec succès est réalisé via les valeurs propres du tenseur de structure local du point [21]. Ensuite, une carte de segmentation basée sur le mouvement [10] permet de déterminer le type de méthode utilisée pour suivre le point considéré. Si le point appartient au support du mouvement dominant global, le modèle impliqué est linéaire, et la trajectoire du point est reconstruite avec le FLC. Dans le cas contraire, si le mouvement du point est décrit par une paramétrisation locale, le système est non linéaire et la méthode mise en œuvre est le FNLC. Ces deux méthodes de trajectographie sont décrites dans les paragraphes suivants.

Trajectographie d'un point avec le FLC. Pour des points se déplaçant avec le mouvement dominant, notre problème de filtrage correspond à un modèle linéaire conditionnel, avec une dynamique de la forme (1) et l'équation de mesure linéaire (7). Un résumé de l'approche correspondante est décrit fig.1.

- modèle (sur $\mathbf{I}_{0:k}$) : calculer A_k , $\mathbf{b}_k$ et Q_k
- prediction : $\hat{\mathbf{x}}_{k/k-1} = A_k \hat{\mathbf{x}}_{k-1/k-1} + \mathbf{b}_k$ $\Sigma_{k/k-1} = A_k \Sigma_{k-1/k-1} A_k^T + Q_k$
- mesure (sur $\mathbf{I}_{0:k}$) : calculer $\mathbf{z}_k$ dans $valid_k$ et estimer R_k
- mise à jour : $K_k = \Sigma_{k/k-1} (\Sigma_{k/k-1} + R_k)^{-1}$ $\hat{\mathbf{x}}_{k/k} = \hat{\mathbf{x}}_{k/k-1} + K_k (\mathbf{z}_k - \hat{\mathbf{x}}_{k/k-1})$ $\Sigma_{k/k} = (I - K_k) \Sigma_{k/k-1}$

FIG. 1 – Filtre linéaire conditionnel pour la trajectographie d'un point

Trajectographie d'un point avec le FNLC. Quand le mouvement du point ne peut être représenté que par une paramétrisation locale, le problème de filtrage est formulé par l'équation de mesure linéaire (7) et la dynamique non linéaire (5), qui implique l'utilisation du FNLC. Ce modèle particulier (bruits gaussiens et équation de mesure linéaire, équation d'état non linéaire) fournit des expressions simples pour les distributions nécessaires à l'utilisation de la fonction de proposition optimale [7]. En effet, on peut remarquer que :

$$p(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_k, R_k + Q_k),$$

ce qui induit une expression simple pour (4). En ce qui concerne la fonction de proposition, nous avons :

$$p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{m}_k(\mathbf{x}_{k-1}), C_k), \quad (11)$$

avec

$$\begin{cases} C_k = (Q_k^{-1} + R_k^{-1})^{-1} \\ \mathbf{m}_k(\mathbf{x}_{k-1}) \\ = C_k(Q_k^{-1} (\mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \boldsymbol{\theta}_k) + R_k^{-1} \mathbf{z}_k). \end{cases}$$

Une récursivité simple pour le calcul des poids et la connaissance de la fonction de proposition optimale représentent deux avantages importants de notre approche. D'une part, d'un point de vue informatique, l'algorithme se révèle très simple à mettre en œuvre puisque les distributions utiles sont gaussiennes. D'autre part, La prise en compte de la mesure dans la diffusion des particules permet, après une zone d'occlusion (période où la covariance du bruit de mesure est élevée), de retrouver la trajectoire réelle dès l'obtention d'une observation précise.

Un autre intérêt du système particulier utilisé est qu'il permet une approximation de la distribution $p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})$ nécessaire à la définition d'une zone de recherche de la mesure. En effet, en remarquant que cette distribution peut être approximée par :

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k}) \simeq \sum_i \tilde{w}_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}),$$

et grâce au résultat (11) on peut facilement avoir accès aux approximations des deux premiers moments de la distribution $p(\mathbf{z}_k | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})$. Ainsi, il est possible de définir une zone de validité $valid_k$ spécifique au cas non linéaire.

Le FNLC pour la trajectographie d'un point est résumé fig.2.

- mesure (sur $\mathbf{I}_{0:k}$) : calculer $\mathbf{z}_k$ dans $valid_k$ et estimer R_k
- modèle (sur $\mathbf{I}_{0:k}$) : pour $i = 1 : N$, évaluer $\boldsymbol{\theta}_k^{(i)}$
- prédiction : pour $i = 1 : N$, $\mathbf{x}_{0:k}^{(i)} = (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})$ avec $\mathbf{x}_k^{(i)} \sim \mathcal{N}(\mathbf{m}_k(\mathbf{x}_{k-1}^{(i)}), C_k)$
- mise à jour : pour $i = 1 : N$ $w_k^{(i)} = w_{k-1}^{(i)} p(\mathbf{z}_k \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})$ avec $p(\mathbf{z}_k \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k}) = \mathcal{N}(\mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \boldsymbol{\theta}_k^{(i)}, R_k + Q_k)$
- rééchantillonnage
- estimation de la trajectoire : évaluer $\hat{T} = \sum_{i=1}^N \mathbf{x}_{0:k}^{(i)} \tilde{w}_k^{(i)}$

FIG. 2 – Filtre non linéaire conditionnel pour la trajectographie d'un point

6 Résultats expérimentaux

Dans cette section, nous présentons les résultats obtenus par les deux méthodes de suivi proposées sur trois séquences réelles. Ces résultats sont comparés avec les résultats de la

méthode de Shi-Tomasi-Kanade (STK) et ceux d'une méthode différentielle robuste (MDR), qui correspond à notre dynamique.

La première séquence, **Corridor**, comporte 7 images (512×512 pixels). À cause des discontinuités de profondeur et des mouvements de large amplitude, elle constitue un cas limite pour une approximation par un modèle global de mouvement affine. Les trajectoires complètes fournies par les algorithmes FLC, STK et MDR sont présentées fig.3. Dans une telle séquence, nous pouvons remarquer que le STK donne de bons résultats pour deux points mais perd les autres dès la deuxième image. Au contraire, les trajectoires de tous les points sont bien reconstruites par le FLC. Nous pensons que d'une part, le modèle dynamique linéaire global utilisé permet de reconstruire des trajectoires complexes, et d'autre part, les observations fournies par une technique de corrélation permettent de corriger une éventuelle déficience de ce modèle de mouvement global. Comme nous pouvons le voir sur les résultats du MDR (fig.3), Corridor présente un mouvement dominant qui n'est pas bien adapté à l'ensemble des points à suivre (voir en particulier les points 1,2,9).

Un résultat obtenu avec le FNLC est présenté sur **Caltra**, une séquence de 40 images (190×180). Cette séquence représente le mouvement de deux boules fixées sur un cercle rigide en rotation, sur un arrière-plan encombré. Il est à noter que, afin de limiter les ressources en temps et en mémoire utilisées par le filtre particulaire, le nombre de particules a été limité à 100. La taille du support d'estimation du vecteur de paramètres de mouvement local a été fixée à une fenêtre 32×32 . Comparé aux résultats apportés par le STK et le MDR, (fig.4), le FNLC réussit à discriminer les boules du papier peint, et fournit les trajectoires exactes. De tels résultats montrent la capacité de notre méthode de suivi à traiter des trajectoires complexes dans un environnement bruité.

La dernière séquence, **Garden**, comporte 27 images (240×360). La séquence subit un mouvement global translationnel, avec un mouvement de parallaxe sur l'arbre en premier-plan. La fig.5 présente les résultats de FNLC pour les points de l'arbre (points 2,3,4) et ceux du FLC pour les autres. Il est à noter que nous avons choisi de suivre un point de l'arrière plan (point 1) avec le FNLC pour prouver sa robustesse aux occlusions. En suivant les trajectoires des points qui passent derrière l'arbre, nous pouvons remarquer que, sans spécifier de schéma particulier d'occlusions, les deux méthodes retrouvent les positions des points après qu'ils aient été cachés. En effet, la covariance du bruit de mesure, estimée de façon adaptative sur la séquence permet aux filtres conditionnels d'être résistants aux occlusions.

L'ensemble de ces résultats, ainsi que des tests sur des séquences supplémentaires sont disponibles sur la page internet <http://www.irisa.fr/prive/earnaud/demos-RFIA04>.

7 Conclusion

Dans cet article, nous avons proposé une nouvelle formulation des filtres stochastiques adaptée à la trajectographie dans des séquences d'images. Ce cadre permet de considérer des systèmes qui dépendent entièrement de la séquence. Pour les deux méthodes de suivi proposées, la combinaison de mesures calculées par une technique de mise en correspondance, et d'un modèle d'état construit sur une estimation de mouvement a donné de très bons résultats. En effet, des trajectoires subissant des changements abrupts, des occlusions, sur des environnements bruités ont été reconstruites avec succès. Enfin, ces algorithmes ne nécessitent pas d'information *a priori* sur la loi d'évolution du point à suivre.

Annexe : expression de $E_W^*[X|Z]$

En remarquant que, pour deux vecteurs aléatoires Y et W , on a :

$$\begin{aligned} E[\|Y\|^2|W] &= E[Y^t Y|W] = E[\text{tr}\{Y Y^t\}|W] \\ &= \text{tr}\{\text{cov}(Y, Y|W)\} + E[Y|W]^t E[Y|W], \end{aligned}$$

où tr signifie la trace de la matrice entre accolades, et en notant pour trois vecteurs aléatoires arbitraires X , Y et W :

$$\begin{aligned} \Sigma_{X,Z|W} &\triangleq \text{cov}(X, Z|W) \\ &= E[(X - E[X|W])(Z - E[Z|W])^t|W] \\ &= E[X Z^t|W] - E[X|W] E[Z|W]^t, \end{aligned}$$

après quelques manipulations, on peut écrire :

$$\begin{aligned} E[\|X - A Z - B\|^2|W] &= \text{tr}\{(A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}) \Sigma_{Z,Z|W} (A - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1})^t\} \\ &\quad + \|E[X|W] - A E[Z|W] - B\|^2 \\ &\quad + \text{tr}\{\Sigma_{X,X|W} - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} \Sigma_{Z,X|W}\}. \end{aligned}$$

Ces trois termes sont semi-définis positifs.

$E[\|X - A Z - B\|^2|W]$ atteint alors son minimum pour :

$$\begin{aligned} A &= \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1}, \\ B &= E[X|W] - \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} E[Z|W]. \end{aligned}$$

On en déduit l'expression de $E_W^*[X|Z]$.

Références

- [1] R. Adrian, Particle imaging techniques for experimental fluid mechanics, *Annal Rev. Fluid Mechanism*, Vol. 23, pp. 261-304, 1991.
- [2] Anderson et Moore. *Optimal Filtering*, Englewood Cliffs, NJ : Prentice Hall, 1979.
- [3] M.S. Arulampalam, S. Maskell, N. Gordon, et T. Clapp, A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking, *TSP*, Vol. 50(2), 2002.

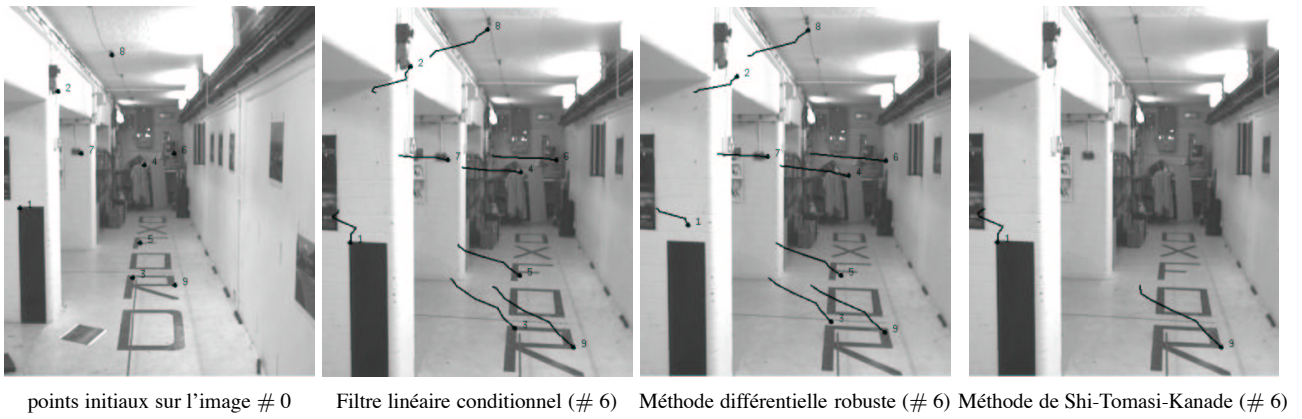


FIG. 3 – Sequence Corridor

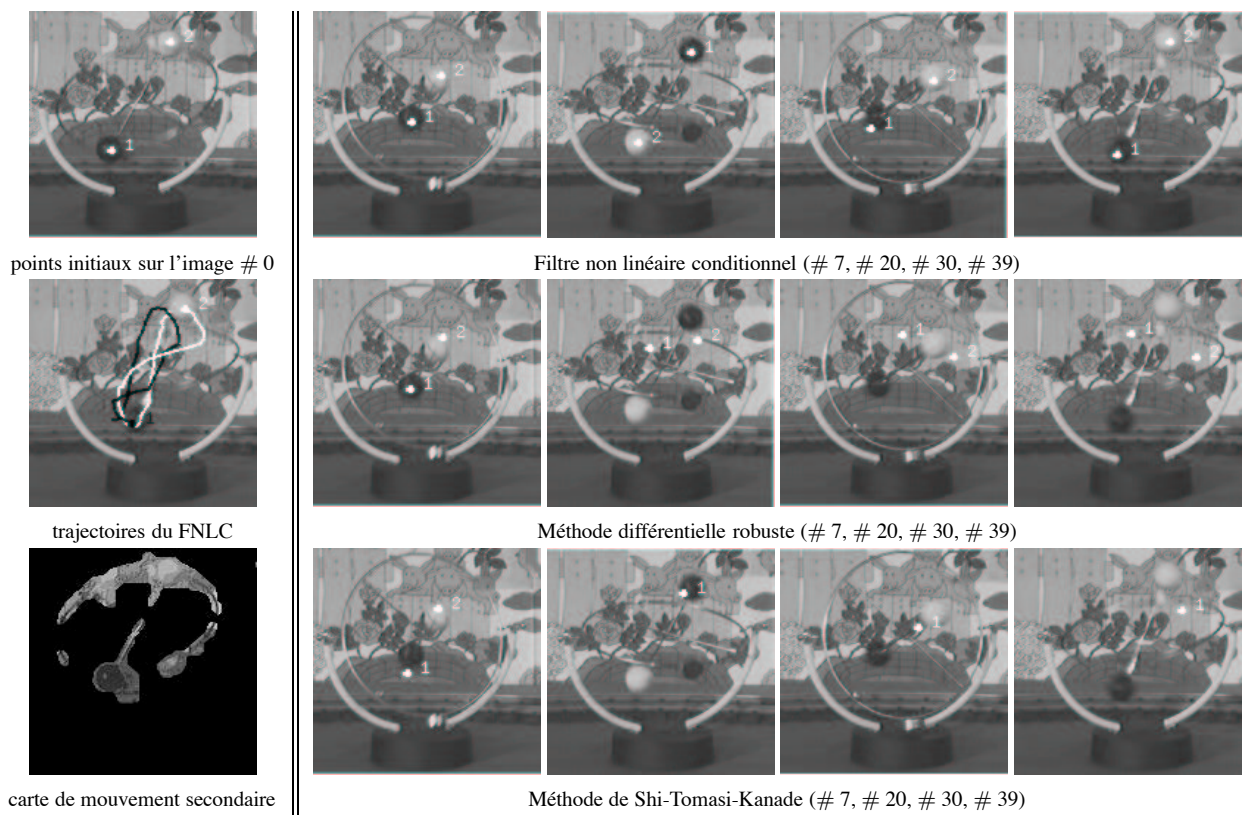


FIG. 4 – Sequence Caltra



FIG. 5 – Sequence Garden

- [4] P. Aschwandten et W. Guggenbühl, Experimental results from a comparative study on correlation-type registration algorithms, *In W. Förstner and St. Ruwiedel, editors, Robust Computer Vision*, pp. 268-289, 1992.
- [5] Y. Bar-Shalom et T.E. Fortmann, *Tracking and Data Association*, Accademic Press, 1988.
- [6] M. Black et P. Anandan, The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, *CVIU*, Vol. 63(1), pp. 75-104, 1996.
- [7] A. Doucet, S. Godsill, et C. Andrieu, On sequential monte carlo sampling methods for bayesian filtering, *Statistics and Computing*, Vol. 10(3), pp. 197-208, 2000.
- [8] N.J. Gordon, D.J. Salmond, et A.F.M. Smith, Novel approach to non-linear/non-gaussian bayesian state estimation, *IEEE Processing-F*, Vol. 140(2), April 1993.
- [9] C. Hue, *Méthodes séquentielles de Monte Carlo pour le filtrage non linéaire multi-objets dans un environnement bûité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D*, PhD thesis, Université de Rennes 1, 2003.
- [10] M. Irani, B. Rousso, et S. Peleg, Computing occluding and transparent motions, *IJCV*, Vol. 12(1), pp. 5-16, 1994.
- [11] M. Isard et A. Blake, Condensation – conditional density propagation for visual tracking, *IJCV*, 1998.
- [12] M. Isard et J. MacCormick, Bramble: a bayesian multiple-blob tracker, *ICCV*, Vol. 2, pp. 34-41, 2001.
- [13] H. Jin, P. Favaro et S. Soatto, Real-time feature tracking and outlier rejection with changes in illumination, *ICCV*, pp. 684-689, 2001.
- [14] A. Kong, J.S. Liu, et W.H. Wong, Sequential imputations and bayesian missing data problems, *Journal of the American Statistical Association*, Vol. 89(425), pp. 278-288, March 1994.
- [15] F. Meyer et P. Bouthemy, Region-based tracking using affine motion models in long image sequences, *CVGIP:IU*, Vol. 60(2), pp. 119-140, 1994.
- [16] K. Nickels et S. Hutchinson, Estimating uncertainty in ssd-based feature tracker, *Image and Vision Computing*, Vol. 20, pp. 47-58, 2002.
- [17] H.T. Nguyen, M. Worring et R. van den Boomgaard, Occlusion robust adaptative template tracking, *ICCV*, Vol. 1, pp. 678-683, 2001.
- [18] N.P. Papanikolopoulos, P.K. Khosla, et T. Kanade, Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision. *IEEE Transactions on Robotics and Automation*, Vol. 9(1), pp. 14-35, 1993.
- [19] P. Pérez, C. Hue, J. Vermaak, et M. Gangnet, Color-based probabilistic tracking, *ECCV*, pp. 661-675, 2002.
- [20] J. Schmetz, K. Holmlund, J. Hoffman, B. Strauss, B. Mason, V. Gaertner, A. Koch, et L. Van de Berg, Operational cloud-motion winds from meteosat infrared images, *Journal of Applied Meteorology*, Vol. 17, pp. 1206-1225, 1993.
- [21] J. Shi and C. Tomasi, Good features to track, *CVPR*, pp. 593-600, 1994.
- [22] A. Singh et P. Allen, Image-flow computation: an estimation-theoric framework and a unified perspective, *CVGIP: IU*, Vol. 56(2), pp. 152-177, 1992.
- [23] T. Tommasini, A. Fusiello, E. Trucco, et V. Roberto, Making good features to track better, *CVPR*, pp. 178-183, 1998.
- [24] L. Torresani et C. Bregler, Space-time tracking, *ECCV*, pp. 801-812, 2002.