

# Filtrage Conditionnel pour le Suivi de Point dans des Séquences d'Images

## *Conditional Filtering For Point Tracking in Image Sequences*

E. Arnaud<sup>1</sup>

E. Mémin<sup>1</sup>

B. Cernuschi-Frias<sup>2</sup>

<sup>1</sup> IRISA, Université de Rennes 1

<sup>2</sup> LIPSIRN, Universidad de Buenos Aires, Argentina

IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France

earnaud@irisa.fr

### Résumé

L'approche proposée pour le suivi de point allie, dans le cadre d'un filtrage stochastique, un modèle dynamique construit sur la contrainte de flot optique, et des mesures fournies par une méthode de mise en correspondance. Sur cette base, deux méthodes simples de suivi, dont la dynamique et les mesures dépendent des données images, sont décrites. La première est une méthode linéaire particulièrement bien adaptée aux séquences présentant un mouvement dominant global. Dans ce contexte, un filtre de Kalman conditionnel est défini par l'utilisation d'un estimateur linéaire conditionnel de variance minimal. La deuxième est une méthode non linéaire construite à partir d'un filtre particulaire conditionnel. Ces deux méthodes permettent de suivre des points dont les trajectoires subissent des changements abrupts. Des résultats expérimentaux sur des séquences réelles sont présentés et comparés avec les résultats donnés par la méthode de suivi de Shi-Tomasi-Kanade.

### Mots Clés

suivi de points, filtre de Kalman, filtrage particulaire

### Abstract

*The approach we investigate for point tracking combines within a stochastic filtering framework a dynamic model relying on the optical flow constraint and measurements provided by a matching technique. Based on this, two kinds of simple trackers are devised whose dynamic and measurements depend on image data. The first one is a linear tracker particularly well-suited to image sequences exhibiting global dominant motion situations. In this context, a conditional Kalman filter is derived through the use of a conditional linear minimum variance estimator. The second one is a nonlinear tracker built from a conditional particle filter. The two trackers allow us to deal with trajectories exhibiting abrupt changes. We present some experimental results on real-world image sequences and compare them to the Shi-Tomasi-Kanade feature tracker.*

### Keywords

Point tracking, Kalman filter, particle filtering

## 1 Introduction

Le suivi de point 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 un problème fondamentalement difficile. En effet, contrairement au suivi de forme structurée, les seules informations sur lesquelles on peut s'appuyer sont les distributions locales de luminance. 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.

L'hypothèse de conservation de la luminance le long d'une trajectoire, ainsi que la sélection des points associée à une caractérisation spatiale et locale de la fonction de luminance ont amené à définir des méthodes de suivi basées sur la corrélation. De telles techniques sont utilisées dans de nombreux domaines pour le suivi de point mais aussi pour estimer les déplacements de formes hautement déformables comme les nuages en imagerie météorologique [1] ou les écoulements fluides en imagerie PIV [1]. Malgré le fait que les fonctions de similarité utilisées ne soient pas invariantes aux transformations géométriques de l'image (changement d'échelle, rotation, distortion perspective), ces méthodes restent très employées pour leur efficacité et leur simplicité. Dans [4], Aschanden et Geggenbühl présentent une étude comparative intéressante entre plusieurs fonctions de similarité. Leurs résultats montrent que les versions à moyenne nulle de la SSD (somme des différences au carré) normalisées ou non, et de la SAD (somme des valeurs absolues) présentent un bon compromis entre vitesse et performance. Pourtant, en cas de transformations géométriques et photométriques trop importantes, ou en cas d'occlusions, l'efficacité de ces méthodes décroît significativement.

L'introduction de la contrainte de flot optique (OFC) [8],

également dérivée de l’hypothèse de conservation de luminance, permet de considérer des méthodes de suivi différentielles. L’algorithme de Shi-Tomasi-Kanade [21] propose d’optimiser un critère reposant sur l’OFC associé soit à un mouvement translationnel en cas de faibles différences inter-images, soit à des déformations affines en cas d’images trop distantes. Pourtant, l’OFC reste sensible aux changements d’illumination. Pour résoudre ce problème, Hager et Belhumeur [7] proposent de définir un modèle photométrique explicite. La solution la plus répandue consiste à inclure des paramètres photométriques d’illumination et/ou de contraste [12]. D’autres adaptations ont été suggérées pour améliorer la qualité des résultats et pour évaluer si le point a été suivi avec succès, telle que l’utilisation de règles de rejet robustes [23].

Une troisième solution consiste à éviter l’hypothèse de conservation de la luminance en incluant une contrainte sur le caractère lisse de la trajectoire [24]. En pratique, cet *a priori* peut paraître trop restrictif dans le cas où la trajectoire à suivre présente un caractère fortement non linéaire.

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 [19, 15] ou des méthodes séquentielles d’approximation de Monte Carlo dans le cas non linéaire [10]. L’état du filtre peut être composé de la position du point, ainsi que d’éventuelles informations complémentaires telles que sa vitesse ou son modèle d’intensité [16]. L’estimation du nouvel état est divisée en deux étapes majeures. Premièrement, une prédiction du nouvel état est obtenue à partir du modèle dynamique. Ce modèle peut être défini *a priori*, ou déduit d’une étape préalable d’apprentissage [10, 22]. Ensuite, la seconde étape consiste à prendre en compte une mesure de l’état à partir de la séquence d’images. Cette mesure, qui se révèle souvent bruitée ou incomplète, permet de corriger la prédiction. Les filtres stochastiques proposent des solutions pour estimer l’état d’un système conditionnellement à l’ensemble des mesures passées.

La méthode que nous proposons associe un modèle dynamique de vitesse reposant sur l’OFC et des mesures fournies par une technique de corrélation invariante aux déformations affines. Cette dernière, basée sur une caractérisation locale des niveaux de gris autour du point cible, est généralement utilisée en appariement d’images [20]. Un filtrage stochastique permet de combiner ces deux techniques de façon probabiliste. L’association d’une méthode de suivi différentielle et d’une méthode de corrélation permet d’améliorer significativement les résultats pour des séquences très diverses. Dans le contexte de cette étude, deux techniques de suivi, dont les structures dépendent du modèle dynamique utilisé sont présentées. La première se place dans un cadre linéaire conditionnellement à la séquence d’images. Elle est particulièrement bien adaptée aux séquences présentant un mouvement dominant. La seconde

se situe dans un cadre non linéaire. Ces deux méthodes permettent de traiter le suivi de point dont les trajectoires subissent des changements abrupts.

Ce papier est organisé comme suit. Après la présentation dans la section 2 du modèle général utilisé en suivi de point, le système construit à partir de la séquence d’images est décrit dans la section 3. Deux cas particuliers de ce système sont développés. D’abord, un filtre de Kalman conditionnel est décrit en section 4 pour un modèle dynamique *global*. Ce filtre se dérive par l’utilisation d’un estimateur linéaire conditionnel de variance minimale. Ensuite, un filtre particulière conditionnel, dont la fonction d’importance est connue, est présenté en section 5. Ce filtre correspond aux situations où le point cible obéit à son propre mouvement. La dernière section présente des résultats expérimentaux sur des séquences réelles. Ces résultats sont comparés à la méthode de Shi-Tomasi-Kanade.

## 2 Suivi par filtrage conditionnel

Nous commençons par décrire le modèle général pour le type de problème de filtrage concerné dans ce papier, ainsi que le déroulement de la solution bayésienne optimale pour le filtrage conditionnel.

### 2.1 Description du modèle général

Soit  $\mathbf{I}_k$  une variable aléatoire représentant l’image acquise au temps  $k$ . Soit  $\mathbf{I}_{0:n}$  la suite finie de variables  $\{\mathbf{I}_k, k = 0, \dots, n\}$ . Cet ensemble correspond à la séquence constituée par la collection d’images de l’instant 0 à l’instant  $k$ . Connaissant une réalisation de  $\mathbf{I}_{0:k}$ , le problème de filtrage est modélisé par une chaîne de Markov cachée. L’évolution de la suite des états  $\mathbf{x}_{0:n} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_n\}$  est donnée par l’équation dynamique :

$$\mathbf{x}_k = f_k(\mathbf{x}_{k-1}, \mathbf{w}_k).$$

L’observation  $\mathbf{z}_k$  est reliée à l’état  $\mathbf{x}_k$  par l’équation de mesure :

$$\mathbf{z}_k = h_k(\mathbf{x}_k, \mathbf{v}_k).$$

A chaque temps  $k$ , une réalisation de  $\mathbf{z}_k$  est obtenue. Elle correspond au résultat d’un processus d’estimation relatif à la séquence d’images  $\mathbf{I}_{0:k}$ . Aucune hypothèse n’est faite sur la forme des fonction  $f_k$  et  $h_k$ . Ces fonctions peuvent dépendre de  $\mathbf{I}_{0:k}$  et être non linéaires. Le bruit sur l’état  $\mathbf{w}_k$  et le bruit sur la mesure  $\mathbf{v}_k$  peuvent également dépendre de  $\mathbf{I}_{0:k}$  et ne sont pas supposés gaussiens. Néanmoins, on suppose que conditionnellement à  $\mathbf{I}_{0:k}$ ,  $\mathbf{w}_k$  et  $\mathbf{v}_k$  sont des bruits blanc, indépendants, de covariance respective  $Q_k$  et  $R_k$ .

Les distributions de probabilité sont supposées être de la forme suivante :

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

De façon identique aux approches classiques de filtrage [3], la variable d’état au temps  $k$  dépend uniquement de l’état

au temps  $k - 1$  (hypothèse Markovienne), et la variable d'observation  $\mathbf{z}_k$  dépend de l'état  $\mathbf{x}_k$ . Néanmoins, il est important de noter que, contrairement aux approches standards, ces deux variables dépendent de la séquence d'images.

## 2.2 Solution bayésienne réursive

A chaque itération, le problème de suivi consiste à avoir une estimation de la densité de probabilité à posteriori de l'état  $\mathbf{x}_k$  par rapport à l'ensemble des informations disponibles, c'est à dire dans notre cas de  $p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$ . Ce problème peut être résolu de façon exacte par une solution bayésienne réursive en deux étapes :

- L'étape de prédiction, qui repose sur l'équation dynamique, permet d'obtenir une première approximation du nouvel état, conditionnellement à l'ensemble des informations disponibles.

En supposant  $p(\mathbf{x}_{k-1} | \mathbf{z}_{1:k-1}, \mathbf{I}_{0:k-1})$  connue, on obtient :

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

- 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 :

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

Le problème de la mise en œuvre de cette solution optimale est l'obtention d'une formulation pour le calcul de ces deux intégrales, avec un bon compromis entre vitesse et précision. Dans le cas où les fonctions  $f_k$  et  $h_k$  sont linéaires, le filtre de Kalman classique [13] propose une solution optimale. Dans le cas d'un modèle qui n'appartient pas à cette classe de problèmes, différents filtres non optimum peuvent être utilisés, en particulier les filtres séquentiels de Monte Carlo.

## 3 Système conditionnel

Il est maintenant nécessaire de définir le modèle de suivi concerné, c'est à dire de définir les expressions de  $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$  et  $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k})$ . Insistons sur le fait que pour le type de problème présenté, les équations dynamique et de mesure sont construites à partir des données images  $\mathbf{I}_{0:k}$ .

### 3.1 Équation de mesure conditionnelle

Un problème majeur dans les applications de suivi en vision est de définir clairement le type de mesures qui pourra être utilisé. Du fait que l'ensemble des informations est contenu dans la séquence d'images, le cas idéal serait de considérer chaque image comme une mesure. Cependant,

en raison de leur structure complexe et de leur trop grande taille, il est, la plupart du temps, impossible de spécifier une relation entre les données images et le système à observer. Un choix raisonnable consiste à s'appuyer sur une information condensée obtenue à partir de la séquence. Dans ce sens, il est possible de construire une observation de l'état hautement non linéaire par rapport aux données images, mais qui possède une forme simple vis à vis de l'état du système. Dans ce travail, nous nous sommes restreint au cadre de systèmes avec une équation de mesure linéaire de la forme :

$$\mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k,$$

où  $\mathbf{z}_k$  est le résultat d'un processus d'estimation sur la séquence d'images  $\mathbf{I}_{0:k}$  et  $\mathbf{v}_k$  est un bruit blanc gaussien conditionnellement à  $\mathbf{I}_{0:k}$ . En résumé,  $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k})$  est supposée être une fonction gaussienne. Il est important de noter que, bien que les données images  $\mathbf{I}_{0:k}$  n'apparaissent pas explicitement dans l'équation d'observation, celles-ci doivent être prises en compte par un conditionnement par rapport à  $\mathbf{I}_{0:k}$  dans les équations de chaque filtre.

**En quoi cette restriction simplifie le mise en œuvre des algorithmes de suivi ?** Dans le cas où l'équation dynamique est linéaire, la forme gaussienne de l'équation de mesure permet d'utiliser le filtre de Kalman pour résoudre le problème de suivi. En effet, même dans le cas de bruits blanc non gaussiens conditionnellement à  $\mathbf{I}_{0:k}$ , des équations similaires au filtre de Kalman peuvent être déduites par un estimateur linéaire conditionnel de variance minimale. Un tel système linéaire est bien adapté au suivi de point dont le mouvement appartient au mouvement *global* dominant de la scène. Le suivi d'un point ayant son propre mouvement *local* est bien représenté par un système dont le modèle dynamique est non linéaire avec un bruit blanc additionnel gaussien. Dans ce cas, la forme gaussienne de l'équation de mesure permet d'utiliser un filtre particulière dont la fonction d'importance est connue.

**En quoi cette restriction est un choix raisonnable en suivi par ordinateur ?** Comme dit précédemment, pour suivre un point dans une séquence d'images, il est nécessaire de faire certaines hypothèses de conservation d'informations en rapport avec ce point. Ces hypothèses peuvent concerner la vitesse du point (position constante, vitesse constante, accélération constante, mouvement périodique), la position du point associée à la construction géométrique de la scène, ou une invariance photométrique de la luminance dans un voisinage du point (luminance constante, invariants photométriques, géométriques, invariants statistiques de la distribution de luminance, etc.). Comme les deux premières solutions apparaissent être à la fois difficiles à établir sans connaissance *a priori*, et trop restrictives en cas de trajectoires avec des changements abrupts, la troisième solution a été préférée. L'approche proposée est construite uniquement sur la base d'une conservation du modèle de luminance autour du point cible. Il est supposé que ce modèle de luminance peut subir des déforma-

tions géométriques affines (changement d'échelle, rotation, translation) ou des changements affines d'intensité (changement de contraste et d'illumination).

Chaque état  $\mathbf{x}_k = (x_k, y_k)^t$  représente la position du point au temps  $k$  correspondant à l'image  $\mathbf{I}_k$ . Au temps  $k$ ,  $\mathbf{x}_k$  est supposé être observable à travers un processus d'estimation dont le but est ici de fournir le point de  $\mathbf{I}_k$  le plus similaire au point initial  $\mathbf{x}_0$  de  $\mathbf{I}_0$ . Soit  $\phi(\mathbf{x}_0, \mathbf{I}_{0:k})$  ce processus d'estimation qui fournit la mesure  $\mathbf{z}_k$ . Il est supposé que cette mesure transporte suffisamment d'information concernant l'état du point suivi pour pouvoir écrire  $\mathbf{x}_k = \mathbf{z}_k$ , à un bruit blanc gaussien près. Ce bruit de mesure modélise une erreur locale d'estimation, dépendante de  $\mathbf{I}_{0:k}$ . Plus précisément, l'équation d'observation considérée s'écrit comme suit :

$$\begin{cases} \mathbf{x}_k = \mathbf{z}_k + \mathbf{v}_k \\ \mathbf{z}_k = \phi(\mathbf{x}_0, \mathbf{I}_{0:k}) \\ p(\mathbf{v}_k | \mathbf{I}_{0:k}) \sim \mathcal{N}(0, R_k). \end{cases} \quad (1)$$

Plusieurs critères de mise en correspondance peuvent être utilisés pour quantifier la similarité entre le point cible et son voisinage, et le point candidat. Le critère utilisé ici est décrit dans [20] pour l'appariement d'images. A chaque point  $\mathbf{p}$  de la grille d'image  $S$  est associé un vecteur  $c_{\mathbf{p}k}$  qui caractérise le signal en niveaux de gris  $\mathbf{I}_k$  autour de  $\mathbf{p}$ . Ce vecteur est construit à partir des dérivées jusqu'à l'ordre trois du signal de luminance. Le processus de mise en correspondance consiste alors à estimer le point qui minimise la distance euclidienne entre le vecteur de caractéristiques initial  $c_{\mathbf{x}_0}$  et les vecteurs  $c_{\mathbf{p}k}$  :

$$\mathbf{z}_k = \phi(\mathbf{x}_0, \mathbf{I}_{0:k}) = \arg \min_{\mathbf{p} \in S} \|c_{\mathbf{x}_0} - c_{\mathbf{p}k}\|^2.$$

### 3.2 Équation dynamique conditionnelle

De façon similaire à l'équation de mesure, l'équation dynamique considérée est obtenue à partir de la séquence d'images, afin d'être réactif à chaque changement de vitesse et de direction du point. Pour cela, il est nécessaire de définir la distribution  $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$ .

L'équation d'état décrit le mouvement du point  $\mathbf{x}_{k-1}$  entre les images  $k-1$  et  $k$ , et fournit une estimation de la nouvelle position  $\mathbf{x}_k$ . Une technique robuste d'estimation paramétrique du mouvement [5, 18] permet d'estimer de façon fiable un modèle paramétrique 2D représentant le mouvement dominant, sur un support donné  $\mathcal{R}$ . L'utilisation d'une telle méthode sur un support approprié  $\mathcal{R}(\mathbf{x}_{k-1})$  autour de  $\mathbf{x}_{k-1}$ , à partir des images  $\mathbf{I}_{k-1}$  et  $\mathbf{I}_k$ , fournit une estimation du vecteur vitesse  $\mathbf{U}(\mathbf{x}_{k-1})$  au point  $\mathbf{x}_{k-1}$  entre les instants  $k-1$  et  $k$ . Ce vecteur de vitesse s'écrit :

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

avec

$$\hat{\boldsymbol{\theta}}_k = \arg \min_{\boldsymbol{\theta}} \int_{\mathcal{R}(\mathbf{x}_{k-1})} \rho \left( \nabla \mathbf{I}_k^t(\mathbf{s}) P(\mathbf{s}) \boldsymbol{\theta} + \frac{\partial \mathbf{I}_k}{\partial t}(\mathbf{s}) \right) d\mathbf{s},$$

où  $\rho$  est une fonction de coût robuste permettant de gérer les points aberrants. Ces points peuvent être identifiés comme appartenant à des régions dont le mouvement ne correspond pas au modèle estimé, ou à des régions où la contrainte de conservation de la luminance n'est pas valide. Rappelons que si on considère un modèle de mouvement affine (i.e.  $\hat{\boldsymbol{\theta}}_k$  est un vecteur à 6 paramètres), la matrice  $P(\mathbf{x}_{k-1})$  est définie telle que :

$$P(\mathbf{x}_{k-1}) = \begin{bmatrix} 1 & x_{k-1} & y_{k-1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_{k-1} & y_{k-1} \end{bmatrix}.$$

Le problème essentiel d'une telle approche est de définir le support d'estimation approprié  $\mathcal{R}(\mathbf{x}_{k-1})$  pour un point donné  $\mathbf{x}_{k-1}$ . Deux cas extrêmes peuvent être distingués.

**Point appartenant au mouvement global.** Quand le mouvement du point à suivre correspond au mouvement dominant global, le support d'estimation est défini comme la grille d'image complète  $\mathcal{R}(\mathbf{x}_{k-1}) = S$ . On appellera  $\hat{\boldsymbol{\theta}}_k^g$  le vecteur de paramètres du modèle global affine de vitesse  $\hat{\boldsymbol{\theta}}_k^g$ , estimé entre les instants  $k-1$  et  $k$ . Comme  $\hat{\boldsymbol{\theta}}_k^g$  ne dépend pas de  $\mathbf{x}_{k-1}$ , la fonction dynamique correspondante est linéaire par rapport à  $\mathbf{x}_{k-1}$  :

$$\begin{aligned} \mathbf{x}_k &= \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \hat{\boldsymbol{\theta}}_k^g + \mathbf{w}_k \\ &= A_k \mathbf{x}_{k-1} + \mathbf{b}_k + \mathbf{w}_k, \end{aligned} \quad (2)$$

où  $A_k$  est la matrice des mouvements de rotation, divergence et 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, il est supposé que, conditionnellement à  $\mathbf{I}_{0:k}$ ,  $\mathbf{w}_k$  est un bruit blanc de moyenne nulle et de covariance  $Q_k$ .

**Point dont le mouvement est local.** Quand le mouvement du point suivi peut être uniquement décrit par un modèle paramétrique local, le support  $\mathcal{R}(\mathbf{x}_{k-1})$  est fixé à un petit domaine centré en  $\mathbf{x}_{k-1}$ . Pour un souci de clarté, on appellera  $\hat{\boldsymbol{\theta}}_k^l$  le vecteur de paramètres d'un tel modèle de mouvement, estimé sur un domaine local autour de  $\mathbf{x}_{k-1}$ , entre les instants  $k-1$  et  $k$ . Même si cela est omis dans la notation, il est important de se rappeler que  $\hat{\boldsymbol{\theta}}_k^l$  dépend de  $\mathbf{x}_{k-1}$ . L'équation dynamique non linéaire correspondante s'écrit alors :

$$\mathbf{x}_k = \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \hat{\boldsymbol{\theta}}_k^l + \mathbf{w}_k. \quad (3)$$

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 de mesure suit une distribution gaussienne centrée conditionnellement à  $\mathbf{I}_{0:k}$ , i.e.  $p(\mathbf{w}_k | \mathbf{I}_{0:k}) \sim \mathcal{N}(0, Q_k)$ .

**Comment décider si le point à suivre a un mouvement global ou local?** L'utilisation d'une méthode robuste d'estimation du mouvement associée à une méthode de détection de mouvement [9, 17], appliquées au temps initial entre les images  $\mathbf{I}_0$  et  $\mathbf{I}_1$ , est un moyen pratique pour déterminer si le point appartient au support de mouvement

dominant ou à une région mobile caractérisée par un modèle de mouvement local.

### 3.3 Critère de sélection de points

Avant de reconstruire la trajectoire d'un point, il est nécessaire de tester si ce point peut être suivi de manière fiable. Pour cela, le critère de sélection présenté dans [21] est utilisé à l'initialisation. Ce critère est basé sur les valeurs propres de la matrice d'auto-corrélation  $M$  :

$$M = \int_{\mathcal{R}(\mathbf{x}_0)} \begin{bmatrix} \mathbf{I}_x^2 & \mathbf{I}_y \mathbf{I}_x \\ \mathbf{I}_x \mathbf{I}_y & \mathbf{I}_y^2 \end{bmatrix}$$

avec  $[\mathbf{I}_x, \mathbf{I}_y] = [\partial \mathbf{I}_0 / \partial x, \partial \mathbf{I}_0 / \partial y]$ . Les deux valeurs propres  $\lambda_1$  et  $\lambda_2$  fournissent des informations sur le profil d'intensité de la fenêtre  $\mathcal{R}(\mathbf{x}_0)$ . Deux faibles valeurs sont associées à un modèle d'intensité constante, alors que deux valeurs importantes indiquent une fonction de luminance qui peut être facilement suivie. Pour être bien conditionnées, les valeurs de la matrice doivent être au-dessus du niveau de bruit. Le point cible correspondant est accepté si  $\min(\lambda_1, \lambda_2) > \lambda$ . En général, la valeur de  $\lambda$  est comprise dans l'intervalle  $[0.1, 1]$ .

## 4 Filtre de Kalman conditionnel

Dans le cas de modèles linéaires gaussiens, le filtre de Kalman donne une solution optimale récursive, par des expressions analytiques des moyenne et covariance de la distribution gaussienne  $p(\mathbf{x}_k | \mathbf{z}_{1:k})$ . De telles expressions peuvent être dérivées par l'utilisation de l'estimateur de variance minimale. Il est connu que cet estimateur est équivalent au meilleur estimateur linéaire dans le cas gaussien, et correspond alors à l'espérance conditionnelle de l'état au temps  $k$  connaissant l'ensemble des mesures disponibles  $E[\mathbf{x}_k | \mathbf{z}_{1:k}]$ . Bien que l'approche de suivi proposée soit linéaire dans le cas d'un descripteur de mouvement global, il a été vu précédemment que le bruit  $\mathbf{w}_k$  n'est pas connu et certainement non gaussien. Du fait qu'il est également nécessaire de prendre en compte la dépendance à la séquence d'images des équations dynamique et de mesure, un estimateur linéaire conditionnel de variance minimale est utilisé. Commençons par introduire cet estimateur.

**Définition 1** Soit 3 variables aléatoires jointes  $X, Y, W$ . On note  $E^*[X|Y, W]$  le meilleur estimateur (au sens de la norme  $L_2$ ) de  $X$ , linéaire en  $Y$ , conditionnel à  $W$  :

$$E_W^*[X|Y] = AY + B$$

avec  $A$  et  $B$  tel que  $E[\|X - AY - B\|^2 | W]$  soit minimum.

$E_W^*[X|Y]$  est appelé estimateur linéaire conditionnel de variance minimale.

Il est à noter que  $E_W^*[X|Y]$  n'est pas une espérance. En définissant  $\Sigma_{X,Y|W} = E[XY^t | W] - E[X|W]E[Y^t | W]$ , le résultat suivant est obtenu :

$$\begin{aligned} E_W^*[X|Y] &= E[X|W] + \Sigma_{X,Y|W} \Sigma_{Y,Y|W}^{-1} (Y - E[Y|W]). \end{aligned}$$

Il peut être vérifié que cet estimateur possède des propriétés équivalentes à l'estimateur linéaire de variance minimale (propriété d'être non biaisé, quantités conditionnelles non corrélées, principe d'orthogonalité, changement de variable conditionnelle)[2]. En supprimant l'hypothèse gaussienne et en conditionnant par  $\mathbf{I}_{0:k}$ , l'utilisation de l'estimateur linéaire conditionnel de variance minimale amène aux mêmes équations que le filtre de Kalman classique, mais l'interprétation de ces expressions est différente. Un tel estimateur fournit uniquement les statistiques du premier et second ordre, et ne fournit que des informations incomplètes sur les moments d'ordre supérieur.

Soit un système général de la forme :

$$\begin{cases} \mathbf{x}_k = A_k \mathbf{x}_{k-1} + \mathbf{b}_k + \mathbf{w}_k \\ \mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{v}_k, \end{cases}$$

où les matrices  $A_k$ ,  $H_k$  et le vecteur  $\mathbf{b}_k$  peuvent dépendre de  $\mathbf{I}_{0:k}$ . Les variables  $\mathbf{w}_k$ ,  $\mathbf{v}_k$  sont des bruits blanc indépendant de moyenne nulle (éventuellement non gaussiens) conditionnellement à  $\mathbf{I}_{0:k}$ , de covariances connues, respectivement notées  $Q_k$  et  $R_k$ .

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 de l'erreur conditionnelle associée. Une formulation récursive de  $\hat{\mathbf{x}}_{k+1|k}$  peut être trouvée par des manipulations similaires au cas gaussien, en considérant les expressions conditionnelles induites par  $E^*$  :

$$\hat{\mathbf{x}}_{k+1|k} = A_{k+1} \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1} + \tilde{K}_k (\mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1}),$$

où la matrice  $\tilde{K}_k$  est définie par la matrice de gain  $K_k$  :

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

Une expression récursive de la covariance de l'erreur conditionnelle  $\Sigma_{k+1|k}$  peut alors être trouvée :

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

Ces équations peuvent être séparées pour distinguer la phase de prédiction et la phase de mise à jour.

### 4.1 Application au modèle de suivi de point

Dans le cas d'un point associé au mouvement global dominant caractérisé par un descripteur linéaire unique, le problème de filtrage combine alors une équation dynamique linéaire (2) et une équation de mesure linéaire (1). L'algorithme correspondant, qui repose sur l'estimateur linéaire conditionnel de variance minimale est décrit fig. 1.

## 5 Filtrage particulière conditionnel

Dans le cas où l'équation d'état ou l'équation de mesure est non linéaire, le problème 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

<ul style="list-style-type: none"> <li>– <b>initialisation</b> soit <math>\hat{\mathbf{x}}_{0 0} = \mathbf{x}_0</math>. définir <math>\Sigma_{0 0}</math>.</li> <li>pour <math>k = 1, \dots, n</math></li> <li>– <b>modèle d'état</b> fixer <math>A_k</math> et <math>\mathbf{b}_k</math> à partir de l'approximation affine du mouvement global <math>\hat{\boldsymbol{\theta}}_k^g</math>.</li> <li>– <b>prédiction</b> <math>\hat{\mathbf{x}}_{k k-1} = A_k \hat{\mathbf{x}}_{k-1 k-1} + \mathbf{b}_k</math> <math>\Sigma_{k k-1} = A_k \Sigma_{k-1 k-1} A_k^t + Q_k</math></li> <li>– <b>mesure</b> <math>\mathbf{z}_k = \phi(\mathbf{x}_k, \mathbf{I}_{0:k})</math></li> <li>– <b>mise à jour</b> <math>K_k = \Sigma_{k k-1} (\Sigma_{k k-1} + R_k)^{-1}</math> <math>\hat{\mathbf{x}}_{k k} = \hat{\mathbf{x}}_{k k-1} + K_k (\mathbf{z}_k - \hat{\mathbf{x}}_{k k-1})</math> <math>\Sigma_{k k} = (I - K_k) \Sigma_{k k-1}</math></li> </ul>
---

FIG. 1 – Filtre de Kalman conditionnel pour le suivi de point dont le mouvement appartient au mouvement dominant global.

mesures jusqu'à  $k$ . Pour surmonter cette difficulté, les techniques de filtrage particulaire proposent de mettre en œuvre de façon récursive une approximation de la densité a posteriori (voir [3, 6] pour une description détaillée de ces méthodes). L'algorithme de filtrage particulaire consiste à propager un nuage de  $N$  particules. Chaque particule  $\mathbf{x}_{0:k}^{(i)}$  correspond à une trajectoire réalisable du système 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 le cas où l'état et la mesure dépendent de la séquence d'images. Une telle dépendance 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  $p(\mathbf{x}_0)$ ,  $p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$  et  $p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{I}_{0:k}) \forall k \geq 1$ , 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 en anglais) permet d'approximer cette intégrale en introduisant une distribution de probabilité  $\pi(\mathbf{x}_{0:k} | \mathbf{z}_{1:k}, \mathbf{I}_{0:k})$  appelée fonction d'importance, pour laquelle on sait obtenir des réalisations.

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

conditionnelle par une somme finie :

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

Les poids non normalisés sont donnés par :

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

Afin de construire une expression récursive de l'espérance conditionnelle, on suppose une forme récursive de la fonction d'importance. Cette formulation introduit une hypothèse de causalité sur les données :

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

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

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

Néanmoins, cette approximation conduit à une augmentation de la variance des poids d'importance 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 dans la littérature.

Une première solution consiste à choisir une fonction d'importance qui minimise la variance des poids d'importance  $w(\mathbf{x}_{0:k})$  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 d'importance 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  $w(\mathbf{x}_{0:k})$  :

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

Le problème de cette approche est que généralement, la distribution de probabilité  $p(\mathbf{x}_k | \mathbf{z}_k, \mathbf{x}_{k-1}, \mathbf{I}_{0:k})$  est totalement inconnue. Nous verrons que ce n'est pas le cas dans le cadre de notre modèle. Remarquons que généralement, dans les applications de vision, la fonction d'importance n'est pas connue et est identifiée au processus de diffusion (i.e.  $\pi(\mathbf{x}_k | \mathbf{x}_{0:k-1}, \mathbf{z}_{1:k}) = p(\mathbf{x}_k | \mathbf{x}_{k-1})$ ) [10].

Une deuxième solution repose sur une méthode de rééchantillonnage qui consiste à supprimer les trajectoires ayant un faible poids d'importance normalisé et à ajouter des copies des trajectoires associées à un poids fort.

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

## 5.1 Application au modèle de suivi de point

Quand le mouvement du point peut être uniquement représenté par une paramétrisation linéaire locale, le problème de filtrage est alors formulé par une équation de mesure linéaire (1) et l'équation dynamique non linéaire décrite (3). Comme il a été expliqué précédemment, dans ce cas,

on suppose que les bruits  $\mathbf{w}_k$  et  $\mathbf{v}_k$  sont des bruits blanc gaussiens indépendants conditionnellement à la séquence d'images  $\mathbf{I}_{0:k}$ . On suppose ainsi que

$p(\mathbf{w}_k|\mathbf{I}_{0:k}) \rightsquigarrow \mathcal{N}(0, Q_k)$  et  $p(\mathbf{v}_k|\mathbf{I}_{0:k}) \rightsquigarrow \mathcal{N}(0, R_k)$ . Les distributions de probabilité impliquées dans le système sont alors :

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{I}_{0:k}) \rightsquigarrow \mathcal{N}(\mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \hat{\boldsymbol{\theta}}_k^l, Q_k),$$

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

Ces distributions gaussiennes conduisent à des expressions simples pour la formule de récursivité des poids et pour la fonction d'importance. En effet, en remarquant que :

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

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

On en déduit que :

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

ce qui conduit à une expression simple pour (4). En ce qui concerne la fonction d'importance, nous avons :

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

et ainsi,  $p(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{z}_k, \mathbf{I}_{0:k}) \rightsquigarrow \mathcal{N}(\mathbf{m}_k(\mathbf{x}_{k-1}), C_k)$  avec

$$\mathbf{m}_k(\mathbf{x}_{k-1}) = C_k(Q_k^{-1}(\mathbf{x}_{k-1} + P(\mathbf{x}_{k-1}) \hat{\boldsymbol{\theta}}_k^l) + R_k^{-1} \mathbf{z}_k),$$

$$C_k = (Q_k^{-1} + R_k^{-1})^{-1}.$$

Un schéma récursif simple pour le calcul des poids et la connaissance de la fonction d'importance optimale constituent deux avantages importants de l'approche proposée. La fig. 2 résume l'approche de suivi de point proposée pour un point dont le mouvement est représenté par une paramétrisation locale linéaire.

## 6 Résultats

Nous présentons en premier lieu des résultats expérimentaux de la méthode de suivi utilisant le filtre de Kalman conditionnel. Ces résultats concernent deux séquences réelles ayant un mouvement global dominant. La première, **Hangars**, qui comporte 10 images ( $512 \times 512$  pixels), est une séquence militaire bruitée présentant un mouvement global chaotique. Une comparaison entre les résultats de la méthode de filtrage proposée et la méthode de Shi-Tomasi-Kanade est présentée fig. 3. Dans une telle séquence, on peut remarquer que la méthode de Shi-Tomasi-Kanade donne de mauvais résultats. Ceci est particulièrement vrai dans le cas de points qui ne peuvent pas être identifiés par un modèle de luminance caractéristique (point sur des coins, etc.). En effet, ceci est un défaut connu d'une telle méthode. Au contraire, dans le cas du filtre de Kalman conditionnel, les trajectoires de chaque point sont bien reconstruites, même si ces dernières présentent des points d'inflexion (voir fig. 3(d)). De plus, il est important de noter que

- |  |
|--|
| <ul style="list-style-type: none"> <li>- <b>initialisation</b><br/>pour <math>i = 1, \dots, N</math>, <math>\mathbf{x}_0^{(i)} \rightsquigarrow p(\mathbf{x}_0)</math>, <math>w(\mathbf{x}_0^{(i)}) = \frac{1}{N}</math><br/>pour <math>k = 1, \dots, n</math></li> <li>- <b>mesures</b><br/>évaluer <math>\mathbf{z}_k = \phi(\mathbf{x}_0, \mathbf{I}_{0:k})</math>,<br/>pour <math>i = 1, \dots, N</math>, évaluer <math>\hat{\boldsymbol{\theta}}_{k-1}^{l(i)}</math></li> <li>- <b>prédiction</b><br/>pour <math>i = 1, \dots, N</math>, <math>\mathbf{x}_{0:k}^{(i)} = (\mathbf{x}_{0:k-1}^{(i)}, \mathbf{x}_k^{(i)})</math><br/>avec <math>\mathbf{x}_k^{(i)} \rightsquigarrow \mathcal{N}(\mathbf{m}_k(\mathbf{x}_{k-1}^{(i)}), C_k)</math></li> <li>- <b>correction</b><br/>pour <math>i = 1, \dots, N</math><br/><math>w(\mathbf{x}_{0:k}^{(i)}) = w(\mathbf{x}_{0:k-1}^{(i)}) p(\mathbf{z}_k \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})</math> avec<br/><math>p(\mathbf{z}_k \mathbf{x}_{k-1}^{(i)}, \mathbf{I}_{0:k})</math><br/><math>\rightsquigarrow \mathcal{N}(\mathbf{x}_{k-1}^{(i)} + P(\mathbf{x}_{k-1}^{(i)}) \hat{\boldsymbol{\theta}}_{k-1}^{l(i)}, R_k + Q_k)</math></li> <li>- <b>rééchantillonnage</b></li> <li>- <b>estimation de la trajectoire</b><br/>évaluer <math>\hat{T} = \sum_{i=1}^N \mathbf{x}_{0:k}^{(i)} \tilde{w}(\mathbf{x}_{0:k}^{(i)})</math></li> </ul> |
|--|

FIG. 2 – Filtre particulaire conditionnel pour le suivi de point dont le mouvement n'est pas conforme au mouvement dominant.

la méthode de Shi-Tomasi-Kanade perd la presque totalité des points à l'image 4, et qu'elle suit les points représentant des coins moins précisément que la méthode proposée (voir fig. 3(d)). 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. Ceci peut être observé à partir des faibles variations de trajectoires entre points voisins.

La deuxième séquence testée, **Avengers**, est une séquence de 34 images ( $288 \times 352$  pixels), prise à partir d'un hélicoptère. Mis à part le mouvement des deux voitures, comme on peut le voir sur la carte binaire de détection du mouvement non conforme au mouvement dominant (fig. 4), cette séquence présente un mouvement globalement translationnel. La fig. 5 montre les trajectoires obtenues par la méthode de filtre de Kalman conditionnel. En suivant les trajectoires des points 1 et 2, on remarque que cette méthode se révèle résistante aux occultations, sans avoir spécifier un schéma d'occlusion particulier. En ce qui concerne le suivi des voitures, la méthode basée sur le filtre particulaire conditionnel a été utilisée, dont le résultat est présenté fig. 6. Au vue de la trajectoire globale (fig. 7), il est à remarquer qu'un *a priori* de type vitesse constante ou accélération constante aurait perturber le suivi.

Un second résultat de la méthode de suivie basée sur le filtre particulaire conditionnel est présenté sur la séquence **Caltra** qui comporte 40 images ( $190 \times 180$ ). Cette séquence représente le mouvement de deux boules fixées sur un cercle rigide en rotation, sur un arrière-plan encombré.

Le mouvement des deux boules est par ailleurs très rapide (en moyenne 10 pixels entre chaque image). 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 est fixée à une fenêtre  $32 \times 32$ . Comparée au résultat apporté par la méthode de Shi-Tomasi-Kanade (fig. 8), la méthode par filtre particulaire conditionnel fournit une trajectoire complète de la boule blanche. La trajectoire donnée par la méthode de Shi-Tomasi-Kanade est à peu près cohérente jusqu'à l'image 6. Un tel résultat montre la capacité de cette méthode à traiter des mouvements rapides et des trajectoires complexes.

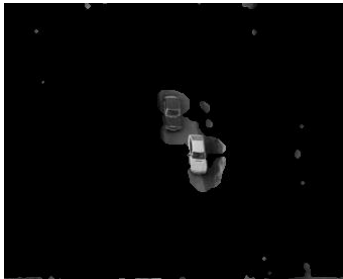


FIG. 4 – Séquence Avengers. carte de détection de région non conforme au mouvement dominant.

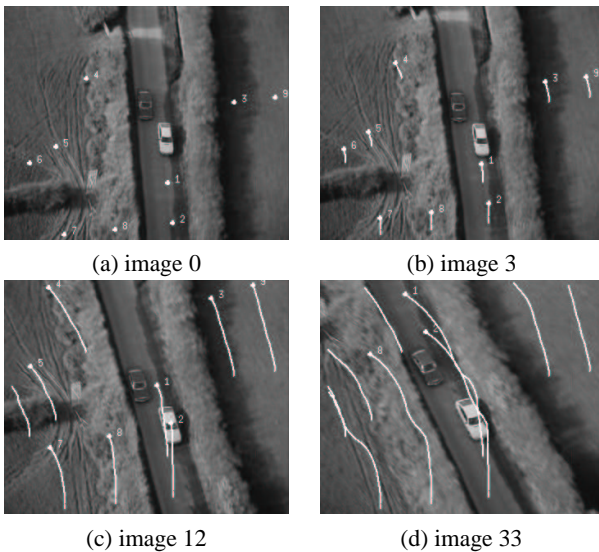


FIG. 5 – Séquence Avengers. (a) points initiaux ; (b,c,d) trajectoires obtenues par la méthode de suivi utilisant le filtre de Kalman conditionnel ; bonne gestion des occlusions rencontrées par les points 1 et 2. les positions finales sont représentées par un point.

## 7 Conclusion

Nous avons proposé deux méthodes de suivi de point reposant sur des techniques de filtrage stochastique. Pour ces

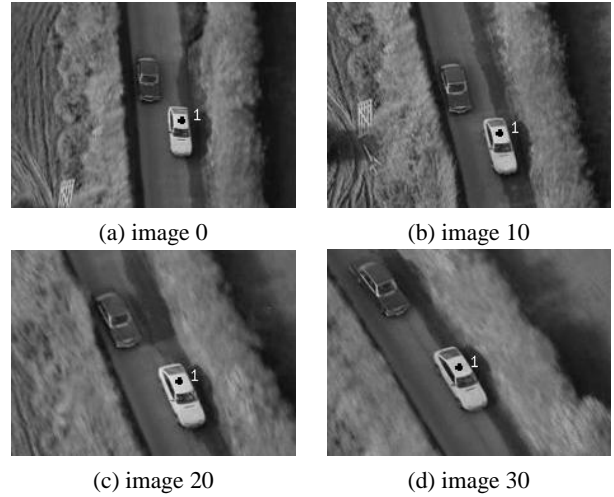


FIG. 6 – Séquence Avengers. (a) point initial ; (b,c,d) résultat du suivi par la méthode utilisant le filtre particulaire conditionnel

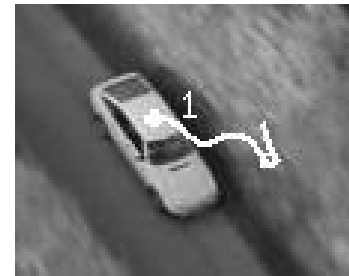


FIG. 7 – Séquence Avengers. trajectoire complète obtenue par le méthode de suivi utilisant le filtre particulaire conditionnel

deux méthodes, l'association d'un modèle dynamique et d'une équation de mesure qui dépendent de la séquence d'images donne de très bons résultats dans des cas difficiles caractérisés par des trajectoires comportant des changements abrupts et des séquences bruitées. Ces algorithmes ne nécessitent pas de spécifications *a priori* sur le point à suivre, ni sur son mouvement. La version linéaire de ces méthodes s'est révélée être résistante aux occlusions. Une extension logique de ce travail serait d'inclure un schéma d'occlusions dans l'équation dynamique non linéaire, ainsi que de considérer une mesure de confiance sur l'observation dans les deux systèmes.

## Références

- [1] R. Adrian. Particle imaging techniques for experimental fluid mechanics. *Annal Rev. Fluid Mechanism*, 23:261–304, 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*, 50(2), 2002.



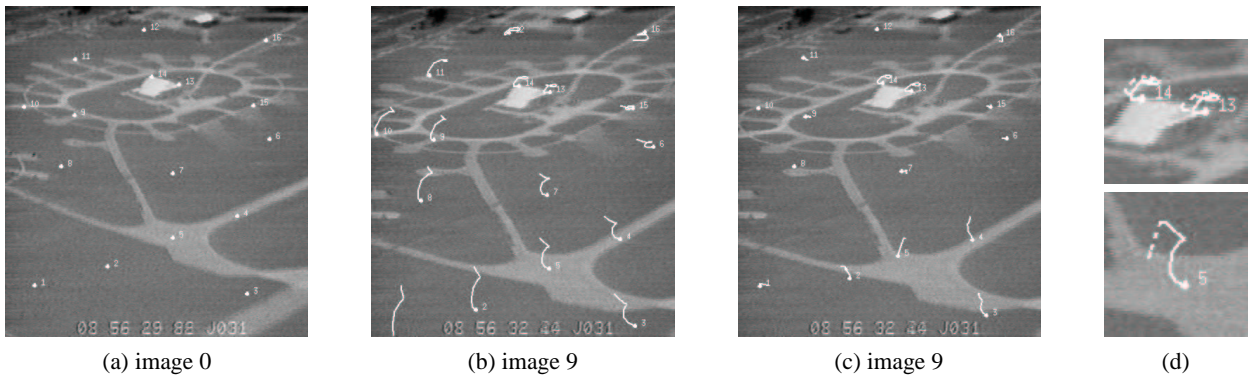


FIG. 3 – Séquence Hangar. (a) points initiaux ; (b) trajectoires obtenues par la méthode de suivi utilisant le filtre de Kalman conditionnel ; (c) trajectoires obtenues par le méthode de suivi Shi-Tomasi-Kanade ; (d) détails de la séquence Hangars. trajectoires de la méthode Shi-Tomasi-Kanade représentées en lignes pointillées ; trajectoires par le filtre de Kalman conditionnel représentées en lignes complètes.

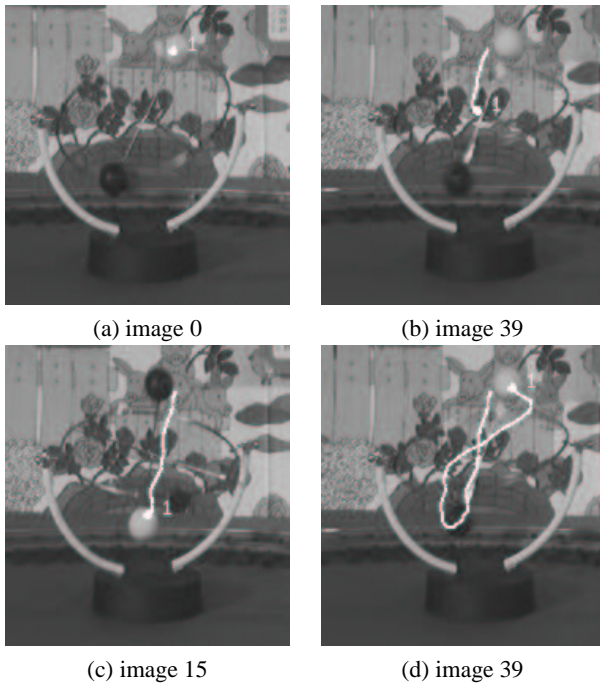


FIG. 8 – Séquence Caltra. (a) point initial ; (b) trajectoire obtenue par le méthode de suivi Shi-Tomasi-Kanade ; (c,d) trajectoire obtenue par le méthode de suivi utilisant le filtrage particulaire conditionnel.

[4] P. Aschwanden et W. Guggenbühl. Experimental results from a comparative study on correlation-type registration algorithms. In W. Förstner and St. Ruwiedel, editors, *Robust Computer Vision*, pages 268–289, 1992.

[5] M. Black et P. Anandan. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *CVIU*, 63(1):75–104, 1996.

[6] A. Doucet, S. Godsill, et C. Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and Computing*, 10(3):197–208, 2000.

[7] G.D. Hager et P.N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 20(10):1025–1039, October 1998.

[8] B. Horn et B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[9] M. Irani, B. Rousso, et S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, 12(1):5–16, 1994.

[10] M. Isard et A. Blake. Condensation – conditional density propagation for visual tracking. *IJCV*, 1998.

[11] K. Holmlund J. Schmetz, J. Hoffman, B. Strauss, B. Mason, V. Gaertner an A. Koch, et L. Van de Berg. Operational cloud-motion winds from meteosat infrared images. *Journal of Applied Meteorology*, 32:1206–1225, 1993.

[12] H. Jin, P. Favaro, et S. Soatto. Real-time feature tracking and outlier rejection with changes in illumination. In *International Conference on Computer Vision*, pages 684–689, 2001.

[13] R.E. Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, pages 35–45, 1960.

[14] A. Kong, J.S. Liu, et W.H. Wong. Sequential imputations and bayesian missing data problems. *Journal of the American Statistical Association*, 89(425):278–288, March 1994.

[15] F. Meyer et P. Bouthemy. Region-based tracking using affine motion models in long image sequences. *CVGIP:IU*, 60(2):119–140, 1994.

[16] H.T. Nguyen, M. Worring, et R. van den Boomgaard. Occlusion robust adaptative template tracking. In *International Conference on Computer Vision*, volume 1, pages 678–683. Vancouver, Canada, 2001.

[17] J.M. Odobez et P. Bouthemy. Motion-based motion segmentation exploiting a 2d motion model robust estimation. In *International Conference on Image Processing*, volume 3, pages 628–632, 1995.

[18] J.M. Odobez et P. Bouthemy. Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4):348–365, 1995.

- [19] 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*, 9(1):14–35, 1993.
- [20] C. Schmid et R. Mohr. Local greyvalue invariants for image retrieval. *PAMI*, 19(5):530–535, 1997.
- [21] J. Shi et C. Tomasi. Good features to track. In *CVPR*, pages 593–600, 1994.
- [22] H. Sidenbladh, M.J. Black, et L. Sigal. Implicit pobabilistic models of human motion for synthesis and tracking. In *European Conference on Computer Vision*, volume 1, pages 784–800, 2002.
- [23] T. Tommasini, A. Fusiello, E. Trucco, et V. Roberto. Making good features to track better. In *IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, California*, pages 178–183, 1998.
- [24] J. Verestóy et D. Chetverikov. Comparative performance evaluation of four point tracking techniques. In *22nd Workshop of the Austrian Pattern Recognition Group*, pages 255–263, 1998.